



รายงานวิจัยฉบับสมบูรณ์

โครงการ “ภาษาศาสตร์ภาษากะเหรี่ยง”

โดย

ศาสตราจารย์ ดร. วีระพันธ์ เหลืองทองคำ และคณะ

31 มกราคม 2556

รายงานวิจัยฉบับสมบูรณ์
โครงการ “ภาษาศาสตร์ภาษากะเหรี่ยง”
(2552-2555)

คณะผู้วิจัย

หัวหน้าโครงการ

ศาสตราจารย์ ดร. ชีระพันธ์ เหลืองทองคำ

ผู้วิจัย

- | | |
|--|-----------------------|
| 1. ศาสตราจารย์ ดร. ชีระพันธ์ เหลืองทองคำ | จุฬาลงกรณ์มหาวิทยาลัย |
| 2. ผู้ช่วยศาสตราจารย์ ดร. ศิริวิไล ชีระโรจนรัตน์ | จุฬาลงกรณ์มหาวิทยาลัย |
| 3. ผู้ช่วยศาสตราจารย์ ดร. ผณิตรา ชีรานนท์ | มหาวิทยาลัยแม่ฟ้าหลวง |
| 4. อาจารย์ ดร. ชมนาด อินทจามรรักษ์ | มหาวิทยาลัยนเรศวร |
| 5. อาจารย์ ดร. พิทยาวัฒน์ พิทยาภรณ์ | จุฬาลงกรณ์มหาวิทยาลัย |
| 6. อาจารย์ ดร. วิชาติ บุรณะประเสริฐสุข | มหาวิทยาลัยธรรมศาสตร์ |

ผู้ช่วยวิจัย

- | | |
|--|-----------------------|
| 1. François Langella | จุฬาลงกรณ์มหาวิทยาลัย |
| 2. ศุจิณัฐ จิตวิริยพันธ์ (นิสิตผู้รับทุน คปก.) | จุฬาลงกรณ์มหาวิทยาลัย |
| 3. กานต์ธิดา เกิดผล (นิสิตผู้รับทุน คปก.) | จุฬาลงกรณ์มหาวิทยาลัย |
| 4. นรินทร สมบัติพันธ์ แบร์ (นิสิตผู้รับทุน คปก.) | จุฬาลงกรณ์มหาวิทยาลัย |
| 5. ญาณินท์ สวณะคุณานนท์ (นิสิตผู้รับทุน คปก.) | จุฬาลงกรณ์มหาวิทยาลัย |
| 6. ศิวพร ทวนไธสง | จุฬาลงกรณ์มหาวิทยาลัย |

ผู้แปลและผู้ตรวจแก้ภาษา

- | | |
|--|-----------------------|
| 1. รองศาสตราจารย์ ดร. พจี ยวชิต | จุฬาลงกรณ์มหาวิทยาลัย |
| 2. Assistant Professor Simon J.P. Wright | จุฬาลงกรณ์มหาวิทยาลัย |
| 3. อาจารย์ ดร. เทพี จรัสจรวงเกียรติ | จุฬาลงกรณ์มหาวิทยาลัย |

สนับสนุนโดยสำนักงานกองทุนสนับสนุนการวิจัย (สกว.)
(ความเห็นในรายงานนี้เป็นของผู้วิจัย สกว.ไม่จำเป็นต้องเห็นด้วยเสมอไป)

สารบัญ

บทคัดย่อ

Abstract

สรุปเนื้อหาของผลงาน

- ภาคผนวก 1 ต้นฉบับบทความภาษาอังกฤษ 10 บทความ
- ภาคผนวก 2 Abstracts และ PowerPoints ของบทความที่เสนอในการประชุมวิชาการ
นานาชาติ
- ภาคผนวก 3 เว็บไซต์ประชาสัมพันธ์โครงการ <http://ling.arts.chula.ac.th/Karen/?p=1>
- ภาคผนวก 4 นิทรรศการผลงานของโครงการในงานจุฬาฯ วิชาการครั้งที่ 14
- ภาคผนวก 5 ข้อมูลเกี่ยวกับ STEDT database
หลักฐานความร่วมมือกับต่างประเทศ

บทคัดย่อ

โครงการวิจัย “ภาษาศาสตร์ภาษากะเหรี่ยง” ประกอบด้วยโครงการย่อย 11 โครงการ ดังนี้ (1) สืบสร้างเสียงและศัพท์ในภาษากะเหรี่ยงดั้งเดิม (2) อิทธิพลเสียงพยัญชนะต้นที่มีต่อพฤติกรรมค่าความถี่มูลฐานของสระธรรมดาและสระนาสิกในภาษากะเหรี่ยงไปบ้านดอย จังหวัด เชียงราย (3) การแปรและการเปลี่ยนแปลงของสระและวรรณยุกต์ภาษากะเหรี่ยงไปบ้านคางใจ จังหวัดแพร่ อันเกิดจากการสัมผัสกับภาษาไทย (4) สมบัติความเป็นเสียง “ร” ของเสียงเสียดแทรกกลืนส่วนหลัง /x/ ในภาษากะเหรี่ยงสะกอบ้านสุดห้วยนา จังหวัดแม่ฮ่องสอน (5) ภาวะหลายหน้าที่ของหน่วยคำเติมหน้า ?à?- (TB *?a-) ในภาษากะเหรี่ยงไปบ้านดงดำ จังหวัด ลำพูน (6) วรรณยุกต์ภาษากะเหรี่ยงสะกอบ้านป่าละอู จังหวัดประจวบคีรีขันธ์: การวิเคราะห์ ด้วยวิธีการหลายมิติ (7) ตัวบ่งชี้ทางกลศาสตร์ที่สำคัญของการจำแนกฐานกรณ์พยัญชนะ นาสิกในภาษากะเหรี่ยงสะกอบ้านป่าละอู จังหวัดประจวบคีรีขันธ์ (8) การจัดทำแผนที่แสดงการ ตั้งถิ่นฐานของชาติพันธุ์กะเหรี่ยง 6 กลุ่มในประเทศไทยโดยใช้ระบบสารสนเทศภูมิศาสตร์ (9) การจัดทำหนังสือสารคดีประกอบภาพ “เรื่องเล่าจากชายแดนไทย-พม่า: ปรากฏไฟทางปัญญา เพื่อการพัฒนาอย่างยั่งยืน” (10) การจัดทำเว็บไซต์โครงการวิจัยภาษาศาสตร์ภาษากะเหรี่ยง และ (11) การจัดนิทรรศการแสดงผลงานของโครงการฯ เพื่อเผยแพร่สู่สาธารณะชน

รายงานผลการวิจัยของโครงการย่อยที่ 1-8 เป็นบทความภาษาอังกฤษเพื่อตีพิมพ์ใน วารสารวิชาการระดับนานาชาติ 10 บทความ และแผนที่สีขนาด 60 ซม. x 84 ซม. เป็น ภาษาไทย 1 แผ่นและภาษาอังกฤษ 1 แผ่น ผลงานของโครงการย่อยที่ 9 เป็นหนังสือสารคดี 2 ภาษา (ไทย-อังกฤษ) ประกอบภาพ 4 สี ผลงานของโครงการย่อยที่ 10-11 เป็นเว็บไซต์ 2 ภาษา (ไทย-อังกฤษ) และนิทรรศการ 2 ภาษา (ไทย-อังกฤษ) ซึ่งจัดในงานจุฬาวิชาการครั้งที่ 14 ระหว่างวันที่ 14-18 พฤศจิกายน 2555

ในเรื่องความร่วมมือกับต่างประเทศนั้น ได้นำเสนอผลงานวิจัยในการประชุมนานาชาติ ในต่างประเทศ 7 ครั้ง ส่วนหนึ่งของรายงานผลการวิจัยที่เป็นรูปสืบสร้างคำศัพท์ภาษากะเหรี่ยง ดั้งเดิมได้รับการบรรจุไว้ในฐานข้อมูล STEDT (Sino-Tibetan Etymological Dictionary and Thesaurus) ของภาควิชาภาษาศาสตร์ มหาวิทยาลัยแคลิฟอร์เนีย (เบอร์กิลีย์) ประเทศ สหรัฐอเมริกา นอกจากนี้ ศาสตราจารย์ ดร. ชีระพันธ์ เหลืองทองคำ และศาสตราจารย์ ดร.เจมส์ เอ แมทธิซอพฟ์ ซึ่งเป็นผู้อำนวยการโครงการฐานข้อมูล STEDT ได้เริ่มทำวิจัยร่วมกันเกี่ยวกับการศึกษาเปรียบเทียบภาษาโลโล-พม่าดั้งเดิมกับภาษากะเหรี่ยงดั้งเดิม เพื่อหาคำตอบเกี่ยวกับ

ความสัมพันธ์ด้านเชื้อสายของภาษา 2 สาขานี้ ซึ่งต่างก็เป็นสาขาของภาษาตระกูลทิเบต-พม่า
ในมหาตระกูลจีน-ทิเบต

Abstract

The Karen Linguistics Project consists of 11 sub-projects: (1) Reconstruction of Proto-Karen Phonology and Lexicon; (2) Oral and Nasal Vowels in the Pwo Karen Dialect of Chiangrai Province: Implications for Tonogenesis Theories; (3) Variation and Change in Phrae Pwo Karen Vowels and Tones Spoken in Ban Khang Chai Induced by Language Contact with Tai; (4) The Class of Rhotic Sounds (/r/ and /ɣ/) in Sgaw Karen Spoken in Ban Sut Huai Na, Mae Hong Son Province; (5) Multi-functions of the Prefix ʔàʔ- (TB *ʔa-) in the Dong Dam Pwo Karen Dialect of Lampun Province; (6) Tones in Sgaw Karen Spoken in Ban Pa La-U, Prachuap Khiri Khan Province: A Multi-dimensional Analysis; (7) Effective Cues to Differentiate the Places of Articulation of Nasals in the Ban Pa La-U Sgaw Karen Dialect of Prachuap Khiri Khan Province; (8) Production of Maps Showing the Settlements of the Six Karen Ethnic Groups in Thailand; (9) Production of an illustrated documentary on “Stories from the Thai-Burmese Border: Intellectual Sparks for Sustainable Development”; (10) Website to publicise the progress of the Karen Linguistics Project; and (11) Exhibition in the 14th Chula Academic Expo to publicise the output of the project.

The output of the sub-projects 1-8 is ten papers written in English to be published in international journals and two coloured maps (size 60 c.m. x 84 c.m.), one in Thai and the other in English. An illustrated documentary on “Stories from the Thai-Burmese Border: Intellectual Sparks for Sustainable Development”, in Thai and English, is the output of the sub-project 9. The aims of the sub-projects 10 and 11 are for social benefit; the output, therefore, consists of a bilingual website, publicising the progress and activities of the Karen Linguistics Project and an exhibition of nine coloured posters, in Thai and English, illustrating the project overview, the output and the social impact of the project displayed at the 14th Chula Academic Expo.

With regard to international co-operation, seven papers were presented at international conferences organised abroad. The reconstructions of Proto-Karen roots were added to the STEDT (Sino-Tibetan Etymological Dictionary and Thesaurus) database of the Linguistics Department at the University of California at Berkley, USA.

Professor James A. Matisoff, the director of the STEDT database project, invited Professor Theraphan Luangthongkum to conduct a co-operative research into the comparative study of Proto-Lolo-Burmese (reconstructed by him) and Proto-Karen in order to investigate the genetic relationship between the two branches, i.e. Lolo-Burmese and Karenic, of the Tibeto-Burman language family of the Sino-Tibetan phylum.

สรุปเนื้อหาของผลงาน

รายงานผลการวิจัยได้เขียนเป็นภาษาอังกฤษรวม 10 บทความ เพื่อตีพิมพ์เผยแพร่ระดับนานาชาติ บทความ 5 เรื่องอยู่ระหว่างขั้นตอนการตีพิมพ์ในวารสาร *Manusya*, Vol. 15.2 ซึ่งคาดว่าจะเสร็จสมบูรณ์ก่อนเดือนพฤษภาคม 2556 เนื่องจากจะมีการประชุมนานาชาติ SEALS 23 ระหว่าง 29-31 พฤษภาคม 2556 ที่กรุงเทพฯ บทความ 5 เรื่องดังกล่าว ได้แก่

Intajamornrak, Chommanad. (in press). Variation and change of the Phrae Pwo Karen vowels and tones induced by language contact with the Tai languages.

Manusya 15.2.

Jitwiriyant, Sujinat. (in press). Ban Pa La-U Sgaw Karen tones: An analysis of semitones, quadratic treadlines and coefficients. *Manusya* 15.2.

Kerdpol, Karnthida. (in press). Formant transitions as effective cues to differentiate the places of articulation of Ban Pa La-u Sgaw Karen nasals. *Manusya* 15.2.

Teeranon, Phanintra. (in press). Initial consonant voicing perturbation of the fundamental frequency of oral vowels and nasal vowels: A controversial case from Ban Doi Pwo Karen. *Manusya* 15.2.

Teerarojanarat, Sirivilai. (in press). Using GIS for exploring Karen settlements: A case of Western and Northern Thailand in the vicinity of the Thai-Burmese border.

Manusya 15.2.

บทความที่ตีพิมพ์เผยแพร่เสร็จเรียบร้อยแล้วสมบูรณ์มี 1 บทความ คือ

Langella, François. (2012). Polyfunctionality in Pwo Karen: The case of ?à?- (<T-B pronominal prefix *?a-). In Tadao Miyamoto, Naoyuki Ono, Kingkarn Thepkanjana and Satoshi Uehara (eds.) *Typological Studies on Languages in Thailand and Japan*. Tokyo: Hituzi Syobo Publishing. 41-55.

บทความของ อาจารย์ ดร. พิทยาวัฒน์ พิทยาภรณ์ เรื่อง Rhoticity and sonorancy of /v/ in Sgaw Karen จะส่งไปตีพิมพ์ในวารสารนานาชาติในประเทศสหรัฐอเมริกา ยุโรป หรือ ออสเตรเลีย จึงต้องใช้เวลามากในกระบวนการกำกับและควบคุมคุณภาพ

บทความ 3 เรื่องของศาสตราจารย์ ดร. ชีระพันธ์ เหลืองทองคำ ซึ่งมีความยาวประมาณ รวมทั้งสิ้น 140 หน้ากระดาษ A4 (บทความเรื่องที่ 1 = 18 หน้า บทความเรื่องที่ 2 = 41 หน้า

บทความเรื่องที่ 3 = 81 หน้า) จะยุบรวมกันเป็น monograph เพื่อส่งให้สำนักพิมพ์พิจารณาตีพิมพ์ ในเรื่องนี้ ได้เคยคุยไว้กับสำนักพิมพ์ White Lotus ไว้บ้างแล้ว บทความ 3 เรื่องนี้ คือ

L-Thongkum, Theraphan. (2011). Numerals and classifiers in Modern Karenic languages and Proto-Karen. Paper presented at the 44th International on Sino-Tibetan Languages and Linguistics, organised by Central Institute of Indian Languages, Mysore, India, October 7-9, 2011.

L-Thongkum, Theraphan. (2012). Proto-Karen (*k-rjaŋ^A) fauna. Paper presented at SEALS 22, co-organised by French laboratories and institutions, Agay, France, May 30-June 2, 2012.

L-Thongkum, Theraphan. (2013). A view on Proto-Karen phonology and lexicon. Paper to be presented at SEALS 23, organised by Faculty of Arts, Chulalongkorn University, Bangkok, May 29-31, 2013.

ในเรื่องเนื้อหาของ 10 บทความดังกล่าวข้างต้น ตีพิมพ์และเผยแพร่ได้ในภาคผนวก 1 นอกจากนี้ ศาสตราจารย์ ดร. วีระพันธ์ เหลืองทองคำ ได้ไปเสนอบทความในการประชุมวิชาการนานาชาติ 5 ครั้ง (ดูบทความย่อและพาวเวอร์พอยต์ในภาคผนวก 2) ผู้ช่วยศาสตราจารย์ ดร. ผณิตรา ชีรานนท์ ได้ไปเสนอผลงานในต่างประเทศ 1 ครั้ง และอาจารย์ ดร. ชมนาด อินทามรรักษ์ ได้ไปเสนอผลงานในต่างประเทศ 1 ครั้ง (การไปเสนอผลงานในต่างประเทศ 7 ครั้ง ได้ใช้งบประมาณจากแหล่งเงินทุนอื่น ไม่ได้ใช้งบประมาณของโครงการกะเหรี่ยง)

นอกจากบทความภาษาอังกฤษแล้ว อาจารย์ ดร. วิชาติ บุรณะประเสริฐสุข และคณะ ได้นำเสนอผลการสำรวจข้อมูลภาคสนามชุมชนกะเหรี่ยงหลายแห่งในหลายจังหวัด รวมทั้งประมวลและสังเคราะห์เรื่องราวเกี่ยวกับชาวกะเหรี่ยงและโครงการตามพระราชดำริ โดยใช้ข้อมูลทุติยภูมิจากเอกสาร ในรูปของหนังสือสารคดีประกอบภาพ 2 ภาษา (ไทย-อังกฤษ) เพื่อเป็นประโยชน์ต่อบุคคลทั่วไปในสังคม ตีพิมพ์และเผยแพร่ในหนังสือ “เรื่องเล่าจากชายแดนไทย-พม่า: ปรากฏภาพทางปัญญาเพื่อการพัฒนาอย่างยั่งยืน” ซึ่งโครงการภาษาศาสตร์ภาษากะเหรี่ยงได้จัดพิมพ์ 700 เล่ม เพื่อทูลเกล้าฯ ถวายสมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี 500 เล่ม สำหรับพระราชทานหน่วยงานและโรงเรียนสังกัดสำนักงานตำรวจตระเวนชายแดน และอีก 200 เล่มได้แจกให้ผู้มีอุปการคุณและบุคคลทั่วไปที่สนใจ

นอกจากบทความภาษาอังกฤษที่ส่งตีพิมพ์ใน Manusya Vol.15.2 ผู้ช่วยศาสตราจารย์ ดร. ศิริวิไล วีระโรจนารัตน์ ยังได้จัดทำแผนที่ 4 สี ขนาด 60 ซม. x 84 ซม. 2 แผ่น เป็น

ภาษาไทย 1 แผ่น และภาษาอังกฤษ 1 แผ่น ซึ่งเป็นผลการวิเคราะห์แบบสอบถามที่ส่งไปยังทุกตำบลใน 15 จังหวัดที่มีชาวกะเหรี่ยง 6 กลุ่มตั้งถิ่นฐานอยู่ (ดูแผนที่ในกระบอกสีดำ)

เว็บไซต์ 2 ภาษา (ไทย-อังกฤษ) ของโครงการภาษาศาสตร์ภาษากะเหรี่ยง แบ่งการนำเสนอออกเป็น 11 หน้าหลัก ได้แก่ (1) หน้าแรก (2) บทความเกิดพระเกียรติสมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี (3) แนะนำโครงการ (4) โครงการย่อย (5) ทีมวิจัย (6) การสำรวจภาคสนาม (7) การประชุมรายงานความก้าวหน้า (8) นิทรรศการงานจุฬาวิชาการครั้งที่ 14 (9) บทความย่อและพาวเวอร์พอยต์ประกอบการนำเสนอบทความในการประชุมวิชาการนานาชาติ (10) ประโยชน์ต่อสังคม และ (11) ติดต่อเรา (ดูตัวอย่างเว็บไซต์ <http://ling.arts.chula.ac.th/Karen> ในภาคผนวก 3)

นอกจากประชาสัมพันธ์เรื่องราวของโครงการภาษาศาสตร์ภาษากะเหรี่ยงทางเว็บไซต์แล้ว เมื่อช่วงวันที่ 14-18 พฤศจิกายน 2556 ยังได้จัดนิทรรศการ 2 ภาษา (ไทย-อังกฤษ) แสดงผลงานของโครงการฯ ในงานจุฬาวิชาการครั้งที่ 14 ซึ่งสมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารีได้เสด็จทอดพระเนตรนิทรรศการซึ่งจัดรวมกับนิทรรศการของชมรมภาษาศาสตร์ภาควิชาภาษาศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ณ ห้อง 301 อาคารบรมราชกุมารี (ดูภาคผนวก 4)

ในด้านความร่วมมือกับต่างประเทศ นอกจากการนำเสนอผลงานในการประชุมวิชาการระดับนานาชาติ 7 ครั้งดังกล่าวข้างต้น ศาสตราจารย์ ดร. วีระพันธ์ เหลืองทองคำ ยังได้ติดต่อแลกเปลี่ยนความคิดเห็นกับอาจารย์ชาวต่างประเทศซึ่งเป็นผู้เชี่ยวชาญด้านภาษาศาสตร์ภาษาตระกูลทิเบต-พม่า คือ Professor James A. Matisoff, University of California at Berkeley ประเทศสหรัฐอเมริกา ซึ่งเป็นผู้อำนวยการโครงการ STEDT (Sino-Tibetan Etymological Dictionary and Thesaurus) database และได้ตกลงกับ Professor James A. Matisoff ว่าจะร่วมงานกันใน 2 ประเด็น คือ (1) นำผลการสืบสร้างศัพท์ในภาษากะเหรี่ยงดั้งเดิมลงใน STEDT database ของ U.C. Berkeley และ (2) ร่วมมือกันศึกษาเปรียบเทียบภาษาสาขาโลโล-พม่า (Lolo-Burmese branch) ซึ่งสืบสร้างโดย Professor James A. Matisoff กับภาษาสาขากะเหรี่ยง (Karenic branch) ซึ่งเป็นผลงานของโครงการภาษาศาสตร์ภาษากะเหรี่ยง เพื่อวิเคราะห์ความสัมพันธ์ทางเชื้อสาย (Genetic relationship) ของทั้ง 2 สาขาซึ่งเป็นภาษาในตระกูลทิเบต-พม่า (Tibeto-Burman, TB) ตระกูลภาษาในมหาตระกูลจีน-ทิเบต (Sino-Tibetan, ST) ผลที่ได้จากการร่วมมือทางวิชาการครั้งนี้ จะช่วยตอบโจทย์ทางภาษาศาสตร์ภาษาเอเชียตะวันออกเฉียงใต้บางประการ (ดูตัวอย่าง STEDT database ในภาคผนวก 5)

ภาคผนวก 1

ต้นฉบับบทความภาษาอังกฤษ 10 บทความ

ศ.ดร. วีระพันธ์ เหลืองทองคำ

- **Numerals and classifiers in Modern Karenic languages and Proto-Karen**
- **Proto-Karen (*k-rjaŋ^A) fauna**
- **A view on Proto-Karen phonology and lexicon**

Numerals and Classifiers in Modern Karenic Languages and Proto-Karen

Theraphan L-Thongkum

Department of Linguistics, Chulalongkorn University, Bangkok
theraphan.l@chula.ac.th

Abstract

A word list of 2,000 items was devised after a brief survey of the Karen languages spoken in Thailand. The data on Pa-O, Kayan-Kayah-Kayaw and Pwo-Sgaw, representing northern Karen (NK), central Karen (CK) and southern Karen (SK), respectively, was collected at many locations by the author between March 2009 and May 2011. A tentative reconstruction of Proto-Karen phonology was made to aid in the reconstruction of Proto-Karen numerals and classifiers. Available documents on the numerals and numeral systems of Karenic languages and the other Tibeto-Burman languages were also consulted to support the author's interpretation.

The research results indicate that the numeral systems of NK and SK are decimal while those of CK (except Kayan) are more complex owing to the use of combined systems, e.g. $so^{33} swa^{55}$ (3x2) 'six' (base-3 system), $lwi^{33} swa^{55}$ (4x2) 'eight' (base-4 system), etc. as in Kayah. Also, there is harmony of the vowel in the word meaning 'one' and the following classifier, e.g. $ta^{11} ba^{11}$ 'a piece of flat object', $to^{11} bo^{11}$ 'a piece of long object', $to^{11} pɔ^{11}$ 'a lump of..., a ball of...' and so forth. This finding suggests that CK numerals and classifiers cannot be separately studied.

The numeral system of Proto-Karen is decimal. The etyma of PK numerals 'one' to 'ten' including 'hundred' and 'thousand' are given. The use of classifiers reflects their world view, i.e. SHAPE is more important than SIZE.

Keywords: Numerals, Numeral systems, Classifiers, Modern Karenic languages, Proto-Karen

1. Introduction

A Proto-Karen (PK) phonology and lexicon (195 items) were reconstructed by Jones (1961). Eight years later, Burling (1969) reanalysed PK using Jones' data.¹ Manson (2009b) with more data on nineteen Karenic languages drawn from his own field notes and the other available sources, both published and unpublished, tentatively reconstructed PK phonology to find some important features of phonological development to be used as the criteria for subgrouping the Karen languages (Mason, 2011).

Since July 2009, I have worked on six Karen languages spoken in Thailand, namely Pa-O (Taungthu), Kayan (Padaung), Kayah, Kayaw, Pwo and Sgaw.² A word list consisting

¹ In footnote 347 of the Sino-Tibetan: a conspectus (Benedict, 1972), Matisoff as the "contributing editor" states that both Jones and Burling paid no attention to the fundamental work by Haudricourt in BSLP 42 (1942-45) and BSLP 49 (1953) who brilliantly solved the problems in the reconstruction of Proto-Karen although having limited data sources. This was also my disadvantage because Haudricourt's works on Karen are not available in Thailand, except for the English version of his work (1961) translated by Court (1972) which contains a few lines of Karen information.

² My major field sites were: Ban Dong Dam in Li district, Lampun province (Pwo), Ban Pa La U in Hua Hin district, Prachuap Khiri Khan province (Sgaw), Ban Huai Suea Thao in Mueang district, Mae Hong Son province (E. Kayah, Kayan, Kayaw) and Wat Thaiwatthanaram in Mae Sot district, Tak province (Pa-O). The

of 2,000 items divided into twenty-one sections based on semantic fields was devised. The data was solely collected by myself to control any variation which could be caused by different linguistic training and the background of the field linguists and also the choice of symbols for transcription. Based on the fresh field data to hand, another tentative reconstruction of the Proto-Karen phonology and lexicon was made. All of the protoforms presented in this paper were drawn from my PK etymological lexicon which can be regarded as a work in progress. I am inclined to believe that Karenic is a branch of the Tibeto-Burman (TB) language family like Lolo-Burmese and Kachinic.

Various aspects of numerals and numeral systems have been studied synchronically and diachronically by linguists in different fields, for example, Przyluski and Luce (1931), Hurford (1975, 1999), Gvozdanović (ed., 1999), Comrie (1990, <http://wals.info/feature/description/131>) and so forth. The numeral systems of the TB languages and dialects were examined by Hodson (1913) using the specimens given in vol. iii of the Report of the Linguistic Survey of India and in the Gazetteer of Upper Burma, vol. i. He reported that “in the Karen group, the dialects are all decimal and prefix the multiplier and the bases are often very obviously associated with Tibeto-Burman forms” (Hodson, 1913:334). He also pointed out that “...the bases in the Karen group for two, three, four and five are akin to one or another of the standard group bases in the Tibeto-Burman languages but at six comes a break, i.e. Karenni, Yintale and Mano have the form three+three, and they form seven as three+three+one and eight as four+three+one” (p.335).

With regard to PK numerals, Jones (1961) reconstructed only five protoforms: *lwiT ‘four’, *ŋaT ‘five’, *nwét ‘seven’, *chih ‘ten’ and *rja ‘hundred’, due to a few unsolved problems. More recent works on the numerals and numeral systems of the TB languages with exhaustive and valuable arguments can be found in Benedict (1972), Matisoff (1997), Bradley (2005) and Mazaudon (2007). These works were my guides when I investigated the numerals in modern Karenic languages and their varieties in order to compare and reconstruct the protoforms. Following Matisoff (1997), I assumed that Karenic is a branch of the Tibeto-Burman languages as Lolo-Burmese, Kachinic and so forth. The Proto Tibeto-Burman (PTB) numerals: one to ten, a hundred and a thousand proposed by Benedict (1972), Matisoff (1997) and Bradley’s PSETB (Proto Southeastern TB or Proto-Loloish, 1979 and 2005) were used as the guidelines when I reconstructed the protoforms of the numerals in PK.

2. Numeral systems in modern Karenic languages

The classification of Karenic languages is somewhat different due to the use of different criteria (see Jones, 1961; Burling, 1969; Bradley, 1997; Shintani, 2003; Kauffman, 1993; Manson, 2009a and 2011). However, I ignored this fact and did my analysis from scratch, i.e. using only my field notes based on the devised word list to avoid bias. Typologically, I have found two major types of numeral system: a decimal system in Pa-O, Kayan, Pwo, Sgaw and a multiplicative and additive in E. Kayah and Kayaw for the big numbers from six to nine.

2.1 Decimal system

	Pa-O	Kayan	Pwo	Sgaw
ONE	ta ²¹	la ¹¹ la ²¹	la ¹¹ la ²¹	tə ¹¹ tə ³³
TWO	ni ^{31..}	ŋi ⁵³	khei ³⁵	khi ³³
THREE	səm ^{31..}	θə ⁵³	sə ³⁵	sə ³³

other available documented material by many authors (see the reference section) and the field notes of my co-researchers in the Karen Linguistics Project funded by the Thailand Research Fund (2009-2012) were also consulted. My present and previous field notes collected at many field sites were among the minor data sources.

FOUR	lip ⁴⁵	lit ⁴⁵	lwi ⁵³	lei ^{11..}	lwi ^{11..}
FIVE	ŋap ⁴⁵	ŋat ⁴⁵	ŋɛ ²¹	jɛ ^{11..}	jɛ ^{11..}
SIX	su ^{31..}		chu ⁵³	xou ³⁵	xi ³³
SEVEN	nit ⁴⁵		nwɛ ⁵³	nwɛ ⁵⁵	nwi ⁴⁵
EIGHT	sət ²¹		chau ⁴⁵	xo ²¹	xo ⁴⁵
NINE	kop ⁴⁵	kot ⁴⁵	khwi ⁵³	khwi ³⁵	khwi ³³
TEN	ɕi ^{31..}		thi ⁵³	chei ³⁵	chi ³³

Even though the numeral systems of Pa-O, Kayan, Pwo and Sgaw are decimal, the patterns of sound correspondence in some numerals seem irregular; for example, the initials or onsets of ‘one’ and ‘two’ and the tones of ‘four’ and ‘five’, due to the internal factors of sound change, e.g. the loss of prefixes, vocalic harmony, tone sandhi, etc. as well as language contact which can be an external factor of sound change.

2.2 Multiplicative and additive system

	E. Kayah	Kayaw
ONE	t- ^{11 3}	t- ^{11 33}
TWO	nə ³³	ki ⁵⁵
THREE	sɔ ³³	sə ⁵⁵
FOUR	lwi ³³	li ⁵⁵
FIVE	ŋɛ ³³	jɛ ³³
SIX	so ³³ swa ⁵⁵	sə ⁵⁵ su ³³
SEVEN	so ³³ swa ⁵⁵ t- ¹¹	sɔ ³³ ne ⁵⁵
EIGHT	lwi ³³ swa ⁵⁵	li ³³ su ³³
NINE	lwi ³³ swa ⁵⁵ t- ¹¹	sɔ ³³ khi ⁵⁵
TEN	chə ³³	ɕi ⁵⁵

With regard to numerals ‘one’, ‘two’, ‘three’, ‘four’, ‘five’ and ‘ten’, all of the Karen languages studied have separate morphemes. However, numerals ‘six’, ‘seven’, ‘eight’ and ‘nine’ are separate morphemes in Pa-O, Kayan, Pwo and Sgaw. In E. Kayah and Kayaw, multiplicative and additive forms based on a ternary system or 3-base system for numeral ‘six’ and a quaternary system or base-4 system for numeral ‘seven’ are used as shown below.⁴

	E. Kayah	Kayaw
SIX	so ³³ swa ⁵⁵ (3x2)	si ⁵⁵ su ³³ (3x2)
SEVEN	so ³³ swa ⁵⁵ t- ¹¹ (3x2+1)	(sɔ ³³) ne ⁵⁵ (...7)
EIGHT	lwi ³³ swa ⁵⁵ (4x2)	li ³³ su ³³ (4x2)
NINE	lwi ³³ swa ⁵⁵ t- ¹¹ (4x2+1)	(sɔ ³³) khi ⁵⁵ (...9)

³- in t-¹¹ stands for any vocalic element which can be predictable from the vowel of the following classifier. A sandhi between the tones of numerals and classifiers can be detected but there is no obvious pattern.

⁴Both systems, decimal and decimal combined with multiplicable and additive, are used in W. Kayah and Bwe. This is not unique, Mazaudon (2007: 5/22) stated that Dzongkha, the national language of Bhutan, has two numeral systems, a decimal system, borrowed from Tibetan, is used on formal occasions, while a vigesimal (base-20) system is used in everyday life.

Kayaw numerals ‘seven’ and ‘nine’ are problematic. There is no explanation for sɔ³³ preceding ne⁵⁵ ‘seven’ and khi³³ ‘nine’; however, it is obvious that ne⁵⁵ and khi⁵⁵ are cognate with numerals ‘seven’ and ‘nine’ in Pa-O (nit⁴⁵), Kayan (nwε⁵³), Pwo (nwε□⁴⁵) and Sgaw (nwi³⁵).

3. Vowel harmony and tone sandhi

The quality of the vowel in numeral ‘one’ is neither definite nor static. Its occurrence is conditioned by the vowel quality in the following classifier. In other words, the vowel of numeral ‘one’ is in perfect harmony with that of the following classifiers. Numeral ‘one’ never appears alone as when we count “one...two...three...”, it is always accompanied by one of the basic classifiers. When uneducated Kayah and Kayaw people are asked to count or to say the word meaning ‘one’ in their languages, they will say something like “one piece/lump...”, for example, to¹¹ ko¹¹ ‘one piece’ (E. Kayah), tɛ¹¹ klɛ³³ ‘one piece’ (Kayaw) and so on. Moreover, four patterns of tone-pair or tone sandhi can be found, especially in Kayaw, i.e. low-mid (LM), low-high (LH), mid-low (ML) and mid-mid (MM). E. Kayah and Kayaw noun phrases (NP) and verb phrases (VP) with the structure N+NUM+CLF⁵ and V+NUM+CLF respectively are given below to illustrate the phenomena of vocalic harmony and tone sandhi in E. Kayah and Kayaw. Tone sandhi in NUM-CLF phrases is not an unusual phenomenon in TB languages, see Bradley (2005) and Michaud (2011) for more information.

3.1 N+NUM+CLF

	E. Kayah	Kayaw
‘one drum’	thɔ ³³ tɛ ¹¹ mɛ ³³ (e-e)	thə ³³ tɑ ¹¹ mɑ ³³ (a-a, LM)
‘one house’	hi ¹¹ tɛ ¹¹ mɛ ³³ (e-e)	ɕi ¹¹ tɑ ¹¹ mɑ ³³ (a-a, LM)
‘one bird’	thu ¹¹ tɛ ¹¹ bɛ ¹¹ (e-e)	thu ¹¹ ta ³³ ba ¹¹ (a-a, ML)
‘one leaf’	le ¹¹ tɛ ¹¹ bɛ ¹¹ (e-e)	la ¹¹ ta ³³ ba ¹¹ (a-a, ML)
‘one eel’	tɛ ³³ ru ¹¹ tɔ ¹¹ bɔ ³³ (-)	tɔ ³³ ru ¹¹ tɔ ¹¹ bɔ ⁵⁵ (o- , LH)
‘one road’	klɛ ⁵⁵ tɔ ¹¹ bɔ ³³ (-)	klɛ ³³ tɔ ¹¹ bɔ ⁵⁵ (- , LH)
‘one string’	su ⁵⁵ plɛ ³³ tɔ ¹¹ bɔ ³³ (-)	si ¹¹ pli ³³ tɔ ¹¹ bɔ ⁵⁵ (- , LH)
‘one dog’	thwi ¹¹ tɔ ¹¹ dɔ ³³ (-)	thi ¹¹ tɔ ¹¹ dɔ ⁵⁵ (ə-ə, LH)
‘one lump of clay’	he ¹¹ klɔ ⁵⁵ tɔ ¹¹ plɔ ¹¹ (ɔ-ɔ)	ha ¹¹ tɔ ³³ khɔ ³³ (ə-ə, MM)
‘one ball of thread’	lo ¹¹ tɔ ¹¹ khɔ ¹¹ (u-u)	lu ¹¹ tɔ ³³ khɔ ¹¹ (u-u, ML)
‘one patch of cloud’	lɔ ⁵⁵ lo ¹¹ tɔ ¹¹ bɔ ¹¹ (i-i)	ta ³³ ɔ ¹¹ tɔ ³³ khɔ ³³ (ə-ə, MM)
‘one slice of cucumber’	də ³³ se ¹¹ tɔ ¹¹ bɔ ¹¹ (i-i)	di ⁵⁵ sa ¹¹ tɔ ³³ khɔ ¹¹ (i-i, ML)

3.2 V+NUM+CLF

	E. Kayah	Kayaw
‘kiss one time’	cwi ⁵⁵ tɔ ¹¹ phɔ ⁵⁵ (-)	ci ³³ ta ³³ pla ¹¹ (a-a, ML)
‘slap one time’	plɛ ¹¹ tɔ ¹¹ phɔ ⁵⁵ (-)	pla ¹¹ ta ³³ pla ¹¹ (a-a, ML)
‘do (sth.) one time’	me ¹¹ tɔ ¹¹ phɔ ⁵⁵ (-)	ma ³³ ta ³³ pla ¹¹ (a-a, ML)

⁵In some cases, the structure of this NP type is changed to N+CLF+NUM in both E. Kayah and Kayaw. When CLF precedes NUM, there is no vocalic harmony, see examples below.

