รายงานวิจัยฉบับสมบูรณ์

Final Research Report

โครงการวิจัยความหลากหลายทางภาษาในจังหวัดน่าน: ความรู้พื้นฐานเพื่อพัฒนาการท่องเที่ยว เฉลิมพระเกียรติสมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี เนื่องในวโรกาสที่ทรงเจริญพระชนมายุ 50 พรรษา วันที่ 2 เมษายน 2548

Research Project on Linguistic Diversity in Nan Province: a Foundation for Tourism Development To Commemorate the Fiftieth Birthday of H.R.H. Princess Maha Chakri Sirindhorn April 2, 2005

โครงการย่อยที่ 3 ระบบงานสารสนเทศภูมิศาสตร์โครงการน่าน

Sub-project 3

The Geographical Information System (GIS) of the Nan Project

โดย

by

พรรณี ชีวินศิริวัฒน์

Pannee Cheewinsiriwat

31 ตุลาคม 2550

October 31, 2007

สัญญาเลขที่ BRG478005

รายงานวิจัยฉบับสมบูรณ์

โครงการวิจัยความหลากหลายทางภาษาในจังหวัดน่าน:
ความรู้พื้นฐานเพื่อพัฒนาการท่องเที่ยว
เฉลิมพระเกียรติสมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี
เนื่องในวโรกาสที่ทรงเจริญพระชนมายุ 50 พรรษา
วันที่ 2 เมษายน 2548

Research Project on Linguistic Diversity in Nan Province:
a Foundation for Tourism Development
To Commemorate the Fiftieth Birthday of H.R.H. Princess Maha Chakri Sirindhorn
April 2, 2005

ที่ปรึกษา **นายแพทย์บุญยงค์ วงศ์รักมิตร**
หัวหน้าโครงการ **ศาสตราจารย์ ดร.ธีระพันธ์ เหลืองทองคำ**

คณะผู้วิจัย	*สังกัด*
1. ธีระพันธ์ เหลืองทองคำ (ล. ทองคำ)	คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
2. พรรณี ชีวินศิริวัฒน์	คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
3. สมเจตน์ วิมลเกษม	โรงเรียนสตรีศรีน่าน จังหวัดน่าน
ผู้ช่วยวิจัย (นักวิจัยหลังปริญญาเอกและนิสิตบัณฑิตศึกษา)	
1. พิณรัตน์ อัครวัฒนากุล	คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
2. ผณินทรา ธีรานนท์	คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
3. ชมนาด อินทจามรรักษ์	คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
4. วิชาติ บูรณะประเสริฐสุข	คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
5. กนิษฐา พุทธเสถียร	คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
6. ยุพาพร ฮวดศิริ	คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
7. สุภาพร ผลิพัฒน์	คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
8. ชนิตา ดวงยิหวา	คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
9. เอกกมล วรรณเมธี	คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
10. ประชาคมน่าน และเจ้าหน้าที่ศูนย์พัฒนาและสงเคราะห์ชาวเขา	

สนับสนุนโดยสำนักงานกองทุนสนับสนุนการวิจัย
(ความเห็นในรายงานนี้เป็นของผู้วิจัย สกว.ไม่จำเป็นต้องเห็นด้วยเสมอไป)

L-Thongkum, Theraphan. 2007. Final fricatives *-s and *-h in Proto-T'in lexicon. **Manusya**, Special issue No.13: 87-93. Presented at the **3rd International Conference on Austroasiatic Linguistics (ICAL 3 pilot picnic)**, organized by École française d'Extrême-Orient (EFEO), June 28 – 29, 2006, Siem Reap, Cambodia.

ABSTRACT

Final /-s/ does not exist in Modern T'in; however, *-s in Proto-T'in can be reconstructed. The distinction between *-s and *-h in Proto-T'in phonology is suggested by two types of final correspondences, i.e., *-s has become /-yh/ in Mal but /-t/ in Pray, and *-h has been kept as /-h/ in both Mal and Pray. This fact was pointed out by Filbeck in 1978; however, no Proto-T'in forms were reconstructed by him.

To show a clearer picture of the phonological history of T'in, 68 Proto-T'in forms were reconstructed, 17 with final *-s and 51 with final *-h. The 68 Proto-T'in reconstructed lexical items, (1) – (17) and (34) – (84), including 16 Proto-Mal reconstructed forms having *-s, (18) – (33), are presented in this paper with Thai and English glosses. Khmu' and Mlabri cognates from the author's corpus are also provided.

Keywords: *final fricatives, Proto Mal-Pray (Mon-Khmer)*

บทคัดย่อ

เสียงพยัญชนะท้าย /-s/ ไม่ปรากฏในภาษา "ถิ่น" (ลัวะน่าน หรือ มัล-ไปร) ปัจจุบัน อย่างไรก็ตาม เสียงพยัญชนะท้าย *-s ในภาษา "ถิ่น" สามารถสืบสร้างได้ในภาษา "ถิ่น" ดั้งเดิม ความแตกต่างที่มีนัยสำคัญระหว่างเสียงพยัญชนะท้าย *-s กับ *-h ในระบบเสียงภาษา "ถิ่น" ดั้งเดิมเห็นได้จากลักษณะการปฏิภาคของเสียง 2 ชุด คือ เสียงพยัญชนะท้าย *-s ที่เปลี่ยนเป็น /-yh/ ในภาษามัล แต่เปลี่ยนเป็นเสียง /-t/ ในภาษาไปร กับเสียงพยัญชนะท้าย *-h ที่ยังคงอยู่ทั้งในภาษามัลและไปร Filbeck (1978) ได้ชี้ให้เห็นความจริงนี้ อย่างไรก็ตาม Filbeck มิได้สืบสร้างคำในภาษา "ถิ่น" ดั้งเดิมที่มีเสียงพยัญชนะท้าย *-s และ *-h

เพื่อแสดงให้เห็นภาพประวัติความเป็นมาของระบบเสียงภาษา "ถิ่น" ชัดเจนขึ้น ผู้วิจัยได้สืบสร้างศัพท์ในภาษา "ถิ่น" ดั้งเดิมจำนวน 68 คำ คำศัพท์ที่สืบสร้างขึ้น 68 คำนี้ 17 คำ (1) - (17) มีเสียงพยัญชนะท้าย *-s และ 51 คำ (34) - (84) มีเสียงพยัญชนะท้าย *-h นอกจากคำในภาษา "ถิ่น" ดั้งเดิม 17 คำที่มีเสียงพยัญชนะท้าย *-s ยังได้เพิ่มเติมคำในภาษามัลดั้งเดิมที่มีเสียงพยัญชนะ *-s เข้ามาด้วย

อีก 16 คำ คือ (18) - (33) ในการนำเสนอคำที่สืบสร้างขึ้นทั้งหมด 84 คำ ได้ให้ความหมายภาษาไทยและความหมายภาษาอังกฤษ รวมทั้งคำร่วมเชื้อสายในภาษาขมุและมลาบรีซึ่งมาจากคลังข้อมูลของผู้วิจัยด้วย

คำสำคัญ: *เสียงพยัญชนะท้ายเสียดแทรก ภาษามัล-ไปรดั้งเดิม (มอญ-เขมร)*

L-Thongkum, Theraphan, Phanintra Teeranon and Chommanad Intajamornrak. 2007. The interaction between vowel length and pitch in SEA languages: an implication for tonal evolution. In **Studies in Tai and Southeast Asian Linguistics**, 225-240, edited by Jimmy G. Harris, Somsonge Burusphat and James E. Harris. Bangkok: Ekphimthai Ltd. Presented at the **38th International Conference on Sino-Tibetan Languages and Linguistics**, organized by Xiamen University, October 28 – 31, 2005, China.

ABSTRACT

Hu, an Angkuic language of the Palaungic branch of the Mon-Khmer language family spoken in Yunnan, has two tones, high and low. The loss of phonological vowel length is the major cause of the birth of the low tone, with original short vowels receiving the high tone in Modern Hu. If this type of tonal evolution is possible, it should be attested phonetically.

The interaction between vowel length and pitch in the Mon-Khmer, Hmong-Mien and Tai languages was studied by measuring the duration, fundamental frequency and intensity values of /a/ and /aa/ in minimal pairs. The results of our investigation indicated that short vowels had higher fundamental frequencies (F0) than long vowels and that the F0 difference between short and long vowels was statistically significant ($p < 0.05$) in both tonal and non-tonal languages. This seems to be a natural phonetic tendency. The interaction between vowel length and loudness (intensity) is less clear.

An attempt to make a pair of short and long vowels of the same quality, e.g., /a/ – /aa/, having “equal prominence”, can cause pitch difference between short and long vowels in SEA languages. The birth of phonological pitch or tone at a later stage may be regarded as compensation for the loss of phonological length at the earlier stage.

Keywords: *vowel length and pitch, SEA languages*

บทคัดย่อ

ภาษาหูซึ่งเป็นภาษาในสาขาปะหล่องอิกของภาษาตระกูลมอญ-เขมรพูดในมณฑลยูนนาน มีวรรณยุกต์ 2 หน่วยเสียง คือ วรรณยุกต์สูง กับ วรรณยุกต์ต่ำ การสูญเสียความยาวของเสียงสระเป็นสาเหตุหลักที่ทำให้เกิดวรรณยุกต์ต่ำในคำภาษาหู ส่วนคำที่มีเสียงสระสั้นมาแต่เดิมจะมีวรรณยุกต์สูงในภาษาหูปัจจุบัน หากวิวัฒนาการของวรรณยุกต์ลักษณะนี้เป็นไปได้ ก็น่าจะพิสูจน์ทราบทางสัทศาสตร์ได้

ปฏิสัมพันธ์ระหว่างความสั้นยาวของเสียงสระกับระดับเสียงในภาษาตระกูลมอญ - เขมร ม้ง - เมี่ยน และภาษาตระกูลไทได้รับการศึกษาวิเคราะห์ในเรื่อง ค่าระยะเวลา ค่าความถี่มูลฐาน และค่าความเข้ม โดยใช้คู่เทียบเสียง /a/ กับ /aa/ ผลการวิจัยชี้ให้เห็นว่าสระสั้นมีค่าความถี่มูลฐานสูงหรือมากกว่าสระยาว ความแตกต่างนี้มีนัยสำคัญทางภาษาศาสตร์ ($p<0.05$) ทั้งในภาษามีวรรณยุกต์และไม่มีวรรณยุกต์ ปรากฏการณ์นี้ดูเหมือนจะเป็นแนวโน้มทางสัทศาสตร์ ส่วนความสัมพันธ์ระหว่างความสั้นยาวของสระกับความดังค่อยหรือความเข้มของเสียงไม่ค่อยมีรูปแบบที่ชัดเจน

คำสำคัญ: *ความสั้นยาวของสระกับระดับเสียง ภาษาเอเชียตะวันออกเฉียงใต้*

Abramson, Arthur S., Patrick W. Nye and Theraphan Luangthongkum. 2007. Voice Register in Khmu': experiments in production and perception. **Phonetica**. Vol 64, No. 2-3, 80-104. Presented at **the 4th Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan**, November, 28 – December, 2, 2006, Honolulu, Hawaii.

ABSTRACT

Some Khmu' dialects have phonologically distinctive voice registers. Auditory observations have claimed a stable distinction between clear voice and high pitch for Register 1 and breathy voice and low pitch for Register 2 in the Khmu' Rawk dialect of northern Thailand. Word-pairs distinguished only by register were recorded by 25 native speakers. Acoustic analysis yielded F0 and overall amplitude contours, frequencies of F1 and F2 in quasi-steady states of the vowels, relative intensities of higher harmonics to that of the first harmonic, and vowel durations. When circumstances caused early attention to perception testing, the words of only eight speakers had been analyzed for properties other than amplitude and F0. Since the only significant factor that had emerged by then was F0

contour, the synthetic stimuli were made with just a series of seven contours. The labeling by 32 native speakers yielded two categories, demonstrating the sufficiency of F0 as an acoustic cue. The completed acoustic analysis showed a significant effect of one of the harmonic ratios for the women only, suggesting a conservative bias. The language has been shifting toward tonality and may have reached it.

Keywords: *phonation type, tonal evolution, Khmu' language (Mon-Khmer)*

บทคัดย่อ

ภาษาขมุบางถิ่นมีความแตกต่างด้านลักษณะน้ำเสียงซึ่งทำให้ความหมายของคำต่างกัน จากการฟังได้มีการกล่าวว่า มีความแตกต่างระหว่างลักษณะน้ำเสียงธรรมดากับระดับน้ำเสียงสูงซึ่งเป็นองค์ประกอบของริจิสเตอร์ 1 และลักษณะน้ำเสียงต่ำทุ้มกับระดับเสียงต่ำซึ่งเป็นองค์ประกอบของริจิสเตอร์ 2 ในภาษาขมุรอกซึ่งพูดในภาคเหนือของประเทศไทย (จังหวัดน่าน) คำคู่เทียบเสียงที่แตกต่างกันเฉพาะริจิสเตอร์ได้ถูกบันทึกเสียงโดยใช้ผู้พูดภาษาขมุเป็นภาษาแม่ 25 คน เพื่อการวิเคราะห์ค่าความถี่มูลฐาน ค่าความเข้มโดยรวม ค่าความถี่ฟอร์เมินท์ 1 และ 2 ของสระช่วงที่ไม่มีการบิดเบน ค่าความเข้มเทียบเคียงของฮาร์มอนิกสูงๆ กับฮาร์มอนิกที่ 1 และค่าระยะเวลาของเสียงสระ ในส่วนการทดสอบการรับรู้ ด้วยสภาวะบางอย่างที่เป็นปัญหา ได้เลือกการออกเสียงคำของผู้พูดภาษาขมุเพียง 8 คนเท่านั้นมาวิเคราะห์ค่าทางกลสัทศาสตร์อื่นๆ ร่วมกับ ค่าความเข้มและค่าความถี่มูลฐาน เนื่องจากได้พบว่า ปัจจัยด้านความถี่มูลฐานมีนัยสำคัญที่สุด การสร้างตัวเร่งเร้าสังเคราะห์เพื่อทดสอบการรับรู้จำกัดอยู่ที่การขึ้นตกของระดับเสียง 7 ลักษณะเท่านั้น จากการทดสอบการรับรู้กับคนขมุ 32 คน ทำให้ทราบว่ามีการรับรู้ 2 ประเภท คือ ส่วนใหญ่รับรู้ค่าความถี่มูลฐานหรือระดับเสียงเป็นสำคัญ ส่วนการรับรู้ความแตกต่างอย่างมีนัยสำคัญของความเข้มเทียบเคียงของฮาร์มอนิกมีแต่ในกลุ่มผู้ทดสอบที่เป็นเพศหญิงเท่านั้น ซึ่งแสดงว่าผู้พูดเพศหญิงยังคงเก็บรักษาลักษณะเก่าแก่ดั้งเดิมของภาษาขมุไว้ คือ ลักษณะน้ำเสียง ปรากฏการณ์นี้แสดงให้เห็นว่าภาษาขมุรอกกำลังเปลี่ยนแปลงจากภาษาลักษณะน้ำเสียงเป็นภาษาวรรณยุกต์จนเกือบจะสมบูรณ์แล้ว

คำสำคัญ: *ลักษณะน้ำเสียง วิวัฒนาการของวรรณยุกต์ ภาษาขมุ (มอญ-เขมร)*

L-Thongkum, Theraphan and Chommanad Intajamornrak. 2008. Tonal evolution induced by language contact: a case study of the T'in (Lua') language of Nan province, Northern Thailand. To appear in **Mon-Khmer Studies 38**. Presented at **Mon-Khmer Workshop**,

organized by the Language and Cognition Group, Max Planck Institute for Psycholinguistics, April 16 – 18, 2007, Nijmegen, The Netherlands.

ABSTRACT

T'in is a Mon-Khmer language; generally, Mon-Khmer languages are non-tonal. The T'in (Lua) language of Bo Kluea District, Nan Province, comprises two major dialects, Mal and Pray. Pray is more conservative and has less speakers, due to geography and communist infiltration in the past. The data on Mal, Pray and Tai Yuan (language of the majority) were collected in 2005 for the research project on "Linguistic Diversity in Nan Province: A Foundation for Tourism Development".

Although different pitches can be heard in both Mal and Pray, there are no consistent patterns in Pray, even in the speech of a single speaker. In Mal, two pitches [high-falling] and [low-rising] have been found distinctive. It can be concluded that the Mal dialect of Bo Kluea District has acquired two tones, /falling/ vs. /rising/ or /high/ vs. /low/. The falling tone occurs mostly in native words, whereas the rising tone mostly occurs in Tai loanwords. However, a few native words can have rising tone and a few Tai loanwords can have falling tone. Acoustical measurements were done to confirm the characteristics of Mal and Thai Yuan tones.

Filbeck (1972) pointed out that a dialect of Mal had two tones, i.e., rising and non-rising, and that the emergence of the rising tone appears to be the result of an independent innovation. The results of our present study indicate that the two rising tones in Tai Yuan, namely, /low-rising/ (A1-2) and /mid-rising/ (A3-4) seem to be the cause of tone birth in Mal, especially, the rising tone. In Tai Yuan, a six-tone language, words having rising tones (A1-2 and A3-4) are out-numbered, thus, frequently heard by Mal people. Perhaps, this auditory factor could have induced tonal evolution in some Mal dialects.

Keywords: *tonal evolution, language contact, Mal and Tai Yuan*

บทคัดย่อ

ภาษา "ถิ่น" เป็นภาษาในตระกูลมอญ-เขมร โดยทั่วไปภาษาตระกูลมอญ-เขมรไม่มีวรรณยุกต์ ภาษา "ถิ่น" (ลัวะ) ที่พูดในอำเภอบ่อเกลือ จังหวัดน่าน แบ่งออกเป็น 2 วิธภาษาหลัก คือ มัล และไปร

ไปรซึ่งมีผู้พูดจำนวนน้อยกว่าสามารถเก็บรักษาลักษณะดั้งเดิมไว้ได้มากกว่า ทั้งนี้เนื่องมาจากสภาพทางภูมิศาสตร์และการแทรกซึมของคอมมิวนิสต์ในอดีต ข้อมูลภาษามัล ไปร ไทยวนหรือคำเมือง (ภาษาของคนส่วนใหญ่) เก็บเมื่อปี 2548 เพื่อการวิจัยของโครงการ “ความหลากหลายทางภาษาในจังหวัดน่าน: ความรู้พื้นฐานเพื่อพัฒนาการท่องเที่ยว”

ถึงแม้ว่าจะได้ยินความแตกต่างของระดับเสียงทั้งในมัลและไปร แต่รูปแบบระดับเสียงที่สม่ำเสมอก็ไม่ปรากฏในภาษาไปร แม้แต่ในคำพูดของผู้พูดเพียง 1 คน ระดับเสียง 2 ลักษณะในมัล คือ [สูง-ตก] และ [ต่ำ-ขึ้น] ก่อให้เกิดความแตกต่างทางความหมาย สรุปได้ว่าวิธภาษามัลที่พูดในอำเภอบ่อเกลือมี 2 วรรณยุกต์ คือ /ตก/ กับ /ขึ้น/ หรือ /สูง/ กับ /ต่ำ/ วรรณยุกต์ตกมักจะเกิดในคำมัลพื้นเมือง ขณะที่วรรณยุกต์ขึ้นมักจะเกิดในคำยืมไท อย่างไรก็ตาม คำมัลพื้นเมืองบางคำก็มีวรรณยุกต์ขึ้นและคำมัลที่เป็นคำยืมไทบางคำก็มีวรรณยุกต์ตก มีการวัดค่าทางกลสัทศาสตร์เพื่อยืนยันสัทลักษณะของวรรณยุกต์ในมัลและไทยวน

ฟิลเบ็ค (1972) ชี้ให้เห็นว่าบางวิธภาษาของมัลมีวรรณยุกต์ 2 หน่วยเสียง คือ วรรณยุกต์ขึ้นกับวรรณยุกต์ไม่ขึ้น การกำเนิดของวรรณยุกต์ขึ้นเป็นนวัตกรรมอิสระ ผลการวิจัยของเราแสดงให้เห็นว่าวรรณยุกต์ขึ้น 2 หน่วยเสียงในภาษาไทยวน คือ /ต่ำ-ขึ้น/ (A1-2) กับ /กลาง-ขึ้น/ (A3-4) น่าจะเป็นที่มาของการกำเนิดวรรณยุกต์ในมัล ภาษาไทยวนหรือคำเมืองถิ่นน่านมีวรรณยุกต์ 6 หน่วยเสียง คำที่มีวรรณยุกต์ขึ้น (A1-2 กับ A3-4) มีเป็นจำนวนมากกว่าวรรณยุกต์อื่นๆ ด้วยเหตุนี้คนมัลทวิภาษาจึงได้ยินและได้ใช้วรรณยุกต์ขึ้นบ่อยๆ จึงรับสัทลักษณะขึ้นเข้าไปในภาษาแม่ของตน อันเป็นปัจจัยต่อวิวัฒนาการของวรรณยุกต์

คำสำคัญ: *วิวัฒนาการของวรรณยุกต์ การสัมผัสภาษา มัลและไทยวน*

L-Thongkum, Theraphan. Language change without collision: a glimpse at linguistic diversity in northern Thailand and southern Laos. Presented at **The Beijing Forum (2007), Panel Session on “Language Identity and Language Change in Collision and Dialogue Between Civilizations”, Theme 3: Clashes of Civilizations and Consequent Language Change**, November 2-4, 2007. Beijing University, China. (To appear in the proceedings)

ABSTRACT

Linguistic diversity can be the cause of serious clashes and collisions in nations with multiethnic groups. Is this always true? Linguistic variation and change due to linguistic

diversity can be subtle and be accepted without collision or negative feelings when induced gradually and naturally by language contact.

Linguistic diversity in the Nan Province of northern Thailand and in the Xekong Province of southern Laos is a good illustration of the latter phenomenon. In the case of Nan and Xekong, though small both in area and population size, these provinces have thirteen and fourteen ethnic groups, respectively, coexisting peacefully. These ethnic minorities have learned to tolerate, to care for and to listen to each other in order to maintain a sense of harmony in the area. Unity is the result of their linguistic tolerance and lack of selfishness. Due to language contact, linguistic variation and change cannot be avoided. Different languages play different roles in a multilingual community. Pluralism and multilingualism should be taken as important keywords for policy making at the national level. However, assimilation should also be encouraged; if people want to speak or conform to the majority instead of maintaining their own identity and speaking their own language, they have the right to choose whatever they think best suits them.

Keywords: *linguistic diversity, language contact, northern Thailand, southern Laos*

บทคัดย่อ

ความหลากหลายทางภาษาสามารถเป็นสาเหตุของความแตกร้าวและแตกแยกในชาติที่มีคนหลายเชื้อชาติเผ่าพันธุ์ เป็นความจริงเสมอไปหรือไม่ การแปรและการเปลี่ยนแปลงของภาษาอันเนื่องมาจากความหลากหลายทางภาษา สามารถที่จะเกิดขึ้นอย่างค่อยเป็นค่อยไปและยอมรับได้โดยไม่ต้องเกิดความแตกแยกหรือความรู้สึกด้านลบ หากเกิดขึ้นช้าๆ และเป็นธรรมชาติโดยการสัมผัสภาษา

ความหลากหลายทางภาษาในจังหวัดน่าน ภาคเหนือของประเทศไทย และในจังหวัดเซกอง ภาคใต้ของประเทศ สปป. ลาว เป็นตัวอย่างที่ดีของการแปรและเปลี่ยนแปลงของภาษาอย่างสันติ กรณีของน่านและเซกองซึ่งเป็นจังหวัดเล็กแต่มีประชากรประกอบด้วย 13 กลุ่มชาติพันธุ์และ 14 กลุ่มชาติพันธุ์ตามลำดับ ก็สามารถอยู่กันได้อย่างสงบ กลุ่มชาติพันธุ์เหล่านี้เรียนรู้ที่จะอดทนอดกลั้น เห็นอกเห็นใจและรับฟังซึ่งกันและกัน เพื่อรักษาความกลมเกลียวในท้องถิ่นของตน เอกภาพเป็นผลจากความอดทนด้านภาษาและความไม่เห็นแก่ตน การสัมผัสภาษาทำให้การแปรและการเปลี่ยนแปลงของภาษาเป็นสิ่งหลีกเลี่ยงไม่ได้ ภาษาต่างๆ มีบทบาทต่างกันในชุมชนพหุภาษา พหุลักษณ์และพหุภาษาควรเป็นกุญแจสำคัญในการจัดทำนโยบายระดับชาติ อย่างไรก็ตาม การผสมกลมกลืนควรได้รับการ

สนับสนุนด้วย หากผู้คนทั้งหลายอยากพูดหรือปฏิบัติตนให้สอดรับกับคนส่วนใหญ่แทนที่จะเก็บรักษาเอกลักษณ์เฉพาะและพูดแต่ภาษาแม่ของตน พวกเขาควรมีสิทธิเสรีภาพในการเลือกสิ่งที่พวกเขาคิดว่าเหมาะสมที่สุดสำหรับพวกเขา

คำสำคัญ: *ความหลากหลายทางภาษา การสัมผัสภาษา ภาคเหนือของไทย ภาคใต้ของลาว*

Intajamornrak, Chommanad. Tai loanwords in Mal: a minority language of Thailand. Presented at the **9th Annual Southeast Asian Studies Graduate Student Conference**, March 16 – 18, 2007, Cornell University, USA.

ABSTACT

Mal is a dialect of Lua/Thin, a Mon-Khmer language of Nan Province. Almost all the speakers are bilingual or multilingual. In a situation of bilingualism or multilingualism, language contact can bring about phonological interference. Mal has borrowed Tai words for a long time. My aim is to study transformational processes of Tai loanword adaptation in Mal.

The data were taken from a wordlist of 2,452 lexical items compiled for investigating the 13 languages of Nan. To separate loanwords from native words, loanwords easily distinguished by phonological form were separated first. Then ambiguous words were checked with the other Tai and Mon-Khmer languages of Nan. Then the percentage of loanwords in the data was calculated.

About one-third of the lexical items are Tai loanwords. Compound nouns and proper nouns are always composed of Mal native words and loanwords, for example, *siəŋ* (M) *khăaŋ* ‘molar teeth’. The transformational processes found for the adaptation of Tai loanwords fall into two main categories; 1) Segmental change and 2) Suprasegmental innovation. Segmental change is the replacement of consonant and vowel phonemes not found in Mal; for example, the phoneme /f/ in the source languages is replaced by /ph/ in the word *fun*^{21} (ST) > *phŭn* (M) ‘dust’. As for suprasegmental innovation, the low-rising tone, found on about half of the Tai loanwords, must be the result of language contact with a tonal language, especially a language with many words on a rising tone. Examples are the Mal words *k□□□ŋ* ‘drum’, *kwăat* ‘to sweep’, and *kə□ət* ‘born’. It seems that the emergence of

the low-rising tone in Mal is a sociolinguistic device for marking loanwords. Loanwords not having low-rising tone vary prosodically, along with all native Mal words, as affected by syllable structure and sentence intonation. The reason for the lack of the low-rising tone on the other half of the Tai loanwords may be the time-depth of borrowing that caused them to be better integrated with the native phonology.

It can be concluded that loanword adaptation in Mal involve the native Mal phonology as well as the emergence of a new feature limited to loanwords. What we seem to have here is an incipient tone language. It remains to be seen whether a full tonal system will come into being over time or whether the present incipient system will die out.

(M) = Mal, (T) = Tai, (ST) = Standard Thai, (TY) = Tai Yuan

Keywords: *bilingualism, multilingualism, borrowing, loanwords*

บทคัดย่อ

ภาษามัลเป็นภาษาในตระกูลมอญ-เขมรที่พูดในจังหวัดน่าน จัดเป็นชุมชนทวิภาษาหรือพหุภาษา ปรากฏการณ์ทางภาษาที่พบในสถานการณ์เช่นนี้ก็คือการสัมผัสภาษา ซึ่งนำไปสู่การแทรกแซงทางเสียง และภาษามัลได้ยืมคำจากภาษาไทมาเป็นเวลานาน วัตถุประสงค์ของการวิจัยนี้เพื่อศึกษากระบวนการแปลงเสียงของคำยืมไทในภาษามัล

ข้อมูลได้จากรายการคำสำหรับเก็บข้อมูล 13 ภาษาที่พูดในจังหวัดน่านจำนวน 2,452 คำ การจำแนกคำยืม พิจารณาจากเสียงหรือหน่วยเสียงเป็นสำคัญ หากพบคำที่สงสัยจะนำไปตรวจสอบกับภาษาไทและภาษามอญ-เขมรอื่นๆ ที่พูดในจังหวัดน่าน จากนั้นจึงคำนวณหาอัตราส่วนของคำยืมที่พบทั้งหมด

ผลการวิจัยพบว่า หนึ่งในสามของรายการคำทั้งหมดเป็นคำยืม โดยคำประสมและคำนามชี้เฉพาะมักประกอบด้วยคำมัลแท้และคำยืม เช่น *siəŋ* (M) *khăaŋ* ‘กราม’ กระบวนการแปลงเสียงของคำยืมไทในภาษามัลแบ่งได้เป็น 2 ประเภท คือ การแปลงหน่วยเสียงเรียง และการเกิดหน่วยเสียงซ้อน การแปลงหน่วยเสียงเรียง ได้แก่การแทนที่เสียงพยัญชนะและสระที่ไม่มีในภาษามัล เช่น เสียง /f/ ในภาษาไทถูกแทนที่ด้วยเสียง /ph/ ในคำว่า *fun*^{21} (ST) > *phŭn* (M) ‘ฝุ่น’ ส่วนการเกิดหน่วยเสียงซ้อนหรือวรรณยุกต์ต่ำ-ขึ้นที่พบในคำยืมไทเป็นจำนวนกว่าครึ่งหนึ่งนั้น เป็นผลจากการสัมผัสภาษากับภาษามีวรรณยุกต์ โดยเฉพาะภาษาที่มีวรรณยุกต์ขึ้นเป็นจำนวนมาก ตัวอย่างคำภาษามัล เช่น *k□□□ŋ* ‘กลอง’, *kwăat* ‘กวาด’, and *kə□ət* ‘เกิด’ การกำเนิดขึ้นของวรรณยุกต์ต่ำ-ขึ้นดูเหมือนเป็นเครื่องมือที่ใช้สำหรับการจำแนกคำยืม ส่วนคำยืมที่ไม่มีวรรณยุกต์ต่ำ-ขึ้นจะแปรไปตามโครงสร้าง

พยางค์และการลงเสียงหนักเบาในประโยค ซึ่งสาเหตุของการหายไปของวรรณยุกต์ต่ำ-ขึ้นของคำยืมนั้นน่าจะมาจากระยะเวลาของการยืม ที่ทำให้คำเหล่านั้นเกิดการผสมผสานกับภาษามัลซึ่งเป็นภาษาแม่

กล่าวได้ว่าการแปลงคำยืมที่พบในภาษามัลประกอบไปด้วยระบบเสียงของภาษามัลและการเกิดขึ้นของคุณลักษณะใหม่ที่เฉพาะเจาะจงแต่ในคำยืม ภาษามัลจึงเรียกได้ว่าเป็นภาษาที่เริ่มมีวรรณยุกต์เกิดขึ้น สิ่งที่ต้องศึกษาต่อไปก็คือ ภาษามัลจะมีการพัฒนาระบบเสียงวรรณยุกต์อย่างเต็มรูปแบบหรือวรรณยุกต์ที่เกิดขึ้นนี้จะสูญหายไป

คำสำคัญ: *ทวิภาษา พหุภาษา การยืม คำยืม*

Akharawatthanakun, Phinnarat. Contact-induced linguistic variation and change: a case study of Tai dialects spoken in language mixture areas of Nan province. Presented at the **1st World Congress on the Power of Language: Theory, Practice, and Development**, May 22 – 25, 2006, Bangkok, Thailand.

ABSTRACT

It is generally accepted that language change can be influenced by both internal and external factors. Language contact is an external factor which has an important role for motivating linguistic change. In Akharawatthanakun (2002, 2004), the tone variation and change which is found in Lao dialects is influenced by both an internal factor, tonal simplification, and an external factor, language contact. This paper aims to show more evidence of phonological variation and change motivated by language contact. Having analyzed the consonants, the vowels, and the tones in five Tai dialects: Tai Nyuan, Tai Lue, Tai Khün, Phuan, and Lao, spoken in the language mixture areas of Nan Province in northern Thailand, the findings reveal that: 1) the variation and change occurs in every phonological aspect: consonants, vowels, and tones and 2) the degree of phonological variation and change in some Tai dialects is higher than in the others. Furthermore, in this study lexical items in the five Tai dialects have also been investigated. Lexical variation and change is also found as evidence proving that, in the language mixture areas, the majority language plays important role in influencing the minority languages. This paper will present the variation and change in only the phonological aspect, not in the lexical aspect, so that the explanation and

discussion will be clearly shown and not confusing. The variation and change in the lexical aspect will be presented in another paper of the author, ***"Contact-Induced Lexical Variation and Change: A Case Study of Tai Dialects Spoken in Language Mixture Areas of Nan Province"*** in which the research result will also the evidence of linguistic variation and change due to an external factor, i.e. language contact.

Keywords: *language variation and change, language contact*

บทคัดย่อ

เป็นที่ยอมรับกันโดยทั่วไปว่า การเปลี่ยนแปลงของภาษาเกิดขึ้นอันเนื่องมาจากทั้งปัจจัยภายในและปัจจัยภายนอก การสัมผัสภาษาเป็นปัจจัยภายนอกประเภทหนึ่งที่มีบทบาทสำคัญต่อการเปลี่ยนแปลงภาษา ในงานวิจัยที่ผ่านมาของผู้เขียน (Akharawatthanakun 2002, 2004) การแปรและการเปลี่ยนแปลงของวรรณยุกต์ที่พบในภาษาลาวถิ่นต่าง ๆ ได้รับอิทธิพลจากทั้งปัจจัยภายใน ได้แก่ การออกเสียงวรรณยุกต์ให้ง่ายขึ้น (tonal simplification) และปัจจัยภายนอก ได้แก่ การสัมผัสภาษา บทความนี้มีวัตถุประสงค์ที่จะนำเสนอให้เห็นหลักฐานเกี่ยวกับการแปรและการเปลี่ยนแปลงของภาษาอันเนื่องมาจากการสัมผัสภาษา จากการวิเคราะห์พยัญชนะ สระ และวรรณยุกต์ในภาษาไทถิ่น 5 ภาษา ได้แก่ ได้แก่ ภาษาไทยวน ไทลื้อ ไทขืน พวน และลาวที่พูดในบริเวณที่มีการปะปนกันของภาษาในจังหวัดน่านทางภาคเหนือของประเทศไทย พบว่า 1) มีการแปรและการเปลี่ยนแปลงทางเสียงเกิดขึ้นทั้งในระดับของพยัญชนะ สระ และวรรณยุกต์ และ 2) ภาษาไทถิ่นบางภาษามีระดับการแปรและการเปลี่ยนแปลงทางเสียงสูงกว่าภาษาไทถิ่นอื่น นอกจากนี้งานวิจัยนี้ยังวิเคราะห์คำศัพท์ในภาษาไทถิ่นทั้ง 5 ภาษาด้วยและพบว่า การแปรและการเปลี่ยนแปลงทางศัพท์นับเป็นอีกหลักฐานหนึ่งที่แสดงให้เห็นว่า ภาษาของชนกลุ่มใหญ่มีบทบาทสำคัญต่อการแปรและการเปลี่ยนแปลงของภาษาของชนกลุ่มน้อย เพื่อให้สามารถอธิบายและอภิปรายผลการวิจัยได้อย่างชัดเจนและไม่สับสน บทความนี้จะนำเสนอเฉพาะการแปรและการเปลี่ยนแปลงทางเสียงเท่านั้น ส่วนการแปรและการเปลี่ยนแปลงในระดับคำศัพท์จะนำเสนอไว้ในอีกบทความเรื่อง ***"Contact-Induced Lexical Variation and Change: A Case Study of Tai Dialects Spoken in Language Mixture Areas of Nan Province"*** ซึ่งเป็นอีกบทความหนึ่งของผู้เขียนที่จะนำเสนอผลการวิจัยเกี่ยวกับการแปรและการเปลี่ยนแปลงของภาษาอันเนื่องมาจากปัจจัยภายนอก หรือได้แก่การสัมผัสภาษานั่นเอง

คำสำคัญ: *การแปรและการเปลี่ยนแปลงของภาษา การสัมผัสภาษา*

FINAL FRICATIVES*-s AND *-h IN PROTO T'IN LEXICON

Theraphan L-Thongkum[1]

Abstract

*Final /-s/ does not exist in Modern T'in; however, *-s in Proto-T'in can be reconstructed. The distinction between *-s and *-h in Proto-T'in phonology is suggested by two types of final correspondences, i.e., *-s has become /-yh/ in Mal but /-t/ in Pray, and *-h has been kept as /-h/ in both Mal and Pray. This fact was pointed out by Filbeck in 1978; however, no Proto-T'in forms were reconstructed by him.*

*To show a clearer picture of the phonological history of T'in, 68 Proto-T'in forms were reconstructed, 17 with final *-s and 51 with final *-h. The 68 Proto-T'in reconstructed lexical items, (1) – (17) and (34) – (84), including 16 Proto-Mal reconstructed forms having *-s, (18) – (33), are presented in this paper with Thai and English glosses. Khmu' and Mlabri cognates from the author's corpus are also provided.*

Introduction

Many aspects of the T'in or Lua' language of Nan Province, Thailand, were studied by Filbeck (1972, 1978, 1991), Ratanakul (1975), Satayawadhna (1987), Singnoi (1988) and Jirananthanaporn (1993). T'in comprises two major dialects, Mal and Pray, and both dialects consist of many sub-dialects or varieties caused by the geography and history of the area where the T'in inhabit.

The author collected data on Mal and Pray spoken in Bo Kluea District in the year 2005. A wordlist of 2,452 lexical items was devised for investigating the 13 languages of Nan. This paper is only a minor output of the research project on "Linguistic Diversity in Nan Province: A Foundation for Tourism Development" sponsored by the Thailand Research Fund (TRF) for three years (May 1, 2004 – April 30, 2007).

Sketch of Proto-T'in phonology

Based on the author's Pray and Mal wordlists, about 440 cognates have been found.[2] Proto-T'in phonology can be reconstructed as shown in Table 1. In comparison with Filbeck's reconstruction (Filbeck, 1978), this present reconstruction is less complex, especially the initials, because prenasalization of the initials of major-syllables, e.g., μβ-, ντ-, Νκη-, is analyzed as the finals of pre-syllables in Proto-T'in, see items (2), (4), (7), (19), (21) and so on.

Reconstructed forms with *-s and *-h

*-s has become /-yh/ in Modern Mal and /-t/ in Modern Pray. In Mal, *V (vowels) before *-s had been diphthongized (*Vs > *Vι9s) and later *-s became h-like sound. This process of sound change yields /-yh/ or /-y∗/ (devoiced palatal fricative [∅∗])

[1] Professor, Department of Linguistics, Faculty of Arts, Chulalongkorn University, Bangkok, Thailand.

[2] The number of cognates is less than expectation. The explanation is that the wordlist devised is for synchronic investigation not for comparative purposes.

in

Table 1 Proto-T'in phonology

Initial

*π-	*πη-	*τ-	*τη-	*χ-		*κ-	*κη-	*/-
*μ-	*ημ-	*ν-	*ην-	*ȷ-	*ηȷ-	*N-	*ηN-	
		*λ-	*ηλ-					
		*ρ-						
		*σ-			*η-			
*πρ-	*πλ-	*τρ-	*τλ-	*χρ-	*κρ-	*κλ-	*κω-	
*πηρ-	*πηλ-	*τηρ-	*τηλ-		*κηρ-	*κηλ-	*κηω-	

Final

*-π	*-τ	*-χ	*-κ	*-/
*-μ	*-ν	*-ȷ	*-N	
*-ω	*-λ	*-ψ		
	*-ρ			
	*-σ			*-η

Vowel

*ι	*◝	*υ	*ιι	*◝◝	*υυ
*ε	*↔	*ο	*εε	*↔↔	*οο
*E	*α	*□	*EE	*αα	*□□
	*ι↔	*ια	*υ↔	*υα	*υ□

Modern Mal dialect and sub-dialects. Khmu' (KM) and Mlabri (MB) cognates are also given for a comparison of Khmuic languages, see Table 2, 3 and 4.

Regarding Pray /-t/ which has been from Proto-T'in *-s, the process of sound change is different from Mal, i.e., voiceless alveolar fricative *-s became voiceless alveolar affricate *ts- in one stage of Pray phonological history. In a more recent stage, this affricative lost its fricative release, then, has become voiceless alveolar plosive /-t/ in Modern Pray. This phenomenon (*-s becoming -t) is not unusual since it also occurs in other Southeast Asian languages spoken in Thailand. In Thai, English loanwords, such as <gas> is pronounced [κE≅τ~κE≅Eτ~κα≅ατ] by speakers who do not know English.

Besides the 17 reconstructed forms and some cognates in Modern Khmu' (KM) and Modern Mlabri (MB) given above, the final /-yh/ in the other 16 Modern Mal words also suggests *-s in Proto-Mal and perhaps Proto-T'in, although the Pray cognates have not been found for a comparison. It is noticeable that some Pray native words have been replaced by Tai loans, e.g., items (19), (20), (22), (23), (25) and (33) in Table 3.

Table 2 Proto-T'in *-s

	Thai gloss	*English gloss*	*Mal*	*Pray*	*Proto-T'in*	
(1)	เข็ม	needle	ναaψη	Nαατ	*( )αασ	
(2)	จาม	to sneeze	Nκηαψη	Nκατ	*N-κ(η)ασ	(KM: ν/α≅ψη; MB: πρΕσ)
(3)	ดีด (นิ้ว)	to snap (fingers)	πηλαψη	πηλατ	*πηλασ	(KM: πλ→≅ψη; MB: πλ℘λη)
(4)	ตั๊กแตน	grasshopper	νσ□□ψη	⟩χη□□τ	*ν-σ□□σ	(KM: η□≅□ψη)
(5)	ถล่ม	tumble down	κλ↔ψη	κλ↔τ	*κλ↔σ	
(6)	นั่ง	to sit	κηυψη	κηυτ	*κηυσ	
(7)	ปาด (ก.)	to slice	νσυ↔ψη	χυ↔τ	*ν-συ↔σ	(MB: σ□□λη)
(8)	ฟัก (ไข่)	to hatch	τηαψη	τηατ	*τηασ	(KM: κηλα∃ψη; MB: τασ)
(9)	ไฟ	fire	πα/ooψη	/oot	*π-/oσ	(MB: /υυλη)
(10)	ราก (น.)	root	κι↔ψη	κι↔τ	*κι↔σ	(KM: ρ→∃↔ψη; MB: ρΕΕλη)
(11)	ริมฝีปาก	lips	νooψη	σανoot	*σ-νooσ	
(12)	ลัก	to steal	λ□□ψη	λ□□τ	*λ□□σ	(MB: λ□□ψη)
(13)	หนี	to escape	παψη	πατ	*πασ	
(14)	หอก	spear	κη□□ψη	κη□□τ	*κη□□σ	(KM: πηλ→⇔↔ψη)
(15)	เห็ด	mushroom	τηιψη	τηιτ	*τηισ	
(16)	หวี (ก.)	to comb	κααψη	καατ	*καασ	(KM: χηρ→⇑↔ψη)
(17)	อีเก้ง	barking deer	πη□□ψη	πη□□τ	*πη□□σ	(KM: πυ⇔↔ψη; MB: πoolη)

Table 3 Proto-Mal *-s

	Thai gloss	*English gloss*	*Mal*	*Proto-Mal*	*Pray*
(18)	กระทืบ	to stamp one's foot	βι↔ψη	*βι↔σ	(⟩χετ)

(19)	ดี (น.)	gall bladder	ντηαψη	*ν-τηασ	(βιι)	
(20)	ตั่ง	stool, bench	νυψη	*νυσ	(ταN)	
	Thai gloss	*English gloss*	*Mal*	*Proto-Mal*	*Pray*	
(21)	ตะแกรง	screen for sifting	γψ↔ψη	*N-κρ↔σ	(Nκι↔N)	
(22)	บวม	to swell	/αψη	*/ασ	(πηυυ, π□□N)	(KM: /α≅ψη)
(23)	เป็น	to be	καψη	*κασ	(πεν)	
(24)	ร่อน	to sift	⟩↔ψη	*⟩↔σ	(κι↔N)	
(25)	ลั่น (ไก)	to pull the trigger	κλαψη	*κλασ	(βιιπ)	(KM: ρλ→∃ψη)
(26)	ลาน (บ้าน)	courtyard, lawn	λααψη	*λααασ	(κηοN τε/)	
(27)	สันเขา	mountain ridge	κηοοψη	*κηοοσ	(νυαN)	
(28)	ส่าย (หัว)	to sway (the head)	⟩↔ψη	*⟩↔σ	(N⟨ψ)	
(29)	แสงแดด	sunlight	/↔ψη	*/↔σ	(⟩χηαN)	
(30)	แสบ	stinging pain	σαψη	*σασ	(σαNααρ)	(MB: β⟨ρσαλη)
(31)	หวี (น.)	comb	νααψη	*νααασ	(Nκραπ)	(KM: ντηρ→⌈↔ψη)
(32)	เหว	chasm, abyss	τ↔ψη	*τ↔σ	(β↔N-β↔N)	
(33)	อับ (ว.)	unventilated	/□□ψη	*/□□σ	(/οπ)	

Many Proto-T'in forms having *-h can be reconstructed with more certainty because *-h has been kept very well in both Modern Mal and Modern Pray as shown in Table 4.

Table 4 Proto-T'in *-h

	Thai gloss	*English gloss*	*Mal*	*Pray*	*Proto-T'in*
(34)	กรอบ, เปราะ	crisp, brittle	Nκηαη	Nκηαη	*N-κηαη
(35)	เกลียด	to hate	σεη	σεη	*σεη
(36)	เกา	to scratch	κιη	κιη	*κιη

(37)	แก่, เฒ่า	old (of people)	βψEη	μπρEη	*μ-πρEη	
(38)	ใกล้	near, close	ⱷοη	ⱷυη	*ⱷυη	
(39)	ขวาก	spikes (of traps)	γψαη	Nκραη	*N-κραη	(KM: σρα≅η)
	Thai gloss	*English gloss*	*Mal*	*Pray*	*Proto-T'in*	
(40)	เขา (พวก~)	they	/αη	/αη	*/αη	
(41)	เขียง	chopping board	νοη	νοη	*νοη	(KM: ννο∃οη)
(42)	ไข, เปิด	to open	πηι↔η	πηι↔η	*πηι↔η	(MB: κρ∝∝η)
(43)	เครื่องดนตรี	musical instrument	πιη~πλιη	πλιη	*πλιη	
(44)	เธอ, มึง	you	μαη	μαη	*μαη	
(45)	เงย	to raise (one's head)	N□η	N□η	*N□η	(KM: N↔∃↔η)
(46)	จมูก	nose	μοη	μυη	*μυη	(KM: μυ∃η; MB: μ□η)
(47)	จุด (ไม้ขีด)	to light (matches)	τηαψη	τηEη	*τηEη	(MB: κλEη)
(48)	ซักผ้า	to wash clothes	πηυη	πηυη	*πηυη	(KM: πυ≅η; MB: σ↔πυη)
(49)	ดอย, ภูเขา	mountain	μβλ□η	λ□η	*μ-πλ□η	(MB: χ↔βοη)
(50)	ด้าน, เบื้อง	side	πηι↔η	πηι↔η	*πηι↔η	(KM: μπηλα∃η)
(51)	เดือด	boiling	ποη	μποη	*μ-ποη	(KM: ντηρο∃η)
(52)	ตก	to drop, to fall	κηλιη	κηλιη	*κηλιη	(KM: κασε≅η)
(53)	ตบ	to clap, to slap	πηοη	μπηαη	*μ-πη( )η	(KM: ντα≅η)
(54)	ตื่น	awake	ψ□η	ρ□η	*ρ□η	(KM: ρ↔∃η; MB: π∝∝η)
(55)	เตา	fireplace	πυ↔η	πυαη	*πυαη	(MB: β□η)
(56)	ถอน	to pull up	τηοη	τηοη	*τηοη	(KM: τρ↔≅η;

						MB: ψοοη)
(57)	นิ้ว	finger	νυ↔η	νυ↔η	*νυ↔η	
(58)	นึ่ง	to steam	/οη	σι/οη	*σ-/οη	(KM: ηο≅η-/ο≅η)
(59)	บอก, เล่า	to tell	δυ↔η	ντυαη	*ν-τυαη	(KM: ρτη□∃η)
(60)	บาน	to bloom	ψαη	ραη	*ραη	(KM: Nκηρα≅η; MB: πλαη)
	Thai gloss	*English gloss*	*Mal*	*Pray*	*Proto-T'in*	
(61)	ปลุก	to wake (someone up)	πηψ□η	πηρ□η	*πηρ□η	(KM: πηρ↔≅η; MB: παπ∝∝η)
(62)	ผลัก	to shove	⟩□η	σα⟩□η	*σ-⟩□η	
(63)	ผ่า	to split (firewood)	πηοη	πηοη	*πηοη	(MB: ποη)
(64)	ผ้าห่ม	blanket	⟩αη	⟩αη	*⟩αη	
(65)	ผืน	clf. for cloth, mat	πηλαη	πηλαη	*πηλαη	(MB: κ↔πλαη)
(66)	มา	to come	λεη	λεη	*λεη	(KM: λε∃η 'near')
(67)	เมีย	wife	/ι↔η	σι/ιαη	*σ-/ιαη	
(68)	ยาว	long	χυ↔η	χυ↔η	*χυ↔η	
(69)	ร้อน	hot	/□η	σα/□η	*σ-/□η	
(70)	ราง (อาหาร)	trough	μπη□η	μπη□η	*μ-πη□η	
(71)	ล้าน (หัว~)	bald	κηλΕη	κηλΕη	*κηλΕη	
(72)	ลืม (ตา)	to open (one's eyes)	βαη	μπαη	*μπαη	
(73)	เลียงผา	mountain goat	κηΕη	κηΕη	*κηΕη	(KM: κε∃ψη; MB: κΕη)
(74)	สว่าง	bright	παη	παη	*παη	(KM: πηα∃η)
(75)	สุก	ripe	κηρΕη	κηρΕη	*κηρΕη	
(76)	สูง	tall, high	τη⟩η	ντη⟩η	*ν-τη⟩η	
(77)	ใส่	to put in/on	σιη	σιη	*σιη	(KM: σε≅η)

(78)	หนัก	heavy	χ\η	χ\η	*χ\η	
(79)	หนัง	skin, leather	τηυ↔η	σατηυαη	*σ-τηυαη	(MB: γυ-γυη)
(80)	หน้าผาก	forehead	ταη	σαταη~νταη	*σν-ταη	(KM: ντηα∃η)
(81)	หว่าน	to sow	σαη	σαη	*σαη	
(82)	เห่า	to bark	ψοη	ροη	*ροη	
(83)	แหว่ง, บิ่น	chipped, nicked	ωεη	ωεη	*ωεη	(KM: ωE∃η; MB: πεεη)
(84)	อ้า (ปาก)	to open (one's mouth)	ωαη	⟩αη	*( )αη	(MB: ψεεη 'to spread apart')

Discussion

As stated earlier in the introductory part, this paper is only a product of our macro-research project on "Linguistic Diversity in Nan Province: A Foundation for Tourism Development". Four Mon-Khmer languages are spoken in Nan: Mal, Pray, Khmu' and Mlabri. However, only three ethic groups have been recognized by Nan people, i.e., T'in or Lua', Khmu' and Mlabri. There are many varieties of the Nan Lua' language. The reconstruction of Proto-T'in phonology by Filbeck (1978) is good; however, there is a lack of Proto-T'in lexicon. This paper is only a small contribution for a more sophisticated reconstruction of Proto-Khmuic by professional Austroasiatic or Mon-Khmer comparativists.

One more interesting thing that should be pointed out is that Mlabri cognates have three groups of fricative-like sounds: a. /-s ~ -yh/, b. /-lh [λ88 ~ ℜ]/ and c. /-h/, see examples (2), (3), (7), (8), (9), (10), (12), (17), (30), (42), (46), (47), (48), (49), (54), (55), (56), (60), (61), (63), (65), (73), (79), (83) and (84). The a. and b. groups correspond with Proto-T'in *-s and the c. group always corresponds with Proto-T'in *-h. This phenomenon suggests the idea that perhaps three final fricatives should be reconstructed at the Proto-Khmuic level.

Acknowledgement–The author would like to express her deep gratitude to the Thailand Research Fund (TRF) for funding the three-year research project on "Linguistic Diversity in Nan Province: A Foundation for Tourism Development". A lot of thanks and appreciation go to the Nan Community and the local authorities of Nan Province. The author also feels very grateful for the kind cooperation of her Mal, Pray, Khmu' and Mlabri informants. Last but not least, without the devotion of a group of research assistants, graduate students in the Linguistics Department at Chulalongkorn University, the Nan Project could not have been possible.

References

Filbeck, David. 1972. "Tone in a dialect of T'in". *Anthropological linguistics*, vol.14, no.4, pp. 111-117.

Filbeck, David. 1978. *T'in: a historical study*. Pacific linguistics series B, no.49. Canberra: The Australian National University.

Filbeck, David. 1991. *Mal-Thai-English dictionary*. Chiangmai: Phayap University.

Jirananthanaporn, Supatra. 1993. *Linguistic geography of minority groups in Nan province: a lexical study of Pray and Mal dialects*. Phitsanulok: Naresuan University. (in Thai with English abstract)

Ratanakul, Suriya. 1975. "A study of Thin: a Mon-Khmer linguistic community in Thailand". Paper presented at the 8th International Conference in Sino-Tibetan Languages and Linguistics, organized by University of California at Berkeley, October 24-26, 1975.

Satayawadhana, Cholthira. 1987. *The Lua' of Nan province*. Bangkok: Mueang Boran Press. (in Thai)

Singnoi, Unchalee. 1988. A comparative study of Pray and Mal phonology. Unpublished M.A. thesis, Mahidol University.

Studies in Tai and Southeast Asian Linguistics

Edited by
Jimmy G. Harris
Somsonge Burusphat
James E. Harris

2007
ISBN 978-974-8130-06-4

The Interaction between Vowel Length and Pitch in Southeast Asian (SEA) Languages: Implications for Tonogenesis

Theraphan L-Thongkum, Phanintra Teeranon and Chommanad Intajamornrak

1. Introduction

When pronouncing some Bangkok Thai words spelled with the low class consonant letters of the Thai (Siamese) writing system, one has to be careful not only about the pronunciation of the tone but also of the vowel length. The tone mark "?èek" signals both the falling tone and the shortening of vowels, as in the examples given below. See especially the middle column.

no tone mark (mid tone)		?èek *tone mark (falling tone)*		*thoo tone mark (high tone)*	
/thɛɛŋ/	'to stab'	/thɛŋ/	'clf. for long objects'	/thɛ́ɛŋ/	'miscarried'
/khɛɛn/	'reed pipe'	/khên/	'to do unwillingly'	/khɛ́ɛn/	'to crave for revenge'
/rɛɛŋ/	'strength'	/rêŋ/	'to sieve'	/rɛ́ɛŋ/	'vulture'
/thɔɔŋ/	'gold'	/thôŋ/	'to learn by heart'	/thɔ́ɔŋ/	'belly'
/yɔɔm/	'subdued'	/yôm/	'reduced, decreased'	/yɔ́ɔm/	'to dye'
/rɔɔŋ/	'to support'	/rôŋ/	'cracks in the ground'	/rɔ́ɔŋ/	'to utter a cry'

The pronunciation of the vowels in the above Thai words will be [ɛɛ] or [ɛ] and [ɔɔ] or [ɔ] depending upon the Thai tone marks. With the "?èek" tone mark, the vowels must be [ɛ] or [ɔ] not [ɛɛ] or [ɔɔ]. In the speech of younger speakers, the phonetic characteristics of the mid tone, falling tone and high tone are [33], [42] and [34], respectively. Before the falling contour, the pitch height of the falling tone is the highest among the five Thai tones. It is also worthwhile pointing out that only *ɛɛ and *ɔɔ were reconstructed among the Proto vowels in the vowel system of Proto-Southwestern Tai, the branch to which Bangkok Thai belongs, by Li (1977), Sarawit (1973) and Jonsson (1991). The short vowels /ɛ/ and /ɔ/ are fairly recent phonological members in comparison with the long vowels /ɛɛ/ and /ɔɔ/.

Gandour (1977: 55) points out that the loss of phonological distinction in vowel length among certain Thai dialects has been conditioned by tone. The changes in the phonological status of vowel length reflect universal phonetic tendencies in the interaction between tone and vowel length. These phonetic tendencies are evident in several geographically and typologically distinct tone languages. Findings from a number of tone languages, e.g., Mandarin, Cantonese, Punjabi, Zapotec, Soyaltepec Mazatec, Chatino, Tenango Otomi, Kutchin and so

on, indicate the fact that vowels with high tone are shorter than those with low or mid tones (Gandour, 1977: 59). Similar types of phenomena can be found in Sre and Korean (Fischer-Jørgensen, 1990).

Svantesson (1991) claims that in Hu, the two tones, i.e., low vs. high, have developed in conjunction with the loss of vowel length. A comparison between two Palaungic languages: Hu which has no vowel length distinction, and Lamet, which has vowel length distinction in its phonological system, suggests the idea that Proto-Palaungic short vowels remained short but long vowels began to lose their length and to merge with the original short vowels in Middle Hu. The distinction of vowel length in Proto-Palaungic has been lost in Modern Hu but has been kept in Modern Lamet. To compensate for the loss of vowel length, two distinctive pitches or two tones have developed in Hu as shown in the following examples (Svantesson, 1991: 72).