	E. Kayah	Kayaw
‘one chicken’	cha ³³ tɛ ¹¹ be ¹¹	ɕi ⁵⁵ ta ³³ ba ¹¹
	chicken 1 CLF	chicken 1 CLF
‘ten chickens’	cha ³³ be ¹¹ chə ³³	ɕi ⁵⁵ ba ¹¹ ɕi ⁵⁵
	chicken CLF 10	chicken CLF 10

The numeral ‘ten’ in the six Karenic languages examined is interesting, i.e. it can occur with numeral ‘one’ in the order ‘one ten’, e.g. ta³³ ɛi^{31..} (Pa-O), la¹¹ thi⁵³ (Kayan) and so on. Due to this fact, Matisoff (1997) analyses ‘ten’ as a classifier. With regard to ‘hundred’ and ‘thousand’, they are like ‘ten’ in the sense that they can also occur with ‘one’ as shown below.

	Pa-O	Kayan	S. Pwo	N. Sgaw	E. Kayah	Kayaw
‘hundred’	ta ²¹ rja ³³	la ¹¹ ja ³³	la ¹¹ ja ^{31..}	tə ¹¹ kə ¹¹ ja ³³	te ¹¹ je ¹¹	ta ¹¹ ja ³³
‘thousand’	ta ²¹ reŋ ^{31..}	la ¹¹ re ⁵³	lə ¹¹ thɔ̃ ^{53 6}	tə ¹¹ kə ¹¹ tho ³³	ti ¹¹ ri ³³	to ³³ tho ³³ , te ⁵⁵ re ⁵⁵

In Kayaw, to³³ tho³³ is used for counting, e.g. people, for example, pra³³ to³³ tho³³ (person one thousand), and te⁵⁵ re⁵⁵ is used for counting e.g. money, for example, ru¹¹ te⁵⁵ re⁵⁵ (money one thousand). It is also worth pointing out that in Sgaw, kə¹¹ can optionally be inserted between ‘one’ and ‘hundred’ and between ‘one’ and ‘thousand’. This kə¹¹ can be viewed as the prefix of ja³³ ‘hundred’ and of tho⁵⁵ ‘thousand’ or as a conjunction. To reach a definite answer, more study is needed.

4. Proto-Karen numerals

Based on my field data, a Proto-Karen phonology and lexicon were tentatively reconstructed. Other relevant documented material was also consulted. See the list of abbreviations in Appendix 1.

There are two types of syllable in PK, smooth or non-checked syllable (*CV, *CVN) and checked syllable (*CVK). The capital N stands for a final nasal and the capital K for a final stop including -ʔ. Unlike the other Karen languages, Pa-O still retains both PK nasal and stop finals. Jones (1961) reconstructed PK final consonants as follows: *-m, *-n, *-ŋ, *-ɲ, *-t, *-th, *-T, *-k, *-K, *-d, *-dh, *-D, *-g, *-gh, *-G, *-, *-h and *-q. Among modern Karen languages, Pa-O has a few finals while Kayan and Pwo have only -ŋ and -ɲ. With regard to Sgaw, E. Kayah and Kayaw, only -ɲ is retained. During my interviews with the Pa-O, I noticed that the three final nasal -m, -n, and -ŋ were in free variation as well as the final stops -p, -t and -k. In my opinion, Jones’s reconstruction of PK final consonants is not convincing. The following is my tentative reconstruction of PK tones⁷, consonants and vowels.

Tones:	*A (smooth tone)	*B (smooth tone)	*D (checked tone)
Initials:	*ph- *p- *b-	*ɲb- *m- *ɲm-	*hm- *w- *ɲw- *hw-
	*th- *t- *d-	*ɲd- *n- *ɲn-	*hn- *l- *ɲl- *hl- *r- *s-
	*ch- *c- *ɲ -	*j- *ɲj- *hj-	

⁶The original Karen morpheme meaning ‘thousand’ was lost in the Pwo variety spoken at Ban Dong Dam in Li district, Lampun province which is my major field site for Pwo. In fact, the Thai word /phan/ ‘thousand’ was borrowed into most of the Pwo varieties spoken in northern Thailand (see examples in Culy, 1993). On the other hand, the Karen morpheme meaning ‘thousand’ is still kept in many Pwo varieties spoken in central Thailand, examples can be found in Phillips (1996). In many Sgaw Karen varieties, the Thai word /mi ɲn/ ‘ten thousand’ is used for ‘thousand’

⁷With the tone-box concept (*A *B *C *D), Shintani (2003) and Manson (2009b) proposed four proto tones. I think that the third non-checked tone (2 or B) which is equivalent to the *C tone in Proto-Tai is redundant. There is no point in saying that this PK tone has completely merged with the other tones in modern Karenic languages. PTB is non-tonal (Benedict, 1972; Matisoff, 2003). In fact, the PK three tones, i.e. *A *B and *D (equivalent to Proto-Tai *D) are adequate to show a possible scenario for the tonal development in the Karen languages and its varieties. See Appendix 2.

*kh- *k- *g- *ŋ- *ɲ- *hŋ-
 *h- *ɲ
 Clusters: *Cw- *Cr- *Cl- *Cj- *Crw- *Clw-
 Vowels: *i *e *ɛ *ɨ *ə *a *u * *ɔ
 *ei/I *əi/ə * u/U *ai *au

Basically, the numeral system of PK is decimal or base-10 even though combined bases can be found in some of the modern Karenic languages. The etyma of PK numerals ‘one’ to ‘ten’ including ‘hundred’ and ‘thousand’ will be given. In addition, the Karen word meaning ‘10,000’ will be discussed. To obtain comparable data, I converted the various ways of marking tones in modern languages and their varieties into a simple one: in smooth syllables, 11/21 = low, 33 = mid, 55/45 = high, 31 = mid-falling, 53/52 = high-falling and in checked syllables, ɲ21 = low and ɲ45 = high. The followings are PK numerals and basic classifiers.

ONE *t-la^T [PTB *it, *kat, *g-t(y)ik (JM); *g-t(y)ik (PB); PSETB *ti² (DB)]
 PK * - (RJ)
Pa-O: taɲ²¹ (TL); tə³³ (RJ); tə³³ (OT)
Kayan: la¹¹ (TL); la¹¹, ta¹¹ (KM); ta³³ (DIC); la³³ (SK)
Kayah: t-¹¹ (TL); ta¹¹, tə¹¹ (DIC); tə³³ (DS)
Kayaw: t-¹¹, t-³³ (TL)
Bwe: də³³, tə³³, to⁵⁵ (EH)
Yinbaw: la¹¹ (HS)
Pwo: lə³³, lə̃³¹, la¹¹, lə¹¹, la^{21'} (TL); lən (RJ); li³¹, lə³¹ (AP); li³³, ni³³ (MC);
 la¹¹, la³³, la^{21'} (PT); na²¹ (CI)
Sgaw: tə³³, tə¹¹, tə⁴⁵, taɲ²¹ (TL); tə³³ (RJ); tə²¹ (SR); tɣ⁴⁵ (PP)

Note: This etymon can have tone *B or *D. In modern Karenic languages, the vowel in numeral ‘one’, although it varies, is always a central vowel: a, ə/ɤ, i, except in E. Kayah and Kayaw where it can be any vowel depending upon the vowel of the following classifier due to vowel harmony. It is noticeable that in Ban Khang Chai Pwo (CI) *l- becomes n-.⁸

TWO *k-hnej^A [PTB *g-nis (JM, PB); PSETB *s-ni² (DB)]
 PK * - (RJ)
Pa-O: ni¹¹, ni³¹ (TL); nɨ (RJ); ni¹¹ (OT)
Kayan: ŋi⁵³ (TL); ŋu⁵⁵ (KM); nu³³ (DIC); ŋu⁵² (SK)
Kayah: nə³³ (TL); nje³³ (DIC); nɛ³³ (DS)
Kayaw: ki⁵⁵, ki³³ (TL); ki³³ (WB)
Bwe: ki⁵⁵ (EH)
Yinbaw: ŋə⁵⁵ (HS)
Pwo: nei³¹, khe³⁵, nə¹¹ (TL); ni¹¹ (RJ); nɨ⁴⁵, nɨ⁴⁵ (AP); khi⁴⁵ (MC);
 khi⁴⁵ (PT); khiŋ⁴⁵ (CI)
Sgaw: khi³³, khi⁵⁵ (TL); khi³³ (SR); khi

Theraphan L-Thongkum –The 44th ICSTLL

Note: In the stem *hnej^A the initial or onset *hn- (voiced nasal) in some modern languages and their varieties but *i has been quite well kept in

⁸Alternations between d, l, n and n, d are rather common in SEA languages since these three sounds share the same place of articulation (alveolar/dental) and the same kind of phonation type (voiced).

almost all. In some cases, the prefix *k- was dropped as in Pa-O, Kayan, Kayah and some Pwo varieties and became kh-, i.e. the pre-aspiration of *hn had shifted to the prefix so *k-hn > *kh-n and later either *kh or *n was lost, one way or another depending on the choice of each language and variety. There had been also a place of articulation (velar) assimilation, i.e. *n > ŋ before *kh- was dropped.

THREE *səm^A [PTB *g-sum (JM, PB); PSETB *c-sum² (DB)]
PK *- (RJ)

Pa-O: səm³¹ (TL); som¹¹ (RJ); som⁵⁵ (OT)

Kayan: θə⁵³ (TL); θjən⁵⁵ (KM); si⁵² (SK)

Kayah: sɔ³³ (TL); θwə³³ (DIC); so³³ (DS)

Kayaw: sɯ⁵⁵, sə⁵⁵ (TL)

Bwe: θo⁵⁵, θɔ⁵⁵ (EH)

Yinbaw: sɤ⁵⁵ (HS)

Pwo: θə̃⁵³, sa³⁵, θəi¹¹ (TL); θən¹¹ (RJ); θə̃⁴⁵, səŋ⁴⁵, siŋ⁴⁵ (AP); sau⁴⁵, sau³³,
saɪ⁴⁵, sə³³, sə⁴⁵, si⁴⁵, sa⁴⁵, sa³³, θa³³ (MC); sau⁴⁵ (PT); səŋ⁴⁵ (CI)

Sgaw: θə⁴⁵, sə⁵⁵, sə^{45'}, θə⁵⁵, θə³³, sə³³ (TL); θə́ (RJ); sə⁴⁵ (SR); sɤ^{45'} (PP)

Note: *səN^A can be *səm^A if the reconstruction of the final nasal is based on -m in Pa-O. In the other Karen languages *-N has been lost, except in some Pwo varieties, *-N/*-m has been retained as -n or -ŋ. In some cases, nasality has been shifted to the preceding vowel and *ə has become diphthongized. Moreover, the PK central vowel *ə seems to have shifted to many positions in the vowel area, resulting in the difference of vowel qualities in modern Karenic languages.

FOUR PNK *lit^D; PCK *hlwi^A; PSK *lwi^B [PTB *b-leɣ, *b-ləy (JM, PB); PSETB *b-le² (DB)]

PK *lwiT (RJ)

Pa-O: lip⁴⁵, lii⁴⁵ (TL); lit⁴⁵ (RJ); lii⁴⁵ (OT)

Kayan: lwi⁵³ (TL); lwe⁵⁴ (KM); lweⁱ (DIC); luy³³ (SK)

Kayah: lwi³³ (TL); lwi³³ (DIC); lwi³³ (DS)

Kayaw: lwi⁵⁵, li⁵⁵ (TL)

Bwe: lu⁵⁵ (EH)

Yinbaw: lwi⁵⁵ (HS)

Pwo: lwi¹¹, lei³³ (TL); lí, li[□] (RJ); lii⁴⁵, lēi⁴⁵ (AP); li³³, lwi³³ (MC); li²¹ (PT); liŋ²¹ (CI)

Sgaw: lwi¹¹, lwi³¹ (TL); lwĩ (RJ); lwĩ (SR); lwi²¹ (PP)

Note: The tone of the PK numeral ‘four’ cannot be reconstructed with certainty due to the different directions of tonal development in modern Karenic languages, i.e. > tone C in Pa-O, > tone A in Kayan, Kayah, and > tone B in Pwo and Sgaw. Based on, Pa-O, Jones (1961) reconstructed *lwiT (checked syllable with a high tone) as the protoform. I think that the reconstructed form should be *hlwi^A or *lwi^B; however, *hlwi^A is more likely. The words meaning ‘four’ in Kayan, Kayah, Kayaw, Bwe and Yinbaw (Central Karen) have tones indicating the development from tone *A.

FIVE *ŋjat^D [PTB *l-ŋa *b-ŋa (JM); l/b-ŋa (PB); PSETB *ŋa² (DB)]
PK *ŋaT (RJ)

Pa-O: $\eta a p^{45}$ $\eta a t^{45}$ (TL); $\eta a t'$ (RJ); $\eta a t^{45}$ (OT)
Kayan: $\eta \varepsilon^{21}$ (TL); $\eta a i^{54}$ (KM); $nga i$ (DIC); $\eta \varepsilon^{33}$ (SK)
Kayah: $\eta \varepsilon^{33}$ (TL); ηa^{33} (DIC); $\eta \varepsilon^{33}$ (DS)
Kayaw: $j \varepsilon^{33}$ (TL)
Bwe: $j \varepsilon^{55}$, $j \varepsilon^{11}$ (EH)
Yinbaw: $\eta a i^{55}$ (HS)
Pwo: $j \varepsilon^{11}$, $j \varepsilon^{33}$, $j e^{55}$ (TL); $j \varepsilon'$, $j a^{11}$ (RJ); $j e^{45}$ (AP); $j \varepsilon^{33}$, $j e^{33}$ (MC); $j \varepsilon^{21}$ (PT);
 $j e^{21}$ (CI)
Sgaw: $j \varepsilon^{11}$, $j \varepsilon^{31}$, $z \varepsilon^{11}$, $\phi \varepsilon^{11}$, $j \varepsilon^{55}$ (TL); $j \varepsilon$ (RJ); $z \varepsilon$ (SR); $z \varepsilon^{21}$ (PP)

Note: Like numeral ‘four’, the proto-tone of this etymon cannot be reconstructed because it becomes tone C in Pa-O and Kayan, Pwo, Sgaw and tone A in Kayah and Kayaw. The protoform $*\eta aT$ (Jones, 1961) is based on $\eta a t^{45}$ in Pa-O. The development of the initial $*\eta$ and the vowel $*a$ is $*\eta a > *\eta a^i$ (off-gliding) $> \eta a i$ or $*\eta a > *\eta^i \varepsilon$ (on-gliding) $> *\eta j \varepsilon > \eta \varepsilon / j \varepsilon$. The change of j - to z - in some Sgaw varieties can be explained, i.e. very often the initial j - in modern Karen is pronounced as a voiced palatal fricative [j], especially when it precedes a front vowel. Phonetically, both [j] and [z] are voiced fricatives.

SIX $*khrow^A$ [PTB $*d-ruk$ (JM, PB); PSETB $*c-krok^L$ (DB)]

PK $*- (RJ)$
Pa-O: su^{31} (TL); $s u$ (RJ); su^{11} (OT)
Kayan: chu^{53} (TL); chu^{11} , shu^{11} (KM); chu^{52} (SK)
Kayah: $shu\omega^{55}$ (DIC)
Kayaw: -
Bwe: xu^{55} (EH)
Yinbaw: $\int u^{55}$ (HS)
Pwo: xou^{35} , xou^{53} , $h\tilde{u}\eta^{31}$ (TL); xu^{11} (RJ); $x\tilde{u}^{45}$, $xu\eta^{55}$ (AP); xu^{45} , xu^{33} (MC); xu^{45} (PT); $xu\eta^{45}$ (CI)
Sgaw: $x i^{33}$, $x i^{55}$ (TL); $x y$ (RJ); $x i^{33}$ (SR); $x u^{45}$ (PP)

Note: The numeral ‘six’ is so^{33} swa^{55} (3x2) in E. Kayah and $s\omega^{55}$ su^{33} or shu^{55} si^{11} (3x2) in Kayaw. This could be an innovation caused by language contact. In some varieties of Pwo Karen spoken in the west of central Thailand and in the Karen State (Am La village) $*u$ becomes nasalized, so $*u > \tilde{u} > \tilde{u}\eta > u\eta$. The cluster $*khr-$ becomes $s-$ in Pa-O, ch / sh - in Kayan, sh - in W. Kayah Li, $\int$ - in Yinbaw, h - in Am La Pwo and x - in most varieties of Pwo and Sgaw. This pattern of sound change, i.e. $*khr- > ch-$, $sh-$, $h-$ or $x-$ can also be found in other etyma such as $*khrwit^D$ ‘bone’, $*khr\omega t^D$ / $*gr\omega t^D$ ‘eight’ and so on.

SEVEN $*\square nwet^D$ / $*\square nwe^A$ [PTB $*s-nis$ (JM, PB); PSETB $*c-\int i(k)^{2/L}$ (DB)]

PK $*nwet$ (RJ)
Pa-O: nit^{45} , $n\omega t^{45}$ (TL); $n\omega t'$ (RJ); $n\omega t^{45}$ (OT)
Kayan: $n\omega \varepsilon^{53}$ (TL); nue^{11} (KM); $nwa i$ (DIC); $n\omega \varepsilon^{33}$ (SK)
Kayah: $n\omega o^{55}$ (DIC)
Kayaw: $(s\omega^{33}) ne^{55}$ (TL); ne^{45} (WB)
Bwe: $n\omega e^{55}$, $n\omega i^{55}$ (EH)
Yinbaw: $n\omega \varepsilon^{55}$ (HS)
Pwo: $n\omega \varepsilon^{55}$, $n\omega \varepsilon^{31}$ (TL); $n\omega e^{31}$ (RJ); $n\tilde{u}e^{45}$, $n\tilde{u}e^{33}$, $n\tilde{u}e^{31}$ (AP); $n\omega e^{45}$, $n\omega \varepsilon^{33}$ (MC); noi^{45} (CI)
Sgaw: $n\omega i^{55}$, $n\omega i^{245}$ (TL); $n\omega i$ (RJ); $n\omega i^{045}$ (SR); $n\omega i^{045}$ (PP)

Note: Based on -t in Pa-O, Jones reconstructed the protoform of numeral ‘seven’ as *ʔnwét. However, the data suggests *A (see Kayan, W. Kayah, Bwe and Kayaw) but *D in Pa-O and some varieties of Pwo and Sgaw. In E. Kayah, a multiplicative and additive form of this numeral, i.e. so³³ swa⁵⁵ t-¹¹ (3x2+1) is used.

EIGHT *khrɔt^D/grɔt^D [PTB *b-r-gyat b-g-ryat (JM); b-r-gyat (PB); PSETB *c-yet^L (DB)]
 PK *- (RJ)
Pa-O: sɔt²¹, sət²¹ (TL, OT); sɔt⁴⁵ (RJ)
Kayan: chau⁴⁵ (TL); chau¹¹, shɔ¹¹ (KM); sɔ (DIC); chaw⁵² (SK)
Kayah: -
Kayaw: -
Bwe: xɔ¹¹ (EH)
Yinbaw: shɔ⁵⁵ (HS)
Pwo: xo²¹, xu⁴⁵, hu⁴⁵ (TL); xo[□], xo[□] (RJ); xu⁴⁵ (AP); xo³³, xo⁴⁵, xo³³, xu³³ (MC); xo³³ (PT); xo²¹ (CI)
Sgaw: xɔ⁴⁵, xo²¹, xo²¹, xɔ²¹, xɔ¹¹ (TL); xɔ[□] (RJ); xɔ⁴⁵ (SR, PP)

Note: The tones of modern Karenic languages suggest that the protoform should have *gr- and tone *D, except Kayan (tone A). In E. Kayah, W. Kayah and Kayaw, numeral ‘seven’ is multiplicative (4x2). With regard to Bwe, Henderson also gives the word lwi⁵⁵ θu³³ (4x2) for this numeral.

NINE PNK *kot^D; PCK, PSK *khwi^A [PTB *d-kəw *d-gaw (JM); d-kəw (PB); PSETB *(□-) go² (DB)]
 PK *- (RJ)
Pa-O: kop⁴⁵ kot⁴⁵ (TL); kút (RJ); kut⁴⁵ (OT)
Kayan: khwi⁵³ (TL); khwi⁵⁵, khwi³³ (KM); khwi, khwi (DIC); khuy⁵² (SK)
Kayah: -
Kayaw: (sɔ³³) khi⁵⁵ (TL)
Bwe: khwi⁵⁵ (EH)
Yinbaw: khwɛi⁵⁵ (HS)
Pwo: khwi³⁵, khwi⁵⁵, khwi⁵³ (TL); khwí, khwi (RJ); khũ³¹ (most varieties), khũ³³ (AP); khwi⁵⁵ (most varieties), khi³³ (MC); khwi³⁵ (PT); khwi⁴⁵ (CI)
Sgaw: khwi⁵⁵, khwi³³ (TL); khwí (RJ); kwi³³ (SR); khwi⁴⁵ (PP)

Note: The final stop in Pa-O which suggests the *D tone is a puzzle. Moreover, the Pa-O tone 45 (D3) also indicates that the initial k- is from *g-.

TEN *chej^A [PTB *ts(y)i(y) (PB); *ts(y)i(y) tsyay (JM); PSETB *tsay¹ (DB)]
 PK *chih (RJ)
Pa-O: ɕi³¹ (TL); chi (RJ); chi¹¹ (OT)
Kayan: thi⁵³ (TL); chi⁵⁵, shi⁵⁵ (KM); su (DIC); thu⁵² (SK)
Kayah: chə³³ (TL); chΛ³³ (DS); shje⁵⁵ (DIC)
Kayaw: ɕi⁵⁵ (TL); ʃhi⁴⁵ (WB)
Bwe: ʃi⁵⁵ (EH)
Yinbaw: ʃu⁵⁵ (HS)
Pwo: chei³⁵, chei⁵³ (TL); shì (RJ); tʃi⁴⁵, tʃi⁵¹ (AP); ɕhi⁵⁵ (most varieties), ɕhe³³, ɕhe³³ (MC); chi³³ (PT); cin⁴⁵ (CI)

Sgaw: chi³³, chi⁵⁵ (TL); shí (RJ); shi³³ (SR); chi⁴⁵ (PP)

Note: Numeral ‘ten’ seems to have a regular sound change in most of the modern Karen languages. Numeral ‘one’ can be added in front of numeral ‘ten’, for example, ta¹¹ ɕi^{31..} (Pa-O), la¹¹ thi⁵³ (Kayan), la¹¹ chei³⁵ (Pwo), ta¹¹ chi⁵⁵ (Sgaw) and so forth. With regard to ‘eleven’, it is 10+1, e.g. ɕi⁵⁵ ta¹¹ ra³³ (10 1 CLF) ‘eleven people’ in Kayaw.

HUNDRED *g-rja^A [PTB *r-gya (JM, PB); PSETB *C-rya¹ (DB)]

PK *rja’ (RJ)

Pa-O: rja⁵³, rja³³ (TL); rja (RJ); rja³³ (OT)

Kayan: ja³³ (TL); ja²², zaⁱ⁵⁴ (KM); ja³³ (DIC); ja¹¹ (SK)

Kayah: je¹¹ (TL); je³³ (DS); zε¹¹ (DIC)

Kayaw: ja³³ (TL)

Bwe: gə³³ jε³³ (EH)

Yinbaw: djwε⁵⁵ (HS)

Pwo: ja⁵⁵, ja³³, ja^{31..} (TL); ja, ja’ (RJ); ja³¹ (AP); jε³³, ja³³ (MC); ja²¹ (PT); ja³³ (CI)

Sgaw: ja³³, ja¹¹, za³³ (TL); ja (RJ); za³³ (SR); za⁴⁵ (PP)

Note: The prefix *g- is reconstructed from gə³³- in gə³³ jε³³ (Bwe) and kə¹¹ in tə¹¹ kə¹¹ ja³³, tə¹¹ kə¹¹ za^{33/55} (Sgaw). Even though the initial of the numeral ‘hundred’ in most of the modern Karen languages is j-, the one of the protoform is reconstructed as a cluster, i.e. *rj- because of the cluster rj- in Pa-O. It is also noticeable that the reconstructed forms in PTB and PSETB also have *ry- (*rj-). Moreover, in Proto-Waic (a branch of AA) which was reconstructed by Diffloth (1980), the etymon *ryah ‘hundred’ also has a cluster *ry- (*rj-). I think that the Waic protoform *ryah can be an old TB loanword.

THOUSAND *hrej^A, *g-thoN^A [PTB *s-tonj, *s-rinj (JM); PSETB *[- tonj (DB)]

PK *- (RJ)

Pa-O: rej^{31..} (TL); reñ (RJ); rej¹¹ (OT)

Kayan: re⁵³ (TL); reñ¹¹, rej¹¹ (KM); reñ, rein (DIC); re⁰²¹ (SK)

Kayah: ri³³ (TL); ri⁵⁵ (DIC); ri⁵⁵ (DS)

Kayaw: re⁵⁵ (for money), tho³³ (for people) (TL); ri³³ (WB)

Bwe: gə³³ tho⁵⁵ (EH)

Pwo: thō⁵³, thoñ³¹ (TL); thōn (RJ); thō⁵¹, thoñ⁵¹ (AP); tho³³ (MC)

Sgaw: tho³³, tho⁵⁵, thu⁵⁵ (TL); kə[□] tho’ (RJ)

Note: Both re⁵⁵ and tho³³ (should be ⁵⁵) are used in Kayaw, the former is for money and the latter is for people. The reconstruction of the prefix *g- is based on gə³³ in gə³³ tho⁵⁵ in Bwe (EH) and kə¹¹ in Sgaw Karen, e.g. tə¹¹ kə¹¹ tho³³ (TL, Huay Khom), tə¹¹ kə¹¹ thu⁵⁵ (TL, Pa La U), tə¹¹ kə¹¹ tho⁵⁵ (TL, Huay Phueng, Nong Khao Klang) and so on.

Interestingly, the Thai words /phan/ ‘thousand’ and /mi ñn/ ‘ten thousand’ have been borrowed into most of the Pwo and Sgaw varieties spoken in Thailand such as phεñ¹¹ (TL, Dong Dam Pwo), phā³³, phē³³, phō³³, pho³³ (MC, Pwo varieties spoken in northern Thailand), phāñ³⁵ (PT, Ban Doi Pwo), phañ¹¹ (CI, Khang Chai Pwo), mā³³, mā³³ (MC, Pwo varieties spoken in northern Thailand), mə³³ (SR, Huay Bong Sgaw), mə³³ (TL, Sgaw varieties spoken at Pha Mon, Yang Kham Nu, Kama Pha Do) and mɤ⁴⁵ (PP, Sut Huay Na Sgaw)

With regard to ‘ten thousands’, the Thai words /mi ñn/ ‘ten thousands’, /sε ñn/ ‘hundred thousands’ and /laan/ ‘million’ have been borrowed by the Pwo and Sgaw living in Thailand. These three words are confusingly used by the Karen. On the other hand, the Pa-O, Kayan, Kayah and Kayaw, although living in northern Thailand, use a non-Thai as shown below.

Pa-O: sɔŋ³³ (TL)

Kayan: θɔ⁵³ (TL); θɔŋ⁵⁵ (KM); tho⁵⁵ n (DIC, th= θ); sɔ³³ (SK)

Kayah: sɔ³³ (TL); θɔ⁵⁵ (DIC)

Kayaw: sɔ⁵⁵ (TL)

4. Proto-Karen classifiers

Solely based on my field notes, five basic PK classifiers were reconstructed.⁹ The use of classifiers in Proto-Karen and modern Karen clearly reflect their world view. To the Karen, SHAPE is more important than SIZE.

CLF [+human] *bra^A

Pa-O: pra⁵³ (Mae Sot), phra³³ (Huay Khan)

Kayan: pra³³ (Huay Suea Thao)

E. Kayah: phre¹¹ (Huay Suea Thao)

Kayaw: ra³³ (Huay Suea Thao)

Pwo: va⁵⁵ (Dong Dam), va¹¹ (Khao Lek)

Sgaw: va³³ (Pa La U, Huay Khom, Yang Kham Nu, Pha Mon, Nong Khao)

CLF [-human +4-legged mammal] *ɗə^A

Pa-O: -

Kayan: dəi⁵³ (Huay Suea Thao)

E. Kayah: do³³ (Huay Suea Thao)

Kayaw: də⁵⁵ (Huay Suea Thao)

Pwo: dəi⁵⁵ (Dong Dam), di¹¹ (Khao Lek)

Sgaw: di⁵⁵ (Pa La U, Nong Khao), di³³ (Huay Khom, Yang Kham Nu, Pha Mon)

Note: The classifier for 4-legged mammals in Pa-O is ba³³ (Mae Sot) or ba⁵⁵ (Huay Khan). This indicates that the world view of the Pa-O as reflected in the use of classifier has been changed, due to earlier separation from the rest of the Karen.

CLF [±animate +flat] *ɗa^B

Pa-O: ba³³ (Mae Sot), ba⁵⁵ (Huay Khan)

Kayan: ba¹¹ (Huay Suea Thao)

E. Kayah: be¹¹ (Huay Suea Thao)

Kayaw: ba¹¹ (Huay Suea Thao)

Pwo: bai³³ (Dong Dam), bei⁵⁵ (Khao Lek)

Sgaw: bi^{45'} (Pa La U), be³¹, be²¹ (Huay Khom, Yang Kham Nu, Pha Mon, Nong Khao)

CLF [±animate +long] *ɗɔN^A

Pa-O: -

Kayan: bɔ⁵³ (Huay Suea Thao)

E. Kayah: bo³³ (Huay Suea Thao)

Kayaw: bo⁵⁵ (Huay Suea Thao)

Pwo: bɔ⁵⁵, bɔŋ⁵⁵ (Dong Dam), bɔ¹¹ (Khao Lek)

Sgaw: bu⁵⁵ (Pa La U, Nong Khao), bo³³ (Huay Khom, Yang Kham Nu, Pha Mon)

⁹For more general information on Karen classifiers, see Ratanakul (1997) for Sgaw and Has Eh Ywar and Manson (2009) for Kayan.

Note: Proto-Karen * $\text{b}\text{ɔ}\text{N}^{\text{A}}$ was lost in Pa-O. The Pa-O living in Thailand use classifier li^{53} for [+long] animate and inanimate objects.

CLF [$\pm$ animate +round] * $\text{phlo}\text{ŋ}^{\text{B}}$

Pa-O: $\text{phlo}\text{ŋ}^{\text{33}}$ (Mae Sot), $\text{phlo}\text{ŋ}^{\text{55}}$ (Huay Khan)

Kayan: $\text{phl}\text{əu}^{\text{11}}$ (Huay Suea Thao)

E. Kayah: $\text{phl}\text{ə}^{\text{11}}$ (Huay Suea Thao)

Kayaw: $\text{phl}\text{ə}^{\text{11}}$ (Huay Suea Thao)

Pwo: $\text{phla}.\text{u}^{\text{33}}$ (Dong Dam), $\text{phl}\text{əu}^{\text{55}}$ (Khao Lek), $\text{phlo}\text{ŋ}^{\text{55}}$ (Am La)

Sgaw: $\text{phli}^{\text{45'}}$ (Pa La U), $\text{phl}\text{ə}\text{ŋ}^{\text{21}}$, $\text{phl}\text{ə}^{\text{31~}}$ (Huay Khom, Yang Kham Nu, Pha Mon, Nong Khao)

5. Conclusion

Karenic as a branch of the Tibeto-Burman language family is similar to the other branches such as SETB (Lolo-Burmese) in the way that numerals show various irregular correspondence patterns which result from vocalic harmony and tonal sandhi in combination with a numeral and a following classifier as in Kayah and Kayaw. With limited sets of data and the rigid comparative method used by Jones (1961), only few PK numerals were reconstructed: * lwi^{T} ‘four’, * ŋa^{T} ‘five’, * $\text{nw}\text{et}^{\text{T}}$ ‘seven’ * chih^{T} ‘ten’ * rja^{T} ‘hundred’; as for classifiers, only one basic classifier, * phronq ‘spherical; sphere (cN)’ was given.

Having more comparable fresh field data collected mostly by one person to hand, a reconstruction of PK numerals and basic classifiers is quite possible, i.e. * t-la^{T} ‘one’, * k-hnej^{A} ‘two’, * səm^{A} ‘three’, * lit^{D} , * hlwi^{A} , * lwi^{B} ‘four’, * ŋjat^{D} ‘five’, * khrow^{A} ‘six’, * ŋnwet^{D} / * ŋnwe^{A} ‘seven’, * khrot^{D} / * grɔt^{D} ‘eight’, * kot^{D} / * khwi^{A} ‘nine’, * chej^{A} ‘ten’, * g-rja^{A} ‘hundred’, * g-thɔN^{A} & * hrej^{A} ‘thousand’, * bra^{A} ‘CLF [+human]’, * $\text{ɔd}\text{ə}^{\text{A}}$ ‘CLF [-human +4-legged mammal]’, * ba^{B} ‘CLF [$\pm$ animate +flat]’, * bɔN^{A} ‘CLF [$\pm$ animate +long]’, * $\text{phlo}\text{ŋ}^{\text{B}}$ ‘CLF [$\pm$ animate +round]’.

In spite of a base-10 (decimal) system of numerals in Proto-Karen, in modern Karen both the inherited decimal system (as in Pa-O, Kayan, Pwo and Sgaw) and mixed or combined systems (as in Kayah and Kayaw) can also be found. Perhaps, this development has resulted from the social, cultural and language contact among the various ethnic groups inhabiting in the TB linguistic area (detailed information on this type of evidence can be found in LaPolla (2001), Matisoff (1983). However, the small numbers from one to five have been preserved quite well in all of the modern Karen languages. According to the Karen world view as suggested by the use of classifiers, SHAPE is more important than SIZE and animals if they are not mammals having four legs are classified by shapes together with inanimate objects.

Due to high technology and globalization, fast changes leading to great loss cannot be avoided. Therefore, more studies of numerals and classifiers especially in endangered languages should be seriously encouraged, otherwise it might be too late.

Acknowledgements

We would like to express our gratitude to the Thailand Research Fund (TRF) for funding the Karen Linguistics Project (2009-2012). We are thankful for the kind co-operation and hospitality of the local authorities and our Karen friends who have helped make our research project possible. Many thanks go to my research assistants, Sujinat Jitwiriyant, Karnthida Kerdpol and Siwaporn Tuantaisong for their assistance in many different ways.

References

- Benedict, Paul K. 1972. *Sino-Tibetan: a conspectus*. New York: Cambridge University Press.
- Bradley, David. 1997. Tibeto-Burman languages and classification. *Papers in Southeast Asian linguistics* 14, David Bradley (ed.), pp 1-71. Canberra: Pacific Linguistics.
- Bradley, David. 2005. Why do numerals show ‘irregular’ correspondence patterns in Tibeto-Burman? Some Southeastern Tibeto-Burman examples. *Cahiers de linguistique-Asie orientale*, vol. 34, no. 2, 2005: 221-238. (http://www.persee.fr/web/revues/home/prescript/article/clao_0153-3320_2005_num_34_2_1736).
- Bumrungkiri, Watchariya. 2002. *The phonological study of Kayaw language*. MA thesis, Mahidol University, Bangkok.
- Burling, Robbins. 1969. *Proto Karen: a reanalysis*. Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics, A.L. Becker (ed.), Publications of the Department of Linguistics, the University of Michigan.
- Comrie, Bernard. 2011. *Chapter 131. Numeral bases*. The world atlas of language structures online (<http://wals.info/feature/description/131>).
- Culy, Martin M. 1993. *A preliminary investigation of Pwo Karen dialects of Northern Thailand*. Chiang Mai: Payap University, PRDI Report 118.
- Grozdanović, Jadranka (ed.). 1999. *Numeral types and changes worldwide*. Berlin: Mouton de Gruyter.
- Haudricourt, André-Georges. 1972. Two-way and three-way splitting of tonal systems in some Far-eastern languages, [Christopher Court’s trans. of Haudricourt, 1961], in *Tai phonetics and phonology*, pp. 58-86. Bangkok: Central Institute of English language.
- Henderson, Eugénie J.A. 1997. *Bwe Karen dictionary with texts and English-Karen word list*, vol.2, Anna J. Allott (ed.). London: School of Oriental and African Studies.
- Hodson, T.C. 1913. Note on the numeral systems of the Tibeto-Burman dialects. *Journal of the Royal Asiatic Society of Great Britain and Ireland*, April 1913: 315-336. Cambridge University Press (<http://www.jstor.org/stable/25188969>).
- Hsa Eh Ywar and Ken Manson. 2009. *Classifiers in Kayan*, paper presented at the 42nd ICSTLL, Payap University, Chiangmai, November 2009.
- Hsar Shee. 2008. *A descriptive grammar of Geba Karen*. MA Thesis, Payap University, Chiangmai.
- Hurford, James R. 1999. Artificially growing a numeral system, in *Numeral types and changes worldwide*, Jadranka Gvozdanović (ed.), pp. 7-41. Berlin: Mouton de Gruyter.
- Jones, Robert B. 1961. *Karen linguistic studies: description, comparison and texts*. Berkeley: University of California Press.
- Karenni Literature Department. 1994. *The modern Western Kayah Li-English lexicon*. Chiang Mai: Payap University and the Summer Institute of Linguistics (SIL).
- Kauffman, William G. 1993. *The great tone split and Central Karen*. MA thesis, the University of North Dakota.
- Khammuang, Sarinya. 1998. *The phonological study of Padaung (Long-Necked-Karen) at Ban Naisoi, Tambol Pang Mu, Muang district, Mae Hong Son province*, MA thesis, Mahidol University, Bangkok.
- LaPolla, Randy J. 2001. The role of migration and language contact in the development of the Sino-Tibetan language family, in *Areal diffusion and genetic inheritance: case studies in language change*, R.M.W. Dixon and A.Y. Aikhenvald (eds.), pp. 225-254. Oxford: Oxford University Press.
- Manson, Ken. 2007. *Pekon Kayan phonology*. Research number#311, Department of Linguistics, Payap University, Chiangmai. (ms.).

- Manson, Ken. 2009a. *The classification of Karen languages*. Paper presented at the 42nd ICSTLL, Payap University, Chiangmai, November 2009.
- Manson, Ken. 2009b. *Prolegomena to reconstructing Proto-Karen*. La Trobe papers in linguistics, vol. 12: 1-26.
- Manson, Ken. 2011. *The subgrouping of Karen*. Paper presented at SEALS 21, organised by Kasetsart University, Bangkok, May 11 – 13, 2011.
- Matisoff, James A. 1983. *Linguistic diversity and language contact, in Highlanders of Thailand*, J. MacKennon and Wanat Bhruksasri (eds.), pp. 56-86. Kuala Lumpur: Oxford University Press.
- Matisoff, James A. 1997. *Sino-Tibetan numeral systems: prefixes, protoforms and problems* [Pacific Linguistics, Series B-114]. Canberra: Research School of Pacific and Asian Studies, the Australian National University.
- Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: system and philosophy of Sino-Tibetan reconstruction*. Berkeley: University of California Press.
- Mazaudon, Martine. 2008. *Number building in Tibeto-Burman languages*, NEILS 2 (North-East India Languages Symposium), Gauhati (Assam, India), 5-9 février 2007. [halshs-00273445, version 1-15 Apr 2008, 22 pages].
- Michaud, Alexis. 2011. The tones of numerals and numeral-plus-classifier phrases: on structural similarities between Naxi, Na and Laze. *Linguistics of the Tibeto-Burman area*, vol. 34.1: 1-26.
- Phillips, Audra. 1996. *Dialect comparison among the Pwo Karen of Central Thailand*, Report II. Bangkok: TU-SIL-LRDP.
- Przyluski, J. and G.H. Luce. 1931. The number “a hundred” in Sino-Tibetan. *Bulletin of the School of Oriental and African Studies*, University of London, vol. 6, no. 3 (1931). Cambridge University Press (<http://www.jstor.org/stable/607200>).
- Ratanakul, Suriya. 1986. *Thai-Sgaw Karen dictionary*. Bangkok: Mahidol University.
- Ratanakul, Suriya. 1997. Numeral classifiers in Sgaw Karen. *Mon-Khmer studies* 28: 101-113.
- Shintani, Tadahiko. 2003. Classification of Brakaloungic (Karenic) languages in relation to their tonal evolution, in *Cross-Linguistics studies of tonal phenomena: historical development, phonetics of tone and descriptive studies*, Shigeki Kaji (ed.), pp. 37-56. Tokyo: Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Solnit, David. 1997. *Eastern Kayah Li: grammar, texts, glossary*. Hawaii: University of Hawaii Press.
- Thanamteun, Orranat. 2000. *A phonological study of Pa-O (Taungthu) at Ban Huay Salop, Tambon Huay Pha, Muang district, Mae Hong Son province*. MA thesis, Mahidol University, Bangkok.
- U Shwe, Matthias. 1999. *Kayan-English dictionary [KAYĀN-ENGLE-KATAN NGÒ ATABAU TABĀN LĪ (DICTIONARY)]*. Taunggyi: St. Xavier Printing Press.

Appendix 1

List of abbreviations

AP	=	Audra Phillips
C	=	one of the initial consonants
CI	=	Chommanad Intajamornrak
CK	=	Central Karen
CLF	=	classifier
DB	=	David Bradley
DIC	=	Dictionary
DS	=	David B. Solnit
E. Kayah	=	Eastern Kayah Li
EH	=	Eugénie J.A. Henderson
HS	=	Hsar Shee
JM	=	James A. Matisoff
K	=	one of the PK final stops
KM	=	Ken Manson
MC	=	Martin M. Culy
N	=	one of the PK final nasals
NK	=	Northern Karen
NP	=	noun phrase
NUM	=	numeral
OT	=	Orranat Thanamteun
PB	=	Paul K. Benedict
PK	=	Proto-Karen
PP	=	Pittayawat Pittayaporn
PSETB	=	Proto-Southeastern-Tibeto-Burman (Proto-Loloish)
PT	=	Phanintra Teeranon
PTB	=	Proto-Tibeto-Burman
RJ	=	Robert B. Jones
SK	=	Sarinya Khammuang
SK	=	Southern Karen
SR	=	Suriya Ratanakul
T	=	one of the PK tones
TB	=	Tibeto-Burman
TL	=	Theraphan Luangthongkum (L-Thongkum)
V	=	one of the vowels
VP	=	verb phrase
W. Kayah	=	Western Kayah Li
WB	=	Watchariya Bumrungkiri

Appendix 2

PK tones and tonal developments in S. Pa-O, Kayan, E. Kayah, Kayaw, N. Pwo and S. Sgaw

S. Pa-O

Initial consonant	*A	*B	*D
1	31ː	33	21 /33
2			
3	53	55	45

Kayan

Initial consonant	*A	*B	*D
1	53	11	45
2			
3	33		21 /33

Mae Sot
(From Pa-An, Mon State)

Huay Suea Thao
(From Loikaw, Karenni State)

E. Kayah

Initial consonant	*A	*B	*D
1	33	11	55
2			
3			33

Kayaw

Initial consonant	*A	*B	*D
1	55	11	33
2			
3	33		

Huay Suea Thao
(Mueang, Mae Hong Son)

Huay Suea Thao
(From Hue Ya, Karenni State)

N. Pwo

Initial consonant	*A	*B	*D
1	35	33	45
2	55		
3		11ː	21 /33

S. Sgaw

Initial consonant	*A	*B	*D
1	55	45	21 /33
2			
3	33	31	11ː

Dong Dam
(Li, Lampun)

Pa La U
(Hua Hin, Prachuap Khiri Khan)

1 = *ph, *hm, *s. etc. 2 = *p, *b, *m, etc. 3 = *b, *m, *w, etc.

Proto-Karen (*k-rjaN^A) Fauna

Theraphan Luangthongkum

Department of Linguistics, Chulalongkorn University, Bangkok

Theraphan.L@chula.ac.th

Abstract

A wordlist consisting of 2,000 items with English and Thai glosses divided into 21 sections based on semantic fields was devised. The data on the six Karenic languages spoken in Thailand, i.e. Pa-O (two varieties), Kayan, Kayah, Kayaw, Pwo (two varieties) and Sgaw (two varieties) was collected by the author at many research sites in Thailand from January 2009 – January 2012. They represent Northern, Central and Southern Karen languages. The cognate words found in the Bwe Karen Dictionary (Henderson, 1997) were added. To analyse the patterns of sound correspondence, the comparative method was applied with an awareness of areal linguistic features due to language contact. A Proto-Karen (*k-rjaN^A) phonology and lexicon were reconstructed. With regard to the Proto-Karen phonology, the relationship between the onsets and tones is quite straightforward, so it is not difficult to reconstruct, unlike the rhymes which, in some cases, are problematical resulting from vowel harmony and the loss of final consonants. The 73 reconstructed forms of Proto-Karen animals are presented and the Proto-Karen tones, onsets and rhymes are provided in the Appendix.

Introduction

The Karenic languages are spoken in Myanmar and Thailand. There are eighteen different groups of Karen in Myanmar: Pa-O, Lahta, Kayan, Bwe, Geko, Geba, Brek, Western Kayah, Eastern Kayah, Yinbaw, Yintale, Manumanaw, Paku, Sgaw, Wewaw, Zayein, Eastern Pwo and Western Pwo (Ethnologue: 16th Edition, 2009, SIL International). In Thailand, only six groups are known of: Pa-O, Kayan, Kayah, Kayaw, Sgaw and Pwo. Among Burmese refugees and labourers, there may also be other groups. The total population of Karen is still debatable due to the lack of good official records; however, 4.5 million seems to be a reasonable estimation.

With regard to the reconstruction of a Proto-Karen phonology and lexicon, there have been a few attempts: Haudricourt, 1946 and 1953; Jones, 1961; Burling, 1969; Solnit, 2001; Manson, 2009). A literature review of the previous works on Karen comparative and historical linguistics can be found in Manson (2009). The classifications of Karen have been done by Jones (1961), Burling (1969), Kauffman (1993), Bradley (1997), Manson (2001) and Shintani (2003). Detailed information on the external and internal classifications of the Karenic languages can also be found in Manson (2009 and 2011). Among these classifications, I have adopted Kauffman's because of its geographical base which quite well suits the starting point of my present research work. Even though Proto-Karen was reconstructed earlier by a few linguists and more data on many Karenic languages is available at present, I still feel that I should start from zero. An attempt to do a comparative study using fresh data solely collected by one experienced field linguist may be able to help obtain a better solution since the data is equal in quality and, thus, compatible.

Data

Funded by the Thailand Research Fund (TRF), the data on the six Karenic languages spoken in Thailand was collected at many research sites in Thailand from January 2009 – January 2012 with a devised wordlist of 2,000 items. The English and Thai glosses were divided into twenty-one sections based on semantic fields: action verbs, stative verbs, bodyparts and body secretions, health and disease, fauna, parts of plants, flora, natural objects and phenomenon, manmade objects and construction, food stuffs, culture and society, kinship terms, numerals, classifiers, measurements, colour terms, time, direction and location, pronouns, question words and miscellaneous. My personal corpus used for the reconstruction of

a Proto-Karen phonology and lexicon consists of Northern Karen: two varieties of Pa-O; Central Karen: Kayan, Eastern Kayah, Kayaw; Southern Karen: two varieties of Sgaw and two varieties of Pwo. Clear cognates were selected for diachronic comparison. The cognate words found in the Bwe Karen Dictionary (Henderson, 1997) were added; therefore, altogether there are four Central Karen languages, i.e. Kayan, E. Kayah, Bwe (Blimaw) and Kayaw. The comparative method was applied with an awareness of language contact and areal linguistic features when analysing the patterns of sound correspondence.

With regard to the Proto-Karen (PK) phonology, the relationship between the onsets and tones is quite straightforward. Proto-Karen has three categories of initial consonants: Class-one consonants (voiceless aspirated stops, voiceless nasals, approximants and fricatives), Class-two consonants (voiceless unaspirated stops, preglottalised voiced stops or implosives, preglottalised nasals and approximants) and Class-three consonants (plain voiced stops, nasals and approximants). See the reconstructed PK three tones (*A *B *D) and onsets (*C- *CC-) in Figure 1 and Table 1, respectively, in the Appendix. The reconstruction of PK rhymes is more problematical resulting from vowel harmony and the loss of final consonants. The examples of the PK rhymes can be found in Table 2 of the Appendix. I have never had the opportunity of reading Haudricourt's original papers on comparative Karen written in French (Haudricourt 1946, 1953 and 1961), however, from the secondary sources, e.g. Matisoff's additional notes in Sino-Tibetan: A conspectus (Benedict, 1972), Luce (1959), Henderson (1979), Court (1972) and so on, I do appreciate Haudricourt's work on Proto-Karen the most.

*k-rjaN^A 'Karen'

Based on the autonames kaꞑjaŋ^{5 3} (Kayan), kɛ¹¹jɛ¹¹ (E. Kayah), kɔ³³jɔ³³ (Kayaw), (pɣa³³) ka ꞑɔ³³ (Sgaw) and names known among Thai people, i.e. Kariang, Karang (Central Thai) and Yang (Northern Thai), the protoform *k-rjaN^A was reconstructed. This etymon has regular sound changes like the other etyma with the *-aŋ rhyme; for example,

Ex.1 PK: *ꞑbaŋꞑ 'bamboo shoot'

Pa-O: baŋ^{5 5} B2 (N.), baŋ³³ B2 (S.)
 Kayan: baŋ¹¹ B
 Kayah: bɛ¹¹ B
 Bwe: ɓa³³ B
 Kayaw: bɔ¹¹ B
 Sgaw: bɔ^{31~} B2 (N.), bɔ^{4 5} B2 (S.)
 Pwo: bɛ³³ B2 (N.), bɔ^{5 5} B2 (S.)

Ex.2 PK: *laŋꞑ 'to descend'

Pa-O: laŋ³³ A3 (N.), laŋ^{5 3} A3 (S.)
 Kayan: laŋ³³ A3
 Kayah: lɛ¹¹ A3
 Bwe: la³³ A3
 Kayaw: lɔ³³ A3
 Sgaw: lɔ³³ A3
 Pwo: lɔ³³ A3 (N.), lɔ³³ A3 (S.)
 lɛ^{5 5} A3 (N.), lɔ³¹ A3 (S.)

Ex.3 PK: *khaŋꞑ 'foot, leg'

Pa-O: khaŋ^{5 5} B1 (N.), khaŋ³³ B1 (S.)
 Kayan: haŋ¹¹ B
 Kayah: khɛ¹¹ B
 Bwe: kha³³ B
 Kayaw: khɔ¹¹ B
 Sgaw: khɔ^{31~} B1 (N.), khɔ^{4 5} B1 (S.)
 Pwo: khɛ³³ B1 (N.), khɔ^{5 5} B1 (S.)

Ex.4 PK: *thaŋꞑ 'to ascend'

Pa-O: thaŋ^{5 5} B1 (N.), thaŋ³³ B1 (S.)
 Kayan: thaŋ¹¹ B
 Kayah: thɛ¹¹ B
 Bwe: tha³³ B
 Kayaw: hɔ¹¹ B
 Sgaw: thɔ^{31~} B1 (N.), thɔ^{4 5} B1 (S.)
 Pwo: thɛ³³ B1 (N.), thɔ^{5 5} B1 (S.)

Tonal development

How many tones should be reconstructed, two, three, four or six? There are different opinions, as follows:

Haudricourt (1946, 1975): two tones (*A *B, *A *B *Bꞑ in smooth syllables)

Jones (1961): two tones (' `)

Burling (1969): six tones (1 2 3 4 ^{□ □} 1 2)

Shintani (2003): four tones (1 2 2' 3)

Manson (2009): four tones (*A *B *B' *C)

To support my view that three tones, namely *A *B *D, should be reconstructed, the tone systems of Modern Karenic languages and their development will be discussed. Both of the Pa-O varieties have four tones in the smooth syllable and two tones in the checked syllable: 31 (A12), 33 (A3), 55 (B12), 53 (B3) and 21 (D12), 45 (D3) in Northern or Highland Pa-O; and 31 (A12), 53 (A3), 33 (B12), 55 (B3), 21 (D12) and 45 (D3) in Southern or Lowland Pa-O. This means that each of the proto-tones (*A *B *D) has split into two tones, i.e. *A > A12-3 or A1-23 *B > B12-3 and *D > D12-3 as shown in Figure 1 of the Appendix.

With regard to Central Karen, synchronically, Kayan and E. Kayah have four tones in the non-checked syllable and two tones in the checked one like Pa-O. However, the patterns of tone split and merger differ. It is also noticeable that the mid-falling tone in Kayan and E. Kayah rarely occurs and when it does, more often than not, the words turn out to be Burmese or Shan loanwords. Among the cognates, this tone has not been found. Perhaps, the mid-falling is a new-born tone in Kayan and E. Kayah. In Kayan, E. Kayah, Bwe and Kayaw, the *A tone has split into two tones, i.e. A12-3, and there is no split in the *B tone column. The *D tone in Kayan, E. Kayah and Bwe has split into *D12-3 while there is no split in the *D tone column in Kayaw. There is a merger between *A3 and *B in E. Kayah and Bwe. In Bwe, interestingly, there is also a merger between *B and *D12, more details can be found in Figure 1. While working with my E. Kayah and Kayaw language consultants at Huai Suea Thao village in Mae Hong Son province, I noticed that their low tones always occurred with a breathy-voiced phonation and that, in the stressed position, the CV syllable always became CV[□], or in other words, the -[□] disappeared in the unstressed position no matter what tone that syllable had.

As for Sgaw, I worked on many varieties of the N. Sgaw especially those spoken in the provinces of Chiangrai, Chiangmai, Lamphun and Mae Hong Son. The varieties of S. Sgaw are the ones spoken in the provinces of Petchaburi and Prachuap Khiri Khan. Since the Sgaw Karen in these areas have lived together with the Pwo Karen (Phlong, Phlou), language contact cannot be avoided. Resulting from this fact, the Pa La-U Sgaw has four tones: 55 (A12), 33 (A3), 45' (B12), 31 (B3), 21 (D12) and 11 (D3), see Sgaw (type 4) in Figure 1. Based on my fieldnotes, four patterns of tonal development have been found: *type 1*, no split in the *A and *B columns while *D > D12-3; *type 2*, no split in the *A column while *B > B12-3 and *D > D12-3; *type 3*: *A > A12-3 and *D > D12-3 while no split in the *B column; and *type 4*: there is a split in every column, i.e. *A > A12-3, *B > B12-3 and *D > D12-3. See the illustration of the patterns of tone split and merger in Figure 1.

Like Pa-O, all Pwo varieties have four tones with an unusual split in the *A column, i.e. *A > A1-23 while the split patterns in the *B and *D columns are typical, i.e. *B > B12-3 and *D > D12-3. In figure 1, Pwo (1) is Northern Pwo while Pwo (2) is Southern Pwo or Western Pwo (Phillips, 2002). The Pwo (2) varieties are mostly spoken in the western areas of Central Thailand, especially in the provinces of Kanchanaburi, Uthai Thani, Ratchaburi, Petchaburi and Prachuap Khirikhan. In the *D tone columns of Pwo (1) and Pwo (2), there is a tonal flip-flop: *D > D12 (high) – 3 (low) in Pwo (1) but *D > D12 (low) – 3 (high) in Pwo (2), as shown in Figure 1.

Fauna

Even though many protoforms in the twenty-one semantic fields (see details in the introductory part) were reconstructed, only the 73 etyma related to fauna or animals are presented in this paper. They are arranged in an alphabetical order A-Z started with the English gloss no. 1 'ant' and ending with no. 73 'weevil'. For each etymon, following the English gloss, is the reconstructed protoform. In some cases, the protoforms of TB cognates drawn from Benedict (1972) or Matisoff (2003) are also given. Additional information is provided under *Note*.

1. ANT *^{၁၁၅}

Pa-O: ^{၁၁၅} (N.), ^{၁၅} (S.)

Kayan: ^{၁၁}

Kayah: tɔ¹¹

Bwe: do³³

Kayaw: tə¹¹

Sgaw: tə¹¹ (N.), tə³¹ (S.)

Pwo: tha¹¹ (N.), tha³³ (S.)

Note: The plain voiced *d- was reconstructed because of the onset d- in Western Bwe (Blimaw, Geba) which has /d/ vs /d/ (Henderson, 1997). It is also noticeable that PK *d- or *d- is always d- in Modern Karen (see no. 30 ‘frog’), while *d- has become t- or th-.

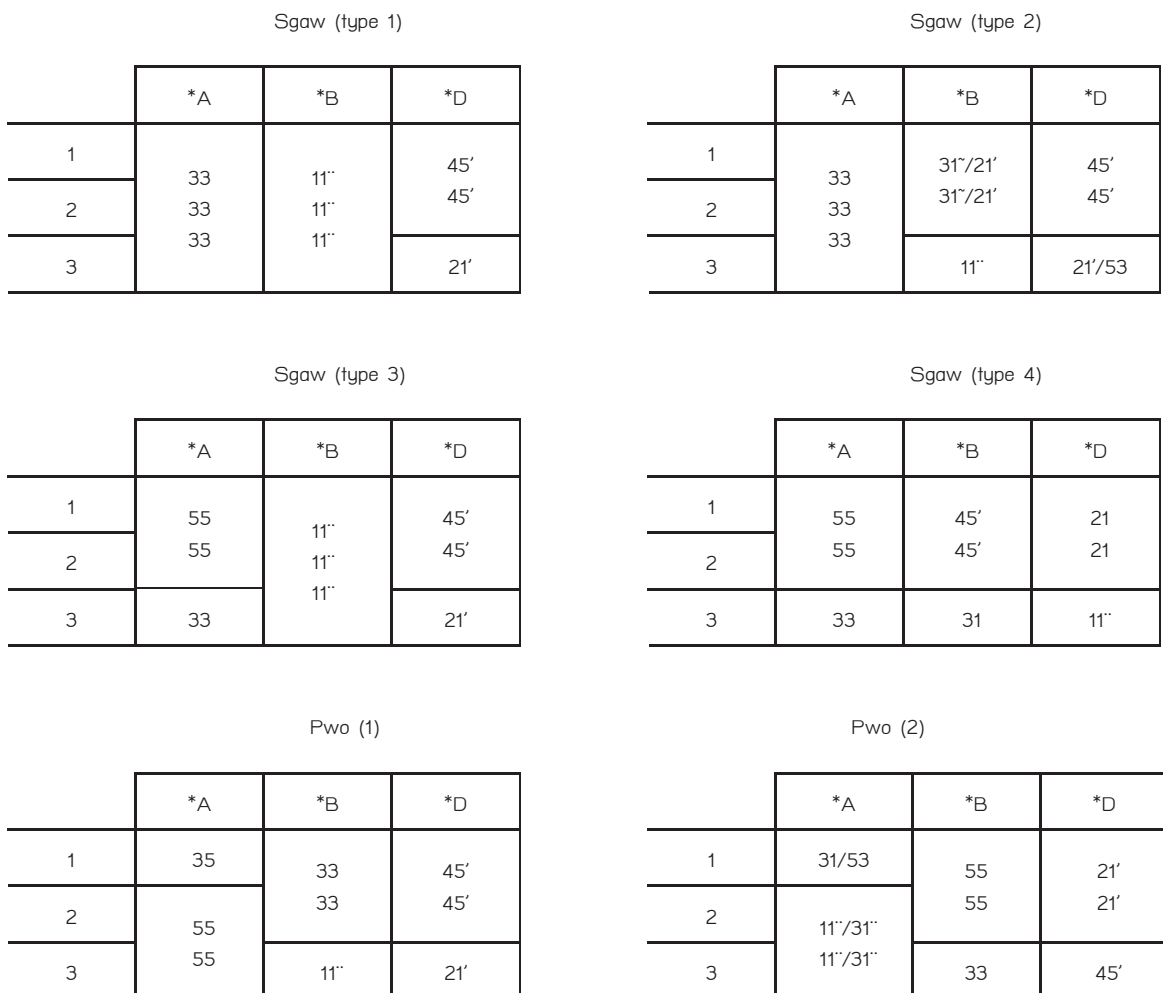


Figure 1 Development of Sgaw and Pwo tones

2. BAMBOO RAT *(jow□) khan□

Pa-O: (ju^{5 3}) khan^{31ˈˈ} (N.), - (S.)

Kayan: (ʃu¹¹) khaŋ^{5 3}

Kayah: -

Bwe: khɛ^{5 5}

Kayaw: (ju¹¹) khɔ̃^{5 5}

Sgaw: khɔ̃³³ (N.), khɔ̃^{5 5} (S.)

Pwo: khɛ̃^{3 5} (N.), khɔ̃^{5 3} (S.)

Note: The Karen regard this rodent as a kind of rat or mouse (*jowꨀ, see no. 47) as can be seen in Pa-O, Kayan and Kayaw. See also no.43 ‘mole’.

3. BAT *plaꨀꨀ

Pa-O: pla^{31ˈˈ} (N., S.)

Kayan: bla¹¹

Kayah: ple¹¹, ple³¹

Bwe: plɛ¹¹

Kayaw: pla¹¹

Sgaw: bla³³ (N.), bla^{5 5} (S.)

Pwo: phla³³ (N.), phla^{5 5} (S.)

Note: Pa-O and Sgaw have tone A while Kayan, Kayah, Kayaw and Pwo have tone B. In this etymon, the PK cluster *pl- has become bl- in Kayan and Sgaw but phl- in Pwo. The low tone (11) in Bwe suggests *D. The initials and tones of the word ‘bat’ in Modern Karen seem irregular. I suspect that its cause is language contact.

4. BEAR *thamꨀ

Pa-O: tham^{31ˈˈ} (N., S.)

Kayan: thaŋ^{5 3}

Kayah: the³³

Bwe: thɛ^{5 5}

Kayaw: thɔ̃^{5 5}

Sgaw: - (N., S.)

Pwo: (phĩ^{11ˈˈ}) thɔ̃^{3 5} (N.), (phou^{5 3}) thɔ̃^{5 3} (S.)

Note: This etymon has been lost in all of the Sgaw varieties. It was replaced by the word $ta^{33} su^{33}$ or $ta^{33} \theta u^{5\ 5}$ which can be a euphemism for ‘bear’. There are curses in Sgaw, for example, “ $ta^{33} su^{33} wa^{4\ 5} phYi^{33} n\theta^{4\ 5} m\epsilon^{4\ 5} sa^{11}$ ”, meaning ‘(I wish) a bear hits your face with its paws’. The generic term for ‘bear’ maybe regarded as a taboo word.

5. BEDBUG *gram

Pa-O: sam^{33} (N.), $sam^{5\ 3} \sim sa\eta^{5\ 3}$ (S.)

Kayan: $ca\eta^{33}$

Kayah: $khre^{11}$

Bwe: -

Kayaw: $c\theta^{33}$

Sgaw: $x\theta^{33}$ (N.), $x\theta^{33}$ (S.)

Pwo: - (N., S.)

Note: This etymon was replaced by $ph\tilde{\theta}^{5\ 3}$ (S. Pwo) and $ph\tilde{a}^{3\ 5}$ (N. Pwo).

6. BEE (Apis cerana) *kwat

Pa-O: wat^{21} (N.), wat^{21} (S.)

Kayan: $hw\epsilon^{4\ 5}$

Kayah: $w\epsilon^{5\ 5}$

Bwe: $w\epsilon^{33}$

Kayaw: $hw\epsilon^{33} \sim khw\epsilon^{33}$

Sgaw: $kw\epsilon^{4\ 5}$, $kw\epsilon^{5\ 5}$ (N.), $kw\epsilon^{5\ 5}$ (S.)

Pwo: $kw\epsilon^{5\ 5}$ (N.), $kw\epsilon^{31''}$ (S.)

Note: Pwo, S.Sgaw and some varieties of N.Sgaw have a smooth syllable with tone A instead of a checked syllable with tone D.

7. BEE (Apis dorsata) *k-hne

Pa-O: $ne^{31''}$ (N., S.)

Kayan: $nai^{5\ 3}$

Kayah: ni^{33}

Bwe: $g\theta-nI^{33}$

Kayaw:	ni ^{5 5}
Sgaw:	kə ¹¹ nɛ ³³ , kə ¹¹ nɛ ^{5 5} (N.), nɛ ^{5 5} (S.)
Pwo:	ni ^{3 5} (N.), ni ^{5 3} (S.)

8. BIRD *tho□

Pa-O:	- (N.), - (S.)
Kayan:	thau ¹¹
Kayah:	thu ¹¹
Bwe:	tho ³³
Kayaw:	thu ¹¹
Sgaw:	tho ^{31~} , tho ¹¹ (N.), thu ^{4 5} (S.)
Pwo:	thu ³³ (N.), thu ^{5 5} (S.)

Note: In Pa-O, it is wa^{5 3} (N.) or wa^{5 5} (S.). This can be regarded as a retention from TB *wa (PB, JM) since the Pa-O branched out earlier than the rest of the Karenic peoples. It can also be interpreted as a loanword from the other TB language.

9. BOAR (wild~) *thɔ□□ mi□

Pa-O:	thɔ□ ²¹ mi ³³ (N.), thɔ□ ²¹ mi ^{5 3} (S.)
Kayan:	thau□ ^{4 5} mi ³³
Kayah:	thɛ ^{5 5} mi ¹¹
Bwe:	-
Kayaw:	tho ³³ mi ³³
Sgaw:	thɔ□ ^{4 5} mi ³³ (N.), thɔ□ ²¹ mi ³³ (S.)
Pwo:	tho□ ^{4 5} mei ^{5 5} (N.), thu□ ²¹ mɛi ³¹ , thu□ ²¹ mɛi ^{31..} (S.)

Note: See also no.53 'pig'.

10. BUFFALO (water~) *p/b-na□

Pa-O:	pa□ ²¹ na ^{5 3} (N.), pa□ ²¹ na ^{5 5} (S.)
Kayan:	bu□ ^{4 5} na ¹¹
Kayah:	pe ^{5 5} ne ¹¹
Bwe:	bə ³³ ne ³³

Kayaw: pa³³ na¹¹

Sgaw: pə³³ na¹¹ː, pa²¹ na¹¹ː, pə^{4 5} na¹¹ː (N.), pə¹¹ na³¹ (S.)

Pwo: pa²¹ nɛ¹¹ː (N.), pə²¹ na³³ (S.)

Note: The reconstructed form *bo²¹ na²¹ is also possible because *bo²¹ was reconstructed for ‘ox’ (see no. 48). The meaning of *bo²¹ could be ‘bovine’ or ‘cattle’.

11. CATERPILLAR *si^D

Pa-O: se²¹ (N.), si²¹ (S.)

Kayan: θi^{4 5}

Kayah: si^{5 5}

Bwe: θa³³ (?)

Kayaw: si³³

Sgaw: si^{4 5}, si^{4 5} (N.), sai²¹ (S.)

Pwo: si^{3 5} (N.), θei³³, θei²¹ (S.)

Note: N.Pwo and some varieties of S.Pwo have tone A1 instead of D1. There is also an irregular vowel correspondence.

12. CATFISH *~ku²¹

Pa-O: khu³¹ː (N., S.)

Kayan: ku^{5 3}

Kayah: ku¹¹

Bwe: -

Kayaw: ku¹¹

Sgaw: - (N.), ku^{4 5}ː (S.)

Pwo: - (N.), ku^{5 5} (S.)

Note: Pa-O and Kayan have tone A while Kayah, Kayaw, S.Sgaw and S.Pwo have tone B.

13. CENTIPEDE *t/da²¹-baŋ

Pa-O: ta^{4 5} baŋ³¹ː (N., S.)

Kayan: ta¹¹ baŋ^{5 3}

Kayah: -

Bwe:	-
Kayaw:	-
Sgaw:	da ^{4 5} bɔ ³³ , da ^{4 5} bɔ ^{5 5} (N.), da ¹¹ bɔ ^{5 5} (S.)
Pwo:	-(N.), da ^{4 5} bɔ̃ ³¹ , da ^{4 5} bɔ̃ ¹¹ (S.)

Note: *baŋ means 'yellow'.

14. CHAMELEON, LIZARD *kwi

Pa-O:	kwi ^{5 5} (N.) kwi ³³ (S.)
Kayan:	khwi ¹¹
Kayah:	khwi ¹¹
Bwe:	khwi ³³
Kayaw:	khi ¹¹
Sgaw:	khwi ^{31~} , khwi ¹¹ (N.), khwi ^{4 5} (S.)
Pwo:	khwi ³³ (N.), khwi ^{5 5} (S.)

Note: See also no.44 and no.45 'monitor lizard'.

15. CHICKEN *chjaX, *chjaN

Pa-O:	ɕja ^{31~} (N., S.)
Kayan:	ɕi ^{5 3}
Kayah:	cha ³³
Bwe:	ɕi ^{5 5}
Kayaw:	ɕi ^{5 5}
Sgaw:	chɔ ³³ , chɔ ^{5 5} (N.), chɔ ^{5 5} (S.)
Pwo:	chɛ̃ ^{3 5} (N.), chɔ̃ ^{5 3} (S.)

Note: The final of this etymon cannot be reconstructed. The dummy *-X became *-N and then nasalised vowels in Pwo but was lost in the other Karenic languages. The pattern of vowel correspondence is rather unusual. If PK had the *-aN rhyme, the normal correspondence would be -aN (Pa-O, Kayan), -e (Kayah), -ɛ (Bwe), -ɔ (Kayaw, Sgaw), -ɛ̃ (N.Pwo) and -ɔ̃ (S.Pwo) as in no.2, no.4 and no.5. The second element of the initial cluster (*-j-) could also be the cause of vowel irregularity in this etymon. *chjaN could be another solution, i.e. it is a loanword from Proto-Monic *tyaaŋ (> *chyaaŋ in Proto-Nyah Kur) and >*cǝi J1 in Proto-Mon) as reconstructed by Diffloth (1984).

16. CICADA *ŋjaɰ

Pa-O:	ŋja ³³ (N.), - (S.)
Kayan:	ji ³³
Kayah:	ja ¹¹
Bwe:	-
Kayaw:	-
Sgaw:	je ³³ , ze ³³ (N.), - (S.)
Pwo:	jai ^{5 5} (N.), jai ³¹ , jai ¹¹ (S.)

Note: In some N. Sgaw varieties, *-j- has become z-. The voiced palatal fricative [ʃ] is, in fact, a variant of the onset j- in Karenic languages.

17. CIVET CAT *thu

Pa-O:	- (N.), - (S.)
Kayan:	thu ¹¹ (mi ³³)
Kayah:	-
Bwe:	thu ³³
Kayaw:	thu ¹¹ (mi ³³)
Sgaw:	- (N.), - (S.)
Pwo:	- (N.), <u>thu</u> ²¹ tho ³¹ , thu ⁵⁵ tho ¹¹ (S.)

Note: Sgaw people call 'civet cat' thɔ^{4 5} se³¹ or thɔ^{4 5} se¹¹ 'tree pig' (N.) or thɔ^{4 5} tu³³ pə¹¹ i^{5 5} 'sticky-rice pig' (S.).