Proto-Palaungic	*Hu*	*Lamet*	
*a	yám	yàm	'to die'
*aa	yàm	yàam	'to cry'

Svantesson thinks that this tonal phenomenon in Hu is unorthodox. In fact, his discovery is not surprising. From the findings reported by Gandour (1977), Fischer-Jørgensen (1990) and our research team at Chulalongkorn University, the interaction between vowel length and pitch which has caused tonal evolution (non-tonal > tonal) and tonal development (fewer tones > more tones) does not seem to be unorthodox as stated by Svantesson (1991). There is a natural phonetic tendency for short vowels, e.g., i, a, u, etc., to have higher intrinsic pitches than their counterpart long vowels ii, aa and uu, respectively. L-Thongkum (1989 and 1991) has also found that in Kui (Suai) and Chong which are register Mon-Khmer languages, phonological short vowels have higher fundamental frequencies than phonological long vowels in both register complexes. This type of phonetic tendency has also been confirmed by the acoustical measurement of phonetic short and long vowels in two Malay dialects spoken in Thailand (Lohde, 2004). Even though these non-tonal Malay dialects do not have vowel length distinction, the interaction between vowel length and pitch seems to work in the same way as that in tonal and register languages having vowel length distinction.

However, the information mentioned above is only a by-product of previous acoustic studies focusing on the other phonetic and phonological aspects. A serious acoustic investigation of the interaction between vowel length and pitch should be conducted. The aim of this paper is to attest scientifically by means of acoustic analysis whether it is plausible that the interaction between the length and pitch of vowels can be a cause of tonal evolution and tonal development in both non-tonal and tonal Southeast Asian (SEA) languages. To

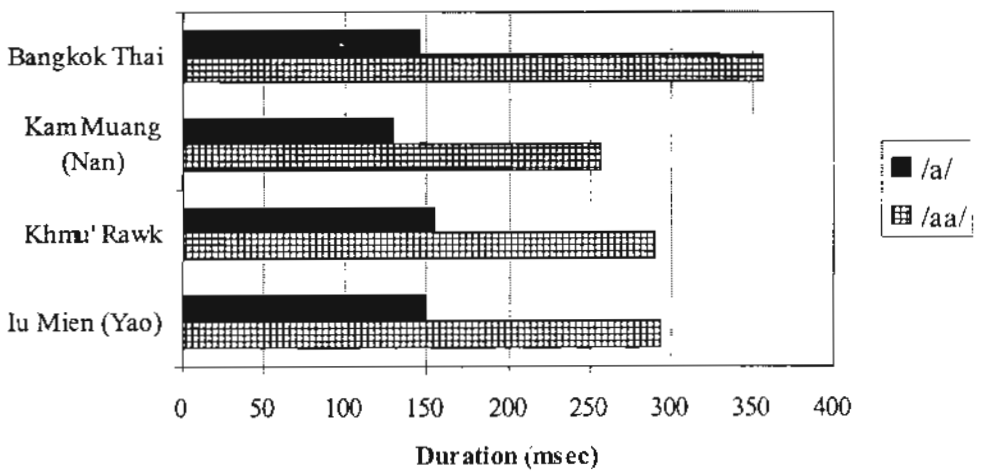
test this hypothesis, the duration, fundamental frequency and intensity values of /a/ and /aa/ in four SEA languages, i.e., Bangkok Thai, Kam Muang, Khmu' Rawk and Iu Mien were measured. The reasons for selecting the test words containing /a/ and /aa/: are a.) in SEA languages having phonological vowel length, in comparison with the other vowel pairs, a large number of minimal pairs with /a/ and /aa/ can be found; b.) in languages which generally do not have vowel length distinction, there is a tendency for [a] and [aa] to become two distinctive vowels in their phonological systems; and c.) in languages which have been losing their phonological length due to contact with more prestigious languages spoken in the same area, the distinction in length between /a/ and /aa/ will be retained longer.

2. Language data

Bangkok Thai was chosen because it is the national language of Thailand spoken widely all over the country. Kam Muang (Tai Yuan), Khmu' Rawk and Iu Mien (Yao) are spoken in Nan and other provinces in the North of Thailand. Kam Muang is a regional Thai dialect whereas Khmu' Rawk (Mon-Khmer) and Iu Mien (Hmong-Mien or Miao-Yao) are regarded as minority languages. All of these four languages have distinctive vowel length in their phonological systems (see Table 1 and Figure 1). The Kam Muang, Khmu' Rawk and Iu Mien data used for this research were drawn from the field notes of the first author collected in Nan Province from May 2004 to March 2005 for the research project "Linguistic Diversity in Nan Province: A Foundation for Tourism Development" sponsored by The Thailand Research Fund (TRF). Regarding Bangkok Thai, five graduate students of the Linguistics Department at Chulalongkorn University kindly acted as our informants.

Language	/a/	/aa/	Diff.	Ratio
Bangkok Thai	145.76	356.60	210.84	2.45
Kam Muang (Nan)	128.56	256.85	128.29	1.99
Khmu' Rawk	154.21	290.68	136.47	1.88
Iu Mien (Yao)	149.55	293.54	143.99	1.96

Table 1: The average durations in ms and the duration ratio of /a/ and /aa/ in four SEA languages: Bangkok Thai, Kam Muang (Nan), Khmu' Rawk and Iu Mien



Figures 1: The average duration in ms of /a/ and /aa/ in four SEA languages: Bangkok Thai, Kam Muang (Nan), Khmu' Rawk and Iu Mien

Bangkok Thai (BT), Kam Muang (KM) and Iu Mien (IM) are tonal languages, whereas Khmu' Rawk (KR) is a register language which is on its way to becoming a tonal language with two tones, high vs. low. For each test language, five female speakers were used. Their average age was 28.45 and the age range was from 20 to 41 years old (BT: 20-29, KM: 23-28, KR: 31-41, IM: 21-30).

Language BT: The BT-wordlist comprises 30 monosyllabic words, 15 words contain /a/ and the other 15 contain /aa/. All of these test words have the BT low tone /21/ and voiceless obstruent finals. They were arranged into 15 minimal pairs, e.g., /tak^{21}/ 'to dip up' - /taak^{21}/ 'to expose in the sun'; /pak^{21}/ 'to embroider' - /paak^{21}/ 'mouth', and so forth. See details in Appendix 1.

Language KM: The KM-wordlist consists of 20 monosyllabic words, 10 words contain /a/ and the other 10 contain /aa/. All of the 20 test words have the KM mid tone /33/ and sonorant finals. The test words were arranged into 10 minimal pairs, e.g., /taŋ^{33}/ 'stool' - /taaŋ^{33}/ 'different'; /khay^{33}/ 'egg' - /khaay^{33}/ 'net' and so on. See details in Appendix 2.

Language KR: The KR-wordlist comprises 20 monosyllabic words, 10 words contain /a/ and the other 10 contain /aa/. The 20 selected test words have a clear register (modal voice) and voiceless obstruent finals. They were arranged into 10 minimal pairs, e.g., /kát/ 'cold' - /káat/ 'market'; /káp/ 'mouse trap' - /káap/ 'chin' and so forth. See details in Appendix 3.

Language IM: The IM-wordlist comprises 20 monosyllabic words, 10 words contain /a/ and the other 10 contain /aa/. All of these test words have the Iu Mien mid tone /33/ and sonorant finals. They were arranged into 10 minimal

pairs as with the other three wordlists mentioned above. The following are some examples: /kan^{33}/ 'to follow, to call for' - /kaan^{33}/ 'fishing rod'; /hlaŋ^{33}/ 'high, tall' - /hlaaŋ^{33}/ 'string, cord, rope' and so on. See details in Appendix 4.

3. Procedure

All of the informants (20 female speakers) were recorded at different research sites and at different times. The BT test words were recorded in the studio of the Center for Research in Speech and Language Processing (CRSLP) at the Faculty of Arts, Chulalongkorn University, in May 2004. The KM language material used for our acoustic studies was recorded in a quiet corner of the library of Sattreeseenan High School, Muang District, Nan Province, in September 2004. The KR test words were recorded at a quiet spot in an open area of Huay Sataeng Village, Thung Chang District, Nan Province, in September 2004. As for the IM language, the test words were recorded at the Center for Tribal Welfare and Development, Pa Klang Field Station, Pua District, Nan Province, in March 2005. To achieve a satisfactory conclusion for the interaction between vowel length and pitch which can be a cause of tonal evolution, the following equipment and programs were used: a Sony IC Recorder ICD-MS515; the Praat program version 4.2.09 for acoustic analysis; a supplemental program for extracting fundamental frequency, formant frequency, duration and intensity prepared by Patavee Charnvivit; Cool Edit Pro for assisting segmentation; Microsoft Excel 2003 for analyzing and plotting Figures of the means ($\bar{x}$) of F_0, duration and intensity values in Hz, ms and dB, respectively; and the SPSS program for statistical analysis.

The five speakers of each language were instructed and had the opportunity to practise before recording. Each speaker was recorded separately in turn. Whilst the recording was taking place, the meaning of each test word was given as a clue by our research assistant, then the informant said the target test word in her language, once. They were instructed to pronounce each word naturally at a moderate tempo. The recording of each word was redone when unusual pronunciation was detected. Three recordings of the whole set of test words were done for each speaker. Different numbers of the test tokens were used as the input for acoustic analysis: 450 test tokens for BT (30 test words x 5 speakers x 3 times) and 300 for KM, KR and IM (20 test words x 5 speakers x 3 times for each language). Thus, the total number of test tokens used for the acoustical measurements in this present study was 1,350. The duration of /a/ and /aa/ in the four languages examined was measured from the onset to the offset in milliseconds (ms). Regarding the fundamental frequency (F_0) and intensity values of /a/ and /aa/, the frequency and intensity at five points of time for each vowel were selected for measurement, i.e., at 0% (vowel onset), 25%, 50%, 75% and 100% (vowel offset).

4. Results

4.1 Language BT

The average durations of /a/ and /aa/ which represent short and long vowels in BT is 145.76 ms and 356.60 ms, respectively, with a difference of 210.84 ms and a ratio of 2.45. The result of this acoustic analysis agrees quite well with those presented in Abramson (1962), Gandour (1984) and Intajamornrak (2002). The fundamental frequencies of /a/ and /aa/ can be found in Table 2 and Figure 2. As expected, /a/ has a higher F_0 than /aa/.

Duration	0%	25%	50%	75%	100%
/a/	206.25	191.76	178.44	165.90	157.22
SD	20.11	16.14	13.65	13.05	13.53
/aa/	201.62	182.52	167.09	155.00	149.57
SD	16.54	13.55	13.21	10.63	21.13
Diff.	4.63	9.24	11.35	10.90	7.65
t-test	Sig	Sig	Sig	Sig	Sig
	$p < 0.05$				

Table 2: The average F_0 values in Hz of /a/ and /aa/ in the BT test words having Tone /21/ (5 female speakers)

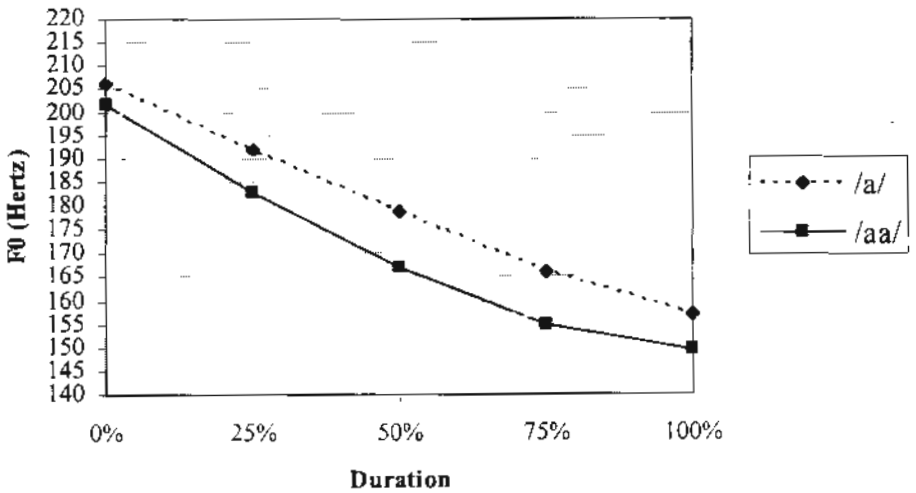


Figure 2: The average F_0 contours of /a/ and /aa/ bearing the BT Tone /21/ (5 female speakers)

It can be seen in Table 2 and Figure 2 that /a/ representing the short vowels and /aa/ representing the long vowels in Bangkok Thai (BT) have different f_0 values at every point of measurement, i.e., at the 0% (vowel onset), 25%, 50%, 75% and 100% (vowel offset) of their durations and that the f_0 value

difference at each point is statistically significant. This means that the pitch of /a/ is higher than the pitch of /aa/ although /a/ and /aa/ carry the same BT tone.

4.2 Language KM

The vowels /a/ and /aa/ in Kam Muang (Nan) have average durations of 128.56 and 256.85 ms, respectively. The duration difference between /a/ and /aa/ which represents the KM short and long vowels is 128.29 ms with a ratio of 1.99. Regarding the fundamental frequency, /a/ has a higher F_0 value than /aa/ as shown in Table 3 and Figure 3.

Duration	0%	25%	50%	75%	100%
/a/	214.58	209.87	206.14	203.22	199.58
SD	18.03	16.26	15.44	15.41	17.44
/aa/	214.98	206.90	202.27	198.73	196.43
SD	18.31	15.56	15.63	16.76	17.44
Diff.	-0.40	2.97	3.87	4.49	3.15
t-test	-	Sig	Sig	Sig	Sig
	$p < 0.05$				

Table 3: The average F_0 values in Hz of /a/ and /a/ in the KM test words having Tone /33/ (5 female speakers)

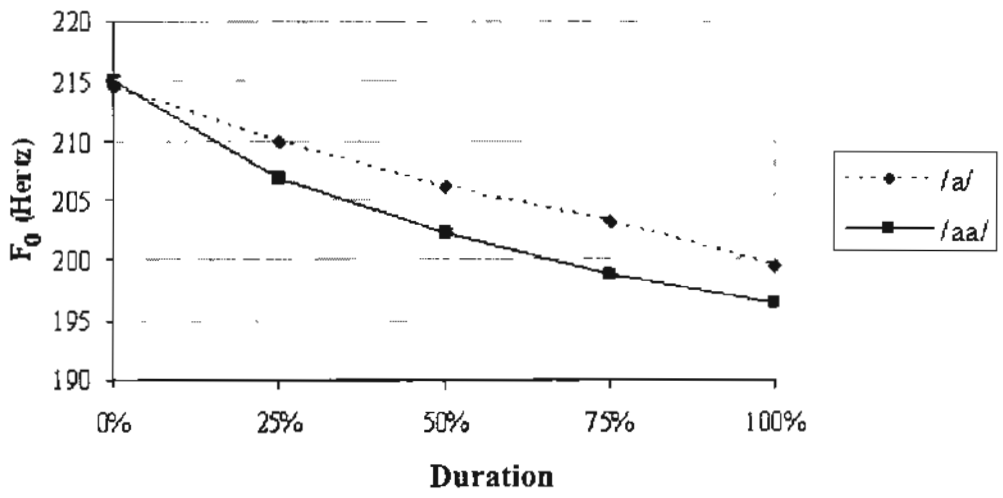


Figure 3: The average F_0 contours of /a/ and /aa/ bearing the KM Tone /33/ (5 female speakers)

Table 3 and Figure 3 suggest the fact that similar to BT, generally speaking, the KM /a/ has a higher pitch than /aa/. At four time points, i.e., 25%, 50%, 75% and 100% (vowel offset), the F_0 values of /a/ are higher than those of /aa/, and the F_0 difference at each of these four points is statistically significant.

The phenomenon found in KM supports well the hypothesis that short vowels tend to have a higher pitch than long vowels.

4.3 Language KM

The Khmu' language comprises many dialects, Khmu' Rawk is one of them. Some dialects have a two-way register contrast (modal voice vs. breathy voice) and some have two tones (high vs. low). The voiced series of Proto-Khmu' initials *b *d *j *g have become devoiced in most of the Modern Khmu' dialects, i.e., p t c k or ph th ch kh. In the Khmu' Rawk dialect spoken in Huay Sataeng Village, the proto voiced plosives have become the voiceless aspirated plosives ph th ch kh (ongoing research of the first author).

Even though KR has register distinction, i.e., two types of register complex (modal voice and higher pitch vs. breathy voice and lower pitch), due to the fact that it is becoming a tonal language, phonation type difference is fading out while pitch difference is becoming more prominent. The vowels /a/ and /aa/ have both modal voice and breathy voice in Khmu' Rawk. However, only the modal or clear voice /a/ and /aa/ were chosen for the acoustical measurements of duration, fundamental frequency and intensity.

The average duration of the KR /a/ is 154.21 ms and of the /aa/ is 290.68 ms, with a difference of 136.47 ms, the ratio being 1.88. Regarding the F_0 difference, the /a/ has a higher F_0 value than the /aa/ at all points of measurement, i.e., at 0% (vowel onset), 25%, 50%, 75% and 100% (vowel offset) of the normalized time. The F_0 difference at each point is statistically significant as shown in Table 4. The difference between the F_0 contours of /a/ and /aa/ in KR can be found in Figure 4.

Duration	0%	25%	50%	75%	100%
/a/	217.05	214.16	219.55	230.31	234.30
SD	16.73	10.82	11.42	14.08	17.65
/aa/	203.22	205.24	215.60	221.25	220.42
SD	12.21	10.31	11.39	13.69	15.14
Diff.	13.83	8.92	3.95	9.06	13.88
t-test	Sig	Sig	Sig	Sig	Sig
	$p < 0.05$				

Table 4: The average F_0 values in Hz of the modal voice /a/ and /aa/ in the KR test words (5 female speakers)

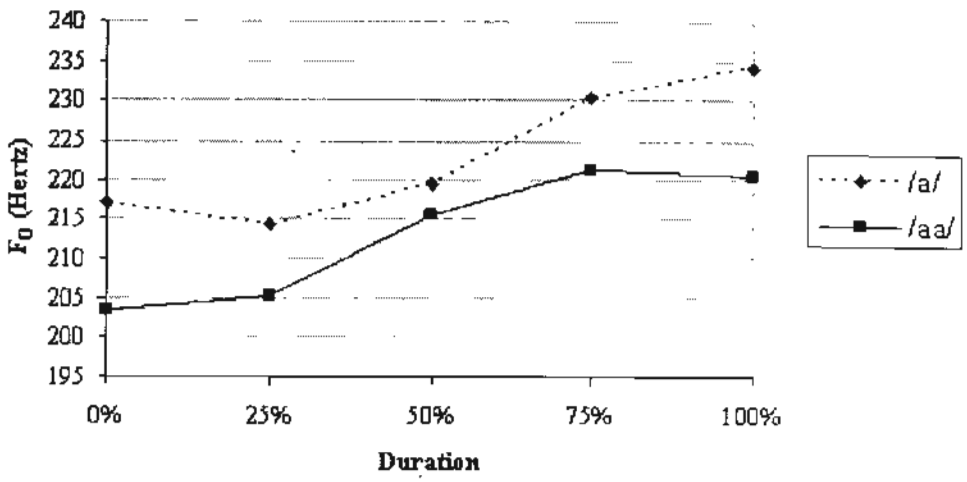


Figure 4: The average F_0 contours of the modal voice /a/ and /aa/ in KR (5 female speakers)

The claim that short vowels have a higher pitch than long vowels can be confirmed by the results of the acoustical measurements presented in Table 4 and Figure 4. The vowels /a/ and /aa/ in a register language like Khmu' Rawk do have different pitches, i.e., high vs. low, respectively.

4.4 Language IM

The duration and F_0 values of /a/ and /aa/ having Tone /33/ in Iu Mien were measured. The average durations of /a/ is 149.55 ms and of /aa/ is 293.54 ms, with a difference of 143.99 ms and a ratio of 1.96. The differences at five measured time points, i.e., at 0% (vowel onset), 25%, 50%, 75% and 100% (vowel offset), are statistically significant as shown in Table 5. The F_0 contours of /a/ and /aa/ can be found in Figure 5.

Duration	0%	25%	50%	75%	100%
	254.08	247.34	245.29	243.60	241.41
SD	24.52	22.50	21.97	22.05	21.74
/aa/	251.58	244.60	240.06	237.03	234.64
SD	23.48	21.91	20.99	20.68	20.85
Diff.	2.50	2.74	5.23	6.57	6.77
t-test	Sig	Sig	Sig	Sig	Sig
	$p < 0.05$				

Table 5: The average F_0 values in Hz of /a/ and /aa/ in the IM test words having Tone /33/ (5 female speakers)

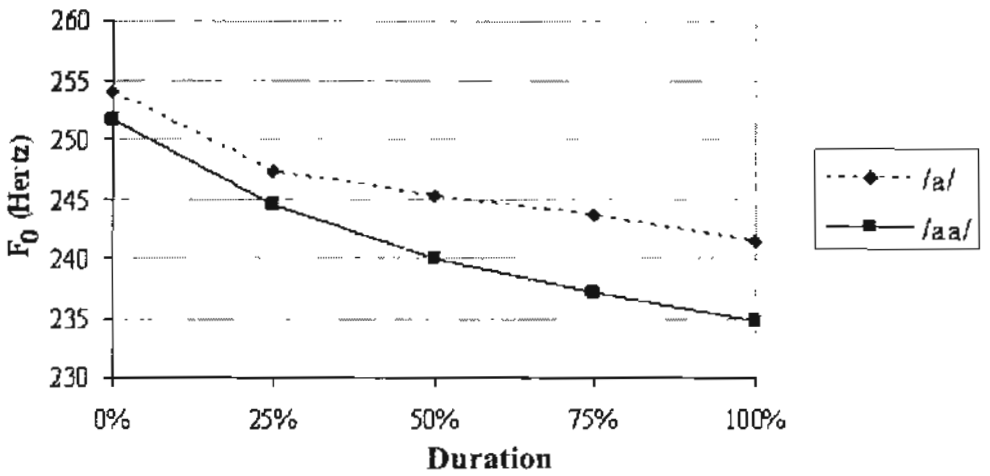


Figure 5: The average F_0 contours of /a/ and /aa/ bearing the IM Tone /33/ (5 female speakers)

The details of the F_0 values given in Table 5 and the F_0 contours shown in Figure 5 imply the fact that the IM /a/ has a higher pitch than the /aa/, the same type of phenomenon that occurs in Bangkok Thai, Kam Muang (Nan) and Khmu' Rawk.

Our findings, based on the results of measurements of the duration and fundamental frequencies of /a/ and /aa/ in four SEA languages: Bangkok Thai, Kam Muang (Nan), Khmu' Rawk and Iu Mien, confirm the idea that the pitch of short vowels is higher than that of long vowels. The interaction between vowel length and pitch seems to be a natural phonetic tendency, at least in the SEA languages examined.

5. Discussion

Even though the /a/ in every test language has a higher F_0 value than the /aa/, and the F_0 difference is significant ($p < 0.05$), it is noticeable that in Khmu' Rawk (KR), a register language, the F_0 difference between the /a/ and the /aa/ seems to be larger than those in the other three tonal languages, BT, KM and IM. See details in Table 6. Since the KR dialect of Huay Sataeng village is becoming a tonal language (modal voice > /high/ and breathy voice > /low/), perhaps, KR speakers are more conscious of pitch difference, a newly acquired feature in their phonological system.

	Duration				
f_0 diff.	0%	25%	50%	75%	100%
BT	4.63	9.24	11.35	10.90	7.65
KM	-0.40	2.97	3.87	4.49	3.15
KR	13.83	8.92	3.95	9.06	13.88
IM	2.50	2.74	5.23	6.57	6.77

Table 6: The F_0 differences in Hz between /a/ and /aa/ in BT, KM, KR and IM

When linguists discuss tonogenesis, tonal evolution or tonal development, we tend to think first of the birth of tonal distinction from initial consonants (voiceless > /high/ and voice > /low/), final consonants (voiceless glottal fricative > /falling/ and glottal stop > /rising/), the voice qualities of vowels (modal voice > /high/ and breathy voice > /low/) and so on, as in a recent survey on "The plausibility of phonetic explanations of tonogenesis" by Abramson (2004). Few of us pay attention to the other important prosodic features which can cause tone birth, e.g., consonant length, vowel length, stress and so on.

With a wider perspective, the loss of distinctive vowel length which has caused the tonal evolution in the Hu language is not unorthodox because the interaction between vowel length and pitch has been attested, i.e., short vowels tend to have higher intrinsic pitches than long vowels in both tonal and non-tonal languages. These natural phonetic tendencies should not be ignored. Therefore, the question that should be asked is: "Why is this so?" Fischer- Jørgensen, (1990: 134) offers the following explanation:

> "One possibility might be that the intensity of short vowels is increased by heightened subglottal pressure in order to compensate for their short duration, and a higher subglottal pressure might also increase the F_0, although not much can be said in favour of this hypothesis.
>
> Another possibility might be an increased F_0 as compensation for the shortness. The higher F_0 might make the vowels stand out more clearly. The differences are sufficiently large to be perceived. This is, however, still less probable than a compensation in intensity. And a problem in both cases is that very little seems to be known about subglottal pressure in short and long vowels."

At present, since we still do not have any explicit and convincing answers to our question, perhaps the best we can do is to assume tentatively that a shorter duration is less prominent than a longer one. To help short and long vowels have equal prominence so that the listeners can differentiate them easily, pitch is assigned as an assistant to length. Generally, since a higher pitch is more prominent and well perceived than a lower pitch, then, the former is assigned to

vowel with a shorter duration and the latter to the one with a longer duration. In terms of perception, this natural phonetic tendency may be regarded as a good compensation. It has a more important role to play when vowel length is lost and tones are born. Thus, the birth of high tone vs. low tone in the Hu language is not unorthodox.

Sound A can be made more prominent than sound B by adjusting its length (longer duration), pitch (higher F_0) or loudness (higher amplitude or intensity). In languages which have distinctive vowel lengths, length must be kept apart as a constant feature. It cannot vary much due to its linguistic role in differentiating lexical meaning. The results of our acoustic investigation presented earlier in this paper confirm the idea that short vowels have higher f_0 values than long vowels. It is also interesting to ask whether amplitude or intensity has any role in making short and long vowels equally prominent. To answer this question, the intensity values of /a/ and /aa/ in the four languages investigated, BT, KM, KR and IM, were measured using the same sets of data.

In KM, KR and IM, /a/ has higher intensity values than /aa/ especially at the 25%, 50%, 75% and 100% (vowel offset) of the durations. The intensity difference at each of these four points is significant ($p < 0.05$). However, at a duration of 0% (vowel onset), the intensity difference is significant only in KR, a register language which is becoming a tonal language. See Table 7 and Figure 6.

Duration	0%	25%	50%	75%	100%
/a/	67.32	75.76	76.90	76.64	64.88
SD	4.40	3.66	2.75	3.24	4.90
/aa/	65.57	74.94	75.92	75.81	63.61
SD	5.19	5.45	4.88	4.76	5.07
Diff.	1.75	0.82	0.98	0.83	1.27
t-test	Sig	Sig	Sig	Sig	Sig
	$p < 0.05$				

Table 7: The intensity values in dB of /a/ and /aa/ in KR (5 female speakers)

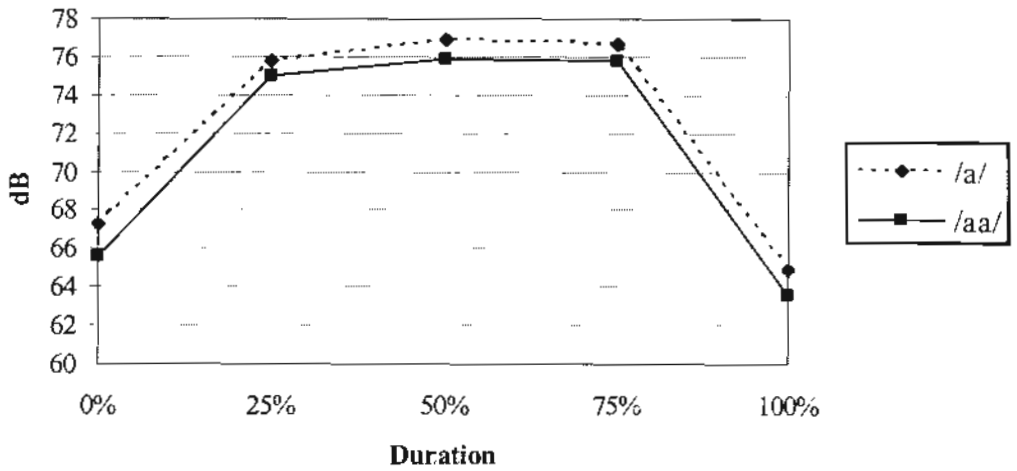


Figure 6: The intensity curves of /a/ and /aa/ in KR (5 female speakers)

It is important to point out that in BT, there are no consistent patterns of intensity difference as in the other SEA languages investigated. We suspect that this is due to the different ways of data elicitation. As for the BT data, the five graduate students of linguistics who acted as informants carefully read the minimal pairs containing /a/ and /aa/ in the wordlist. They were aware of the fact that we wanted good recordings done in a quiet studio for our acoustical measurements. However, we did not inform them about the purpose of our acoustic analysis. The KM, KR and IM informants said the words in their own languages after hearing the Thai words or phases with the same or similar meanings. They had no idea at all of what we were doing. However, they were willing to cooperate and to follow our instructions when the recordings were made at the research sites. Perhaps, they responded and behaved in a more natural way since the recording situation was closer to a situation when real verbal communication normally takes place.

6. Conclusion

Based on our findings, we can conclude that there is an interaction between vowel length and pitch, i.e. short vowels have higher intrinsic pitches than long vowels. It seems to be a natural phonetic tendency. Since length in some languages has a role to play in the phonological domain, only pitch (f_0) and loudness (intensity) have been left as phonetic features to help make short vowels as prominent as long vowels. Between pitch and loudness, pitch seems to have a more clear role so far as compensation is concerned, especially in languages which have a distinctive vowel length. This helps pave the way for tonogenesis. When tones are born, vowel length can become a less important feature in some

languages, such as the Hu language of Yunnan. Therefore, the loss of phonological length can be regarded as a sacrifice to enable the birth of phonological pitch or tone. In other words, the birth of phonological pitch or tone is compensation for the loss of phonological length. Perception studies may help confirm our acoustic findings.

Acknowledgements

We would like to express our gratitude to The Thailand Research Fund (TRF) for sponsoring our three-year research project on "Linguistic Diversity in Nan Province: A Foundation for Tourism Development". Many thanks must go to our friends in Nan Province: Prachakhom Nan (Nan Community), the district authorities of Muang, Thung Chang and Pua, and our Kam Muang (Nan), Khmu' Rawk and Iu Mien informants. Last but not least, we appreciate the cooperation of the graduate students in the Linguistics Department of Chulalongkorn University.

References

Abramson, A. S. (1962). *The Vowels and Tones of Standard Thai: Acoustical Measurements and Experiments*. Indiana University Research Center in Anthropology, Folklore, & Linguistics, Pub. 20: Bloomington. [International Journal of American Linguistics 28. No.2 Part III.].

Abramson, A. S. (2004). The plausibility of phonetic explanations of tonogenesis. In G. Fant, H. Fujisaki, J. Cao, and Y. Xu (Eds.), *From Traditional Phonology to Modern Speech Processing: Festschrift for Professor Wu Zongji's 95th birthday* (pp. 17-29). Beijing: Foreign Language Teaching and Research Press.

Fischer-Jørgensen, E. (1990). Intrinsic f_0 in tense and lax vowels with special reference to German. *Phonetica, 47*, 99-140.

Gandour, J. (1977). On the interaction between tone and vowel length: evidence from Thai dialects. *Phonetica, 34*, 54-65.

Gandour, J. (1984). Vowel duration in Thai. *Crossroads: an Interdisciplinary Journal of Southeast Asian Studies, 2*(1), 59-64.

Intajamornrak, C. (2002). The acoustic characteristics of vowels produced by Thai tracheoesophageal and normal speakers and the perception of tracheoesophageal vowels. In T. L-thongkum (Ed.), *The Accent of Thai tracheoesophageal Speakers* (pp. 73-95). Bangkok: The Thailand Research Fund (TRF), printed by Chulalongkorn University House (in Thai).

Jonsson, N. L. (1991). Proto Southwestern Tai. Unpublished doctoral dissertation, The State University of New York at Albany.

L-Thongkum, T. (1989). An acoustic study of the register complex in Kui (Suai). *Mon-Khmer Studies,* 15, 1-19.

L-Thongkum, T. (1991). An instrumental study of Chong registers. In J. H. C.S. Davidson (Ed.), *Austroasiatic Languages: Essays in Honour of H.L. Shorto* (pp. 141-160). London: SOAS.

Li, F.-K. (1977). *A Handbook of Comparative Tai.* Honolulu: The University Press of Hawaii.

Lohde, K. (2004). A comparison of the interaction between the fundamental frequency and duration of vowels and final consonants in Pattani Malay spoken in the provinces of Pathumthani and Pattani: an acoustic study. Unpublished M.A. Thesis, Chulalongkorn University (in Thai).

Sarawit, M. E. S. (1973). The Proto-Tai vowel system. Unpublished doctoral dissertation. The University of Michigan at Ann Arbor.

Svantesson, J.O. (1991). Hu: A language with unorthodox tonogenesis. In J. H. C. S. Davidson (Ed.), *Austroasiatic Languages: Essays in Honour of H.L. Shorto* (pp. 67-79). London: SOAS.

Appendix 1: BT-wordlist

'suddenly'	(pup^{21}) pap ^{21}	'to conquer (praap^{21}>)'	paap^{21}
'to brush off'	pat^{21}	'to slice'	paat^{21}
'to embroider'	pak^{21}	'mouth'	paak^{21}
'liver'	tap^{21}	'kind of jewelry'	taap^{21}
'to cut'	tat^{21}	'kind of cloth'	taat^{21}
'to dip up'	tak^{21}	'to dry'	taak^{21}
'to catch'	cap^{21}	'to be rude'	(cuəŋ^{42}) caap^{21}
'to split (bamboo)'	cak^{21}	'from'	caak^{21}
'to arrange'	cat^{21}	'fish name'	caat^{21}
'mouse trap'	kap^{21}	'layers of banana trunk'	kaap^{21}
'to bite'	kat^{21}	'Chinese cabbage'	(phak^{21}) kaat^{21}
'to obstruct'	kak^{21}	'pieces of chaff'	kaak^{21}
'small box'	ʔap^{21}	'to bathe'	ʔaap^{21}
'to compress'	ʔat^{21}	'may be'	ʔaat^{21}
'to trap'	dak^{21}	'prolapse'	daak^{21}

Appendix 2: KM-wordlist

'stool'	taŋ^{33}	'different'	taaŋ^{33}
'to measure the depth of water'	jaŋ^{33}	'(this) way'	jaaŋ^{33}
'to weep continuously'	dan^{33}	'customs'	daan^{33}
'to cut'	han^{33}	'goose'	haan^{33}
'to spin'	pan^{33}	'hemp'	paan^{33}
'hilt of a knife'	kan^{33}	'white'	kaan^{33}
'chicken'	kay^{33}	'to cross over'	kaay^{33}
'egg'	khay^{33}	'net'	khaay^{33}
'shoulder'	lay^{33}	'other side'	laay^{33}
'to crawl'	tay^{33}	'rabbit'	taay^{33}

Appendix 3: KR-wordlist

'pulp found in fruit'	klák	'clump'	kláak
'duck'	pát	'to slice'	páat
'cold'	kát	'market'	káat
'to fish'	nták	'tongue'	ntáak
'to break'	pák	'big spoon'	páak
'trap'	káp	'chin'	káap
'sewing machine'	cák	'to tear'	cáak
'to adhere to'	ták	'to spit out'	táak
'to prick'	táp	'to repair clothes'	táap
'to hold in one's hand'	cáp	'kind of plate'	cáap

Appendix 4: IM-wordlist

'to cause to move'	pan^{33}	'generation (clf.)'	paan^{33}
'a moment ago'	(ʔa~) tɕaŋ^{33}	'neck'	tɕaaŋ^{33}
'to perform two duties'	tɕaw^{33}	'sticklac'	tɕaaw^{33}
'to follow'	kan^{33}	'fishing rod'	kaan^{33}
'sound of small bells'	(ɲiŋ^{33}) naŋ^{33}	'snake'	naaŋ^{33}
'high, tall'	hlaŋ^{33}	'cord, string, rope'	hlaaŋ^{33}
'earth, soil'	daw^{33}	'equal'	daaw^{33}
'to answer'	taw^{33}	'kind of palm'	taaw^{33}
'hook'	ŋaw^{33}	'cry of cats'	ŋaaw^{33}
'kind of fruit'	(ma~) tan^{33}	'alone'	taan^{33}

Voice Register in Khmu': Experiments in Production and Perception

Arthur S. Abramson[a b] Patrick W. Nye[a] Theraphan Luangthongkum[c]

Running head: Voice Register in Khmu'

[a] Haskins Laboratories, New Haven, Conn., USA

[b] Department of Linguistics, University of Connecticut, Storrs, Conn., USA

[c] Department of Linguistics, Faculty of Arts. Chulalongkorn University, Bangkok, Thailand

Abstract

Some Khmu' dialects have phonologically distinctive voice registers. Auditory observations have claimed a stable distinction between clear voice and high pitch for Register 1 and breathy voice and low pitch for Register 2 in the Khmu' Rawk dialect of northern Thailand. Word-pairs distinguished only by register were recorded by 25 native speakers. Acoustic analysis yielded F0 and overall amplitude contours, frequencies of F1 and F2 in quasi-steady states of the vowels, relative intensities of higher harmonics to that of the first harmonic, and vowel durations. When circumstances caused early attention to perception testing, the words of only eight speakers had been analyzed for properties other than amplitude and F0. Since the only significant factor that had emerged by then was F0 contour, the synthetic stimuli were made with just a series of seven contours. The labeling by 32 native speakers yielded two categories, demonstrating the sufficiency of F0 as an acoustic cue. The completed acoustic analysis showed a significant effect of one of the harmonic ratios for the women only, suggesting a conservative bias. The language has been shifting toward tonality and may have reached it.

I. Introduction

Voice register as a phonological unit in a number of languages is interesting for its phonetic complexity and variability, as well as for its possible role in tonogenesis. The concept of register itself was apparently first formulated by Eugénie Henderson [1952], who described it as a complex of laryngeal and supralaryngeal properties, one of which may be dominant. Traditional, largely impressionistic descriptions of register complexes have included such characteristics as voice quality (phonation type), vowel quality, vowel length, pitch, and perhaps others. Although for the most part,

such languages have two registers, there are some that have more, even as many as four [Thongkum, 1991].

Our interest in the topic has been stimulated not only by published instrumental studies [e.g., Gordon and Ladefoged, 2001; Thongkum, 1989; Wayland and Jongman, 2003; Brunelle, 2005] and its possible relevance to tonogenesis [Thongkum, 1990; Thurgood, 2003, 2007; Svantesson, 2001; Premsrirat, 2003; Abramson, 2004] but also by the apparent lack or scarcity of perceptual studies of register complexes in the public domain. Indeed, there have been perceptual studies of dialects of Khmu', but these dialects have been taken to be either tonal or non-tonal. Thus, on the assumption of the existence of phonemically distinct tones in a northern dialect of Khmu', perceptual experiments with three sets of F0 patterns were run with three native speakers to explore the possible role of differences in shape of the contours [Gandour et al., 1978]; this was a sequel to a study of the F0 contours of its tones [Gårding and Lindell, 1977]. This dialect, however, is a member of a subgroup considered to have undergone tonogenesis perhaps as early as 700 years ago [Svantesson and House, 2006, p. 312]. In a recent study just called to our attention [Svantesson and House, 2006], the investigators used synthetic F0 contours for perception testing with speakers of non-tonal Eastern Khmu' and speakers of tonal Western and Northern Khmu'. As far as we can tell, in the latter study too, only the absence or probable presence of distinctive tones was recognized; presumably no Western dialect conventionally described as having register complexes [Premsrirat, 2003] was involved. Rather, two tones differentiated essentially by pitch were the units to be studied.

For our first foray into this area [Abramson et al., 2004], we synthesized stimuli that varied incrementally along dimensions traditionally said to be relevant for the register distinction of the Kuai dialect of Suai, a Mon-Khmer language, and we did acoustic analyses of those properties in a large sample of speakers. We found the distinction to be in a state of flux both in production and perception.

To the extent that it was still viable, the registers were mainly distinguished by fundamental frequency (F0) contours with some contribution from phonation type.

Responding to reports of greater stability in the register distinctions of some Western dialects of another Mon-Khmer language, Khmu' [Premsrirat, 2002, 2003], we have undertaken a study of the Khmu' Rawk dialect, the home language of 3,197 people in ten villages of the northern Nan Province of Thailand. Specifically, we have worked with speakers in the village of Huay Sataeng, which has 482 speakers, in the Thung Chang District. The Khmu' language belongs to the Khmuic branch of the Mon-Khmer language family. Khmu' in its many dialects is spoken in southern China and in northern areas of Vietnam, Laos, and Thailand. In Thailand about 10,000 Khmu' speakers live in the provinces of Chiangrai and Nan near the Lao border. The inventory of Khmu' Rawk phonemes is given in the Appendix. Other dialects have been described [e.g., Gårding and Lindell, 1977; Premsrirat, 1987, 2002; Smalley, 1961].

With regard to our research topic, it is important to note that the languages of the Mon-Khmer family are of three types [Ferlus, 1980]: (1) Non-register and non-tonal, (2) Register (usually two register complexes), and (3) Tonal (two tones). Indeed, the dialects of Khmu' itself are distributed across the three types. Type 1 is historically the most conservative with maintenance of the old voicing contrasts in consonants, while type 2 can be seen as a transitional stage of a movement toward tonality. Bringing linguistic findings and laryngeal mechanisms to bear on the matter, Graham Thurgood [1993, 2007] has argued plausibly for the probable role of voice register as a stage of tonogenesis in general. Evidence for its role in the Mon-Khmer family has been published [Thongkum, 1990] and specifically for its role in Khmu' [Premsrirat, 2003, Table 4]. Premsrirat's [2003] listing of dialects places Khmu' Rawk in the subgroup of Western dialects with register complexes [See also L-Thongkum et al., 2007, p.232]. Preliminary observations of the Khmu' Rawk

spoken in the village of Huay Sataeng, accessible through the third author's project Linguistic Diversity in Nan Province: A Foundation for Tourism Development, did indeed strike the ear as belonging to type 2 with variation among its speakers as to how much Register 1 is characterized by modal (clear) voice or by high pitch and how much Register 2 is characterized by breathy voice or by low pitch. For some speakers interviewed, impressions included utterances with both phonetic properties present for each register, while for others it was just pitch.

Our research had two goals. First, we wished to determine what acoustic properties serve to differentiate the two voice registers in speech production. Are these registers indeed complexes of more or less equally important properties, such as phonation type and fundamental frequency, or, rather, does one property dominate in the differentiation even if one or more other concomitant properties participate? If, in the latter case, the dominant property is fundamental frequency (the principal physical correlate of pitch), is tonogenesis in progress? Our second goal was to design experiments to determine the perceptual efficacy of any significant properties found in our acoustic analysis. Although we could not conduct physiological or aerodynamic studies in the village, we nevertheless expected to be able to rationalize our findings in the light of what is known about the acoustic consequences of articulatory gestures.

II. Methods

Production Studies

Recorded Materials

With a few refinements noted herein, the acoustic recording and analysis methods closely followed those employed in an earlier study [Abramson et al., 2004]. The recordings we analyzed were of nine word-pairs; one member of each was uttered in Register 1 (clear voice, high pitch) and

the other member in Register 2 (breathy voice. Low pitch) (see table 1). The words were recorded in randomized sequences by 25 native speakers of Khmu' (identified by the letters A through Z excluding D due to the excessively poor quality of his recording). Five of the 25 speakers were employed in two recording sessions that generated six tokens of each word; meanwhile the remaining speakers were engaged in one recording session that yielded three tokens per word. The speakers were of both sexes (9 men and 16 women), and they ranged in age from 19 to 72 years. To minimize the influence of surrounding languages of the Tai family (Standard Thai, Northern Thai, or Lao) on the voice registers of the Khmu' dialect it was necessary to obtain data from speakers who have spent their lives in the dialect region of interest. Such people are plentiful in rural villages, remote from the linguistic mélange of an urban environment and also, it must be noted with some regret, frequently unavoidably remote from such basic technical amenities as a sound-damped room or walled enclosure that, in an urban area, can be relied upon to attenuate ambient noise. Indeed, our recordings had to be made in an open-walled Buddhist worship hall. Although it was up the mountain slope at the edge of the village, people in vehicles, domestic animals, wild birds, and insects contributed to the ambient noise. Consequently, despite the use of a directional microphone, our recordings from time to time were contaminated by some low level domestic noises typical of those that arise in a farming community. All the recordings, originally sampled at a rate of 44,100 per sec., were down-sampled to 11,025 per sec. prior to analysis. The total number of analyzed tokens exceeded 1,670 whereas in the earlier study [Abramson et al., 2004] only 288 were ultimately examined.

[TABLE 1 ABOUT HERE]

Noise contamination, exacerbated by the habit of our Southeast Asian subjects to speak softly in face-to-face exchanges and when confronted with a microphone, undermined the robustness of several computer-implemented analysis procedures, particularly those involving formant frequency

extraction. Formant frequencies extracted by computer analysis were among the least reliable data. Consequently, all formant frequencies were obtained by visual inspection. We might note here that while we lack any systematic evidence of a general cultural tendency among people of the region to favor passive manners of inter-personal communication, we believe that we have much anecdotal evidence to support such a claim.

Each word was inspected in a combined waveform and spectrographic format, and its visible beginning and ending points were identified and stored. A similar procedure was used to visually identify and store the onset and offset points of "steady state" regions of the spectrum, regions in which the variation in first and second formants was less than the difference limens of 70 and 170Hz [Flanagan, 1955a; Mermelstein, 1978].

Praat, Matlab®, and StatView© software was used throughout to perform acoustic and statistical analyses in the manner described in greater detail in an earlier paper [Abramson et al., 2004]. Some empirical modifications were made to portions of the Matlab software in an effort to improve the precision of formant extraction. However, we ultimately chose to rely upon formant-frequency data that we obtained by the combined visual inspection of broad-band and narrow-band spectrograms and narrow-band cross-sections of steady-state regions of the speech data.

Amplitude and F0 Contours

Overall amplitude values in dB and glottal frequency (F0) values over a 75 to 600Hz range were obtained at 10-ms intervals for each utterance. Visual inspection was employed to eliminate some instances of frequency doubling, and the resulting frequency values were then converted from Hz to semitones using the formula $P_{semitones} = 3.32 \times 12 \times \log_{10}((F0_{Hz})/base)$. This conversion, done as part of

the normalization across speakers, was meant to more satisfactorily approximate the listeners' sensation of pitch. In the case of each speaker the "base" was the minimum F0 frequency measured for that speaker across all utterances in both voice registers. Thus, for each speaker, a single base value was computed and applied in the conversion of all F0 contours in both registers. Next, both the frequency and amplitude data were time-normalized to fit 100-point scales, and, using these scales, the amplitude and F0 data for different subggroups of subjects (males vs. females; old speakers vs. youthful speakers) and the entire coterie of speakers were averaged. Graphs of these data are shown in figures 1 through 10. Finally, to help us examine the differences between the time-normalized contours in each register, we borrowed a procedure from the previous study. Basically, this involved the examining of plots of the averaged data and selecting sectors along the normalized time scale where the largest differences in rate of change of amplitude or F0 between the two Registers were apparent. Thus, in this instance, we divided the time scale into four sectors (1-25, 26-50, 51-85, and 86-100) and computed the gradients of the Register 1 and Register 2 contours in each sector. Thus, for each utterance by each speaker, four gradients were calculated by linear regression. Then for each of the gradients 1 through 4, each speaker's repetitions of a given word in a given register were averaged. This procedure led to the formation of a data set of 1,800 measurements (25 speakers x 9 words x 2 registers x 4 gradients) that was subsequently submitted to a repeated-measures analysis of variance (ANOVA).

Formant Frequencies

Shortcomings in the quality of the speech, primarily due, we believe, to the generally soft speech of the speakers but also partially due to background noise, led to numerous small errors in the results produced by a usually effective Linear Predictive Coding (LPC) method of formant extraction. To correct these errors, a Matlab program was created to display broad- or narrow-band spectrograms

and narrow band cross-sections of each member of the entire steady-state data set. Using this display, both first (F1) and second (F2) formant frequencies were identified by eye, and their values were stored in a file linked to each utterance of origin. These data were subsequently used in statistical tests of the stability of the F1 and F2 values as a function of voice register and also in calculating harmonic ratio values. That is, mismatched formant frequencies between the members of a word pair would invalidate any measure of spectral tilt for that pair.

Ratios of Harmonic Intensities

Following the procedure adopted earlier [Abramson et al., 2004], for each steady-state region of the vowel formants the intensity of the fundamental (H_1) and its second harmonic (H_2) were identified and combined with the intensity of the F1 peak harmonic (H_{F1}) to form two ratios, H_2/H_1 and H_{F1}/H_1. For this study, however, a third ratio involving F2, H_{F2}/H_1, was added. The harmonic extraction procedure applied a Hamming window to each steady-state data segment and extended it to a length of 1,024 points by the addition of zeros. A Fast Fourier Transform (FFT) of the 1,024-point array led to the production of a spectrum vector with a 21.5Hz frequency resolution. The harmonic frequency was extracted from the cepstrum and a sine wave with that frequency was aligned or synchronized by autocorrelating the sine wave with the harmonic structure of the spectrum vector. Finally, peaks in the sine wave representing the fundamental and its second harmonic were used to locate those specific frequencies in the harmonic spectrum and extract their intensities.

Vowel Duration

To determine whether differences in vowel length might be being used to differentiate the two voice registers [Thongkum, 1988; Gordon and Ladefoged, 2001] we repeated our vowel-duration measurement procedure based on a gestural criterion for defining the span [Abramson et. al., 2004, pp. 152–153]. Thus, the vowel was taken to run from the release of the initial consonant to the closing

gesture of any final consonant. For the one pair with initial /w/, the onset of the vocalic glide was marked as the beginning, and for the one pair with no final consonant, the last detectable voice excitation of at least one formant was marked as the end. All vowel onset and offset points were identified by visually inspecting a wide-band spectrogram of each utterance. The positions of a hairline cursor placed at the onset and offset boundaries of each vowel were recorded.

Perception Studies

Natural Speech Perception

Whatever the phonetic properties of the two traditional voice registers might be today, we had no reason to doubt the validity of the phonological distinction itself in Khmu' Rawk; nevertheless, as a basis for our perceptual experiment, we wished to have an assessment of its auditory robustness. In addition, dealing with a similar populace with little schooling in our earlier study of Suai, we had encountered a small number of listeners who lacked either the ability to hear the distinction or an understanding of the task, and we wished to verify that each Khmu' listener understood and could carry out the instructions correctly. The natural stimuli were two utterances of the words /ráaŋ/ 'tooth' in Register 1 and /ràaŋ/ 'flower' in Register 2; one pair of utterances was spoken by a male (Speaker J) and the other pair by a female (Speaker K). Chosen for the salience of the register distinction, each of the selected utterances was presented seven times in one of two random orders to 32 native speakers (10 men and 22 women) for identification. Pictures of teeth and flowers lay in front of the listeners, and, to indicate what they had heard, they made their selections known by pointing at the appropriate picture. The pointing responses were observed by the experimenter and recorded via the computer's keyboard.

Speech Synthesis

As a starting point for synthesis, a production of the word /ráaŋ/ in (clear voice) spoken by Speaker A was selected as the model utterance. The SynthWorks® software package was used to extract its synthesis parameters, among which was included an F0 contour. Although in our previous study of Suai the behavior of F0 was found to be the most significant factor in a multifactor analysis of listeners' Register 1 vs. Register 2 responses to synthetic stimuli, we had no reason at the outset of our work to believe that the same would be true of our dialect of Khmu'. We were led to focus our attention solely on the F0 parameter in this study because of the exigencies of scheduling. When it was nearly time for the first author to take up residence in Thailand to prepare the test orders and logistics for carrying out the perceptual experiments, we had not done much of the analysis of harmonic ratios. Thus, at this stage of the work, we had no analytic basis for assuming that phonation type is relevant to the phonological distinction. Therefore, it seemed prudent at the time to proceed only with F0, the significance of which we had found through acoustic analysis. The averaged F0 contours obtained from measurements of natural productions of the two registers (figure 6) provided the basis for synthesizing a set of seven different F0 contours representing steps along a hypothetical F0 continuum. The interpolated contours were obtained by computing increments representing one-sixth of the differences between the frequency values of Registers 1 and 2 at each time sample. These increments were used to create seven contours that, in addition to those of the original averaged F0 Register 1 and Register 2 contours, included four intermediate contours and one super stimulus that followed a trajectory lying one increment above that of Register 2 (see figure 13).

Perception of Synthetic Speech

The same 32 native speakers of Khmu were paid to listen to sequences of synthesized words presented in one of two random orders and to identify them. Each random order contained seven tokens of each stimulus. Again, pictures of teeth and flowers lay in front of the listeners and, to indicate what they had heard, they made their selections known by pointing at the appropriate picture. The synthesized words were reproduced by a Macintosh laptop computer in sequences controlled by a program written for PsyScope [Cohen et al., 1993].

[Figure 1 about here]

III. Results

Amplitude and F0

Figure 1 shows the overall averages of the time-normalized amplitude contours of 25 speakers' productions of nine utterances in both Register 1 and Register 2. With the issue of register stability in mind, subsets of the data were plotted. They are shown in figures 2 through 5 and represent the collective performances of male, female, young, and old speakers respectively. In every instance, the peak of the Register 1 contour is consistently higher than that of Register 2. Nevertheless, to address the question of whether the rate of change of amplitude might differ, we took our examination of the data a step further and divided each amplitude contour into four sectors along the time dimension and conducted an ANOVA. The results of this analysis are summarized in table 2, which shows in the 25-speaker group an inconsistent pattern of register-effect probabilities that reaches significance[1] only for sector 3. This inconsistency is also apparent for the smaller subgroups.

[1] In this paper, an effect is termed significant when the probability p that a null hypothesis is true is less than 0.005 (i.e., $p < 0.005$).

[Figures 2–5 about here]

[Table 2 about here]

[Figure 6 about here]

In figure 6 the time-normalized average F0 contours of the Register 1 and Register 2 productions of 25 speakers are shown plotted on a semitone scale. A mean difference of 3.27 semitones is apparent over the entire duration scale. The maximum difference of 4.2 semitones occurs at time datum 18. Plots of data in the age and sex categories are shown in figures 7 through 10. The female contours are a few semitones higher in pitch than those of the males and exhibit larger standard deviations. The average difference across all 100 points of the normalized time scale ranges from a maximum of 3.54 semitones (Young Speakers) to 2.56 semitones (Old Speakers). Meanwhile, the maximum Register 1–Register 2 difference (4.37 semitones) occurs at datum 26 of the young dataset, while the minimum difference (0.97 semitones) is found in datum 9 of the old speakers' dataset. To examine differences in F0 contour shape, we computed the gradients of the two contours in each of the four normalized time periods or sectors. The results, which appear in table 3, reveal that for the group consisting of all 25 speakers there are significant differences between Register 1 and Register 2 in each of the four sectors. For a group composed exclusively of male speakers, only the two mid sectors of the normalized time scale yield significant differences. In contrast, the corresponding results for the female group show significant effects across all four sectors. Elderly speakers as a group produce significant mid-section results like those of male speakers, while younger speakers have somewhat less clear results, given the probability $p = 0.009$ in sector 2. However, the fact that $p << 0.005$ in both sectors 1 and 3 suggests that it may be reasonable to conclude that the shape of the F0 contour over all of the first three sectors is important.

[Figures 7–10 about here]

[Table 3 about here]

Analysis of Vowel Duration

Vowel duration data derived from onset and offset times established by the visual inspection of spectrograms were analyzed by an ANOVA. The test was designed to determine whether the two registers were indistinguishable with regard to vowel duration. The results, shown in table 4, reveal that Register-2 vowels are longer with a mean difference in duration of 22 ms for the full 25-speaker dataset. Moreover, within both age and sex sub-categories, the difference, ranging from 28 to 21 ms, persists. An analysis to determine whether this duration difference might have arisen by chance shows that, with the exception of the old and young datasets that are marred by small sample sizes (N = 6 and 5 respectively), the probability is exceedingly small ($p < 0.0003$).

[Table 4 about here]

Formant Frequencies

To determine whether productions of the two registers might be distinguished by systematic variations in formant frequencies, we performed independent repeated-measures ANOVAs on the visually-extracted F1 and F2 data. The results are shown in table 5. The probabilities that the populations of F1 and F2 frequencies in Register 1 are indistinguishable from their corresponding populations in Register 2 are very high for all speaker categories. Thus, no matter the age or sex of the speaker, Registers 1 and 2 do not differ significantly with respect to either their F1 or F2 frequencies.

[Table 5 about here]

Harmonic Intensity Ratios

As a result of the presence of recorded noise coupled with the speakers' tendency toward soft speech, the periodicity of the harmonic spectrum was sometimes buried in the noise, and the autocorrelation procedure failed to find a plausible alignment. Therefore, graphs of all the spectra and their autocorrelation results were visually examined for instances of synchronization failure which, when found, led to the omission of that ratio value. The frequency distributions of the surviving Register 1 and Register 2 ratio values of all types (H_2/H_1, H_{F1}/H_1, and H_{F2}/H_1) were positively skewed and, to achieve closer approximations to the gaussian or normal distributions underlying repeated-measures ANOVAs, a log transformation of the ratio data derived from each repetition of each of the words was applied prior to analysis. Thus, for each of the three ratio types, the log ratio for each repetition of a given word uttered by a given speaker was calculated and averaged over all repetitions of that word to yield one datum for each ratio type, word, and register. As a result of rejecting all the ratios affected by synchronization failure, a total of 3.05 percent of the Register 1 data derived from individual word repetitions and 5.00 percent of the corresponding Register 2 data was eliminated before averaging. Histograms of Register 1 and Register 2 log-ratio data obtained from all 25 speakers are shown in figure 11.

[Figure 11 about here]

[Table 6 about here]

Table 6 summarizes the results of an analysis of the logarithmically-transformed harmonic intensity ratios. The pooled data of the entire group of 25 speakers reveal significant differences ($p < 0.005$) for *Register* for the first two ratios but not for the third. Significant differences are found for *Word* in all three cases as expected, since the nine words were all chosen to be distinctively different from one another. The *Register x Word* interaction, meanwhile, lacks significance for all three ratios. Harmonic ratios are a measure of the slope of the voice spectrum, and significant differences between

the two voice registers suggests that the durations of the open and closed quotients of the glottal cycle undergo change [Stevens, 1998, pp. 85–92]. As a consequence of their different social roles in the community, male and female speakers are exposed to different influences; females may spend more time within their village, while males may travel further afield and become exposed to greater linguistic variation. Thus, it is of interest to analyze the data of males and females independently. In this case the analysis shows that, in male speakers, the glottal duty cycle makes no contribution toward the production of voice register distinctions. The situation is shown to be different for females whose data achieve the $p < 0.005$ level of significance for the first of the three log ratios. Less reliable, due to the small quantity of data available for analysis, are the results of old vs. young speakers. All probability levels for register fall substantially short of significance and do not permit any conclusions to be drawn as to the effects of changes in linguistic norms over time.

[Figure 12 about here]

Perception of Control Stimuli

Figure 12 shows the overall percentage of responses given by 32 listeners to utterances spoken in Register 1 and Register 2 by speakers J and K. At a level of 96.9 percent versus 93.1 percent, the Register 1/Register 2 distinction was apparently perceived more readily in the productions of the female speaker (K) than in those of the male speaker. Eleven of the 32 listeners made errors. Eight of those listeners made three errors or fewer, while the remaining three listeners made six errors or more. When the responses of the three lowest performing listeners are removed, the overall response levels to the productions of male and female speakers rise to 97.2 and 100 percent respectively. Accordingly, the responses of the three least acute listeners were excluded from our analysis of the perception test using synthesized stimuli.

[Figure 13 about here]

Perception of Synthetic Stimuli

The same group of 32 listeners was engaged to judge the synthetic utterances. These utterances differed only in their F0 contours and were made to match the overall F0 contours of male speakers in the two registers (figure 13), with the addition of five interpolated F0 trajectories to bring the number of intervals along the F0 continuum to a total of seven. The response data, omitting the responses of three of the 32 listeners for the reasons noted above, are shown plotted in figure 14. The plot lines are quasi-sigmoidal in shape and indicate that, at the extreme, Register 1 performance lay at the 90 percent level. At the Register 2 end of the scale, the percentage of correct responses is slightly lower than that for Register 1, and it is evident that no enhancement of recognition performance was achieved by the F0-elevated super stimulus.

[Figure 14 about here]

IV. Discussion

Amplitude

Although the amplitude peak in at least one sector for all speakers, as well as for all the subgroups, is significantly greater in Register 1 (table 2), the small difference makes us believe that it can have little or no value in differentiating the two registers. These differences are seen to be smaller than the standard deviations of the individual utterances. In each graph, dark and light shaded areas indicate the span of standard deviations for Registers 1 and 2 respectively. The graphs of data in all categories reveal a broadly identical pattern, a predominantly positive Register 1–Register 2 difference, whose mean ranges from a maximum of 2.05dB (Young Speakers) to a minimum of 1.10dB (Female Speakers). The largest single difference (3.4dB) occurs in the young speakers' plot at

datum 23 on the normalized time scale. Flanagan (1955b) measured the perceptible difference limen for vowel amplitude and concluded that, for a synthesized neutral vowel of fixed duration, the threshold lay in the neighborhood of 1.5dB. Natural vowels extracted from spoken words have variable durations and a variety of vocal damping and excitation characteristics that would be expected to give rise to a higher limen that exceeds the differences found here. In brief, the difference limens of steady-state synthetic vowels are smaller than those of the typically dynamic vowels of natural speech [O'Shaughnessy, 2000, pp. 134–136]. Thus, it appears plausible that, notwithstanding the persistence of an absolute Register 1–Register 2 difference, it is unlikely to be perceptible to native speakers of Khmu'. That is, as a parameter of the speech mechanism, it is not here under voluntary control. Rather, it seems to be tightly correlated with the F0 contours.

Despite the foregoing, the question arises as to why the amplitude should be consistently higher for Register 1 along with a higher F0. If we take the position that the language has given up its register system and has shifted to a tonal system with F0 contours as the paramount differentiators, as implied at least by the male data, one possibility is that higher subglottal air pressure, the major contributor to sound pressure level, makes enough of a contribution to the higher F0 to account for the higher output amplitude. That is, even though most of the control of the rate of glottal pulsing resides in the laryngeal muscles, subglottal pressure may play a sufficient role to account for the small but statistically significant differences in amplitude. Indeed, a lengthy exposition [Fant et al., 2000, pp. 122–125] of covariation of subglottal pressure, F0, and sound pressure level explains that up to approximately the mid point of a speaker's voice range the voice-excitation amplitude increases with F0. Of course, aside from these trends toward covariation, subglottal pressure and F0 can vary independently. In a follow-up study [Fant and Kruckenberg, 2004, pp. 79–80] we find, "…we may now predict that in connected speech there should exist conformity between intonation contours and

intensity contours, with 1 semitone in F0 corresponding to about 1 dB in SPL. This rule has a support in our data on prose reading."

To this may be added another possibility on the assumption of a state of linguistic flux in which a transition from registers to tones is not complete. To the extent that breathy voice *is* a property of Register 2, especially for the women, one would expect weak medial compression of the vocal folds with a resultant small opening along much of the glottis, to cause, through wasted air in less efficient phonation, a drop in the amplitude of the radiated signal [Laver, 1980, pp. 132–135]. No support, however, appears for such a difference between the men and women in table 2. A study of the clear and breathy vowels of Gujarati [Fischer-Jørgensen, 1970, pp. 101–103] yields inconsistent differences in amplitude between the two categories.

The foregoing material gives us little reason to believe that our significantly different amplitude contours give evidence of voluntary control over an independent mechanism in the production of the registers.

Fundamental frequency

Our acoustic analysis of the speech of a good sampling of the Khmu' populace of Huay Sataeng revealed F0 to be the most salient and consistent property differentiating the two voice registers. Register 1 is uttered on a higher pitch than that of Register 2. As for the shapes of the contours, the gradual drop toward the end of each of them is most likely merely a sentence-intonational fall, since the words were recorded one by one as citation forms. Perceptual validation of this pitch difference was obtained from the results of our experiment with synthetic speech. The fact that the peaks of the two identification functions did not quite reach 90% is, we believe, either an artifact of the substandard testing conditions or, perhaps, our failure to make the stimuli natural enough. In particular, to the extent that breathy voice still occurs as a property of Register 2 in the community,

the complete absence of it in the stimuli at that end of the pitch series may have been dissatisfying. The same subjects achieved somewhat higher peaks in our control test with natural speech, especially with the utterances of Speaker K, a woman.

Vowel Duration

The vowels of Register 2 are longer than those of Register 1. Although the difference is significant, it is so small that it is psychoacoustically unlikely to have value as an acoustic cue. At first blush, one might assume that this is simply a matter of covariation of F0 height and duration like the apparent covariation of F0 and amplitude. Indeed, as reported by Gandour [1977], several studies have demonstrated the effects on duration of the height and shape of F0 contours. One tone language in which low F0 is found with longer vowels is Thai [Abramson, 1962, p. 107; Gandour, 1977]. This is an especially interesting example, because Thai has a phonological distinction between short and long vowels. This distinction apparently suffers no interference in the current state of the language from the slightly greater duration of vowels with low F0, although, as shown by Gandour, historically there has been interaction between tones and vowel length in some dialects of Thai.

We are concerned here with registers that are traditionally described as differing not only in pitch but also in phonation type; therefore, it is interesting to note that Fischer-Jørgensen [1970, pp. 92–95] finds the murmured vowels of Gujarati to be longer than the clear ones. This is, she says [p, 138] "probably due to the historical origin of the murmured vowels as a fusion of vowel with [h]…" On the other hand, Thongkum [1989, 1991] thinks that length distinctions between voice registers are much less likely in languages with phonemic distinctions between short and long vowels. Thus, in the Kui dialect of Suai she finds the difference limited to short vowels, while in Chong there is no length difference between the clear and breathy registers. In addition, in our recent study [Abramson et al., 2004] of the Kuai dialect of Suai, which has distinctive vowel length, once our duration data had been

corrected for the presence of consonantal aspiration in Register 1 but not Register 2, there was no length difference between the registers.

In an interesting paper Alan C.L. Yu [to appear] presents evidence for an explanation of the duration phenomenon. Drawing on work in the psychoacoustic literature with pure tones as well as his own experiments with synthetic speech, he reports that stimuli with a low F0 are heard as shorter than those of the same duration produced with a higher F0. Consequently, he explains the contrary finding of longer durations at lower F0s to be the effect of hypercorrection on the part of speakers to make syllables of rather different pitches sound equally long. There are some complications when it comes to pitting "flat tones" against "dynamic tones." The latter would seem irrelevant to our Khmu' contours, which show little movement before the final intonational dip. If Yu is right, we can only suppose that speakers start making this adjustment as children during the acquisition of language and that it eventually becomes automatic.

It is true, of course, that a definitive way of testing the hypothesis that vowel duration is indeed relevant to the register distinction would be by means of perceptual experiments employing durational variants. In the light of the foregoing discussion together with the smallness of the mean difference of 22 ms in our data, and the presence of a phonemic vowel-length difference independent of register in Khmu', the probability of length as perceptually relevant to the register distinction is remote.

Formant Frequencies

In a register language one of the properties emerging over time to distinguish the lexical items in the registers could be diverging vowel quality. In the extreme case, a vowel of an earlier stage of the language could split into two vowel phonemes. In a less extreme case, the divergence could yield subphonemic differences that are concomitant with other properties of each register. For our Khmu' Rawk data the vowels in the minimal pairs studied are phonologically the same; however, even a

significant subphonemic difference would have rendered invalid our method of determining acoustic indices of phonation-type differences through intensity ratios of harmonics. Our finding of no significant difference between the formant frequencies of the two registers failed to invalidate this metric.

Harmonic Intensity Ratios

Our analysis shows a significant difference in one of the three acoustic indices of spectral tilt for the women in our study but none for the men. In the sociolinguistic literature one comes across references to linguistic conservatism among women. For the most part, the women of Huay Sataeng limit their activities to the village, while the men largely work elsewhere in the vicinity where they are in contact with Northern Thai, Standard Thai, or Lao. Even so, the men do chat with the women, and it seems likely that they are sensitive to such phonation-type differences as remain. Thus, phonation type has weakened as a differentiator of the voice registers but seems not to have disappeared altogether.

V. Conclusion

Among the regional dialects of Khmu', Khmu' Rawk belongs to the subgroup known to have developed phonologically relevant voice registers. It, however, at least as represented by the variety spoken in Huay Sataeng, appears to be in a late transitional stage. It seems that the registers are dying out and being replaced by tones. Perhaps a better way of putting it is that one phonetic property of the traditional registers, pitch, now dominates for the community as a whole. Indeed, although we are not native speakers of the language, the two of us who are experienced field workers have listened many times to the recordings of our 25 speakers and are struck *mainly* by the pitch difference, even in the speech of the women, who seem not to have entirely lost phonation-type differences. The men and

women of the village do certainly talk to each other and may be aware of qualitative differences between male and female speech. Pitch is what we normally take to be the essential property of a tonal system, although for one or more of the tones in a language taken to be tonal there may be other concomitant properties. The frequent difficulty in specifying exactly at what point along the evolutionary scale of tonogenesis a language stands, calls our attention to the unclear typological boundary between voice-register languages and tone languages. We were not able to find enough very old subjects to make our examination of subgroups by age very fruitful.

Unfortunately, at the time when we were preparing our perceptual experiments, we had not yet found the differences in one harmonic intensity ratio among the women, so we did not test for phonation type in our synthetic stimuli. We now hope to follow up this study with appropriate new listening tests. Pending further research, we must conclude then that Khmu' Rawk, at least the variety spoken in Huay Sataeng, has either shifted from a register-complex system to a tonal system or is on the brink of doing so.

VI. Acknowledgments

This work was supported by NIH Grant DC-02717 to Haskins Laboratories and a grant from the Thailand Research Fund to the third author. We thank Ms. Chommanad Intajamornrak and Ms. Supakorn Panichkul for their making of the digital sound recordings in rather difficult circumstances. The excellent help of Ms. Chommanad Intajamornrak and Dr. Phinnarat Akharawatthanakun in handling the logistics of the perception testing in Huay Sataeng is much appreciated. We also wish to acknowledge the assistance of Dr. Steven Frost in the development of our Psyscope designs and Dr. Christine Shadle from whom we gained the benefit of a discussion about the interactions between air pressure, turbulence, amplitude, and fundamental frequency. Two anonymous reviewers made very

helpful comments. As always, the first author's strong ties to the members of the Department of Linguistics, Faculty of Arts, Chulalongkorn University in Bangkok greatly facilitated matters during his stay for part of the research.

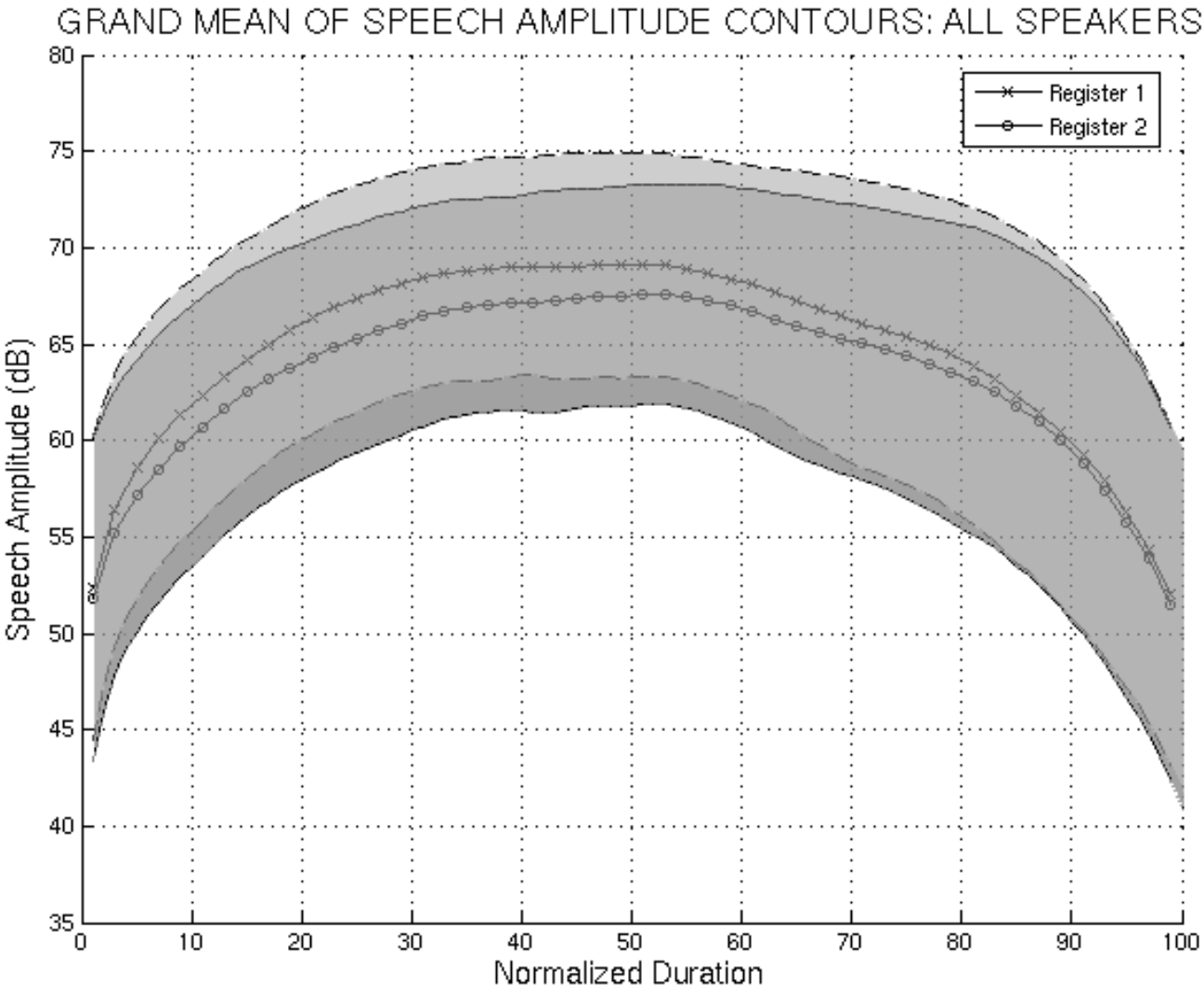


Fig. 1. Average amplitude contours for all Register 1 and Register 2 utterances produced by 25 Khmu' speakers of both sexes. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

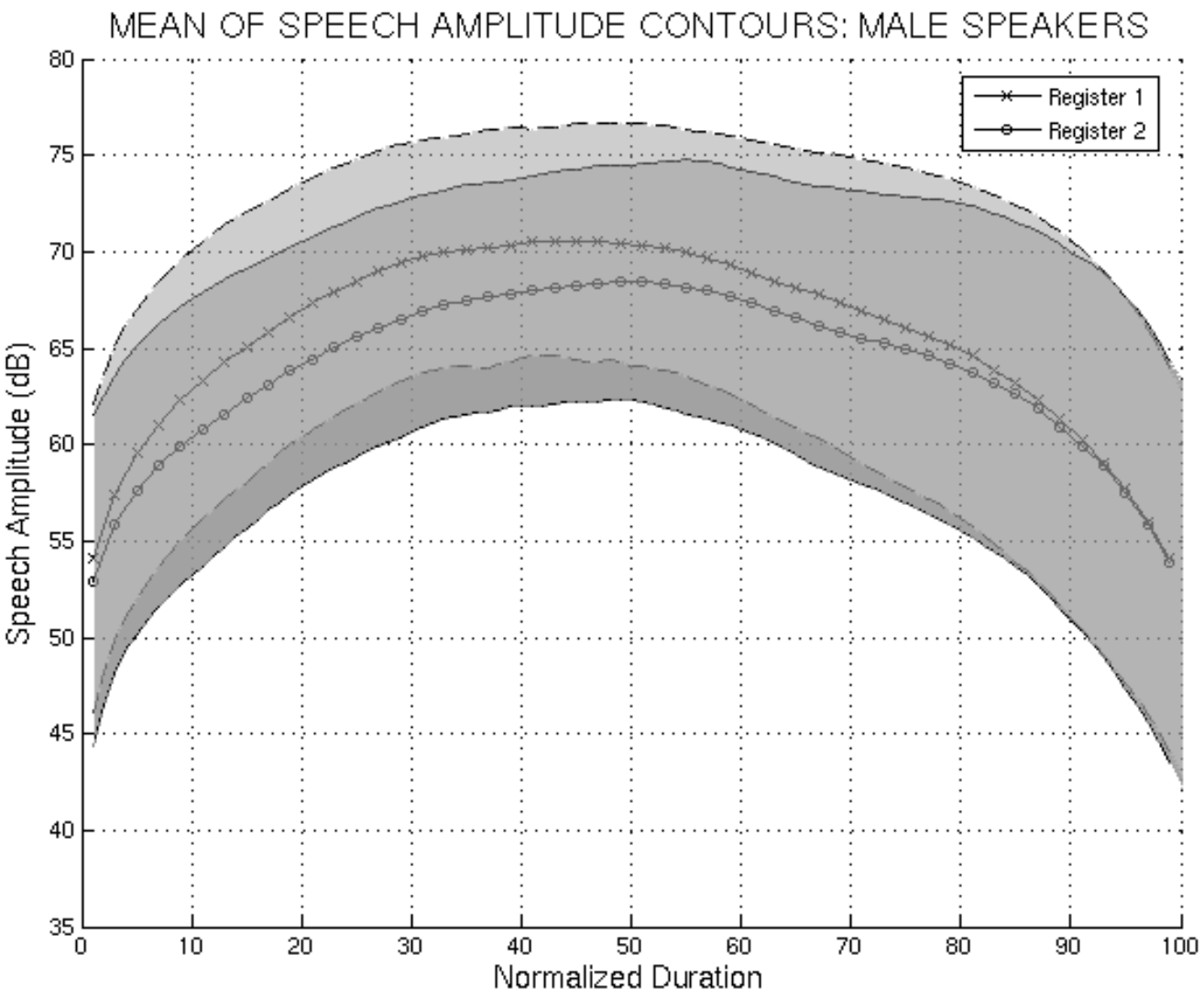


Fig. 2. Average amplitude contours for all Register 1 and Register 2 utterances produced by 9 male Khmu' speakers. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

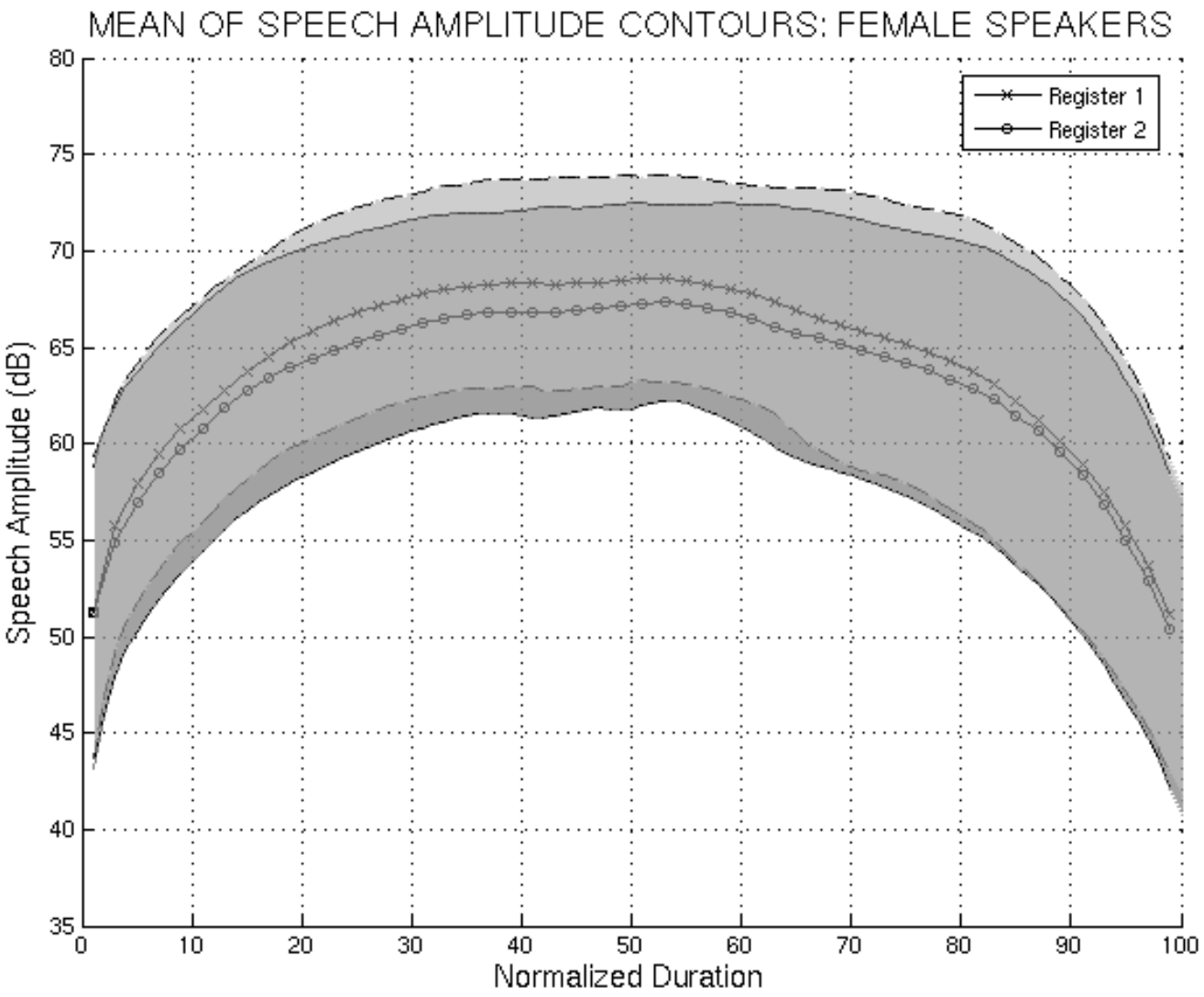


Fig. 3. Average amplitude contours for all Register 1 and Register 2 utterances produced by 16 female Khmu' speakers. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

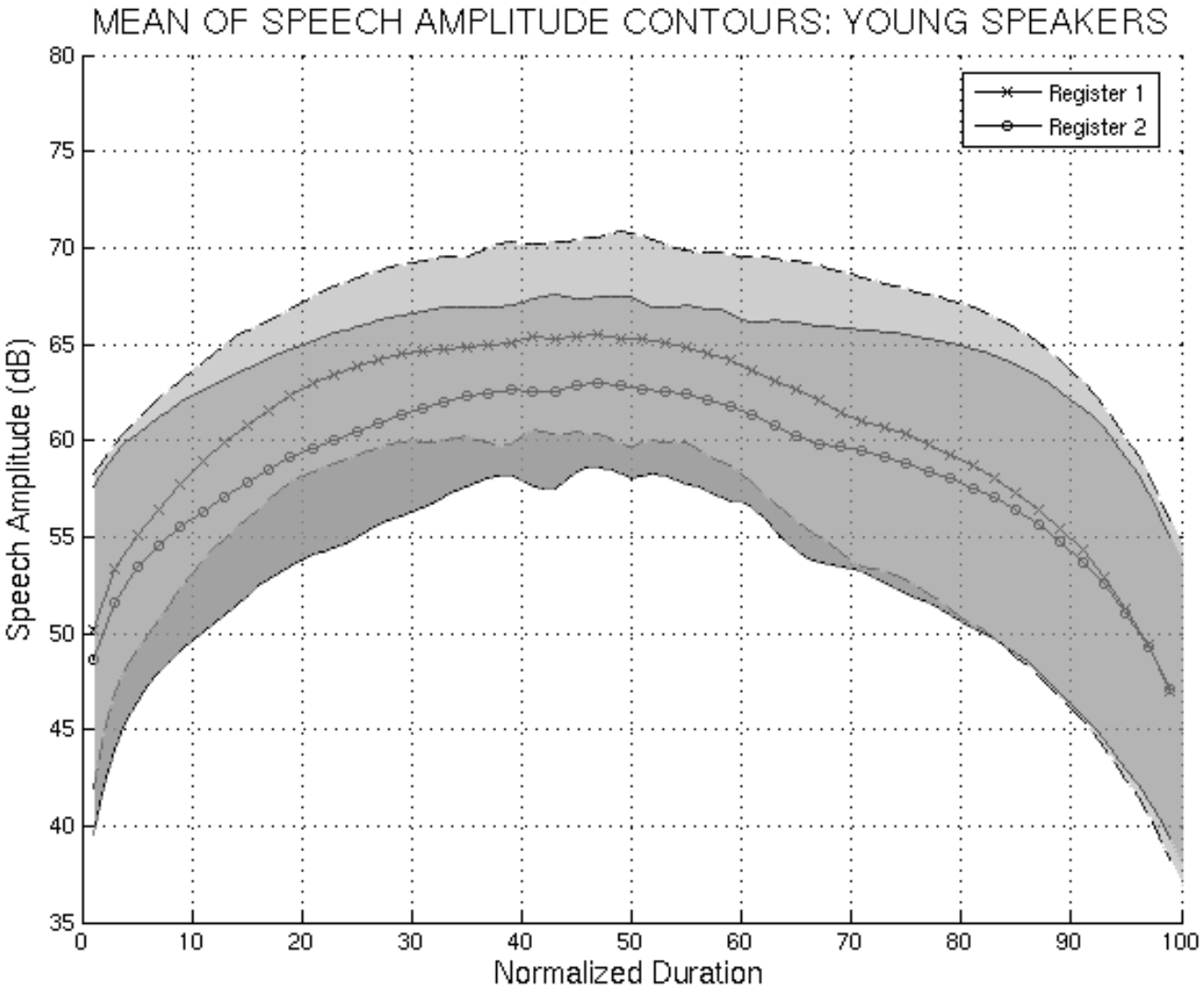


Fig. 4. Average amplitude contours for all Register 1 and Register 2 utterances produced by 5 young Khmu' speakers The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

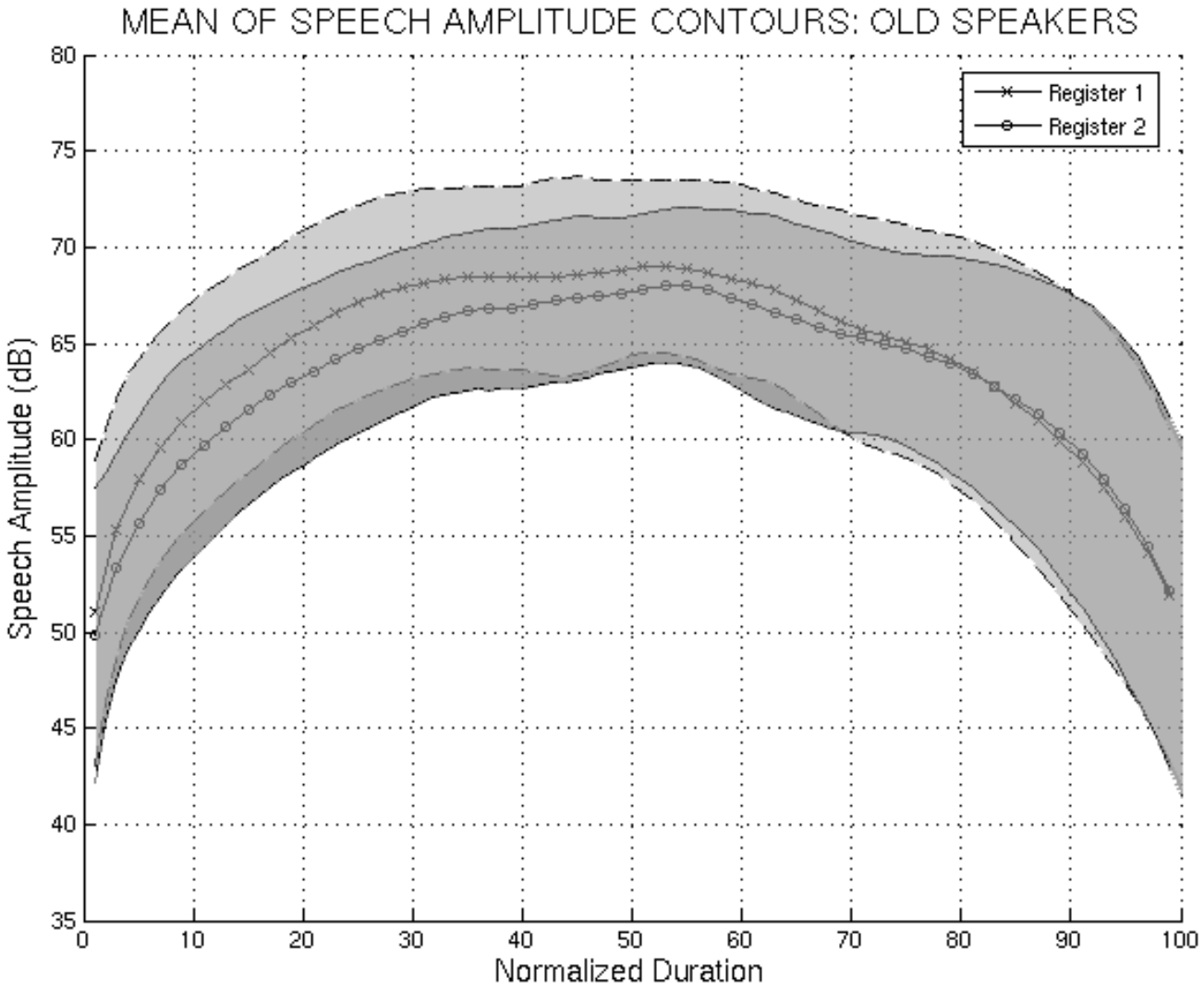


Fig. 5. Average amplitude contours for all Register 1 and Register 2 utterances produced by 6 older Khmu' speakers. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

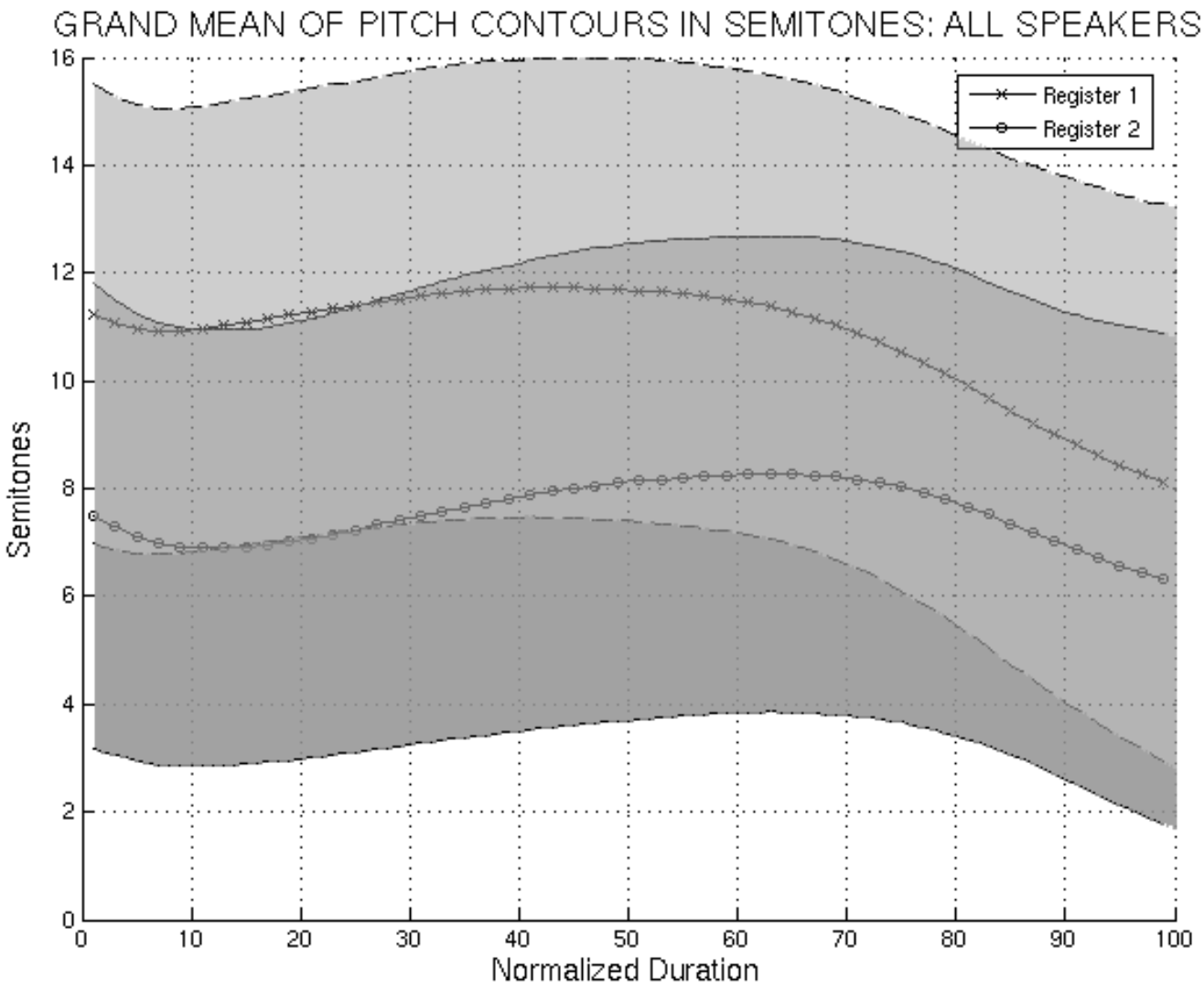


Fig. 6. Normalized pitch contours averaged over the Register 1 and Register 2 utterances of 25 Khmu' speakers of both sexes. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

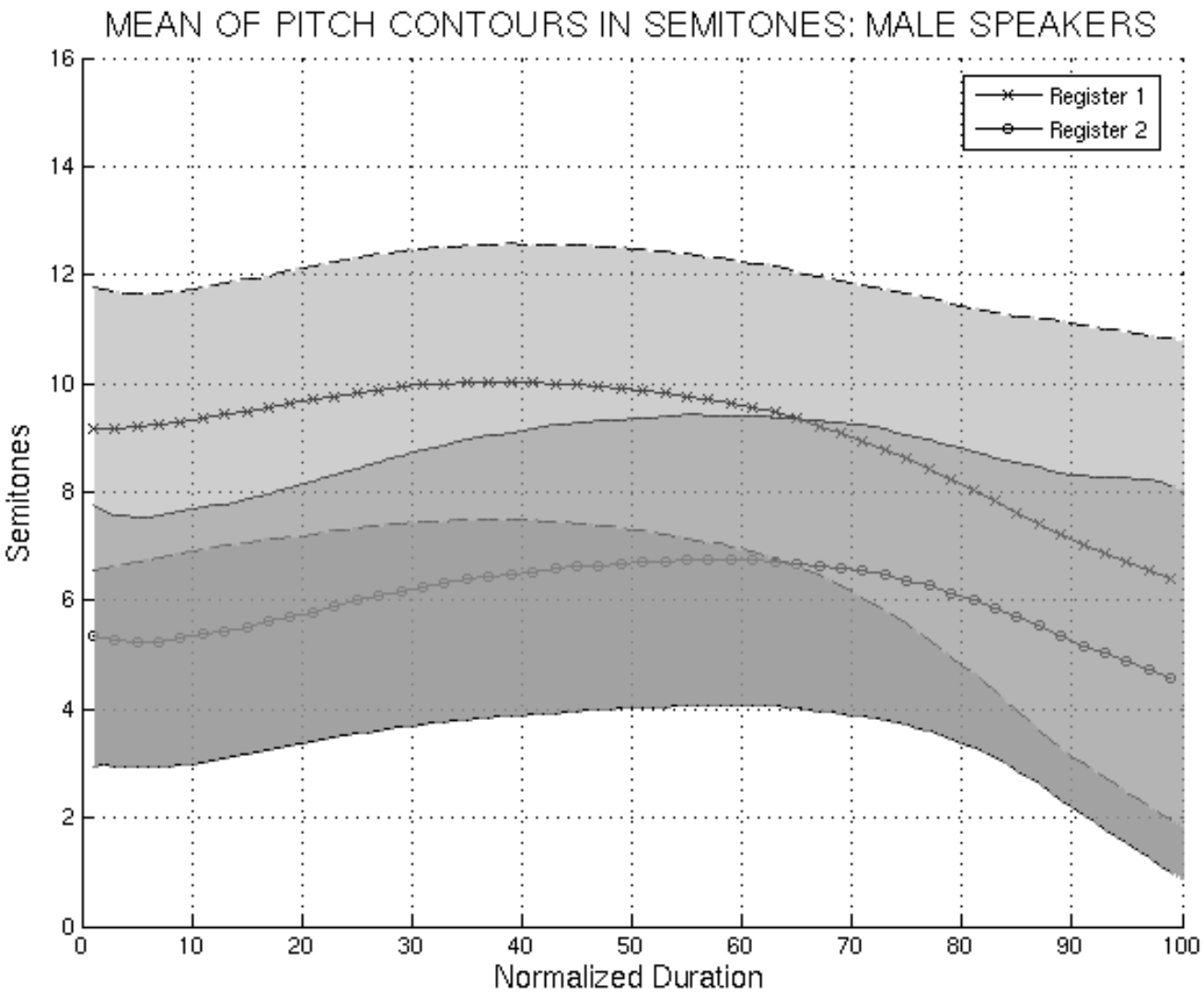


Fig. 7. Normalized pitch contours averaged over the Register 1 and Register 2 utterances of 9 male speakers of Khmu'. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

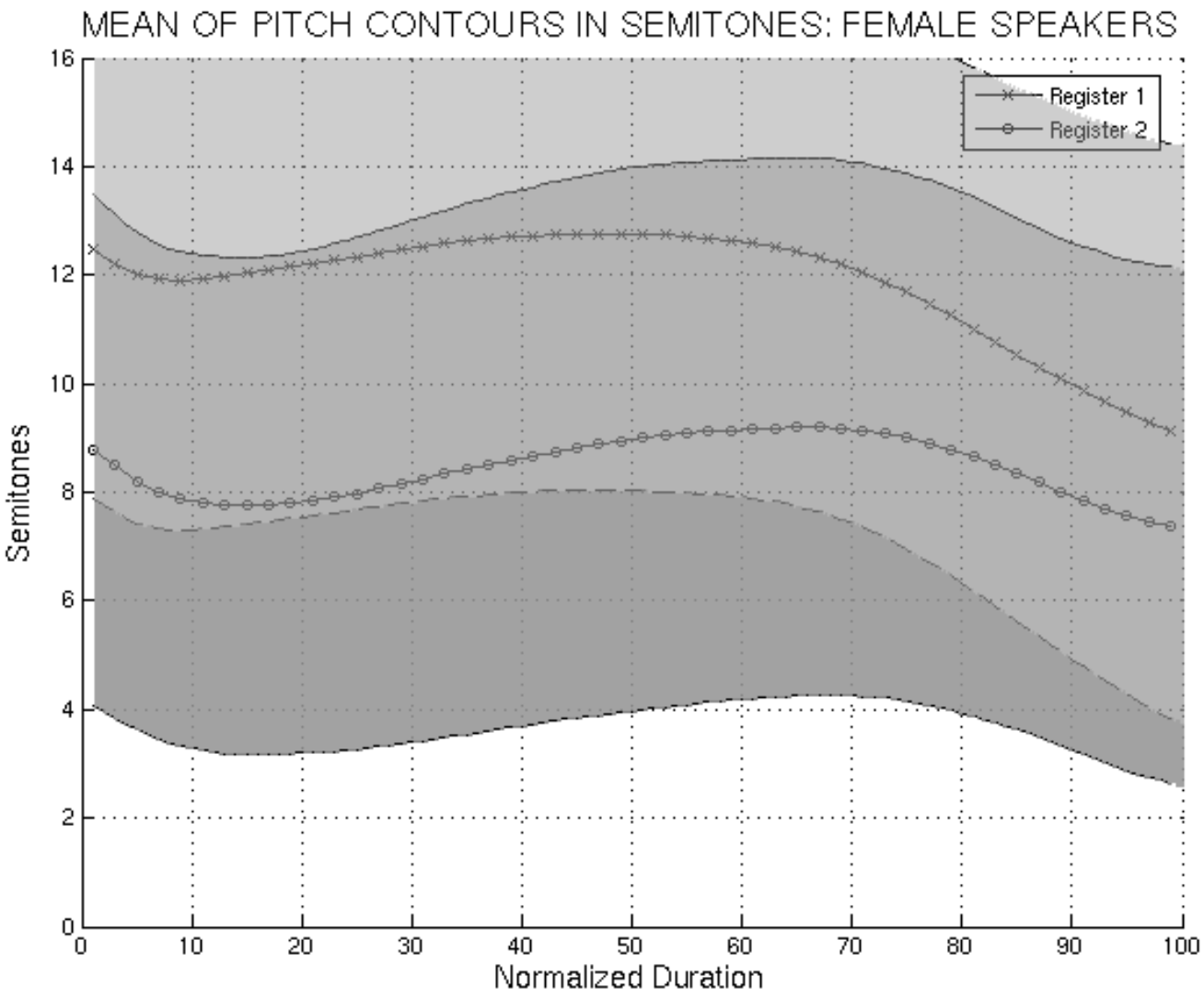


Fig. 8. Normalized pitch contours averaged over the Register 1 and Register 2 utterances of 16 female speakers of Khmu'. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

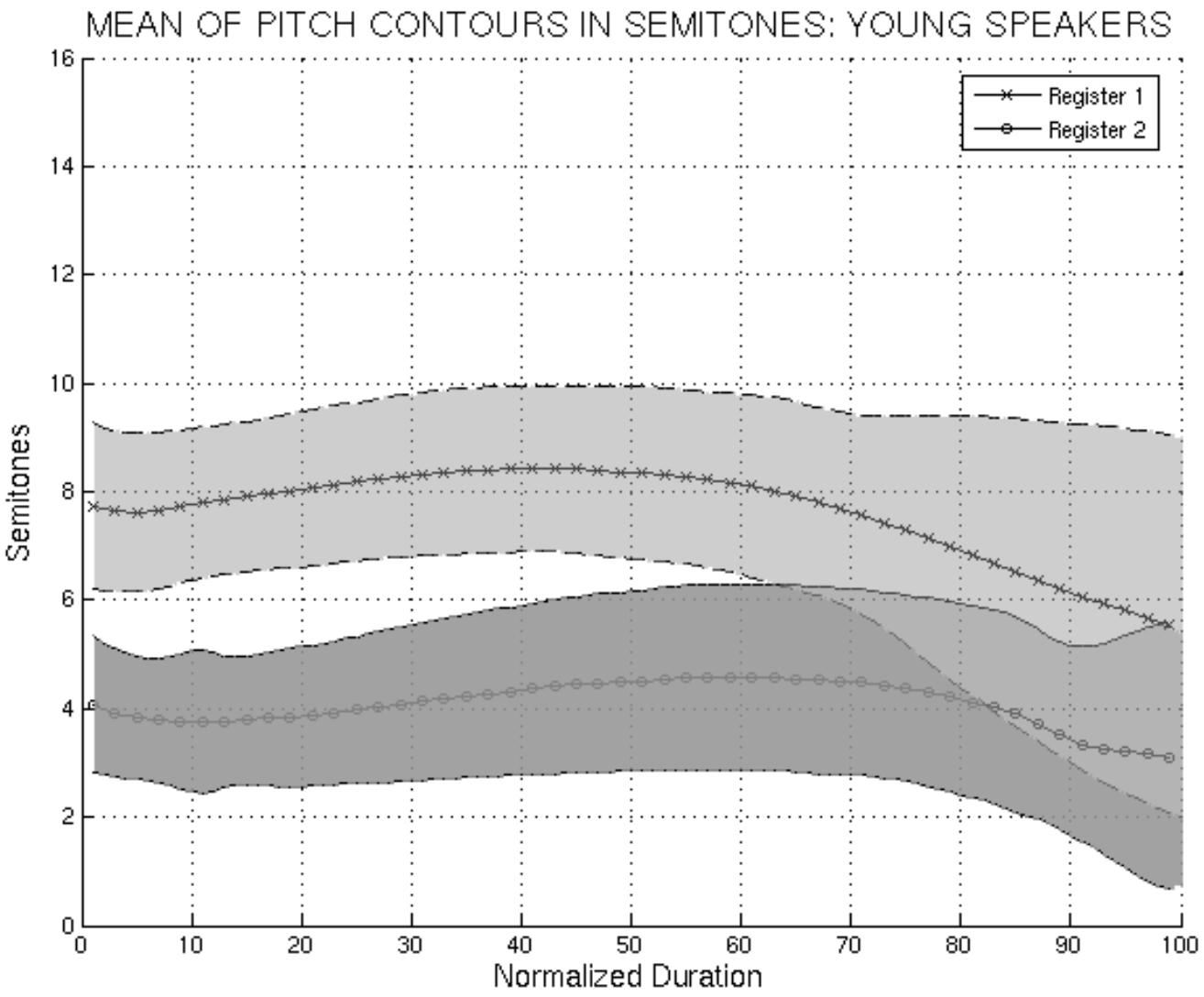


Fig. 9. Normalized pitch contours averaged over the Register 1 and Register 2 utterances of 5 young speakers of Khmu'. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

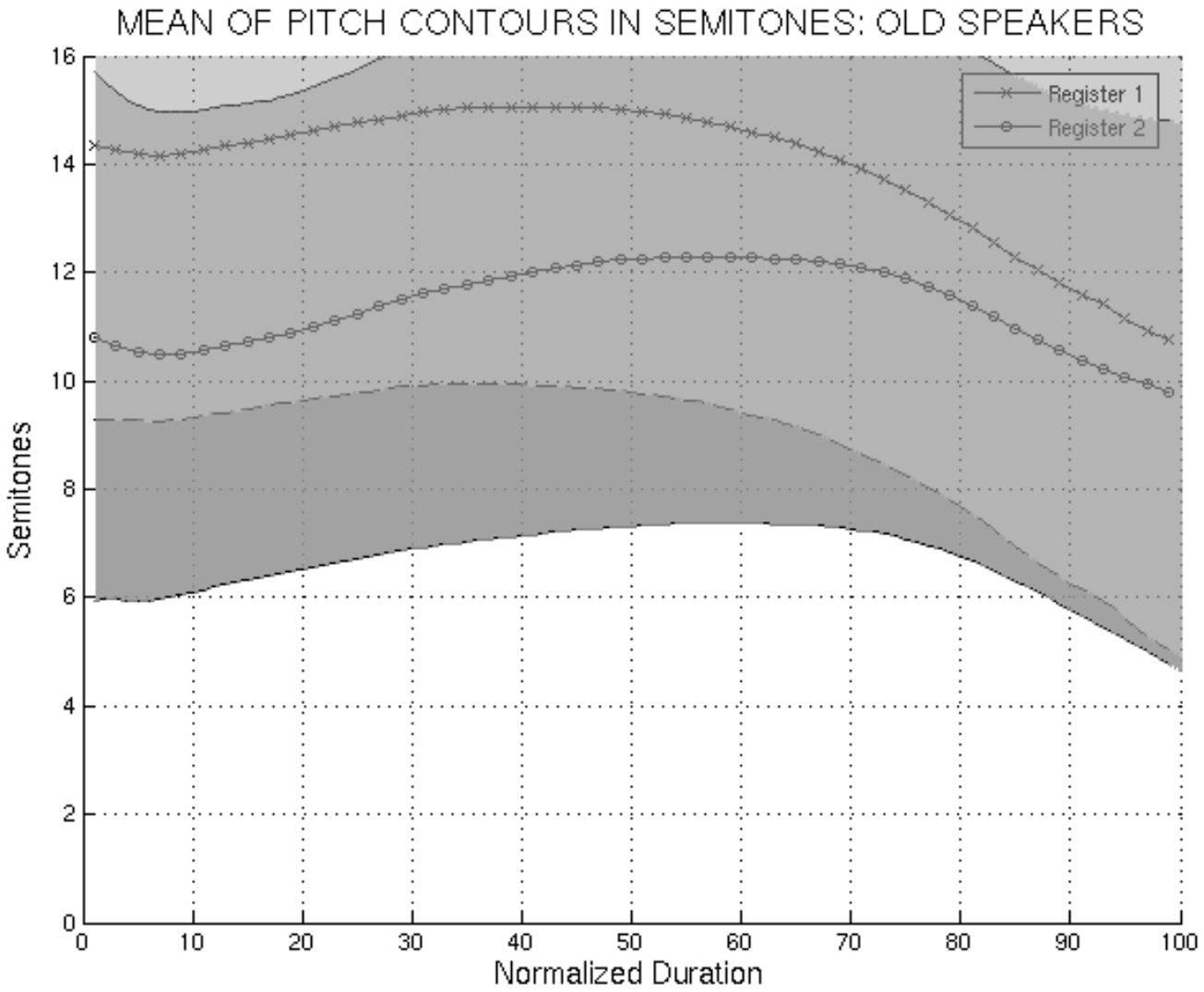


Fig. 10. Normalized pitch contours averaged over the Register 1 and Register 2 utterances of 6 elderly speakers of Khmu'. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

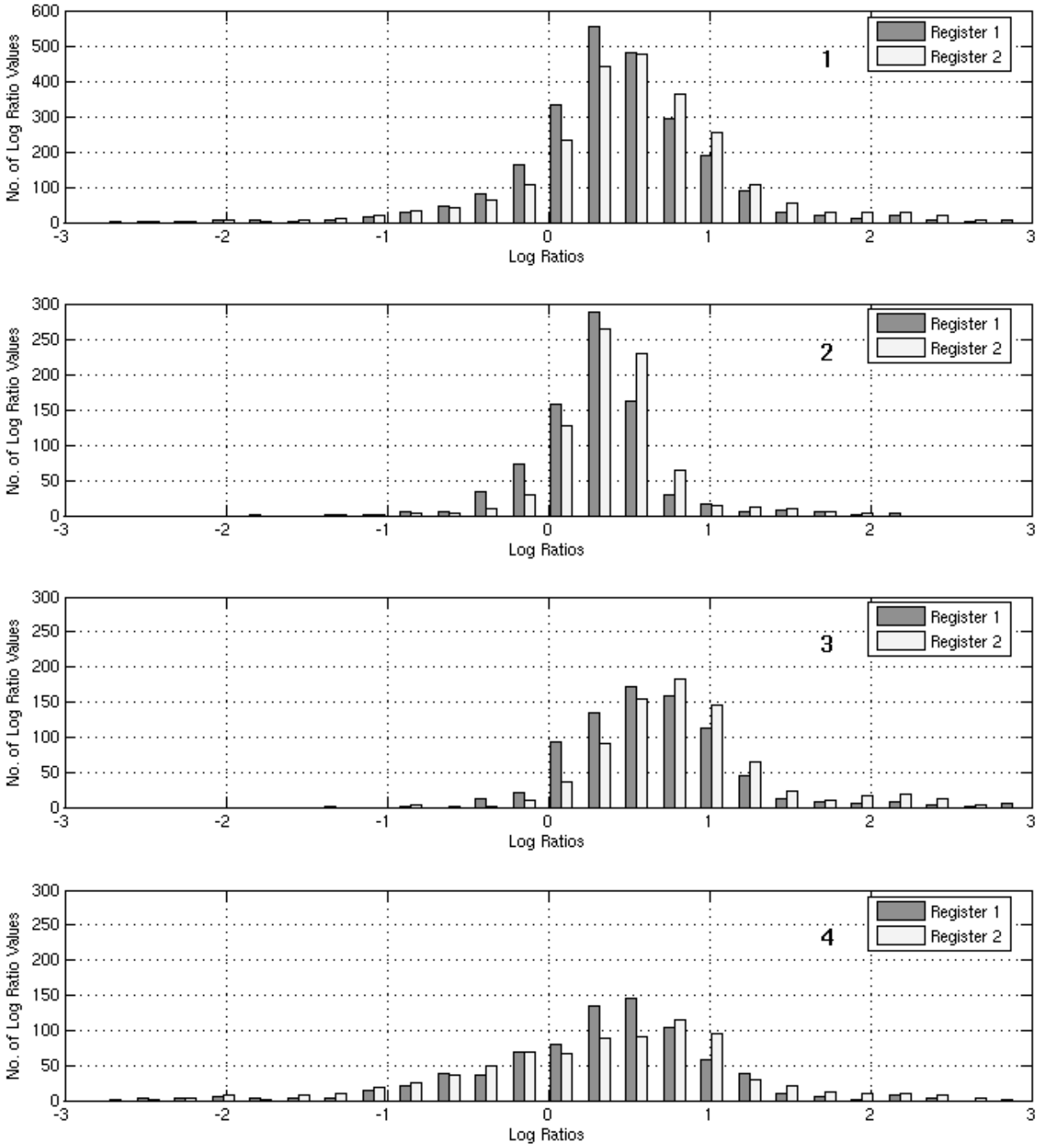


Fig. 11. Distributions of log ratio values after the elimination of all instances of synchronization failure prior to the application of an ANOVA. From top to bottom are shown in Graph 1 the combined distribution of all ratio values, in Graph 2 the distribution of H_2/H_1 ratio values, in Graph 3 the distribution of H_{F1}/H_1 ratio values, and in Graph 4 the distribution of H_{F2}/H_1 ratio values.

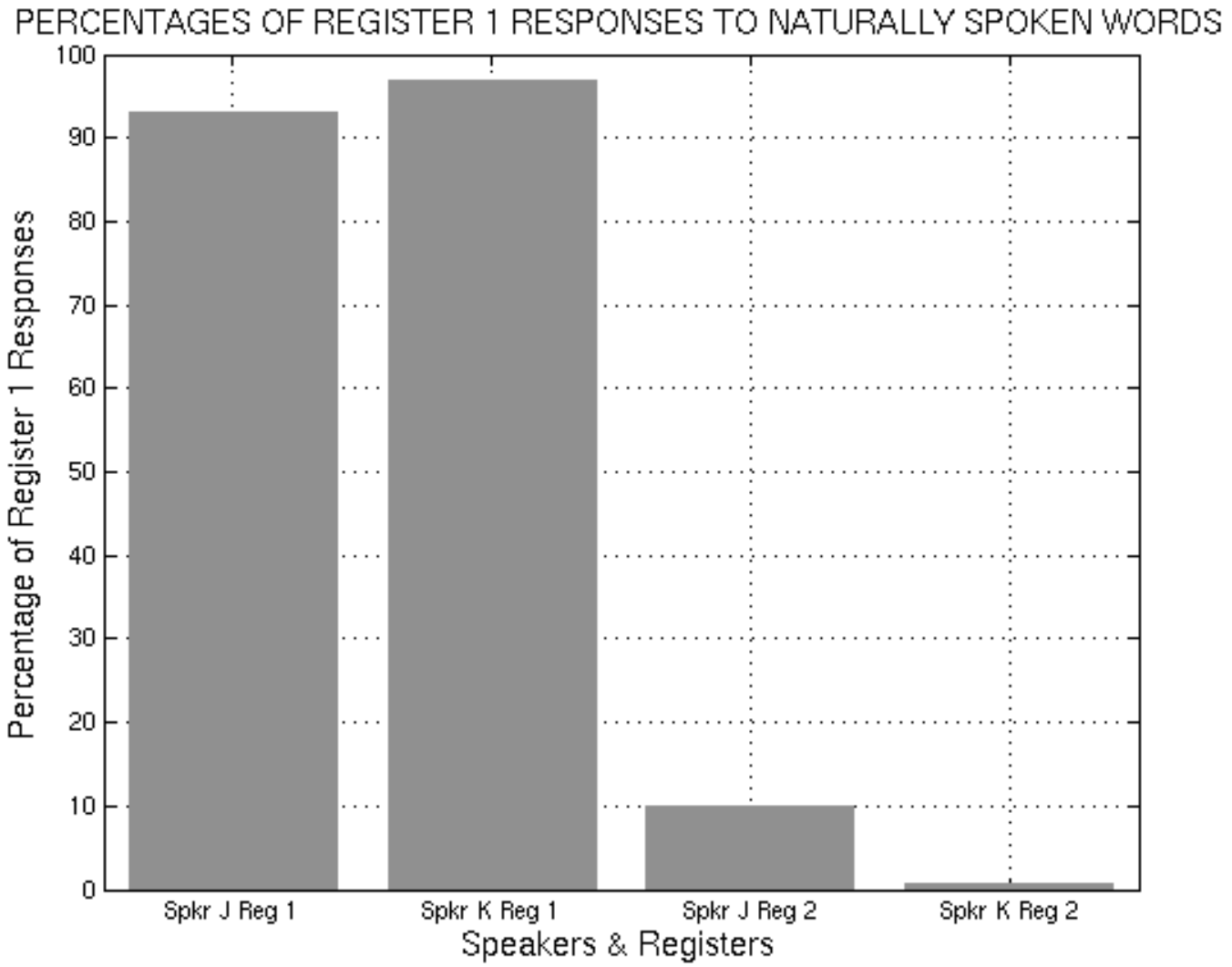


Fig. 12. The register identification scores of 32 listeners who heard a randomized presentation of natural exemplars of the two registers produced by male and female speakers.