18. CRAB *chwɛ

Pa-O:	chwɛ ^{5 5} (N.), chwɛ ³³ (S.)
Kayan:	chwɛ ¹¹
Kayah:	chwa ¹¹
Bwe:	ɲwɛ ³³
Kayaw:	chɛ ¹¹
Sgaw:	chwɛ ³¹ , chwɛ ¹¹ (N.), chwɛ ^{4 5} (S.)
Pwo:	chwe ³³ (N.), chwe ^{5 5} (S.)

19. CRICKET *s-ki

Pa-O:	ki ^{31..} (N.), ki ^{31..} (S.)
Kayan:	ki ⁵³
Kayah:	ki ³³
Bwe:	-
Kayaw:	dɛ ¹¹ ki ^{5 5}
Sgaw:	sə ¹¹ ki ³³ , θa ²¹ ki ³³ , sə ²¹ ki ³³ (N.), tə ¹¹ ki ^{5 5} (S.)
Pwo:	- (N., S.)

Note: N.Pa-O has tone A3 while the rest have tone A2. It is kha²¹ rai^{4 5} in S. Pwo which is not cognate.

20. CROCODILE *s-hma

Pa-O:	- (N.), - (S.)
Kayan:	-
Kayah:	-
Bwe:	θə ³³ mɛ ³³
Kayaw:	ma ¹¹
Sgaw:	sə ³³ ma ^{31~} , θa ³³ ma ¹¹ , sə ^{5 5} ma ¹¹ , sə ^{4 5} ma ¹¹ (N.), ma ^{4 5} (S.)
Pwo:	ma ³³ (N.), ma ^{5 5} (S.)

21. DEER (barking ~) *d-khej (TB *d-key, *d-kəy)

Pa-O:	khi ^{31..} (N., S.)
Kayan:	kh ⁱ ^{5 3}
Kayah:	khə ³³
Bwe:	do ¹¹ khi ^{5 5}
Kayaw:	khi ^{5 5}
Sgaw:	- (N.), - (S.)
Pwo:	- (N.), khi ⁵³ bɔ ^{31..} (S.)

Note: In Sgaw, the etymon *d-khej had been lost and was replaced by the word ta³³ho³³, ta³³ho^{5 5} (N. Sgaw).

22. DEER (sambha~) *t-khro (TB *d-yuk)

Pa-O:	khjo ²¹ (N.), kjo ²¹ (S.)
-------	---

Kayan:	khjɔ̌ ^{4 5}
Kayah:	khɾɔ̌ ^{5 5}
Bwe:	kho ³³
Kayaw:	khə̌ ³³
Sgaw:	ta ³³ xɔ̌ ³³ (N.), thɔ̌ ³³ xɔ̌ ⁵⁵ (S.)
Pwo:	ta ³³ xɔ̌ ^{5 5} (N.), chə̌ ³³ xo ³¹ ː, chə̌ ³³ xo ¹¹ (S.)

Note: Northern and Central Karen have tone D (D12) while Southern Karen (Sgaw, Pwo) has tone A (A3) which suggests a voiced initial.

23. **DOG *thwi** (TB *kwəy, *kwiŋ)

Pa-O:	thwi ^{5 5} (N.), thwi ³³ (S.)
Kayan:	thwi ¹¹
Kayah:	thwi ¹¹
Bwe:	thwi ³³
Kayaw:	thi ¹¹
Sgaw:	thwi ³¹ ː, thwi ¹¹ , chwi ³¹ ː, chwi ¹¹ (N.), thwi ^{4 5} ˊ (S.)
Pwo:	thwi ³³ (N.), thwi ^{5 5} (S.)

24. **DOVE *(tho) lwi**

Pa-O:	(wa ^{5 3}) lwi ^{5 3} (N.), (wa ^{5 5}) lwi ^{5 5} (S.)
Kayan:	thau ¹¹ lwi ¹¹
Kayah:	thu ¹¹ lwi ¹¹
Bwe:	tho ³³ lwi ³³
Kayaw:	thu ¹¹ li ¹¹
Sgaw:	tho ³¹ ː lwi ¹¹ ː, tho ¹¹ lwi ¹¹ (N.), thu ^{4 5} ˊ lwi ³¹ (S.)
Pwo:	thu ³³ lei ¹¹ ː (N.), thu ^{5 5} lwi ³³ (S.)

25. **EARTHWORM *jɛ**

Pa-O:	jɛ̌ ^{4 5} (N., S.)
Kayan:	ɟɛ̌ ²¹ (ːcǎ ²¹ krɔ̌ ^{5 3})
Kayah:	ja ³³ (ːkhro ³³)

Bwe:	-
Kayaw:	je ³³ (tha ^{5 5} ~)
Sgaw:	- (N.), - (S.)
Pwo:	- (N.), - (S.)

Note: The first parts of the compounds in Sgaw and Pwo, i.e. tho^{4 5} or tho²¹; chɛ̃^{3 5} or chɔ̃^{5 3} mean ‘pig’ and ‘chicken’, respectively.

26. ELEPHANT *k-chaŋ (TB *tshaŋ, *tsaŋ)

Pa-O:	chaŋ ³¹ (N., S.)
Kayan:	chaŋ ^{5 3}
Kayah:	chɛ ³³
Bwe:	gə-ŋa ^{5 5}
Kayaw:	rɔ ¹¹ chɔ ^{5 5}
Sgaw:	kə ¹¹ chɔ ³³ , kə ¹¹ chɔ ^{5 5} (N.), kə ¹¹ chɔ ^{5 5} (S.)
Pwo:	ka ¹¹ chɛ̃ ^{3 5} (N.), kə ¹¹ chɔ̃ ^{5 3} (S.)

Note: This etymon is an Austroasiatic loanword. The etyma meaning ‘elephant’ were reconstructed as *ksaŋ in Proto-Waic (Diffloth, 1980) and *ciiŋ in Proto-Monic (Diffloth, 1984).

27. FISH *da (TB *da)

Pa-O:	tha ^{4 5} (N.), tha ^{4 5} (S.)
Kayan:	ta ²¹
Kayah:	te ³³
Bwe:	da ¹¹ (-pho ³³)
Kayaw:	tɔ ³³
Sgaw:	-
Pwo:	-

Note: This etymon has been kept in Northern and Central Karen but has been lost in Southern Karen. A shared innovation in Sgaw and Pwo for ‘fish’ is ‘meat (animal) -offspring’, i.e. ɲa³¹ pho³³, ɲa¹¹ pho³³ (N. Sgaw), ɲa^{4 5}, phu^{5 5} (S. Sgaw); ja³³ phu³³ (N. Pwo), ja^{5 5} phu^{5 5} (S. Pwo).

28. FLEA *kli (TB *s-liy, *s-ləy)

Pa-O:	khli ³¹ (N.), - (S.)
-------	---------------------------------

Kayan:	khli ^{4 5}
Kayah:	kla ^{5 5}
Bwe:	kle ³³
Kayaw:	kle ³³
Sgaw:	kli ^{5 5} (N., S.)
Pwo:	khlei ^{3 5} (N.), khlei ^{5 3} (S.)

Note: The tones of Northern Karen (Pa-O) and Southern Karen (Sgaw, Pwo) suggest PK *A while the ones of Central Karen (Kayan, Kayah, Bwe, Kayaw) suggest PK *D.

29. FOX, WOLF *thwi³³ mi³³

Pa-O:	thwi ^{5 5} mi ³³ (N.), thwi ³³ mi ^{5 3} (S.)
Kayan:	thwi ¹¹ mi ³³
Kayah:	thwi ¹¹ mi ¹¹
Bwe:	-
Kayaw:	thi ¹¹ mi ³³
Sgaw:	thwi ^{31~} mi ³³ , chwi ¹¹ mi ³³ (N.), thwi ^{4 5,} mi ³³ (S.)
Pwo:	- (N.), thwi ^{5 5} mɛi ^{31~} (S.)

Note: This etymon is a compound meaning 'wild dog': thwi or chwi (B12) 'dog' and mi (A3) 'wild'. See no. 23.

30. FROG *de³³

Pa-O:	de ^{5 5} (N.), de ³³ (S.)
Kayan:	dai ¹¹
Kayah:	di ¹¹
Bwe:	d̥i ³³
Kayaw:	di ¹¹
Sgaw:	de ^{31~} , de ²¹ (N.), de ^{4 5,} (S.)
Pwo:	di ³³ (N.), di ^{5 5} (S.)

Note: See no. 69.

31. GRASSHOPPER *dwɛ³³

Pa-O:	dwɛ ^{5 5} (N.), dwɛ ³³ (S.)
-------	---

Kayan:	dwɛ ¹¹
Kayah:	-
Bwe:	dɛ ³³
Kayaw:	dɛ ¹¹
Sgaw:	dwɛ ^{31~} , dwɛ ²¹ (N.), dwɛ ^{4 5} (S.)
Pwo:	thwe ^{11~} (N.), thwe ³³ (S.)

Note: There are some innovations in Pwo, i.e. PK *ɲd > *d > th and *B > B3 while *B>B12 in the other Karenic languages.

32. HAWK *lek

Pa-O:	le ^{4 5} (N.), lek ^{4 5} (S.)
Kayan:	li ²¹
Kayah:	la ³³
Bwe:	le ¹¹
Kayaw:	le ³³
Sgaw:	li ²¹ , li ^{4 5} (N.), lai ¹¹ (S.)
Pwo:	lai ²¹ (N.), lai ^{4 5} (S.)

33. HORNET *phrɪn

Pa-O:	phrɪm ^{31~} (N.), phrɪn ^{31~} (S.)
Kayan:	phrɪ ^{5 3}
Kayah:	phlɪ ³³
Bwe:	phlu ^{5 5} 'wasp'
Kayaw:	phrɪ ^{5 5}
Sgaw:	phlɔ ³³ , phlɔ ^{5 5} (N.), phlɪ ^{5 5} (S.)
Pwo:	phlɔ̃ ^{3 5} (N.), phlɔ̃ ^{5 3} (S.)

Note: In this etymon, *phr- remains phr- in Northern Karen (Pa-O) but becomes phl- in Southern Karen (Sgaw and Pwo). With regards to Central Karen, both clusters, phr- and phl- can be found.

34. HORSE *k-sre

Pa-O:	se ³¹ (N., S.)
-------	---------------------------

Kayan:	θi ^{4 5}
Kayah:	si ^{5 5}
Bwe:	θə-rI ³³ , θri ³³
Kayaw:	si ³³ ri ³³ (?)
Sgaw:	kə ³³ se ¹¹ , ka ²¹ θe ¹¹ (N.), kə ¹¹ se ^{4 5} (S.)
Pwo:	ka ²¹ si ³³ (N.), kə ¹¹ θi ^{5 5} (S.)

Note: This etymon is an Austroasiatic loanword. The reconstructed form for ‘horse’ in Proto-Monic is *ksɛh (Diffloth, 1984). It is worth pointing out that the Modern Karenic languages have different tones, i.e. tone A in Northern Karen (Pa-O), tone D in Central Karen (Kayan, Kayah, Bwe, Kayaw) but tone B in Southern Karen (Sgaw, Pwo). This fact suggests the idea that *-h was dropped in NK and SK but became -ɿ in CK.

35. LEECH (land ~) *s-wa²¹

Pa-O:	wa ²¹ (N.), wa ²¹ (S.)
Kayan:	θu ^{4 5}
Kayah:	swa ^{5 5}
Bwe:	-
Kayaw:	su ³³
Sgaw:	su ^{4 5} , θu ^{4 5} (N.), sɔu ²¹ (S.)
Pwo:	wa ²¹ (N.), wa ^{4 5} , βwa ^{4 5} (S.)

Note: Pa-O, Kayan, Kayah and Sgaw have tone D1-2 which suggests a voiceless onset, while Pwo has tone D3 which indicates a voiced one, so *sw- seems to be a good solution for both ways of tonal development, i.e. *D > *D12 due to the first voiceless element *s- and *D > D3 due to *w-, the onset of the root, which is a voiced sound.

36. LEECH (water~) *k/s-lej³³

Pa-O:	leu ³³ (N.), ljeu ^{5 3} (S.)
Kayan:	-
Kayah:	-
Bwe:	gə ³³ lI ³³
Kayaw:	su ¹¹ li ³³
Sgaw:	sə ¹¹ li ³³ , θa ³³ li ³³ , θə ¹¹ li ³³ (N.), tə ¹¹ li ³³ (S.)
Pwo:	lei ^{5 5} (N.), lei ³¹ , lei ¹¹ (S.)

Note: Pa-O seems to have a strange rhyme.

37. LEMUR (flying~) *p/ble

Pa-O:	phli ²¹ (N.), - (S.)
Kayan:	-
Kayah:	pla ³³
Bwe:	bli ³³
Kayaw:	ple ³³
Sgaw:	phli ^{4 5} , phli ²¹ (N.), plai ¹¹ (S.)
Pwo:	phlai ²¹ (N.), phlai ^{4 5} (S.)

Note: Tone D12 in N. Pa-O, Bwe and some N. Sgaw varieties suggests PK *pl- but tone D3 in Kayah, both N. and S. Pwo and also S. Sgaw indicates *bl-. It is noticeable that both *pl- and *bl- become phl- in Modern Karen while *phl- is likely to remain phl-.

38. LORIS *k-ch()N

Pa-O:	- (N.), - (S.)
Kayan:	chəŋ ^{5 3}
Kayah:	chɔ ³³
Bwe:	-
Kayaw:	chə ^{5 5}
Sgaw:	kə ¹¹ che ³³ (N.), ta ¹¹ chi ^{5 5} (S.)
Pwo:	chāi ^{3 5} (N.), chēi ^{5 3} , kə ¹¹ chēi ^{5 3} (S.)

Note: The rhyme of this etymon is problematical, due to the lack of data from Pa-O.

39. LOUSE, INSECT *gra

Pa-O:	sa ^{5 3} (N.), ɕa ^{5 5} (S.)
Kayan:	ca ¹¹
Kayah:	khre ¹¹
Bwe:	-
Kayaw:	ca ¹¹
Sgaw:	xa ¹¹ (N.), xa ³¹ (S.)
Pwo:	xa ¹¹ (N.), xa ³³ (S.)

Note: In Modern Karen, this etymon never occurs alone but always in compounds, such as chɔ̃^{5 3} xa³³ ‘chicken louse’ (S.Pwo), ta¹¹ pho³³ xa¹¹ ‘insect’ (N. Sgaw) and so forth.

40. LOUSE (head~) *swiX^B, *swiN^B (TB *sar, śar)

Pa-O:	sɪ̃ ^{5 5} (N.), sɪ̃ ³³ (S.)
Kayan:	θi ¹¹
Kayah:	so ¹¹
Bwe:	θo ¹¹
Kayaw:	sɪ̃ ¹¹
Sgaw:	su ³¹ ~, su ¹¹ (N.), su ^{4 5} (S.)
Pwo:	sã ³³ (N.), θãi ^{5 5} (S.)

Note: Similar to the etymon ‘chicken’ (no. 15), the reconstructed final of the etymon ‘head louse’ is the dummy *-X. This PK *-X became *-N and then –Ṽ∅ in Pwo Karen while it was dropped in the other Karenic languages. The second element of the cluster onset (-w-) causes the lip rounding and backness of the vowels in Kayah, Bwe, Sgaw and N. Pwo.

41. MAGGOT *hloŋ

Pa-O:	loŋ ^{5 5} (N.), loŋ ³³ (S.)
Kayan:	lo ¹¹
Kayah:	lɔ̃ ¹¹
Bwe:	lo ³³
Kayaw:	lɔ̃ ¹¹
Sgaw:	lɔ̃ ³¹ ~, lɔ̃ ¹¹ (N.), lɔ̃ ^{4 5} (S.)
Pwo:	lãu ³³ (N.), lãu ^{5 5} (S.)

Note: With regard to the *B column, there is no split as in Kayan, Kayah, Bwe, Kayaw and some Sgaw varieties, or a split between B12 and B3 as in Pa-O and Pwo. Due to the pattern of tone split in the *A column of Pwo Karen, i.e. between A1 and A23, three types of lateral sounds were reconstructed: voiceless (*hl), preglottalised (*ɰl) and voiced (*l). As for the split pattern of the *B tone, it is impossible to tell whether the onset of this etymon is *hl- or *ɰl.

42. MILLIPEDE *k/s-waj

Pa-O:	cɔ̃ ^{5 5} wɛ ³³ (N.), cɔ̃ ¹¹ wɛ ^{5 3} (S.)
Kayan:	θa ²¹ kɔ̃ ¹¹ wi ³³
Kayah:	si ^{5 5} kɔ̃ ^{5 5} wi ¹¹

Bwe:	-
Kayaw:	te ¹¹ we ³³
Sgaw:	si ^{4 5} wɔ ¹¹ de ³³ (N.), si ^{5 5} wa ³¹ de ^{5 5} (S.)
Pwo:	wai ¹¹ (N.), wai ³¹ , wai ¹¹ , βai ¹¹ (S.)

Note: Pa-O, Central Karen and S. Pwo have tone A3 while Sgaw and N.Pwo have tone B3.

43. **MOLE *(jow) wi** (TB *bwiɣ, *bwəy)

Pa-O:	(ju ^{5 3}) wi ³³ (N.), (ju ^{5 5}) wi ^{5 3} (S.)
Kayan:	-
Kayah:	(jo ¹¹) wi ¹¹
Bwe:	wi ³³
Kayaw:	(ju ¹¹) wi ³³
Sgaw:	wi ³³ (N.), wi ³³ (S.)
Pwo:	wei ^{5 5} (N.), wei ³¹ , wei ¹¹ , βei ¹¹ (S.)

Note: To the Karen, a 'mole' is a kind of rat. See also no. 2 'bamboo rat' and no. 47 'mouse, rat'. The Kayan use Ju¹¹ ta¹¹ ha^{4 5} which is not a cognate word.

44. **MONITOR LIZARD (land ~) *khwo**

Pa-O:	- (N.), - (S.)
Kayan:	(khwi ¹¹ re ³³) khɔ ^{4 5}
Kayah:	khwɔ ^{5 5}
Bwe:	-
Kayaw:	(to ³³) ko ³³
Sgaw:	(tə ¹¹) khu ^{4 5} (N.), khɔu ²¹ (S.)
Pwo:	khu ^{4 5} (N.), khɔu ²¹ (S.)

Note: The highland Pa-O (N. Pa-O) use lɛm^{5 3}~lɛn^{5 3} and the lowland Pa-O (S. Pa-O) use lɛm³³ which is a Tai loan. Unlike the other Karenic languages, Kayaw has unaspirated onset.

45. **MONITOR LIZARD (water~) *k-re^A/*tre^A**

Pa-O:	- (N.), - (S.)
Kayan:	(khwi ¹¹) re ³³

Kayah:	(ta ¹¹ khwa ^{5 5} ta ¹¹) re ³³
Bwe:	tre ^{5 5}
Kayaw:	(te ^{5 5}) re ^{5 5}
Sgaw:	kre ^{4 5} , kə ¹¹ re ^{5 5} (N.), re ^{5 5} (S.)
Pwo:	- (N.), ʔei ³¹ , ʔei ¹¹ (S.)

Note: Some varieties of N. Sgaw have tone D while most of the Karenic languages have tone A. In Bwe, *k-re becomes *t-re and the animal prefix *k- is dropped (*k-r- > r-) in Kayan, Kayah, Kayaw, Pwo and S. Sgaw.

46. MONKEY *k-jo

Pa-O:	jo ^{4 5} (N.), ju ^{4 5} (S.)
Kayan:	jɔ ²¹
Kayah:	jɔ ³³
Bwe:	jo ¹¹
Kayaw:	jə ³³
Sgaw:	-
Pwo:	-

Note: There are lexical innovations in Southern Karen, i.e. mɔ¹¹ li^{4 5}, pha^{3 5} lai²¹ and chə^{5 5} ɔu³¹ in N. Sgaw, N. Pwo and S. Pwo, respectively, for 'monkey'. Probably, the etymon *k-jo is related to the word 'gibbon' in Sgaw and Pwo: kə¹¹ ju³³, zu³³ pʔa¹¹ (N. Sgaw), kə¹¹ ju³³ pha³¹ (S. Sgaw), ka²¹ ji³³ pha³³ (N. Pwo) and kə³³ ji^{5 3} pha³³ (S. Pwo). In N. Sgaw, ju³³ (A) means 'to swing' and pʔa¹¹ (B) means 'forest'; therefore, the literal meaning of the word 'gibbon' is 'animal (*k-) swinging in the forest'.

47. MOUSE, RAT *jow (TB *b-yuw, *b-yəw)

Pa-O:	ju ^{5 3} (N.), ju ^{5 5} (S.)
Kayan:	ʔu ¹¹
Kayah:	jo ¹¹
Bwe:	ju ¹¹
Kayaw:	ju ¹¹
Sgaw:	ji ¹¹ , zi ¹¹ (N.), ji ³¹ (S.)
Pwo:	jou ¹¹ (N.), jou ³³ (S.)

Note: See also no. 2 'bamboo rat' and no. 42 'mole'.

48. OX *bo

Pa-O: pho⁵³ (N.), po⁵⁵ (S.)

Kayan: pau¹¹

Kayah: pu¹¹

Bwe: bo³³

Kayaw: pu¹¹

Sgaw: -

Pwo: -

Note: Some groups of Pwo people call an ox “something having horns”, e.g. ch⁵⁵ n⁵³ (S. Pwo). In N. Sgaw, ‘ox’ is ta¹¹ t⁴⁵ or kl¹¹. The word kl¹¹ could be a loanword from Mon, i.e. khl¹¹ w (< Proto-Mon *klɛ ɔ (Diffloth, 1984)).

49. PANGOLIN, ANTEATER *jo

Pa-O: ju³³ (N.), - (S.)

Kayan: Jau³³

Kayah: ju¹¹, ju³¹

Bwe: j³³ (-θ³³)

Kayaw: ju³³

Sgaw: jo³³ (h^{31~}, h¹¹), zo³³ (ho¹¹) (N.), ju³³ (h⁴⁵) (S.)

Pwo: - (N.), j^{31~}, j¹¹ (S.)

50. PARROT *(tho) ki

Pa-O: ki⁵⁵ (N.), - (S.)

Kayan: khi¹¹

Kayah: khwi¹¹ (?)

Bwe: k(h)³³

Kayaw: ki¹¹, khi¹¹

Sgaw: ki²¹ (N.), ki⁴⁵ (S.)

Pwo: kei³³ (N.), kei⁵⁵ (S.)

Note: This etymon should be ki^{31~} or ki¹¹ (tone B12) instead of ki²¹ (tone D3) as in N. Sgaw.

51. PEACOCK *(tho) bra

Pa-O:	- (N.), - (S.)
Kayan:	pra ²¹
Kayah:	phre ³³
Bwe:	(tho ³³ -) ba ¹¹
Kayaw:	prɔ ³³
Sgaw:	bɣa ²¹ , phɣa ²¹ , phra ²¹ , pwa ²¹ (N.), pɣa ¹¹ (S.)
Pwo:	sja ²¹ (N.), ɕja ⁴⁵ (S.)

52. PHEASANT *(tho²¹) re²¹

Pa-O:	- (N.), - (S.)
Kayan:	-
Kayah:	ra ³³
Bwe:	-
Kayaw:	re ³³
Sgaw:	ɣi ²¹ (N.), ɣai ¹¹ (S.)
Pwo:	ɣai ²¹ (N.), ɣai ^{4 5} (S.)

Note: It is ɕja³³ tham³¹ in S. Pa-O. ɕja³³, the first part of the compound, is in fact ɕja³¹ 'chicken' (31>33). This suggests that Southern Pa-O people view 'pheasants' as 'chickens' not 'birds'.

53. PIG *thɔ²¹

Pa-O:	thɔ ²¹ (N., S.)
Kayan:	thau ^{4 5}
Kayah:	thɛ ^{5 5}
Bwe:	thɔ ³³
Kayaw:	tho ³³
Sgaw:	thɔ ^{4 5} (N.), thɔ ²¹ (S.)
Pwo:	tho ^{4 5} (N.), thu ²¹ (S.)

Note: See also no. 9 'wild boar'.

54. PORCUPINE (big kind) *sun²¹

Pa-O:	sun ^{5 5} (N.), sɔm ³³ , sɔn ³³ (S.)
-------	---

Kayan:	θwaŋ ¹¹
Kayah:	sɿ ¹¹
Bwe:	θu ³³
Kayaw:	su ¹¹
Sgaw:	- (N.), su ^{4 5} (S.)
Pwo:	- (N.), - (S.)

Note: The lexical innovations in N. Sgaw, N. Pwo and S. Pwo are pɔ³³ dɔ^{4 5}, pa²¹ do^{4 5} and chə³³ ʔau^{4 5} chũ^{5 5} ‘something shaking off its hair (spine)’, respectively.

55. PORCUPINE (small kind) *s-^{ba}

Pa-O:	- (N.), - (S.)
Kayan:	bɿ ^{5 3} (?)
Kayah:	se ³³ be ³³
Bwe:	-
Kayaw:	sa ^{5 5} ba ^{5 5}
Sgaw:	ba ³³ (N.), ba ^{5 5} (S.)
Pwo:	- (N.), θə ^{5 5} ba ³¹ , θə ^{5 5} ba ¹¹ (S.)

Note: Kayan has an irregular vowel.

56. QUAIL *(tho) hrwi (?)

Pa-O:	rɿ ^{4 5} (N.), - (S.)
Kayan:	rwi ^{4 5}
Kayah:	rwi ^{5 5}
Bwe:	-
Kayaw:	ri ³³
Sgaw:	pə ¹¹ wi ³³ (N.), pə ¹¹ wi ^{5 5} (S.)
Pwo:	phu ³³ wei ³³ (N.), wei ^{5 5} (S.)

Note: The reconstruction of the tone of this etymon is problematical, due to tonal variation, i.e. tone D in Northern and Central Karen, tone A in Sgaw but tone B in Pwo.

57. RABBIT, HARE *p-də

Pa-O:	- (N.), - (S.)
Kayan:	dɛ ^{5 3}
Kayah:	dɛ ³³
Bwe:	pə ³³ dɛ ³³
Kayaw:	dɛ ^{5 5}
Sgaw:	pə ¹¹ dɛ ³³ , pə ¹¹ dɛ ^{5 5} (N.), pə ¹¹ thɛ ²¹ (S.)
Pwo:	pa ³³ dɛ ^{5 5} , pə ¹¹ de ^{5 3} (N.), pə ³¹ dɛ ³¹ , pə ³³ thɛ ²¹ (S.)

Note: In some S. Sgaw and S. Pwo varieties, *ɲd > *d > th- and tone *A > D.

58. SEROW, MOUNTAIN GOAT *jaŋ (TB *kye.l, *kyi[.])

Pa-O:	jaɪ ³³ (N.), - (S.)
Kayan:	ɰaɪ ³³
Kayah:	je ¹¹
Bwe:	ji ³³
Kayaw:	ji ³³
Sgaw:	- (N.), - (S.)
Pwo:	- (N.), - (S.)

Note: There is a lexical innovation in Southern Karen: chə³³ pha²¹ (S. Pwo) and ta³³ pha^{4 5} (N. Sgaw).

59. SHEEP *so^{A/B}

Pa-O:	sə ^{5 5} (N.), so ³³ (S.)
Kayan:	θɔ ³³
Kayah:	-
Bwe:	θu ^{5 5}
Kayaw:	so ^{5 5}
Sgaw:	so ³³ , so ^{5 5} (N.), su ^{5 5} (S.)
Pwo:	- (N.), θu ^{5 5} (S.)

Note: This etymon has tone A in Kayan, Bwe, Kayaw and Sgaw but tone B in Pa-O and Pwo.

60. SKINK *biɛ

Pa-O:	plɛ̃ ^{4 5} (N.), plɛ̃ ^{4 5} (S.)
Kayan:	plɛ̃ ²¹
Kayah:	plɑ ³³
Bwe:	bli ¹¹ (ca ^{5 5})
Kayaw:	ple ³³
Sgaw:	plɛ̃ ²¹ , phlɛ̃ ²¹ (N.), plɛ̃ ¹¹ (S.)
Pwo:	phlẽ ²¹ (N.), phlẽ ^{4 5} (S.)

61. **SNAIL (land~) *khlo^{A/B}**

Pa-O:	khlo ³¹ (N.), - (S.)
Kayan:	khlo ^{5 3} (?)
Kayah:	-
Bwe:	khlo ³³
Kayaw:	-
Sgaw:	khlo ³¹ , khlo ¹¹ (N.), khlu ^{4 5} (S.)
Pwo:	khlu ³³ (N.), khlu ^{5 5} (S.)

Note: In Kayan, this etymon has tone A instead of tone B like the rest. Also, the vowel should be au not o. Language contact could be a cause of vowel and tone irregularity in Kayan. See also no. 62 'water snail'

62. **SNAIL (water~) *s-ŋwĩ**

Pa-O:	ŋwĩ ^{5 5} (N.), ŋwi ³³ (S.)
Kayan:	ŋwi ¹¹
Kayah:	-
Bwe:	θə̃ ^{5 5} mj ³³
Kayaw:	si ¹¹ mi ¹¹
Sgaw:	- (N.), - (S.)
Pwo:	- (N.), - (S.)

Note: This etymon has not been kept in Southern Karen. The word khlu (B12) is used for both 'water snail' and 'land snail' with different modifiers, for example, khlu^{5 5} mĩ̃^{5 5} 'water snail' and khlu^{5 5} klaĩ^{4 5} 'land snail' in S. Pwo. See also no. 60 'land snail'.

63. SNAKE *rowŋ (TB *b-ru.l)

Pa-O:	ru ^{5 3} (N.), ru ^{5 5} (S.)
Kayan:	rəu ¹¹ (?)
Kayah:	ro ¹¹ , ru ³¹
Bwe:	Ru ³³
Kayaw:	ru ¹¹
Sgaw:	ɣi ¹¹ (N.), ɣĩ ³¹ (S.)
Pwo:	ɣou ¹¹ , ɣou ³¹ (N.), βou ³³ , vũ ^{5 5} (S.)

Note: Mostly, the PK final nasals were reconstructed from those that have been retained in Pa-O. This etymon in Pa-O has the CVØ syllable structure or open syllable. However, it is quite unusual that in some S. Sgaw and S. Pwo varieties, the word 'snake' has a nasalised vowel $\tilde{i}$ (S. Sgaw) or the nasal vowel $\tilde{u}$ (S. Pwo). Perhaps, the final *-w became nasal (*-N) and then a nasalised vowel ($\tilde{v}$).

64. SPIDER *gaŋ

Pa-O:	(kuŋ ³³) kaŋ ³³ (N.), (jɔŋ ³¹) kaŋ ^{5 3} (S.)
Kayan:	kaŋ ³³
Kayah:	-
Bwe:	ga ¹¹ - gu ¹¹
Kayaw:	kɔ ³³
Sgaw:	- (N.), - (S.)
Pwo:	pu ¹¹ khɛ̃ ^{5 5} , phu ^{5 5} khã̃ ^{5 3} (N.), khɔ̃ ¹¹ , khɔ̃ ³¹ (S.)

Note: The etymon *gaŋ has been retained in almost all of the Karenic languages, except in Sgaw, 'spider' is an innovated word, i.e. kɔ¹¹ pɔ³³.

65. SQUIRREL *hli

Pa-O:	li ^{5 5} (N.), li ³³ (S.)
Kayan:	-
Kayah:	-
Bwe:	li ¹¹ (?)
Kayaw:	-
Sgaw:	li ^{4 5} , li ^{4 5} lu ¹¹ (N.), li ^{4 5} (S.)

Pwo: lei³³ (N.), lei^{5 5} (S.)

Note: N. Sgaw has tone D12 and Bwe has tone D3 (suggesting *l-), while the rest have tone B12. Tone B12 and tone D12 indicate a voiceless onset, *hl- or *ɿl.

66. TERMITE (winged~) *bi

Pa-O: phi^{5 3} (N.), pi^{5 5} (S.)

Kayan: pi¹¹

Kayah: pi¹¹

Bwe: pə³³ ɿ³³

Kayaw: pi¹¹ ɿ¹¹

Sgaw: pi¹¹ (N.), pi³¹ (S.)

Pwo: phai³³, phɛi³³ (N.), phɛi³¹ (S.)

Note: The B3 tone in Pa-O, S. Sgaw, some N. Sgaw varieties and S. Pwo indicate that the onset of this etymon should be *b-. Pwo vowels are irregular. S. Pwo has tone A23, while N. Pwo has tone B12. The reduction of the original compound into a monosyllabic word could have been the cause of these irregularities.

67. TICK *khej

Pa-O: ki^{5 5} (N.), - (S.)

Kayan: kh i¹¹

Kayah: khə¹¹

Bwe: khi³³

Kayaw: khi¹¹

Sgaw: khi³¹~, khi¹¹ (N.), khi^{4 5} (S.)

Pwo: khei³³ (N.), khei^{5 5} (S.)

68. TIGER *k(h)e (TB *d-key, *d-kəy, *k-key)

Pa-O: ke³¹ (N., S.)

Kayan: khai^{5 3}

Kayah: khi³³

Bwe: khi^{5 5}

Kayaw: khi^{5 5}

Sgaw: khe³³ j i¹¹.. 'lion' (N.), - (S.)

Pwo: khi^{3 5} (N.), khi^{5 3} (S.)

Note: The word meaning 'tiger' in Sgaw Karen is bɔ̌²¹ so²¹, bɔ̌²¹ θo¹¹, bɔ̌²¹ so^{5 3}, bɔ̌²¹ sɔ̌²¹ ɔ̌²¹ or bɔ̌¹¹ so^{4 5} ɔ̌²¹ depending upon each variety; however, khe³³ (A) is found in the compound meaning 'lion', i.e. khe³³ ji¹¹.

69. TOAD *(de) sow

Pa-O: de^{5 5} su^{5 5} (N.), de³³ su³³ (S.)

Kayan: dai¹¹ θu¹¹

Kayah: di¹¹ so¹¹

Bwe: di³³ θu³³

Kayaw: di¹¹ su¹¹

Sgaw: de^{31~} si^{31~} (N.), de^{45'} si^{45'} (S.)

Pwo: di³³ sou³³, di³³ θu³³ (N.), di^{5 5} sou^{5 5}, di^{5 5} θou^{5 5} (S.)

70. TORTOISE *k(h)li

Pa-O: kle²¹ (N.), kli²¹ (S.)

Kayan: khli^{4 5}

Kayah: kli^{5 5}

Bwe: khli³³

Kayaw: khli³³

Sgaw: khli^{4 5} (N.), khlai²¹ (S.)

Pwo: khlai^{4 5} (N.), khlai²¹, khlai³¹ (S.)

71. VULTURE *hlaŋ^{A/B} k-da

Pa-O: lɛŋ^{5 5} ta^{4 5} (N.), ta³³ lɔŋ³³ ta^{4 5} (S.)

Kayan: -

Kayah: lɛ¹¹ ta³³

Bwe: lɔ³³ da^{5 5} (?)

Kayaw: la¹¹ ta³³

Sgaw: lɔ³³ ka²¹ ta²¹, lɔ³³ kɔ¹¹ ta²¹ (N.), lɔ³³ kɔ^{11~} ta^{11~} (S.)

Pwo: lɛ^{3 5} ka²¹ tha²¹ (N.), lɔ̃^{5 3} ka²¹ tha²¹ (S.)

Note: In some N. Sgaw varieties, vultures are called “tho^{31~} ta¹¹ ꨑi^{31~}, tho¹¹ ta¹¹ ꨑi^{4 5} or tho¹¹ ta²¹ ꨑi²¹”, literally this means ‘rotten-thing birds’. This could be an Austroasiatic loanword. In Praok-Wa, the word meaning ‘vulture’ is (sim) klaŋ cuꨑ (from the author’s fieldnotes collected in December, 1995).

72. WASP *(dəŋꨑ) ꨑdeꨑ

Pa-O: lak^{4 5} kaꨑ²¹ de^{31~} (N.), - (S.)

Kayan: dauꨑ^{4 5} dai^{5 3}

Kayah: di³³

Bwe: -

Kayaw: di^{5 5}

Sgaw: tə^{11~} de³³, tə¹¹ de^{5 5} (N.), ti³¹ de^{5 5} (S.)

Pwo: thã^{11~} di^{5 5}, thə¹¹ di^{5 3} (N.), thə³³ di^{31~}, thə³³ di¹¹ (S.)

Note: To the Sgaw and Pwo, ‘wasp’ is a sub-species of ‘ant’. See also no. 1 ‘ant’.