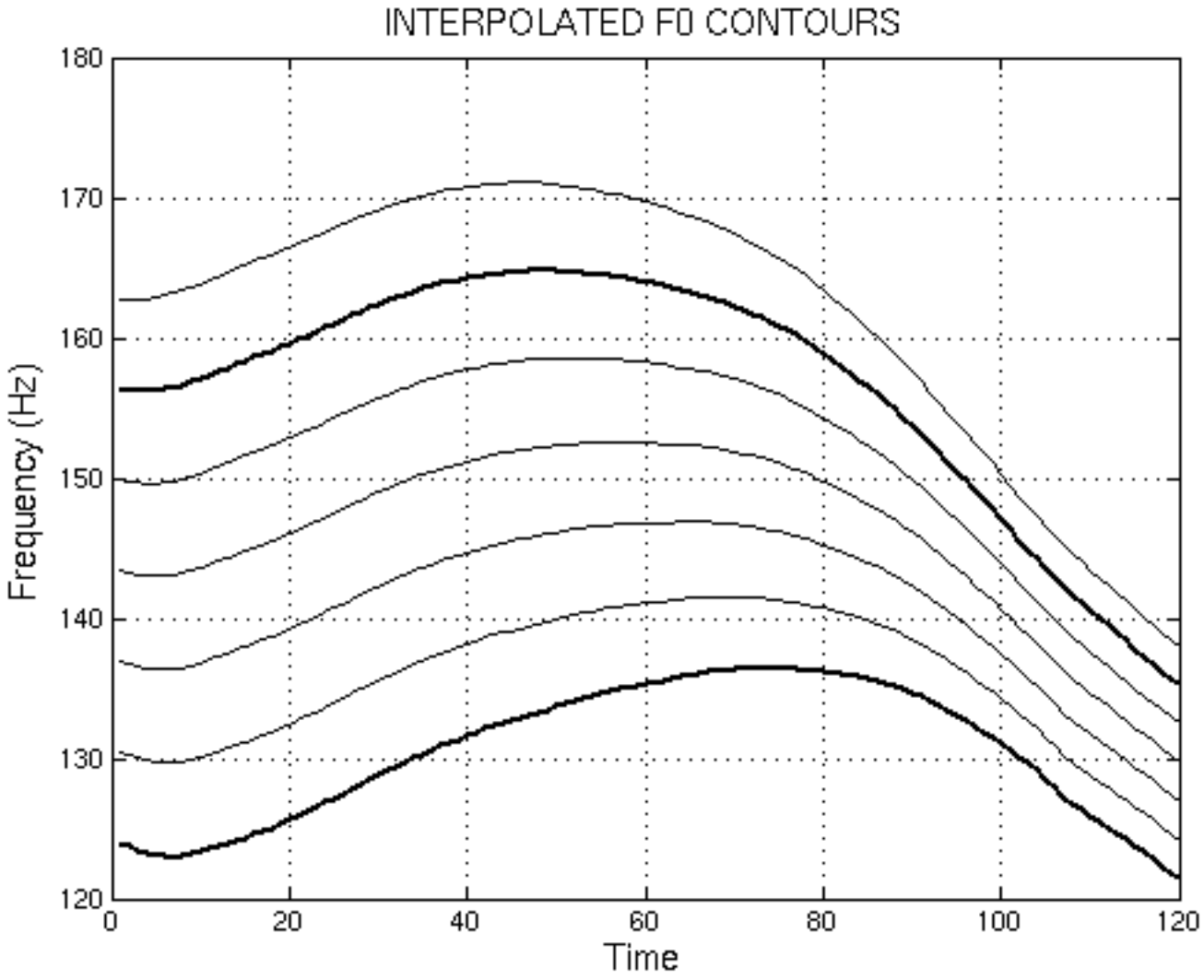


Fig. 13. The series of F0 contours imposed on the synthesized syllable [raːŋ]. The thick lines show the limits of the set before the addition of the super stimulus (upper line).

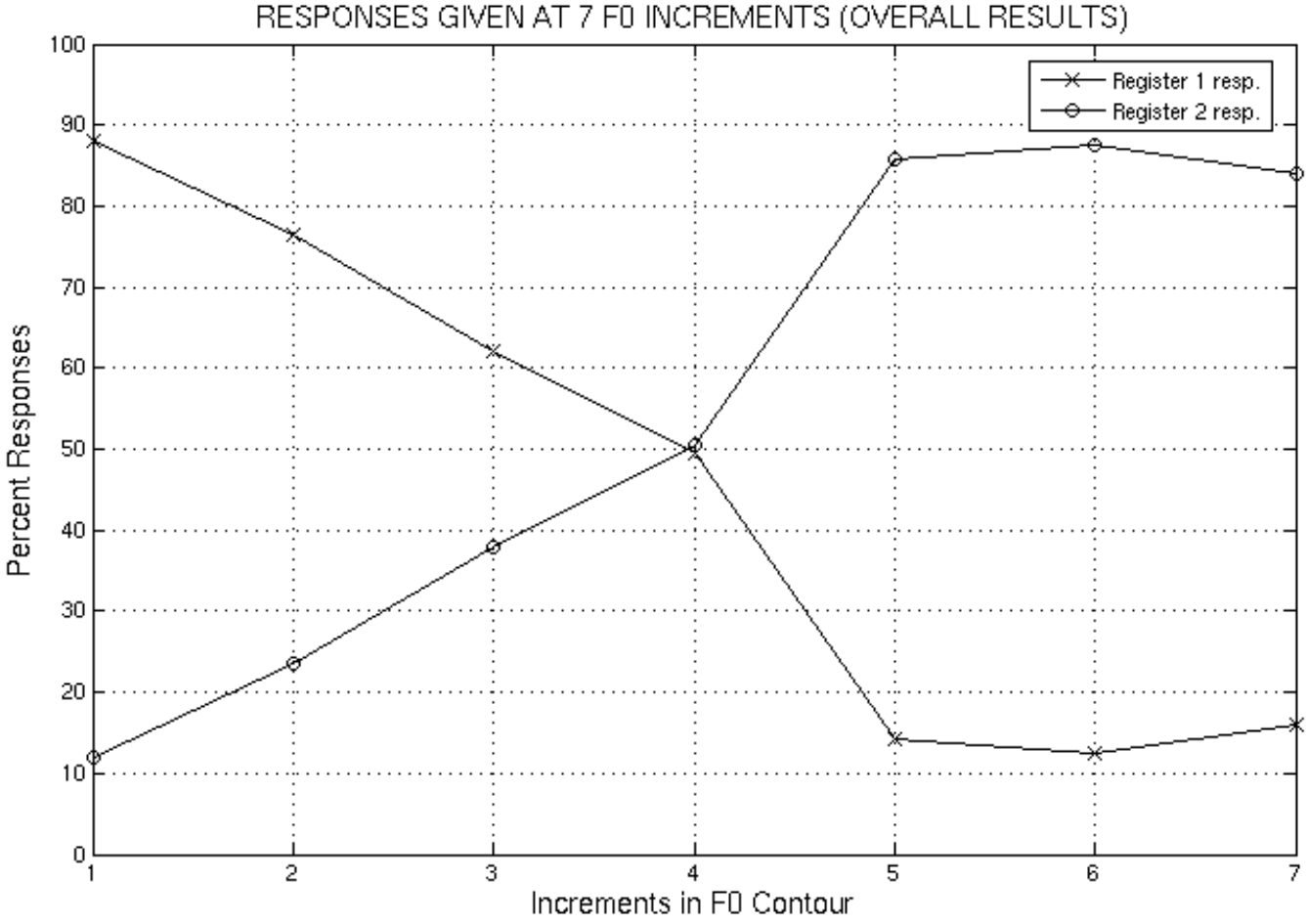


Fig. 14. The percentages of Register 1 vs. Register 2 responses of 29 listeners to incremental changes in the F0 contours of synthesized versions of the Khmu' syllable [raːŋ].

TABLES

Table 1. A list of Khmu' word-pairs spoken from 3 to 6 times each by 25 native speakers

Register 1		Register 2	
Transcription	Gloss	Transcription	Gloss
khɛ́ɛm	Johnson grass	khɛ̀ɛm	waterway edge
lɔ́ɔk	haunted	lɔ̀ɔk	to peel off
már	sterile	màr	snake
míəŋ	fermented tea leaves	mi̖əŋ	to chew
phák	vegetable	phàk	to ride
phɛ́ɛ	wound	phɛ̀ɛ	raft
phlúʔ	firework	phlùʔ	thigh
ráaŋ	tooth	ràaŋ	flower
wák	to hang	wàk	abyss

Table 2. The probability *p* that the hypothesis of a zero difference between the overall amplitude gradients of Registers 1 & 2 is true; analysis performed in four sectors of time-normalized utterance duration. Results are shown for a group of 25 speakers and four sub-groups of those speakers.

Groups	Source	Sector 1				Sector 2			
		F(d.f.)	F	MSe	p	F(d.f.)	F	MSe	p
All Spkrs	Register	(1,24)	8.529	0.206	0.0075	(1,24)	5.647	0.055	0.0258
	Word	(8,192)	19.503	0.791	0.0001	(8,192)	93.487	1.507	0.0001
	Reg x Word	(8,192)	2.001	0.032	0.0482	(8,192)	2.349	0.016	0.0295
Male	Register	(1,8)	7.431	0.125	0.0260	(1,8)	18.067	0.058	0.0028
	Word	(8,64)	18.993	0.594	0.0001	(8,64)	31.556	0.533	0.0001
	Reg x Word	(8,64)	0.898	0.014	0.5236	(8,64)	1.668	0.009	0.1570
Female	Register	(1,15)	3.153	0.092	0.0961	(1,15)	0.972	0.012	0.3398
	Word	(8,120)	8.526	0.333	0.0001	(8,120)	63.357	0.991	0.0001
	Reg x Word	(8,120)	1.872	0.032	0.0883	(8,120)	2.040	0.015	0.0473
Old	Register	(1,5)	1.168	0.013	0.3292	(1,5)	41.177	0.088	0.0014
	Word	(8,40)	4.141	0.166	0.0011	(8,40)	22.509	0.313	0.0001
	Reg x Word	(8,40)	0.832	0.016	0.5796	(8,40)	3.220	0.015	0.0064
Young	Register	(1,4)	15.358	0.182	0.0173	(1,4)	14.316	0.025	0.0194
	Word	(8,32)	10.976	0.458	0.0001	(8,32)	40.282	0.450	0.0001
	Reg x Word	(8,32)	0.512	0.009	0.8381	(8,32)	0.890	0.005	0.5361

Groups	Source	Sector 3				Sector 4			
		F(d.f.)	F	MSe	p	F(d.f.)	F	MSe	p
All Spkrs	Register	(1,24)	15.090	0.091	0.0007	(1,24)	0.102	0.005	0.7525
	Word	(8,192)	103.308	0.178	0.0001	(8,192)	10.629	1.182	0.0001
	Reg x Word	(8,192)	2.346	0.008	0.0244	(8,192)	1.954	0.079	0.0886
Male	Register	(1,8)	33.005	0.075	0.0004	(1,8)	1.089	0.030	0.3272
	Word	(8,64)	38.519	0.672	0.0001	(8,64)	4.088	0.224	0.0065
	Reg x Word	(8,64)	1.018	0.003	0.4317	(8,64)	2.142	0.067	0.1136
Female	Register	(1,15)	3.910	0.030	0.0667	(1,15)	0.027	0.002	0.8724
	Word	(8,120)	67.303	1.252	0.0001	(8,120)	18.944	1.779	0.0001
	Reg x Word	(8,120)	2.578	0.009	0.0124	(8,120)	1.212	0.054	0.2978
Old	Register	(1,5)	4.641	0.026	0.0838	(1,5)	0.100	0.001	0.7652
	Word	(8,40)	26.624	0.417	0.0001	(8,40)	0.760	0.073	0.6395
	Reg x Word	(8,40)	1.496	0.006	0.1893	(8,40)	2.280	0.079	0.0410
Young	Register	(1,4)	34.575	0.065	0.0042	(1,4)	4.076	0.126	0.1136
	Word	(8,32)	35.724	0.510	0.0001	(8,32)	4.120	0.414	0.0018
	Reg x Word	(8,32)	1.354	0.004	0.2540	(8,32)	0.433	0.024	0.8923

Table 3. The probability *p* that the hypothesis that of a zero difference between the F0 gradients of Registers 1 & 2 is true. The analysis was performed in four sectors of each time-normalized utterance duration. Results are shown for a group of 25 speakers and four sub-groups of those speakers.

Groups	Source	Sector 1				Sector 2			
		F(d.f.)	F	MSe	p	F(d.f.)	F	MSe	p
All Spkrs	Register	(1,24)	15.601	0.046	0.0006	(1,24)	107.74	0.068	0.0001
	Word	(8,192)	16.973	0.020	0.0001	(8,192)	20.482	0.015	0.0001
	Reg x Word	(8,192)	10.553	0.009	0.0001	(8,192)	35.905	0.010	0.0001
Male	Register	(1,8)	0.014	2.7E-5	0.9073	(1,8)	54.564	0.022	0.0001
	Word	(8,64)	9.510	0.008	0.0001	(8,64)	9.665	0.005	0.0001
	Reg x Word	(8,64)	4.823	0.003	0.0001	(8,64)	12.362	0.003	0.0001
Female	Register	(1,15)	32.787	0.070	0.0001	(1,15)	58.819	0.047	0.0001
	Word	(8,120)	9.421	0.013	0.0001	(8,120)	17.146	0.012	0.0001
	Reg x Word	(8,120)	8.366	0.008	0.0001	(8,120)	27.657	0.008	0.0001
Old	Register	(1,5)	0.013	3.1E-5	0.9141	(1,5)	73.418	0.025	0.0004
	Word	(8,40)	7.075	0.006	0.0001	(8,40)	3.007	0.003	0.0097
	Reg x Word	(8,40)	5.460	0.004	0.0001	(8,40)	5.507	0.002	0.0001
Young	Register	(1,4)	237.24	0.013	0.0001	(1,4)	21.960	0.005	0.0094
	Word	(8,32)	3.035	0.003	0.0116	(8,32)	8.880	0.003	0.0001
	Reg x Word	(8,32)	5.889	0.003	0.0001	(8,32)	14.587	0.002	0.0001

Groups	Source	Sector 3				Sector 4			
		F(d.f.)	F	MSe	p	F(d.f.)	F	MSe	p
All Spkrs	Register	(1,24)	257.02	0.217	0.0001	(1,24)	10.794	0.055	0.0031
	Word	(8,192)	118.80	0.178	0.0001	(8,192)	20.645	0.193	0.0001
	Reg x Word	(8,192)	13.367	0.006	0.0001	(8,192)	1.669	0.006	0.1083
Male	Register	(1,8)	86.231	0.052	0.0001	(1,8)	0.505	0.002	0.4975
	Word	(8,64)	73.970	0.062	0.0001	(8,64)	12.643	0.103	0.0001
	Reg x Word	(8,64)	7.719	0.003	0.0001	(8,64)	0.247	0.001	0.9799
Female	Register	(1,15)	227.6	0.170	0.0001	(1,15)	12.766	0.069	0.0028
	Word	(8,120)	64.722	0.009	0.0001	(8,120)	12.993	0.116	0.0001
	Reg x Word	(8,120)	7.419	0.003	0.0001	(8,120)	3.032	0.011	0.0039
Old	Register	(1,5)	42.921	0.052	0.0012	(1,5)	2.177	0.030	0.2001
	Word	(8,40)	25.711	0.048	0.0001	(8,40)	7.171	0.078	0.0001
	Reg x Word	(8,40)	3.391	0.001	0.0046	(8,40)	0.449	0.003	0.5930
Young	Register	(1,4)	102.27	0.034	0.0005	(1,4)	1.743	0.006	0.2573
	Word	(8,32)	21.119	0.026	0.0001	(8,32)	3.470	0.033	0.0055
	Reg x Word	(8,32)	4.818	0.003	0.0006	(8,32)	0.151	0.001	0.9956

Table 4. Given an hypothesis that the durations of vowels in Register 1 are drawn from the same population as corresponding values in Register 2, the column headed by the symbol p shows the probabilities that the hypothesis is true for different groups of speakers.

Groups	Avg Reg 1 (sec)	Avg Reg 2 (sec)	Mean Diff (sec)	Variables	F(d.f.)	F	MSe	p
All Spkrs	0.2491	0.2714	0.022	Register	(1,25)	79.871	0.058	0.0001
				Word	(8,200)	469.60	0.595	0.0001
				Reg x Word	(8,200)	12.242	0.003	0.0001
Male	0.2126	0.2374	0.025	Register	(1,8)	38.084	0.025	0.0003
				Word	(8,64)	268.688	0.167	0.0001
				Reg x Word	(8,64)	3.971	0.001	0.0007
Female	0.2696	0.2905	0.021	Register	(1,15)	37.522	0.031	0.0001
				Word	(8,120)	303.169	0.418	0.0001
				Reg x Word	(8,120)	8.30	0.002	0.0001
Old	0.2408	0.2685	0.028	Register	(1,5)	44.035	0.021	0.0012
				Word	(8,40)	85.208	0.155	0.0001
				Reg x Word	(8,40)	3.028	0.001	0.0093
Young	0.2388	0.2597	0.021	Register	(1,4)	22.001	0.010	0.0094
				Word	(8,32)	68.421	0.107	0.0001
				Reg x Word	(8,32)	7.028	0.001	0.0001

Table 5. Given an hypothesis that Formant 1 and Formant 2 frequencies in Register 1 are indistinguishable from their corresponding values in Register 2, the columns headed by the symbol p show the probabilities that the hypothesis is true for different groups of speakers.

Groups	Analysis of Variance							
	Variables	F(d.f.)	Formant 1			Formant 2		
			F	MSe	p	F	MSe	p
All Spkrs	Register	(1, 24)	2.60	5287.21	0.1197	1.185	121407	0.2871
	Word	(8, 192)	163.86	21E+5	0.0001	160.496	19E+6	0.0001
	Reg x Word	(8, 192)	8.75	16629.4	0.0001	0.779	82740.1	0.6213
Male	Register	(1, 8)	1.56	598.965	0.2469	0.245	523.297	0.6337
	Word	(8, 64)	157.27	70E+5	0.0001	293.95	46E+5	0.0001
	Reg x Word	(8, 64)	1.55	603.80	0.1568	6.531	10462.1	0.0001
Female	Register	(1, 15)	4.66	11934.9	0.0475	1.100	175047	0.3109
	Word	(8, 120)	80.56	14E+5	0.0001	96.098	15E+6	0.0001
	Reg x Word	(8, 120)	9.18	21440.9	0.0001	0.866	142372	0.5469
Old	Register	(1, 5)	0.34	260.183	0.5868	0.753	32E+4	0.4253
	Word	(8, 40)	25.80	39E+4	0.0001	14.226	51E+5	0.0001
	Reg x Word	(8.40)	1.59	2905.683	0.1574	0.868	37E+4	0.5508
Young	Register	(1, 4)	1.11	1123.03	0.3513	1.362	2316.18	0.3081
	Word	(8, 32)	34.60	49E+4	0.0001	65.92	37E+5	0.0001
	Reg x Word	(8, 32)	3.23	1395.06	0.0084	2.947	4686.05	0.0138

Table 6. Given an hypothesis that the log-transformed ratios H_2/H_1, H_{F1}/H_1, & H_{F2}/H_1 are independent of voice register, the column s headed by the symbol p show the probabilities that the hypothesis is true for different groups of speakers.

Groups	Source	F(d.f.)	H_2/H_1			H_{F1}/H_1			H_{F2}/H_1		
			F	MSe	p	F	MSe	p	F	MSe	p
All Spkrs	Register	(1,23)	14.681	1.851	0.0009	12.049	4.452	0.0017	1.327	0.385	0.2611
	Word	(8,184)	5.268	0.349	0.0001	12.769	1.542	0.0001	78.207	13.820	0.0001
	Reg x Word	(8,184)	1.601	0.085	0.1269	0.840	0.043	0.8247	3.790	0.324	0.0068
Male	Register	(1,7)	0.621	0.073	0.4566	4.806	0.754	0.0645	0.802	0.138	0.4002
	Word	(8,56)	6.092	0.184	0.0001	14.406	0.599	0.0001	56.952	4.983	0.0001
	Reg x Word	(8,56)	3.306	0.044	0.0037	2.412	0.045	0.0259	3.962	0.116	0.0009
Female	Register	(1,15)	19.442	2.176	0.0005	8.508	3.882	0.0106	0.678	0.247	0.4232
	Word	(8,120)	5.508	0.396	0.0001	7.740	1.000	0.0001	41.816	9.044	0.0001
	Reg x Word	(8,120)	1.394	0.099	0.2061	0.646	0.069	0.7380	2.771	0.307	0.0404
Old	Register	(1,5)	1.522	0.181	0.2721	2.636	1.195	0.1654	0.123	0.034	0.7397
	Word	(8,40)	1.994	0.097	0.0723	5.137	0.414	0.0002	17.682	2.599	0.0001
	Reg x Word	(8,40)	1.185	0.019	0.3319	0.792	0.048	0.6127	0.959	0.060	0.4809
Young	Register	(1,3)	16.768	0.820	0.0263	6.658	1.296	0.0818	1.537	0.510	0.3033
	Word	(8,24)	1.125	0.065	0.3821	8.627	0.587	0.0001	25.362	3.097	0.0004
	Reg x Word	(8,24)	2.037	0.080	0.0850	1.741	0.074	0.1400	3.225	0.198	0.0124

REFERENCES

Abramson, A. S.: The vowels and tones of Standard Thai: acoustical measurements and experiments. Indiana University Research Center in Anthropology, Folklore, & Linguistics, Pub.20, Bloomington [=International Journal of American Linguistics, *28,* No.2, Part III]. (1962).

Abramson, A. S.: The plausibility of phonetic explanations of tonogenesis; in Fant, Fujisaki, Cao, Xu, From traditional phonology to modern speech processing: Festschrift for Professor Wu Zongji's 95th birthday, pp. 17–29 (Foreign Language Teaching and Research Press, Beijing 2004).

Abramson, A.S.; L-Luangthongkum, T.; Nye, P.W.: Voice register in Suai (Kuai): an analysis of perceptual and acoustic data. Phonetica *61:* 147–171 (2004).

Brunelle, M.: Register in Eastern Cham: phonological, phonetic and sociolinguistic approaches; doct. diss. Cornell University, Ithaca (unpublished, 2005).

Cohen, J.D.; MacWhinney, B.; Flatt, M.; Provost, J.: PsyScope: A new graphic interactive environment for designing psychology experiments. Behavioral Research Methods, Instruments and Computers *25:* 257–271 (1993). See also http://psyscope.psy.cmu.edu/.

Fant, G.; Kruckenberg, A.; Liljencrantz, J.: The source-filter frame of reference. Phonetica *57,* 113–127 (2000).

Fant, G.; Kruckenberg, A.: Analysis and synthesis of Swedish prosody with outlooks on production and perception; in Fant, Fujisaki, Cao, Xu, From traditional phonology to modern speech processing: Festschrift for Professor Wu Zongji's 95th birthday, pp. 73–95 (Foreign Language Teaching and Research Press, Beijing 2004).

Ferlus, M.: Formation des registres et mutations consonantiques dans les langues Mon-Khmer. Mon-Khmer Studies *8,* 1–76 (1980).

Fischer-Jørgensen, E.: Phonetic analysis of breathy (murmured) vowels in Gujarati. Indian Linguistics *28:* 71–139 (1970).

Flanagan, J.L.: A difference limen for vowel formant frequency. J. acoust. Soc. Am. *27:* 149–174 (1955a) (Letter).

Flanagan, J.L.: Difference limen for the intensity of a vowel sound. J. acoust. Soc. Am. *27:* 1223–1225 (1955b) (Letter).

Gandour, J.: On the interaction between tone and vowel length: evidence from Thai dialects. Phonetica *34:* 54–65 (1977).

Gandour, J.T.; Gårding, E.; Lindell, K.: Tones in Northern Kammu: a perceptual investigation. Acta Orientalia *39:* 181–189 (1978).

Gårding, E.; Lindell, K.: Tones in Northern Kammu: a phonetic investigation. Acta Orientalia *38:* 321–332 (1977).

Gordon, M.; Ladefoged, P.: Phonation types: a cross-linguistic overview. J. Phonet. *29:* 382–406 (2001).

Henderson, E.: The main features of Cambodian pronunciation. Bull. School Orient. Stud. *17:* 149–174 (1952).

Laver, J.: The phonetic description of voice quality (Cambridge University Press, Cambridge, 1980).

L-Thongkum,[2] T.; Teeranon, P.; Intajamornrak, C.: The interaction between vowel length and pitch in Southeast Asian (SEA) languages: implications for tonogenesis; in Harris, Burusphat, Harris, Studies in Tai and Southeast Asian linguistics, pp. 335–240 (Ekphim Thai, Bangkok, 2007).

Mermelstein, P.: Difference limens for formant frequencies of steady-state and consonant-bound vowels. J. acoust. Soc. Am. *63:* 572–580 (1978).

[2] In the References, L-Thongkum, T. and Thongkum, T.L. are the same person, Theraphan Luangthongkum.

O'Shaughnessy, D.: Speech communications: human and machine (IEEE Press, Piscataway, N.J., 2000).

Premsrirat, S.: Khmu, a minority language of Thailand (a Khmu grammar and a study of Thai and Khmu cutting words). Papers in Southeast Asian Linguistics, No. 10, Pacific Linguistic Series A-No. 75 (1987).

Premsrirat, S.: Dictionary of Khmu in Thailand (Institute of Language and Culture for Rural Development, Mahidol University, Nakhon Pathom, Thailand, 2002).

Premsrirat, S.: Khmu dialects: a case of register complex and tonogenesis; in Kaji, Proceedings of the symposium Cross Language Studies of Tonal Phenomena: Historical Development, Phonetics of Tone, and Descriptive Studies, Dec. 17–19, 2002, pp. 13–36 (Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies, Tokyo, 2003).

Smalley, W.A.: Outline of Khmuʔ structure (American Oriental Society, New Haven, Connecticut, 1961).

Stevens, K.N.: Acoustic phonetics (The MIT Press, Cambridge, Massachusetts, 1998).

Svantesson, J-O.: Tonogenesis in Southeast Asia—Mon-Khmer and beyond; in Kaji, Proceedings of the symposium Cross Language Studies of Tonal Phenomena: Tonogenesis, Japanese accentology, and other topics, Dec. 12–14, 2000, pp. 58 (Research Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies, Tokyo, 2001).

Svantesson, J-O.; House, D.: Tone, production, tone perception and Kammu tonogenesis. Phonology *23:* 309–333 (2006).

Thongkum, T.L.: Phonation types in Mon-Khmer languages; in Fujimura, Vocal fold physiology: voice productions, mechanisms and functions, vol.2, pp. 319–333 (Raven Press, New York, 1988).

Thongkum, T.L.: An acoustical study of the register complex in Kui (Suai). Mon-Khmer Stud. *15:* 1–19 (1989).

Thongkum, T.L.: The interaction between pitch and phonation type in Mon: implications for a theory of tonogenesis. Mon-Khmer Studies, *16–17,* 11–24 (1990).

Thongkum, T.L.: An instrumental study of Chong registers; in Davidson, Austroasiatic languages: essays in honor of H.L. Shorto, pp. 141–160 (School of Oriental and African Studies. University of London, London, 1991).

Thurgood, G.: Vietnamese and tonogenesis: revising the model and the analysis. Diachronica *19:* 333–363 (2003).

Thurgood, G.: Tonogenesis revisited: reviving the model and the analysis; in Harris, Burusphat, Harris, Studies in Tai and Southeast Asian linguistics, pp. 241–262 (Ekphim Thai, Bangkok, 2007)

Wayland, R.; Jongman, A.: Acoustic correlates of breathy and clear vowels: the case of Khmer. J. Phonet. *21:* 181–201 (2003).

Yu, A.C.L.: Tonal effects on perceived vowel duration; in Fougeron, Kühnert, D'Imperio, Vallée, Papers in Laboratory Phonology X: Variation, phonetic detail and phonological modeling (Mouton de Gruyter, Dordrecht to appear).

Appendix

The phonemes of Khmu' Rawk

Consonants

Initials					Examples of initial consonant clusters
p	t	c	k	ʔ	C_1C_2
ph	th	ch	kh		mp nr ns sʔ
ʔb/ʔm	ʔd/ʔn				$C_1C_2C_3$
m	n	ɲ	ŋ		ntr nsr ŋkw mpl
f	s			h	
	r				
	l				
w		j			
		ʔj			

Finals

p	t	c	k	ʔ
m	n	ɲ	ŋ	
w		j		
		ç/jh		h
	r			
	l			

Vowels

i	e	ɛ	ɨ	ə	a	u	o	ɔ
ii	ee	ɛɛ	ɨɨ	əə	aa	uu	oo	ɔɔ
iə	ɨə	uə						

Registers
R1=clear voice, high pitch R2=breathy voice, low pitch
(Considerable variation as to presence of phonation types)

Figure Legends

Fig. 1. Average amplitude contours for all Register 1 and Register 2 utterances produced by 25 Khmu' speakers of both sexes. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

Fig. 2. Average amplitude contours for all Register 1 and Register 2 utterances produced by 9 male Khmu' speakers. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

Fig. 3. Average amplitude contours for all Register1 and Register 2 utterances produced by 16 female Khmu' speakers. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

Fig. 4. Average amplitude contours for all Register 1 and Register 2 utterances produced by 5 young Khmu' speakers. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

Fig. 5. Average amplitude contours for all Register 1 and Register 2 utterances produced by 6 older Khmu' speakers. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

Fig. 6. Normalized pitch contours averaged over the Register 1 and Register 2 utterances of 25 Khmu' speakers of both sexes. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

Fig. 7. Normalized pitch contours averaged over the Register 1 and Register 2 utterances of 9 male speakers of Khmu'. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

Fig. 8. Normalized pitch contours averaged over the Register 1 and Register 2 utterances of 16 female speakers of Khmu'. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

Fig. 9. Normalized pitch contours averaged over the Register 1 and Register 2 utterances of 5 young speakers of Khmu'. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

Fig. 10. Normalized pitch contours averaged over the Register 1 and Register 2 utterances of 6 elderly speakers of Khmu'. The shaded regions within dashed lines and continuous lines indicate the standard deviations of the Register 1 and Register 2 contours respectively.

Fig. 11. Distributions of log ratio values after the elimination of all instances of synchronization failure prior to the application of an ANOVA. From top to bottom are shown in Graph 1 the combined

distribution of all ratio values, in Graph 2 the distribution of H_2/H_1 ratio values, in Graph 3 the distribution of H_{F1}/H_1 ratio values, and in Graph 4 the distribution of H_{F2}/H_1 ratio values.

Fig. 12. The register identification scores of 32 listeners who heard a randomized presentation of natural exemplars of the two registers produced by male and female speakers.

Fig. 13. The series of F0 contours imposed on the synthesized syllable [raːŋ]. The thick lines show the limits of the set before the addition of the super stimulus (upper line).

Fig. 14. The percentages of Register 1 vs. Register 2 responses of 29 listeners to incremental changes in the F0 contours of synthesized versions of the Khmu' syllable [raːŋ].

Tonal Evolution Induced by Language Contact:

A Case Study of the T'in (Lua') Language of Nan Province,

Northern Thailand[1]

Theraphan L-Thongkum and Chommanad Intajamornrak

Chulalongkorn University

0. Abstract

T'in is a Mon-Khmer language; generally, Mon-Khmer languages are non-tonal. The T'in (Lua') language of Bo Kluea District, Nan Province, comprises two major dialects, Mal and Pray. Pray is the more conservative and has fewer speakers, due to its geography and communist infiltration in the past. The data on Mal, Pray and Tai Yuan (language of the majority) was collected in 2005 for the research project on "Linguistic Diversity in Nan Province: A Foundation for Tourism Development".

Although different pitches can be heard in both Mal and Pray, there are no consistent patterns in Pray, even in the speech of a single speaker. In Mal, two pitches [high-falling] and [low-rising] have been found to be distinctive. It can be concluded that the Mal variety of Bo Kluea District has acquired two tones, /falling/ vs. /rising/ or /high/ vs. /low/. The falling tone occurs mostly in native words, whereas the rising tone mostly occurs in Tai loanwords. However, a few

[1] Paper presented at the Mon-Khmer Workshop organized by the Language and Cognition Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands, April 16th-18th, 2007

native words can have the rising tone and a few Tai loanwords can have the falling tone. Acoustical measurements were done to confirm the characteristics of Mal and Thai Yuan tones.

Filbeck (1972) pointed out the fact that a sub-dialect of Mal had two tones, i.e. rising and non-rising, and that the emergence of the rising tone appeared to be the result of independent innovation. The results of our present study indicate that the two rising tones in Tai Yuan, namely, the /low-rising/ (A1-2) and /mid-rising/ (A3-4), seem to induce tone birth in Mal, especially, the rising tone. In Tai Yuan, a six-tone language, words having the two rising tones (A1-2 and A3-4) outnumber the rest and are, thus, frequently heard and borrowed by Mal people, almost all of whom are bilingual. Perhaps, the auditory factor could have inspired and induced tonal evolution in some Mal sub-dialects.

1. Introduction

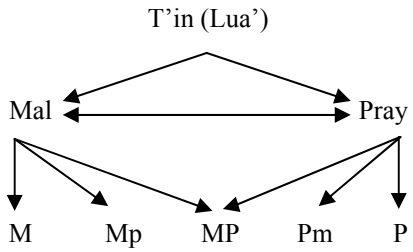
The T'in or Lua' live in Nan Province, Thailand and in the adjacent Sayabouri Province of Lao PDR. The T'in language comprises two major dialects, Mal and Pray, and each dialect consists of a few sub-dialects or varieties. T'in belongs to the Khmuic branch of the Mon-Khmer language family. In Nan Province, northern Thailand, there are about 34,600 speakers of the T'in language. Ten districts (Amphoes) are inhabited by the T'in: Bo Kluea, Pua, Chaloem Pra Kiat, Chiang Klang, Thung Chang, Santi Suk, Mueang Nan, Mae Charim, Wiang Sa and Song Khwae. See further details in Table 1.

Table 1 The number of T'in or Lua' (Mal-Pray) speakers in Nan province (a survey conducted by the research project on "Linguistic Diversity in Nan province: A Foundation for Tourism Development, 2004-2006")

District	*No. of sub-district*	*No. of village*	*No. of speakers*
Bo Kluea	4	31	10,808
Pua	3	20	8,483
Chaloem Pra Kiat	3	14	5,848
Chiang Klang	4	14	3,987
Thung Chang	4	7	2,398
Santi Suk	1	2	1,504
Mueang Nan	2	3	777
Mae Charim	1	1	370
Wiang Sa	1	1	347
Song Khwae	2	5	78
Total	*25*	*98*	*34,600*

Linguists who have worked on T'in do not agree on the number of sub-dialects or varieties of the Mal and Pray dialects. Filbeck (1978) says that Mal has three sub-dialects while Pray has only two. Jirananthanaporn (1993) who did a lexical study of Mal and Pray claims that Mal comprises three sub-dialects, the same analysis as Filbeck (1978), but Pray has five instead of two. The difference between their analyses is caused by the different criteria used by them, i.e. phonological criteria by Filbeck and lexical by Jirananthanaporn. Based on our Village Database (GIS) and a supplement questionnaire devised for a survey of

the Mal-Pray language in 2006, the Mal-Pray language can be classified approximately into five intricate groups as illustrated in Figure 1.



M	=	Typical Mal
Mp	=	Basically Mal with some Pray elements
MP	=	Unable to identify clearly whether Mal or Pray
Pm	=	Basically Pray with some Mal elements
P	=	Typical Pray

Figure 1 A synchronic classification of the T'in (Mal-Pray) language

Proto T'in which was reconstructed by Filbeck (1978) had no tones (non-tonal). However, two tones, namely, the non-rising (or level) and rising tones have been developed in some of the modern Mal sub-dialects. He concludes that the development of tones in Mal is an independent innovation due to language contact with Thai. In this paper, we would like to argue and prove that the two rising tones, /23/ and /35/, in the Kham Mueang dialect of Nan Province have induced tonal evolution in Mal. An acoustic study of the Kham Mueang and Mal tones was done to confirm their tonal characteristics; and the Tai-Thai loanwords in Mal spoken at Yot Doi Watthana Village in Bo Kluea Tai Sub-district, Bo Kluea District were compared with the original Tai-Thai words.

2. Language scenario of Nan Province

Nan is a small northern province of mostly mountainous areas (85%) and a population of 478,080 (Department of Provincial Administration, Ministry of Interior, 2005). The population of Nan comprises thirteen ethnic groups: Khon Mueang (Tai Yuan), Lue, Khuen, Lao, Phuan, Haw' (speaking Kham Mueang), Hakka Chinese, Mpi (Kaw), Hmong, Mien (Yao), Khmu', T'in (Mal-Pray) and Mlabri (Tong Lueang). The number of population of each ethnic group can be found in Table 2.

Table 2 The languages and number of population of the 13 ethnic groups of Nan Province (a survey conducted by the research project on "Linguistic Diversity in Nan province: A Foundation for Tourism Development, 2004-2006")

Language family	*Ethnic group*	*No. of population*	*Percentage*
Tai	Khon Mueang (Tai Yuan)	315,382	65.97
	Lue	54,400	11.38
	Lao	9,672	2.02
	Phuan	2,296	0.48
	Haw'	1,047	0.22
	Khuen	807	0.17
Sino-Tibetan	Hakka Chinese	2,135	0.45
	Mpi (Kaw)	508	0.11
Hmong-Mien	Hmong	22,037	4.16

	Mien (Yao)	11,415	2.39
Mon-Khmer	T'in/Lua' (Mal-Pray)	34,600	7.24
	Khmu'	7,708	1.61
	Mlabri	148	0.03
Unable to identify	Miscellaneous	15,925	3.33
Total		*478,080*	*100*

From Table 2 as shown above, one can see clearly that Nan is a multilingual society. Kham Mueang (Tai Yuan or Northern Thai dialect) is the lingua franca of the province. Most Nan people can understand Standard Thai which is the national language of Thailand. At home and in the village where the whole or majority of the villagers belongs to the same ethnic group, they speak their own mother tongue. Kham Mueang is used by the members of all ethnic groups at the markets and with outsiders. Standard Thai is used as the medium of instruction in all schools and also as the mass media language. School children speak Standard Thai with their teachers and Kham Mueang or their own mother tongue with their schoolmates. With officials, strangers and tourists, Standard Thai seems to be their preferred choice. When the minority groups want to hide their ethnic identity, Kham Mueang is spoken.

So far as language selection is concerned, T'in people behave in the same way as the other ethnic or ethnolinguistic groups in Nan Province. Generally, ethnicity consists of life style, culture, language, identity and so on. For the T'in, language seems to be the most important ethnicity component and unlike the Hmong, Mien and Lue, their cultural aspects, e.g. clothing, have no role to play in the tourism business of Nan Province. T'in people are bilingual, i.e. they can

speak Kham Mueang as well as their own mother tongue. Some of them, especially the well-educated ones, can speak Standard Thai quite well.

In a language-contact situation as stated above, linguistic borrowings, especially, lexical borrowing cannot be avoided. As a result, T'in (recipient language) has a lot of Tai-Thai loanwords. This means that Kham Mueang or Tai Yuan, Standard Thai and probably Lao are the source languages or donor languages. Tai-Thai words have been loaned to T'in at different periods of time for many reasons and in various circumstances. The T'in can be regarded as the local inhabitants of the highland areas of Nan Province. They have lived in the remote mountainous areas of many districts of Nan for a long time. However, during the Indo-China war and the Communist infiltration forty years ago, a large number of the T'in were moved down to resettlement camps in the lowlands and those who joined the Communist Party of Thailand stayed on in the mountains. Later, when the base of the Communist Party in Nan was devastated and the areas became peaceful, some of the T'in returned to their mountain homeland. The so-called "modernization" and "rural development" schemes of Thai governments and the private sector have reinforced both language and cultural borrowings. As viewed by the functionalists, "languages are primarily complex tools for referring to the world and for communication, and these tools easily adapt to new communicative and referential needs" (Appel and Muysken, 1988:154).

3. Tai-Thai loanwords

Filbeck (1972) points out the fact that some Mal sub-dialects have developed prosodic features to the point that a tonal analysis is necessary (p.111); "the non-contrastive intonation pitches would form the contrastive unit to the rising tone" (p.112). He lists a few Tai-Thai loanwords in Mal which have the rising tone. In

1997, Filbeck wrote another paper entitled “Thai loanwords in the Lua Language of Muang Nan” (unpublished) to bring attention to some of the implications for historical study.

Using the wordlist devised by our research project, 2,452 Mal lexical items were collected at Yot Doi Watthana Village by the first author in 2005. Although our major aim is to compile a basic vocabulary used in everyday life, a lot of Tai-Thai loanwords with the rising tone have been found. The following are some examples:

Standard Thai	*Kham Mueang (Nan)*	*Mal*	*Gloss*
(ŋaw^{24})	ŋ□□m^{31}	ŋ□□□m	‘lonely’
phaan^{21}	(kaay^{23})	phăan	‘to pass’
tun^{21}	tun^{33}	tŭun	‘bamboo rat’
kl□□ŋ^{33}	k□□ŋ^{23}	k□□□ŋ	‘drum’
kliip^{21}	kiip^{33}	klĭip	‘segment (of an orange)’
k□□t^{21}	(wan^{23})	k□□□t	‘to embrace’
kəən^{33}	(lam^{43})	kə□ən	‘too much’
khluy^{21}	khuy^{33}	khŭy	‘flute’
khaam^{42}	khaam$^{44•}$	khăam	‘to cross over’
khaaw^{21}	khaaw^{33}	khăaw	‘news’
ŋaam^{42}	ŋaam^{31}	ŋăam	‘fork (of a tree)’
caan^{33}	caan^{23}	căan	‘plate, dish’
ch□□^{42}	c□□^{31}	cĕe	‘to soak’
(din^{33} p□□n^{33})	faw^{33}	phăw	‘gunpowder’

daap^{21}	daap^{33}	dăap	‘sword’
(du□^{21} raay^{45})	suək^{33}	sŭək	‘fierce’
takuə^{21}	(c□□n^{35})	takŭə	‘lead (metal)’
faay^{24}	faay^{23}	phwăay	‘dam’
ruəm^{42}	huəm^{31}	yŭəm	‘to join in’
yaak^{42}	□aak^{31}	□ăak	‘difficult’

From the above examples, it looks as if the low-rising pitch has been assigned to Tai-Thai loanwords. However, instead of the low-rising pitch, some Tai-Thai loanwords can have a falling tone like most of the native words, for example,

Standard Thai	*Kham Mueang (Nan)*	*Mal*	*Gloss*
khəəy^{33}	kəəy^{35}	kə□əy	‘used to’
leŋ^{33}	leŋ^{35}	lêeŋ	‘to take aim’
loon^{45}	loon^{43}	lôon	‘hairless’
woŋ^{33}	woŋ^{35}	wôŋ	‘circle’
suu^{42}	suu$^{44•}$	sûu	‘to fight’
m□□ŋ^{24}	(saw$^{44•}$)	hm□□□ŋ	‘gloomy’
maay^{24}	maay^{23}	hm□aay	‘to mark’
buəŋ^{21}	buəŋ^{33}	bûəŋ	‘loop, noose’
h□ŋ^{42}	h□ŋ$^{44•}$	h□□□ŋ	‘room’
rien^{33}	hiən^{35}	lîən	‘to study’
maa^{45}	maa^{43}	mâa	‘horse’
y□□m^{45}	□□□m^{43}	□□□□m	‘to dye’

yay^{33}	□ay^{35}	□ây	‘fiber’
f□□n^{45}	f□□n^{43}	ph□□□n	‘to dance’
than^{33}	tan^{35}	tân	‘in time’
faay^{42}	faay$^{44•}$	phwâay	‘cotton’
looŋ^{33}	looŋ^{35}	lôoŋ	‘coffin’
siəŋ^{24}	siəŋ^{23}	sîəŋ	‘sound, voice, noise’
tiəŋ^{33}	tiəŋ^{23}	tîəŋ	‘bed’
h□□p^{21}	h□□p^{33}	h□□□p	‘hoarse, rasping’

Regarding consonants and vowels, Mal does not need serious adaptation because the Mal consonant and vowel systems are more complex than those of the Tai-Thai languages. However it is worth pointing out that some Tai-Thai sounds can be a problem to Mon-Khmer speaking people, for example, the voiceless labio-dental fricative /f/ does not exist in the Mal phonological system, therefore, when Tai-Thai words having the initial /f/ are loaned to Mal, /f/ changes to /ph/. Since all of the Tai-Thai languages are tonal, adaptation is needed for a non-tonal language like Mal. In fact, Mal native words have different pitches--levels or contours--depending upon syllable structures and consonant types in the syllable. More details can be found in Filbeck (1972) and Singnoi (1988). We have tried to construct some rules for the adjustment of tones in Tai-Thai loanwords but our attempts have always failed. Perhaps, the present Tai-Thai loanwords with the falling tone are early loans whose tones were adapted to suit the pitch patterns of Mal. To avoid homophones which can cause lexical-meaning confusion, the rising tone has been assigned to recent loanwords.

4. Contact-induced tonal evolution

Filbeck (1972) says that the tonal system of a Mon-Khmer language can develop from some internal factors or by means of contact with a tonal language such as Thai (p.111). He postulates two contrastive tones, namely, the rising tone and the non-rising or level tone. The non-rising tone is described as "intonation" or "speech rhythm" (p.113). The Mal variety which he studied comprises three degrees of stress and three types of pitch which can be summarized as three units. See figure 2.

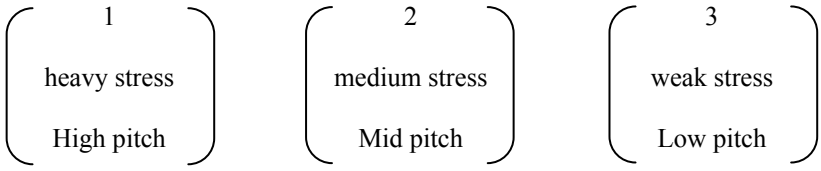


Figure 2 Interaction between stress and pitch in Mal

The rising tone stands in contrast to the above three allotones of the non-rising or level tone. He concludes that the Mal two tones have emerged through contact with Thai. Because of the difference in tone between the source language (Thai) and the recipient language (Mal), the emergence of the rising tone appears to have been the result of some independent development which is called "independent innovation" by Filbeck (p.115).

Singnoi (1988) makes a synchronic comparison of pitches occurring in different environments in Mal and Pray. However she does not conclude whether pitches in these two dialects of the T'in language are distinctive or not.

Based on the Mal data collected by the first author at Yot Doi Watthana Village in Bo Kluea Tai Sub-district, we must conclude that this Mal variety has two distinctive pitches or two tones, i.e. /FALLING/ vs. /RISING/ or /HIGH/ vs.

/LOW/. An acoustic study was done to confirm our analysis of the Mal tones. In this study, two female speakers, age 56 and 38, were asked to pronounce 20 minimal pairs. (See the Appendix.) The language data used for our acoustical measurements was recorded on a computer notebook and the acoustic analysis of pitches (fundamental frequencies) was done using the Praat program, version 4.5.06. The result of our acoustic study is shown in figure 3.

Suspecting that the two rising tones of Kham Mueang might be the inspiration or source of the rising tone in Mal, Kham Mueang data was also collected using the same wordlist. It has been found that the Kham Mueang or Northern Thai dialect spoken in Nan Province has six tones:

Tone 1	Low-rising	[23]	(A1-2)
Tone 2	Mid-rising	[35]	(A3-4, DS 1-2-3)
Tone 3	Mid-level	[33]	(B1-2-3, DL1-2-3)
Tone 4	Mid-falling	[31]	(B4, DL4)
Tone 5	High-level	[44’]	(C1-2-3), [44] (DS4)
Tone 6	High-falling	[43]	(C4)

The acoustic characteristics of these six tones can be found in figure 4.

It is noticeable that in the Kham Mueang dialect of Nan, words having the two rising tones (T1 and T2) outnumber the rest. Making a survey using the Matisoff’s 200-wordlist for SEA languages, 57% of Kham Mueang words bear the two rising tones. Frequently hearing and using these two rising tones in everyday life, the Mal-Kham Mueang bilinguals have borrowed the rising contour and assigned it to the more recent Tai-Thai loanwords no matter what the original tones of those words, since the old Tai-Thai loanwords have obtained typical native pitches. At a later stage, to avoid homophones and to make lexical distinctions, not only the Tai-Thai loanwords but also some of the native words can have the rising tone, and some of the loanwords can have the falling tone.

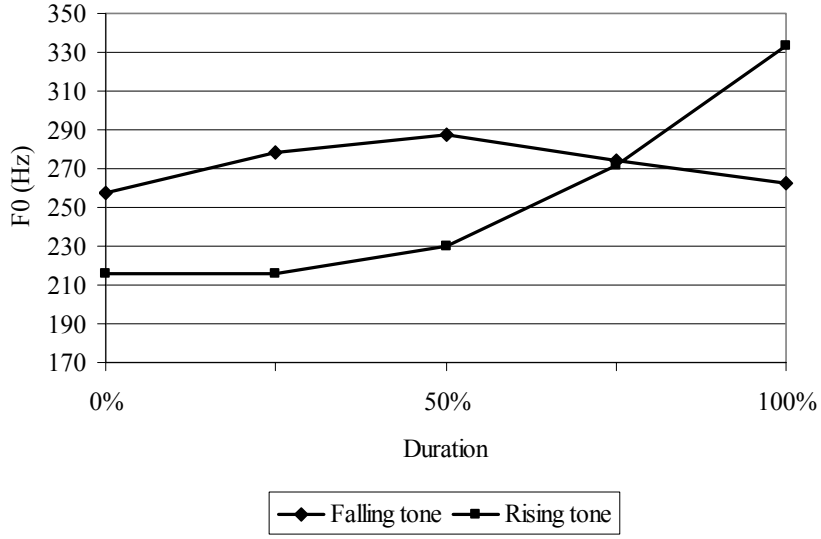


Figure 3 The acoustic characteristics of the two tones in Mal

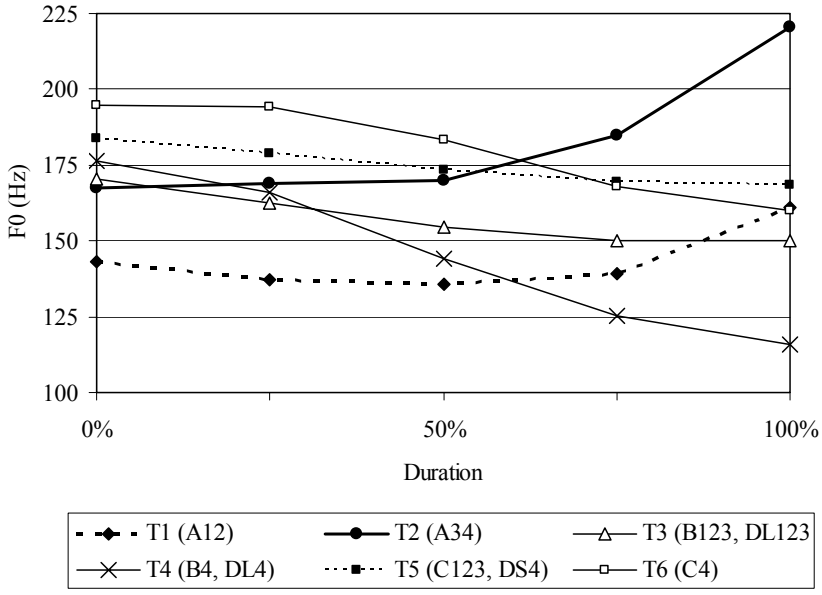


Figure 4 The acoustic characteristics of the six tones in Kham Mueang

The borrowing and adaptation processes could be a cause of the tonal distinctions in Mal. Consequently, several minimal pairs bearing the falling tone vs. the rising tone can be found in some Mal varieties. Some examples of the minimal

pairs in the Yot Doi Watthana variety are given in the Appendix. Although the falling tone in the other Mal varieties has a few variants, due to syllable-structure types, there is always a two-way tonal distinction, i.e. /Falling/ vs. /Rising/ or /High/ vs. /Low/.

5. Discussion and conclusion

An evidence of contact-induced tonogenesis or tonal evolution in Southern Qiang (Tibeto-Burman languages) has been presented by Evans (2001). He says that some of the SQ dialects becomes tone-prone due to lexical borrowings from Sichuanese Mandarin and some dialects, e.g. Longxi, Taoping, etc. have developed tonal systems. Our findings presented in this paper confirm the possibility of non-tonal languages becoming tonal due to language contact. Tone borrowing seems to be quite common in SEA language areas. A case study of the tonal evolution of Mal helps prove the hypothesis of contact-induced tonogenesis. The results of our present study indicate that the two rising tones of Kham Mueang, i.e. the low-rising tone (A1-2) and the mid-rising tone (A3-4), are the inspiration or inducement of tonogenesis in the Mal variety spoken in Yot Doi Watthana Village. Although different pitches can be heard in both Mal and Pray, the two major dialects of the T'in language, there are no consistent pitch patterns in Pray, even in the speech of a single speaker. Some varieties of modern Mal have become tonal with two-way distinction or two tones: /Falling/ vs. /Rising/ or /High/ vs. /Low/, depending upon the interpretation of each analyst.

Moreover, language contact can also induce tonal variation and change in SEA languages (L-Thongkum, 1997; Akharawatthanakun, 2002).

Acknowledgements

This paper is a research report of the project on "Linguistic Diversity in Nan Province: A Foundation for Tourism Development" funded by the Thailand Research Fund (TRF) from May 2004 to April 2007. We would like to thank the Nan Community and the local authorities of Nan Province for their kind assistance. A lot of thanks go to our Mal-Pray friends for their cooperation and, last but not least, to Supaporn Phalipat and Kanitha Putthasatien who helped type the manuscript.

References

Akharawatthanakun, Phinnarat. 2002. Tonal variations and changes in a language mixture area: a case study of Northeastern Thailand (Isan). *Manusya* 5.2 (September 2002): pp. 30-51.

Appel, René and Pieter Muysken. 1988. *Language contact and bilingualism*. London: Edward Arnold, a division of Hodder & Stoughton.

Department of Provincial Administration, Ministry of Interior. 2005. *Thailand population and house statistics*, *December 2005.* (Website: http://www.dopa.go.th/xstat/pop48-1.html)

Evans, Jonathan P. 2001.Contact-induced tonogenesis in Southern Qiang.*Language and Linguistics* 2.2: 63-110.

Filbeck, David. 1972. Tone in a dialect of T'in. *Anthropological linguistics*, vol.14, no.4, pp. 111-117.

Filbeck, David. 1978. *T'in: a historical study*. [Pacific linguistics series B-no.49]. Canberra: Department of Linguistics, Research School of Pacific Studies, the Australian National University.

Filbeck, David. 1997. *Thai loanwords in the Lua language of Muang Nan*. Seminar document paper for the conference "Suepsan Adit An

Rueangrong Khong Mueang Nan: Khomun Lae Mummong Mai Thang Borankhadi: Prawattisat Lae Chattiphan", organized by the Tai Studies Project, Payap University and the Nan Community, December 18-21, 1997. (Unpublished)

Jirananthanaporn, Supatra. 1993. Linguistic geography of minority groups in Nan Province: a lexical study of Pray and Mal dialects. Faculty of Humanities and Social Sciences, Naresuan University. (Unpublished research report)

L-Thongkum, Theraphan. 1997. Tone change and language contact: a case study of Mien-Yao and Thai. *Southeast Asian linguistic studies in honor of Vichin Panupong*, edited by Authur S.Abramson, 153-160. Bangkok: Chulalongkorn University Press.

Matisoff, James A. 1978. *Variational semantics in Tibeto-Burman: the "organic" approach to linguistic comparison*. Philadelphia, PENN: Institute for the Study of Human Issues.

Singnoi, Unchalee. 1988. A comparative study of Pray and Mal phonology. M.A. thesis, Faculty of Graduate Studies, Mahidol University. (Unpublished)

Appendix

1.	yâa	‘cigarette’	2.	hâa	‘five’
	yǎa	‘cholera’		hǎa	‘clf. for heavy rain’
3.	lâayh	‘to stir steamed sticky rice’	4.	s□□□t	‘sticky’
				s□□□t	‘to travel for pleasure’
	lǎayh	‘lawn’			
5.	wâaŋ	‘year’	6.	mân	‘female-in-law’
	wǎaŋ	‘to have free time’		măn	‘sterile’
7.	kâw	‘first, oldest (child)’	8.	môoŋ	‘o’clock (e.g. 6 ~)’
	kǎw	‘equal (e.g. length)’		mŏoŋ	‘to hit (a ball) with the forehand’
9.	pîi	‘Job’s tears (plant)’	10.	□âan	‘saddle’
	pǐi	‘oboe’		□ăan	‘to read’
11.	lôoŋ	‘coffin’	12.	sîət	‘odd (~ people)’
	lŏoŋ	‘clear and clean’		sĭət	‘kind of tree’
13.	sîəŋ	‘tooth’	14.	sây	‘to pinch off (e.g. small leaves)’
	sĭəŋ	‘to take a risk’		săy	‘to saw’
15.	kh□□□p	‘to pick up with	16.	h□□□p	‘dry (of the throat)’

		pincers’		h□□□p	‘tweezers’
	kh□□□p	‘kind of creeping plant’			
17.	pûut	‘liquor, wine’	18.	yâaŋ	‘flower’
	pŭut	‘kind of bird’		yăaŋ	‘deserted’
19.	th□□□t	‘to stab’	20.	kûut	‘to enter’
	th□□□t	‘to fry (e.g. a fish)’		kŭut	‘quail’

Paper to be presented at the Beijing Forum (2007), Panel Session on "Language Identity and Language Change in Collision and Dialogue Between Civilizations", Theme 3: Clashes of Civilizations and Consequent Language Change, organized by Beijing University, November 2-4, 2007.

Language Change without Collision: A Glimpse at Linguistic Diversity in Northern Thailand and Southern Laos

Theraphan L-Thongkum
Department of Linguistics, Faculty of Arts, Chulalongkorn University
Theraphan.L@chula.ac.th

Abstract

Linguistic diversity can be the cause of serious clashes and collisions in nations with multiethnic groups. Is this always true? Linguistic variation and change due to linguistic diversity can be subtle and be accepted without collision or negative feelings when induced gradually and naturally by language contact.

Linguistic diversity in the Nan Province of northern Thailand and in the Xekong Province of southern Laos is a good illustration of the latter phenomenon. In the case of Nan and Xekong, though small both in area and population size, these provinces have thirteen and fourteen ethnic groups, respectively, coexisting peacefully. These ethnic minorities have learned to tolerate, to care for and to listen to each other in order to maintain a sense of harmony in the area. Unity is the result of their linguistic tolerance and lack of selfishness. Due to language contact, linguistic variation and change cannot be avoided. Different languages play different roles in a multilingual community. Pluralism and multilingualism should be taken as important keywords for policy making at the national level. However, assimilation should also be encouraged; if people want to speak or conform to the majority instead of maintaining their own identity and speaking their own language, they have the right to choose whatever they think best suits them.

Introduction

Thailand, formally the Royal Kingdom of Siam, and Lao PDR, which is a socialist country, are parts of Mainland Southeast Asia or "Usa Kanay", the popularly used term among present-day scholars of pre-civilization. The two countries are located in the monsoon zone where monsoon winds bring a lot of heavy rain for about three months between July and September. Thailand and Laos are alike in that each of these two countries contains more than fifty ethnic groups.

Based on language and cultural differences, Thailand can be divided into four major regions: north, northeast, central and south. Each of these four regions has its own prominent characteristics. Standard Thai is the official or national language of Thailand. Lao PDR comprises three regions: north, central and south. Three cities, Luang Pra Bang, Vientiane and Champasak are the centres of northern, central and southern Laos, respectively. Similar to Thailand, each region has its own prominent characteristics due to the inhabitants of the areas. Nan is a small province in northern Thailand whereas Xekong is a small province in southern Laos.

From 1997-2007, we conducted two big research projects to investigate linguistic diversity in Xekong Province (1997-2001) and in Nan Province (2004-2007) funded by the Thailand Research Fund (TRF). Many linguistic aspects were studied. The two

provinces were selected because they have small populations but many ethnic minorities or tribes speaking their own native languages. Our interest was to discover how they communicate in everyday life and live happily together.

The contents presented in this paper are a by-product of our research findings which have already been published and presented at national and international conferences. Here, I would like to illustrate briefly two examples of language diversity and change without collision in two multilingual areas, one in southern Laos (Xekong Province) and the other in northern Thailand (Nan Province).

Xekong Province: Demography

Xekong Province (see Map 1) is divided into four districts, Tha Taeng, Lamam, Kaluem and Dak Chueng. Lamam is the centre of the province where all of the provincial offices are located. The Xekong River is the heart and soul of the people inhabiting Xekong Province.

According to official documents, there are fourteen "tribes" (in their own word: "Chon Phao") in Xekong Province: Lao Lum (lowland Lao), Taliang, Yae, Taliw, Dak Kang, Katu, Ngae, Chatong, Alak, Laven, Suai, Ta Oi, Lavi and Kaseng. The total population is 64,170 (Population Census, 1995), of which only 5% is Lao Lum and the remaining 95% is Lao Thoeng, comprising thirteen tribes or Mon-Khmer speaking peoples. Regarding their religious beliefs, the people living in towns or developed areas are Buddhists and those living in remote mountainous areas are animists. These thirteen Mon-Khmer ethnic groups played an important role during the two liberation wars against the French and Americans and some of their ancestors became national heroes. At present, many of them hold high administrative positions. Besides their own mother tongues, they also speak or at least comprehend Lao Lum (a Tai dialect), which is the official language of the province and also of the nation, and the other Lao Thoeng (Mon-Khmer) languages spoken in nearby villages. There follows a brief look at the multilingual society of Xekong Province in southern Laos. All the facts and figures presented in this paper are from the unpublished official documents given to us by the local authorities of Xekong Province.

Lao Lum

The population is 3,360. There are only three old Lao Lum villages in Xekong: Don Chan, Dan and Phon. They have recently moved from the other provinces. The Lao Lum population comprises mostly government employees, businessmen and the spouses of local people. Even though they belong to the majority ethnic group of the country, they are the minority in Xekong Province. However, Lao Lum is the lingua franca of the "tribes" or ethnic groups living in Xekong. It is used as the medium of instruction in all schools and government offices.

Katu

The population is 13,705. The Katu people of Xekong speak either the Katu or Kantu dialect. Each of these two dialects has many varieties. Most of the Katu live in Kaluem District. After the liberation war, some of them migrated from their homeland in the mountains to the districts of Tha Taeng and Lamam. Among the seventy-two Katu villages in Laos, fifty villages are located in Kaluem District, nineteen in Tha Taeng District and three in Lamam District. As well as in Xekong Province, the Katu also live in Saravan Province, which is adjacent to Xekong, and in Kwang Nam Province, Vietnam.

Triang (Tariang, Taliang)

The population is 12,399. There are sixty-seven Triang villages in Xekong Province, fifty-nine villages in Dak Chueng District, four in Tha Taeng District and four in Lamam District. The Triang language is the lingua franca of Dak Chueng District. In Lao PDR, the Triang live only in Xekong Province. However, the Triang can be found in Kwang Ngai Province, Vietnam, which is adjacent to Xekong. In fact, the Triang living in southern Laos call themselves "Khon Dak Chueng" which simply means 'the people of Dak Chueng'. They obtained the tribal name "Triang" by accident, due to a misunderstanding by French explorers in the past.

Harak (Halrak, Harlak, Alak)

They are mostly called "Alak". The population is 11,231. The Alak are the native people of Lamam District. In Xekong Province, there are altogether fifty-five Alak villages, thirty-five villages in Lamam and twenty in Tha Taeng. In the olden days, Tha Taeng used to be a sub-district of Lamam. Based on the sub-dialects or varieties spoken, the Alak can be divided into four sub-groups: the clusters of Ban Pak Thon, Ban Mo, Ban Kasang Klang and Ban Saen Nuea. Each of these four clusters consists of many villages. The Alak language is used as the lingua franca of Lamam District.

Since the Alak live in "Thung Phiang" or lowland areas, they have had a greater opportunity to come into contact with more civilized peoples. As a consequence, a lot of Khmer (Cambodian) and Lao words have been borrowed into Alak.

Kriang (Ng-Kriang, Ngae', Ngae)

The population is 6,784. Most of the Kriang live in Kaluem District. There are 37 Kriang villages, twenty-nine in Kaluem seven in Tha Taeng and one in Lamam. They call themselves "Kriang" or "Ng-Kriang" but are known as "Ngae" by the officials. The origin of the name "Ngae" is also the result of a misunderstanding by French explorers. Kriang women are good at weaving cotton cloth with animal motifs.

Yaeh (Yae', Yae)

The population is 4,916. The majority of the Yaeh live on both sides of the Xekaman River in Dak Chueng Province. Thirty Yaeh villages are located in Dak Chueng. There is only one Yaeh village in Lamam District, Ban Non Nong Va or Ban Thung Krachiaw. Yaeh people also live in Attapue, which is a province adjacent to Xekong, and in Charaikongtum, Vietnam.

Ta Oi

The population is 2,631. They are not native to Xekong but migrated from the Ta Oi District of Salavan Province, especially the group living at Ban Palaeng Nuea, Ban Palaeng Klang and Ban Palaeng Tai. The Ta Oi speak the Katuic branch of the Mon-Khmer language. Seven Ta Oi villages are located in Tha Taeng District and six Ta Oi villages are to be found in Kaluem District.

Suai

The population is 2,603. There are only four Suai villages in Xekong Province, all of which are located in Tha Taeng District. The Suai can be divided into two separate groups based on the languages spoken. The Suai of Ban Yok Thong and Ban Chunla speak the Bahnaric branch of the Mon-Khmer language whereas the Suai living in the

other two villages speak the Katuic branch of the Mon-Khmer language, which is similar to that spoken in Sukuman District, Salavan Province.

Triw (Tariw, Taliw)

The population is 1,482. In Dak Chueng District, ten Triw villages are to be found. According to the map showing the administrative divisions of Xekong Province, these ten villages are located near Kaluem District where Katuic languages are spoken. The Triw speak the Katuic branch of the Mon-Khmer language, even though the lingua franca of Dak Chueng District is Triang, a Bahnaric language.

Jru' (Laven)

The population is 1,301. The Bolaven Plateau is the homeland of the Jru'. I suspect that the tribe name "Laven" may be from the place name "Bolaven". Most of the Jru' live in Pak Song District, which is a rich area of Champasak Province. In Xekong, there are only three Jru' villages, Ban Hua Xe, Ban Nong Nok and Ban Nong Lau which are located in Tha Taeng District. In comparison with the other Mon-Khmer speaking tribes in Xekong, the Jru' living in Tha Taeng District are quite well off. They own coffee plantations and orchards. In each Jru' village, there is a Buddhist temple. The Jru' feel that they are superior because their ancestors were great warriors. Ong Kommadam organized an army consisting of several Mon-Khmer tribes and fought bravely against the French. He also invented a secret writing system for communicating among themselves.

Dak Kang

The population is 2,386. This Mon-Khmer speaking tribe lives in eleven villages around the Kang River in Dak Chueng District. Outsiders do not know much about them. However, based on my language corpus and analysis, they speak a Katuic language although they live among speakers of Bahnaric languages.

Chatong

The population is 580. In Kaluem District, there are four Chatong villages, Ban Tang Kat, Ban Tang Plang, Ban Liny and Ban Sathon. Ban Tang Kat is the oldest and largest village. The Chatong can speak Kriang, which is the lingua franca of Kaluem District. In the past, they were called "Bru" meaning 'mountain people' by the Jru' or Laven. This suggests that the other tribes living in Xekong Province used to look down upon the Chatong.

Lavi (Havi, Savoeng)

The population is 492. The three Lavi villages, Ban Lavi Fang Daeng, Ban Lavi Nam Phan and Ban Lavi Noi, are located in Lamam District. In the past, this tribe used to live around the Yuet Stream in the Phu Luang Mountains. Ban Lavi Fang Daeng is about five kilometres from the municipality of Lamam District. Lavi is a Mon-Khmer language of the Bahnaric branch. There are many Alak loanwords in Lavi.

Kaseng

The population is about 300. In the past, there used to be four Kaseng villages in Dak Chueng District, Ban Dak Nong, Ban Dak Nyok, Ban Dak Troep and Ban Dak Uen Dro. At present, Ban Dak Troep is the only Kaseng village to be located in Dak Chueng. Most of the Kaseng villages, totalling about ten, are in San Sai District, Attapue Province which is adjacent to Dak Chueng District, Xekong Province. The Kaseng speak the

Bahnaric branch of the Mon-Khmer language, which is closely related to the language spoken by the Triang.

Xekong Province, southern Laos: Multilinguistic setting

Fifteen languages are spoken by the fourteen ethnic groups living in Xekong Province in southern Laos. It is interesting to note that the Lao Lum, who are the majority of the nation, are only an ethnic minority of Xekong Province. The population of Lao Lum in Xekong is smaller than some of the other Mon-Khmer ethnic groups. Lao is the official or national language of the whole nation. It is used as the medium of instruction in all schools and in the institutions of higher education, in official meetings and also in the various types of mass media. Lao script and its writing system are used for all literacy purposes. Even though Lao is the lingua franca of Xekong Province, most of the districts, except Tha Taeng District, have their own lingua franca. For example, Alak, Kriang and Triang are the lingua franca of the districts of Lamam, Kaluem and Dak Chueng, respectively.

It can be seen clearly that Xekong Province is a multilingual community. People in Xekong speak not only their own mother tongue but also can speak Lao which is the national language of Lao PDR. Moreover, they speak the languages of the other tribes living in nearby villages so that there can be good relationships and mutual understanding between them. There is no doubt that most of them are multilingual speakers. Since different languages have different functions in a multilingual society, collision is out of the question.

Cultural and linguistic diversity in Xekong Province is harmless. Cultural and linguistic contact cannot be avoided and code switching and language blending are natural. Linguistic borrowing can cause various types of linguistic variation and change at many levels, for example, the lexical, phonological, morphological and syntactical levels.

When I collected data on the thirteen Mon-Khmer languages spoken in Xekong Province for reconstructing Proto-Katuic and Proto-Bahnaric, a wordlist of 2,300 lexical items was devised. Due to language change, only 1,406 Proto-Katuic lexical forms could be reconstructed. I noticed that a large number of Bahnaric, Khmer and Lao words had been borrowed into the Katuic languages spoken in Xekong Province. In respect of Proto-Bahnaric, about 1,361 lexical forms could be reconstructed. Again, a lot of Katuic, Khmer and Lao loanwords were found in the Bahnaric languages.

The fourteen tribes of Xekong have never fought against each other but helped fight in the liberation wars against the French and Americans. We did not detect any racial or tribal problems when we conducted our field research in Xekong Province. Inter-marriages among members of different tribes are common and this encourages bilingualism within households and, later, multilingualism in villages.

More detailed information on the Katuic and Bahnaric peoples of Xekong and their languages can be found in L-Thongkum (2001). Various anthropological and sociological details of the ethnic groups or tribes of Lao PDR are presented with maps in Chazée (1999) and Goudineau (1997).

Nan Province, northern Thailand: Demography

Nan is a province in the northern region of Thailand, which shares a border with central Laos in the north and east. About 14% of the province is lowland and river basin, the rest, about 86%, are highlands and mountains. Nan Province is rich in natural resources, such as jungle, dense forests, teak woods, waterfalls, wild animals, wild orchids and so on. In the past, the northern and north-eastern areas of Nan were an

important base for the Communist Party of Thailand. Buddhist temples and sacred spiritual shrines can be found all over the province. Nan has a long history going as far back as the Kingdom of Sukhothai and it also has a few pre-historic sites.

At present, Nan province (see Map 2) is divided into 15 districts, 99 sub-districts and 902 villages (Village Database of the research project on "Linguistic Diversity in Nan Province: A Foundation for Tourism Development," 2007). There are good highways connecting all of the fifteen districts of Nan Province. The total population is 462,155 (Thailand Population and House Statistics, December 2005). The inhabitants of Nan comprise thirteen ethnic groups (see Map 3): Khon Mueang (Tai Yuan), Lue, Khuen, Lao, Phuan, Haw', Chinese (mostly the Hakka), Mpi, Hmong (Miao), Mien (Yao), Khmu', Lua' (Thin), Mlabri (Tong Lueang or the spirits of the Yellow Leaves). In addition, there is a non-identified mixed group speaking Central Thai and other languages. The following are demographic sketches of the thirteen ethnic groups living in Nan Province, a quiet and peaceful place.

Khon Mueang (Tai Yuan)

The population is 315,382. They represent the majority of the province. Khon Mueang people speak Kham Mueang or the Nan-Phrae variety of the northern Thai dialect. They have their own script and writing system called "Tua Mueang". Tua Mueang script is taught only in Buddhist monasteries for religious, historical and folk medicinal purposes. However, the population of Nan Province is literate in Thai, the official and national language of the country. Almost all of the Khon Mueang, especially the younger generation, are bilingual. The culture of Khon Mueang people is different from that of central Thai people; for example, they eat steamed sticky rice while central Thai people eat boiled plain rice for everyday meals. More information on the dialects and varieties of Kham Mueang can be found in Smalley (1994).

Lue (Tai Lue)

The population is 54,400. The Lue are religious and very good at weaving cotton cloth. The Lue style of weaving and motifs, "Pha Lai Nam Lai", are famous and are in great demand. Ban Nong Bua and Nong Bua Temple in Tha Wang Pha District are the centres for the Lue people living in Nan. About one hundred and one Lue villages are scattered over many districts. The Lue migrated from many places in the old Lue Kingdom, e.g., Mueang Chiang Rung, Mueang La and so on, which are in Sipsongpanna, China. Some of them have recently migrated from Lao PDR. They still maintain their traditional way of life. The older generation can perform Lue folk dances, sing and play Lue music. A few old men know Lue script and the writing system, so they are literate in their own language. The younger generation is trilingual, i.e. they can speak Lue, Kham Mueang and Thai.

Khuen (Khoen)

The population is 807. This small ethnic group migrated from Chiang Tung State in Burma. There are only two Khuen villages in Nan. Their language and culture are quite similar to those of the Yuan. Even though they have their own Buddhist temple, the Khuen have a lesser role to play in the multicultural society of Tha Wang Pha District, Nan Province.

Lao

The population is 9,672. There are about twenty-seven Lao villages in Nan Province. Nam Pua Sub-district in Wiang Sa District is the centre for Lao people living in Nan. This ethnic group migrated from northern areas of Lao PDR, such as Luang Pra Bang. The Lao in Nam Pua have their own Buddhist temple. They still follow their traditional way of life and speak Lao, their own mother tongue, at home. The younger generation is trilingual. As well as Lao, they can speak Kham Mueang and Thai well. Lao script and its writing system are taught in temples, so that some of them can read old documents and books written in Lao and are able to keep in touch with their relatives living in Lao PDR.

Phuan (Lao Phuan)

The population is 2,296. There are three Phuan villages in Tha Wang Pha District. They are in regular contact with the Lue and Yuan living in nearby villages. The Phuan migrated from Chiang Khwang Province, which is in the north of Lao PDR. Phuan men are very good at forging iron to make knifes, machetes, hoes and other iron utensils and tools. There are many Lao and Kham Mueang linguistic elements in the Phuan language.

Chinese (mostly the Hakka)

The population is about 2,135. Most of the Hakka Chinese live and do business of all kinds in the municipalities of many districts in Nan Province, especially in the municipality of Mueang District. They are the wealthiest ethnic group of Nan Province. The Chinese run their own school where Mandarin Chinese is taught to Chinese children and outsiders who want to learn Mandarin. They also have traditional spiritual shrines and cemeteries. A foundation has been set up to help fund cultural activities and social organizations. As in the other big cities of Thailand Nan Province's modern hotel businesses, plant and shops are owned by this ethnic group. The younger generation is literate in Thai and can speak Kham Mueang well.

Mpi (Kaw)

The population is 508. The single Mpi village is located in Yot Sub-district, Song Khwae District. The Mpi migrated from southern China to the provinces of Phrae and Nan about 150 years ago. This Tibeto-Burman ethnic group has lived among the Lue, therefore, a large number of Lue words have been borrowed into their language. Later, when the mountainous area where the Mpi live was opened to outsiders, they started to learn Kham Mueang. As a result, the younger generation can speak Kham Mueang and Thai. A lot of linguistic variations and changes can be detected in the Mpi language.

Hmong (Miao)

The population is 22,037. There are twenty-seven Hmong villages in Nan Province. There are two sub-groups of the Hmong, White Hmong and Black Hmong, and they can be differentiated by the clothing of the women. Hmong people used to live in the mountains where they practiced a slash-and-burn type of agriculture but, nowadays, they live permanently in the highlands of Nan Province. Instead of moving around and growing opium poppies, they grow and sell cabbages and lychees. They also raise chickens, pigs and cattle for sale and for sacrifice in their traditional rituals. Some of them have become Christians and can read the Bible written in Roman-based or Thai-based Hmong script. The conservative Hmong remain animists and continue to practise their traditional beliefs. Their New Year festival is very important so all of the Hmong, both

Christians and animists, return to their villages to celebrate this special occasion. Generally, Hmong people can speak Kham Mueang (Tai) as well as their own native language, which belongs to the Hmong-Mien language family. The well-educated ones can speak, read and write Thai.

Mien (Yao)

The population is 11,415. There are thirty Mien villages in Nan Province. The Mien are highlanders like the Hmong. Both ethnic groups migrated from southern China. Most of the Mien believe in Daoism and some of them have been converted into Christianity. There are two types of Mien orthography, Roman-based and Thai-based. Mien who are literate in their native language can read the Bible translated into Mien. The ones who know the Roman-based orthography are able to contact their relatives and friends who live in Laos, Vietnam, China, America and Canada via ordinary mail and e-mail. Mien women are famous for their eye-catching embroidery, clothing and silverware and Mien men are skilful silversmiths. The Mien living in Pa Klang Sub-district, Pua District own rubber plantations and lychee orchards. Many of them earn their living through business.

Khmu' (Khammu')

The population is 7,708. There are twenty-two Khmu' villages. The Khmu' of Nan Province can be divided into 3 sub-groups, Khmu' Yuan, Khmu' Lue and Khmu' Rawk. The twenty-two Khmu' villages are scattered over the districts of Song Khwae, Tha Wang Pha, Thung Chang and so on. Khmu' people speak the Khmuic branch of the Mon-Khmer language. They are either Buddhists or animists. They prefer hilly areas to lowlands because they are good at dry field cultivation. In their free time, Khmu' men weave bamboo and rattan baskets for their own use and also for sale. Some young men become soldiers in the Royal Thai Army or come to big cities to work as labourers.

Lua' (Thin)

The population is 34,600. The Lua' are the indigenous people of Nan Province. They comprise two major sub-groups, the Mal and Pray. They are called "Thin" by the Khon Mueang, which is taken as an insult because the name "Thin" has negative connotations. Actually, the word "Thin" means 'local, locality, pertaining to the region'. During the communist infiltration forty years ago, many of the Lua' living in the mountainous areas of Nan joined the Communist Party of Thailand and fought against the state's army. The Lua' language consists of two major dialects and each dialect has many varieties. The ninety-eight Lua' villages are located in many districts: Mueang Nan, Chaloem Pra Kiat, Bo Kluea, Thung Chang, Chiang Klang, Pua, Mae Chrim, Santi Suk and Song Khwae. Lua' people are Buddhists and animists. Local government and the private sector have set up various kinds of development schemes and programmes to provide the Lua' with better living conditions. This Mon-Khmer speaking ethnic group is found only in Nan and the areas adjacent to Nan in Lao PDR.

Mlabri (Tong Lueang, Spirits of the Yellow Leaves)

The population is 148. The Mlabri used to be a hunting and gathering ethnic group. Recently, they have been encouraged to stay permanently at Ban Huay Yuak (Hamlet 5), which is a large Hmong village in Wiang Sa District. Trying very hard to adjust themselves to this new way of life, they have been unhappy. Their freedom has become limited and they have had to learn to cope with modernization. They have to earn

their living by clearing the jungle for the Hmong so they can have enough money to buy food and send their children to school. In order to promote tourism in Nan, the Mlabri have been advertised by travel agencies as the most primitive of the jungle people. In order to survive in a new environment, this small ethnic group has had to struggle very hard. Their future is rather bleak in comparison with the other ethnic minorities.

Haw'

The population is 1,047. The Haw' have rather peculiar origins. They are a mixed ethnic group that may have been slaves who ran away from their masters and hid themselves in remote jungles and mountains. During the communist infiltration, they were forced to move and settle in the lowlands before the area was devastated. At present, the Haw' live with the other ethnic groups in three villages located in the district of Thung Chang. Generally, they are not friendly to outsiders. Even though many generations of the Haw' people have been born and have lived in Nan, some of them still do not have proper identity cards. This means that, legally, they are not citizens of Thailand. However, all Haw' speak Kham Mueang as their native language and most of them can understand Thai. They are Buddhists or animists.

Detailed information on the languages and cultures of the thirteen ethnic groups of Nan Province can be found in the research reports of the project on "Linguistic Diversity in Nan Province: A Foundation for Tourism Development" (L-Thongkum et al, 2007).

Nan Province, northern Thailand: Multilinguistic setting

Generally speaking, northerners are friendly, polite and generous. The thirteen ethnic groups living in Nan Province all display the above characteristics. Cultural and language differences have never caused any conflict in Nan Province. The majority (Khon Mueang) and the other twelve ethnic minorities (Lue, Khuen, Lao, Phuan, Chinese, Mpi, Hmong, Mien, Khmu', Lua', Mlabri, Haw') are able to maintain harmony and peace in all areas, both in the lowlands and highlands. Inter-marriage is common and readily accepted. Problems, if they accidentally occur, are handled in a subtle manner to avoid further conflict. When we were in Nan, it was rather hard for us to imagine the warfare between the Thai army and the communist guerrillas about 40 years before.

The habitants of Nan can speak or understand many of the language and dialects used in the province including Standard Thai ,which is the national language of Thailand. Nan Province has an interesting multilinguistic setting. The languages spoken in Nan belong to four language families: Tai (Thai, Kham Mueang, Lue, Khuen, Lao, Phuan), Sino-Tibetan (Chinese, Mpi), Hmong-Mien (Hmong, Mien) and Mon-Khmer (Khmu', Lua', Mlabri). The linguistic scenario of Nan is as follows:

Standard Thai or Thai (Tai): National or official language
Kham Mueang (Tai): Lingua franca of the province
Lue (Tai): Minority language
Khuen (Tai): Minority language
Lao (Tai): Minority language
Phuan (Tai): Minority language (enclave language)
Hakka Chinese (Sino-Tibetan): Minority language (town or trade language)
Mpi (Sino-Tibetan): Minority language (enclave language)
Hmong (Hmong-Mien): Minority language
Mien (Hmong-Mien): Minority language
Khmu' (Mon-Khmer): Minority language

Lua' or Mal (Mon-Khmer): Indigenous language
Lua' or Pray (Mon-Khmer): Indigenous language
Mlabri (Mon-Khmer): Indigenous language (enclave language)

With regard to their religions, they believe in Buddhism, Christianity, Daoism and Animism. Ceremonies and rituals are usually performed in their own native languages. There are no religious conflicts among the ethnic groups in Nan.

To illustrate a case of multilingualism in Nan Province, the language use of the Mlabri can be used as an example:

At home and among themselves, the Mlabri speak their own language, Mlabri (Mon-Khmer). When they want to find jobs, they negotiate with the Hmong by using Hmong (Hmong-Mien). Outside Huay Yuak village where they live and in the market places, they converse with the outsiders in Kham Mueang (Tai). At meetings where provincial officials are present, they speak Thai (Tai). In some situations, when the members of many ethnic groups are together, Kham Mueang (Tai) is their language choice. It can be seen from this example that each language has its own function. Comprehension and mutual understanding can be created in a natural way, appropriate to their needs. On the other hand, if the provincial government tries to set up a language policy based on its own taste, nobody knows what may happen. Resentment may be a cause of conflict, which could develop into wars.

Multilinguistic settings are common, especially in the north of Thailand, where there are multiethnic groups living together both in the highland and lowland areas. People have learned to adjust themselves to their neighbours' ways of life and have reciprocated their good will to maintain harmony. As a consequence, though cultural and linguistic changes occur it is in a subtle way, in some cases unconsciously.

Conclusion

Linguistic bias and linguistic hierarchies derived from preferences and prejudices attaching to the speakers of various languages, such as *superior* vs. *inferior*, *delicate* vs. *harsh*, *primitive* vs. *civilized*, *clear* vs. *ambiguous*, *pleasant* vs. *unpleasant*, and so forth, can be a cause of conflict, clashes and collisions (Edwards, 1994). However, it is highly fortunate that violence caused by linguistic diversity does not exist in southern Laos and northern Thailand.

Linguistic variation and change due to linguistic diversity can be subtle and accepted without conflict or negative feelings when it is induced gradually and naturally through language contact. Cultural and linguistic change cannot be avoided when the inhabitants of an area are multicultural and multilingual. Harmony and peace can be maintained as long as people are willing to suppress their selfishness and are ready to tolerate and compromise. This way, conflict can be avoided to a certain extent.

Different languages have different functions and play different roles in a multilingual community. Pluralism and multilingualism should be encouraged and taken as an important key for policy making at both provincial and national level. However, assimilation resulting from cultural and language contact should also be taken into account. If the members of various ethnic groups living in the same area prefer to conform to the majority or to be like the other ethnic groups instead of maintaining their own ethnicity, they should have the right to do so. They ought to be allowed to select whatever they believe best suits them, so that harmony, peace and unity can be maintained in the area and the nation.

Acknowledgement

I would like to express my gratitude to the Thailand Research Fund (TRF) for funding the two research projects conducted in Xekong Province, southern Laos (1997-2001) and in Nan Province, northern Thailand (2004-2007). Much thanks goes to the provincial governments of Xekong and Nan, local officials and language assistants and informants for their kind assistance in various ways. Last but not least, I would like to thank my research assistants, Chommanad Intajamornrak, Kanitha Putthasatien, Supaporn Phalipat and Yupaporn Huadsiri, for accompanying me on my fieldwork, arranging my field trips and typing the manuscript of this paper.

References

Chazée, Laurent. 1999. *ThePeoples of Laos: Rural and Ethnic Diversity*. Bangkok: White Lotus Press.

Department of Provincial Administration, Ministry of Interior. 2005. *Thailand Population and House Statistics, December 2005.* (Website:http://www.dopa.go.th/xstat/pop48_1.html)

Edwards, John. 1994. *Multilingualism*. London: Routledge.

Goudineau, Yves. 1997. *Resettlement and Social Characteristics of New Villages: Basic needs for Resettled Communities in Laos PDR, an Orstom Survey, Vol. 1-2.* Vientiane: UNDP.

L-Thongkum, Theraphan. 2001. *Languages of the Tribes in Xekong Province, Southern Laos*. Bangkok: The Thailand Research Fund (TRF). [Printed by Chulalongkorn University Printing House].

L-Thongkum, Theraphan, Wichat Booranaprasertsook, Pachee Yuvajita and Simon J.P. Wright. 2005. *From Nan to the Metropolis: the Phu Fa*. Bangkok: Faculty of Arts, Chulalongkorn University and the Thailand Research Fund (TRF). [Printed by Chulalongkorn Printing House].

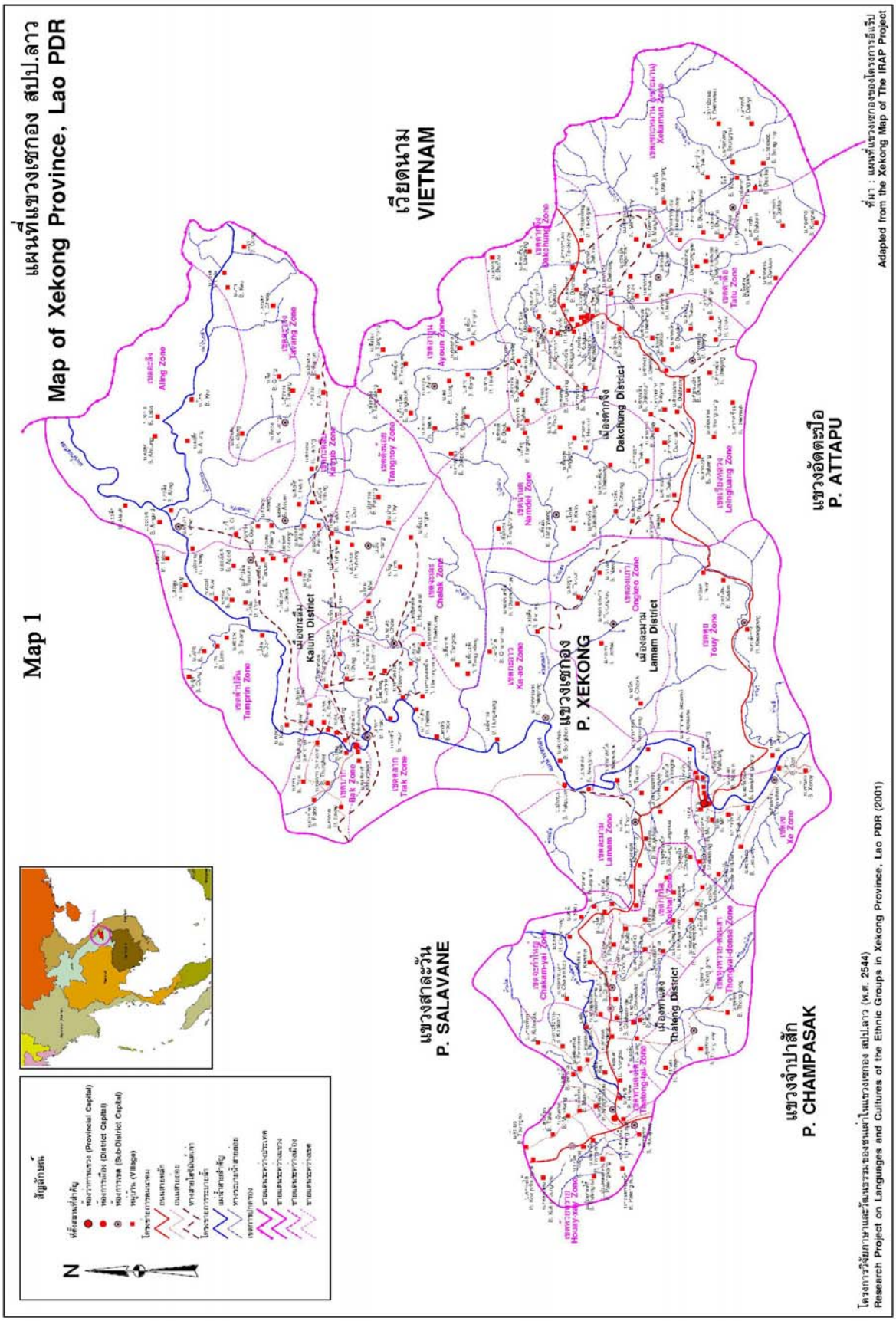
L-Thongkum, Theraphan, Phinnarat Akharawatthanakun, Chommanad Intajamornrak, Khanitha Putthasatien, Supaporn Phalipat, and Yupaporn Huadsiri. (2007). *Dictionary of the Fifteen Languages Spoken in Nan Province*. Bangkok: Faculty of Arts, Chulalongkorn University and the Thailand Research Fund (TRF).

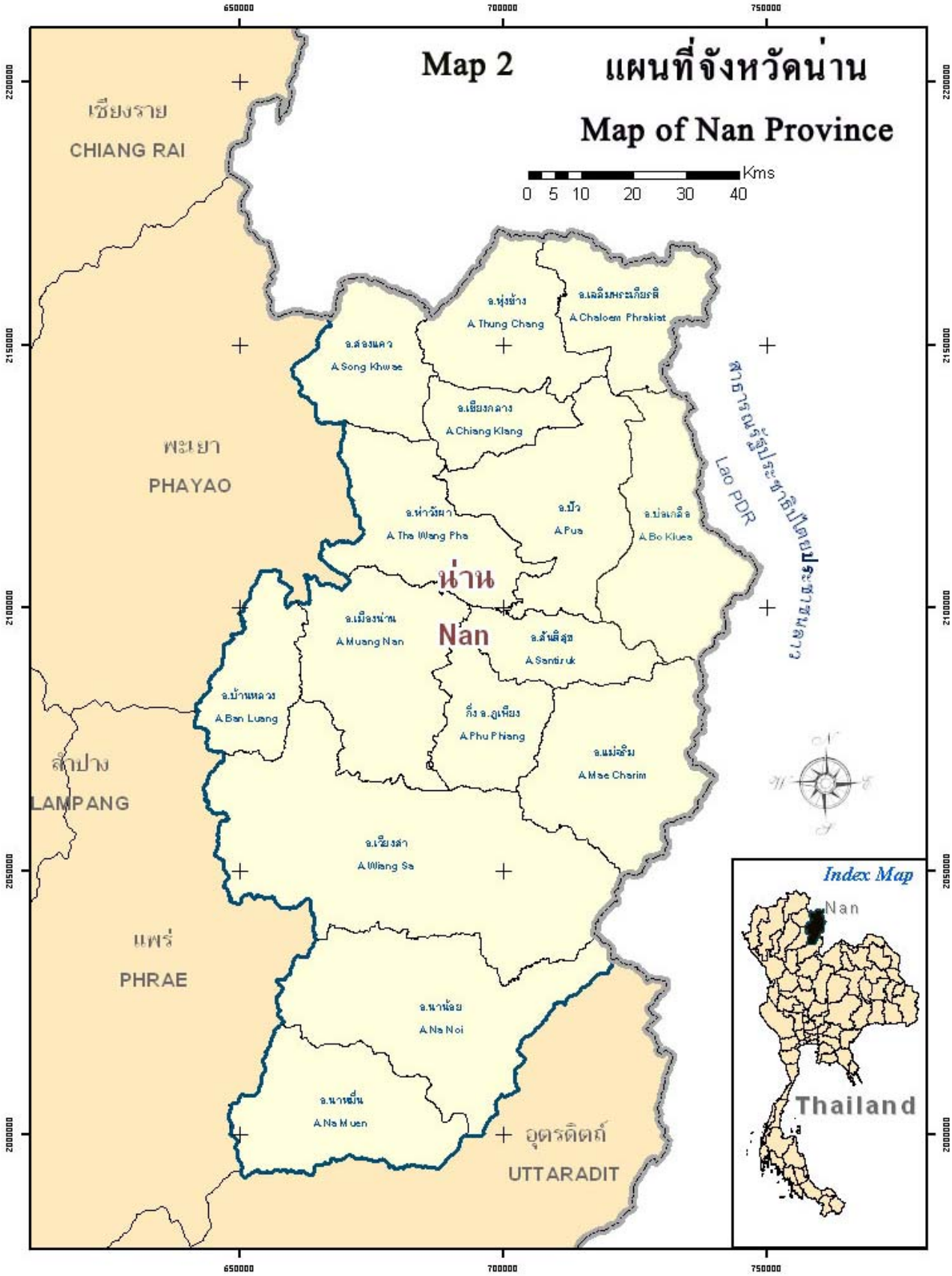
Linguistic Diversity in Nan Province Project. 2007. *Village Database*. Bangkok: Faculty of Arts, Chulalongkorn University and the Thailand Research Fund (TRF).

National Geographic Department. 1996. *Administrative Map of Lao PDR.* Vientiane.

Smalley, William A. 1994. *Linguistic Diversity and National Unity*. Chicago: The University of Chicago Press.

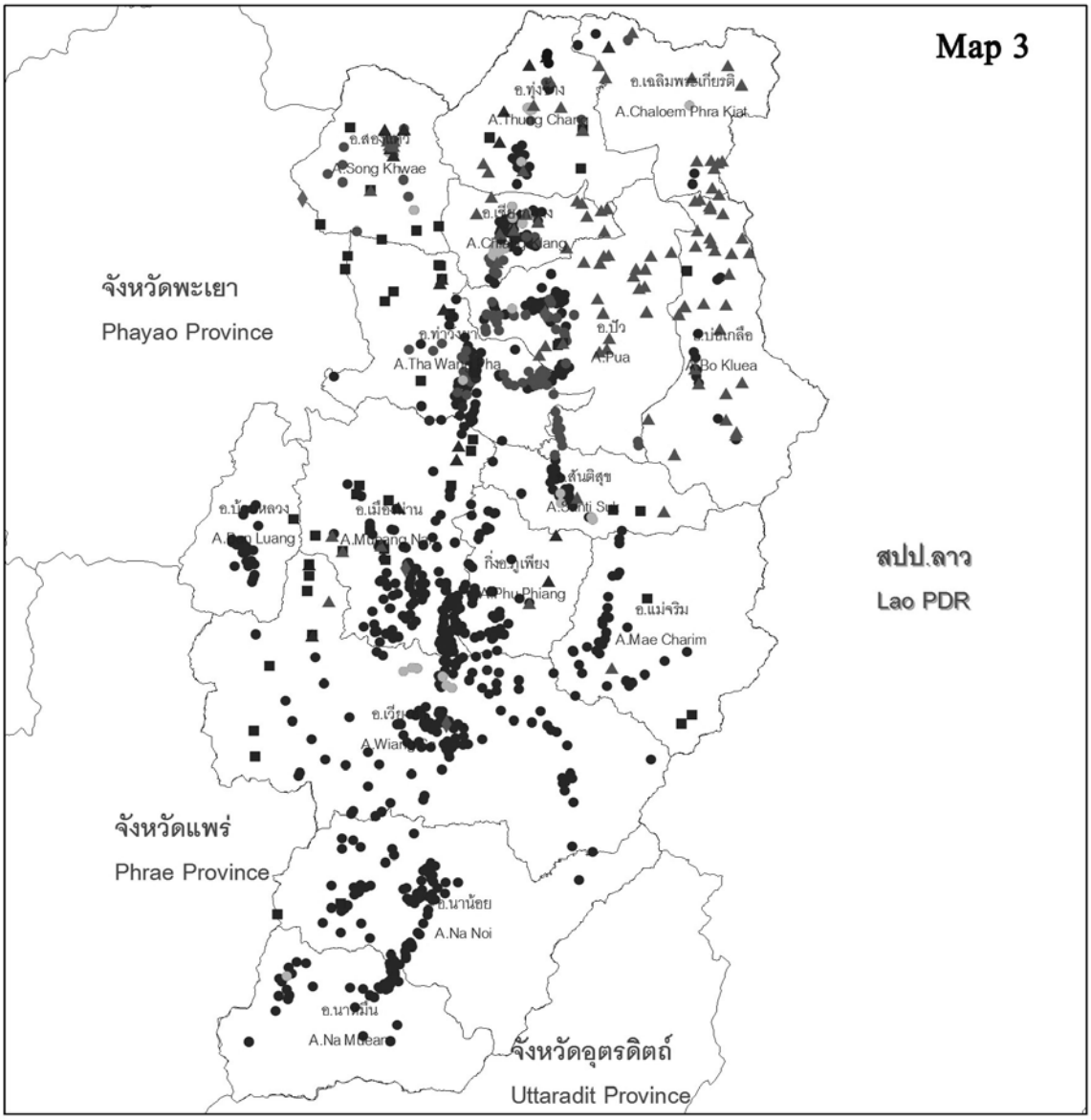
State Planning Committee, National Statistical Center. 1997. *Results from the Population Census* 1995. Vientiane.


แผนที่แขวงเซกอง สปป.ลาว
Map of Xekong Province, Lao PDR
Map 1
เวียดนาม
VIETNAM
แขวงอัตตะปือ
P. ATTAPU
แขวงเซกอง
P. XEKONG
แขวงสาละวัน
P. SALAVANE
แขวงจำปาสัก
P. CHAMPASAK
Kalum District
Dakchung District
Lamam District
Thateng District
N
ที่มา : แผนที่นำมาจากผลของโครงการนี้เอง
Adapted from the Xekong Map of The IRAP Project
โครงการวิจัยภาษาและวัฒนธรรมของชนเผ่าในแขวงเซกอง สปป.ลาว (พ.ศ. 2544)
Research Project on Languages and Cultures of the Ethnic Groups in Xekong Province, Lao PDR (2001)


Map 2
แผนที่จังหวัดน่าน
Map of Nan Province
Kms
0 5 10 20 30 40
เชียงราย
CHIANG RAI
พะเยา
PHAYAO
ลำปาง
LAMPANG
แพร่
PHRAE
อุตรดิตถ์
UTTARADIT
Lao PDR
สาธารณรัฐประชาธิปไตยประชาชนลาว
น่าน
Nan
อ.ทุ่งช้าง
A.Thung Chang
อ.เฉลิมพระเกียรติ
A.Chaloem Phrakiat
อ.สองแคว
A.Song Khwae
อ.เชียงกลาง
A.Chiang Klang
อ.ท่าวังผา
A.Tha Wang Pha
อ.ปัว
A.Pua
อ.บ่อเกลือ
A.Bo Kluea
อ.เมืองน่าน
A.Muang Nan
อ.สันติสุข
A.Santisuk
อ.บ้านหลวง
A.Ban Luang
กิ่ง อ.ภูเพียง
A.Phu Phiang
อ.แม่จริม
A.Mae Charim
อ.เวียงสา
A.Wiang Sa
อ.นาน้อย
A.Na Noi
อ.นาหมื่น
A.Na Muen
Index Map
Nan
Thailand
650000
700000
750000
2200000
2150000
2100000
2050000
2000000

แผนที่แสดงกลุ่มชาติพันธุ์ในจังหวัดน่าน

Map of Ethnic Groups in Nan Province



คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย และประชาคมน่าน
Faculty of Arts, Chulalongkorn University and the Nan Community

ได้รับทุนสนับสนุนจากสำนักงานกองทุนสนับสนุนการวิจัย (สกว.)
Funded by the Thailand Research Fund (TRF)

Tai Loanwords in Mal: A Minority Language of Thailand

Chommanad Intajamornrak
Department of Linguistics, Chulalongkorn University
Bangkok, Thailand

I. Introduction

Mal is a dialect of Lua/Thin, a Mon-Khmer language spoken in Nan Province of Thailand. It is in the Khmuic sub-branch of the Northern Mon-Khmer branch. According to Filbeck (1978), the Lua/Thin language has two main dialects, Mal and Pray. There are three varieties of Mal that agree in vocabulary and sound change. Mal A is spoken only in one village in Thung Chang District. The other two varieties are spoken in several villages in Pua District. For Pray, there are only two varieties found in Thailand. They share similar vocabulary and sound change with the Mal dialects.

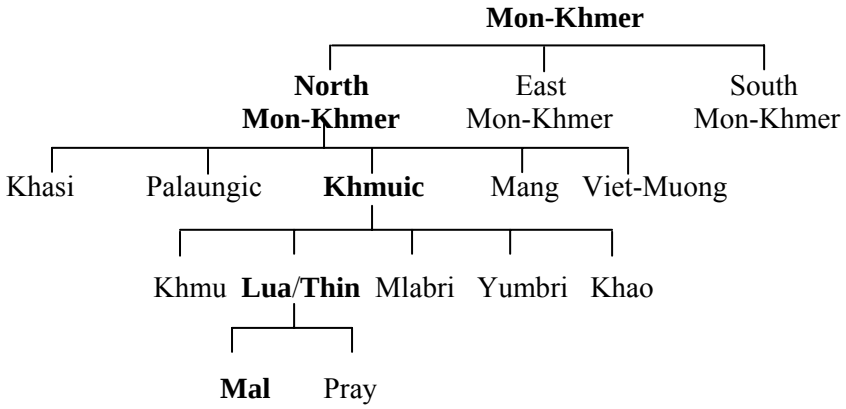


Figure 1. Lua/Thin language classification (adapted from Diffloth, 1985 and Filbeck, 1978)

Examples of cognates within Lua/Thin from Filbeck (1978) are given below:

Mal			**Pray**		
Mal A	Mal B	Mal C	Pray A	Pray B	Gloss
phram	*phjam*	*pham*	*khram*	*khjam*	'person'
pr□□ŋ	*pj□□ŋ*	*p□ŋ*	*ŋkr□□*	*ŋkj□□*	'morning'

From the examples, we can see that Mal A keeps the proto-sound */r/, but it has been changed to /j/ in Mal B and has disappeared in Mal C. It also occurs in Pray A, but it has been changed to /j/ in Pray B.

Almost all of the speakers of the Mal language are bilingual or multilingual. They can speak Tai Yuan (Kammuang Nan) fluently as well as their native language, because they usually communicate in it with Nan, Hmong, and Mien people nearby (Smalley, 1994: 230). Nowadays, some of them who are educated can also speak Standard Thai very well, especially the young generation. Linguistic research on Mal in Thailand has been done by David Filbeck. One of his works that is of interest to me is about loanwords (Filbeck, 1997).

In a situation of bilingualism or multilingualism, a natural linguistic phenomenon is language contact, which brings about phonological interference. It was found that Mal has borrowed a lot of Tai words for many years. The Tai loanwords found in Filbeck's study caused him to classify the Lua villages into three groups (Filbeck, 1997) depending on how the words are used and changed.

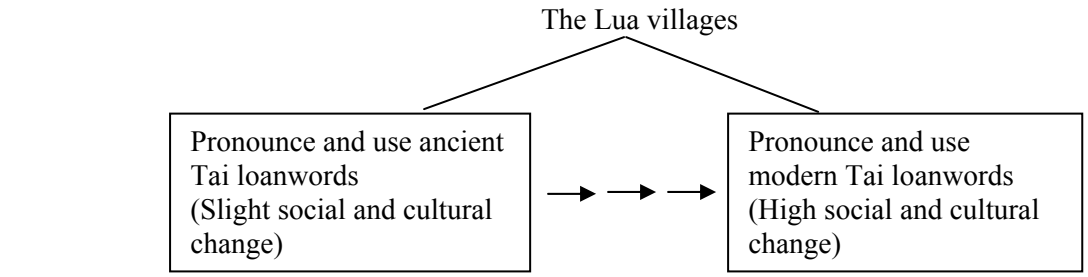


Figure 2: The classification of Pray and Mal villages by their use of Tai loanwords (adapted from Filbeck, 1997)

Examples of Tai loanwords having voiceless sonorant as an initial consonant are shown below:

Kwet Village (Mal)	**Wen Village (Pray)**	Gloss
hm□□	*m□□*	'doctor'
hlaŋ	*laŋ*	'classifier for house'

It is indeed true that some villages still keep voiceless sonorant sounds which they must have borrowed from ancient Tai words, while the others don't. There are some villages that are undergoing change. From my fieldwork at Yoddoy Wattana Village, Bo Kluea Tai Sub-District of Bo Kluea District in Nan Province,[1] I have come to realize that other aspects of the movement of Tai loanwords into Mal must be studied, not just the comparison of the history of Tai and Mal.

Due to the typological differences between the Mon-Khmer and Tai language families, transformational processes of loanword adaptation between two languages of the same family are likely differ from those involved in loanwords from another family. Loanword adaptation is a controversial linguistic phenomenon of interest to many phonologists as well as other linguists. Recent studies on loanword adaptation can be found for many languages, such as Thai (Kenstowicz and Suchato, 2006), Norwegian (Rice, 2006), Mandarin Chinese (Ching-yu Hsieh, 2005), Japanese (Dohlus, 2005), and Korean (Kenstowicz, 2005). The aim of this paper is to study Tai loanwords and to classify transformational processes of loanword adaptation in Mal.

II. Theoretical considerations and previous research

Borrowing is one of the linguistic mechanisms occurring in language-contact situations. It means the incorporation of foreign features into a group's native language by speakers of that language; the native language is maintained but is changed by the addition of the incorporated features (Thomason and Kaufman, 1988) In addition, Coetsem (1988) also considers who the beginner of the borrowing is. Borrowing means features in a source language are transferred to a recipient language by the speaker. A feature first transferred is lexicon, that is, lexical borrowing. Then, phonological borrowing, morphological borrowing and syntactic borrowing may occur when a contact situation is more intense. Thai has borrowed a lot of words from various languages such as:

Mon:	*kwiən*^{32}	'a cart'	*khrua*^{32}	'kitchen'	*ch□□n*^{34}	'spoon'	
Pali:	*cee*^{32}*dii*^{32}	'pagoda'	*karu*^{21}*naa*^{32}	'kind'	*cin*^{32}*ta.naa*^{32}*kaan*^{32}	'imagination'	
Chinese:	*kau*^{42}*□ii*^{42} 'wonton'		'chair'	*tuu*^{42}	'cabinet'	*kiew*^{34}	

[1] For the research project on "Linguistic Diversity in Nan Province: A Foundation for Tourism Development" directed by Professor Theraphan Luangthongkum.

Loanword adaptation is transformations that apply to words when they are borrowed into a foreign language. Two approaches to loanword adaptation have been in dispute as to whether the input of adaptation is phonetic or phonological (Heffernan, 2005, Kenstowicz, 2005). Moreover, perception is proposed to have a crucial role in loanword adaptation (Peperkamp & Dupoux, 2003; Miao, 2005).

Kenstowicz and Suchato (2006) study the adaptation of loanwords from English into Thai. Their interpretation of their data is that the adaptation of consonants lacking counterparts in the Thai phonemic inventory, such as the adaptation of [v] to [w], can be explained in terms of auditory similarity, while the change from interdentals to dentals is rather articulatory-based. The next section is the mapping of English voiceless and voiced stops to the Thai three-way, voiceless aspirated, voiceless unaspirated and voiced series. It seems to be that Thai speakers assign English word-initial voiceless stops to the Thai aspirated stops. This means that the adaptation is based on surface phonetics. Then the adaptation of prosodic structure and tone is investigated. The repair strategies of vowel lengthening and truncating of final clusters are applied to loanwords. Since Thai is a tonal language, tone assignment found in this study confirms what Gandour (1979) has found. There are two major rules for tonal assignment; syllables ending with a sonorant have the mid tone and syllables ending with an obstruent have the high tone.

There has not been as much research on loanwords in Mon-Khmer languages as in some other language families. Berenhult (2001) has studied loanword phonology in Jahai, a language of the Northern Aslian Subgroup of the Aslian branch of the Mon-Khmer language family, spoken in Malaysia. Jahai speakers have long been in contact with Tamiar, a Central Aslian language, and also with Malay, the Austronesian majority language. He found four main phonological changes in loanword phonology in Jahai, phonetic adaptation, phonemic replacement, reorganization of syllabic structure, and relocation of stress. The crucial point is that the several patterns are the result of indigenous innovation. He also concludes that Jahai loanword phonology is like other Aslian languages with respect to the phonological treatment of Malay loanwords.

If we ignore the question of whether the inputs are phonetic or phonemic, and focus instead on transformational processes, it can be explained that the processes can be segmental and suprasegmental change, deletion, epenthesis, and so on.

Loanwords from English to Thai are examples for segmental and suprasegmental change (Kenstowicz and Suchato, 2006).

	English	Thai
Segmental	***g**oal*	***k**oo*^{32}
	***v**isa*	***w**ii*^{32}*saa*^{42}
Suprasegmental	*camp*	*kh m*^{34}
	brand	*br n*^{32}

And loanwords from Malay to Jahai are examples of deletion and epenthesis (Burenhult, 2001).

	Malay	Jahai	
Deletion	*lembu*	*l.**m**u*	‘cattle’
	rendah	*r.**n**ah*	‘low’
Epenthesis	*ayam*	***h**ajam*	‘poultry’
	nasi	nasi	‘cooked rice’

III. Method

This paper will focus on transformational processes of Tai loanwords found in Mal. This means not only Thai, but also other Tai languages spoken in Nan Province. It is clear that Thai is the standard language of Thailand, but in this setting of language contact, Tai Yuan is the majority language of Nan Province. Therefore, lexical borrowing occurring in this situation may be from both Thai and Tai Yuan.

The data were taken from a wordlist of 2,452 lexical items devised for investigating the 13 languages of Nan[2]. This is an output of the research project on "Linguistic Diversity in Nan Province: A Foundation for Tourism Development" sponsored by the Thailand Research Fund. To separate loanwords from native words, loanwords which are easily distinguished by phonological form will be separated first. Then loanwords which are ambiguously separated will be rechecked with the other Tai and Mon-Khmer languages spoken in Nan Province. After that the percentage of loanwords found in the data will be calculated.

Loan-translations will not be treated in this paper, for example:

Tai Yuan	Mal	Gloss
kop^{35} kin^{23} d□ən^{35}	*teey p□ŋ thuə□*	'lunar eclipse'
dii^{35} caj^{23}	*l□□ □□□m*	'delighted, greatly pleased'

IV. Analysis

A. Tai Phonological systems

Although, Standard Thai and Tai Yuan are in the same language family, there are some phonological differences between these two languages. The first category is the tonal system. Standard Thai has five tones, while Tai Yuan has six tones.

		Standard Thai	Tai Yuan (Nan)
Word	Gloss	Tone	Tone
khaa	'remain embedded'	32	35
khaa	'galangal'	21	33
khaa	'value'	42	31
khaa	'to trade'	34	43
khaa	'leg'	323	23
khaa	'servant'		44□

The consonants and vowels are very similar in both languages; however, there are some systematic differences, especially in the consonants. The aspirated sounds in Standard Thai correspond to two sounds in Tai Yuan, aspirated and unaspirated. The correspondences are shown here (adapted from Smalley, 1994: 75):

Standard Thai	Tai Yuan	Gloss
phaa^{42}	*phaa$^{44'}$*	'cloth'
thaam^{323}	*thaam^{23}*	'to ask'
chom^{32}	*chom^{35}*	'to admire, to praise'
khaa^{323}	*khaa^{23}*	'leg'
phan^{32}	*pan^{35}*	'thousand'
thaaŋ^{32}	*taaŋ^{35}*	'road'
ch□□^{42}	*c□□^{31}*	'name'
kham^{32}	*kam^{35}*	'word'

Also, the phoneme /j/ in Standard Thai corresponds to two consonants, [j] and [□], in Tai Yuan; the latter does not occur in Standard Thai, for example:

Standard Thai	Tai Yuan	Gloss
juu^{21}	*juu^{33}*	'be at'
jaaw^{32}	*□aaw^{35}*	'long'

The other consonantal difference between these two languages is the correspondence between Standard Thai /r/ and consonants of Tai Yuan. To begin with, the stability of the contrast between the traditionally accepted consonants /l/ and /r/ in Standard Thai is rather dubious. In colloquial Standard

[2] The data were collected in the field by Professor Theraphan Luangthongkum. I am grateful to her for letting me use them for my research.