73. WEEVIL *roŋꨑ

Pa-O: roŋ^{5 3} (N.), roŋ^{5 5} (S.)

Kayan: ro¹¹

Kayah: rɔ¹¹

Bwe: -

Kayaw: rə¹¹

Sgaw: ɤə^{11~} (N.), ɤi³¹ (S.)

Pwo: ɤãu^{11~} (N.), βõu³³, ɤõ³³ (S.)

Conclusion and Discussion

Among the seventy-three etyma, twenty-one etyma, i.e. no. 1 ‘ant’, no.7 ‘bee’ (*Apis dorsata*), no. 8 ‘bird’, no. 10 ‘buffalo’, no. 14 ‘chameleon’, no. 17 ‘crab’, no. 23 ‘dog’, no.24 ‘dove’, ‘no. 26 ‘elephant’, no. 28 ‘flea’, no. 30 ‘frog’, no.32 ‘hawk’, no. 47 ‘rat’, no. 49 ‘pangolin’, no. 51 ‘peacock’, no. 53 ‘pig’, no. 57 ‘rabbit’, no. 60 ‘skink’, no. 67 ‘tick’, no. 69 ‘toad’ and no. 70 ‘tortoise’, are not problematical because of the availability of data for reconstruction from all of the languages selected as representatives of the three major branches of the Karenic languages. In addition, the Karen seemed to be familiar with these animals and could quickly recognise them when the pictures were shown to them during the interviews. It is also possible that some of the etyma are loanwords from Austroasiatic languages, especially Mon and Wa; for example, no. 15 ‘chicken’, no. 26 ‘elephant’, no. 34 ‘horse’ and no. 71 ‘vulture’. The names of some aggressive wild animals, such as no. 4 ‘bear’, no. 68 ‘tiger’ etc., have become taboo words and have been replaced by euphemisms as in Sgaw Karen. In the modernised mountainous areas of Thailand, I noticed that children remember fewer and fewer animal names in their own languages,

especially non-domesticated animals, due to compulsory education with Thai as the medium of instruction, good transportation, the mass media and tourism. Sooner or later a similar phenomenon will occur in Myanmar. In the near future, only a few Karen animal names will be in the lexicon. A decrease of Karen words with an increase of Thai loanwords for animal names could provide a good case for studying “language change in progress” or “change in apparent time”, with regard to vocabulary loss and morphological change as well as sound change.

Acknowledgements

I would like to express my gratitude to the Thailand Research Fund (TRF) for funding the Karen Linguistics Project for three years from 2009-2012. Many thanks go to my Karen friends and the local authorities for their kind co-operation. I also feel thankful to my research assistants, Karnthida Kerdpol and Sujinat Jitwiriyant, for typing the manuscript and providing all kinds of assistance. Lastly, useful comments and suggestions from James A. Matisoff, David Bradley and David B. Solnit made me be more careful when drafting my research report.

References

- Benedict, Paul K. 1972. **Sino-Tibetan: A conspectus**. New York: Cambridge University Press.
- Bradley, David. 1997. Tibeto-Burman languages and classification. **Papers in Southeast Asian linguistics** 14, ed. David Bradley, 1-71. Canberra: Pacific Linguistics.
- Burling, Robbins. 1969. **Proto-Karen: A reanalysis**. (Occasional papers of the Wolfenden Society on Tibeto-Burman linguistics vol 1). Ann Arbor, MI: Department of Linguistics, University of Michigan.
- Court, Christopher [trans.]. 1972. Two-way and three-way splitting of tonal systems in some Far Eastern languages. **Tai phonetics and phonology**, eds. Jimmy G. Harris and Richard B. Noss, 58-86. Bangkok: Central Institute of the English Language (CIEL), Ministry of University Affairs. [Translation and expansion of Haudricourt 1961]
- Diffloth, G. 1980. The Wa languages. **Linguistics of the Tibeto-Burman Area** 5. 2: 1-182.
- Diffloth, G. 1984. **The Dvaravati Old Mon language and Nyah Kur**, 402 pp., maps. Chulalongkorn University Printing House [on behalf of the Mon and Nyah Kur Linguistics Project].
- Haudricourt, André-Georges. 1946. Restitution du Karen commun. **Bulletin de la Société Linguistique de Paris** 42: 103-111.
- Haudricourt, André-Georges. 1953. À propos de la restitution du Karen commun. **Bulletin de la Société Linguistique de Paris** 49.1:129-132.
- Haudricourt, André-Georges. 1961. Bipartition et tripartition des systèmes de tons dans quelques langues d'Extrême-Orient. **Bulletin de la Société de Linguistique de Paris** 56.1:163-180.
- Henderson, Eugénie J. A. 1973. Bwe Karen as a two-tone language? An Inquiry into the Interrelation of pitch, tone and initial consonant. Paper presented at the 6th International Conference on Sino-Tibetan Languages and Linguistics, San Diego.

- Henderson, Eugénie J. A. 1997. **Bwe Karen dictionary: with texts and English-Karen word list** (2 vols). London: School of Oriental and African Studies.
- Jones, Robert B. 1961. **Karen linguistic studies: Description, comparison and texts**. Berkeley: University of California Press.
- Kauffman, William G. 1993. **The great tone split and Central Karen**. University of North Dakota: MA thesis.
- Lewis, M. Paul (ed.), 2009. **Ethnologue: Languages of the World**, Sixteenth edition. Dallas, Tex.: SIL International. [Online] Available at: <http://www.ethnologue.com> [Accessed 14 May 2012].
- Luce, Gordon H. 1959a. Introduction to the comparative study of Karen languages. **Journal of the Burma Research Society** 42.1:1-18.
- Manson, Ken. 2001. Review of David Solnit's Eastern Kayah Li: Grammar, texts, glossary. **Notes on Linguistics** 4.4:234-236.
- Manson, Ken. 2009. Prolegomena to reconstructing Proto-Karen. **LaTrobe Working Papers in Linguistics** 12.
- Manson, Ken. 2011. The subgrouping of Karen. Paper presented at SEALS 21, organised by Kasetsart University, Bangkok, May 11 – 13, 2011.
- Matisoff, James A. 2003. **Handbook of Proto-Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction**. Berkeley: University of California Press.
- Phillips, Audra. 2002. The West-Central Thailand Pwo Karen people. **Minority language orthography in Thailand: Five case studies**, eds. TU-SIL-LRDP, 69-83. Bangkok: TU-SIL-Language Research and Development Project.
- Shintani, Tadahiko. 2003. Classification of Brakalaungic (Karenic) languages in relation to their tonal evolution. **Cross-Linguistics studies of tonal phenomena: historical development, phonetics of tone and descriptive studies**, Shigeki Kaji (ed.), pp. 37-56. Tokyo: Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Solnit, David B. 2001. Another look at Proto-Karen. Paper presented at the 34th International Conference on Sino-Tibetan Languages and Linguistics, Kunming.

Table 1 Proto-Karen onsets

Proto-Karen	N. Pa-O	S. Pa-O	Kayan	Kayah	Bwe	Kayaw	N. Sgaw	S. Sgaw	N. Pwo	S. Pwo	Gloss
*pɪ-	phɛ³¹ (A12)	phɛ³¹ (A12)	phɛ³¹ (A12)	phɛ³³ (A12)	--	phɛ⁵ ⁵ (A12)	phɛ³³ (A)	phɛ⁵ ⁵ (A12)	phɛ⁵ ⁵ (A1)	phɛ⁵ ⁵ (A1)	‘charff, bran’
*tɪ-	thɛ³¹ (A12)	thɛ³¹ (A12)	thɛ³¹ (A12)	thɛ³³ (A12)	chɛ⁵ ⁵ (A12)	thɛ⁵ ⁵ (A12)	thɛ³³ (A)	thɛ⁵ ⁵ (A12)	thɛ⁵ ⁵ (A1)	thɛ⁵ ⁵ (A1)	‘water’
*tʰɪ-	chɛ³¹ (A12)	chɛ³¹ (A12)	chɛ³¹ (A12)	chɛ³³ (A12)	ʃɛ⁵ ⁵ (A12)	chɛ⁵ ⁵ (A12)	chɛ³³ (A)	chɛ⁵ ⁵ (A12)	chɛ⁵ ⁵ (A1)	chɛ⁵ ⁵ (A1)	‘ill, painful’
*kɪ-	kʰɛ⁵ ⁵ (B12)	kʰɛ³³ (B12)	kʰɛ³¹ (B)	kʰɛ¹¹ (B)	kʰɛ³³ (B)	kʰɛ¹¹ (B)	kʰɛ³³ (B12)	kʰɛ⁴ ⁵ (B12)	kʰɛ³³ (B12)	kʰɛ⁵ ⁵ (B12)	‘bitter’
*hm-	mɛ³¹ (A12)	mɛ³¹ (A12)	mɛ³³ (A12)	mɛ³³ (A12)	mɛ⁵ ⁵ (A12)	mɛ⁵ ⁵ (A12)	mɛ³³ (A)	mɛ⁵ ⁵ (A12)	mɛ³³ (A1)	mɛ⁵ ⁵ (A1)	‘wife’
*hn-	nɛ³¹ (A12)	nɛ³¹ (A12)	nɛ³³ (A12)	nɛ³³ (A12)	gɛ⁹ nɛ³³ (A3)	nɛ⁵ ⁵ (A12)	kɛ¹¹ nɛ³³ (A)	nɛ⁵ ⁵ (A12)	nɛ⁵ ⁵ (A1)	nɛ⁵ ⁵ (A1)	‘bee (Apis dorsata)’
*ɲɪ-	ʝɛ³¹ (A12)	ʝɛ³¹ (A12)	ʝɛ³¹ (A12)	ʝɛ³³ (A12)	ʝɛ⁵ ⁵ (A12)	ʝɛ⁵ ⁵ (A12)	ʝɛ³³ (A)	ʝɛ⁵ ⁵ (A12)	ʝɛ³³ (A1)	ʝɛ⁵ ⁵ (A1)	‘easy’
*ɲɪ/ɲɪ-	ɲɛ³¹ (A12)	ɲɛ³¹ (A12)	ɲɛ³³ (A12)	ɲɛ³³ (A12)	--	--	ɲɛ³³ (A)	ɲɛ⁵ ⁵ (A12)	--	--	‘front’
*hw-	wɛ⁵ ⁵ (B12)	wɛ³³ (B12)	hwɛ¹¹ (B)	wɛ¹¹ (B)	hwɛ³³ (B)	--	wɛ³¹ (B12)	wɛ⁴ ⁵ (B12)	wɛ³³ (B12)	wɛ⁵ ⁵ (B12)	‘bamboo’
*hɪ/ɲɪ-	ɛɲ³¹ (A12)	ruɛ n³¹ (A12)	rwɛɲ⁵ ⁵ (A12)	rɪ³³ (A12)	hwɛ⁵ ⁵ (A12)	ru⁵ ⁵ (A12)	--	--	--	--	‘silver, money’
*hɪ/ɲɪ-	lɛ⁵ ⁵ (B12)	lɛ³³ (B12)	lɛ¹¹ (B)	lɛ¹¹ (B)	lɛ³³ (B)	lɛ¹¹ (B)	lɛ³¹ (B12)	lɛ⁴ ⁵ (B12)	lɛ³³ (B12)	lɛ⁵ ⁵ (B12)	‘leaf’
*s-	sɛ³¹ (A12)	sɪ³¹ (A12)	θɪ⁵ ⁵ (A12)	sɛ³³ (A12)	θɪ⁵ ⁵ (A12)	sɪ⁵ ⁵ (A12)	sɪ³³ (A)	sɪ⁵ ⁵ (A12)	sɛɟ⁵ ⁵ (A1)	θɛɪ⁵ ⁵ (A1)	‘dead, to die’
*ʃ-	chɛ⁵ ⁵ (B12)	chɛ³³ (B12)	chɛ¹¹ (B)	chɛ¹¹ (B)	ʃɛ³³ (B)	chɛ¹¹ (B)	chɛ³¹ (B12)	chɛ⁴ ⁵ (B12)	ʃɛ³³ (B12)	ʃɛ⁵ ⁵ (B12)	‘star’
*h-	~hɛ³¹ (A12)	~hɛ³³ (A12)	~hɛ¹¹ (B)	~hɛ³³ (A12)	~hɛ³³ (B)	~hɛ⁵ ⁵ (A12)	hɛ³³ (A)	~hɛ⁵ ⁵ (A12)	--	Yɛ⁵ ⁵ (B12)	‘last night, yesterday’
*p-	pɛ⁵ ⁵ (B12)	pɛ³³ (B12)	pɛ¹¹ (B)	pɛ¹¹ (B)	bɛ³³ (B)	pɛ³¹ (B)	--	--	--	--	‘to look after things’
*t-	ɛ⁵ ⁵ (B12)	ɛ³³ (B12)	ɛ¹¹ (B)	ɛ¹¹ (B)	dɛ³³ (B)	ɛ³¹ (B)	--	--	--	--	‘to arrive’
*c-	cɛ⁵ ⁵ (B12)	cɛ³³ (B12)	cɛ¹¹ (B)	cɛ¹¹ (B)	cɛ³³ (B)	cɛ¹¹ (B)	cɛ³¹ (B12)	cɛ⁴ ⁵ (B12)	cɛ³³ (B12)	cɛ⁵ ⁵ (B12)	‘wet’
*k-	kɛ⁵ ⁵ (B12)	kɛ³³ (B12)	kɛ¹¹ (B)	kʰɛ¹¹ (B)	kʰɛ³³ (B)	kʰɛ¹¹ (B)	kʰɛ³¹ (B12)	kʰɛ⁴ ⁵ (B12)	kʰɛ³³ (B12)	kʰɛ⁵ ⁵ (B12)	‘chin’
*ɲ-	ɲɛ⁵ ⁵ (B12)	ɲɛ³³ (B12)	ɲɛ¹¹ (B)	ɲɛ¹¹ (B)	ɲɛ³³ (B)	ɲɛ¹¹ (B)	ɲɛ³¹ (B12)	ɲɛ⁴ ⁵ (B12)	ɲɛ³³ (B12)	ɲɛ⁵ ⁵ (B12)	‘dung, excrement’
*ɲb-	bɔɲ⁵ ⁵ (B12)	bɔɲ³³ (B12)	bɔɲ¹¹ (B)	bɛ¹¹ (B)	ɲɛ³³ (B)	bɔ¹¹ (B)	bɔ³¹ (B12)	bɔ⁴ ⁵ (B12)	bɛ³³ (B12)	bɔ⁵ ⁵ (B12)	‘bamboo shoot’
*ɲd-	dɔ³¹ (A12)	dɔ³¹ (A12)	dɔ⁵ ⁵ (A12)	dɛ³³ (A12)	ɲɛ⁵ ⁵ (A12)	dɔ⁵ ⁵ (A12)	dɔ³³ (A)	dɔ⁵ ⁵ (A12)	dɔ³¹ (A23)	dɔ⁵ ⁵ (A23)	‘to spread out (mats)’
*hm/ɲh-	mɛ⁵ ⁵ (B12)	mɛ³³ (B12)	mɛ¹¹ (B)	mɪ¹¹ (B)	mɛ³³ (B)	mɪ¹¹ (B)	mɛ³¹ (B12)	mɛ⁴ ⁵ (B12)	mɪ³³ (B12)	mɪ⁵ ⁵ (B12)	‘fine’
*ɲɪ-	nɛ⁵ ⁵ (B12)	nɛɲ³³ (B12)	nɛ¹¹ (B)	nɛ¹¹ (B)	ɲɛ³³ (B)	dɛ¹¹ (B)	nɪ³¹ (B12)	nɪ⁴ ⁵ (B12)	nɛ³³ (B12)	nɛ⁵ ⁵ (B12)	‘year’

Table 1 Proto-Karen onsets (continued)

Table 1 Proto-Karen onsets (continued)

Proto-Karen	N. Pa-O	S. Pa-O	Kayan	Kayah	Bwe	Kayaw	N. Sgaw	S. Sgaw	N. Pwo	S. Pwo	Gloss
*khl-	khaul ^{5 5} (B12)	klaul ³³ (B12)	--	--	khlul ³³ (B)	--	khlul ^{31~} (B12)	khlul ^{4 5} (B12)	khlul ³³ (B12)	khlul ^{5 5} (B12)	'mat'
*pl-	plal ^{1~} (A12)	plal ¹ (B)	plal ¹ (B)	plal ¹ (B)	plal ¹¹ (D37)	plal ¹ (B)	plal ³³ (A)	plal ^{5 5} (A12)	plal ³³ (B12)	plal ^{5 5} (B12)	'bat'
*kl-	klral ³¹ (A12)	klal ³¹ (A12)	--	kle ³³ (A12)	--	ka ^{5 5} (A12)	klal ^{4 5} (D12)	klal ^{5 5} (A12)	klal ^{5 5} (A23)	klal ³¹ (A23)	'forest, jungle'
*bl-	plal ³¹ (A12)	--	blal ^{5 3} (A12)	--	plal ^{5 5} (A12)	--	blal ³³ (A)	blal ^{5 5} (A12)	blal ^{5 5} (A23)	blal ³¹ (A23)	'tasteless, pale'
*bl-	plral ^{5 3} (B3)	plal ¹ (B)	plal ¹ (B)	plal ¹ (B)	blal ³³ (B)	plal ¹ (B)	plal ¹¹ (B3)	plal ³¹ (B3)	plal ³³ (B3)	plal ³¹ (B3)	'to wash (face)'
*gl-	--	klal ^{5 3} (A3)	klal ^{5 3} (A3)	kle ³³ (A12)	kle ³³ (A3)	kle ³³ (A3)	kle ³³ (A)	kle ³³ (A3)	kle ^{5 5} (A23)	--	'path, way'
*ml-	--	mul ^{5 3} (A3)	mul ³³ (A3)	mol ¹ (A3)	-blul ¹¹ (D37)	--	klul ¹¹ mlul ³³ (A)	mlul ³³ (A3)	mlul ^{5 5} (A23)	mlul ³¹ (A23)	'trunk (of an elephant)'
*ph-	phwal ³¹ (A12)	phwal ³¹ (A12)	phwal ^{5 3} (A12)	phwal ³³ (A12)	phwal ^{5 5} (A12)	phl ^{5 5} (A12)	phl ^{31~} (A)	phl ^{5 5} (A12)	khwl ³³ (A1)	khwl ^{5 3} (A12)	'light (adj.)'
*hw-	thwl ^{5 5} (B12)	thwl ³³ (B12)	thwl ¹¹ (B)	thwl ¹¹ (B)	thwl ³³ (B)	thl ¹¹ (B)	thwl ^{31~} (B12)	thwl ^{4 5} (B12)	thwl ³³ (B12)	thwl ^{5 5} (B12)	'dog'
*chw-	chwel ^{5 5} (B12)	chwel ³³ (B12)	chwel ¹¹ (B)	chwul ¹ (B)	chwel ³³ (B)	chel ¹¹ (B)	chwel ^{31~} (B12)	chwel ^{4 5} (B12)	chwel ³³ (B12)	chwel ^{5 5} (B12)	'crab'
*khw-	khul ³¹ (A12)	khul ^{5 3} (A12)	khul ^{5 3} (A12)	khul ³³ (A12)	~khu ^{5 5} (A12)	kuh ^{5 5} (A12)	khul ³³ (A)	khul ^{5 5} (A12)	khul ³³ (A1)	khul ^{5 3} (A1)	'man'
*hw- *hw- *hw-	nhw ^{5 5} (B12)	nhw ³³ (B12)	nhw ¹¹ (B)	--	nhw ³³ (B)	nhel ¹¹ (B)	nhw ¹¹ (B37)	nhw ^{4 5} (B12)	nhel ³³ (B12)	nhel ^{5 5} (B12)	'yam'
*sw-	swil ^{5 5} (B12)	swil ³³ (B12)	thwl ¹¹ (B)	swil ¹ (B)	thwl ³³ (B)	su ¹¹ (B)	swil ^{31~} (B12)	swil ^{4 5} (B12)	swil ³³ (B12)	thwl ^{5 5} (B12)	'blood'
*cw-	cul ²¹ (D12)	cul ²¹ (D12)	cwl ^{4 5} (D12)	cwl ^{5 5} (D12)	--	cul ³³ (D)	(sul ¹¹) wul ^{4 5} (D12)	--	(thul ¹¹) wul ^{31~} (D12)	(thul ¹¹) wul ^{31~} (D12)	'to suck'
*k(w)-	~kl ^{5 5} (B12)	--	~khl ¹¹ (B)	~khwil ¹¹ (B)	~k(h) ³³ (B)	~kl ¹¹ (B)	~k(h) ²¹ (D37)	~kl ^{4 5} (B12)	~kel ³³ (B12)	~kel ^{5 5} (B12)	'parrot'
*bw-	bwal ³¹ (A12)	bwal ³¹ (A12)	bwal ^{5 3} (A12)	bul ³³ (A12)	bu ^{5 5} (A12)	bu ^{5 5} (A12)	wal ³³ (A)	wal ^{5 5} (A12)	kwul ³¹ (A2-3)	kwul ³¹ (A2-3)	'white'
*dw-	dwal ^{5 5} (B12)	dwal ³³ (B12)	dwal ¹¹ (B)	--	(dul ³³ (B)	del ¹¹ (B)	dwal ^{31~} (B12)	dwal ^{4 5} (B12)	thwel ¹¹ (B3)	thwel ³³ (B3)	'grasshopper'
*dw-	thel ^{4 5} (D3)	thel ^{4 5} (D3)	thul ²¹ (D3)	thel ³³ (D3)	thel ³³ (D12)	thul ³³ (D3)	thel ²¹ (D3)	thel ¹¹ (D3)	thel ²¹ (D3)	thel ^{4 5} (D3)	'torn'
*gw-	wal ³³ (A3)	wal ³³ (A3)	gwul ³³ (A3)	wel ¹ (A3)	wul ³³ (A3)	wul ³³ (A3)	wul ³³ (A)	wul ³³ (A3)	wul ^{5 5} (A23)	wul ³¹ (A23)	'husband'
*mw-	mwal ^{5 3} (B3)	mwal ^{5 5} (B3)	mwal ¹¹ (B)	mal ¹ (B)	--	mel ¹¹ (B)	mel ¹¹ (B3)	mel ³¹ (B3)	mel ¹¹ (B3)	mwal ³³ (B3)	'yes'
*jw-	jlwl ^{5 5} (B12)	jlwl ³³ (B12)	jlwl ¹¹ (B)	--	(thul ^{5 5}) ml ³³ (B)	(sil ¹¹) ml ¹¹ (B)	--	--	--	--	'water snail'
*rw-	rwil ^{5 3} (B3)	rwil ^{5 5} (B3)	rwil ¹ (B)	rwil ¹ (B)	(khu ³³) wul ³³ (B)	rl ¹¹ (B)	Yl ¹¹ (B3)	Yl ³¹ (B3)	Yel ¹¹ (B3)	Yel ³³ (B3)	'roof'
*lw-	~lwl ^{5 5} (B3)	~lwl ¹¹ (B)	~lwl ¹¹ (B)	~lwl ¹¹ (B)	~lwl ³³ (B)	~l ¹¹ (B)	~lwl ¹¹ (B3)	~lwl ³¹ (B3)	~l ¹¹ (B3)	~lwl ³³ (B3)	'dove'

Table 1 Proto-Karen onsets (continued)

Proto-Karen	N. Pa-O	S. Pa-O	Kayan	Kayah	Bwe	Kayaw	N. Sgaw	S. Sgaw	N. Pwo	S. Pwo	Gloss
*ɕj-	ɕja³¹ˣ (A12)	ɕja³¹ˣ (A12)	ɕj¹⁵³ (A12)	cha³³ (A12)	j¹⁵⁵ (A12)	ɕ¹⁵⁵ (A12)	chɔ³³ (A)	chɔ⁵⁵ (A12)	chɕ³⁵ (A1)	chɔ⁵⁵ (A1)	'chicken'
*kj-	khjaᵀ²¹ (D12)	khjaᵀ²¹ (D12)	khjᵀ⁴⁵ (D12)	khja⁵⁵ (D12)	--	--	kiᵀ⁴⁵ (D12)	kaᵀ²¹ (D12)	kaᵀ³¹ (D12)	kaᵀ³¹ (D12)	'astirigent'
*ᵀbj-	pja⁵³ (B12)	pja³³ (B12)	bja¹ˣ (B)	--	(ᵀa³³) ᵀa³³ (B)	--	--	--	pa³³ (B12)	ba¹⁵⁵ (B12)	'flat'
*ŋj-	ŋjaᵀ⁴⁵ (D3)	ŋjaᵀ⁴⁵ (D3)	ŋjᵀ²¹ (D3)	ŋja³³ (D3)	(ɔ³³,ɔ) j¹ˣ¹¹ (D3)	j¹³³ (D)	jjᵀ²¹ (D3)	jɔ¹¹ (D3)	jɔᵀ²¹ (D3)	jɔᵀ⁴⁵ (D3)	'long (of time)'
*rj-	rja³³ (A3)	rja⁵³ (A3)	ja³³ (A3)	je¹¹ (B)	(gɔ)jɕ³³ (B)	ja¹¹ (B)	ja³³ (A)	ja³³ (A3)	ja⁵⁵ (A23)	ja³¹ (A23)	'hundred'
*khrw-	chut²¹ (D12)	chut²¹ (D12)	chwᵀ⁴⁵ (D12)	khrw¹⁵⁵ (D12)	khwi³³ (D12)	chu³³ (D)	xi³³ (A)	(ta¹¹) xi⁵⁵ (A12)	xe³³ (B12)	xwi⁵⁵ (B12)	'bone'

Table 2 Proto-Karen rhymes

Proto-Karen	N. Pa-O	S. Pa-O	Kayan	Kayah	Bwe	Kayaw	N. Sgaw	S. Sgaw	N. Pwo	S. Pwo	Gloss
*ɿ	mɪ ³³ (A3)	mɪ ^{5_3} (A3)	mɪ ³³ (A3)	mɪ ¹¹ (A3)	--	mɪ ³³ (A3)	mɪ ³³ (A)	mɪ ³³ (A3)	mɛɪ ^{5_5} (A23)	mɛɪ ³¹ (A23)	'wild, e.g. ~boar'
*e	ce ^{5_5} (B12)	ce ³³ (B12)	ceɪ ³³ (B12)	ci ¹¹ (B)	ci ^{1_33} (B)	ci ¹¹ (B)	ce ^{31~} (B12)	ce ^{4_5} (B12)	ci ^{5_5} (B12)	ci ^{5_5} (B12)	'left side'
*ɛ	lɛɪ ^{5_5} (B3)	lɛɪ ^{5_5} (B3)	lɛɪ ¹¹ (B)	lɛɪ ¹¹ (B)	lɛ ³³ (B)	lɛ ¹¹ (B)	lɛɪ ¹¹ (B3)	lɛɪ ³¹ (B3)	lɛɪ ³³ (B3)	lɛɪ ³³ (B3)	'wide'
*i	mɪ ^{5_3} (B3)	mɪ ³³ (B12)	mɪ ¹¹ (B)	mɔɿ ¹¹ (B)	mɔ ³² (B)	mɪ ¹¹ (B)	mɪ ^{31~} (B12)	mɪ ^{4_5} (B12)	mɪ ³³ (B12)	mɪ ^{5_5} (B12)	'woman, female'
*ɔ	pɪə ^{5_5} (B12)	pɪə ³³ (B12)	pɪəɪ ¹¹ (B)	pɪoɿ ¹¹ (B)	--	pɪə ¹¹ (B)	pɪɪ ^{31~} (B12)	pɪɪ ^{4_5} (B12)	pɪəɪ ³³ (B12)	pɪɪ ^{5_5} (B12)	'short, low'
*a	ce ³¹ (A12)	ce ³¹ (A12)	che ^{5_3} (A12)	che ³³ (A12)	ʃɛ ^{5_5_5} (A12)	che ^{5_5} (A12)	che ³³ (A)	che ^{5_5} (A12)	che ^{3_5} (A1)	che ^{5_3} (A1)	'to sell'
*u	ɭu ³¹ (A12)	ɭu ³¹ (A12)	ɭu ^{5_3} (A12)	ɭu ³³ (A12)	u ³ (B)	ɭu ^{5_5} (A12)	ɭu ³³ (A)	ɭu ^{5_5} (A12)	ɭu ^{5_5} (A23)	ɭu ³¹ (A23)	'to blow'
*o	tho ^{5_3} (B3)	to ^{5_5} (B3)	--	luɿ ¹¹ (B)	do ³³ (B)	luɿ ¹¹ (B)	toɿ ¹¹ (B3)	luɿ ³¹ (B3)	thu ³³ (B3)	thu ³³ (B3)	'handle (n.)'
*ɔ	so ^{5_5} (B12)	ɕau ³³ (B12)	θau ¹¹ (B)	soɿ ¹¹ (B)	θo ³³ (B)	so ¹¹ (B)	ɕYɔ ^{31~} (B12)	ɕYɔ ^{4_5} (B12)	ɕo ³³ (B12)	θo ^{5_5} (B12)	'hemp'
*ej	dɪ ^{5_5} (B12)	dɪ ³³ (B12)	ɿ ¹¹ (B)	dəɿ ¹¹ (B)	dɿ ³³ (B)	dɪ ¹¹ (B)	dɪ ^{31~} (B12)	dɪ ^{4_5} (B12)	deɪ ³³ (B12)	deɪ ^{5_5} (B12)	'egg'
*oj	--	loɪ ³³ (B12)	loɿ ¹¹ (B)	--	lɪ ³³ (B)	lɪ ¹¹ (B)	leɿ ^{31~} (B12)	lɪ ^{4_5} (B12)	--	laɪ ^{5_5} (B12)	'to lick'
*aw	tho ³¹ (A12)	tho ³¹ (A12)	thou ^{5_3} (A12)	thu ³³ (A12)	tho ^{5_5} (A12)	thu ^{5_5} (A12)	tho ³³ (A)	tho ^{5_5} (A12)	tho ^{5_3} (A1)	tho ^{5_3} (A1)	'long, tall'
*ow	lu ^{5_5} (B12)	lu ³³ (B12)	luɿ ¹¹ (B)	loɿ ¹¹ (B)	lu ³³ (B)	luɿ ¹¹ (B)	lɪ ^{31~} (B12)	lɪ ^{4_5} (B12)	lou ³³ (B12)	lou ^{5_5} (B12)	'cotton thread'
*ɪm	ci ¹¹ (B12)	ci ¹¹ (B12)	ci ¹¹ (B)	ci ¹¹ (B)	ci ³³ (B)	ci ¹¹ (B)	ci ¹¹ (B3)	ci ³¹ (B3)	chəɪ ¹¹ (B3)	chɛɪ ³³ (B3)	'to squeeze, to ooze'
*ɪn	mɪn ³¹ (A12)	mɪn ³¹ (A12)	mɿ ^{5_3} (A12)	mɪ ³³ (A12)	mɪ ^{5_5} (A12)	mɪ ^{5_5} (A12)	mɪ ³³ (A)	mɪ ^{5_5} (A12)	mɿ ^{3_5} (A1)	mɿ ^{5_5} (A1)	'ripe, cooked'
*ɪŋ	~kɪɪŋɿɿ	kɪɪ ^{5_3} (B12)	kɪɿ ^{5_3} (A12)	kɪɿ ³³ (A12)	--	kɪɿ ^{5_5} (A12)	~kɪɿ ³³ (A)	--	kɪ ^{5_5} (A1)	~kɪɛɿ ^{5_3} (A1)	'filled'
*eŋ	neŋ ^{5_5} (B12)	neŋ ³³ (B12)	neɿ ¹¹ (B)	noɿ ¹¹ (B)	ɲe ³³ (B)	deɿ ¹¹ (B)	nɪ ^{31~} (B12)	nɪ ^{4_5} (B12)	ɲəɪ ³³ (B12)	ɲɛɪ ^{5_5} (B12)	'year'
*ɛm	ɛm ^{5_5} (B12)	ɛm ³³ (B12)	joŋɿ ¹¹ (B)	joɿ ¹¹ (B)	--	lɪ ¹¹ (B)	leɿ ^{31~} (B12)	lɪ ^{4_5} (B12)	lɿ ^{5_5} (B12)	laɿ ^{5_5} (B12)	'to put out (~the tongue)'
*ɛn	ɭeŋ ^{5_5} (B12)	ɭeŋ ^{5_5} (B12)	ɭɿ ¹¹ (B)	ɭuɿ ¹¹ (B)	--	--	ɭeɿ ^{31~} (B12)	ɭɪ ^{4_5} (B12)	ɭəɪ ³³ (B12)	ɭəɪ ^{5_5} (B12)	'to bite'
*ɛŋ	thɛŋ ³¹ (A12)	thɛŋ ³³ (A12)	thɿ ^{5_3} (A12)	tho ³³ (A12)	the ^{5_5} (A12)	thɿ ^{5_5} (A12)	the ³³ (A)	thɿ ^{5_5} (A12)	thəɪ ^{5_5} (A1)	thəɪ ^{5_3} (A1)	'to weave (baskets)'
*im	thom ³³ (A3)	ɪə ^{5_3} (A3)	ɪi ³³ (A3)	ɪi ¹¹ (A3)	--	ɪɿ ³³ (A3)	ɪə ³³ (A)	ɪi ³³ (A3)	ɪɿ ^{5_5} (A23)	ɪə ^{5_3} i ³¹ (A23)	'a linear measure (½ yard)'
*in	pɪɾi ^{m31} (A12)	pɪɾi ^{5_3} (A12)	pɪɿi ³³ (A12)	pɪɿi ³³ (A12)	pɪlu ^{5_5} (A12)	pɪɾi ^{5_5} (A12)	pɪɿi ³³ (A)	pɪɿi ^{5_5} (A12)	pɪɿi ^{5_5} (A1)	pɪɿi ^{5_3} (A1)	'horne'
*iŋ	nəŋ ³³ (A3)	nəŋ ^{5_3} (A3)	nɪ ³³ (A3)	nɪ ¹¹ (A3)	--	--	nə ³³ (A)	nɪ ³³ (A)	ɲə ^{5_5} i ³¹ (A23)	ɲə ^{5_3} i ³¹ (A23)	'to win'

Table 2 Proto-Karen rhymes (continued)