Thai, there is much vacillation between the two. Here, however, with this situation in mind, we can look at examples of words in Tai Yuan that are cognates of words conventionally viewed as beginning with /r/ in Standard Thai.

Standard Thai	Tai Yuan	Gloss
r□ə^{32} ~ *l□ə*^{32}	*h□ə*^{35}	'boat'
r□□j^{34} ~ *l□□j*^{34}	*l□□j*^{43}	'hundred'

Moreover, Tai Yuan doesn't have consonant clusters with /l/ and /r/ as in Standard Thai.

Standard Thai	Tai Yuan	Gloss
khr□□p^{42}*-khruə*^{32}	*kh□□p*^{31}*-khuə*^{35}	'family'
plaa^{32}	*paa*^{23}	'fish'

B. Mal Phonological system

The phonological system of Mal, the Yoddoy Wattana dialect, is not too complicated compared to other Mon-Khmer languages, especially in final consonants and vowels. There are 27 consonant phonemes of which only 11 can be final consonants. The consonant phoneme inventory is shown in Table 1.

Table 1: The consonant system of Mal

	Bilabial	Alveolar	Alveolo-Palatal	Palatal	Velar	Glottal
Stops	p ph b	t th d			k kh g	□
Nasals	m m□	n n□		□ □□	ŋ ŋ□	
Fricatives		s				h
Affricates			c □			
Laterals		l l□				
Approximants	w w□			j j□		

Final consonants: -p, -t, -k, -□, -h, -m, -n, -ŋ, -l, -w, -j

The vowel system of Mal distinguishes only three degrees of height in front, central, and back position. Besides, length has phonemic significance. Table 2 shows the vowel phoneme inventory in Mal.

Table 2: The vowel system of Mal

	Short vowels			Long vowels		
	Front	Central	Back	Front	Central	Back
High	i	□	u	ii	□□	uu
Mid	e	ə	o	ee	əə	oo
Low	□	a	□	□□	aa	□□
Diphthongs with a centering offglide		iə	uə			

Mal has two tones, low-rising tone and non low-rising tone. The low-rising tone is mainly limited to Tai loanwords. For example:

k□□□t (LW) 'to embrace'
khăam (LW) 'to cross over'

b□□□k (LW) ‘to carry on the shoulder’
păan (LW) ‘birthmark’
l□□□k (LW) ‘to exchange’

C. Transformational processes

Since Mal has been settled in this area for a long time and has been in contact with Tai Yuan, the majority language, Standard Thai, as well as other minority languages, it is not surprising that many loanwords can be found.

Out of my corpus of 2,452 words, 813 or about one-third are loanwords. Some words, especially compound nouns, are composed of Mal native words and loanwords, for example:

siəŋ (M) *khăaŋ* (LW) ‘molar teeth’
phŭn (LW) *puəh* (M) ‘ashes’
kh□□□ŋ (LW) *p□ŋ* (M) ‘condiments eaten with rice’
th□□ŋ (LW) *soo* (M) ‘copper’
□uəŋ (M) *kwa□at* (LW) ‘broom’

Some compound nouns contain a native Mal word to designate a category followed by a loanword to specify the member of the category. Some categories are /thuu/ for vegetable, /phl□□/ for fruit, /jaaŋ/ for flower, /lam/ for tree, and /phjam/ for person. For example:

thuu (M) *ka-phaw* (LW) ‘sweet basil (kind of)’
phl□□ (M) *mŭəŋ* (LW) ‘mango’
jaaŋ (M) *□iəŋ* (LW) ‘orchid’
lam (M) *təəj* (LW) ‘Pandanus amaryllifolius Roxb (kind of plant)’
phjam (M) *cĭin* (LW) ‘Chinese person’

Some Tai loanwords can be described in terms of semantic shift. It means that the meanings of the loanwords are related to those of the words of the source language but are somewhat different in Mal. For example:

Standard Thai	Gloss	Mal	Gloss
hoo^{21}	‘to cheer’	*hŏo*	‘to shout’
leew^{32}, *waj*^{32}	‘fast, quick’	*leew, waj*	‘to hurry’
phun^{21}	‘dust’	*phŭn*	‘fine particles’
liəŋ^{323}	‘yellow’	*l□iəŋ*	‘pale’
kwaj^{32}	‘to swing’	*kwăj*	‘to wave’

Words borrowed from Tai become adapted to the native phonological system of Mal. The transformational processes fall into two main categories, segmental change and suprasegmental innovation.

1. Segmental change

A. Tai phonemes with no Mal equivalents

The consonant phoneme /f/ in Tai loanwords, which is not found in Mal, is replaced by /ph/ or /phw/. For example:

Standard Thai	Tai Yuan	Mal	Gloss
fa-raŋ^{21}	*khon*^{33} *fa-laŋ*^{33}	*phjam pha-lăŋ*	‘westerner’
fun^{21}	*khii*$^{44'}$ *fun*^{33}	*phŭn*	‘dust’
fin^{21}	*fin*^{33}	*phĭn*	‘opium’
faaj^{42}	*faaj*$^{44'}$	*phwaaj*	‘cotton’
faak^{21}	*faak*^{33}	*phwăak*	‘to entrust’

As for the vowel phonemes, the diphthong with a centering offglide /□ə/ in loanwords is always replaced by /iə/. Mal has only /iə/ and /uə/, while Tai has /iə/, /□ə/ and /uə/. That is Mal chooses the one starting with a front vowel. For example:

Standard Thai	Tai Yuan	Mal	Gloss
l□ək^{42}	*l□ək*^{31}	*lĭək*	'to choose'
l□əŋ^{323}	*l□əŋ*^{23}	*l□iəŋ*	'yellow'
□□əŋ^{42}	*□□əŋ*^{43}	*jaaŋ □iəŋ*	'wild orchid'
ch□ə^{42}	*c□ə*^{31}	*ciə*	'to believe'
th□ən^{21}	*thiən*^{33}	*thĭən*	'illegal, unauthorized'

B. The same distinction in Tai and Mal

The most interesting case is that the phonological differences between the source languages, Standard Thai and Tai Yuan, make the adaptation process of loanwords more complicated. The stop consonant phonemes take into account this process. Both the source languages and the native language have three-way distinctions of stop consonants, but the phonemes /ph/, /th/, /kh/ in Standard Thai correspond to /p/ - /ph/, /t/ - /th/, and /k/ - /kh/ respectively in Tai Yuan cognate words.

The first case is that the phonemes of the source languages are the same: /p/, /t/ and /k/. In this case, it is interesting that they can be sometimes adapted to the sounds [ph], [th], and [kh] respectively. For example:

Standard Thai	Tai Yuan	Mal	Gloss
pet^{21}	*pet*^{35}	*pet*	'duck'
piŋ^{42}	*piŋ*$^{44'}$	*phiiŋ*	'to roast, to barbecue'
t□□n^{32}	*t□□n*^{23}	*t□□□n*	'to castrate'
tuəŋ^{32}	*tuəŋ*^{35}	*thŭəŋ*	'to measure'
kut^{21}	*kut*^{35}	*kut*	'cut off'
kooŋ^{32}	*kooŋ*^{33}	*khooŋ*	'humpbacked'

On the other hand, if the phonemes in the source languages are different, varied substitutions can be found, especially for the phoneme /kh/. For example:

Standard Thai	Tai Yuan	Mal	Gloss
phon^{34}	*pon*^{43}	*pŏn*	'free from (danger, misfortune)'
phan^{32}	*pan*^{35}	*phan*	'thousand'
phaaj^{32}	*phaaj*^{35}	*phaaj*	'to paddle'
khap^{34}	*kap*^{44}	*kap*	'tight'
khan^{42}	*kan*$^{44'}$	*khăn*	'to separate'
khuə^{42}	*khuə*$^{44'}$	*khuə*	'to parch, to roast'
kh□□^{323}	*kh□□*^{23}	*khj□□*	'to ask for, to beg'

If the Tai Yuan words begin with /l/, there is no adaptation; the Mal word also begins with /l/. Tai Yuan initial /h/ is replaced by Mal /l/ or /j/.

Standard Thai	Tai Yuan	Mal	Gloss
r□□j^{34}	*l□□j*^{43}	*l□□j*	'hundred'
rəəm^{42}	*ləəm*^{31}	*lə□əm*	'to start'
rak^{34}	*hak*^{44}	*jak*	'to love'
ruup^{42}	*huup*^{31}	*jŭup*	'picture, photography'
riən^{32}	*hiən*^{35}	*liən*	'to study'
r□ə^{32}	*hiə*^{35}	*liə*	'boat'

The same process occurs with the phoneme /j/ in Standard Thai which corresponds to [j] and [□] in Tai Yuan. This means that loanwords having [j] in both Standard Thai and Tai Yuan still keep the same consonant without adaptation, while loanwords having /j/ in Standard Thai but /□/ in Tai Yuan keep the same sound as in Tai Yuan.

Standard Thai	Tai Yuan	Mal	Gloss
j□t^{34}	*j□t*^{44}	*j□t*	'to hold back'

j□□m[32]	*j□□m*[35]	*j□□m*	‘to borrow’
jaak[42]	*□aak*[31]	*□ăak*	‘difficult’
j□□m[32]	*□□□m*[35]	*□□□m*	‘consent’
jok[34]	*□ok*[44]	*□□k*	‘to lift, to rise’

The next segmental change deals with the vowel phonemes. The monophthongal vowels in Mal are exactly the same as in Standard Thai and Tai Yuan, but they are also adapted inconsistently, for example, the substitution of [□] for [a], [o] for [□], [ə] for [o], and [e] for [□].

Tai Yuan	Mal	Gloss
miit[33] *□ap*[44]	*ma-□□p*	‘scissor’
naŋ[35] *w□ŋ*[43]	*jaaŋ woŋ*	‘rubber band’
□ok[35]	*□ək*	‘chest’
c□□[31]	*cĕe*	‘to soak’

However, the outstanding vowel change is vowel lengthening, especially when the words are borrowed from Tai Yuan.

Tai Yuan	Mal	Gloss
w□n[31]	*w□□□n*	‘mirror’
ma-nun[23]	*phl□□ nŭun*	‘jack fruit’
kiw[43]	*kiiw mat*	‘eyebrow’
ŋ□j[31]	*ŋ□□j*	‘lame’
n□ŋ[31]	*n□□□ŋ*	‘net for catching fish’

C. A more complex case

For the affricate phoneme /ch/, it was shown above that Standard Thai /ch/ also corresponds with two consonants in Tai Yuan, /ch/ and /c/. If the words of Standard Thai and Tai Yuan both have /ch/, it will be replaced by /s/, while the words with /ch/ in Standard Thai but /c/ in Tai Yuan have the same consonant as Tai Yuan.

Standard Thai	Tai Yuan	Mal	Gloss
chom[32]	*chom*[35]	*som*	‘to admire, to praise’
chaam[32] *□aaŋ*[21]	*chaam*[35]	*saam sək*	‘basin’
chaŋ[42]	*caŋ*[31]	*căŋ*	‘to weigh’
chuu[34]	*cuu*[43]	*cuu*	‘lover’
chaw[42]	*caw*[31]	*căw*	‘to rent’

2. Suprasegmental innovation

The low-rising tone is an innovation found mostly in Tai loanwords. It is surely the result of contact with a tonal language, especially a language with many words on a rising tone. It seems likely that the emergence of the low-rising tone in Mal is a sociolinguistic device for marking loanwords.

Out of 813 Tai loanwords, 346 (42.5%) have the low- rising tone. It was found that 654 words of Tai Yuan origin (80.4%) carry the rising tone, and 95 words of Standard Thai origin (11.7%) carry the rising tone. As a result, when Mal speakers hear Tai words, the most striking feature is the rising pitch on so many words, especially in Tai Yuan, which has two rising tones. Apparently, it causes them to pronounce Tai words with the low-rising tone regardless of what tone the original words have. For example:

Standard Thai	Tai Yuan	Mal	Gloss
ta-kh□□[323]	*kh□□*[23]	*kh□□□*	‘hook’
kl□□ŋ[32]	*k□□ŋ*[23]	*k□□□ŋ*	‘drum’
kra-den[32]	*sa-den*[35]	*sa-dĕn*	‘to splash, to fly off’
khaap[42]	*kaap*[31]	*kăap*	‘to seize (in the teeth or beak)’
taŋ[42] *t□□*[21]	*taŋ*[44’] *t□□*[33]	*tăŋ*	‘from’
phon[34]	*pon*[43]	*pŏn*	‘free from’

Most Mon-Khmer languages are classified typologically as non-tonal. The one prominent exception is Vietnamese, which is clearly tonal. Proto-Thin has been reconstructed as a non-tonal language (Filbeck, 1976). It means that Mal developed tone later on. Filbeck (1972) suggested that contact with Tai might be a cause of this linguistic phenomenon. However, he did not explicitly state why it had to have rising pitch no matter what tone that loanword had in the source language.

In trying to find the causes of this tonal evolution, Filbeck looked for internal factors and considered the possibility of a mechanism for marking loanwords; however, he feels that the failure of loanwords to be consistent in taking the rising tone in Mal makes the second factor unlikely. In my view, the reason for the lack of the low-rising tone on the other half of the Tai loanwords may be the time-depth of borrowing that caused them to be better integrated with the native phonology. A good explanation is found in the loanwords having voiceless sonorants, which don't carry the low-rising tone. For example:

Standard Thai	Tai Yuan	Mal	Gloss
ph□□^{42} *maaj*^{42}	*p□□*^{31} *maaj*$^{44'}$	*□aw m□aaj*	'widower'
m□□ŋ^{323}	-	*m□□□ŋ*	'gloomy, depressed'
nep^{21}	*nep*^{35}	*hep*	'to stick in'
lop^{21}	*lop*^{35}	*l□op*	'to avoid'
l□ə^{323}	*liə*^{23}	*l□iə*	'remaining, left over'

The adaptation of these words with Mal voiceless sonorants or /h/ strongly suggests contact with very early stages of Tai, which has been reconstructed as having voiceless sonorants (Li, 1977).

V. Discussion and Conclusion

The transformational processes of loanword adaptation found in Mal can be classified into two categories, segmental change and suprasegmental innovation. Although Standard Thai is the national language, Tai Yuan, which is the majority language of Nan Province, seems to have much more influence on Mal.

It has been accepted that loanwords are adapted by undergoing phonological analysis in which the native sounds repair the illegal non-native forms to agree with the native phonology. In some cases, more than one sound can be substituted. It means that there is more than one possibility to replace a non-native segment by a native one by single feature change (Peperkamp and Dupoux, 2003), as in the adding of aspiration to unaspirated stop consonants. Moreover, unnecessary adaptations occur during the transformation of Tai loanwords. Peperkamp (in press) points out that unnecessary adaptations might be called generalizations since they apply to foreign forms that are well-formed in the native (borrowing) language but do not conform to some default pattern. For example, Tai and Mal have the same monophthongs, nevertheless, Mal is inconsistent in its adaptation of Tai loanwords with monophthongs.

We may conclude then that loanword adaptations in Mal involve the native Mal phonology as well as the emergence of a new feature that is predominantly limited to loanwords. What we seem to have here is an incipient tone language. It remains to be seen whether a full tonal system will come into being over time or whether the present incipient system will die out.

VI. Acknowledgement

I would like to express my gratitude to Naresuan University and The Royal Golden Jubilee Ph.D. Program for sponsoring my research scholarship, and The Thailand Research Fund for sponsoring the research project on "Linguistic Diversity in Nan Province: A Foundation for Tourism Development". I also would like to express my sincerest appreciation to Professor Theraphan Luangthongkum, my advisor, and Professor Arthur S. Abramson, my co-advisor, for their kind support in my research.

References

Berenhult, N. 2001. Loanword phonology in Jahai. Lund University, Department of Linguistics, *Working Papers* 48 (2001): 5–14.

Coetsem, F. V. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.

Diffloth, Gérard. 1985. Austro-Asiatic Languages. *Encyclopaedia Britannica*. Chicago/London/ Toronto/Geneva, Encyclopaedia Britannica Inc. pp 480-484.

Dohlus, K. 2005. Phonetics or Phonology: Asymmetries in Loanword Adaptations - French and German Mid Front Rounded Vowels in Japanese. Zentrum für Allgemeine Sprachwissenschaft *Papers in Linguistics 42: 117 – 135*.

Filbeck, D. 1978. T'in: a historical study. *Pacific Linguistics, Series B*: 49. Canberra: Australian National University.

Filbeck, D. 1997. Thai loanwords in the Lua language of Muang Nan. Seminar document paper for the conference "Suepsan Adit An Rueangrong Khong Mueang Nan: Khomun Lae Mummong Mai Thang Borankhadi: Prawattisat Lae Chattiphan". Tai Studies Project, Payap University and Nan Community, December 18-21, 1997.

Heffernan, K. 2005. Phonetic similarity and phonemic contrast in loanword adaptation. *Toronto Working Papers in Linguistics* 24**:** 117–123.

Hsieh, Ching-yu S. 2005. Taiwanese Loanwords in Mandarin Chinese: Language Interaction in Taiwan. *Taiwan Papers 5*: October 2005. [available at http://www.ruhr-uni-bochum.de/slc/ taiwan/TaiwanPapers5.pdf]

Kenstowicz, M. 2005. The phonetics and phonology of Korean loanword adaptation. Paper presented at First European Conference on Korean Linguistics, Leiden University, February 2005.

Kenstowicz, M. and Suchato, A. 2006. Issues in loanword adaptation: A case study from Thai. *Lingua* 116: 7, 921-949

L-Thongkum, T. 1984. Tonal evolution and tonal development in Mon-Khmer. Warasarn Aksornsart 16: 2, pp. 53-72. (in Thai)

L-Thongkum, T. and Intajamornrak, C. 2007. Tonal evolution caused by language contact: A case study of the T'in language of Nan Province, Northern Thailand. To be presented at the Mon-Khmer Workshop, Max Planck Institute for Psycholinguistics, Nijmegen, The Natherlands.

L-Thongkum, T. and Puengpa, S. 1980. Bru-Thai-English Dictionary. Department of Linguistics. Chulalongkorn University.

Li, F-K. 1977. *A Handbook of Comparative Tai*. Honolulu: The University Press of Hawaii.

Miao, R. 2005. *Loanword Adaptation in Mandarin Chinese: Perceptual, phonological and sociolinguistic factors.* Doctoral dissertation. Department of Linguistics. State University of New York at Stony Brook.

Peperkamp, S. (in press). A psycholinguistic theory of loanword adaptations. *Proceedings of the 30th Annual Meeting of the Berkeley Linguistic Society*.

Peperkamp, S. and Dupoux E. 2003. Reinterpreting loanword adaptations: the role of perception. Barcelona: Universitat Autonoma de Barcelona. *Proceedings of the 15th International Congress of Phonetic Sciences*, 367-370.

Rice, C. 2006. Norwegian stress and quantity: The implications of loanwords. *Lingua* 116: 7, 1171-1194.

Smalley, W. A. 1994. *Linguistic Diversity and National Unity.* Chicago and London: The University of Chicago Press

Singnoi, Unchalee. 1988. A comparative study of Pray and Mal phonology. Master Degree Thesis. Mahidol University.

Thomason, S. G. and Kaufman, T. 1988. *Language contact, Creolization and Genetic Linguistics*. Los Angeles: University of California Press.

Contact-Induced Phonological Variation and Change: A Case Study of Tai Dialects Spoken in Language Mixture Areas of Nan Province[1]

Phinnarat Akharawatthanakun[2]
Department of Linguistics, Faculty of Humanities
Payap University

ABSTRACT

It is generally accepted that language change can be influenced by both internal and external factors. Language contact is an external factor which has an important role for motivating linguistic change. In Akharawatthanakun (2002, 2004), the tone variation and change which is found in Lao dialects is influenced by both an internal factor, tonal simplification, and an external factor, language contact. This paper aims to show more evidence of phonological variation and change motivated by language contact. Having analyzed the consonants, the vowels, and the tones in five Tai dialects: Tai Nyuan, Tai Lue, Tai Khün, Phuan, and Lao, spoken in the language mixture areas of Nan Province in northern Thailand, the findings reveal that: 1) the variation and change occurs in every phonological aspect: consonants, vowels, and tones and 2) the degree of phonological variation and change in some Tai dialects is higher than in the others. Furthermore, in this study lexical items in the five Tai dialects have also been investigated. Lexical variation and change is also found as evidence proving that, in the language mixture areas, the majority language plays important role in influencing the minority languages. This paper will present the variation and change in only the phonological aspect, not in the lexical aspect, so that the explanation and discussion will be clearly shown and not confusing. The variation and change in the lexical aspect will be presented in another paper of the author, ***"Contact-Induced Lexical Variation and Change: A Case Study of Tai Dialects Spoken in Language Mixture Areas of Nan Province"*** in which the research result will also the evidence of linguistic variation and change due to an external factor, i.e. language contact.

บทคัดย่อ

เป็นที่ยอมรับกันโดยทั่วไปว่า การเปลี่ยนแปลงของภาษาเกิดขึ้นอันเนื่องมาจากทั้งปัจจัยภายในและปัจจัยภายนอก การสัมผัสภาษาเป็นปัจจัยภายนอกประเภทหนึ่งที่มีบทบาทสำคัญต่อการเปลี่ยนแปลงภาษา ในงานวิจัยที่ผ่านมาของผู้เขียน (Akharawatthanakun 2002, 2004) การแปรและการเปลี่ยนแปลงของวรรณยุกต์ที่พบในภาษาลาวถิ่นต่าง ๆ ได้รับอิทธิพลจากทั้งปัจจัยภายใน ได้แก่ การออกเสียงวรรณยุกต์ให้ง่ายขึ้น (tonal simplification) และปัจจัยภายนอก ได้แก่ การสัมผัสภาษา บทความนี้มีวัตถุประสงค์ที่จะนำเสนอให้เห็นหลักฐานเกี่ยวกับการแปรและการเปลี่ยนแปลงของภาษาอันเนื่องมาจากการสัมผัสภาษา จากการวิเคราะห์พยัญชนะ สระ และวรรณยุกต์ในภาษาไทถิ่น 5 ภาษา ได้แก่ ได้แก่ ภาษาไทยวน ไทลื้อ ไทขืน พวน และลาวที่พูดในบริเวณที่มีการปะปนกันของภาษาในจังหวัดน่านทางภาคเหนือของประเทศไทย พบว่า 1) มีการแปรและการเปลี่ยนแปลงทางเสียงเกิดขึ้นทั้งในระดับของพยัญชนะ สระ และวรรณยุกต์ และ 2) ภาษาไทถิ่นบางภาษามีระดับการแปรและการเปลี่ยนแปลงทางเสียงสูงกว่าภาษาไทถิ่นอื่น นอกจากนี้งานวิจัยนี้ยังวิเคราะห์คำศัพท์ในภาษาไทถิ่นทั้ง 5 ภาษาด้วยและพบว่า การแปรและการเปลี่ยนแปลงทางศัพท์นับเป็นอีกหลักฐานหนึ่งที่แสดงให้เห็นว่า ภาษาของชนกลุ่มใหญ่มีบทบาทสำคัญต่อการแปรและการเปลี่ยนแปลงของภาษาของชนกลุ่มน้อย เพื่อให้สามารถอธิบายและอภิปรายผลการวิจัยได้อย่างชัดเจนและไม่สับสน บทความนี้จะนำเสนอเฉพาะการแปรและการเปลี่ยนแปลงทางเสียงเท่านั้น ส่วนการแปรและการเปลี่ยนแปลงในระดับคำศัพท์จะนำเสนอไว้ในอีกบทความเรื่อง ***"Contact-Induced Lexical Variation and Change: A Case Study of Tai Dialects Spoken in Language Mixture Areas of Nan Province"*** ซึ่งเป็นอีกบทความหนึ่งของผู้เขียนที่จะนำเสนอผลการวิจัยเกี่ยวกับการแปรและการเปลี่ยนแปลงของภาษาอันเนื่องมาจากปัจจัยภายนอก หรือได้แก่การสัมผัสภาษานั่นเอง

---

[1] This paper is a part of the research project ***"Linguistic Diversity in Nan Province: A Foundation for Tourism Development"*** by Faculty of Arts, Chulalongkorn University in cooperation with Nan Community, granted by the Thailand Research Fund (TRF). The project director is Professor Dr. Theraphan Luangthongkum. The other paper which is related to this paper is ***"Contact-Induced Lexical Variation and Change: A Case Study of Tai Dialects Spoken in Language Mixture Areas of Nan Province"***.

[2] The author had been in the research team of the project ***"Linguistic Diversity in Nan Province: A Foundation for Tourism Development"*** while receiving a postdoctoral fellowship from Chulalongkorn University during November 2004–June 2006. Some parts of this paper were presented in the First World Congress on the Power of Language: Theory, Practice, and Development, 22-25 May 2006 Bangkok, Thailand.

1. BACKGROUND

Nan, a province located in the east of the northern region bordering Laos in the north and Chiang Rai and Phayao in the south, can be considered an interesting example of a language mixture area since several groups of people speaking different dialects/languages are scattered all over the province. These languages are in different language families as shown below.

Tai: Nyuan/Tai Nyuan, Lue/Tai Lue, Khün/Tai Khün, Phuan, Lao
Mon-Khmer: Khmu, Luaʔ (Mal), Luaʔ (Pray), Mlabri
Hmong-Mien: Hmong/Meo, Mien/Yao
Sino-Tibetan: Hakka Chinese, Mpi

The Tai language family is the largest group and this study is focused on investigating the variation and change in the five Tai dialects: Nyuan/Tai Nyuan, Lue/Tai Lue, Khün/Tai Khün, Phuan, and Lao. Since Nan province is in the northern part of Thailand, Nyuan/Tai Nyuan which is normally called "Northern Thai dialect" or "Kham Mueang" (literally 'city language') has been spoken as the majority language. It has also been used as the lingua franca among the people speaking different languages/dialects. Thus many of minority people are bilingual or multilingual can speak not only their own mother tongue but also Nyuan/Tai Nyuan and Standard Thai which has been used as the official language in Thailand.

In order to lead to clear explanations and discussions related to the findings in this study, the general background and the classification of Southwestern Tai group will be summarized as follows.

The Classification of Southwestern Tai languages[3]

On the basis of the distinction of vocabulary and the sound development, Fang Kuei Li (1960), one of the well-known Tai comparativists, classified the Tai language family into three subgroups: Northern Tai Group, Central Tai Group, and Southwestern Tai Group.

Nyuan, Lue, Khün, Phuan, and Lao are all Southwestern Tai dialects. James R. Chamberlain (1975) classified the Southwestern Tai into two subgroups on the basis of the initial development; P group in which the proto-initial voiced stops developed to be unaspirated sounds - *b, *d, *g > p, t, k and PH group in which the proto-initial voiced stops developed to be aspirated sounds - *b, *d, *g > ph, th, kh. Nyuan, Lue, and Khün are classified into the P group while Phuan and Lao are in PH group as shown in Figure 1.

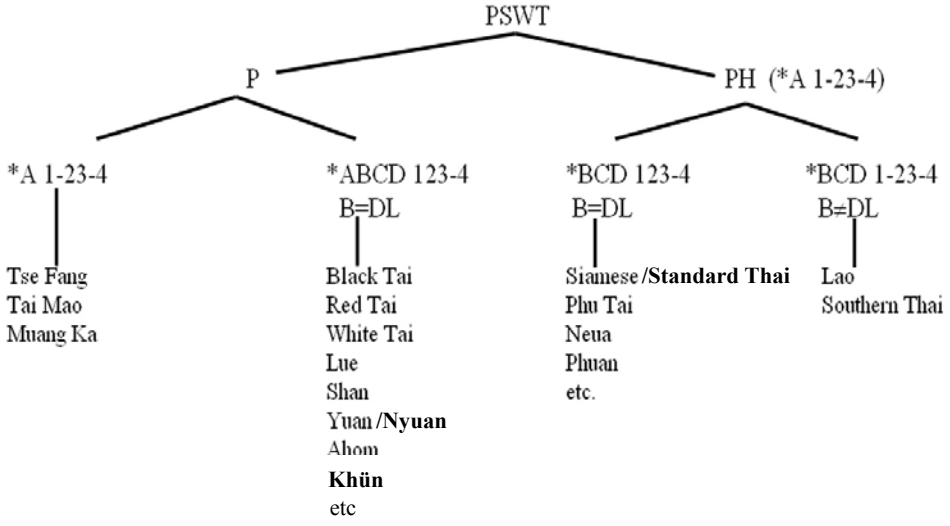


Figure 1 Classification of Southwestern Tai Dialects (Chamberlain 1975: 50) [4]

[3] The symbol ' > ' means 'to change to'.

[4] Khün has been classified the same subgroup as Yuan/Nyuan and Lue but Chamberlain (1975: 50) does not put Khün into the figure of the classification of Southwestern Tai dialects. Since Khün is one of the Tai dialects studied in this project, the author added Khün in Figure 1 in order to show the relationship between Khün and

Figure 1 shows that the Southwestern Tai dialects in P and PH groups can also be divided into different subgroups by using the pattern of tonal development as the criteria. Nyuan, Lue, and Khün which are in the P group have the pattern of proto-tones development as *ABCD 123-4 and B=DL (i.e. the tones in column *A, *B, *C and *D split into two ways: 123-4). Phuan and Lao are in the PH group but they are in different subgroups as they have different patterns of tone development, i.e. Phuan has a *BCD 123-4, B=DL pattern but Lao has a *BCD 1-23-4, B≠DL one.[5]

Considering the relationship between the Southwestern Tai dialects, Figure 1 shows that Nyuan (or "Yuan" in Chamberlain 1975), Lue, and Khün are closer to each other than to Phuan and Lao. Since Phuan and Lao are in the same PH group, they are closer to each other than to the other three Tai dialects. The relationship between the dialects is one of the factors which influence the degree of language variation and change. The differences and similarities between the dialects/languages in contact also need to be considered as stated in Weinreich (1953: 2):

> *Great or small, the differences and similarities between the languages in contact must be exhaustively stated for every domain – phonic, grammatical, and lexical – as a prerequisite to an analysis of interference. It might even be fruitful to draw up general canons of differential description. These should probably include, among others, a picture of the degree of apriori structural homogeneity of the languages in contact and the precedents for borrowing in each.*

There are both similarities and differences in every linguistic domain between the five Tai dialects investigated in this study. A summary of such similarities and differences is shown below.

- **The similarities**:
  - **Basic vocabulary:** all five Tai dialects share a number of basic vocabulary. Some examples are shown in Table 1.

Table 1 Examples of the basic vocabulary, sharing among the five Tai dialects

	Nyuan	**Lue (Nong Bua)**	**Lue (Donchai)**	**Khün (Nong Muang)**	**Phuan (Fay Mun)**	**Lao (Nam Pua)**
'field shack'	**tuup^{33}**	**tuup^{24}**	**tuup^{35}**	**tuup^{22}**	**tuup^{22}**	**tuup^{44}**
'crisp'	**phɔɔy^{33}**	**phɔɔy^{24}**	**phɔɔy^{21}**	**phɔɔy^{22}**	**phɔɔy^{22}**	**phɔɔy^{21}**
'bicycle'	**lot^{44} thiip^{33}**	**lot^{33} thiip^{24}**	**lot^{32} thiip^{21}**	**lot^{44} thiip^{22}**	**lot^{22} thiip^{22}**	**lot^{44} thiip^{44}**
'small, little (in size)'	**nɔɔy$^{44'}$**	**nɔɔy$^{22'}$**	**nɔɔy$^{52'}$**	**nɔɔy$^{41'}$**	**nɔɔy$^{41'}$**	**nɔɔy$^{35'}$**
'bright, lighted'	**cɛɛŋ$^{44'}$**	**cɛɛŋ$^{213'}$**	**cɛɛŋ$^{41'}$**	**cɛɛŋ$^{44'}$**	**cɛɛŋ$^{33'}$**	**cɛɛŋ$^{35'}$**
'highland field'	**hay^{31}**	**hay^{33}**	**hay^{32}**	**hay^{31}**	**hay^{31}**	**hay^{21}**
'suffocated'	**sa-mak^{35}**	**sa-mak^{33}**	**sa-mak^{32}**	**sa-mak^{44}**	**sa-mak^{35}**	**sa-mak^{35}**

the other Tai dialects. "Nyuan" and "Standard Thai" are also added to show the other ethnic names of "Yuan" and "Siamese" respectively.

[5] A, B, C, DL, DS represent the tones in Proto-Tai reconstructed by the Tai comparativists, e.g. Li (1960, 1977) and Gedney (1972) on the basis of the relationship between the voicing states of initial consonants and the mergers and splits of tones in Tai dialects. A, B, and C are the tones in the live syllables. DL (or D-long) and DS (or D-short) refer to the tones in dead syllables with long vowels and short vowels, respectively. The concept of Proto-Tai tones was successfully applied in the 'Tone Box', a tool for studying the tones in Tai dialects, proposed by Gedney (1972). (See more detail in 5.3 and Appendix 1)

- **Consonants:** the consonantal phonemes which are found in all of the five Tai dialects are shown in Table 2.

Table 2 The consonantal phonemes found in all five Tai dialects

	Bilabial	Labio-dental	Alveolar	Palatal	Velar	Glottal
Stop	p		t	c	k	ʔ
	ph		th		kh	
	b		d			
Nasal	m		n		ŋ	
Fricative		f	s			h
Lateral			l			
Approximant		w		y		

- **Vowels:** both long and short vowels, shown in Table 3 are found in all five Tai dialects.

Table 3 The vowel phonemes found in all five Tai dialects

	Front	Central	Back
High	i, ii	ɨ, ɨɨ	u, uu
Mid	e, ee	ə, əə	o, oo
Low	ɛ, ɛɛ	a, aa	ɔ, ɔɔ

- **The differences:**
  - **Basic vocabulary:** some Tai dialects have different vocabulary from the others. The examples are shown in Table 4.

Table 4 Examples of the different basic vocabulary in some Tai dialects

	Nyuan	**Lue (Nong Bua)**	**Lue (Donchai)**	**Khün (Nong Muang)**	**Phuan (Fay Mun)**	**Lao (Nam Pua)**
'bangle, anklet'	**wɔŋ^{43}**	**kham^{24} may^{31}**	**kiŋ^{21} may$^{52'}$, wɔŋ$^{52'}$ khɛɛn^{35}, kam^{35} lay^{44}**	**kiŋ^{22} may^{35}**	**pɔɔʔ^{22} hɛɛn^{23}**	**kɔɔŋ^{35} khɛɛn^{434}**
'to flatter'	**tuəʔ^{35}**	**ʔoot^{24}**	**toʔ^{35}**	**ʔoot^{22}**	**ʔuət^{22}**	**ɲɔɔŋ$^{35'}$**
'to smile'	**ɲim^{43}**	**yim$^{22'}$**	**yim$^{52'}$**	**yim$^{41'}$**	**hoo^{23}**	**huə^{434}**
'between'	**cɔŋ^{31} kaaŋ^{23}**	**waaŋ^{24} kaaŋ^{35}**	**hɔm^{21} kaaŋ^{35}**	**waaŋ^{22} kaaŋ^{23}**	**thəŋ^{31} kaaŋ^{35}**	**waaŋ^{21}**
'to belch'	**ʔəʔ^{35}**	**ʔəʔ^{33}**	**ʔəəʔ$^{41'}$**	**ʔəʔ^{35}**	**ʔəʔ^{35}**	**ʔiəm$^{35'}$**
'to run'	**lon^{31}**	**lɛɛn^{33}**	**lɛɛn^{32}**	**loon^{31}**	**lɛɛn^{31}**	**pop^{35}, lɛɛn^{21}**
'resemble, look alike'	**mɨən^{23}**	**məən^{35}**	**məən^{35}**	**məən^{23}**	**khɨɨ^{35}**	**khɨɨ^{23}**

- **Consonants:** there are consonantal phonemes found in some Tai dialects but not in the others. The different consonantal phonemes between the Tai dialects are shown in Table 5. Note that the Tai dialects in which each consonantal phoneme

is found are shown in the parentheses after each phoneme. In the case of /ph/ and /d/ in Khün, they occur in free variation with [f] and [l], respectively.

Table 5 The different consonantal phonemes in some Tai dialects

	Bilabial	**Labio-dental**	**Palatal**
Stop	ph (~ f) (Khün)		ch (Nyuan, Lue, Khün, Phuan)
		d (~ l) (Khün)	
Nasal			ɲ (Nyuan, Phuan, Lao)

- **Vowels:** a number of vowels, especially diphthongs and cases of vowel shifts are found in some Tai dialects but not in the others as shown in Table 6.

Table 6 The different vowel phonemes in some Tai dialects[6]

	Lao	**Phuan**	**Lue (Nong Bua, Donchai), Khün**
Diphthongs:	iə, ɨə, uə	iə	
Vowel shift:	*aə > aə	*uə > oo, *ɨə > iə	*o > u, *e > i, *uə > oo, *ɨə > əə, *iə > ee

Having investigated the linguistic characteristics in each domain of the five Tai dialects in contact, it is noticeable that the variation and change mostly occur with the items which are different between the Tai dialects. In other words, in the case that linguistic characteristics, especially the consonants and vowels, are found to be shared among all of the Tai dialects, they tend to be maintained. This paper aims to show each case of the variation and change in the phonological domain, i.e. consonants, vowels, and tones, in the Tai dialects in contact. The percentage of the variation and the donor dialects/languages which play important role in motivating such variation and change will be discussed.

2. STUDY LOCATIONS

There are six study locations – one village for each Tai dialect as follows.

1. **Nyuan:** Mueang District[7]
2. **Lue (Nong Bua):** Nong Bua Village, Tha Wang Pha District
3. **Lue (Donchai):** Donchai-Rai Oy Village, Pua District
4. **Khün:** Nong Muang Village, Tha Wang Pha District
5. **Phuan:** Fay Mun Village, Tha Wang Pha District
6. **Lao**: Nam Pua Village, Wiang Sa District

The six study locations scattered in four districts of Nan Province as shown in Figure 2.

[6] The vowels marked with the asterisk (*) represent the vowel in Proto-Tai. The symbol " > " is used as 'to be changed to, to be developed to', for example, *aə > aə means that the proto-vowel *aə is maintained while *uə > oo means that the proto-vowel *uə changed to [oo], and so on. (See details in 5.2)

[7] The data of vocabulary in Nyuan (Mueang Nan) was collected by Professor Dr. Theraphan Luangthongkum, the research project director, and the tone data was collected by the author.

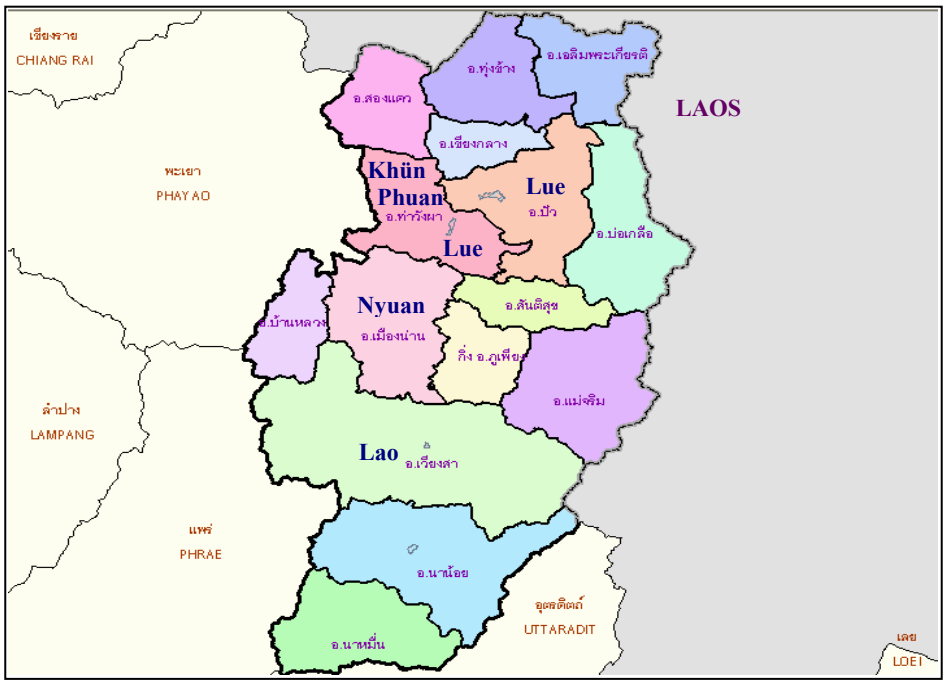


Figure 2 The study locations of Tai dialects in Nan Province

3. LINGUISTIC DATA

The data in each Tai dialect consist of two main sets; the first one is the list for checking the vocabulary and the second one is the tone checklists. In the first set, there are a total of 14,712 lexical tokens (i.e. 2,452 lexical items in each studied Tai dialects: Nyuan, two varieties of Tai Lue, Khün, Phuan, and Lao) which were collected for the main purpose of producing a dictionary of fifteen languages in the research project.[8] The lexical items for checking the variation and change in the studied Tai dialects are selected from this list.

As for the tone data, they were collected by using tone checklists which consisted of two sets of wordlists:[9]

1. 80 test words – adapted from William J. Gedney's word list (Gedney 1972). This set of test words is suitable for auditory analysis.
2. 100 test words from an analogous set – (20 words X 5 repetitions = 100 test words). This set of test words is suitable for instrumental analysis.

In order to get the most accurate data, both lexical and tonal checklists were elicited by interviewing at least 3-5 speakers of each Tai dialect; one or two older speakers who could speak their own mother-tongue fluently were selected as the main language resource persons (LRPs). Then the data was rechecked with at least three speakers of each Tai dialect in order to gain more accuracy. Both lexical and tonal data were recorded with a Sony ICD-MS515 IC Recorder and an Aiwa Stereo Radio Cassette Recorder HS-JM400. For the purpose of storing the recorded sound into separate electronic files for analysis by a computerized program, the recorded data were transfered into Cool Edit Pro (Version 1.0), an Audio Editing Software.[10]

---

[8] One of the objectives of the research project ***"Linguistic Diversity in Nan Province: A Foundation for Tourism Development"*** is to produce a dictionary of fifteen languages – Thai, English, Nyuan, Lue, Khün, Phuan, Lao, Mien (Yao), Hmong (Meo), Hakka Chinese, Lua? (Mal), Lua? (Pray), Khmu, Mlabri, and Mpi.

[9] See the two sets of test words in Appendix 1.

[10] Cool Edit Pro (Version 1.0) was designed by David Johnston, Syntrillium Software Coporation, USA.

4. DATA ANALYSIS

The phonetic and phonological variation/change of consonants and vowels was determined by comparing the consonants and vowels in this study to the consonants and vowels in Proto-Tai (Li, 1977) and in the Tai dialects presented in previous studies. For the tone analysis, the two tone checklists were separately analyzed. The total 480 test words adapted from Gedney's word list (80 test words X 2 varieties of Tai Lue and the other 4 Tai dialects) were firstly analyzed so as to gain the patterns of tonal merger and split. The other total 600 test words from the analogous set (100 test words X 2 varieties of Tai Lue and the other 4 Tai dialects) were also auditorily analyzed to confirm the exact patterns of tonal merger and split. All of the tokens from both two tone checklists were finally analyzed by Praat (Version 4.4.19)[11] in order to get the results of tonal characteristics and to confirm the results of auditory analysis.

5. RESULTS

Having analyzed and organized the 2,452 lexical items in each Tai dialect for the dictionary, the author noticed that variation and change occur in some Tai dialects in some degree not only at the lexical level but also at the phonetic and phonological level. Furthermore, such variation and change occur in several linguistic aspects: consonants, vowels, and tones as shown in the following sections.

5.1 The variation of initial consonants in Tai dialects

As can be seen in Figure 1, the Southwestern Tai languages are classified into two subgroups on the basis of the development of initial consonants and the tones: P and PH groups. Three Tai dialects in this study; Nyuan, Lue, and Khün, are in the P group while the other two ones: Phuan and Lao, are in the PH group. It can be said that the Tai dialects in each subgroup have a distinction between some sets of initial consonants, especially the sets of plosive consonants, i.e. [p, t, k, ph, th, kh] as shown in Table 7.

Because of the fact that Standard Thai (classified in the PH group) has been used as the official language and taught in schools in Thailand, Tai speakers can speak not only their own mother-tongue but also Standard Thai. Even though some of them cannot speak Standard Thai as fluently as native speakers, they can understand it perfectly. Thus, Standard Thai is considered in this study to be one main donor language, inducing variation and change in the studied Tai dialects. The initial consonants in Standard Thai are then also included in Table 7.[12]

Table 7 The distinction of some initials in Tai dialects[13]

Proto-Tai (Li, 1977)	Standard Thai	Nyuan (Mueang Nan)	Lue (Nong Bua, Donchai)[14]	Khün (Nong Muang)	Phuan (Fay Mun)	Lao (Nam Pua)
***b-**	**ph-** [phɔɔ^{42}] *'father'*	**p-** [pɔɔ^{31}]	**p-** [pɔɔ^{33}]	**p-** [pɔɔ^{31}]	**ph-** [phɔɔ^{31}]	**ph-** [phɔɔ^{21}]
***d-**	**th-** [thaa^{33}] *'to smear'*	**t-** [taa^{35}]	**t-** [taa$^{31/44}$]	**t-** [taa^{35}]	**th-** [thaa^{35}]	**th-** [thaa^{23}]

[11] Praat (Version 4.4.19) is a computerized program, designed by Pauls Boersma and David Weenink, Institute of Phonetic Sciences, University of Amsterdam, The Netherlands (1992-2006).

[12] The village or the district names of the studied locations are put into the parentheses following the ethnic names.

[13] The symbol ' → ' means 'to vary to'.

[14] In the case that some tones in the two varieties of Lue have different characteristics, the tones in Lue (Nong Bua) will be put preceding the tones in Lue (Donchai), for example, in [taa$^{31/44}$] 'to smear', in Lue (Nong Bua) the word has mid-falling tone [31] but in Lue (Donchai) it has high level tone [44].

Table 7 (continued)

Proto-Tai (Li, 1977)	Standard Thai	Nyuan (Mueang Nan)	Lue (Nong Bua, Donchai)	Khün (Nong Muang)	Phuan (Fay Mun)	Lao (Nam Pua)
*g-	**kh-** [khon^{33}] *'to stir'*	**k-** [kon^{35}]	**k-** [kun$^{31/44}$] **Note:** Lue (Nong Bua) uses [khun^{31}] which means that there is the variation of initial consonant: k → kh.	**k-** [kon^{35}]	**kh-** [khon^{35}]	**kh-** [kon^{23}] **Note:** This word is expected to be pronounced [khon^{23}] in Lao (Nam Pua) but it is [kon^{23}], i.e. there is variation of initial consonant: kh → k.
*ǰ-	**ch-** [chaaŋ^{45}] *'elephant'*	**c-** [chaaŋ^{43}]	**c-** [chaaŋ$^{22'/52}$]	**c-** [chaaŋ$^{41'}$]	**s-** [saaŋ$^{41'}$]	**ch-** [caaŋ$^{35'}$] **Note:** This word is expected to be pronounced [chaaŋ$^{35'}$] in Lao (Nam Pua) but it is [caaŋ$^{35'}$], i.e. there is variation of initial consonant: ch → c.
*x-[15]	**kh-** [khɨn^{42}] *'to rise'*	**kh-** [khɨn$^{44'}$]	**kh-** [khɨn$^{213'/31}$]	**kh-** [khɨn$^{44'}$]	**h-** [hɨn$^{33'}$]	**kh-** [khɨn$^{44'}$]

On the basis of using Li's reconstruction of Proto-Tai (Li, 1977) as the starting point, the development of voiced plosive initial consonants in Tai dialects in P and PH groups is different as mentioned before, that is *b, *d, *g, *ǰ > p, t, k, c in P group and *b, *d, *g, *ǰ > ph, th, kh, ch in PH group. Some of the studied Tai dialects, for example, Lue (Nong Bua, Donchai), has the variation (k > kh) while Lao (Nam Pua) has (kh > k) and (ch > c). These kinds of variation will be discussed in section 5.1.1.

Phuan is a Tai dialect in PH group which has a different development of proto-initial *ǰ-, considered not to be the variation, i.e. the development of (*ǰ > s) instead of (*ǰ > ch). This is one of the common patterns of development in Phuan and also in most of the Lao dialects studied in the author's previous research (Akharawatthanakun, 2002, 2004).

The other initial consonant which has an interesting development is *x. The Tai dialects in both P and PH groups have the development pattern as (*x > kh) but Phuan has (*x > h) which is also commonly found in some other Tai languages, for example, Phu Thai which is a Southwestern Tai language in PH group and Po-ai, a Northern Tai language. The development of each proto-initial consonant discussed here is shown in Table 8.

[15] It is stated in Li (1989) that in *A Handbook of Comparative Tai*, Li (1977) reconstructed two Proto-Tai consonants *kh- and *x-. These two consonants merge into kh- in most of the Southwestern and Central Tai languages, such as Siamese and Lungchow, but they are distinguished in the Northern group of Tai languages, such as Po-ai, where *kh- becomes k- (unaspirated) and *x- becomes h-. A few examples may be given as follows: *kh-: 'to kill' Siamese [khaa C1], Lungchow [khaa C1], Po-ai [kaa C1]; *x-: 'to enter' Siamese [khau C1], Lungchow [khau C1], Po-ai [hau C1]. There is some evidence from the Southwestern group of languages which makes Li (1989) reconsider the whole problem of reconstruction, particularly the problem of aspiration in Proto-Tai. In White Tai and some of the Lue dialects (Li spells "Lü"), and also, Li believes, in Old Siamese, i.e. of the Sukhothai period or about that time, there is a split of both *kh- and *x-. For example, A. *kh- 'to kill' Siamese [khaa C1], White Tai [xa C1], Po-ai [kaa C1]; *x- 'white' Siamese [khaau A1], White Tai [xau A1], Po-ai [haau A1]. B. *kh- 'ditch, pit' Siamese [khum A1], White Tai [khum A1], Po-ai [kum A2]; *x- 'bitter' Siamese [khom A1], White Tai [khum A1], Po-ai [ham A2]. Note that the symbol A1, A2, and C1 refer to the Proto-Tai tones which are also reconstructed in Li (1977).

Table 8 The development of some initial consonants in the studied Tai dialects

Proto-Tai (Li, 1977)	**Standard Thai**	**Nyuan (Mueang Nan)**	**Lue (Nong Bua, Donchai)**	**Khün (Nong Muang)**	**Phuan (Fay Mun)**	**Lao (Nam Pua)**
*b-	ph-	p-	p-	p-	ph-	ph-
*d-	th-	t-	t-	t-	th-	th-
*g-	kh-	k-	k-	k-	kh-	kh-
			Note: There is the variation, k → kh.			**Note:** There is the variation, kh → k.
*ǰ-	ch-	c-	c-	c-	s-	ch-
						Note: There is the variation, ch → c.
*x-	kh-	kh-	kh-	kh-	h-	kh-

The variation and change of the initial consonants in some Tai dialects investigated in this study will be discussed in Sections 5.1.1-5.1.2.

5.1.1 The variation of /ph-, th-, kh-, s-/ in Phuan

Phuan is spoken in Fay Mun village which is not far from Nong Bua, a village where most of the population is Tai Lue. Phuan speakers can speak not only Phuan, their own mother-tongue, but also Tai Lue and Nyuan/Northern Thai dialect which is the lingua franca.

Having compared the initial consonants /ph-, th-, kh-, s-/ in Phuan to the Proto-Tai initial consonants and to the available previous studies, it is found that there is variation of some initial consonants to different degrees. The usage percentage of initial variants in Phuan is shown in Figure 3.[16]

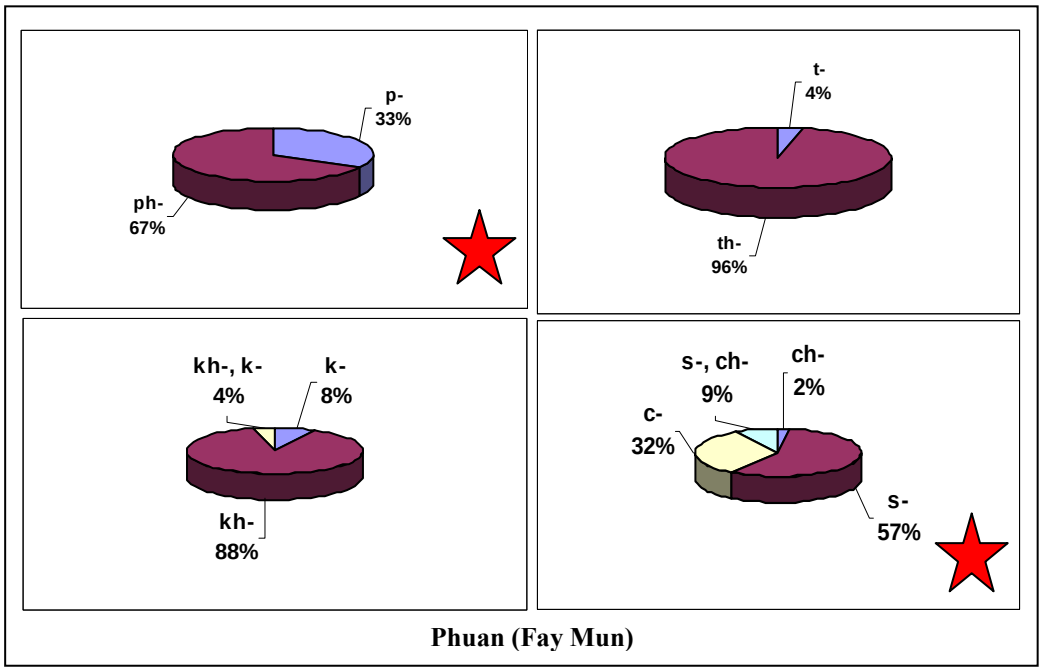


Figure 3 The usage percentage of the variants of stop initials in Phuan (Fay Mun)

[16] The percentage of variant usage is indicated after each variant. The red star symbol shown in the figure refers to the variation of the initial consonants which occur in higher percentage than the others.

Examples of the lexical items, pronounced with different variants are shown in Table 9.

Table 9 Examples of lexical items, pronounced with different variants of /ph-, kh-, th-, s-/ in Phuan (Fay Mun)[17]

Initial consonants /ph-, th-, kh-, s-/ in Phuan (Fay Mun)			
no variation		**with variation**	
Gloss	**/ph-/, /th-/, /kh-/, /s-/**	**Gloss**	**ph- → p- = Nyuan/Lue th- → t- = Nyuan/Lue kh- → k- = Nyuan/Lue s- → c-, ch- = Nyuan/Lue, Standard Thai**
'chili pepper'	***ph***ik^{22}	*'elder sibling'*	***p***ii^{31} (← ***ph***ii^{31})
'father'	***ph***ɔɔ^{31}	*'harp'*	***p***iin^{35} (← ***ph***iin^{35})
'to smear'	***th***aa^{35}	*'noon'*	***t***iəŋ^{31} (← ***th***iəŋ^{31})
'person'	***kh***on^{35}	*'to bend over'*	***k***ooŋ$^{41'}$ (← ***kh***ooŋ$^{41'}$)
'mortar'	***kh***ok^{22}	*'lamp, lantern'*	***k***oom^{35} (← ***kh***oom^{35})
'to hate'	***s***aŋ^{35}	*'slow, late, delayed'*	***c***aa$^{41'}$ (← ***ch***aa$^{41'}$)
'to help'	***s***ɔɔy^{31}	*'name'*	***s***ɨɨ^{31}, ***ch***ɨɨ^{31} (← ***s***ɨɨ^{31})
'elephant'	***s***aaŋ$^{41'}$	*'to taste'*	***ch***im^{35} (← ***s***im^{35})

The variation of the initial consonants in Phuan as shown in Figure 3 can be hierarchically arranged from high to low degree as /ph-, s-, kh-, th-/. Each initial consonant has different variants as follows:

1) /ph/ has two variants: [ph-] and [p-]. Though [ph-] is still used most frequently (67%), the variant [p] is used to a low degree (33%).
2) /s-/ has three variants: [s], [c], and [ch]. There are also some lexical items pronounced with both initial [s-] and the variant [ch-], which means that the speakers used these two variants even when they were pronouncing the same lexical item in different time. It is noticeable that the variant which is used in quite high degree is [c-] (32%), but not as high as [s-].
3) /kh-/ has two variants: [kh-] and [k-] but the former variant [kh-] is still used in high degree (88%). Thus, it can be said that the variation of /kh-/ in Phuan does not occur to a high degree.
4) /th-/ has two variants: [th-] and [t-]. Similarly to the variation of /kh-/, the former variant [th-] is still used in high degree (96%) which means that it faces very little degree of variation.

[17] The symbol ' = ' is used to show that the variant of each consonant is found in the language/dialect at the end of ' = ' and it is hypothesized that such variation is influenced by such language/dialect. In the case that there is variation in Phuan (Fay Mun), the expected pronunciation will be put in the parenthesis after the variant form, for example, 'elder sibling' ***p***ii^{31} (← ***ph***ii^{31}) means that in Phuan (Fay Mun), the lexical items for 'elder sibling' is expected to be pronounced ***ph***ii^{31} but it is pronounced ***p***ii^{31}.

As can be seen in Table 9, the variation of each initial consonant is suspected to have been induced by an external factor, language contact. Nyuan, Lue, and Standard Thai are hypothesized to be the donor dialects/languages which influence the variation of initial consonants in Phuan. It can also be stated that the phonic interference[18] is the main mechanism, leading to such variation. Phuan speakers can speak Northern Thai rather fluently and they use Northern Thai as the lingua franca when they communicate the people speaking a different mother-tongue (for example, Lue and Khün), but not many of them can speak Lue or Standard Thai. Thus it can be hypothesized that Lue might have played more important role than the other two languages for inducing the variation of the stop initials in Phuan, especially ph- > p- and s- > c-, ch- as shown in Figure 3 and Table 9.