Proto-Karen	N. Pa-O	S. Pa-O	Kayan	Kayah	Bwe	Kauaw	N. Sgaw	S. Sgaw	N. Pwo	S. Pwo	Gloss
*ɔm	*səmꞤꞤ (A12)	səm ³¹ (A12)	θə ^{5 3} (A12)	sɔ ³³ (A12)	θə ^{5 5} (A12)	sə ^{5 5} (A12)	sə ³³ (A)	sə ^{5 5} (A12)	sə ^{5 5} (A1)	sə ^{5 3} (A1)	‘three’
*ɔŋ	*thəŋꞤꞤ ^{5 5} (B12)	thəŋ ³³ (B12)	thə ¹¹ (B)	thɔ ¹¹ (B)	~thə ³³ (B)	ɪə ¹¹ (B)	~thə ³¹ (B12)	thə ^{4 5} (B12)	thə ^{5 5} (B12)	thə ^{5 5} (B12)	‘to stand’
*am	*Ꞥəm ^{5 5} (B12)	Ꞥəm ³³ (B12)	Ꞥəŋ ¹¹ (B)	Ꞥə ¹¹ (B)	Ꞥə ³³ (B)	Ꞥɔ ³³ (B)	Ꞥɔ ³¹ (B12)	Ꞥɔ ^{4 5} (B12)	Ꞥə ^{5 5} (B12)	Ꞥɔ ^{5 5} (B12)	‘to eat’
*an	*~khan ³¹ (A12)	~khan ³¹ (A12)	~khan ^{5 3} (A12)	--	khɛ ^{5 5} (A12)	~khɔ ^{5 5} (A12)	khɔ ³³ (A)	khɔ ^{5 5} (A12)	khɛ ^{5 5} (A1)	khə ^{5 3} (A1)	‘bamboo rat’
*aŋ	*laŋ ³³ (A3)	laŋ ^{5 3} (A3)	laŋ ³³ (A3)	lɛ ¹¹ (A3)	la ³³ (A3)	lɔ ³³ (A3)	lɔ ³³ (A)	lɔ ³³ (A3)	lɛ ^{5 5} (A23)	lɔ ³¹ (A23)	‘to descend’
*um/n	*chum/nꞤꞤ ^{5 5} (B12)	chum ³³ (B12)	chwəŋ ¹¹ (B)	chɪ ¹¹ (B)	ɟɔ ³³ (B)	ɟu ¹¹ (B)	chwə ³¹ (B12)	ɟu ^{4 5} (B12)	chə ^{5 5} (B12)	chə ^{5 5} (B12)	‘hair, fur, feather’
*uŋ	*ꞤbɪəŋꞤꞤ ^{5 5} (B12)	pwaŋ ³³ (B12)	bwaŋ ¹¹ (B)	bɪ ¹¹ (B)	ɟɔ ³³ (B)	bɪ ¹¹ (B)	bɔ ³¹ (B12)	bɔ ^{4 5} (B12)	bə ^{5 5} (B12)	bə ^{5 5} (B12)	‘fat (adj.)’
*om	*ꞤlomꞤꞤ	pom ³³ (B12)	blo ¹¹ (B)	plɪ ¹¹ (B)	plə ³³ (B)	plə ¹¹ (B)	plə ³¹ (B12)	pu ^{4 5} (B12)	pəu ^{5 5} (B12)	pəu ^{5 5} (B12)	‘pile (clt.)’
*oŋ	*loŋꞤꞤ ^{5 3} (B3)	loŋ ^{5 5} (B3)	ləu ¹¹ (B)	lɔ ¹¹ (B)	lo ³³ (B)	lə ¹¹ (B)	lɔ ¹¹ (B3)	lɪ ³¹ (B3)	ləu ¹¹ (B3)	ləu ³³ (B3)	‘stone, rock’
*ɔŋ	*dɔŋꞤꞤ ^{5 3} (B3)	ɔŋ ^{5 5} (B3)	ɔ ¹¹ (B)	--	dɔ ³³ (B)	lo ¹¹ (B)	lo ¹¹ (B3)	lu ³¹ (B3)	thə ³³ (B3)	thə ³³ (B3)	‘to pound’
*ɪt	*kɪnwiŋꞤꞤ	ɟuɪ ²¹ (D12)	ɟwiŋꞤ ^{4 5} (D12)	khwi ^{5 5} (D12)	khwi ³³ (D12)	ɟuɪ ³³ (D)	xɪ ³³ (A)	xɪ ^{5 5} (A12)	xəp ³³ (B12)	xwɪ ^{5 5} (B12)	‘bone’
*ɪk	*cwɪkꞤꞤ	cɪk ²¹ (D12)	cwɪŋꞤ ^{4 5} (D12)	cwɪ ^{5 5} (D12)	--	ɔɪ ³³ (D)	s-wɪŋꞤ ^{4 5} (D12)	--	--	θ-waɪŋ ³¹ (D12)	‘to suck’
*ɪ	*kɪɪꞤꞤꞤ	khɪŋꞤ ²¹ (D12)	khɪŋꞤ ^{4 5} (D12)	khɪ ^{5 5} (D12)	khɪ ³³ (D12)	khɪ ³³ (D)	khɪŋꞤ ^{4 5} (D12)	khəɪŋꞤ ²¹ (D12)	khəɪŋꞤ ^{4 5} (D12)	khəɪŋ ³¹ (D12)	‘dark’
*ek	*ɪekꞤꞤ	ɪek ^{4 5} (D3)	ɪɪŋꞤ ²¹ (D3)	ɪə ³³ (D3)	ɪe ¹¹ (D3)	ɪe ³³ (D)	ɪɪŋꞤ ²¹ (D3)	ɪəɪ ¹¹ (D3)	ɪəɪŋꞤ ²¹ (D3)	ɪəɪŋ ^{4 5} (D3)	‘hawk’
*ə	*ꞤdeꞤꞤꞤ	dəŋꞤ ²¹ (D12)	--	də ^{5 5} (D12)	dɛ ³³ (D12)	də ³³ (D)	dɪŋꞤ ^{4 5} (D12)	dəɪŋꞤ ²¹ (D12)	dəɪŋꞤ ^{4 5} (D12)	dəɪŋ ³¹ (D12)	‘wing’
*ɛ	*dɛꞤꞤꞤ	thɛŋꞤ ^{4 5} (D3)	ɪɛŋꞤ ²¹ (D3)	ɪə ³³ (D3)	də ¹¹ (D3)	ɪə ³³ (D)	ɪɛŋꞤ ²¹ (D3)	ɪɛ ¹¹ (D3)	thəŋꞤ ^{4 5} (D3)	thəŋ ^{4 5} (D3)	‘to fall’
*ə	*ɔꞤꞤꞤ	rəŋꞤ ^{4 5} (D3)	rəŋꞤ ²¹ (D3)	rɔ ³³ (D3)	--	rə ³³ (D)	Yrə ¹¹ (D3)	Yrə ¹¹ (D3)	YrəɪŋꞤ ²¹ (D3)	Yrəɪŋ ^{4 5} (D3)	‘flank (n.)’
*ap/t	*ŋɔpɔp/nꞤꞤ	ŋɔp ^{4 5} (D3)	ŋɛŋꞤ ²¹ (D3)	ŋɛ ³³ (D3)	ɟɛ ¹¹ (D3)	ɟɛ ³³ (D)	ɟɛ ¹¹ (B3)	ɟɛ ³¹ (B3)	ɟɛ ¹¹ (B3)	ɟɛ ³³ (B3)	‘five’
*ək	*~dəkꞤꞤ	(tə ³³) dək ^{4 5} (D3)	(kə ¹¹) dəɪŋꞤ ²¹ (D3)	(kɛ ^{5 5}) də ³³ (D3)	--	(kha ¹¹) dɔ ³³ (D)	(kɔɔŋꞤ ²¹) ɪəŋꞤ ²¹ (D3)	(thə ^{5 5}) ɪə ¹¹ (D3)	(kəɪ ³³) thəɪŋꞤ ²¹ (D3)	(kəɪŋꞤ ²¹) ɪəŋꞤ ^{4 5} (D3)	‘palate, to click’
*ə	*səꞤꞤꞤ	səŋꞤ ²¹ (D12)	θəŋꞤ ^{4 5} (D12)	sə ^{5 5} (D12)	θə ³³ (D12)	sɔ ³³ (D)	səŋꞤ ²¹ (D12)	səŋꞤ ²¹ (D12)	səŋꞤ ^{4 5} (D12)	θəŋꞤ ²¹ (D12)	‘heart’
*uk	*kukꞤꞤ	--	khɪ ^{5 5} (D12)	khɪ ³³ (D)	θə ^{5 5} khɪ ³³ (D12)	khɪ ^{4 5} (D12)	kɔuŋꞤ ²¹ (D12)	kɔuŋꞤ ²¹ (D12)	kəuŋꞤ ^{4 5} (D12)	kɔuŋ ³¹ (D12)	‘to caught’
*u	*thɪuꞤꞤꞤ	--	--	thɪ ^{5 5} (D12)	--	thɪ ³³ (D)	thɪuŋꞤ ^{4 5} (D12)	thɪuŋꞤ ²¹ (D12)	thəuŋꞤ ^{4 5} (D12)	thəuŋ ³¹ (D12)	‘want’
*ək	*~nokꞤꞤ	~nok ^{4 5} (D3)	~nɔŋꞤ ²¹ (D3)	~nɔ ³³ (D3)	~no ¹¹ (D3)	~nə ³³ (D)	~nɔŋꞤ ²¹ (D3)	~nɔ ¹¹ (D3)	~nəuŋꞤ ²¹ (D3)	~nɔuŋꞤ ^{4 5} (D3)	‘brain’

Table 2 Proto-Karen rhymes (continued)

Proto-Karen	N. Pa-O	S. Pa-O	Kayan	Kayah	Bwe	Kayaw	N. Sgaw	S. Sgaw	N. Pwo	S. Pwo	Gloss
*oŋ	*boŋŋŋ	boŋ ²¹ (D12)	bɔŋ ^{4 5} (D12)	bɔ ^{5 5} (D12)	--	bɔ ³³ (D)	bɪŋ ^{4 5} (D12)	bɔuŋ ²¹ (D12)	bouŋ ^{4 5} (D12)	bɔuŋ ³¹ (D12)	'to reach into'
*ɔt	*ktɪv/ɡrɔtŋ	sɔt ²¹ (D12)	chɔuŋ ^{4 5} (D12)	--	xɔ ¹¹ (D3)	--	xɔŋ ^{4 5} (D12)	xɔŋ ²¹ (D12)	xɔŋ ²¹ (D3)	xuŋ ^{4 5} (D3)	'eight'
*ɔŋ	*prɔŋŋ	phrɔŋ ²¹ (D12)	phrauŋ ^{4 5} (D12)	phrɛ ^{5 5} (D12)	pɔ ³³ (D12)	pro ³³ (D)	bʏɔŋ ^{4 5} (D12)	bʏoŋ ²¹ (D12)	pjoŋ ^{4 5} (D12)	pjuŋ ²¹ (D12)	'to vomit'

A view on Proto-Karen phonology and lexicon¹

Theraphan Luangthongkum

Department of Linguistics, Chulalongkorn University

Bangkok, Thailand

E-mail: theraphan.l@chula.ac.th

0. Abstract

Even though the reconstruction of Proto-Karen (PK) has previously been attempted and presented in different ways by a few scholars: Haudricourt (1946, 1953), Jones (1961), Burling (1969), Benedict (1972, 1979, 1983), Solnit (2001) and Manson (2009); some serious disagreements among them on some major points are to be found. To offer another new look at PK based on fresh data collected by myself (except Bwe), the PK phonology and lexicon (341 entries)² were reconstructed. Deliberately, available documented materials on the Karenic languages since 1799 onwards were not used for this reconstruction although they were consulted. Therefore, the reconstruction is based on the selected ten Karenic varieties spoken in Thailand, i.e. Northern Pa-O and Southern Pa-O (Northern branch, NK); Kayan, Kayah, Bwe (from Henderson 1997) and Kayaw (Central branch, CK); Northern Sgaw, Southern Sgaw, Northern Pwo and Southern Pwo (Southern branch, SK). A word list consisting of 2,000 items with English and Thai glosses was devised. For comparative purposes, only the obvious cognates found in the three branches (NK, CK, SK) or at least in two branches (NK&CK, NK&SK, CK&SK) were used. In following this method, most of the items in my field notes had to be eliminated. The correspondence patterns of the onsets, rhymes and tones were investigated and, then, the protoforms were reconstructed. The results were compared with the previous PK reconstructions and with the PTB forms reconstructed by Benedict (1972) and Matisoff (2003). Some important points of the findings are discussed.

1. Introduction

Karenic is a distinct cluster of languages (Van Driem, 2001) or a branch of Tibeto-Burman (Matisoff 1991, 2003; Bradley, 1997) of the Sino-Tibetan language family. They are spoken in the border area of Thailand and Myanmar, a long strip of land from the north to the south. Some Christian Sgaw Karen have migrated to the Andaman Islands and also to the United States of America, Europe and Australia due to the wars with the Burmese. In Myanmar, there are at least sixteen groups of the Karen: Pa-O, Lahta, Kayan, Bwe, Geko, Geba, Brek, Kayah, Yinbaw, Yintale, Manumanaw, Paku, Sgaw, Wewaw, Zayein and Pwo (Ethnologue, 2009). Only two groups, i.e. Pwo and Sgaw, permanently live in the northern and western areas of Thailand. Most of them are Thai citizens. The total Karen population scattered in fifteen provinces is about 500,000. Those who live on the highlands still use their native languages (Pwo and Sgaw) in everyday life. The ones who live or work in towns and cities are unable to speak their own languages well anymore or have shifted to the languages spoken by the majority of the areas. The Sgaw living in Thailand prefer to be called /pʸa³³ kə³³ ɰɔ³³/ meaning 'human being', not "Kariang", "Karang" nor "Yang" which they think of as having a bad connotation. Pwo people call themselves "Phlong" or "Phlow" meaning 'human being'. A large number of Karen refugees can be found in many refugee camps situated near the Thailand-Myanmar border area, such as those in the provinces of Tak, Mae Hong Son, Ratchaburi and so on. Besides the Sgaw and Pwo, four more groups have been recorded: Pa-O, Kayah, Kayan and Kayaw. See Figure 1. Most of the Pa-O in Thailand are factory labourers in Central Thailand. The Kayah do dry-field cultivation in Mae Hong Son province, while the Kayan and Kayaw are part of the tourism business in the North. A rough and vague estimation of the total population of the Karen ethnic groups in Myanmar, Thailand, Europe, America and Australia is between 6-12 million (Manson, 2009).

¹ This paper is a research report of the sub-project "Proto-Karen" which comes under the project "Karen Linguistics" funded by the Thailand Research Fund (TRF) from 2009-2012.

² For the public use, these 341 reconstructed protoforms can be found on the online STEDT database of the University of California at Berkeley.

According to the Bibliography of Karen Linguistics compiled by Manson (2004, 2010) and my knowledge of the more recent works done in Canada at the University of Ottawa by Brunelle and his students³ and by our research team in Thailand at Chulalongkorn University, one can see that Karenic languages have widely been studied in comparison with the other indigenous languages of Southeast Asia, especially Sgaw Karen and Pwo Karen. Due to Manson's contribution, a brief overview of Karen linguistic studies from the past to the present, about two centuries



Sgaw



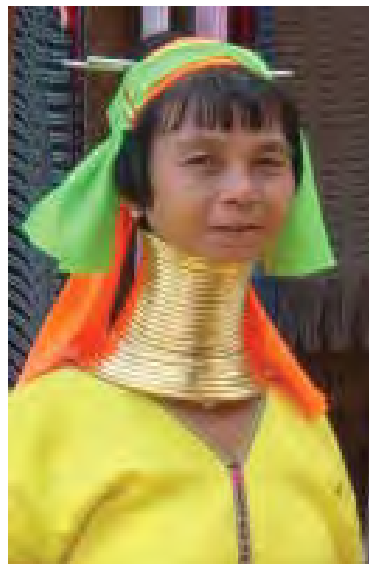
Pwo



Pa-O



Kayah



Kayan



Kayaw

Figure 1 Ethnic Karen in Thailand (with courtesy of the "Karen Linguistics Project", Chulalongkorn University and the Thailand Research Fund)

³ See more information (mostly term papers resulting from their Linguistic Field Methods class) in the References.

(1799-2010), can be made. The older works seem to focus on five aspects: language survey & usage, handbook & primer, vocabulary & comparative vocabulary, dictionary & thesaurus and grammar book & grammatical sketch. From 1946 onwards, more modern linguistic-oriented papers, research monographs, M.A. theses, Ph.D. dissertations and so forth on various linguistic aspects, i.e. phonetics,⁴ phonology, morphology & syntax, discourse analysis and comparative & historical study have appeared. The research works both published and unpublished on Karen linguistics from 1799-2010 are to be found in the Bibliography of Karen Linguistics by Manson (2004, 2010). In this paper, I will mainly focus on those related to my present study, especially the more recent ones.

Unfortunately, I have no access to Haudricourt's important papers written in French. However, his valuable views on Proto-Karen can be synthesised from a few secondary sources, e.g. Luce (1959), Benedict (1972), Henderson (1979), Mazaudon (1977) and Weidert (1987). Because of his wide-ranging interest in the many language families of Mainland Southeast Asia, he seems to have a good sense of what SEA languages should be like from both the synchronic and diachronic perspective. For example, his reconstruction of the three categories of PK initials, the preglottalised set (e.g. *ʔb, *ʔd, etc.) which comes under the voiceless category (Category II), the two tones, *A and *B, and later, an addition of one more tone (B ✓) to cope with the greater number of patterns of tonal correspondences pointed out by Luce (1959) and Jones (1971). Mazaudon (1977) and Weidert (1987) have expanded Haudricourt's opinions on the Proto-Karen tones and their development. To clarify his views, both authors devote one section of their monographs to discussing the development of the tonal systems in some modern Karenic languages from the Proto-Karen tones.

After his lengthy negative comments on Jones (1961) which I agree with, Burling uses Jones's data on Pa-O, Pwo (two dialects), Palaichi and Sgaw (two dialects) to reanalyse and present a new version of the Proto-Karen phonology and lexicon. Instead of two tones, i.e. *high and *low in both non-checked and checked syllable types as proposed by Jones, he reconstructs six proto-tones, namely, *1, *2, *3, *4, *ʔ1 and *ʔ2, the first four tones occur in non-checked syllables while the last two tones occur in checked syllables.⁵ His PK reconstruction looks much simpler and easier to follow than that of Jones (1961) who uses more complex methods and ways of handling problems. Burling reconstructs more tones but fewer of the other aspects. His solution makes the PK tonal system look like that of some Modern Karenic languages which have four tones in non-checked or smooth syllables and two tones in the checked ones,⁶ such as Pa-O, Pwo, some varieties of Sgaw and so forth. This suggests that the four-tone languages and varieties are conservative, i.e. their tonal systems are exactly like that of PK. However, some modern Karenic languages do have two tones, e.g. the Sgaw Karen varieties spoken in Pai district, Mae Hong Son province (L-thongkum, 2012), Geba and Thalebwa (Shintani, 2003), or three tones, e.g. Sgaw (Ratanakul, 1986; L-thongkum, 2012). Since PK has six tones, it implies that the "Great Tone Split" never occurred in Karenic languages. According to Matisoff (2003), Proto-Tibeto-Burman (PTB) is non-tonal; therefore, it is unlikely that the six tones in PK were born at the same time. With regard to tonal development, there is only a merger between tones, never a tone split, i.e. from the PK six tones to five or four tones in modern Karenic languages spoken nowadays. In the other tonal languages of Southeast Asia, splitting occurred first and was then followed by merging, or in other words, fewer tones become more tones due to the devoicing of the voiced initial obstruents (e.g. *b > p/ph) or the voicing of the voiceless initial sonorants (e.g. *hm > m). Moreover, Burling's reconstruction of the PK lexicon, which consists of many hundred protoforms, is heavily based on Sgaw and Pwo which are members of the Southern-Karen branch (Bradley, 1997; Kauffman, 1993) or the Sgaw-Pwo branch (Shintani, 2003). Even though Pa-O is included, it seems to play a marginal role due to the lack of cognates. The data on Central Karen languages is completely ignored in both Jones (1961) and Burling (1969). It looks as if the valuable works on Proto-Karen by Haudricourt have mostly been appreciated in Europe, see Luce (1959), Henderson (1979), Mazaudon (1977) and Weidert (1987).

⁴ With regard to the acoustic and experimental studies of Sgaw and Pwo Karen, the contributors are: Abramson (1995), Brunelle and Finkeldey (2011), Intajamornrak (in press), Teeranon (in press), Jitwiriyant (in press), Kerdpol (in press) and Pittayaporn (manuscript). Also, there are online term papers by Thomas & Alves-Soares (2011) and Finkeldey (2011).

⁵ Burling's tones *1 and *2 equal *B, *3 and *4 equal *A and *ʔ1 and *ʔ2 equal *D.

⁶ Synchronically analysed, these two tones which are sometimes called "glottalised tones", can be regarded as the allotones of the high and low tones occurring in non-checked syllables.

Henderson (1979) refers to Benedict (1972) and the items which show the reflexes of the PST tone *A, the PTB tone *A and the PK tone *A, for example, in the protoforms meaning ‘water, new, bear (n.), white, name’ and so on. As for the PST tone *B, the PTB tone *B and the PK tone *B, the protoforms such as the ones meaning ‘bitter, carry (on the back), dog, eat, ear, tail, female, mother’ are listed.⁷

Bennett (1992) briefly gives an overview of the comparative and historical phonology of the Karen languages in the last section of his seven-page paper (handout) on Karen phonetics and phonology: the loss and retention of final sonorants and the tonal developments caused by the register split of initial consonants including the developments of the “minor tones” from *B ✓ and the developments of the D tone in stopped or checked syllables. Finally, he ends his explorations by talking about the vocalic developments (vowel raising in Central Karen) and the consonantal developments ([w] [w] [r]) in the words meaning ‘bark fibre, bone, bowels, dry, side of body, snake’.

In his six-page paper “Another look at Proto-Karen” (looking like the handout distributed at ICSTLL 34, organised in Kunming, 2001), Solnit divides the contents into six sections: (1) The Introduction which consists of the etymon *k-ŋaŋ [k-ŋaŋ] ‘Karen’, the Karenic languages from which his data was drawn, the locations where they are spoken and the previous works by Haudricourt, Luce, Jones, Burling, Benedict and Mazaudon; (2) Tones and initial consonants: the proto-tones *A *B *B ✓ *D and the relationship between these tones and the laryngeal features of initial consonants conditioning tone splits and mergers, examples of the PK tone *A and the developments of stops; (3) PK phonology: an initial consonant inventory, the preservation of the laryngeal contrasts in obstruents in Blimaw⁸, the problem of mid-series labial and dental obstruents; (4) Rhymes: the PK basic vowel system and rhymes, the Central Karen vowel shift (*a, *e, *o, *i); (5) Subgrouping: the north-to-south transition of the seven words (spider, behind, fish, heavy, cooked rice, aim/point at, shadow); and (6) Relations: genetic relationship and contact relationship. Solnit’s work can be viewed as a speculation on PK. It looks like a rough sketch of his PK monograph which he would like to complete in the future. Interestingly, Bennett (1992) and Solnit (2001) seem to base their PK analyses on Haudricourt’s previous views.

Besides the Bibliography of Karenic Linguistics (2004) and the revised and updated version (2010), Manson has contributed approximately fourteen papers (mostly online) of his synchronic and diachronic studies of Karen linguistics from 2001-2011. However, only his works on the reconstruction of Proto-Karen, in his own words a “preliminary discussion” and the subgrouping or classification of the Karenic languages, which are relevant to my present study will be reviewed here.

In his 26-page article (Manson, 2009) he focuses his presentation on the following four major points: (1) external and internal classification of Karen; (2) summaries and comments of the previous reconstructions (Haudricourt, 1946 & 1953; Jones, 1961; Burling, 1969);⁹ (3) tonal development;¹⁰ and (4) a preliminary reconstruction of PK initials and rhymes with the correspondence patterns in eleven Karenic languages drawn from Luce (1985), Kauffman (1993) and his own field notes: Pwo (T), Pwo (D), Sgaw, Paku, W.Bwe, Geba, N.Pa-O, Kayan, W.Kayah, E.Kayah and E.Pho.¹¹

According to Manson’s preliminary reconstruction, the reflexes of the 31 PK initials and 24 PK rhymes are as follows: *ph, *th, *ch, *kh, *p, *t, *c, *k, *ʔ, *b, *d, *ɓ, (*g), *hm, *hn, *hŋ, *m, *n, *ŋ, *s, *h, *hw, *hl, *hj, *hr, *ʔw, *ʔl, *w, *l, *j, *r, *i, *e, *ɛ, *ə, *a, *u, *o, *ɔ, *am, *an, *aŋ, *aʔ, *ɛm, *ɛn, *ɛŋ, *ɛʔ, *eŋ, *eʔ, *iŋ, *oŋ, *ɔn, *ɤn, *ɤŋ and *uʔ. With regard to consonant clusters, they are: *phl, *pl, *bl, *br, *thw, *khl, *khw, *sw.

⁷ My reconstruction of the protoforms of these words agrees very well. For the PK *A, see no. 324, 202, 15, 330, 198 and for the PK *B, see no. 24, 40, 87, 98, 95, 289, 112, 113 in the lexicon part of this paper.

⁸ Henderson (1979) mentions this point and gives some examples. Many more examples are to be found in her *Bwe Karen dictionary* (1997).

⁹ Manson (2009: 11) concludes, “Comparing the three published reconstructions, Haudricourt’s analysis continues to be the most natural and predictive of the three reconstructions, ...”.

¹⁰ This section, which consists of the reviews of previous works, the Great Tone Split, the development of tones in Karen and Karen tone-box, seems to be the heart and soul of the article.

¹¹ For a full account of his reconstruction, he refers to Manson (in preparation). As yet, I have not seen it.

Even though Manson (2009) thinks that Haudricourt's analysis is the most "natural" and "predictive" among the existing reconstructions, he does not reconstruct $*\beta/*\beta_b$, $*d/*d_b$, $*x$ and $*\gamma$ as does Haudricourt. Moreover, $*h_j$, $*h_r$, $*\beta_w$ and $*\beta_l$ are reconstructed while in Haudricourt they are not.

With regard to the classification of Karen languages, Manson (2009, 2011) provides a comparison of the classifications done by Jones (1961), Burling (1969), Kauffman (1993), Bradley (1997), Manson (2002) and Shintani (2003). Based on shared innovations with the emphasis on rhyme development, Manson (2011) proposes a new classification of Karen languages as shown in Figure 2.

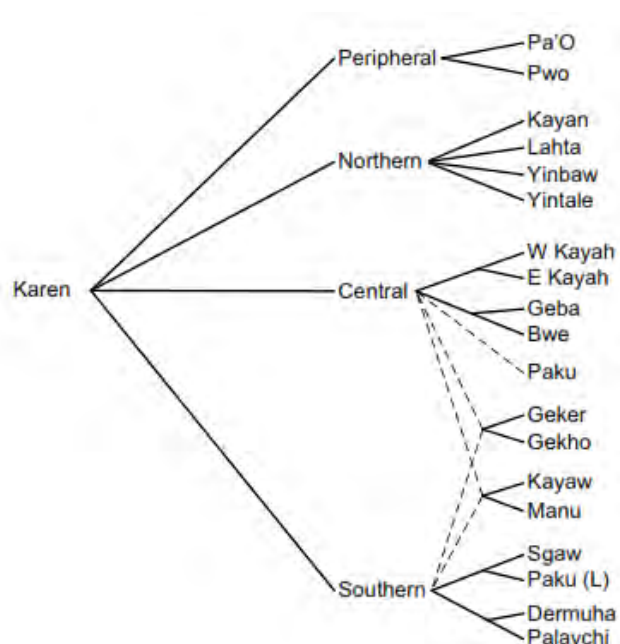


Figure 2 Classification of Karen languages (Manson, 2011: 8)

To summarise, one of the most controversial aspects of previous PK reconstructions seems to have been the number of the PK tones and the development of tones in Karenic languages:

Jones (1961): 2 tones ($*high$ and $*low$ in both non-checked and checked syllables)

Burling (1969): 6 tones (4 in non-checked syllables and 2 in checked syllables)

Haudricourt (1946): 3 tones (2, i.e. $*A$ and $*B$ in non-checked syllables and $*C$ in checked syllables)¹²

(1975): 4 tones (3, i.e. $*A$ [level], $*B$ [falling] and $*B \checkmark$ [rising] in non-checked syllables and $*C$ in checked syllables)

Solnit (2001): 4 tones (3, i.e. $*A$, $*B$ and $*B \checkmark$ in non-checked syllables and $*D$ in checked syllables)

Shintani (2003): 4 tones (3, i.e. $*1$, $*2$ and $*2 \checkmark$ in non-checked syllables and $*3$ in checked syllables)

Manson (2009, 2011): 4 tones (3, i.e. $*A$ [high], $*B$ [low] and $*B \checkmark$ [mid] in non-checked syllables and $*C$ [mid+glottal] in checked syllables)

¹² A detailed discussion and expansion of Haudricourt's views is to be found in Mazaudon (1977).

It is interesting to point out that Haudricourt's proposal (1975) has been the most accepted by current Karen linguistic specialists.¹³

Being aware of the previous reconstructions of Proto-Karen, somewhat different methods and different sets of Karen language data were used for my comparative and historical study and its results presented in this paper. The main contents are divided into six sections: methodology, field sites and language data, the correspondences of onsets, rhymes (vowel and vowel+final) and tones, PK phonology, PK lexicon (341 items) and remarks on the internal and external classifications.

2. Methodology

The previous reconstructions of Proto-Karen have mostly been based on the available documented materials in various forms, e.g. language learning materials, dictionaries, personal field notes and so on. More or less the same materials have been used for drawing suitable cognates in modern Karenic languages for the PK reconstructions which have been proposed from 1961 to the present.¹⁴ In the recent studies of PK, e.g. Bennett (1992), Solnit (2001), Shintani (2003) and Manson (2009, 2011), the analysts' personal field notes have been added.¹⁵ This method is good when the old materials are regarded as the written record of the conservative pronunciation of the past that is assumed to be closer to the proto-language, provided that the transcription done by untrained field workers and non-native language teachers is accurate or good enough. On the other hand, we have to accept the fact that the data on present-day pronunciation even though collected by well-trained linguists (the authors) has gone through various kinds of language variation and change due to language and cultural contact. It is doubtful whether the two types of data are comparable. Being aware of the problems mentioned above, for my reconstruction of the Proto-Karen phonology and lexicon presented in this paper, I deliberately used only fresh data collected by myself from 2009-2011. However, I skimmed through the available written materials which I could get hold of before devising a word list consisting of 2,000 items with Thai and English glosses. This word list was used as a guideline for data eliciting when interviewing my language consultants during my fieldwork in the North and West of Thailand. The elicited data obtained from ten Karen languages and/or their varieties (see the next section) was transcribed using IPA symbols. The minute and unnecessary phonetic differences were eliminated for the sake of a generalisation of the consonant, vowel and tone systems in each Karenic language variety before searching for cognates. For example, the high tones, mid tones and low tones were specified as 55, 33 and 11, respectively, no matter what the detailed phonetic characteristics of these three tones are. As for the contour tones, the more common ones are the falling tones, i.e. 53 or 31. Only a few of the Karen varieties, for example, the variety of Northern Pwo spoken at Ban Dong Dam in Li district of Lamphun province has a rising tone (35) which is rather uncommon. Therefore, the minimisation of phonetic differences is not difficult to achieve.

Even though Karenic languages have been classified with different criteria by different linguists as summarised in the first section of this paper, in selecting the Karenic languages to be used as the representatives of each branch, I adopted Kauffman's classification which is geographically based. I do not think that a definite subgrouping can be done with certainty until many more Karenic languages spoken in Myanmar are carefully studied. In my study, two varieties of Pa-O (N.Pa-O and S.Pa-O) were chosen to represent Northern Karen (NK); Kayan, E.Kayah, W.Bwe (Henderson's Blimaw) and Kayaw for Central Karen (CK); two varieties of Sgaw (N.Sgaw and S.Sgaw) and two varieties of Pwo (N.Pwo and S.Pwo) for Southern Karen (SK). For comparative purposes, only the obvious cognates found in the three major branches (NK, CK, SK) or at least in two branches (NK & CK, NK & SK, CK & SK) were used. In following this method, most of the items obtained from my language consultants during the interviews had to be eliminated. The correspondence patterns of the onsets (initial consonants), rhymes (vowels and vowels+final consonants) and tones were analysed, and then, the PK phonology and protoforms (341 items) were reconstructed.

¹³ For more information on the development of Karen tones, see Mazaudon (1977), Weidert (1987) and Manson (2009).

¹⁴ They are to be found in Manson (2004 and 2009).

¹⁵ Kato (2009) does the same thing when reconstructing the Proto-Pwo Karen roots.

3. Field sites and language data

The data on the selected ten Karenic language varieties was solely collected by myself with the devised word list prepared as a guideline when working with my Karen language consultants. The “Word List for Investigating Karenic Languages” comprises twenty-three sections:¹⁶ Action verbs; Stative verbs; Body parts and secretion; Health and diseases; Fauna; Parts of plant; Flora; Natural objects and phenomena; Manmade objects and construction; Food stuff; Culture and society; Kinship terms; Numerals; Classifiers; Measurements; Colours; Time; Direction and location; Pronouns; Question words and Miscellaneous.

The field sites are as follows:

N.Pa-O: Huai Khan village, Mok Champae sub-district, Mueang district, Mae Hong Son province

S.Pa-O: Wat Thaiwatthanaram, Mae Sot district, Tak province

Kayah, Kayan, Kayaw: Huai Suea Thao village, Pha Bong sub-district, Mueang district, Mae Hong Son province

N.Sgaw: Huai Khom village, Mae Yao sub-district, Mueang district, Chiangrai province

S.Sgaw: Pa La-U village, Huai Sat Yai sub-district, Hua Hin district, Prachuap Khiri Khan province

N.Pwo: Dong Dam village, Dong Dam sub-district, Li district, Lamphun province

S.Pwo: Chao Wat Yang Daeng village, Kaen Magrut sub-district, Mueang district, Uthai Thani province

W.Bwe: [Henderson’s Bwe Karen dictionary: with texts and English-Karen word

list (1997)]¹⁷

It is noticeable that the *N.Pa-O* variety has a lot of Tai borrowings since its speakers have a good relationship with the Shan living in Myanmar and Thailand, while those who speak *S.Pa-O* were in contact with the Mon before migrating to Thailand. The main distinction between the two varieties: (1) the phonetic characteristics of the B12 and B3 tones (<*B), i.e. B12 = 55, B3 = 53 in *N.Pa-O* but B12 = 33 and B3 = 55 in *S.Pa-O*; and (2) mostly, clear monophthongs in *N.Pa-O* while on-gliding and off-gliding vocalic quality plus a heavier phonation type in *S.Pa-O* due to the influence of Mon. Both varieties have final stops (p, t, k, ʔ) and final nasals (m, n, ŋ) but sometimes there is no agreement, i.e. different kinds of stops and nasals. Basically, *E.Kayah*, *Kayaw* and *W.Bwe* have three tones: high (55), mid (33) and low (11). I have noticed that in *Kayah* and *Kayaw*, monosyllabic citation forms tend to be pronounced with [-ʔ], whereas in compounds, phrases and sentences, especially in a stressed position, the final [-ʔ] always disappears. Vowel harmony and tone sandhi are common features. Synchronically analysed, *Kayan* and some varieties of *Kayah* can be said to have four tones in non-checked syllables; however, the fourth tone, i.e. mid-falling (31) has not been found in cognates. This tone occurs in a few words and I suspect that most of them are loanwords, especially from Burmese and/or Tai. The final -ŋ in *Kayan* comes from PK *-m, *-n or *-ŋ. It is fortunate that some *W.Bwe* languages have implosives or preglottalised stops and voiceless sonorants, which I think are a retention not an innovation, in their phonological systems.

¹⁶ The data on Karen animal names (fauna), numerals and classifiers has also been used for my other papers: L-thongkum 2011 and 2012.

¹⁷ *W.Bwe Karen* (Blimaw, Geba) is very important for PK reconstruction because it has the implosives or preglottalised obstruents: ɓ/ʔb and d/ʔd, including voiceless sonorants, e.g. hn, hl, and so forth.

The nasal vowels in Pho which come from the final nasals in PK are now becoming oral vowels plus final **-ŋ** in the speech of young Pho speakers. Most or all of the Pho varieties have four tones in non-checked syllables, i.e. PK *A > A1-23 and PK *B > B12-3. The unusual split pattern in the A-tone column is useful for the reconstruction of some glottalised sonorants. The phonetic characteristics or tone shapes of the A1, A23, B12, B3, D12 and D3 tones in N.Pwo and S.Pwo are different: 35, 55, 33 and 11'' in N.Pwo and 53, 31'', 55 and 33 in S.Pwo, respectively.

Based on the development of tones A and B, Sgaw has at least four major varieties. The varieties which I have had the opportunity to work on have two, three or four tones in the non-checked syllable, see Figure 3. Among the ethnic Karen in Thailand, the Sgaw are the majority, possibly one-third. Therefore, Sgaw has widely been studied in comparison with the other Karenic languages spoken in Thailand as can be seen in the Bibliography of Karen Linguistics (Manson, 2010). However, I only used my own field data for the PK reconstruction presented here. Among the Karenic languages, I think that Sgaw is the easiest to work on because of its simpler phonetics and phonology. More information on the profiles of the six ethnic Karen groups existing in Thailand is to be found in Schliesinger (2000).

(1)

	*A	*B	*D
1	33	11"	45'
2		11"	45'
3		11"	21'

(3)

	*A	*B	*D
1	55	11"	45'
2		11"	45'
3		11"	21'

(2)

	*A	*B	*D
1	33	31"/21'	45'
2		31"/21'	45'
3		11"	21'/53

(4)

	*A	*B	*D
1	55	45'	21'
2		45'	21'
3	33	31	11"

Figure 3 Tones in the four varieties of Sgaw

In addition to the four tonal patterns illustrated above, the Sgaw variety spoken in Pai district, Mae Hong Son province, has an unusual split in the B column: B1 = 11'', B2 = 21' and B3 = 11'', D12 = 45' while D3 = 53. Even though there are two tones in the B column, the non-checked syllables having the Category-I initials (e.g. *ph, *hm, *s) and the Category-III initials (e.g. *b, *m, *l) have the same tone, i.e. 11'', while those having the Category-II initials (e.g. *p, *ɓb, *ɓ) have tone 21'. This means that the non-checked syllables in PK, i.e. *CV□ has become CVɓ²¹ (checked-syllable) in this modern Sgaw variety, provided that their initials or onsets are p-, t-, c-, k-, d- (<*ɓd-), b- (<*ɓb-) and ɓ; for example, *ce□ > ceɓ²¹ 'left (side)', *ɓe□ > ɓeɓ²¹ 'dung, excrement', *ɓde□ > deɓ²¹ 'frog', *ɓbaŋ□ > bɓɓ²¹ 'bamboo shoot' and so on. Instead of the more common B12-3 like the split-pattern in (2), it has become B13-2.

4. Sound correspondences

The correspondences of the initials or onsets as well as the ones of tones seem to have fewer problems, see Tables 1 and 3. With regards to vowel correspondences, the picture is not so clear, consequently, the reconstruction of PK vowels is more problematical, difficult and time consuming. If we assume that PK vowels should be more or less similar to the spellings found in old Sgaw and Pwo texts and dictionaries written in the Burmese-based and Roman-based scripts invented about two hundred years ago, then, a reconstruction of the vowels is not difficult. However, it is not so easy when the reconstruction is solely based on the data drawn from many spoken Karenic language varieties collected in the present. Perfect patterns of sound correspondences, especially of vowels in the Karen case, cannot be expected for many reasons, e.g. vowel harmony phenomena and borrowings which are the result of various layers of language contact from the past to the present in the areas where Sino-Tibetan languages are spoken as pointed out by LaPolla (2001). The causes of the irregular sound changes, in some cases, cannot be explained. This fact should be accepted; therefore, we must keep in mind that the reconstruction of a proto language is tentative. Reconstructions which are based on different sets of data can yield different results, more or less.

Some of the unusual patterns of vowel correspondences could have stemmed from the loss of some final consonants which have no traces in the present-day languages. Thus, it is likely that the correspondences of rhymes should receive more of our attention, at least in Karen, see Table 2.

Table 1 Correspondence patterns of the onsets

PROTO-KAREN	N. Pa-O	S. Pa-O	Kayan	Kayah	W.Bwe	Kayaw	N. Sgaw	S. Sgaw	N. Pwo	S. Pwo	No.
*ph	ph	ph	ph	ph	ph	ph	ph	ph	ph	ph	6, 12, 34, 36, 103, 121, 182, 256, 281
*phl	phl	phl	phl	phl	-	phl	phl	phl	phl	phl	57
*phr/*phw	phr/phw	phw	phw/ph	phw/ph	phr/phw/ph	ph	phɣ	phɣ/ph	ph/khw	ph/khw	30, 39, 148, 166
*th	th	th	th	th	th	th/h	th	th	th	th	5, 15, 22, 27, 51, 115, 143, 167, 175, 218, 221, 226, 243, 252, 270, 277, 287, 298, 319, 324, 326, 327
*thw	thw	thw	thw	thw	thw	th	thw/chw	thw	thw	thw	87, 238, 258, 292
*thr	-	-	th	th	thr	th	th	th	th	th	237
*ch	ch/ç	ch/ç	ch/th	ch	ʃ	ch/ç	ch	ch	ch	ch	102, 110, 152, 177, 228, 251, 278, 291, 314
*chw	chw	chw	chw	chw	fw	chw	chw	chw	chw	chw	64, 325
*chj	çj	çj	ç	ch	ʃ	ç	ch	ch	ch	ch	47, 268
*chr	ch	ch	ch	ch	-	ch	chɣ	chɣ	-	-	322
*kh	kh	kh	kh/h	kh	kh	kh	kh	kh	kh	kh	9, 24, 56, 68, 73, 79, 83, 162, 184, 255, 301, 304
*khw	khw/kh	kh	khw/kh	khw/kh	khw/kh	kh	khw/kh	khw/kh	khw/kh	khw/kh	61, 183, 190, 203
*khl	khl	khl	khl	khl	khl	-	khl	khl	khl	khl	170, 186, 264
*khr	khr/khj/s	khr/khj/s	khj/ch	khr	kh/x	kh/c/x	x	x	k/x	k/x	29, 80, 101, 259
*khrw	ch	ch	chw	khrw	khw	ch/x	x	x	x	xw	31
*s	s	s	θ	s	θ	s	s	s	s	θ	41, 52, 76, 128, 142, 156, 180, 202, 222, 229, 242, 244, 253, 300, 305
*sw	sw/s	sw/s	θw/θ	sw/s	θw/θ	s/β	sw/s	sw/s	sw/s	θw/θ	25, 46, 173, 179
*sl	l	l	l	l	l	l	s	s	s	-	136
*sr	s	s/ç	θ	s	θr	s	çɣ/s	çɣ	çj/sj/s	θ	10, 32, 144, 149

Table 1 (continued)

PROTO-KAREN	N. Pa-O	S. Pa-O	Kayan	Kayah	W.Bwe	Kayaw	N. Sgaw	S. Sgaw	N. Pwo	S. Pwo	No.
*h	h	h	h	h	h	h	h	h	-	ɣ	48
*hm	m	m	m	m	m	m	m	m	m	m	66, 112, 114, 165, 197, 239, 246, 271, 332, 336
*hn	n	n	n	n	n	n	n	n	n	n	18, 67, 71, 171, 254, 313, 338
*hnw	n	n	nw	-	nw	n	nw	nw	nw	nw	337
*hn	j	j	ʃ	j	j	j	ɲ	ɲ	j	j	97, 120, 288
*hn	ŋ	ŋ	ŋ	ŋ	-	-	ɲ	ɲ	-	-	127
*hw	w	w	hw	w	h	-	w	w	w	w	8
*hl	l	l	l	l	l	l	l	l	l	l	14, 62, 92, 158, 164, 181, 231, 276, 317
*hr	r	r	r	r	h	r	-	-	-	-	298
*hrw	r	-	rw	rw	-	r	w	w	w	w	232
*p	p	p	p/ph	p/ph	p/ph	p/ph	ph	ph	p/ph	p/ph	36, 40, 106, 176
*pw	bw	bw	bw	b	pw	b	pɣ	-	-	-	129
*pl	pl/l/p	pl/bl/p	pl/bl	pl	pl	pl/p	bl/l/p	bl/l/p	phl/l/p	phl/l/p	13, 55, 219, 257, 262, 272, 290, 295
*pj	pj	pj	bj	-	b	-	b	-	p	b	118
*pr	pr/phr	pr	phr	phr	phr/p	pr/phr	bɣ	pr/bɣ	pj/phj	pr/pj	273, 316
*t	t	t	t/th	t/th	d	t/th	t/th	t/th	th	th	3, 216
*tr	-	-	-	r	tr	tr	kr	r	-	ɣ	191
*c	c	c	c/ch	c/ch	c/ʃ	c/ch	c/ch	c/ch	c/ch	c/ch	161, 220, 248, 302, 329
*cw	c	c	cw	cw	-	c	~w	-	-	~w	284
*k	k/kh	k/kh	k/kh	k/kh	k/kh	k/kh	k/kh	k/kh	k/kh	k/kh	42, 43, 49, 65, 111, 212, 303

Theraphan Luangthongkum 2

Table 1 (continued)

PROTO-KAREN	N. Pa-O	S. Pa-O	Kayan	Kayah	W.Bwe	Kayaw	N. Sgaw	S. Sgaw	N. Pwo	S. Pwo	No.
*kw	kw/w	kw/w	khw/hw	khw/w	kw/khw/w	khw/hw/kh	kw/khw	kw/khw	kw/khw	kw/khw	17, 45
*kl	kl/khl/k	kl/k	kl/khl	kl/khl/k	kl/khl/k	kl/khl/k	kl/khl	kl/kh	kl/khl	kl/khl	28, 69, 119, 138, 309
*kj	khj	-	kh	khj	-	-	k	k	k	k	7
*?b	b	b	b	b	ḃ	b	b	b	b	b	11, 44, 53, 58, 60, 201, 223, 234, 236, 247, 297, 339
*?bw	bw	bw	b	b	ḃ	b	kw (?)	w	-	kw (?)	330
*?bl	pl	pw	bw	b	ḃ	b	b	b	b	b	109
*?d	d	d	d	d	d/d	d	d	d	d	d	21, 59, 72, 74, 84, 99, 126, 200, 233, 274, 305, 318, 323, 335
*?dw	dw/d	dw/d	dw	d	d	d	dw	dw	thw	thw	132, 296
*?l	l	l	l	l	l	l	l	l	l	l	193
*?j	ʔ	ʔ	ʔj	ʔ	ʔ	ʔ	ʔ	ʔ	ʔ	ʔ	199
*b	ph	p	p	p	b/p	p	p	p	ph	ph	209, 215, 224, 294
*bl	pl/phl	pl	pl/p	pl	bl	pl	pl/phl	pl	pl/phl	phl	4, 168, 260, 283, 307, 321
*br	phr	phr/pr	pr	phr	bw/b	pr/r	by/py/phy	py/y	y/xw/sj	y/xw/c	38, 54, 154, 205, 214
*d	th/t	t	t	t	d	t	t	t	th	th	2, 108, 115, 137, 146, 169, 225, 240, 299, 308
*j	c	c	c	c	c	c	c	c	ch	ch	275, 280, 341
*g	k/kh	k	k	k	g	k	k	k	kh	kh	100, 269, 286
*gw	w	w	gw	w	w	w	w	w	w	w	151
*gl	-	kl/kw	khl/k	kl	kl	kl/k	kl/k	kl/k	kw	kw	145, 213
*gr	s/ç	s/ç	c/ch	chr	-	c/x	x	x	x	x	16, 101, 153, 178

Table 1 (continued)

PROTO-KAREN	Pa-O	S. Pa-O	Kayan	Kayah	W.Bwe	Kayaw	N. Sgaw	S. Sgaw	N. Pwo	S. Pwo	No.
*m	m	m	m	m	m	m	m	m	m	m	27, 86, 88, 90, 107, 113, 194, 198, 267, 285, 289, 311, 333
*mw	mw/m	mw/m	mw/m	m	m	m	m	m	m	mw/m	93, 340
*ml	-	m	m	m	bl	-	ml	ml	ml	ml	312
*n	n	n	n	n	n	n	n	n	n	n	33, 35, 37, 75, 95, 104, 147, 204, 261, 263, 334
*ɲ	j	j	ʃ	j	-	-	ɲ	-	j	j	210
*ŋw	ŋw	ŋw	ŋw	-	m	m	-	-	-	-	265
*ŋj	ŋj/ŋ	ŋj/ŋ	j/ŋ	ŋj/ŋ/j	j	j	j/z	j	j	j	50, 116, 174
*w	w	w	w	w	w	w	w	w	w	w	187, 188
*l	l	l	l	l	l	l	l	l	l	l	1, 20, 82, 133, 134, 135, 139, 157, 160, 206, 227, 230, 282, 315, 320, 331
*lw	lw	lw	lw/l	lw	lw/l	l	lw/l	lw/l	l	lw/l	89, 131
*j	j	j	j/ʃ	j	j	j	j	j	j	j	19, 78, 96, 192, 195, 211, 249, 310
*r	r	r	r	r	R	r	ɣ	r/ɣ	ɣ	w/β	117, 217, 266
*rw	rw	rw	rw	rw	w	r	ɣ	ɣ	ɣ	w	241
*rj	rj	rj	j	j	j	j	j	j	j	j	150

Table 2 Correspondence patterns of the rhymes

PROTO-KAREN	N. Pa-O	S. Pa-O	Kayan	Kayah	W.Bwe	Kayaw	N. Sgaw	S. Sgaw	N. Pwo	S. Pwo	No.
*i	i	i	i	i	i/ɪ	i (u)	i	i	ei (i)	ei (i)	25, 27, 45, 60, 65, 67, 81, 87, 89, 119, 125, 166, 188, 203, 212, 241, 265, 276, 280, 294, 333
*e	e	e	ai (iʔ)	i (e)	i (ɪ,e)	i (e)	e (ɛ)	e (ɛ)	i (ɛ)	i (ei)	18, 34, 38, 94, 114, 126, 149, 157, 161, 191, 200, 204, 216, 235, 250, 262, 289, 303, 307, 318, 323
*ɛ	ɛ (ai)	ɛ (ai)	ɛ	a (ɛ)	ɛ (e)	ɛ	ɛ/e	ɛ/e	ɛ/e	ɛ/e	64, 74, 129, 131, 132, 213, 233, 238, 331, 337, 340
*i	i	i	i (o)	ɔ	o (u)	i	i (iʔ)	i	ɤi (ou)	i (u)	40, 43, 112, 285
*ə	ə (i)	ə (i)	əi	o	o	ə	i	i	əi (ɔ)	i (o)	3, 59, 143, 176, 194, 234, 256
*a	a (i)	a	a (au)	e (a,u)	ɛ (e,a,u)	a (e,u)	a	a	a (ɛ)	a (ai)	4, 6, 8, 10, 13, 24, 37, 46, 49, 53, 54, 66, 86, 88, 95, 118, 120, 123, 127, 128, 135, 150, 151, 152, 153, 158, 167, 168, 171, 178, 182, 183, 185, 193, 205, 207, 210, 223, 244, 248, 272, 274, 278, 290, 321, 326, 330, 332, 336, 341
*u	u	u	u	u (o)	u (u)	u	u (o)	u	u (a,əu)	u (ə)	26, 35, 42, 51, 83, 113, 136, 220, 270, 273, 287
*o	o (ə,u)	o (u)	au (ɔ)	u	o (u,u)	u (o)	o (ɔ)	u (ɔ)	u (oʔ)	u (i,uʔ)	21, 22, 56, 70, 89, 92, 134, 137, 180, 209, 211, 253, 264, 286
*ɔ	au	au	ɔ/au	o	ɔ (o)	o	ɔ (o)	ɔ (u)	ɔ	ɔ (o)	30, 91, 121, 140, 144, 186, 201, 306, 329
*im	im	im	i	i	i	i	i	i	āi	ēi	275
*in	in	in	(əŋ)	i	i	i	i	i	āi	ēi	198, 239, 315
*iŋ	iŋ	iŋ	i	i	-	i	i	-	āi	ēi	304
*eŋ	eŋ	eŋ	eŋ (e)	a/ai	e	e (ə)	e (i)	e (i)	āi	ēi	77, 122, 130, 247, 257, 298, 338
*eN	en	-	e	ai	e	e	e	e	-	ēi	240
*em	em	em	aŋ	-	-	i	e	i	āi	āi	231
*en	en	en	aŋ	i	-	i	i	i	āi	āi	199

Table 2 (continued)

PROTO-KAREN	N. Pa-O	S. Pa-O	Kayan	Kayah	W.Bwe	Kayaw	N. Sgaw	S. Sgaw	N. Pwo	S. Pwo	No.
*eŋ	eŋ	eŋ	i	a	-	i	e	i	äi	äi	325, 327
*eN	em/eŋ	em/en	i/aŋ	a,i	-	-	e/i	i	äi	äi	23, 170
*im	um/om	əm/uəm	i	i/i	u	i	ə	i	ä	äi	19, 169, 245
*(w)in	in	uən	i	o	u	u	u	u	ä	äi	173
*iŋ	əŋ	əŋ	i	i	-	-	ə	i	ä	äi	334
*iN	in/im/un	in/uəm	i	i/o	u/o	i/u	ə/u	i/u	ä	äi	93, 148, 179, 258, 296
*əm	əm	əm	ə	ɔ	o	ə	ə	ə	ä	ä	300
*əŋ	əŋ	əŋ	əŋ/ə	ɔ	o	ə	ə	ə	ä (ä)	ä (äi)	2, 224, 242, 277, 302, 311
*əN	əŋ	-	ə	ɔ	-	ə	ə	ə	ä	-	85
*am	am	am	aŋ/a	e/a/e	e/a	ɔ/a	ɔ	ɔ	ē/ō	ō	15, 16, 98, 145, 310, 320
*an	an	-	aŋ	-	e	ɔ	ɔ	ɔ	ē	ō	9
*aŋ	aŋ	aŋ	aŋ	e	a	ɔ	ɔ	ɔ	ē	ō	5, 11, 44, 82, 90, 102, 162, 252, 269, 317, 339
*aN	an/aŋ/a	an/a	aŋ (i,i)	e/a/e	e/a/i	ɔ (a,i)	ɔ	ɔ	ē	ō	47, 68, 100, 196, 202, 208, 228, 268, 271, 292, 293
*un	un/ən	uən	aŋ	i	u/u	u	-	u	-	-	189, 222
*uŋ	əŋ	əŋ	aŋ	i	u	u	ɔ	ɔ	ä	ö	109
*uN	un/in/ən	uəm	aŋ	i	u/u	u	u	u	ä (ä)	ö	39, 105, 110, 221, 227
*-om	om	om	o	u	o	ə	u	u	äu	öu	219
*-oŋ	oŋ	oŋ	o (əu)	ɔ	o	ə	ə (ɔ)	ə (i,u)	äu (ə)	öü	57, 146, 147, 181, 282, 295, 328
*-ɔŋ	ɔŋ	ɔŋ	ɔ	o/ɔ	ɔ	o	o	u	ō	ö	55, 215, 225

Table 2 (continued)

PROTO-KAREN	N. Pa-O	S. Pa-O	Kayan	Kayah	W.Bwe	Kayaw	N. Sgaw	S. Sgaw	N. Pwo	S. Pwo	No.
*-ɔN	oŋ	oŋ/ɔ	ɔ/au	o	ɔ/u	o	o/ɔ	u	õ	õ	58, 71, 172, 312
*-ej	i	i	i (ai)	ɔ	i (e)	i	i (e)	i	ei	ei	28, 29, 32, 52, 69, 72, 75, 76, 79, 99, 133, 160, 165, 197, 206, 283, 291, 301, 313, 314, 324
*-ej	ɛ	ɛ	i	i	-	e	e	e	ai	ai	187
*-aj	ai (a)	ai	ai (i)	ai (a,e)	ɪ	i (a)	e	i	ai	ai	50, 164, 249, 261
*-aw	o	o	au	u	o	u	ɔ	ɔ	ɔ	o	97, 175
*-ow	u	u	u (əu)	o	u	u	i (u)	i	ou	ou	9, 14, 62, 195, 259, 266, 279, 281, 299, 305, 322
*-iʔ	eʔ	iʔ	iʔ/iʔ	i	i (a,i)	i	iʔ	aiʔ	aiʔ (i)	aiʔ (eiʔ)	41, 61, 73, 106, 119, 232, 309
*-it	ut	əut	iʔ	i	i	u	i	i	ei	i	31
*-(w)ik	uk	uʔ	iʔ	i	-	i	iʔ	-	-	aiʔ	284
*-eʔ	eʔ/iʔ	iʔ	iʔ	a	e (i,i)	e	iʔ	aiʔ/ai	aiʔ	aiʔ	12, 154, 163, 217, 230, 335
*-ek	eʔ	ek	iʔ	a	e	e	iʔ	ai	ai	?aiʔ	139
*-ɛʔ	ɛʔ	ɛʔ	ɛʔ	a	e/i	e	ɛʔ	ɛʔ, ɛ	eʔ	eʔ	96, 107, 108, 226, 246, 260, 308
*-əʔ	əʔ	əʔ	əʔ	ɔ	o	ə	uʔ	əi	əəʔ	əəʔ	111, 117, 141
*-aʔ	aʔ	aʔ	aʔ (i,u)	a/e	a (ɛ,i)	ɔ (a,u)	aʔ (uʔ)	aʔ/ai	aʔ	aʔ	7, 20, 36, 44, 115, 142, 155, 156, 159, 174, 214, 251, 255, 267, 317
*-at	at	at	ɛʔ	ɛ	ɛ	ɛ	ɛʔ/ɛ	ɛ	ɛ	ɛ	17, 116
*-aK	at	ap	ɛʔ	ɛ	ɛ	ɛ	ɛ	ɛ	-	ɛ	48
*-uʔ	-	-	-	i	-	i	uʔ	ɔuʔ	auʔ	ɔuʔ	243, 319

Table 2 (continued)

PROTO-KAREN	N. Pa-O	S. Pa-O	Kayan	Kayah	W.Bwe	Kayaw	N. Sgaw	S. Sgaw	N. Pwo	S. Pwo	No.
*-uk	uk	-	uʔ	i	u	i	uʔ	ɔuʔ	auʔ	ɔuʔ	63
*-oʔ	oʔ	uʔ (oʔ)	ɔʔ	ɔ	o	ə	ɔʔ/iʔ/uʔ	ɔuʔ	auʔ	ɔuʔ	1, 80, 190, 192, 236, 237
*-ok	ok	uʔ	ɔʔ	ɔ	o	ə	uʔ	ɔ	auʔ	ɔuʔ	33
*-ɔʔ	ɔʔ (əʔ)	ɔʔ	auʔ	ɛ	ɔ	o	ɔʔ	ɔʔ (oʔ)	oʔ	uʔ	27, 78, 103, 184, 218, 229, 254, 288, 316
*-ɔt	ɔt	ɔt	auʔ	-	-	ɔ	ɔʔ	oʔ	oʔ	uʔ	101

Table 3 Correspondence patterns of the tones

PROTO-KAREN	N. Pa-O	S. Pa-O	Kayan	Kayah	W.Bwe	Kayaw	N. Sgaw	S. Sgaw	N. Pwo	S. Pwo	No.
*A > A1	{ 31'' }	{ 31'' }	{ 53 }	{ 33 }	{ 55 }	{ 55 }	{ 33 }	{ 55 }	{ 35 }	{ 53 }	9, 10, 15, 18, 30, 34, 39, 47, 67, 68, 76, 79, 85, 92, 97, 102, 105, 121, 127, 148, 152, 165, 166, 167, 170, 175, 177, 182, 183, 189, 202, 203, 238, 239, 242, 245, 252, 253, 258, 259, 261, 263, 268, 270, 271, 281, 291, 298, 300, 304, 313, 317, 324, 325, 327, 332
> A2	{ 31'' }	{ 31'' }	{ 53 }	{ 33 }	{ 55 }	{ 55 }	{ 33 }	{ 55 }	{ 55 }	{ 31'' }	26, 28, 43, 58, 59, 65, 72, 77, 91, 119, 123, 129, 130, 185, 191, 193, 200, 201, 208, 220, 223, 233, 248, 257, 262, 273, 274, 290, 293, 296, 297, 302, 303, 307, 318, 323, 330, 339
> A3	33	53	33	11	33	33	{ 33 }	33	{ 55 }	{ 31'' }	16, 27, 38, 50, 54, 75, 82, 88, 93, 100, 122, 125, 131, 133, 134, 135, 147, 150, 151, 160, 169, 187, 188, 198, 206, 210, 211, 213, 215, 224, 240, 249, 269, 283, 306, 312, 320, 333, 334
*B > B1	{ 55 }	{ 33 }	{ 11 }	{ 11 }	{ 33 }	{ 11 }	{ 31'' }	{ 45' }	{ 33 }	{ 55 }	5, 6, 8, 14, 22, 24, 25, 29, 32, 46, 51, 52, 56, 57, 62, 64, 66, 71, 83, 87, 110, 112, 114, 120, 125, 128, 136, 143, 144, 158, 162, 164, 171, 173, 179, 180, 181, 186, 197, 221, 222, 228, 231, 244, 256, 264, 276, 277, 278, 279, 287, 292, 301, 305, 314, 322, 326, 336, 337, 338
> B2	{ 55 }	{ 33 }	{ 11 }	{ 11 }	{ 33 }	{ 11 }	{ 31'' }	{ 45' }	{ 33 }	{ 55 }	3, 11, 21, 23, 40, 45, 49, 53, 55, 60, 69, 70, 74, 81, 84, 94, 98, 99, 109, 118, 126, 132, 138, 140, 161, 172, 176, 199, 212, 216, 219, 234, 247, 272, 295, 305, 329
> B3	53	55	{ 11 }	{ 11 }	{ 33 }	{ 11 }	11''	31	11''	33	2, 4, 19, 35, 37, 86, 89, 90, 95, 113, 137, 145, 146, 153, 157, 168, 178, 194, 195, 196, 204, 205, 209, 225, 227, 235, 241, 261, 265, 266, 275, 280, 282, 285, 286, 289, 294, 299, 310, 311, 315, 321, 328, 331, 340, 341
*D > D1	{ 21' }	{ 21' }	{ 45' }	{ 55 }	{ 33 }	{ 33 }	{ 45' }	{ 21' }	{ 45' }	{ 21'/31' }	12, 27, 31, 41, 48, 61, 73, 80, 101, 103, 142, 155, 156, 159, 184, 190, 218, 226, 229, 232, 237, 243, 246, 251, 254, 255, 288
> D2	{ 21' }	{ 21' }	{ 45' }	{ 55 }	{ 33 }	{ 33 }	{ 45' }	{ 21' }	{ 45' }	{ 21'/31' }	7, 36, 63, 106, 111, 119, 141, 163, 236, 284, 309, 316, 335
> D3	45'	45'	21'	33	11	{ 33 }	21'	11''	21'	33'/45'	1, 20, 33, 78, 96, 101, 104, 106, 107, 108, 115, 116, 117, 139, 154, 174, 192, 214, 217, 230, 260, 267, 308, 317

Onsets

Category I (High series):

*Vl. aspirated stops	*ph	*th	*ch	*kh
*Vl. nasals	*hm(m)	*hn(n)	*hɲ(ɲ)	*hŋ(ŋ)
*Vl. fricatives		*s		*h
*Vl. approximants	*hw(w)	*hl(l)		
*Vl. rhotic		*hr(r)		

Category II (Mid series):

*Vl. unaspirated stop		*p	*t	*c	*k
*Glottal and glottalised stops	*ʔb(ɓ)	*ʔd(ɗ)			*ʔ
*Glottalised nasal			*ʔn		
*Glottalised approximants	*ʔw	*ʔl	*ʔj		

Category III (Low series):

*Vd. Stops	*b*	*d	*ɗ	*g
*Vd. Nasals	*m	*n	*ɲ	*ŋ
*Vd. Approximants	*w	*l	*j	
*Vd. Rhotic		*r		

Consonant clusters

*Cw-	*phw	*thw	*chw	*khw	*sw	*hnw	*hrw		
	*pw	*cw	*kw	*ʔbw	*ʔdw	*ʔnw			
	*mw	*dw	*gw	*ŋw					
*Cl-	*phl	*khl	*sl	*pl	*kl	*ʔbl	*bl	*gl	*ml
*Cr-	*phr	*thr	*chr	*khr	*sr	*pr	*tr	*br	*gr
*Cw-	*khrw								
*Cj	*chj	*pj	*kj	*ŋj	*rj				

It is noticeable that the reconstructed onsets or initials are quite complex with a large number of voiceless sonorants, glottalised sonorants and consonant clusters. This is due to the fact that some of the onsets derive from the reductions of the PTB prefixes, infixes, or the preceding part of compounds, in order to become monosyllabic words with tones (Matisoff, 1973), which is a process of tonal evolution and tonal development in SEA languages. Having more solid data on Central Karen languages in the future, perhaps in some cases, consonant clusters could be reduced and replaced by a reconstruction of sesquisyllabic words. The following are some of the examples of the reductions mentioned above:¹⁸

PTB *g/b/m-la-y > PK *bla^B 'arrow'

PTB *s-hywəy, *s-hwi-y > PK *swi^B 'blood'

PTB *b-r-gyat > *b-g-ryat > PK *grɔt^D 'eight'

PTB *b-wa > PK *ɔbwa^A 'white'

PTB *m-/s-lay 'tongue' > PK *hlaj^B 'lick'

PTB *g-p^wa, *r-p^wa, *r-wa > PK *hwa^B 'bamboo'

PTB *b-rey > PK *bre^A 'buy'

PTB *s-hwi-y > PK *swi^B 'blood'

PTB *s-le-y > PK *hli^B 'squirrel'

PTB *s-ni(:)ŋ > PK *hneŋ^B 'year'.

Rhymes

Two types of rhyme were reconstructed, i.e. *-V and *-VC; and among *-C are : *-m, *-n, *-ŋ, *-N¹⁹, *-j, *-w (*-p),²⁰ *-t, *-k, *K, *-ɿ

Based on my own field data and the data drawn from Henderson (1997), 57 rhymes were reconstructed. They can be divided into four types, as follows :²¹

Type I : *-i *-e *-ɛ *-i̇ *-ə *-a *-u *-o *-ɔ

Type II : *-im *-in *-iŋ *-eŋ *-eN *-ɛm *-ɛn *-ɛŋ *-ɛN *-im̄ *-in̄ *-iŋ̄ *-iN̄

*-ə̄m *-ə̄ŋ *-ə̄N *-am̄ *-aŋ̄ *-aN̄ *-un̄ *-uŋ̄ *-uN̄ *-om̄ *-oŋ̄ *-ɔ̄ŋ̄ *-ɔ̄N̄

Type III : *-ej *-ɛj *-aj *-aw *-ow

¹⁸ More examples are to be found in the seventh section of this paper.

¹⁹ *-N and *-K in the rhymes *-VN and *-VK are neither a uvular nasal [N] nor a uvular stop [-q]; -K means 'one of the nasals, i.e. *-m, *-n or *-ŋ' and -K means one of the final stops, i.e. *-p, *-t or *-k.

²⁰ The rhyme -ap only occurs in S.Pa-O. It corresponds to -at in N.Pa-O.

²¹ See more detail in Table 2.

Type IV : *iɿ̌ *it *ik *eɿ̌ *ek *ɛɿ̌ *əɿ̌ *aɿ̌ *at *uɿ̌ *uk *oɿ̌

*ot *ok *ɔɿ̌ *ɔt

Among the 57 PK rhymes listed above, the rhymes having the *a vowel were reconstructed with more confidence due to the neat patterns of correspondences.

Tones

Three tones, i.e. *A, *B and *D were reconstructed. The correspondences of tones in modern Karenic language varieties are illustrated in Table 3 and Figure 4. The *A and *B tones occur in non-checked syllables while the *D tone occurs in checked syllables. Resulting from the eight patterns of tonal correspondences presented in Luce (1959) and the concept of tone-box adopted, the *B ✓ tone (equivalent to the *C tone in Tai languages) was reconstructed to solve the tonal problems in a small number of words, e.g. *'pus, paddy, blow, breathe, many, child,*' etc., which have tone A in Sgaw but tone B in Pwo. In my opinion, perhaps it is too early to reconstruct the *B ✓ tone (or *C). When carefully look at the tone boxes in Shintani (2003) and Manson (2009), it is amazing to see that the so-called B ✓ tone in modern Karenic languages has completely merged with the A, B or D tone, unlike the *C tone in Tai languages and dialects spoken both inside and outside Thailand and in the Mien or Yao languages spoken in Thailand and Guangxi province, southern China, which I had the opportunity to work on during our several field trips in China, Laos and Vietnam (see L-Thongkum, 1991,1993,1997; Kullavanijaya & L-Thongkum, 2000). I feel that the development of the PK *B ✓, based on the tone-box concept proposed by Shintani and Manson, does not sound convincing, even though their views can be traced back to Haudricourt's proposal in 1975. The uncommon type of tone splitting of the B tone in some Sgaw varieties, e.g. the PK tone *B > B13-2 (CVɿ̌²¹) and *D > D12 (CVɿ̌⁴⁵) – 3 (CV⁵³) may be a reason why the *B ✓ tone was added. In my opinion, the reconstructed three tones, namely, *A, *B and *D are sufficient to handle an unusual development of tones B and D in some Sgaw varieties. Even though the addition of the *B ✓ tone can help make the reconstruction of the proto tones in some PK roots possible, I am still reluctant to accept this idea. There might be a better solution, if we could reach more fresh data, especially on the Karenic languages spoken in Myanmar.

N. Pa-O

	*A	*B	*D
1	31''	55	21'
2	31''	55	21'
3	33	53	45'

S. Pa-O

	*A	*B	*D
1	31''	33	21'
2	31''	33	21'
3	53	55	45'

Kayan

	*A	*B	*D
1	53	11	45'
2	53		45'
3	33		21'

E. Kayah

	*A	*B	*D
1	33	11	55
2	33		55
3	11	11	33

W. Bwe

	*A	*B	*D
1	55	33	33
2		33	33
3	33	33	11

Kayaw

	*A	*B	*D
1	55	11	33
2	55	11	33
3	33	11	33

N. Sgaw

	*A	*B	*D
1	33	31''	45'
2	33	31''	45'
3	33	11''	21'

S. Sgaw

	*A	*B	*D
1	55	45'	21'
2	55	45'	21'
3	33	31	11''

N. Pwo

	*A	*B	*D
1	35	33	45'
2	55		45'
3	55	11''	21'

S. Pwo

	*A	*B	*D
1	53	55	21'
2	31''		21'
3	31''	33	45'

Figure 4 Proto-Karen tones and their tonal developments in modern Karenic languages

6. PK Lexicon

Based on the correspondences of the onsets, rhymes and tones as shown in Figures 1-3 and the reconstructed PK phonology in the fifth section, a PK lexicon consisting of 341 roots was reconstructed as presented in this section. The PK roots are arranged according to the alphabetical order of the English glosses (A-Z) with 'all' as the first item and 'young' as the last item. For comparison, the reconstructed forms by Burling (1969) are also given since his reconstruction, although different, equates to mine in some respects, such as the tones: 1=B3; 2=B12; 3=A2/A3; 4=A1/A12; 1̂=D3; 2̂=D12 and so forth.

Abbreviations and symbols

PK = Proto-Karen	KW = Kayaw
NK = Northern Karen	N.SG. = Northern Sgaw
CK = Central Karen	S.SG = Southern Sgaw
SK = Southern Karen	N.PW = Northern Pwo
N.PO = Northern Pa-O	S.PW = Southern Pwo
S.PO = Southern Pa-O	RB = Robins Burling
KN = Kayan	*- CV= Part of a compound (only Bwe)
KH = Kayah	*- CV= Preceded by an affix /sesquisyllable
BW = Western Bwe	*[CV/ *CV] Part of a compound (except Bwe)

Lexicon

- 'all, all gone, whole' PK: *lo[CV]
 NK: lo[CV]^{4 5} (N.PO), lu[CV]^{4 5} (S.PO)
 CK: lo[CV]³³ (KN); lo[CV]³³ (KH); - (BW); lo[CV]³³ (KW)
 SK: (khɛ³³)lo[CV]²¹ (N.SG), lo[CV]¹¹ (S. SG); lau[CV]²¹ (N.PW), lo[CV]^{4 5} (S.PW)
- 'ant' PK: *daŋ[CV]
 NK: tɔŋ^{5 3} (N.PO), tɔŋ^{5 5} (S.PO)
 CK: tɔ¹¹ (KN); tɔ¹¹ (KH); do³³ (aw); tɔ¹¹ (KW)
 SK: tɔ¹¹ (N.SG), tɔ³¹ (S. SG); thã¹¹ (N.PW), thã¹¹ (S.PW)
- 'arrive' PK: *tɔ[CV] *thy³ (RB)
 NK: tɔ^{5 5} (N.PO), tɔ³³ (S.PO)
 CK: tɔ¹¹ (KN); tɔ¹¹ (KH); do³³ (BW); tɔ¹¹ (KW)
 SK: tɔ³³ (N.SG), tɔ^{4 5} (S.SG); thɔ³³ (N.PW), tho³¹ (S.PW)

Note N.Sgaw and S.Pwo have tone A instead of tone B. The two varieties of Pwo have irregular vowels.

4. 'arrow, dart' PK: *blaꠞꠞ *phla¹ (RB)

NK: pla^{5 5} (N.PO), pla³³ (S.PO)

CK: pla¹¹ (KN); pla¹¹ (KH); blɛ³³ (BW); ple¹¹ (KW)

SK: pla¹¹ (N.SG), pla³¹ (S.SG); pla¹¹ (N.PW), phla³³ (S.PW)

Note Irregular vowels in Kayah (KH) and Kayaw (KW), i.e. they should be e and a in KH and KW, respectively. S.Karen (SK) tones suggest *bl-, while those in N.Karen (NK) indicate *pl-.