5.1.2 The variation of /h-/ in Phuan

One of the initial consonants in Phuan which is different from the other Tai dialects spoken in Nan Province in the voiceless glottal fricative /h-/ derived from the proto-consonant *x, reconstructed in Li (1977).

Li (1977: 207) stated that the proto-voiceless fricative *x- has merged with Proto-Tai *kh- in most Southwestern and Central Tai dialects, but is represented by /h-/ in most Northern Tai dialects, occasionally /ɣ-/ or /v-/ in some Pu-i dialects. The Northern Tai dialects then are crucial in determining whether the Proto-Tai consonant is *kh- or *x-. Among the Southwestern Tai dialects, it is represented in Lue (or 'Lü' in Li, 1977) and White Tai by either /x-/ or /kh-/, just as Proto-Tai *kh-. The /x-/ form occurs more often in words where there is no shift of tone from series 1 to series 2, while the kh- form occurs in words where series 1 of the tones is shifted to series 2 in the Northern Tai dialects. As series 1 indicates voiceless initials, and series 2 voiced initials, there is correlation between the two representations in Lue and White Tai and the alternation of voiced and voiceless consonants in the Northern Tai dialects. If so, the alternation of voiceless and voiced consonants must not be considered as merely an aberrant feature of the Northern Tai dialects, as it has certain correlations with the Southwestern Tai dialects.

The Sukhothai inscriptions have two letters, the normal letter (ข) which is corresponding to Indic kh- to represent kh- and modified letter (ฃ) to represent x-, agreeing on the whole with Lue and White Tai. Some examples of the development of Proto-Tai *x- stated in Li (1977: 208) are shown in Table 10.

Table 10 Examples of the development of Proto-Tai *x- in some Tai dialects (adapted from Li, 1977: 208)[19]

Gloss	Tone	Southwestern Tai			Central Tai	Northern Tai
		Siamese	Lao	Lue/Lü	Lungchow	Po-ai
'to crow as a cock'	A1	khan	-	xăn	khan	han
'bitter'	A1	khom	-	khum	khum	ham (A2)
'ginger'	A1	khiŋ	khiŋ, hiŋ	xiŋ	khiŋ	hiŋ

[18] The term 'interference' was introduced by Weinreich, Uriel (1953: 1) as a neutral term: "those instances of deviation from the norms of either language which occur in the speech of bilinguals as a result of their familiarity with more than one language, i.e. as a result of language contact".

[19] The columns of Lao and Lue/Lü are added in Table 10. The dash (-) in Lao column means that the words in Lao are not shown in Li (1977).

Table 10 (continued)

Gloss	Tone	Southwestern Tai			Central Tai	Northern Tai
		Siamese	Lao	Lue/Lü	Lungchow	Po-ai
'white'	A1	khaau	-	xau	khaau	haau
'to cross, step across'	B1	khaam	khaam, haam	xam	khaam	haam
'to ascend, rise'	C1	khf̌n	-	xf̌n	khf̌n	hf̌n
'to enter'	C1	khau	-	xă̱u	khau	hau
'slave'	C1	khɔɔi	khɔɔi	xɔi	khooi	hooi (B1)
'excrement'	C1	khii	khii	khi	khii	hai (C2)
'tooth'	C1	khiau	khɛɛu	xɛu	kheeu	heeu
'to knot'	DL1	khɔɔt	-	xɔt	khoot	hoot
'year (of age of children)'	DL1	khuap	-	khop	khuup	hoop
'guest, stranger'	DL1	khɛɛk	-	khɛk	kheek	heek
'to bite'	DS1	khop	-	khop	khoop (DL1)	hap (DS2)
'to dig, hollow out'	DS1	khut	-	khut (DS1), kut (DS2)	kut (DS2)	hut (DS2)

Considering the development of the Proto-Tai *x in Table 10, it can be noticed that the Proto-Tai *x developed to /kh-/ in some lexical items in Lue/Lü while in some the Proto-form *x- is still maintained. Furthermore, there are also some lexical items in Lao showing that the Proto-Tai *x- mostly developed to /kh-/ but some lexical items use both initial /kh-/ and /h-/.

Based on the lexical data in Phuan, studied in this research, it is noticed that in some but not all live syllables, i.e. the syllables with no final consonant or with sonorant final consonants, e.g. [-m, -n, -ŋ], if the initial is [h-], it will be [kh-] in the other Tai dialects as shown in Table 11.

Table 11 The examples of the lexical items in Phuan (Fay Mun), pronounced with the initial [h-] in comparison with the initial [kh-] in the other Tai dialects

Gloss	Phuan	The other Tai dialects			
		Nyuan	Lue[20]	Khün	Lao
'to carry off'	hon^{23}	khon^{23}	khun^{35}	khun^{23}	khon^{434}
'thick (as of liquid)'	hon$^{33'}$	khun$^{44'}$	khun$^{213'/41'}$	khun$^{44'}$	khun$^{44'}$
'to cross over'	haam$^{44'}$	khaam$^{44'}$	khaam$^{213'/41'}$	khaam$^{44'}$	khaam$^{44'}$
'to rise, to ascend'	hɨn$^{44'}$	khɨn$^{44'}$	khɨn$^{213'/41'}$	khɨn$^{44'}$	khɨn$^{44'}$

[20] The first tone refers to the tones of Lue (Nong Bua Village) and the second one refers to the tone in Lue (Donchai Village). If the lexical item in the two varieties of Lue has the same tone, only one tone symbol is shown.

To compare the lexical items in Phuan in Table 11 to the development of Proto-Tai *x- in Table 10, it can be said that the Proto-Tai *x- developed to h- in Phuan, similar to the case of some words in Lao and the development of *x in Po-ai, a Northern Tai dialect. A number of lexical items in Phuan studied in this research show the variation of initial consonant h- to some degree as shown in Figure 4, and examples of lexical items pronounced with different variants are shown in Table 12.

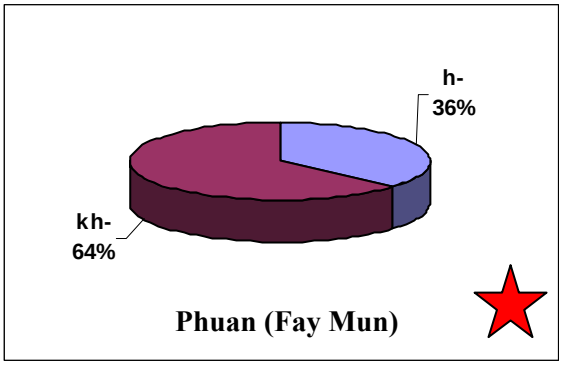


Figure 4 The percentage of the usage of the voiceless glottal fricative initials' variant [h > kh] in Phuan (Fay Mun)

Based on the lexical data of Phuan in this research, the voiceless glottal fricative [h-] is still maintained in lexical items around 36% of the time while [kh-], its variant, is used 64%, a difference of 28%. Some examples of the lexical items which still have the initial consonant [h-] and some which have the initial [kh-] are shown in Table 12.

Table 12 Examples of lexical items in Phuan (Fay Mun) pronounced with the initial [h-] and its variant [kh-]

Initial consonant /h-/ in Phuan (Fay Mun)			
no variation		**variation**	
Gloss	**/h-/**	**Gloss**	**h- → kh- = Nyuan/Lue/Khün/ Standard Thai**
'body hair, to carry off'	***h***on^{23}	*'bitter'*	***kh***om^{23} (← ***h***om^{23})
'thick (as liquid)'	***h***un$^{33'}$	*'leg'*	***kh***aa^{23} (← ***h***aa^{23})
'to cross over'	***h***aam$^{33'}$	*'galangal'*	***kh***aa^{22} (← ***h***aa^{22})
'to twist into a knot'	***h***ɔɔt^{22}	*'to write'*	***kh***iən^{23} (← ***h***iən^{23})
'arm'	***h***ɛɛn^{23}	*'fish basket'*	***kh***ɔɔŋ$^{33'}$ (← ***h***ɔɔŋ$^{33'}$)

Comparing to the *Dictionary of Thai Phuan*[21] (Thai Phuan Foundation, 1989), the lexical items pronounced with the initial [h-], derived from Proto-Tai *x- can be shown in Table 13.

[21] Phuan has been spoken in some areas of Thailand: central, north, and northeast. Phuan has been also called "Thai Phuan" (ไทยพวน/ไทพวน) and "Lao Phuan" (ลาวพวน). In the available dictionary, Phuan is called "Thai Phuan".

Table 13 Variation of initial [h-] in Phuan in the present study in comparison with the Dictionary of Thai Phuan

Gloss	Phuan	
	From Dictionary of Thai Phuan[22]	From the present study
'body hair, to carry off'	hon^{LF} (หน)	hon^{23}
'to cross over'	haam^{HR} (ห้าม)	haam$^{44'}$
'to twist into a knot'	-	hɔɔt^{22}
'to rise, to ascend'	-	hɨn$^{44'}$
'needle'	hem^{LF} (เห็ม)	hem^{23}
'horn'	haw^{LF} (เหา)	haw^{23}
'to enter'	haw^{HR} (เห้า)	haw$^{33'}$
'chopping block, cutting board'	hiəŋ^{LF} (เหียง)	hiəŋ^{23}
'arm'	khɛɛn^{LF} (แขน)	hɛɛn^{23}
'thick (as of liquid)'	hun^{HR}, khun^{HR} (หุ้น, ขุ้น)	hon$^{33'}$
'fang, canine tooth'	hɛɛw^{HR}, khɛɛw^{HR} (แห้ว, แข้ว)	hɛɛw$^{33'}$ fɔɔŋ^{35}
'log'	hɔɔn^{LR}, khɔɔn^{LR} (หอน, ขอน)	khɔɔn^{23}

As can be seen from Table 13, even in the *Dictionary of Thai Phuan* some lexical items have both [h-] and [kh-] as the initial consonants. A number of lexical items in the dictionary have the initial [h-] as in Phuan in this study. From this evidence, it may be plausible to hypothesize that the variation of [h-] might have been occurring for a long time and it is still ongoing process of change. Since Phuan, spoken in Fay Mun village where the people have contact with Northern Thai, Lue, and Khün people, it can be hypothesized that [kh-], the variant of [h-] which has been used in some lexical items, has been borrowed from the surrounding Tai dialects. The other possibility is the influence of Standard Thai which Phuan people have a chance to perceive from the mass media: television, radio, newspaper, and so on. Further study is needed to make a conclusion about this issue.

5.2 The variation of vowels in Tai dialects

Mary Sarawit (1973) reconstructed the Proto-Tai vowels by using the data from 115 dialects in Northern, Central, and Southwestern Tai classified in Li (1960). The vowels in Proto-Tai reconstructed in Sarawit (1973) are shown in Table 14.

[22] The tones in the Dictionary of Thai Phuan are described in comparison with Standard Thai, for example, there are six tones and they are mostly similar to Standard Thai tones, except for some cases. That is, in the lexical items with the initial consonants [p-, t-, k-, c-, b-, d-, ʔ-] and written with falling tone marker or [máay thoo] (◌๊), for example, ป้า [paa^{42}] 'aunt', ต้ม [tom^{42}] 'to boil', บ้าน [baan^{42}] 'house', etc. in Standard Thai they will have high-falling tone but in Phuan they have high-rising tone (HR). In the lexical items which have low-rising tone in Standard Thai, they have low-falling tone (LF) in Phuan. In the dead syllables with short vowels, they have high-rising (HR) or high-falling (HF), different from Standard Thai.

Table 14 Proto-Tai vowels (Sarawit, 1973)

a) Monophthongs:	*i	*iː	*ɨ	*ɨː	*u	*uː
	*e	*eː	*ə	*əː	*o	*oː
			*a	*aː		
b) Diphthongs:	*ia	*iaː	*ɨa	*ɨaː	*ua	*uaː
	(*eu	*euː)			(*oi	*oiː)

As for Proto-Southwestern Tai, Sarawit reconstructed 12 monophthongs and 3 diphthongs as shown in Table 15.

Table 15 Proto-Tai vowels (Sarawit, 1973)

a) Monophthongs:	*i	*iː	*ɨ	*ɨː	*u	*uː
	*e				*o	
		*ɛː	*a	*aː		*ɔː
b) Diphthongs:	*ia	*ɨa		*ua		

In *A Handbook of Comparative Tai* Li (1977) reconstructed Proto-Tai vowels in three categories: monophthongs, diphthongs, and triphthongs as shown in Table 16.

Table 16 Proto-Tai vowels (Li, 1977)

a) Monopthongs:	*i	*ɨ	*u				
	*e	*ə	*o				
	*ɛ	*a	*ɔ				
b) Diphthongs:	*ie	*ɨe	*ue				
	*əɨ	*ei/*ɛi		*a̯u		*iu	
c) Triphthongs:	*iau	*iəu	*uai	*uəi	*i̯ai	*u̯əi	*i̯əu

In some of the five Tai dialects investigated in this study, there is variation of some vowels, i.e. /o, iə, ɨə, uə/.[23] Table 17 shows the distinction of such vowels by using the vowels in Proto-Tai reconstructed by Sarawit (1973) and Li (1977) as the starting point.

Table 17 The vowel distinction between Standard Thai and other five Tai dialects

Proto-Tai (Li, 1977)	**Standard Thai**	**Nyuan (Mueang Nan)**	**Phuan (Fay Mun)**	**Lao (Nam Pua)**	**Lue (Nong Bua, Donchai)**	**Khün (Nong Muang)**
*e	e [khem^{23}] *'needle'*	e [khem^{23}]	e [khem^{23}]	e [khem^{434}]	i [khim^{35}]	i [khim^{23}]
*o	o [kon^{42}] *'buttocks'*	o [kon$^{44'}$]	o [kon$^{33'}$]	o [kon$^{44'}$]	u [kun$^{213'/31}$]	u [kun$^{44'}$]

[23] The diphthongs /iə, ɨə, uə/ in this research are transcribed by using /ə/ as the second element while they are reconstructed as *ia, *ɨa, ua in Sarawit (1973) and *ie, *ɨe, *ue in Li (1977). The distinction of such second elements in the Proto-Tai vowels will not be considered when the vowel variation in this research is determined.

Table 17 (continued)

Proto-Tai (Li, 1977)	Standard Thai	Nyuan (Mueang Nan)	Phuan (Fay Mun)	Lao (Nam Pua)	Lue (Nong Bua, Donchai)	Khün (Nong Muang)
*ie	iə [tiəŋ^{33}] *'bed'*	iə [tiəŋ^{23}]	iə [tiəŋ^{35}]	iə [tiəŋ^{23}]	ee [teeŋ^{35}]	ee [teeŋ^{23}]
*i̯e	i̯ə [ri̯ə^{33}] *'boat'*	iə ?? [hiə^{35}] **Note:** 'salt' [ki̯ə^{23}][24]	iə [hiə^{35}]	i̯ə [hi̯ə^{23}]	əə [həə$^{31/44}$]	əə [həə^{35}]
*ue	uə [huə^{24}] *'head'*	uə [huə^{23}]	uə [huə^{23}]	uə [huə^{434}]	oo [hoo^{35}]	oo [hoo^{35}]

5.2.1 The variation of /iə/ in Nyuan and Phuan and the variation of /i̯ə/ in Lao

Li (1977) reconstructed the Proto-Tai *ie and stated that this diphthong with the accent on i has the following vowel lowered and relaxed giving /iə/ or /ia/, which are the pronunciations of certain dialects. In Lungchow (Central Tai Group) and Po-ai (Northern Tai Group) this diphthong is further monophthongized to /ii/; in some Southwestern Tai dialects, such as Shan, Lue/Lü, White Tai, etc., it is monophthongized to e. The development of the Proto-Tai *ie in Siamese, Lungchow, and Po-ai is shown in Table 18.

Table 18 The development of Proto-Tai *ie (Li, 1977: 280-281)[25]

Gloss	Tone	Siamese (SWT)	Lungchow (CT)	Po-ai (NT)
'to lick'	A2	lia	lii	lii
'to change'	B1	plian	piin	piin
'to feed'	C2	liaŋ	čii̯n (?)	šiiŋ
'chopping block'	A1	khiaŋ	khiiŋ	heeŋ (?)
'tripod'	A2	-	khiiŋ	čiiŋ
'to stretch'	DL1	jiat	jiit	jiit

Among the five Tai dialects investigated in this study, the [i] of the diphthong *ie is still maintained while the second element [e] is weakened to [ə], giving /iə/ in Nyuan, Phuan, and Lao. In Lue and Khün, the diphthong *ie is monophthongized and lengthened to /ee/ (see the examples in Table 17). From comparing the vowels among the five Tai dialects, it is found that there is the variation of the diphthong /iə/ in two Tai dialects, Nyuan and Lao as shown in Figure 5.

[24] In Nyuan (Mueang Nan), the lexical item for 'salt' is pronounced [ki̯ə^{23}] which means that there is vowel variation [iə → i̯ə].

[25] The ' ? ' which is shown after some cases of cognates in the Tai languages in Li (1977) presented the questionable development in such Tai languages.

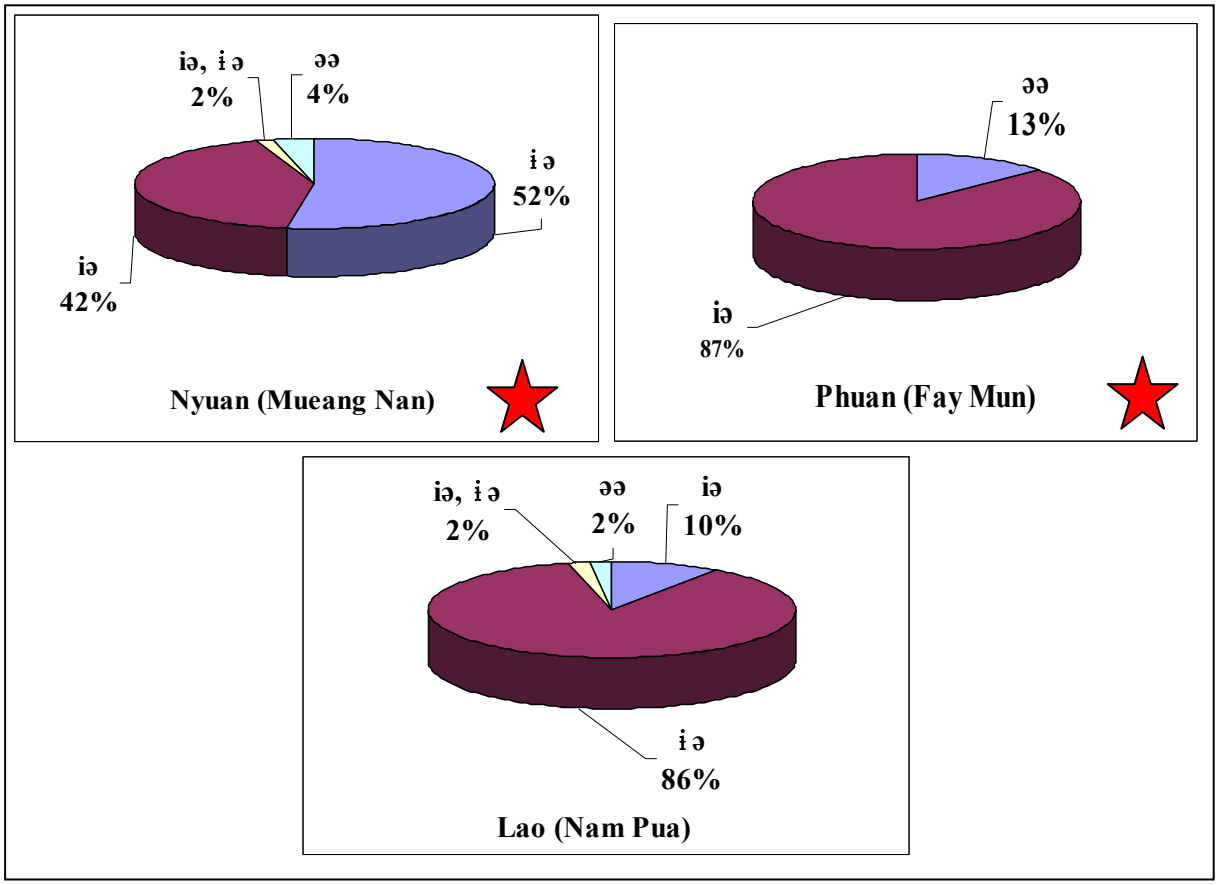


Figure 5 The variation of /iə/ in Nyuan (Mueang Nan) and Phuan (Fay Mun) and of /iə/ in Lao (Nam Pua)

The Proto-Tai *ie is normally developed to /iə/ in Nyuan (Mueang Nan) and Phuan (Fay Mun) but it is found in this study that the diphthong /iə/ in Nyuan has three variants, i.e. [iə, ɨə, əə] and two variants in Phuan (Fay Mun): [iə, əə]. The usage of each variant of the diphthong /iə/ in Nyuan has a different percentage, arranged in order from high to low as follows: [ɨə], [iə], [əə], as well as the usage of both [iə] and [ɨə], i.e. some lexical items have been pronounced with both diphthong [iə] and the monophthong [əə]. As for the case of Phuan, the usage of the variant [əə] is found only 13%, less than the other variant [iə]. The examples of the lexical items in which there is the variation of /iə/ in Nyuan and Phuan are shown in Table 19.

Table 19 Examples of variation of /iə/ in Nyuan (Mueang Nan) and Phuan (Fay Mun)

Diphthong /iə/ in Nyuan (Mueang Nan) and Phuan (Fay Mun)			
no variation		**variation**	
Gloss	**The diphthong**	**Gloss**	**The Variation of the diphthong /iə/ in Nyuan (Mueang Nan)**
	/iə/		**iə → ɨə = Standard Thai** **→ əə = Lue**
'annoyed, irritated'	kh*iə*ŋ^{35}	*'salt'*	k*iə*^{23} (← k*iə*^{23})
'knot'	ŋ*iə*n^{31}	*'classifier of banana'*	kh*iə*^{35} (← kh*iə*^{35})
'to poison'	b*iə*^{33}	*'to believe'*	c*iə*^{31} (← c*iə*^{31})

Table 19 (continued)

Diphthong /iə/ in Nyuan (Mueang Nan) and Phuan (Fay Mun)			
no variation		**variation**	
Gloss	**/iə/**	**Gloss**	**Nyuan (Mueang Nan)**
			iə → iə = Standard Thai **→ əə = Lue**
'house'	h*iə*n^{35}	*'boiling'*	d*iə*t^{33} (← d*iə*t^{33})
'friend'	ph*iə*n^{31}	*'moon, month'*	d*iə*n^{35} (← d*iə*n^{35})
'boat'	h*iə*^{35}	*'yellow'*	l*iə*ŋ^{23} (← l*iə*ŋ^{23})
'blood'	l*iə*t^{31}	*'taro'*	ph*iə*k^{33} (← ph*iə*k^{33})
'gum (of teeth)'	ŋ*iə*k^{35}	*'to choose, to select'*	l*iə*k^{31}, l*iə*k^{31} (← l*iə*k^{31})
'bedbug'	h*iə*t^{31}	*'saw'*	l*əə*y^{31} (← l*iə*y^{31} ← l*iə*y^{31})
'remaining, left over'	l*iə*^{23}	*'to crawl (of snakes), to creep (of plants)'*	l*əə*y^{43} (← l*iə*y^{43} ← l*iə*y^{43})
Gloss	**/iə/**	**Gloss**	**Phuan (Fay Mun)**
			iə → (iə →) əə = Lue
'to return'	m*iə*^{35}	*'naked'*	p*əə*y^{35} (← p*iə*y^{35} ← p*iə*y^{35})
'salt'	k*iə*^{35}	*'stiff (of muscles)'*	m*əə*y^{31} (← m*iə*y^{31} ← m*iə*y^{31})
'moon, month'	d*iə*n^{35}	*'Job's tears'*	ma-d*əə*y^{35} (← ma-d*iə*y^{35} ← ma-d*iə*y^{35})
'bed bug'	h*iə*t^{31}	*'to roll the eyes up'*	l*əə*ʔ^{22} (← l*iə*ʔ^{22} ← l*iə*ʔ^{22})

Since there is no /iə/ used in the Tai dialects spoken in the same area of Nyuan and Phuan, it is suspected that the variant [iə] is influenced by Standard Thai. One of the noteworthy case is that of the variant [əə] which is suspected to be influenced by Lue[26]. The long monophthong vowel /əə/ in Lue is basically derived from Proto-Tai *ie (see the examples in Table 17) and it is borrowed into both Nyuan and Phuan. There are two possibilities of the borrowing strategies, firstly, the vowel [iə] may be borrowed from Standard Thai into Nyuan and derived to be [əə] due to the influence of Lue, one of the main Tai dialects in the studied area which has the vowel /əə/, derived from Proto-Tai *ie. Secondly, it is possible that the [əə] is borrowed directly from Lue into Nyuan and Phuan without any stage of derivation.

As for the case of the diphthong /iə/, Li (1977) reconstructed Proto-Tai *ie and stated that this diphthong with the accent on /i/ has the following vowel lowered and relaxed, giving /iə/ or /ia/ in some dialects. In Lungchow and Po-ai it is further monophthongized to /ii/. In some Northern Tai dialects (i.e. Pu-i) it remains a diphthong in open syllables, but is

[26] The Proto-Tai *ie is also derived to [əə] in Khün but the variant [əə] in Lao (Nam Pua) is not suspected to be influenced by Khün because Khün is spoken by small amount number of speakers in the villages very far away from Lao (Nam Pua). Thus it is hypothesized that Khün is not the donor language of the variant [əə] in Lao (Nam Pua).

monophthongized in closed syllables. In some Southwestern Tai dialects, such as Shan, Lue/Lü, White Tai, etc., it is monophthongized to /ə/, a mid unrounded back vowel. The development of the Proto-Tai *ɨe is shown in Table 20.

Table 20 The development of Proto-Tai *ɨe (Li, 1977: 281-282)

Gloss	**Tone**	**Siamese (SWT)**	**Lungchow (CT)**	**Po-ai (NT)**
'time'	B2	mɨa	mɨɨ	mɨɨ
'disgusted, bored'	B1	bɨa	bii	mii
'month'	A1	dɨan	bɨin	nɨin
'blood'	DL2	lɨat	lɨɨt	lɨɨt
'taro'	DL1	phɨak	phɨɨk	piik
'mythological sea monster, dragon'	DL2	ŋɨak	phɨɨk	piik

Based on the author's previous studies, if the lexical items in Standard Thai are pronounced with the vowel [ɨə], they will be pronounced with the vowel [iə] in several Lao dialects, especially the ones spoken in northeastern Thailand. In other words, the Proto-Tai *ɨe is derived to /ɨa/ or /ɨə/ in Standard Thai but to /ie/ or /iə/ in several Lao dialects. As for the case of Lao (Nam Pua) which is analyzed to be Lao Luang Prabang accent in this study (see more details in Section 5.3.4), the reflex of the Proto-Tai *ɨe is /ɨə/ which means that the proto-diphthong *ɨe is still maintained in Lao (Nam Pua) with very little change, i.e. the weakening of the second element: e > ə.

Similar to the case of the /iə/ variation in Nyuan, the variation of /ɨə/ in Lao (Nam Pua) is also found to be influenced by the surrounding Tai dialects, especially Nyuan and Lue. As can be seen in Figure 5, the variation of /ɨə/ in Lao (Nam Pua) occurs to a lower degree than the variation of /iə/ in Nyuan but the variants of such vowels, i.e. /iə/ in Nyuan and /ɨə/ in Lao (Nam Pua), seem to be similar. The variants [iə] and [əə] in Lao (Nam Pua) are suspected to be influenced by Nyuan and Lue respectively.

It is noticeable that the percentage of [iə] usage in Lao (Nam Pua) is higher than the percentage of [əə] usage. This is due to the fact that most of Lao (Nam Pua) speakers, including the LRPs of this study, can also speak Nyuan, which has been used as the lingua franca among people speaking different dialects/languages in Nan province as aforementioned. There is no doubt that the vowel [iə] is borrowed from Nyuan (which is probably influenced by Lue) into Lao (Nam Pua) to a higher degree than [əə]. The examples of /ɨə/ variation in Lao (Nam Pua) are shown in Table 21.

Table 21 Examples of lexical items in Lao (Nam Pua), pronounced with the variants of the diphthong [iə]

Diphthong /iə/ in Lao (Nam Pua)			
no variation		**variation**	
Gloss	**/iə/**	**Gloss**	**iə → iə = Nyuan** **→ əə = Lue**
'salt'	k*iə*^{23}	*'annoyed, irritated'*	kh*iə*ŋ^{35} (← kh*iə*ŋ^{35})
'classifier for banana'	kh*iə*^{35}	*'knot'*	ŋ*iə*n^{31} (← ŋ*iə*n^{31})
'boiling'	d*iə*t^{44}	*'overhanging, rock'*	ŋ*iə*m^{43} (← ŋ*iə*m^{43})
'moon, month'	d*iə*n^{23}	*'leprosy'*	khii$^{44'}$ c*iə*n^{35}, khii$^{44'}$ h*iə*n^{35} (← khii$^{44'}$ h*iə*n^{35})
'to poison'	b*iə*^{21}	*'Jub's tears'*	mak^{35} d*əə*y^{23} hiin^{434} (← mak^{35} d*iə*y^{23} hiin^{434} ← mak^{35} d*iə*y^{23} hiin^{434})

5.2.2 The variation of /uə/ in Phuan

The other interesting case of vowel variation is that the vowel /uə/ in Phuan. Li (1977) reconstructed Proto-Tai *ue and stated that this diphthong has the accent on the first element and the vowel [u] is retained. The second element, being unaccented, is relaxed to [ə] or [a], becoming Siamese and Lao /ua/. It is monophthongized to /uu/ in Lungchow and Po-ai. In some Southwestern Tai dialects, it is monophthongized to /o/, as in Shan, Lue/Lü, White Tai, etc. In Po-ai, it is further shortened to u in closed syllables. The examples of the Proto-Tai *ue different Tai dialects are shown in Table 22.

Table 22 The development of Proto-Tai *ue (Li, 1977: 281-282)

Gloss	**Tone**	**Siamese (SWT)**	**Lungchow (CT)**	**Po-ai (NT)**
'classifier of animals'	A1	tua	tuu	tuu
'bean'	B1	thua	thuu	tuu (B2)
'head'	A1	hua	huu	-
'to leak'	B2	rua	łuu	luu
'deaf'	DL1	nuak	nuuk	nuk (DS1)
'to overflow, inundate'	C1	thuam	thuum	tum (B2)

Based on the lexical data of the five Tai dialects in this study, the Proto-Tai *ue is developed to /uə/ in Nyuan (Mueang Nan), Phuan (Fay Mun) and Lao (Nam Pua), but to the long monophthong /oo/ in Lue (Nong Bua, Donchai) and Khün (Nong Muang) as shown in Table 17. More examples of lexical items, showing the development of the Proto-Tai *ue in the studied five Tai dialects are shown in Table 23.

Table 23 The development of Proto-Tai *ue in five Tai dialects

Gloss	*ue > uə			*ue > oo	
	Nyuan (Mueang Nan)	Phuan (Fay Mun)	Lao (Nam Pua)	Lue (Nong Bua, Donchai)[27]	Khün (Nong Muang)
'classifier for round shapes or objects'	d***uə***ŋ^{35}	d***uə***ŋ^{35}	d***uə***ŋ^{23}	d***oo***ŋ^{35}, d***o***ŋ^{44}	d***oo***ŋ^{35}
'to check, to inspect, to examine'	k***uə***t^{33}	k***uə***t^{22}	k***uə***t^{44}	k***oo***t$^{24/21}$	k***oo***t^{22}
'mud hole'	b***uə***k^{33}	b***uə***k^{22}	b***uə***k^{44}	b***oo***k$^{24/21}$	b***oo***k^{24}
'ear (of paddy, corn)'	ŋ***uə***ŋ^{35}	h***uə***ŋ^{35}	h***uə***ŋ^{23}	h***oo***ŋ$^{31/31}$	h***oo***ŋ^{35}
'to pan-fry (as meat)'	kh***uə***$^{44'}$	kh***uə***$^{33'}$	kh***uə***$^{44'}$	kh***oo***$^{213'/41'}$	kh***oo***$^{44'}$
'to leak'	h***uə***^{31}	h***uə***^{31}	h***uə***^{21}	h***oo***$^{33/32}$	h***oo***^{31}
'to scald, to parboil'	l***uə***k^{31}	l***uə***ʔ^{31}	l***uə***k^{35}	l***oo***k$^{33/32}$	l***oo***k^{31}
'courtyard, lawn'	kh***uə***ŋ^{33}	kh***uə***ŋ^{22}	kh***uə***ŋ^{21}	kh***oo***ŋ$^{24/21}$	kh***oo***ŋ^{22}
'tadpole'	ʔi-h***uə***k^{31}	ʔi-h***uə***k^{31} (Note: There is variation of -ʔ → -k)	ʔii^{21} h***uə***k^{35}	ʔi-h***oo***k$^{33/32}$	ʔi-h***oo***k^{31}
'ox, cow'	ŋ***uə***^{35}	ŋ***uə***^{35}	ŋ***uə***^{23}	ŋ***oo***$^{31/44}$	ŋ***oo***^{35}
'to have fun'	m***uə***n^{31}	m***uə***n^{31}	m***uə***n^{21}	m***oo***n$^{33/32}$	m***oo***n^{31}
'things '	kh***uə***^{35}	kh***uə***^{35}	kh***uə***^{23}	kh***oo***$^{31/44}$	kh***oo***^{35}

As can be seen in Table 23, Proto-Tai *ue is supposed to develop to /uə/ in Phuan, i.e. only the second element [e] is relaxed to [ə]. But from investigating the usage of /uə/ in Phuan, spoken in Fay Mun village, it is found that there are two variants of the diphthong /uə/; [uə] and [oo] which have been used in different percentage as shown in Figure 6.

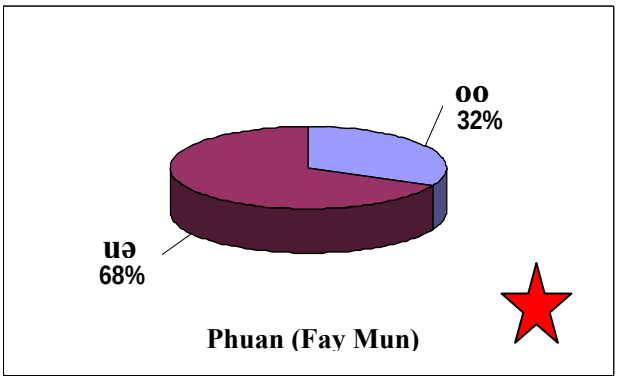


Figure 6 The variation of /uə/ in Phuan (Fay Mun)

It is clearly shown that even though the diphthong [uə] is still used in Phuan (Fay Mun) in higher percentage (68%), the variant [oo] is used 32%. This is considered to be not a small

[27] In the case that the lexical items in Lue, spoken in Nong Bua and Donchai villages are different, the lexical item of Lue (Nong Bua) will be shown firstly and the one of Lue (Donchai) will be in the second order. In the case that both Lue dialects use the same lexical items but different tones, the tones of Lue (Nong Bua) and Lue (Donchai) will be shown respectively.

degree of usage when compared to the variation of the other vowels in Phuan, for example the variation of ie > əə, described in 5.2.1. The examples of the lexical items which still maintain the former variant [uə], compared to the ones with the new variant [oo] are shown in Table 24.

Table 24 The variation of /uə/ in Phuan (Fay Mun)

Diphthong /uə/ in Phuan (Fay Mun)			
no variation		**variation**	
Gloss	**The diphthong /uə/**	**Gloss**	**uə → oo = Lue, Khün**
'spoon'	b***uə***ŋ^{22}	*'banana'*	k***oo***y$^{33'}$ (← k***uə***y$^{33'}$)
'to check'	k***uə***t^{22}	*'bowl'*	th***oo***y$^{33'}$ (← th***uə***y$^{33'}$)
'mud hole'	b***uə***k^{22}	*'net for catching fish (big size)'*	ʔ***oo***n^{23} (← ʔ***uə***n^{23})
'ox, cow'	ŋ***uə***^{35}	*'fence'*	h***oo***$^{41'}$ (← h***uə***$^{41'}$)
'ear (of paddy, corn)'	h***uə***ŋ^{35}	*'to smile'*	h***oo***^{23} (← h***uə***^{23})
'to scald, to parboil'	l***uə***ʔ^{31}	*'classifier for egg, fruit, pebble'*	n***oo***y^{22} (← n***uə***y^{22})
'to have fun'	m***uə***n^{31}	*'coconut-shell dipper'*	nam$^{41'}$ b***oo***y^{35} (← nam$^{41'}$ b***uə***y^{35})
'loufah gourd, sponge gourd'	ma-b***uə***p^{22}	*'weaver's shuttle'*	ma-s***oo***y^{23} (← ma-s***uə***y^{23})
'to leak'	h***uə***^{31}	*'to cover with'*	kh***oo***p^{31} (← kh***uə***p^{31})

Since Fay Mun, Nong Bua, and Nong Muang, the villages of Phuan, Lue, and Khün respectively, are in the same sub-district (Pa Kha) and they are not far from each other, the people in these three ethnic groups have been in contact for a very long time. According to information from Phuan LRPs, some Phuan people can speak Lue and/or Khün. Thus, it is plausible to hypothesize that the variant [oo] used in Phuan language has been borrowed from Lue and/or Khün. Even though the borrowed vowel [oo] has been used in a low percentage of words at present, it will be interesting to investigate the tendency of the vowel change in Phuan in the future.

5.2.3 Notes on the variation of the other vowels in the five Tai dialects

Besides the variation of the diphthongs: /iə/ in Nyuan (Mueang Nan) and Phuan (Fay Mun), /ɨə/ in Lao (Nam Pua), and /uə/ in Phuan (Fay Mun), there are some other vowels having different variants which will perhaps lead to change in the future. The other vowels in the Tai dialects which have variants are briefly described in 5.2.3.1-5.2.3.3.

5.2.3.1 The variation of /u/ in Lue (Donchai) and Khün (Nong Muang)

The monophthong /u/ in Lue (Donchai) and Khün (Nong Muang) which is developed from the Proto-Tai *o (see Table 17) has two variants; [u] and [o]. In other words, some lexical items of these two Tai dialects are expected to be pronounced with [u] but they are pronounced with [o] which is hypothesized to be influenced by the surrounding dialects, especially Nyuan or perhaps by Standard Thai via the mass media. The other noticeable point is that of Proto-Tai *o, raised to /u/ particularly in short live syllables ending with nasal consonants, for example, in Lue (Donchai): k***u**n$^{213'}$ 'buttocks', th***u**m^{24} 'to spit out', s***u**ŋ^{24} 'to send (a piece of news)'. Based on the data in this study, there is only one example of dead syllable ending with a stop consonant [p], e.g. l***u**p^{33} 'to erase' in Lue (Donchai).

The examples of some lexical items in Lue (Donchai) and Khün (Nong Muang), pronounced with the vowel [u] and the borrowed variant [o] are shown in Table 25. Note that the lexical items in Lue (Sipsongpanna), presented in Kullavanijaya, et al (1984) are used as the starting point to show the clearer evidence of the /u/ variation. Furthermore, the lexical items in Lue (Nong Bua) with no variation of /u/ as well as the lexical items in Nyuan (Mueang Nan) in which the Proto-vowel *o is still maintained are also presented in Table 25. This is to show an evidence that the variant [o] in Lue (Donchai) and Khün (Nong Muang) has probably been borrowed from Nyuan.

Table 25 The variation of /u/ in Lue and Khün[28]

Gloss	***o > o**	***o > u**			
	Nyuan (Mueang Nan)	**Lue (Sipsong Panna)**[29]	**No variation**	**Variation: u → u, o**	
			Lue (Nong Bua)	**Lue (Donchai)**	**Khün (Nong Muang)**
'buttocks'	k***o**n$^{44'}$	-	k***u**n$^{213'}$	k***u**n$^{41'}$	k***u**n$^{44'}$
'body hair, feather'	kh***o**n^{23}	x***u**n^{55}	kh***u**n^{35}	kh***u**n^{35}	kh***u**n^{23}
'winnowing basket'	d***o**ŋ$^{44'}$	d***u**ŋ^{13}	d***u**ŋ$^{213'}$	d***u**ŋ$^{41'}$	d***u**ŋ$^{44'}$
'round'	m***o**n^{35}	m***u**n^{51}	m***u**n^{31}	m***u**n^{44}	m***u**n^{35}
'to erase'	l***o**p^{44}	-	l***u**p^{33}	l***u**p^{32}	l***u**p^{44}
'to bend the body or head'	k***o**m$^{44'}$	-	k***u**m$^{213'}$	k***o**m$^{41'}$ (← k***u**m$^{41'}$)	k***u**m$^{44'}$
'to spit out'	th***o**m^{33}	th***u**m^{35}	th***u**m^{24}	th***o**m^{21} (← th***u**m^{21})	th***u**m^{22}
'to boil'	t***o**m$^{44'}$	t***u**m^{13}	t***u**m$^{213'}$	t***o**m$^{41'}$ (← t***u**m$^{41'}$)	t***u**m$^{44'}$
'milk'	n***o**m^{35}	nam^{11} n***u**m^{51}	n***u**m^{31}	n***o**m^{44}	n***u**m^{31}

[28] If in any Tai dialect has different lexical item from the others, such lexical item will be shown in the parentheses, for example, the lexical item for 'to send (a piece of news)' in Khün (Nong Muang) is [bɔɔk^{22}] which is different from the others.

[29] The data of Lue (Sipsong Panna) is from Kullavanijaya, et al (1984). The Vocabulary of Six Tai Languages (**ศัพท์ไท 6 ภาษา**). Faculty of Arts, Chulalongkorn University (in Thai).

Table 25 (continued)

Gloss	*o > o	*o > u			
	Nyuan (Mueang Nan)	Lue (Sipsongpanna)	No variation	Variation: u → u, o	
			Lue (Nong Bua)	Lue (Donchai)	Khün (Nong Muang)
'to send (a piece of news)'	s***o***ŋ^{33}	s***u***ŋ^{35}	s***u***ŋ^{24}	s***o***ŋ^{21}	(bɔɔk^{22})
'vinegar'	nam^{43} s***o***m$^{44'}$	-	nam$^{22'}$s***u***m$^{213'}$	nam$^{52'}$ s***o***m$^{41'}$	nam$^{22'}$s***u***m$^{213'}$
'flag'	t***u***ŋ^{35}, th***o***ŋ^{35}	(thi^{11})	th***u***ŋ^{31}	t***u***ŋ^{44}, th***o***ŋ^{44} (← t***u***ŋ^{44})	th***u***ŋ^{31}
'to cover over'	th***o***m^{23}	-	th***u***m^{35}	th***u***m^{35}	th***o***m^{23} (← th***u***m^{23})
'to sink'	c***o***m^{23}	ts***u***m^{55}	c***u***m^{35}	c***u***m^{35}	c***o***m^{23} (← c***u***m^{23})
'to bump against'	c***o***n^{35}	-	c***u***n^{31}	c***u***n^{44}	c***o***n^{35} (← c***u***n^{35})
'sour'	s***o***m$^{44'}$	s***u***m^{13}	s***u***m$^{213'}$	s***u***m$^{41'}$	s***u***m$^{44'}$, s***o***m$^{44'}$ (← s***u***m$^{44'}$)

Comparing the vowel variation, u > o in Lue (Donchai) and Khün (Nong Muang) as shown in Table 25, it can be clearly seen that the variation of /u/ in these two Tai dialects occur in similar degree, i.e. the number of lexical items pronounced with [o] instead of [u] in these two Tai dialects are not significantly different. The [i] usage in Lue (Sipsong Panna) can be taken as evidence to suggest that the lexical items in Table 25 are expected to be pronounced with the vowel [u] in Lue but in Lue (Donchai) some are pronounced with [o].

It can be hypothesized that such variation occurs in Lue (Donchai) and Khün (Nong Muang) but not in Lue (Nong Bua) due to the fact that the Lue people in Donchai village and Khün in Nong Muang village are smaller group than Lue (Nong Bua), especially Khün which is the smallest group, compared to the other Tai dialects. Nyuan is suspected to be the donor language of the [o] variant in both Lue (Donchai) and Khün (Nong Muang), as the speakers of these two Tai dialects can speak Nyuan very fluently.

5.2.3.2 The variation of /i/ in Lue (Donchai) and Khün (Nong Muang)

Li (1977: 269) reconstructed the mid vowels in Proto-Tai and stated that it the Proto-Tai *e is found only in closed syllables and is always short. It is raised to /i/ in Lungchow, but lowered to /ɛ/ or /a/ in Po-ai. In some Southwestern Tai dialects, it is also raised to i before a nasal, such as Lue/Lü and White Tai. It is rounded to /o/ before /-m/ and /-p/ in Siamese and /u/ in Lunchow, but becomes /a/ in Po-ai. The development of the Proto-Tai *e in different Tai dialects is shown in Table 26.

Table 26 The development of Proto-Tai *e (Li, 1977: 269)

Gloss	Tone	Siamese (SWT)	Lungchow (CT)	Po-ai (NT)
'seven'	DS1	čet	čit	šɛt
'fish scales'	DS1	klet	kit	čɛt
'louse'	A2	len	min	nan
'to be'	A1	pen	pin (A2)	pan (A2)
'to see'	A1	hen	han (?)	han
'to fall, topple'	C2	lom	lum (B1)	lam

In Lue and Khün, investigated in this study, the Proto-Tai *e is raised to /i/ in live syllables ending with final nasal consonants, for example, 'needle': [khim^{35}] (Lue Nong Bua), [khim^{35}] (Lue Donchai); 'to jump up and down': [tin$^{213'}$] (Lue Nong Bua), [tin$^{44'}$] (Khün Nong Muang).

From considering the lexical items ending with nasal consonants and pronounced with the vowel [i] in Lue and Khün, it is found that there is a variation of /i/ in Khün and Lue (Donchai), but none in Lue (Nong Bua) as shown in Table 27.

Table 27 The variation of /i/ in Lue and Khün[30]

Gloss	*e > e	*e > i			
	Nyuan (Mueang Nan)	Lue (Sipsong Panna)[31]	No variation	Variation: i → i, e	
			Lue (Nong Bua)	Lue (Donchai)	Khün (Nong Muang)
'needle'	kh***e***m^{23}	x***i***m^{55}	kh***i***m^{35}	kh***i***m^{35}	kh***i***m^{23}
'full, filled'	t***e***m^{23}	t***i***m^{55}	t***i***m^{35}	t***i***m^{35}	t***i***m^{23}
'to stink, to smell bad'	m***e***n^{23}	m***i***n^{55}	m***i***n^{35}	(ʔoo^{21})	m***i***n^{23}
'to jump up and down'	t***e***n$^{44'}$	-	t***i***n$^{213'}$	t***e***n$^{41'}$ (← t***i***n$^{41'}$)	t***i***n$^{44'}$
'to take aim'	l***e***ŋ^{35}	-	l***i***ŋ^{31}	l***e***ŋ^{44} (← l***i***ŋ^{44})	l***e***ŋ^{35} (← l***i***ŋ^{35})
'line, strand'	s***e***n$^{44'}$	s***i***n^{13}	s***i***n$^{213'}$	s***i***n$^{41'}$	s***i***n$^{44'}$, s***e***n$^{44'}$ (← s***i***n$^{44'}$)
'to exert force in trying to expel'	b***e***ŋ^{33}	-	b***i***ŋ^{24}	b***i***ŋ^{21}	b***e***ŋ^{24} (← b***i***ŋ^{24})
'salty, salted'	k***e***m^{35}	ts***i***m^{51}	c***i***m^{31}	k***i***m^{44}	k***e***m^{35} (← k***i***m^{35})
'to play'	l***e***n$^{44'}$	d***i***n^{13}	l***i***n$^{213'}$	l***i***n$^{41'}$	l***e***n$^{44'}$ (← l***i***n$^{44'}$)
'to splash, to fly off'	sa-d***e***n^{35}	-	si?^{35} d***i***n^{35}	sa-d***i***n^{44}	sa-d***e***n^{35} (← sa-d***i***n^{35})
'to omit, to skip'	w***e***n^{43}	-	w***i***n$^{22'}$	w***i***n$^{52'}$	w***e***n$^{41'}$(← w***i***n$^{41'}$)

[30] If in any Tai dialect has different lexical item from the others, such lexical item will be shown in the parentheses, for example, the lexical item for 'to stink, to smell bad' in Lue (Donchai) is [ʔoo^{21}] which is different from the others.

[31] The data of Lue (Sipsong Panna) is from Kullavanijaya, et al (1984). The Vocabulary of Six Tai Languages (**ศัพท์ไท 6 ภาษา**). Faculty of Arts, Chulalongkorn University (in Thai).

Similar to the case of the /u/ variation described in 5.2.3.1, the lexical items which are live syllables ending with nasal consonants in Lue would be expected to be pronounced with [i] as same as Lue (Sipsongpanna) but some are pronounced with the variant [e]. As can be seen in Table 27, the [i] variation occurs in more lexical items in Khün (Nong Muang) than in Lue (Donchai). Even though the variation of [i] in both Lue (Donchai) and Khün (Nong Muang) occurs to a low degree, at least it can lead to the hypothesis of vowel change in these two Tai dialects in the future if there are more lexical items pronounced with [e] instead of [i]. Both Nyuan and Standard Thai could be suspected to motivate this variation.

5.3 Some problematic cases of tones in five Tai dialects

It has been generally known that tones can be used as one of the important criteria to distinguish the Tai languages into different groups. On the basis of the relationship between the initial consonants, the syllable types, and the tones, Fang Kuei Li (1977) reconstructed the Proto-Tai tones: A, B, C in the live syllables and D in the dead syllable which can be further divided into two categories on the criteria of the vowel length distinction, long dead syllable (DL) and short dead syllable (DS). Based on the Proto-Tai tone reconstruction by Li (1977) and the different patterns of tone development in modern Tai dialects, William J. Gedney (1972) proposed the 'tone box', a practical useful tool for collecting tone data in Tai dialects as shown in Figure 7.

		Proto-Tai Tones				
		A	**B**	**C**	**DL (D-long)**	**DS (D-short)**
1.	**Voiceless friction sounds:** ***s-, *hm-, *hn-, *ph-, *th-, *kh-**	ขา [khǎa] 'leg'	ไข่ [khàì] 'egg'	ห้า [hâa] 'five'	ขวด [khùət] 'bottle'	หก [hòk] 'six'
2.	**Voiceless unaspirated stops:** ***p-, *t-, *k-**	ปี [pii] 'year'	ปี่ [pìi] 'flute'	ป้า [pâa] 'aunt'	ปีก [pìik] 'wing'	กบ [kòp] 'frog'
3.	**Glottal sounds:** ***ʔ-, *ʔb-, *ʔd-**	บิน [bin] 'to fly'	ด่า [dàa] 'to scold'	อ้า [ʔâa] 'to open mouth'	ดีด [dìit] 'to flick'	ดิบ [dìp] 'raw, uncooked'
4.	**Voiced sounds:** ***b-, *d-, *g-, *m-, *n-, *l-, *r-**	นา [naa] 'rice field'	ไร่ [râi] 'farm'	น้ำ [náam] 'water'	มีด [mîit] 'knife'	มด [mót] 'ant'
		Smooth Syllable (Live Syllable)			**Checked Syllable (Dead Syllable)**	

Figure 7 The tone box (adapted from Gedney, 1972: 434)

The tone box can be used to show not only the patterns of tonal mergers and splits but also the characteristics of the tones. The lexical items shown in Figure 7 are taken from Standard Thai. To study the tones in different Tai/Thai dialects, at least four lexical items are needed to fill in each box of the tone box.[32]

5.3.1 The similar tones of Nyuan and Khün

Using the concept of tone box as the starting point to study the tones in the five Tai dialects, it is found that Nyuan (Mueang Nan) and Khün (Nong Muang) have exactly the same pattern of tonal mergers and splits; A12-34, BCD123-4, B=DL as shown in Figure 8.

[32] See the test words for the tones in the five Tai dialects in Appendix 2.

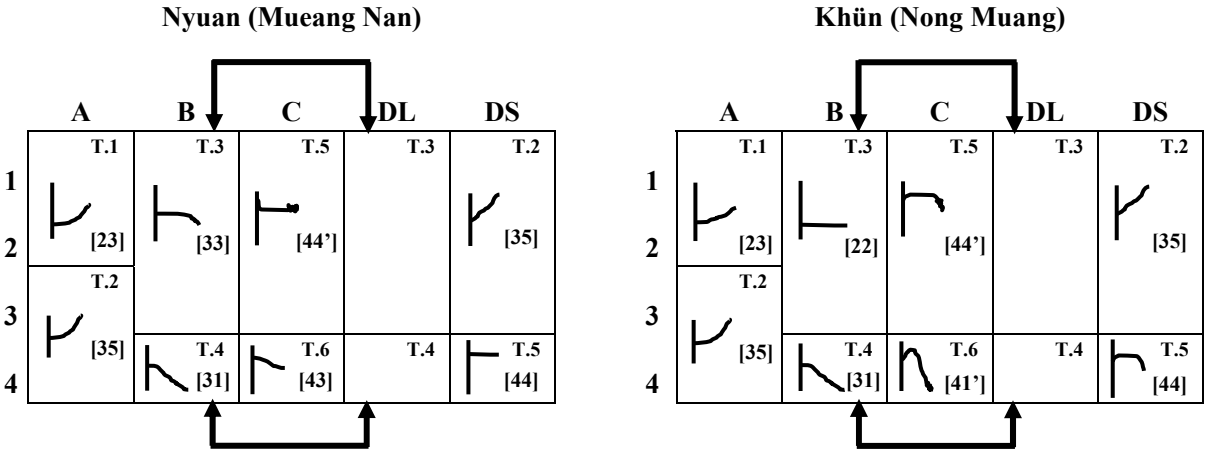


Figure 8 The tonal systems of Nyuan (Mueang Nan) and Khün (Nong Muang)

It can be seen in Figure 8 that there are a total of six tones in both Nyuan (Mueang Nan) and Khün (Nong Muang) and almost all of them have the same characteristics as shown in Table 28.

Table 28 The tonal characteristics in Nyuan (Mueang Nan) and Khün (Nong Muang)

	Occurrence	**Nyuan (Mueang Nan)**	**Khün (Nong Muang)**
Tone 1	A12	Low-Rising [23]	Low-Rising [23]
Tone 2	A34/DS123	Mid-Rising [35]	Mid-Rising [35]
Tone 3	B123/DL123	**Mid Level [33]**	**Low Level [22]**
Tone 4	B4/DL4	Mid-Falling [31]	Mid-Falling [31]
Tone 5	C123/DS4	High Level with glottalization [44']	High Level with glottalization [44']
Tone 6	C4	**High-Falling [43]**	**High-Falling with glottalization [41']**

It can be seen that only two tones in Nyuan (Mueang Nan) and Khün (Nong Muang), i.e. Tone 3 and Tone 6, have different characteristics but the distinction between the two tones in these two Tai dialects is only the level of tones. That is, Tone 3 in Nyuan is mid level [33] while in Khün it is low level [22], and Tone 6 in Nyuan is clear high-falling tone [43] while in Khün it is glottalized high-falling [41'].

The similarity of the tonal systems in Nyuan (Mueang Nan) and Khün (Nong Muang) leads to two possible hypotheses: firstly, it is possible that such similarity is caused by language contact, i.e. the tones in Khün have been influenced by the tones in Nyuan, not only phonetically (the tone characteristics) but also phonologically (the number of tones and the pattern of tone mergers and splits). Secondly, the similarity between the tones in Nyuan and Khün is not the outcome of language contact, but the genetic relationship.

It has been widely known among the Tai linguists that Khün (Khuen/Kuen) is the language of the city and state of Keng Tung in Eastern Shan State of Myanmar. It is closely related to Nyuan (Northern Thai/Kham Mueang) in Chiang Mai region and Lue in Sipsongpanna. Khuen, like the other two, uses a writing system similar in form to Mon-Burmese but in usage to Siamese (Egerod 1959: 124 cited in Gedney 1994: 977).

According to Gedney (1994), Khün (Klang Muong Khuen, Keng Tung)[33] has the following tonal system:

[33] Gedney worked with two Khuen (or Khün) speakers, the first LRP was a man born in Keng Tung who spoke Klang Muong Khuen, the dialect of that city. The second one was a man from the village of Baan Veng, located about five miles east of Keng Tung, Burma. Both LRPs referred to their language as Khuen (khɯn[35]), but the second LRP distinguished his dialect from Klang Muong Khuen. The interviewing occurred in Mae Sai,

On open syllables (those ending with a final vowel, semivowel, or nasal), Klang Muong Khuen has six tones. These tones, along with pitch levels and contours based on the Chao 5-level pitch scale (Chao 1930 cited in Gedney 1994: 978), include the following:

1. Mid-Rising [35]: nay^{35} 'where'
2. High Level [44]: nay^{44} 'inside'
3. Low with a slight rise [12]: kay^{12} 'chicken'
4. Low Level [22]: hay^{22} 'dry field'
5. Mid Level with glottalization [33']: khay$^{33'}$ 'fever'
6. High-Falling with glottalization [51']: may$^{51'}$ 'wood'

Checked syllables (those ending in -p, -t, -k, -ʔ) with short vowels have tones phonetically similar to tones 2 and 6 of open syllables. Checked syllables with long vowels have tones phonetically similar to tones 3 and 4 of open syllables. The historical origins of this tonal system can be represented in Figure 9.

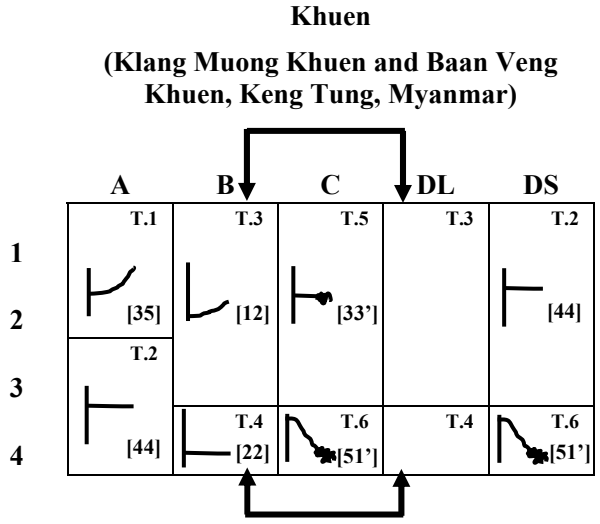


Figure 9 The tonal system of Khuen (Klang Muong, Keng Tung) (adapted from Gedney 1994: 978)[34]

The occurrence of each tone in Khuen (Klang Muong, Keng Tung) in comparison with Khün (Nong Muang) is shown in Table 29.

Table 29 The tonal characteristics in Khuen (Klang Muong and Baan Veng, Keng Tung) and Khün (Nong Muang)

	Occurrence	**Khuen (Klang Muong and Baan Veng, Keng Tung) (Gedney 1994)**	**Khün (Nong Muang)**
Tone 1	A12	Mid-Rising [35]	Low-Rising [23]
Tone 2	A34/DS123	High Level [44]	Mid-Rising [35]
Tone 3	B123/DL123	Low with a slight rise [12]	Low Level [22]
Tone 4	B4/DL4	Low Level [22]	Mid-Falling [31]
Tone 5	C123	Mid Level with glottalization [33']	High Level with glottalization [44']
Tone 6	C4/DS4	High-Falling with glottalization [51']	High-Falling with glottalization [41']

Thailand. Klang Muong Khuen was from 30 November to 1 December 1964 and Baan Veng Khuen was from 1-2 December 1964. (Gedney 1994: 977-1055)

[34] The pitch scales in Gedney (1994: 978) are adapted to be represented as tone sticks and put into the tone box by the author so that they can be compared with Khün (Nong Muang) in the present study.

It can be seen from Figures 9 and Table 29 that Khuen (Klang Muong Khuen and Baan Veng Khuen, Keng Tung) in Gedney (1994) and Khün (Nong Muang) have the same distinctive pattern of tonal mergers and splits, especially, A12-34 and B=DL. The only distinction is the mergers of C123=DS4 in Khün (Nong Muang) and C4=DS4 in Khuen (Klang Muong and Baan Veng, Keng Tung). As for the tonal characteristics, it is shown in Figure 9 that Tones 1, 5, and 6 in these two dialects of Khün are similar but Tones 2, 3, and 4 are different. Such similarities between Khuen (Klang Muongand Baan Veng, Keng Tung) and Khün (Nong Muang) may be taken as an evidence to support the close relationship between Nyuan and Khün.

5.3.2 The distinction of tones in Lue (Nong Bua) and Lue (Donchai)

From investigating the tones in two dialects of Lue spoken in two different villages, Nong Bua village (Pa Kha Sub-District, Tha Wang Pha District) and Donchai (Pua Sub-District, Pua District), it is found that the tonal systems of these two dialects of Lue are different, especially the merger and split pattern of Column A as shown in Figure 10.

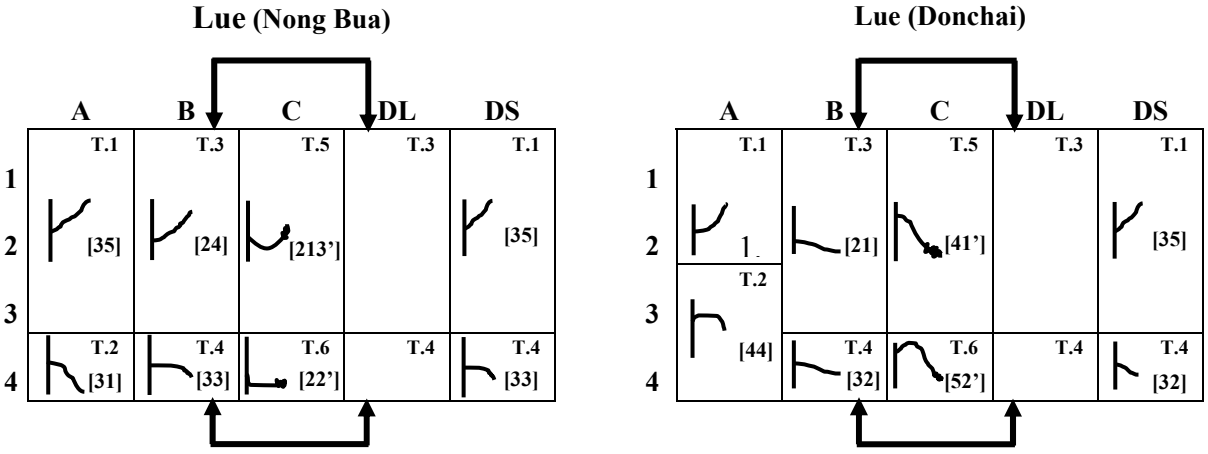


Figure 10 The tonal systems of Lue (Nong Bua) and Lue (Donchai)

There are six tones in both dialects of Lue and most of the tones have different characteristics as shown in Table 30.

Table 30 The tonal characteristics in Lue (Nong Bua) and Lue (Donchai)

	Occurrence	Lue (Nong Bua)	Lue (Donchai)
Tone 1	A123/DS123 (Nong Bua), A12/DS123 (Donchai)	Mid-Rising [35]	Mid-Rising [35]
Tone 2	A4 (Nong Bua), A34 (Donchai)	Mid-Falling [31]	High Level [44]
Tone 3	B123/DL123	Low-Rising [24]	Low-Falling [21]
Tone 4	B4/DL4	Mid Level [33]	Mid-Falling [32]
Tone 5	C123	Low-Falling-Rising with glottalization [213']	High-Falling with glottalization [44']
Tone 6	C4	Low Level with glottalization [22']	Extra-High-Falling with glottalization [52']

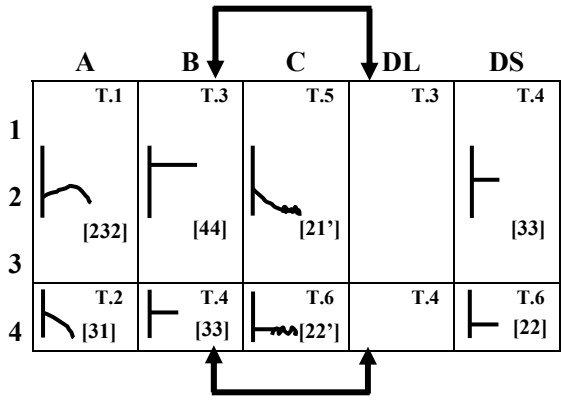
As shown in Figures 10 and Table 30, the patterns of tonal mergers and splits in Columns B, C, DL, and DS in both dialects of Lue are a two-way split, i.e. B/C/DL/DS 123-4, and there are also the mergers between Columns B and DL, i.e. B123=DL123 and B4=DL4. The distinction between these two systems is that of the pattern of merger and splits of Column A, that is A123-4 in Lue (Nong Bua) and A12-34 in Lue (Donchai). It can be clearly seen that the pattern A12-34 of Lue (Donchai) is the same as Nyuan (Mueang Nan) and the characteristics of some tones are similar as can be seen in the following Table 31.

Table 31 The comparision of tonal characteristics in Nyuan (Mueang Nan) and Lue (Donchai)

	Occurrence	**Nyuan (Mueang Nan)**	**Lue (Donchai)**
Tone 1	A12	Low-Rising [23]	Mid-Rising [35]
Tone 2	A34/DS123	Mid-Rising [35]	High Level [44]
Tone 3	B123/DL123	Mid Level [33]	Low-Falling [21]
Tone 4	B4/DL4	Mid-Falling [31]	Mid-Falling [32]
Tone 5	C123/DS4	High Level with glottalization [44’]	High-Falling with glottalization [41’]
Tone 6	C4	High-Falling [43]	Extra-High-Falling with glottalization [52’]

To compare the pattern of tonal mergers and splits of Lue (Nong Bua) and Lue (Donchai) with the other Lue dialects presented in the available previous studies may lead to a clearer explanation of the distinction between the tonal systems of the two dialects of Lue in this study. The patterns of tonal mergers and splits of Lue from the previous studies can be shown in Figure 11.

1. **Lue**: Yo Village (Mueang Bang, Sipsong Panna) (from L-Thongkum’s fieldnotes 1982)



2. **Lue**: Na Yang Tai Village (Nam Bak, Luang Prabang) (from L-Thongkum’s fieldnotes 1982)

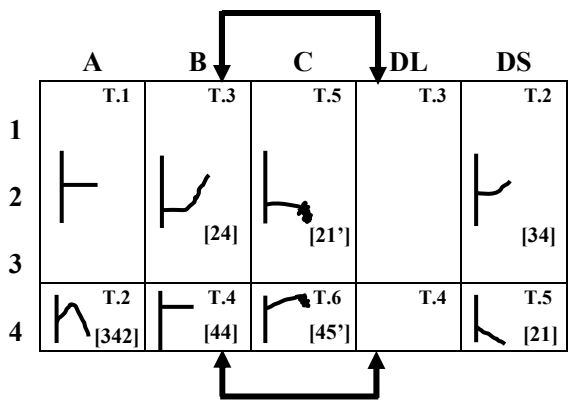
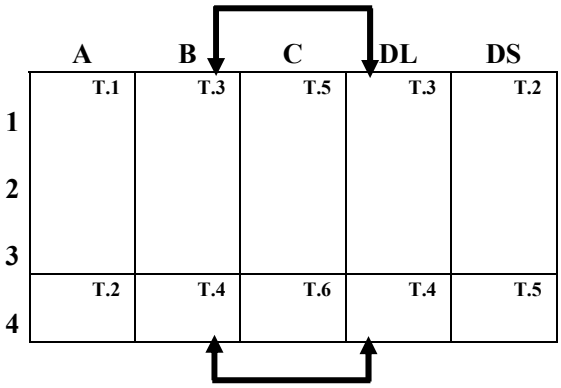
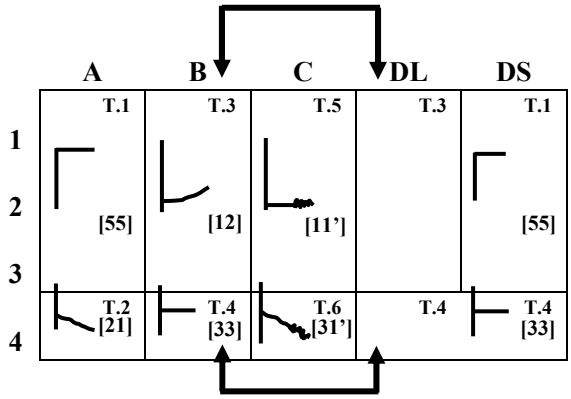


Figure 11 The patterns of tonal merger and split in Lue dialects

3. **Lue**: Ceng Tong, Mueang Sing, Mueang Long, Ou Neua, Kanlampa, Sop Tiek, Houei Lao (Chamberlain 1975: 55)[35]



4. **Lue**: Chieng Hung (from Gedney's fieldnotes in Hudak (ed.), 1994)[36]



5. **Lue**: Muong Yong (from Gedney's fieldnotes in Hudak (ed.), 1994)

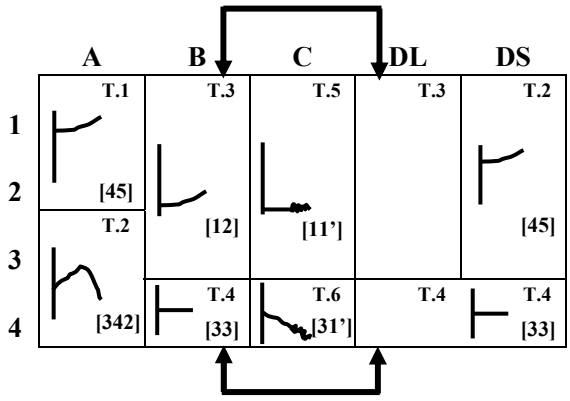


Figure 11 (continued)

Considering the pattern of tonal mergers and splits of Column A in Lue dialects in the available previous studies as shown in Figure 11, Lue dialects can be classified into two subgroups as follows:

1) **Group I (A123-4):** Lue Yo Village (Mueang Bang, Sipsong Panna), Lue Na Yang Tai Village (Nam Bak, Luang Prabang) (from L-Thongkum's fieldnotes 1982), Lue Ceng Tong, Muang Sing, Muang Long, Ou Neua, Kanlampa, Sop Tiek, Houei Lao (Chamberlain1975), and Lue Chieng Hung (from Gedney's fieldnotes in Hudak (ed.), 1994).
2) **Group II (A12-34):** Lue Muong Yong (from Gedney's fieldnotes in Chamberlain 1975).

As shown in Figure 11, Lue (Nong Bua) has the same pattern (A123-4) as Lue dialects in Group I and Lue (Donchai) has the same pattern (A12-34) as Group II. Thus, it can be

[35] The tonal characteristics are not provided in Chamberlain (1975).

[36] Lue, or Tai Lue, belongs to the Southwestern branch of the Tai language family, and Lue speakers can be found in a number of areas: 1) Chiang Hung, the capital city of Sipsong Panna, and most of southern Yunnan province; 2) in Muong Yong and vacinity in northeastern Myanmar; 3) in Chiangrai and Chiang Mai provinces in northern Thailand; 4) in Muong Sing and Luang Prabang in Laos; and 5) in northwest Vietnam around Binh Lue. Gedney began his studies of Lue on 27 November 1964 and pursued them off and on over the next fifteen years, synthesizing much of his data in April, May, and June, 1969. All of Gedney's Lue material presented in ***'William J. Gedney's the Lue Language: Glossary, Texts, and Translations'*** (edited by Thomas John Hudak) was provided by speakers from two localities: 1) Chiang Hung, where Lue is the dialect of the capital city and the ruling family, and 2) Muong Yong, to the west, across the border in northeastern Myanmar. (adapted from Hudak (ed.) 1994: xiii)

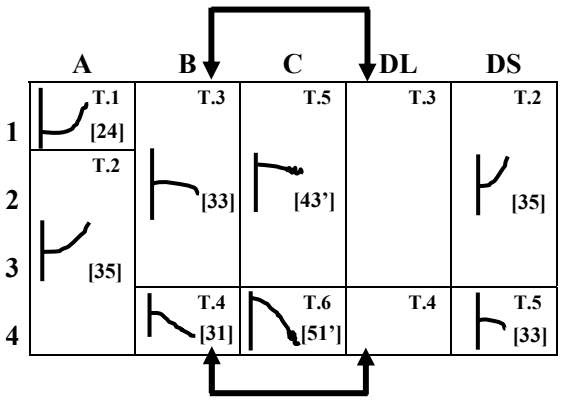
hypothesized that Lue people in Nong Bua village and in Donchai village might belong to different groups of Lue and migrated from different places. The patterns of tonal merger and split can be used as an evidence to confirm the history of migration of Lue (Nong Bua), that their ancestors migrated from Sipsong Panna as they insisted. As for Lue (Donchai), its pattern of tonal mergers and splits is the same as Lue Muong Yong (from Gedney's fieldnotes in Chamberlain 1975), therefore it can be hypothesized that they are the same group as Lue (Muong Yong) and their ancestors might probably migrated from Muong Yong.

Even though the same pattern of tonal mergers and splits between Lue (Donchai) and Lue (Muong Yong) can lead to the hypothesis that they may be the same Lue dialects or perhaps the ancestors of Lue (Donchai) had migrated from Muong Yong, there may be another possible hypothesis. It is also possible that Lue (Donchai) has been in contact with Nyuan (Mueang Nan) for such a long time that the tones of Nyuan influenced the tones in Lue (Donchai). In other words, the pattern A12-34 in Lue (Donchai) may be influenced by the pattern A12-34 in Nyuan (Mueang Nan). Further studies are needed for clear explanations and/or the exact conclusion related to this issue.

5.3.3 The tone variation of Phuan

Similar to the case of the consonant and vowel variation, there is also the tone variation in Phuan (Fay Mun). The tone data of Phuan in this study were collected from two main LRPs, an 85-year-old woman and a 68-year-old man. These two LRPs insisted that they spoke Phuan as their mother-tongue but their tonal systems do not look exactly the same as each other as shown in Figure 12.

1. **Phuan** (Fay Mun): 85-year-old LRP



2. **Phuan** (Fay Mun): 68-year-old LRP

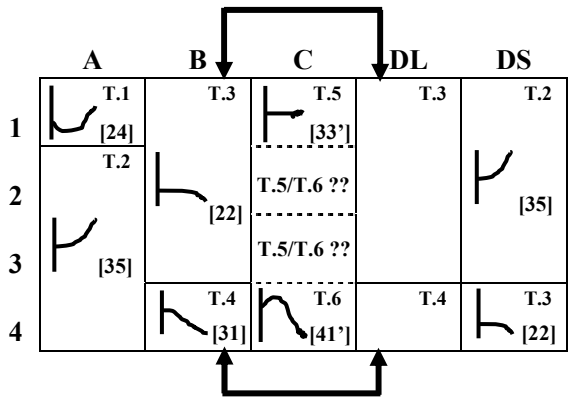


Figure 12 The comparison of tonal systems in Phuan (Fay Mun) between two LRPs

The 68-year-old LRP pronounced some lexical items in C2 and C3 with two variants of the tones, i.e. some with Tone 5 [33'] and some with Tone 6 [41']. Thus, the pattern of tonal mergers and splits in Column 4 of the 68-year-old LRP could not be determined. Considering the usage of these two variants in percentage, it is found that the glottalized mid level tone [33'] is used less than the glottalized high-falling tone [41'] as shown in Figure 13.

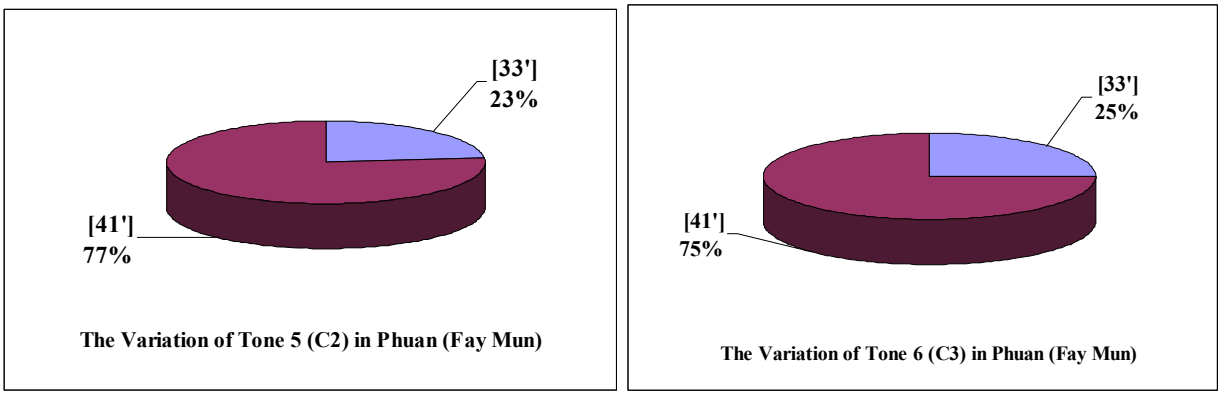


Figure 13 The percentage of the tone variation in Phuan (Fay Mun)

The examples of the lexical items in Phuan (Fay Mun) in Columns C2 and C3, pronounced with two different variants, glottalized high-falling tone [41'] and glottalized mid level tone [33'], are shown in Table 32.

Table 32 The tone variation in Phuan (Fay Mun)

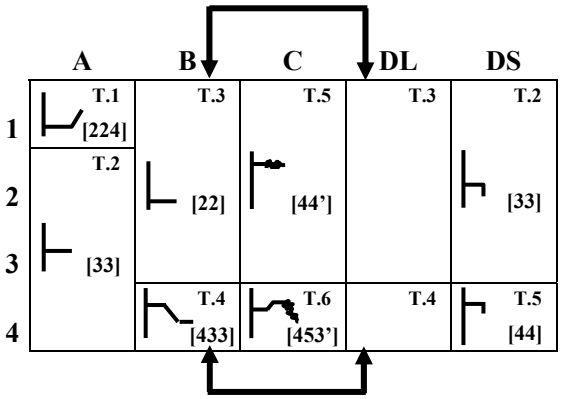
Tones C2 and C3 in Phuan (Fay Mun)			
no variation		**variation**	
Gloss	**Tone C2**	**Gloss**	**Tone C2**
'buttocks'	kon$^{41'}$	*'nine'*	kaw$^{33'}$ (← kaw$^{41'}$)
'fish bone'	kaaŋ$^{41'}$	*'step (clf.)'*	kaaw$^{33'}$ (← kaaw$^{41'}$)
'to boil'	tom$^{41'}$	*'cupboard, cabinet, case'*	tuu$^{33'}$ (← tuu$^{41'}$)
'to set up'	taŋ$^{41'}$	*'to tickle'*	cii$^{33'}$ (← cii$^{41'}$)
'elder aunt'	paa$^{41'}$	*'clear, unobscured'*	cɛɛŋ$^{33'}$ (← cɛɛŋ$^{41'}$)
'to mold (e.g. clay)'	pan$^{41'}$	*'to hold close to'*	cam$^{33'}$ (← cam$^{41'}$)
'dirty'	piən$^{41'}$		
'flour'	pɛɛŋ$^{41'}$		
'target'	paw$^{41'}$		
Gloss	**Tone C3**	**Gloss**	**Tone C3**
'handle, holder'	daam$^{41'}$	*'thread'*	daay$^{33'}$ (← daay$^{33'}$)
'to wriggle'	diin$^{41'}$	*'to speak, to talk, to say'*	ʔuu$^{33'}$ (← ʔuu$^{41'}$)
'to get, to obtain, to receive'	day$^{41'}$	*'to scoop up'*	sɔɔn$^{33'}$ (← sɔɔn$^{41'}$)
'hairy caterpillar'	buŋ$^{41'}$	*'bamboo rat'*	ʔon$^{33'}$ (← ʔon$^{41'}$)
'cowrie shell'	biə$^{41'}$	*'wild orchid'*	ʔiəŋ$^{33'}$ (← ʔiəŋ$^{41'}$)
'sugar cane'	ʔɔɔy$^{41'}$		
'giant reed'	ʔɔɔ$^{41'}$		
'to carry (a child, an animal)'	ʔum$^{41'}$		

As for the tonal system of the 85-year-old LRP, it is clearly seen in Figure 12 that there are six tones the same as the tonal system of the 68-year-old but the pattern of merger and split in Column C is different. While the pattern of merger and split in Column C of the 85-year-old LRP is C123-4, there are three possible patterns of tonal merger and split in Column C of the 68-year-old: C123-4, C1-234, and C12-34. The first two patterns seem to be more normal than

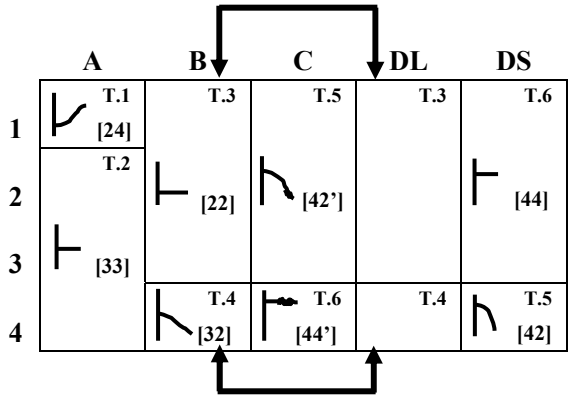
the third one. Even though there seems to be no variation in the tonal system of the 85-year-old Phuan LRP as far as the complete pattern of merger and split can be determined, it is suspected that this system has already been completely changed. In other words, it is possible that the tonal system of the 85-year-old LRP has already passed the stage of variation. The other possible hypothesis is that the tonal system of this LRP has not changed and the pattern found in this study is original. To find clearer explanation, it is worthwhile to investigate the tonal systems of Phuan in previous studies to compare with the ones in this study.

Based on the available previous studies, the tonal systems of Phuan spoken in the different areas are shown in Figure 14.

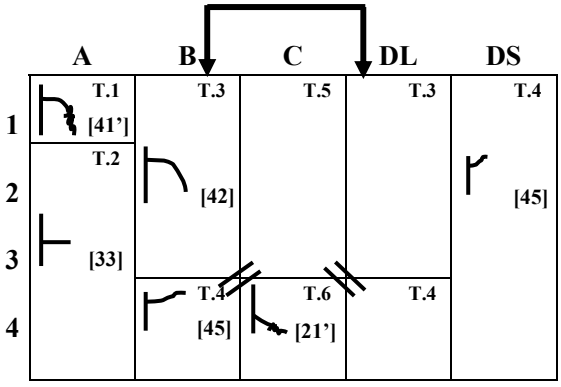
1. **Phuan**: Ban Mi District, Lopburi Province (adapted from Brown 1965)



2. **Phuan**: Saraburi, Singburi, Lopburi Provinces (Khanittanan 1973)



3. **Phuan**: Pak Seng (Chamberlain 1975)



4. **Phuan**: Chiang Khwang (Chamberlain 1971)

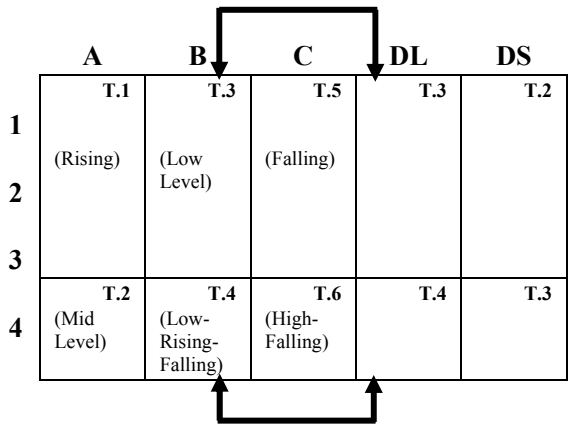


Figure 14 The tonal systems of Phuan from some earlier literature

To compare the pattern of tonal merger and split in Phuan (Fay Mun) in Figure 12 and the other Phuan dialects (in Figure 14), it can be clearly seen that Phuan (Fay Mun) has the same main pattern of tonal merger and split, i.e. B123=DL123, B4=DL4, as Phuan Ban Mi District, Lopburi Province in Brown (1965); Phuan Saraburi, Singburi, Lopburi Provinces in Khanittanan (1973); and Phuan Chiang Khwang in Chamberlain (1971). From this evidence, it can be hypothesized that Phuan (Fay Mun) and the aforementioned Phuan dialects probably belong to the same group or perhaps migrated from the same area. As for Phuan (Pak Seng) in Chamberlain (1975), it has the pattern of tonal merger and split, B4=C123=DL4 which is different from Phuan (Fay Mun), therefore, it may be hypothesized that these two Phuan dialects may be the different group or migrated from different areas.

5.3.4 The tonal system of Lao (Nam Pua): An evidence showing the real history of migration

Based on Brown (1965), the Lao language can be classified into three subgroups: Luang Prabang, Vientiane, and Nyo. Each subgroup of the Lao language is composed of different dialects of Lao as shown in Figure 15.

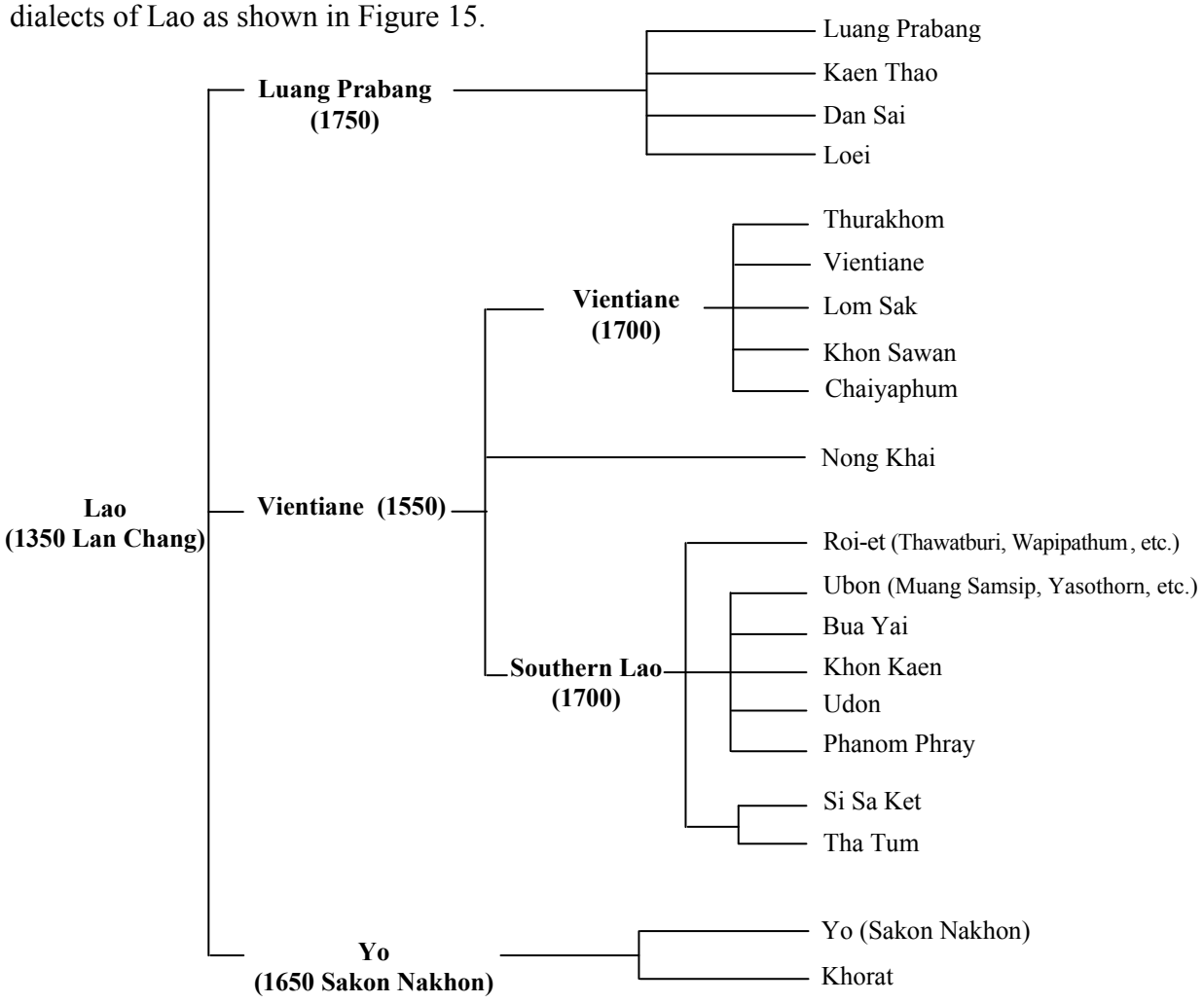


Figure 15 The classification of the Lao language (adapted from Brown 1965: 144)

As far as the patterns of tonal merger and split are concerned, the distinctive pattern of all three Lao groups is C1=DL123 and C234=DL4. As for the patterns which can be used as the criteria to distinguish the Luang Prabang group from Vientiane group, they are mainly in Column A. The distinctive pattern of the Luang Prabang group is A1-234 whereas in the Vientiane group, it is A1-23-4, A123-4, and A1-234 (this pattern appears in Si Sa Ket and Tha Tum dialects only).

On the basis of tonal development, the Lao language group studied in the author's dissertation (Akharawatthanakun 2004 unpublished)[37], can be classified into two main groups: Central Lao-Southern Lao group and Northern Lao group. Each group has distinctive pattern of tonal merger and split as shown in Figure 16.

[37] A part of the dissertation, "*Tone Change: A Case Study of the Lao Language Group"* was published in Akharawatthanakun. 2004. "*Tonal variation and change in dialects in contact: A case study of Lao.*" **Manusya** 7.1 (March 2004).

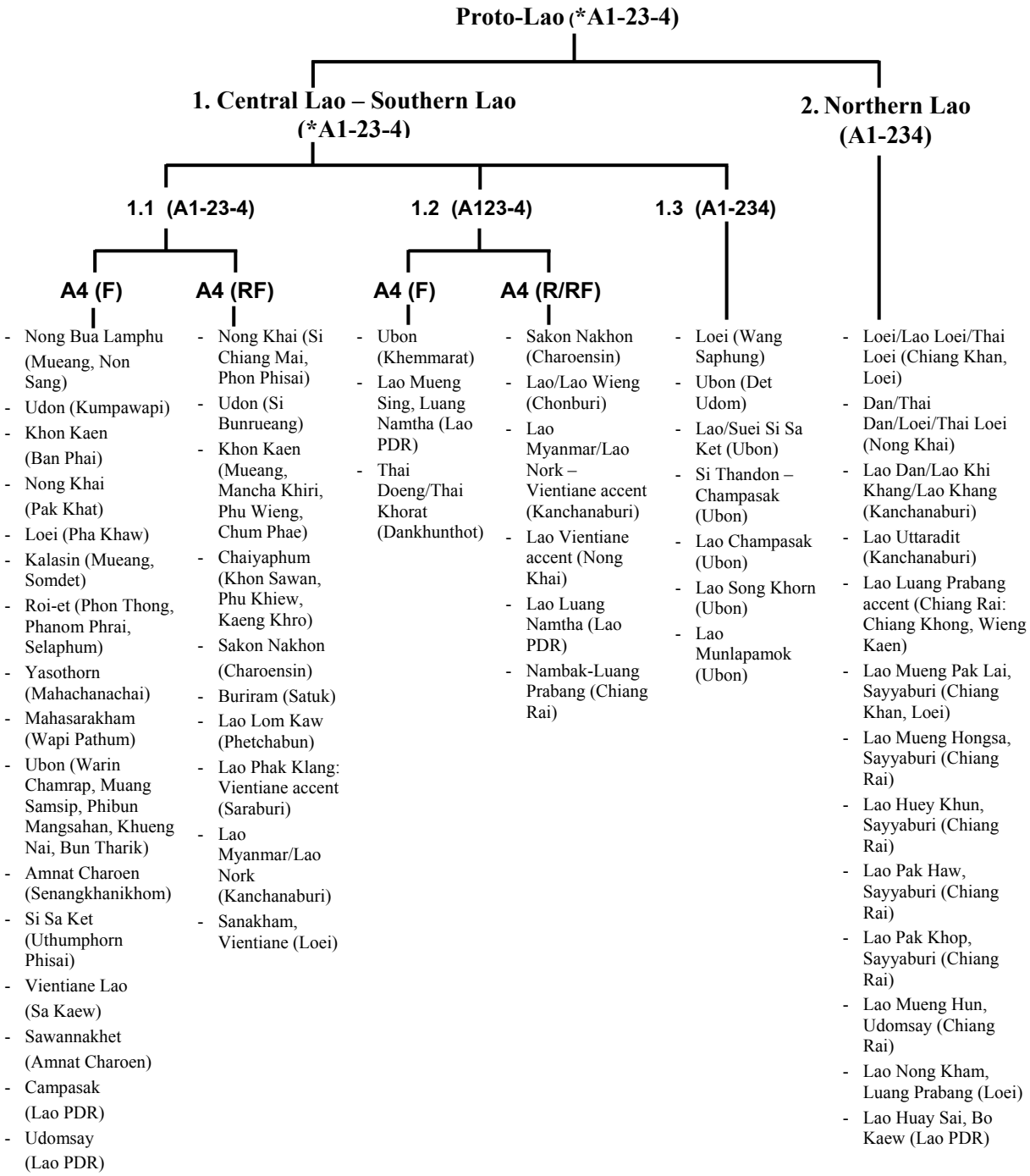


Figure 16 The classification of the Lao language group (adapted from Akharawatthanakun 2004: 332)[38]

The main patterns of tonal merger and split shared between among the two main groups of Lao language are B1234, C1=DL123, and C234=DL4. The distinction of the tonal development between them is the pattern of tonal merger and split in Column A, i.e. *A1-23-4 in the Central Lao-Southern Lao group and *A1-234 in the Northern Lao group. As shown in Figure 16, most of the Lao Vientiane dialects have A1-23-4 and A123-4 patterns while the Lao Luang Prabang dialects have A1-234 pattern the same as Lao Luang Prabang in Brown (1965).

[38] The name in the parentheses, following the name of Lao dialect refers to the location where each Lao dialect is spoken, for example, Vientiane Lao (Sa Kaew) = Vientiane Lao dialect, spoken in Sa Kaew Province (Thailand); Lao Nong Kham, Luang Prabang (Loei) = Lao Nong Kham, Luang Prabang dialect, spoken in Loei Province (Thailand); Udomsay (Lao PDR) = Udomsay Lao dialect, spoken in Lao PDR, and so on.

Therefore, it can be said that the distinctive pattern of tonal merger and split of the Lao Luang Prabang group is A1-234 while in the Lao Vientiane group it is A1-23-4 and A123-4.

While the author was eliciting the lexical data in Nam Pua village, it was noticeable that most of the Lao people that the author interviewed had been confused with their own history of migration. Based on the information from the field work at Nam Pua village, some Lao people in the village believed that their ancestors migrated from Vientiane (Lao PDR) and they speak Lao Vientiane while some said that they were from Luang Prabang. According to the book of the village history recorded by the oldest man in the village from his own memory about what his ancestors told him, the homeland of Lao people in Nam Pua village was in Vientiane and their spoken language was called "Lao Vientiane".

In order to help clarify which group that Lao Nam Pua should be classified into and which accent of Lao that the people speak, the tone data in Lao Nam Pua was also elicited on the basis of using the tones as the criteria for classifying the language group. The pattern of tonal merger and split of Lao Nam Pua can be shown in comparison to Lao Vientiane and Lao Luang Prabang from the previous studies in Figure 17.

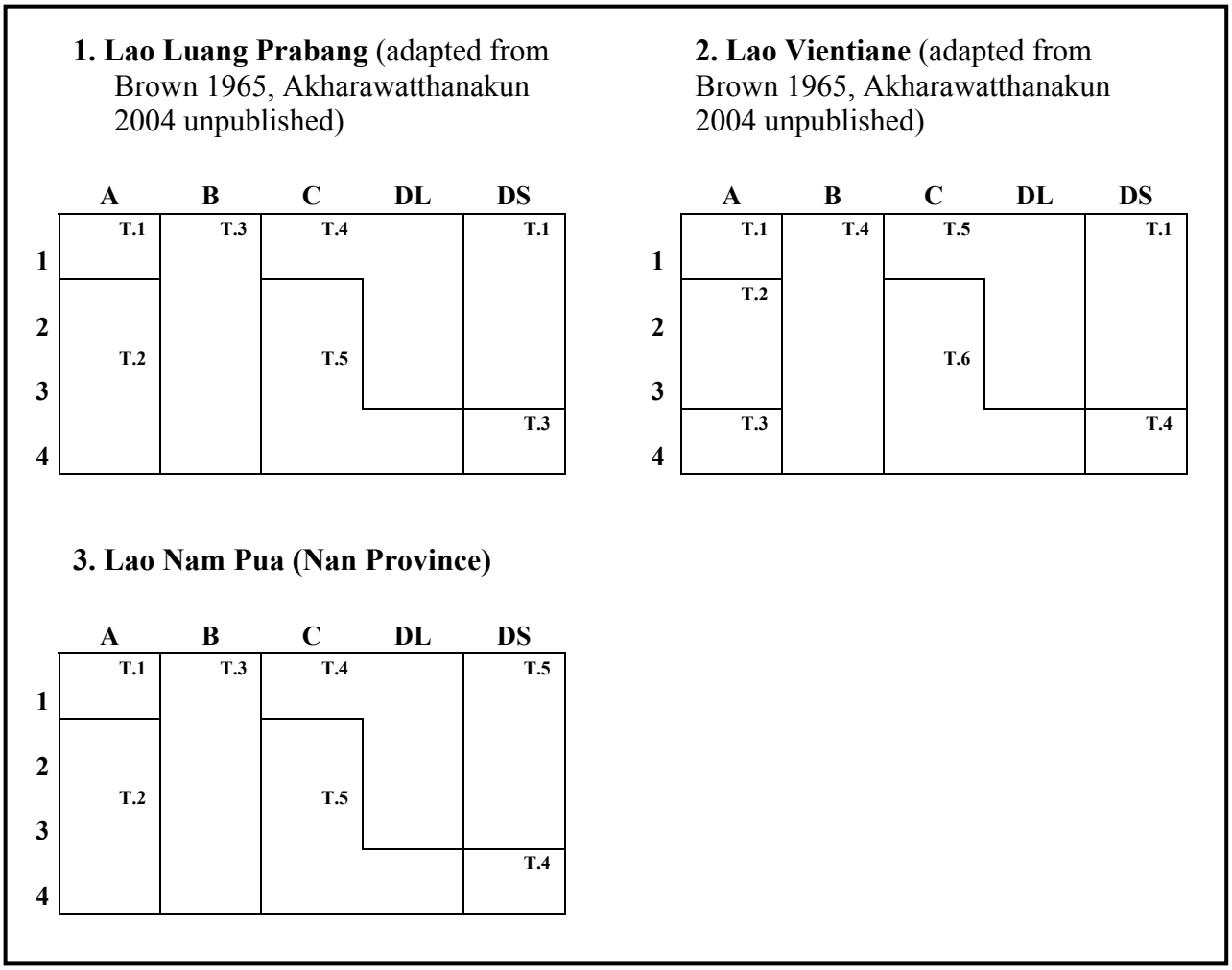


Figure 17 The patterns of tonal merger and split in Lao Nam Pua in comparison to Lao Luang Prabang and Lao Vientiane

If the pattern A1-234 is used as the criteria, it can be said that Lao Nam Pua is in the Luang Prabang group, in other words, it is a Lao Luang Prabang accent, not Vientiane. Even though the pattern of tonal merger and split can be used as an evidence to clarify the history of migration and the classification of a language, more data on tones and the other aspects, for example, consonants, vowels, and lexical items need to be further studied.

6. CONCLUSION AND DISCUSSION

Having analyzed the consonants, vowels, and tones in five Tai dialects: Nyuan (Mueng Nan), Lue (Nong Bua), Lue (Donchai), Khün (Nong Muang), Phuan (Fay Mun), and Lao (Nam Pua), it is found that variation and change occur in all of the phonological aspects, i.e. consonants, vowels, and tones. One of the main mechanism of language contact which induces language variation and change is the **phonic interference**, for example, the variation of consonants and vowels /ph-, th-, kh-, s-, h-, uə, iə/ in Phuan and the vowel /u/ in Lue (Donchai) which is considered to be mainly influenced by Nyuan. The other case of phonic interference is found in the tonal aspect, i.e. the tonal variation in Phuan (Fay Mun) which is also influenced by Nyuan. The other suspicious issue is that of the pattern of tonal merger and split: A12-34 in Lue (Donchai) which is exactly the same as Nyuan. Even though such pattern is also found in Lue (Chieng Hung, Mueng Yong) in Chamberlain (1975) and it can be hypothesized that they may probably be the same group, more data need to be further studied in order to find a conclusion.

If comparing the degree of language contact, it can be said that Phuan (Fay Mun), Khün (Nong Muang), Lao (Nam Pua), and Lue (Donchai) seem to have higher degree of language variation and change due to language contact than Lue (Nong Bua) and Nyuan (Mueang Nan). The main dialect and language which induce such variation and change in the Tai dialects are Nyuan (Mueang Nan) and Standard Thai. However, Nyuan (Mueang Nan), which has been the majority language and used as the lingua franca among the people who speak different dialects/languages, also has variation and change, influenced by Standard Thai. The direction of language contact among the five Tai dialects can be hypothesized as shown in Figure 18.

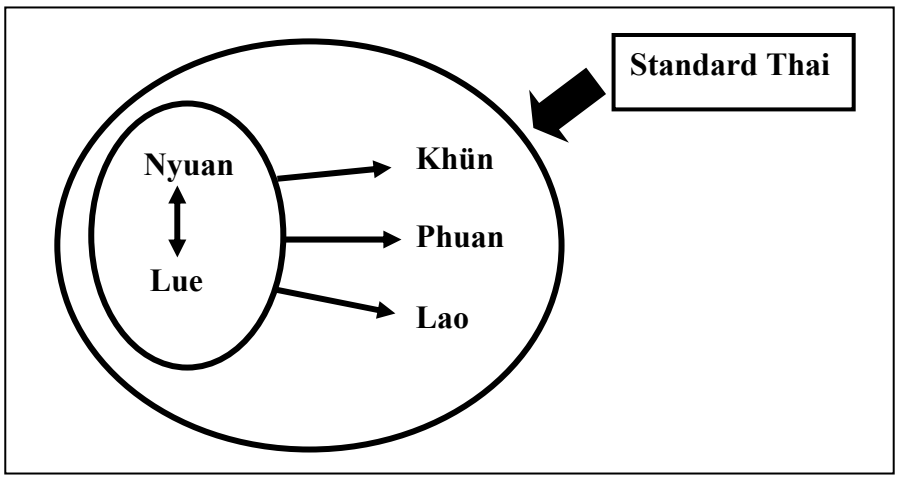


Figure 18 The hypothesized direction of language contact among Tai dialects

As can be seen in Figure 18, Nyuan and Lue are considered to influence each other and both of them influence Khün, Phuan, and Lao. All of the Tai dialects have been influenced by Standard Thai.

Since the findings in this study are the indirect outcome from the author's observation during the lexical elicitation for the purpose of producing a dictionary, the explanation related to some aspects of phonological variation and change in each Tai dialect is not complete yet. However, the variation and change which is found to appear in the five Tai dialects can lead to further study in some areas. The other interesting issue which is worthwhile to investigate is the intensity of language contact and the outcome of language contact in different aspects of each Tai dialect.

7. REFERENCES

Akharawatthanakun, Phinnarat. 2002. *"Tonal variations and changes in a language mixture area: A case study of Northeastern Thailand (Isan).*" **Manusya** 5.2 (September 2002): pp. 30-51.

Akharawatthanakun, Phinnarat. 2004. *Tone Change: A Case Study of the Lao Language Group*. Unpublished Ph.D. dissertation. Department of Linguistics, Chulalongkorn University.

Akharawatthanakun, Phinnarat. 2004. *"Tonal variation and change in dialects in contact: A case study of Lao.*" **Manusya** 7.1 (March 2004).

Brown, J.M. 1965. **From Ancient Thai to Modern Dialects**. Bangkok: Social Science Association Press.

Chamberlain, J.R. 1975. *"A new look at the history and classification of the Tai languages"*. In J.G. Harris and J.R. Chamberlain (eds.), **Studies in Tai Linguistics in Honor of William J. Gedney**, pp. 49-66. Bangkok: Central Institute of English Language, Office of State University.

Gedney, William J. 1972. "A Checklist for determining tones in Tai dialects". In M. Estellie Smith (ed.). **Studies in Linguistics in Honor of George L. Trager.** The Hague: Mouton, pp. 423-437.

Kullavanijaya, et al (1984). The Vocabulary of Six Tai Languages (ศัพท์ไท 6 ภาษา). Faculty of Arts, Chulalongkorn University (in Thai).

Li, Fang Kuei. 1960. *A tentative classification of Tai dialects*. In S. Diamond (ed.). **Culture in History: Essays in Honor of Paul Radin**, pp. 951-959. New York: Columbia University Press.

Li, Fang Kuei. 1977. **A Handbook of Comparative Tai**. Honolulu: The University Press of Hawaii.

Sarawit, Mary E. 1973. **The Proto-Tai Vowel System**. Doctoral Dissertation, The University of Michigan.

Thomason, Sarah G. 2001. **Language Contact: An Introduction**. Washington D.C.: Georgetown University Press.

Thomason, Sarah G. and Kaufman, Terrence. 1988. **Language Contact, Creolization, and Genetic Linguistics**. Berkeley and Los Angeles: University of California Press.

Weinreich, Uriel. 1953. **Languages in Contact: Findings and Problems**. The Netherlands: Mouton & Co.

8. APPENDICES

Appendix 1: The Tone Box (adapted from Gedney, William J. 1972: 434)

		Proto-Tai Tones				
		A	**B** (◌́)	**C** (◌̌)	**DL (D-long)**	**DS (D-short)**
1.	**Voiceless friction sounds:** *s-, *hm-, *hn-, *ph-, *th-, *kh-	1	5	9	13	17
2.	**Voiceless unaspirated stops:** *p-, *t-, *k-	2	6	10	14	18
3.	**Glottal sounds:** *ʔ-, *ʔb-, *ʔd-	3	7	11	15	19
4.	**Voiced sounds:** *b-, *d-, *g-, *m-, *n-, *l-, *r-	4	8	12	16	20
			Smooth Syllable (Live Syllable)		**Checked Syllable (Dead Syllable)**	

The tone checklist (Gedney 1972: 434)

Column A:

Box 1.	*huu*	‘ear’
	khaa	‘leg’
	hua	‘head’
Box 2.	*pii*	‘year’
	taa	‘eye’
	kin	‘to eat’
Box 3.	*bin*	‘to fly’
	dɛɛŋ	‘red’
	daaw	‘star’
Box 4.	*mɨɨ*	‘hand’
	khwaay	‘water buffalo’
	naa	‘ricefield’

Column B:

Box 5.	*khay*	‘egg’
	phaa	‘to split’
	khaw	‘knee’
Box 6.	*paa*	‘forest’
	kay	‘chicken’
	kɛɛ	‘old’
Box 7.	*baa*	‘shoulder’
	baaw	‘young man’
	daa	‘to scold’
Box 8.	*phii*	‘older sibling’
	phɔɔ	‘father’
	ray	‘dry field’

Column C:

Box 9.	*khaw*	‘rice’
	sɨa	‘shirt’
	***khaa*	‘to kill’
	khay	‘fever’
	haa	‘five’
Box 10.	*paa*	‘aunt (older sister of either parent’
	klaa	‘rice seedlings’
	tom	‘to boil’
Box 11.	*baa*	‘crazy’
	baan	‘village’
	ʔaa	‘to open (the mouth)’
Box 12.	*nam*	‘water’
	nɔɔŋ	‘younger sibling’
	may	‘wood’
	maa	‘horse’

Column D-long (DL):

Box 13.	*khaat*	‘broken, torn’
	ŋɨak	‘the gums’
	haap	‘to carry on a shoulder pole’
Box 14.	*pɔɔt*	‘the lungs’
	piik	‘wing’
	tɔɔk	‘to pound’
Box 15.	*dɛɛt*	‘sunshine’
	ʔaap	‘to bathe’
	dɔɔk	‘flower’
Box 16.	*miit*	‘khife’
	luuk	‘(one’s) child’
	lɨak	‘blood’
	nɔɔk	‘outside’

Column D-short (DS):

Box 17.	*mat*	‘flea’
	suk	‘cooked, ripe’
	phak	‘vegetable’
Box 18.	*kop*	‘frog’
	tap	‘liver’
	cep	‘to hurt’
Box 19.	*bet*	‘fishhook’
	dip	‘raw, unripe’
	ʔok	‘the chest’
Box 20.	*nok*	‘bird’
	mat	‘to tie up’
	lak	‘to steal’

Appendix 2: Test Words for the Tones in Tai dialects (adopted from Akharawatthanakun 2004)

Set 1: Test words modified from Gedney's wordlist

	A	B	C	DL	DS
1	หู [hŭu] 'ears' ขา [khăa] 'legs' หัว [hŭa] 'head' สอง [sŏɔŋ] 'two'	ไข่ [khày] 'egg' ผ่า [phàa] 'to cut' ข่า [khàa] 'ginger' ถั่ว [thùa] 'bean'	ข้าว [khâaw] 'rice' เสื้อ [sîa] 'shirt' หม้อ [mɔ̂ɔ] 'pot' ห้า [hâa] 'five'	ขวด [khùat'bottle' หมวก [mùak] 'hat' ศอก [sɔ̀ɔk] 'elbows' สาก [sàak] 'pestle'	หมัด [màt] 'flea' ผัก [phàk] 'vegetable' หก [hòk] 'six' สิบ [sìp] 'ten'
2	ปู [pūu] 'crab' ตา [tāa] 'eyes' กิน [kīn] 'to eat' กา [kāa] 'crow'	ไก่ [kày] 'chicken' เป่า [pàw] 'to blow' เต่า [tàw] 'turtle' ปี่ [pìi] 'flute'	เก้า [kâaw] 'nine' ป้อน [pɔ̂ɔn] 'to feed' ก้าง [kâaŋ] 'fish-bone' ตู้ [tûu] 'wardrobe'	ปีก [pìik] 'wing' กอด [kɔ̀ɔt] 'to hug' ตอก [tɔ̀ɔk] 'to hammer' ปาก [pàak] 'mouth'	กบ [kòp] 'frog' เจ็ด [cèt] 'seven' เตะ [tè] 'to kick' จับ [càp] 'to touch'
3	บิน [bīn] 'to fly' แดง [dǣæŋ] 'red' บาน [bāan] 'to blossom' ดาว [dāaw] 'star'	บ่า [bàa] 'shoulders' อึ่ง [ʔɨŋ] 'bull frog' บ่าว [bàaw] 'single man' ด่า [dàa] 'to scold'	ด้าย [dâay] 'thread' อ้อย [ʔɔ̂ɔy] 'sugar cane' ด้าม [dâam] 'handle' อ้า [ʔâa] 'to open mouth'	แดด [dæ̀æt] 'sunshine' บีบ [bìip] 'to squeeze' บอด [bɔ̀ɔt] 'to be blind' ดีด [điit] 'to flick'	เบ็ด [bèt] 'fish hook' ดิบ [đip] 'to be raw' อก [ʔòk] 'chest' เด็ด [dèt] 'to pick'
4	มือ [mɯ̄ɯ] 'hands' ควาย [khwāay] 'buffalo' นา [nāa] 'rice field' งู [ŋūu] 'snake'	พ่อ [phɔ̂ɔ] 'father' ว่าว [wâaw] 'kite' เลื่อย [lɯ̂ay] 'saw' ม่าย [mâay] 'widow'	น้า [náa] 'aunt' คิ้ว [khíw] 'eyebrows' น้ำ [náam] 'water' ลิ้น [lín] 'tongue'	มีด [mîit] 'knife' เลือด [lɯ̂at] 'blood' เชือก [chɯ̂ak] 'rope' มอด [mɔ̂ɔt] 'woodmite'	นก [nók] 'bird' มด [mót] 'ant' ครก [khrók] 'mortar' เล็บ [lép] 'nail'

Set 2: The analogous set

	A	B	C	DL	DS
1	ขา [khăa] 'legs'	ข่า [khàa] 'ginger'	ข้าว [khâaw] 'rice'	ขาด [khàat] 'to be torn'	ขัด [khàt] 'to brush'
2	ปา [pāa] 'to throw'	ป่า [pàa] 'forest'	ป้า [pâa] 'aunt'	ปาด [pàat] 'to cut'	ปัด [pàt] 'to wipe'
3	บาน [bāan] 'to blossom'	บ่า [bàa] 'shoulders'	บ้า [bâa] 'to be mad'	บาด [bàat] 'to be cut'	บัตร [bàt] 'card'
4	คา [khāa] 'thatch grass'	ค่า [khâa] 'price'	ค้า [kháa] 'to trade'	คาด [khâat] 'to tie a belt'	คัด [khát] 'to handwrite'

Test words randomized from the analogous set.

1. ขา [khăa]	13. ป่า [pàa]	25. ค้า [kháa]	37. ป่า [pàa]	49. ป่า [pàa]
2. ปา [pāa]	14. บาน [bāan]	26. ค้า [kháa]	38. ขา [khăa]	50. บาน [bāan]
3. บาน [bāan]	15. ขา [khăa]	27. บ้า [bâa]	39. ป่า [pàa]	51. บ่า [bàa]
4. คา [khāa]	16. บ่า [bàa]	28. ป่า [pàa]	40. บ่า [bàa]	52. ป่า [pàa]
5. ข่า [khàa]	17. ข้าว [khâaw]	29. ข้าว [khâaw]	41. ข่า [khàa]	53. ข่า [khàa]
6. ป่า [pàa]	18. คา [khāa]	30. คา [khāa]	42. ค้า [kháa]	54. ขา [khăa]
7. บ่า [bàa]	19. ค้า [kháa]	31. บาน [bāan]	43. บ้า [bâa]	55. ปา [pāa]
8. ค่า [khâa]	20. ป่า [pàa]	32. ข่า [khàa]	44. ค่า [khâa]	56. ข้าว [khâaw]
9. ข้าว [khâaw]	21. ปา [pāa]	33. ค่า [khâa]	45. ปา [pāa]	57. บ้า [bâa]
10. ป้า [pâa]	22. ข่า [khàa]	34. ขา [khăa]	46. คา [khāa]	58. คา [khāa]
11. บ้า [bâa]	23. บ้า [bâa]	35. ป่า [pàa]	47. ข้าว [khâaw]	59. ค่า [khâa]
12. ค้า [kháa]	24. ค่า [khâa]	36. ปา [pāa]	48. บาน [bāan]	60. ค้า [kháa]

1. ขาด [khàat]	5. ปาด [pàat]	9. คาด [khâat]	13. ขาด [khàat]	17. คาด [khâat]
2. ปาด [pàat]	6. ขาด [khàat]	10. บาด [bàat]	14. ปาด [pàat]	18. ปาด [pàat]
3. คาด [khâat]	7. คาด [khâat]	11. ขาด [khàat]	15. คาด [khâat]	19. ขาด [khàat]
4. บาด [bàat]	8. บาด [bàat]	12. ปาด [pàat]	16. บาด [bàat]	20. บาด [bàat]

21. ขัด [khàt]	25. ขัด [khàt]	29. ขัด [khàt]	33. ขัด [khàt]	37. บัตร [bàt]
22. ปัด [pàt]	26. คัด [khát]	30. ปัด [pàt]	34. บัตร [bàt]	38. ขัด [khàt]
23. คัด [khát]	27. ปัด [pàt]	31. คัด [khát]	35. คัด [khát]	39. ปัด [pàt]
24. บัตร [bàt]	28. บัตร [bàt]	32. บัตร [bàt]	36. ปัด [pàt]	40. คัด [khát]