5. 'ascend, go up, rise' PK: *thaŋꠞꠞ *thaŋ² (RB)

NK: thaŋ^{5 5} (N.PO), thaŋ³³ (S.PO)

CK: thaŋ¹¹ (KN); thɛ¹¹ (KH); tha³³ (BW); hɔ¹¹ (KW)

SK: thɔ³¹ (N.SG), thɔ^{4 5} (S.SG); thɛ³³ (N. PW), thɔ^{5 5} (S.PW)

6. 'ashes, fireplace' PK: *phaꠞꠞ

NK: pha^{5 5} (N.PO), pha³³ (S.PO)

CK: pha¹¹ (KN), phe¹¹ (KH); - (BW), pha¹¹ (KW)

SK: pha³¹ (N.SG), pha^{4 5} (S.SG); - (N.PW), - (S.PW)

Note In some N.Sgaw varieties pha¹¹ ~pha¹¹ means 'fireplace'.

7. 'astringent' PK: *kjaꠞꠞ *khiꠞꠞ¹ (RB)

NK: khjaꠞ²¹ (N.PO), - (S.PO)

CK: khiꠞ^{4 5} (KN); khja^{5 5} (KH); - (BW); - (KW)

SK: kiꠞ^{4 5} (N.SG), kaiꠞ²¹ (S.SG); kaiꠞ^{4 5} (N.PW), kaiꠞ³¹ (S.PW)

8. 'bamboo' PK: *hwaꠞꠞ *wa² (RB)

NK: wa^{5 5} (N.PO), wa³³ (S.PO)

CK: hwa¹¹~hau¹¹ (KN); we¹¹ (KH); hɔ³³ (BW); - (KW)

SK: wa³¹ (N.SG), wa^{4 5} (S.SG); wa³³ (N.PW), wa^{5 5} (S.PW)

9. 'bamboo rat' PK: *(jowꠞꠞ)khanꠞꠞ

NK: (ju^{5 5})khan³¹ (N.PO), - (S.PO)

CK: Ju¹¹khaŋ^{5 3} (KN); - (KH); khɛ^{5 5} (BW); (ju¹¹)khɔ^{5 5} (KW)

SK: khɔ³³ (N.SG), khɔ^{5 5} (S.SG); khɛ^{3 5} (N.PW), khɔ^{5 3} (S.PW)

10. 'bamboo strip' PK: *sraꠞꠞ

NK: sa³¹ (N.PO), sa³¹ (S.PO)

CK: $\theta a^{5\ 3}$ (KN); se^{33} (KH); $\theta r\varepsilon^{5\ 5} \sim \theta \tilde{a}^{5\ 5} r\varepsilon^{5\ 5}$ (BW); $sa^{5\ 5}$ (KW)

SK: $\zeta \gamma a^{33}$ (N.SG), $\zeta \gamma a^{5\ 5}$ (S.SG); $sja^{3\ 5}$ (N.PW), $\theta a^{5\ 3}$ (S.PW)

11. 'bamboo shoot' PK: * baŋ

NK: $baŋ^{5\ 5}$ (N.PO), $baŋ^{33}$ (S.PO)

CK: $baŋ^{11}$ (KN); $b\varepsilon^{11}$ (KH); $\bar{b}a^{33}$ (BW), $b\partial^{11}$ (KW)

SK: $b\partial^{31} \sim$ (N.SG), $b\partial^{4\ 5}$ (S.SG); $b\tilde{\varepsilon}^{33}$ (N.PW), $b\tilde{\partial}^{5\ 5}$ (S.PW)

12. 'bark (n.), leather, skin' PK: * phe * phi^2 (RB)

NK: phe^{21} (N.PO), - (S.PO)

CK: - (KN); $pha^{5\ 5}$ (KH); phe^{33} (BW); phe^{33} (KW)

SK: $phi^{4\ 5}$ (N.SG), $phai^{21}$ (S.SG); $phai^{4\ 5}$ (N.PW), $phai^{31}$ (S.PW)

13. 'bat' PK: * pla

NK: pla^{31} (N.PO), pla^{31} (S.PO)

CK: bla^{11} (KN); ple^{11} (KH); $pl\varepsilon^{11}$ (BW); pla^{11} (KW)

SK: bla^{33} (N.SG), $bla^{5\ 5}$ (S.SG); $phla^{33}$ (N.PW), $phla^{5\ 5}$ (S.PW)

Note Pa-O and Sgaw have tone A while C.Karen (CK) languages and Pwo have tone B. *pl- has become bl- in Sgaw.

14. 'bathe' PK: * hlow

NK: $lu^{5\ 5}$ (N.PO), lu^{33} (S.PO)

CK: - (KN); lo^{11} (KH); - (BW); lu^{11} (KW)

SK: $l\dot{i}^{11} \sim lu^{11}$ (N.SG), $l\dot{i}^{4\ 5}$ (S.SG); lou^{33} (N.PW), $(\tilde{\partial}^{5\ 5})lou^{5\ 5}$ (S.PW)

Note N.Sgaw has the B3 tone while S. Sgaw has the B1 tone as in Pwo and Pa-O.

15. 'bear (animal)' PK: * tham * tham^4 (RB)

NK: $tham^{31}$ (N.PO), $tham^{31}$ (S.PO)

CK: $thaŋ^{5\ 3}$ (KN); the^{33} (KH); $th\varepsilon^{5\ 5}$ (BW); $th\partial^{5\ 5}$ (KW)

SK: - (N.SG), - (S.SG); $(ph\dot{i}^{11}) th\tilde{\partial}^{3\ 5}$ (N.PW), $(phou^{5\ 3}) th\tilde{\partial}^{5\ 3}$ (S.PW)

16. 'bedbug' PK: * gram

NK: $sam^{33} \sim \zeta am^{33}$ (N.PO), $sam^{5\ 3} \sim saŋ^{5\ 3}$ (S.PO)

CK: $caŋ^{33}$ (KN); $kh\varepsilon^{11}$ (KH); - (BW); $c\partial^{33}$ (KW)

SK: $x\partial^{33}$ (N.SG), $x\partial^{33}$ (S.SG); - (N.PW), - (S.PW)

17. 'bee (*Apis cerana*)' PK: *kwat□

NK: wat²¹ (N.PO), wat²¹ (S.PO)

CK: hwε□^{4 5} (KN); wε^{5 5} (KH); wε³³~θ^{5 5}kwε^{5 5} (BW); hwε³³~khwε³³ (KW)

SK: kwε□^{4 5} (N.SG), kwε^{5 5} (S.SG); kwε^{5 5} (N.Pwo), kwε³¹ (S.PWO)

Note This etymon has irregular tone change, i.e. *D has become tone A in S.Sgaw and Pwo. Perhaps, the final *-t had been dropped much earlier, then the *D tone became the *A tone.

18. 'bee (*Apis dorsata*)' PK: *k-hne□

NK: ne³¹ (N.PO), ne³¹ (S.PO)

CK: nai^{5 3} (KN); ni³³ (KH); (gθ) nI³³ (BW); ni^{5 5} (KW)

SK: kθ¹¹nε³³ (N.SG), nε^{5 5} (S.SG); ni^{3 5} (N.PW), ni^{5 3} (S.PW)

19. 'believe' PK: *jim□

NK: jum^{5 3} (N.PO), jəm^{5 5}~jən^{5 5} (S.PO)

CK: ʃi¹¹ (KN); ji¹¹ (KH); - (BW); ji¹¹ (KW)

SK: - (N.SG), - (S.SG); - (N.PW); - (S.PW)

20. 'below, under, underneath' PK: *~la□□

NK: (ka³³) la□^{4 5} (N.PO), la□^{4 5} (S.PO)

CK: la□^{4 5} (KN); le³³ (KH); gɔ¹¹lε¹¹, -lε¹¹ (BW); lɔ³³ (KW)

SK: phɔ³³la□²¹ (N.SG), ka¹¹la¹¹ (S.SG); cha□^{4 5}la□²¹ (N.PW), phɔ^{5 3}la□^{4 5} (S.PW)

21. 'big, large' PK: *do□ *do² (RB)

NK: - (N.PO), - (S.PO)

CK: dau¹¹ (KN); du¹¹ (KH); do³³ (BW); du¹¹ (KW)

SK: do³¹~ (N.SG), du^{4 5}, (S.SG); du³³ (N.PW), du^{5 5} (S.PW)

22. 'bird' PK: *tho□ *tho² (RB)

NK: - (N.PO), - (S.PO)

CK: thau¹¹ (KN); thu¹¹ (KH); tho³³ (BW); thu¹¹ (KW)

SK: tho³¹~ (N.SG), thu^{4 5}, (S.SG); thu³³ (N.PW), thu^{5 5} (S.PW)

Note The words meaning 'bird' in N.PO and S.PO are wa^{5 3}(B3) and wa^{5 5}(B3), respectively.

23. 'bite (v.)' PK: *εN□ *εη² (RB)

NK: εη^{5 5} (N.PO), εm³³~εn³³ (S.PO)

CK: i¹¹ (KN); a¹¹ (KH); - (BW); - (KW)

- SK: $\text{e}^{31}\sim$ (N.SG), $\text{i}^{4\ 5'}$ (S.SG); $\tilde{\text{ai}}^{33}$ (N.PW), $\tilde{\text{ai}}^{5\ 5}$ (S.PW)
24. 'bitter' PK: *kha $\square$ *kha² (RB)
- NK: kha^{5\ 5} (N.PO), kha³³ (S.PO)
- CK: kha¹¹ (KN); khe¹¹ (KH); khe³³ (BW); kha¹¹ (KW)
- SK: kha^{31}\sim} (N.SG), kha^{4\ 5'}} (S.SG); kha³³ (N.PW), kha^{5\ 5} (S.PW)
25. 'blood' PK: *swi $\square$ *swi² (RB)
- NK: swi^{5\ 5} (N.PO), swi³³ (S.PO)
- CK: θ wi¹¹ (KN); swi¹¹ (KH); θ wi³³, θ u³³ (BW); su¹¹ (KW)
- SK: (ta¹¹)swi^{31}\sim} (N.SG), swi^{4\ 5'}} (S.SG); swi³³ (N.PW), θ wi^{5\ 5} (S.PW)
26. 'blow (a fire)' PK: * $\square$ u $\square$ * $\square$ u⁴ (RB)
- NK: $\square$ u³¹ (N.PO), $\square$ u³¹ (S.PO)
- CK: $\square$ u^{5\ 3}} (KN); $\square$ u³³ (KH); u³³ (BW); $\square$ u^{5\ 5}} (KW)
- SK: $\square$ u³³ (N.SG), $\square$ u^{5\ 5}} (S.SG); $\square$ u^{5\ 5}} (N.PW), $\square$ u³¹ (S.PW)
27. 'boar (wild~)' PK: *tho $\square$ mi $\square$
- NK: tho $\square$ ²¹mi³³ (N.PO), tho $\square$ ²¹mi^{5\ 3}} (S.PO)
- CK: tho $\square$ ^{4\ 5}}mi³³ (KN); th ε ^{5\ 5}}mi¹¹ (KH); - (BW); tho³³mi³³ (KW)
- SK: tho $\square$ ^{4\ 5}}mi³³ (N.SG), tho $\square$ ²¹mi³³ (S.SG); tho $\square$ ^{4\ 5}}mei^{5\ 5}} (N.PW), thu $\square$ ²¹mei³¹ (S.PW)
28. 'boat' PK: *klej $\square$ *khli⁴ (RB)
- NK: phri³¹ (N.PO), pli³¹ (S.PO)
- CK: - (KN); kl $\mathfrak{e}$ ³³ (KH); khli³³ (BW); kli^{5\ 5}} (KW)
- SK: khli³³ (N.SG), khli^{5\ 5}} (S.SG); khlei^{5\ 5}} (N.PWO), khlei^{5\ 3}} (S.PWO)
- Note *kl- has become phr- and pl- in N.Pa-O and S.Pa-O. respectively.
29. 'body dirt' PK: *khrej $\square$ *xi² (RB)
- NK: khri^{5\ 5}} (N.PO), khri³³ (S.PO)
- CK: - (KN); khr $\mathfrak{e}$ ¹¹ (KH); - (BW); ci¹¹ (KW)
- SK: xi^{31}\sim} (N.SG), xi^{4\ 5'}} (S.SG); kei³³ (N.PW), kei^{5\ 5}} (S.PW)
30. 'boil (v.)' PK: *phr $\square$
- NK: - (N.PO), - (S.PO)
- CK: ph $\square$ ^{5\ 3}} (KN); pho³³ (KH); a^{5\ 5}}phr $\square$ ³³ (BW); pho^{5\ 5}} (KW)

SK: khɿɯ³³ (N.SG), phɯ^{5 5} (S.SG); phɯ^{3 5} (N.PW), pho^{5 3} (S.PW)

31. 'bone' PK: *khrwit¹ *xwi⁴ (RB)

NK: chut²¹ (N.PO), chə¹ ut²¹ (S.PO)

CK: chwi^{4 5} (KN); khrwi^{5 5} (KH); -khw³³ (BW); chu³³ (KW)

SK: xi³³ (N.SG), (ta¹¹)xi^{5 5} (S.SG); (a²¹) xei³³ (N.PW), xwi^{5 5} (S.PW)

Note This etymon has irregular tone changes in SK, i.e. tone A in Sgaw and tone B in Pwo, while it is still tone D in NK and CK. The final *-t might have been dropped very early; consequently, the tones in Sgaw and Pwo have developed differently.

32. 'braid' PK: *srej¹ *si² (RB)

NK: si^{5 5} (N.PO), si³³ (S.PO)

CK: - (KN); sɿ¹¹ (KH); θre³³ (BW); si¹¹ (KW)

SK: si³¹ (N.SG), - (S.SG); sei³³ (N.PWO), θei^{5 5} (S.PWO)

33. 'brain' PK: *nok¹ *kho²nu¹ (RB)

NK: (u^{5 5})nok^{4 5} (N.PO), (a²¹)nu^{4 5} (S.PO)

CK: nɯ²¹ (KN), (ku¹¹)nɯ³³ (KH); -no¹¹ (BW); (kə³³)nə³³ (KW)

SK: (kho³¹)nu²¹ (N.SG), (khu^{4 5})nɯ¹¹ (S.SG); (khu³³)nau²¹ (N.PW), (khu^{5 5})nɯ^{4 5} (S.PW)

34. 'bran, chaff' PK: *phe¹ *phe⁴ (RB)

NK: phe³¹ (N.PO), phe³¹ (S.PO)

CK: phai^{5 3} (KN); phi³³ (KH); - (BW); phi^{5 5} (KW)

SK: phe³³ (N.SG), phe^{5 5} (S.SG); phi^{3 5} (N.PW), phi^{5 3} (S.PW)

35. 'breasts' PK: *nu¹ *ny¹ (RB)

NK: - (N.PO), - (S.PO)

CK: nu¹¹ (KN); (i¹¹)nu¹¹ (KH); nu³³ (BW); nu¹¹ (KW)

SK: nu¹¹ (N.SG), nu³¹ (S.SG); nə¹ u¹¹ (N.PW), nu³³ (S.PW)

36. 'broken, split, separate' PK: *pa¹ *pha² (RB)

NK: pa²¹ (N.PO), pa²¹ (S.PO)

CK: pha^{4 5} (KN); phe^{5 5} (KH); pha³³ (BW); phɯ³³ (KW)

SK: pha^{4 5} (N.SG), pha²¹ (S.SG); pha^{4 5} (N.PW), pha²¹ (S.PWO)

37. 'buffalo' PK: *p/b-na¹ *pə²na¹ (RB)

NK: pa²¹na^{5 3} (N.PO), pa²¹na^{5 5} (S.PO)

CK: bu^{4 5}na¹¹ (KN); pe^{5 5}ne¹¹ (KH); bə³³ne³³ (BW); pa³³na¹¹ (KW)

SK: pə³³na¹¹ (N.SG), pə¹¹na³¹ (S.SG); pa²¹nə¹¹ (N.PW); pə²¹na³³ (S.PWO)

38. 'buy' PK: *bre¹¹ *phye³ (RB)

NK: phre³³ (N.PO), phre^{5 3} (S.PO)

CK: prai³³ (KN); phri¹¹ (KH), a^{5 5}bwI³³ (BW); pri³³ (KW)

SK: pə¹¹ye³³~pə³³ye³³ (N.SG), pə³³ye³³ (S.SG); - (N.PW); - (S.PW)

39. 'caress' PK: *phruN¹¹

NK: phri^{5 5} (N.PO), pwəm³³ (S.PO)

CK: phwaŋ^{5 3} (KN); phi³³ (KH); - (BW); phu^{5 5} (KW)

SK: phyu³³ (N.SG), phu^{5 5} (S.SG); phã^{3 5} (N.PW), phõ^{5 5} (S.PW)

Note It is noticeable that NK has tone B while CK and SK have tone A.

40. 'carry (on the back)' PK: *pi¹¹ *phy² (RB)

NK: - (N.PO), - (S.PO)

CK: - (KN); - (KH); phu³³ (BW); phi¹¹ (KW)

SK: - (N.SG), phi^{4 5} (S.SG); pə³³i³³ (N.PW), pi^{5 5} (S.PW)

41. 'caterpillar' PK: *si¹¹

NK: se²¹ (N.PO), si²¹ (S.PO)

CK: θi^{4 5} (KN); si^{5 5} (KH); θa³³ (BW); si³³ (KW)

SK: si^{4 5} (N.SG), sai²¹ (S.SG); si^{3 5} (N.PW), θai²¹, θei²¹ (S.PW)

Note N.Pwo has tone A which is irregular.

42. 'catfish' PK: *~ku¹¹

NK: khu³¹ (N.PO), khu³¹ (S.PO)

CK: ku^{5 3} (KN); ku¹¹ (KH); - (BW), ku¹¹ (KW)

SK: - (N.SG), ku^{4 5} (S.SG); - (N.PW), ku^{5 5} (S.PW)

Note NK and Kayan (CK) have tone A while the rest have tone B.

43. 'carapace, shell' PK: *ki¹¹

NK: - (N.PO); - (S.PO)

CK: ko^{5 3} (KN); ɔa¹¹kɔ³³ (KH); ko^{5 5} (BW); ɔa¹¹ki^{5 5} (KW)

SK: ki^{4 5} (N.SG), ki^{5 5} (S.SG); ɔa¹¹ko^{5 5} (N.PW), ku³¹ (S.PW)

Note Kayan has an irregular vowel, i.e. *o*, instead of *i*. N.Sgaw also has an irregular tone, i.e. tone D.

44. 'centipede' PK: *t/daꣳꣳꣳbaŋꣳ *daꣳ²baŋ³ (RB)
 NK: taꣳ^{4 5}baŋ³¹ (N.PO), taꣳ^{4 5}baŋ³¹ (S.PO)
 CK: ta¹¹baŋ^{5 3} (KN); - (KH); - (BW); - (KW)
 SK: daꣳ^{4 5}bɔ³³ (N.SG), da¹¹bɔ^{5 5} (S.SG); - (N.PW), daꣳ^{4 5}bɔ̃³¹ (S.PW)
45. 'chameleon' PK: *kwiꣳ
 NK: kwi^{5 5} (N.PO), kwi³³ (S.PO)
 CK: khwi¹¹ (KN); khwi¹¹ (KH); khwi³³ (BW); khi¹¹ (KW)
 SK: khwi³¹̃ (N.SG), khwi^{4 5} (S.SG); khwi³³ (N.PW), khwi^{5 5} (S.PW)
46. 'charcoal' PK: *swaꣳ *swa² (RB)
 NK: swa^{5 5} (N.PO), swa³³ (S.PO)
 CK: θau¹¹ (KN); su¹¹ (KH); - (BW); su¹¹ (KW)
 SK: swa³¹̃ (N.SG), swa^{4 5} (S.SG); swa³³ (N.PW), θwa^{5 5} (S.PW)
47. 'chicken' PK: *chjaNꣳ *chjaN⁴ (RB)
 NK: ɕja³¹ (N.PO), ɕja³¹ (S.PO)
 CK: ɕi^{5 3} (KN); cha³³ (KH); ɕi^{5 5} (BW); ɕi^{5 5} (KW)
 SK: chɔ³³ (N.SG), chɔ^{5 5} (S.SG); chɛ^{3 5} (N.PW), chɔ̃^{5 3} (S.PW)
Note [Proto-Monic: *tɣaaŋ], Proto-Mon: *cəiŋ] (Diffloth, 1984)]
48. 'chili pepper, pungent' PK: *hatꣳ
 NK: hat²¹ (N.PO), hap²¹ (S.PO)
 CK: hɛꣳ^{4 5} (KN); hɛ^{5 5} (KH); hɛ³³ (BW); hɛ³³ (KW)
 SK: hɛ³³ (N.SG), hɛ^{5 5} (S.SG); - (N.PW), ʔɛ^{5 5} (S.PW)
Note NK and CK have tone D while SK languages have irregular tone changes,
 i.e. *D>A in Sgaw but >B in Pwo.
49. 'chin, jaw' PK: *kaꣳ *kha² (RB)
 NK: ka^{5 5} (N.PO), ka³³ (S.PO)
 CK: ka¹¹ (KN); khe¹¹ (KH); khe³³ (BW); kha¹¹ (KW)
 SK: kha³¹̃ (N.SG), kha^{4 5} (S.SG); kha³³ (N.PW), kha^{5 5} (S.PW)
Note [Thai: khaaŋ³³]

50. *'cicada'* PK: ***ŋjaŋ**
 NK: **ŋja**³³ (N.PO), - (S.PO)
 CK: **ji**³³ (KN); **ja**¹¹ (KH); - (BW); - (KW)
 SK: **je**³³, **ze**³³ (N.SG), - (S.SG); **jai**^{5 5} (N.PW), **jai**³¹ (S.PW)
51. *'civet cat'* PK: ***thu**
 NK: - (N.PO), - (S.PO)
 CK: **thu**¹¹ (KN); - (KH); **thu**³³ (BW); **thu**¹¹ (KW)
 SK: - (N.SG), - (S.SG); - (N.PW), **thu**²¹**thō**³¹, **thu**^{5 5}**thō**¹¹ (S.PW)
52. *'classifier (hand of bananas)'* PK: ***sej**
 NK: **si**^{5 5} (N.PO), **si**³³ (S.PO)
 CK: **θi**¹¹ (KN); **sə**¹¹ (KH); - (BW); **si**¹¹ (KW)
 SK: **si**³¹ (N.SG), **si**^{4 5} (S.SG); **sei**³³ (N.PW), - (S.PW)
53. *'classifier for flat objects'* PK: ***ba**
 NK: **ba**^{5 5} (N.PO), **ba**³³ (S.PO)
 CK: **ba**¹¹ (KN); **be**¹¹ (KH); - (BW); **ba**¹¹ (KW)
 SK: **be**³¹ (N.SG), **bi**^{4 5} (S.SG); **bai**³³ (N.PW), **bai**^{5 5} (S.PW)
Note The correspondence pattern of rhymes in SK suggests ***baŋ** [Thai: **baj**³³]
54. *'classifier for people, persons'* PK: ***bra** ***ʔa**³ (RB), ***phwa**³ *'person'* (RB)
 NK: **phra**³³ (N.PO), **pra**^{5 3} (S.PO)
 CK: **pra**³³ (KN); **phre**¹¹ (KH); - (BW); **ra**³³ (KW)
 SK: **pʔa**³³ (N.SG), **ʔa**³³ (S.SG); **ʔa**^{5 5} (N.PW), **ʔa**³¹ (S.PW)
55. *'classifier for a piece of meat, cake, etc.'* PK: ***plɔŋ** ***lɔŋ**² (RB)
 NK: **lɔŋ**^{5 5} (N.PO), - (S.PO)
 CK: - (KN); **plɔ**¹¹ (KH); - (BW); **plɔ**¹¹ (KW)
 SK: **lɔ**³¹ (N.SG), **lu**^{4 5} (S.SG); **lɔ**³³ (N.PW), **lɔ**^{5 5} (S.PW)
56. *'classifier for plants, e.g. bamboo, rice, banana etc. (clumps, clusters)'*
 PK: ***kho**
 NK: - (N.PO), - (S.PO)
 CK: **khau**¹¹ (KN); - (KH); - (BW); - (KW)

- SK: kho³¹~ (N.SG), khu^{4 5} (S.SG); khu³³ (N.PW), khu^{5 5} (S.PW)
57. 'classifier for round objects' PK: *phloŋ¹ *phloŋ² (RB)
- NK: phloŋ^{5 5} (N.PO), phloŋ³³ (S.PO)
- CK: phloŋ¹¹ (KN); phloŋ¹¹ (KH); - (BW); phloŋ¹¹ (KW)
- SK: phloŋ³¹~ (N.SG), phloŋ^{4 5} (S.SG); phloŋ³³ (N.PW), phloŋ^{5 5} (S.PW)
58. 'classifier for long objects' PK: *boŋ¹ *boŋ³ (RB)
- NK: - (N.PO), - (S.PO)
- CK: boŋ^{5 3} (KN); boŋ³³ (KH); - (BW); boŋ^{5 5} (KW)
- SK: boŋ³³ (N.SG), bu^{5 5} (S.SG); boŋ^{5 5} (N.PW), bō³¹ (S.PW)
59. 'classifier for four-legged animals' PK: *doŋ¹
- NK: - (N.PO), - (S.PO)
- CK: doŋ^{5 3} (KN); doŋ³³ (KH); - (BW); doŋ^{5 5} (KW)
- SK: doŋ³³ (N.SG), doŋ^{5 5} (S.SG); doŋ^{5 5} (N.PW), doŋ³¹ (S.PW)
60. 'close (one's eyes)' PK: *bi¹ *bin² (RB)
- NK: - (N.PO), - (S.PO)
- CK: bi¹¹ (KN); - (KH); - (BW); bi¹¹ (KW)
- CK: bi³¹~ (N.SG), bi^{4 5} (S.SG); bei³³ (N.PW), bei^{5 5} (S.PW)
61. 'comb (v.)' PK: *khwi¹ *khwi⁴ (RB)
- NK: khwi²¹ (N.PO), - (S.PO)
- CK: khi^{4 5} (KN); khwi^{5 5} (KH); khwi³³ (BW); khi³³ (KW)
- SK: khwi³³ (N.SG), khwi^{5 5} (S.SG); khwi^{5 5} (N.PW), khwi^{5 3} (S.PW)
- Note The SK tones suggest tone *A in PK.
62. 'cotton thread' PK: *hloŋ¹ *ly² (RB)
- NK: ly^{5 5} (N.PO), ly³³ (S.PO)
- CK: ly¹¹ (KN); ly¹¹ (KH); ly³³ (BW); ly¹¹ (KW)
- SK: ly³¹~ (N.SG), ly^{4 5} (S.SG); lou³³ (N.PW), lou^{5 5} (S.PW)
63. 'cough (v.)' PK: *kuk¹ *ku² (RB)
- NK: (ta²¹he^{5 5}) khuk²¹ (N.PO), - (S.PO)
- CK: khu^{4 5} (KN); kh^{5 5} (KH); θ^{5 5}ku³³ (BW); kh³³ (KW)

- SK: ku^{4 5} (N.SG), kɔu²¹ (S.SG); kau^{4 5} (N.PW), kɔu³¹ (S.PW)
64. 'crab' PK: *chwɛ̃ *chwɛ² (RB)
- NK: chwɛ^{5 5} (N.PO), chwɛ³³ (S.PO)
- CK: chwɛ¹¹ (KN); chwa¹¹ (KH), ɲwɛ³³ (BW); chɛ¹¹ (KW)
- SK: chwɛ³¹̃ (N.SG), chwɛ^{4 5,} (S.SG); chwɛ³³ (N.PW), chwɛ^{5 5} (S.PW)
65. 'cricket' PK: *s-kĩ
- NK: ki³¹ (N.PO), ki³¹ (S.PO)
- CK: ki^{5 3} (KN); ki³³ (KH); - (BW); (dɛ¹¹) ki^{5 5} (KW)
- SK: sɔ̃¹¹ki³³ (N.SG), tɔ̃¹¹ki^{5 5} (S.SG); - (N.PW), - (S.PW)
- Noteki^{5 3} in Kayan is a kind of grub.
66. 'crocodile' PK: s-hmã *ma² (RB)
- NK: - (N.PO), - (S.PO)
- CK: - (KN); - (KH); θɔ̃³³mɛ³³ (BW); ma¹¹ (KW)
- SK: sɔ̃³³ma³¹̃ (N.SG), ma^{4 5,} (S.SG); ma³³ (N.PW), ma^{5 5} (S.PW)
67. 'crop (of fowl), gizzard' PK: *k-hnĩ *nin⁴ (RB)
- NK: - (N.PO), - (S.PO)
- CK: - (KN), kɛ³³ni³³ (KH); kɔ̃^{5 5}ni^{5 5} (BW); ki¹¹ni^{5 5} (KW)
- SK: kɔ̃¹¹ni³³ (N.SG), ni^{5 5} (S.SG); nei^{3 5} (N.PW), nei^{5 3} (S.PW)
68. 'cross (v.)' PK: *khaÑ
- NK: - (N.PO), - (S.PO)
- CK: - (KN); khɛ³³ (KH); - (BW); khɔ̃^{5 5} (KW)
- SK: khɔ̃³³ (N.SG), khɔ̃^{5 5} (S.SG); khɛ^{3 5}̃ (N.PW), khɔ̃^{5 3} (S.PW)
69. 'crossbow' PK: *klej̃ *khli² (RB)
- NK: khli^{5 5} (N.PO), kli³³ (S.PO)
- CK: - (KN); klɔ̃¹¹ (KH); khli³³ (BW); khli¹¹ (KW)
- SK: khli³¹̃ (N.SG), khli^{4 5,} (S.SG); khlei³³ (N.PW), khlei^{5 5} (S.PW)
70. 'crow (v.)' PK: *õ
- NK: u^{5 5} (N.PO), u³³ (S.PO)
- CK: - (KN); u¹¹ (KH); ɔ̃³³ (BW); u¹¹ (KW)

SK: ɔ̃³¹ ɔ̃³³ (N.SG), ɔ̃^{5 5} (S.SG); ɔ̃³³ (N.PW), ɔ̃³¹ ɔ̃^{5 5} (S.PW)

Note Sgaw has tone A while Pwo, NK and CK have tone B.

71. 'crush (v.)' PK: *k-hnɔ̃N

NK: - (N.PO), - (S.PO)

CK: - (KN); no¹¹ (KH); kə³³nɔ̃³³ (BW); no¹¹ (KW)

SK: no³¹ ɔ̃³¹ (N.SG), nu^{4 5} (S.SG); nɔ̃³³ (N.PW), nɔ̃^{5 5} (S.PW)

72. 'cucumber' PK: *ɔ̃dej ɔ̃ *di³ (RB)

NK: - (N.PO), di³¹ (S.PO)

CK: - (KN); də³³ (KH); dɪ^{5 5} (BW); di^{5 5} (KW)

SK: di³³ (N.SG), di^{5 5} (S.SG); thei^{5 5} (N.PW); thei³¹ (S.PW)

Note Pwo has an irregular initial, i.e. it should be d- instead of th-.

73. 'dark, late evening' PK: *khi ɔ̃ ɔ̃ *khi ɔ̃² (RB)

NK: khe²¹ (N.PO), khi²¹ (S.PO)

CK: khi^{4 5} (KN); khi^{5 5} (KH); khi³³ (BW); khi³³ (KW)

SK: khi^{4 5} (N.SG), khai²¹ (S.SG); khai^{4 5} (N.PW), khai³¹ (S.PW)

74. 'daughter-in-law' PK: *ɔ̃dɛ ɔ̃ *dɛ² (RB)

NK: - (N.PO), - (S.PO)

CK: dɛ¹¹ (KN); dɛ¹¹ (KH); -dɛ³³ (BW); dɛ¹¹ (KW)

SK: dɛ³¹ ɔ̃³¹ (N.SG), dɛ^{4 5} (S.SG); dɛ³³ (N.PW), dɛ^{5 5} (S.PW)

75. 'day' PK: *nej ɔ̃ *ni³ (RB)

NK: ni³³ (N.PO), ni^{5 3} (S.PO)

CK: - (KN); nɔ̃¹¹ (KH); ni³³ (BW); ni³³ (KW)

SK: ni³³ (N.SG), ni³³ (S.SG); nei^{5 5} (N.PW), nei³¹ (S.PW)

76. 'dead, die' PK: *sej ɔ̃ *si⁴ (RB)

NK: si³¹ (N.PO), si³¹ (S.PO)

CK: θi^{5 3} (KN); sə³³ (KN); θi^{5 5} (BW); si^{5 5} (KW)

SK: si³³ (N.SG), si^{5 5} (S.SG); sei^{5 5} (N.PW), θei^{5 3} (S.PW)

77. 'deaf' PK: *ɔ̃əŋ ɔ̃

NK: ɔ̃əŋ³¹ (N.PO), ɔ̃əŋ³¹ (S.PO)

CK: ɲ̥^{5 3} (KN); ɲ̥³³ (KH); tɕ'o^{5 5} (BW); ɲ̥^{5 5} (KW)

SK: ɲ̥³³ (N.SG), ɲ̥^{5 5} (S.SG); ɲ̥̃^{5 5} (N.PW), ɲ̥̃³¹ (S.PW)

78. 'deep' PK: *jɔɲ̥ɲ̥ *ŋoɲ̥¹ (RB)

NK: jɔɲ̥^{4 5} (N.PO), jɔɲ̥^{4 5} (S.PO)

CK: ɲ̥auɲ̥²¹ (KN); jɛ³³ (KH); jɔ¹¹ (BW); - (KW)

SK: jɔɲ̥²¹ (N.SG), jo¹¹ (S.SG); joɲ̥²¹ (N.PW), juɲ̥^{4 5} (S.PW)

79. 'deer (barking~)' PK: *d-khejɲ̥

NK: khi³¹ (N.PO), khi³¹ (S.PO)

CK: khĩ^{5 3} (KN); khɛ³³ (KH); do¹¹khi^{5 5} (BW); khi^{5 5} (KW)

SK: - (N.SG), - (S.SG); - (N.PW), - (S.PW)

80. 'deer (sambha ~)' PK: *t-khroɲ̥ɲ̥

NK: khjoɲ̥²¹ (N.PO), khjoɲ̥²¹ (S.PO)

CK: khjɔɲ̥^{4 5} (KN); khɾɔ^{5 5} (KH); kho³³ (BW); khɛ³³ (KW)

SK: tɔ³³xɔ³³ (N.SG), thɔ³³xɔ^{5 5} (S.SG); tɔ³³xɔ^{5 5} (N.PW), chɛ³³xo³¹ (S.PW)

Note In SK, the rhyme *-oɲ̥ɲ̥ has become -ɔɲ̥ (SG, N.PW) and -oɲ̥ (S.PW).

81. 'delicious' PK: *wiɲ̥

NK: wi^{5 5} (N.PO), wi³³ (S.PO)

CK: wi¹¹ (KN); wi¹¹ (KH); - (BW); wi¹¹ (KW)

SK: kwi²¹ (N.SG), - (S.SG); wi³³ (N.PW), wi^{5 5} (S.PW)

82. 'descent, go down, set' PK: *laŋɲ̥ *laŋ³ (RB)

NK: laŋ³³ (N.PO), laŋ^{5 3} (S.PO)

CK: laŋ³³ (KN); lɛ¹¹ (KH); la³³ (BW); lɔ³³ (KW)

SK: lɔ³³ (N.SG), lɔ³³ (S.SG); lɛ̃^{5 5} (N.PW), lɔ̃³¹ (S.PW)

83. 'dig' PK: *khuɲ̥ *khun² (RB)

NK: khu^{5 5} (N.PO), khu³³ (S.PO)

CK: - (KN); khĩ¹¹ (KH); - (BW); khu¹¹ (KW)

SK: khu³¹ (N.SG), khu^{4 5} (S.SG); khɔ³³ (N.PW), khɛ^{5 5} (S.PW)

84. 'dip up' (water) PK: *d(ɲ̥)ɲ̥ *dɔm² (RB)

NK: - (N.PO), - (S.PO)

CK: du¹¹ (KN); - (KH); d^{5 5} (BW); du¹¹ (KW)

SK: də³¹ (N.SG), d^{4 5} (S.SG); dā³³ (N.PW), dā^{5 5} (S.PW)

85. 'Dipterocarpus tuberculatus Roxb' PK: *thəN

NK: tə³¹ (N.PO), - (S.PO)

CK: thə^{5 3} (KN); th³³ (KH); - (BW); thə^{5 5} (KW)

SK: thə³³ (N.SG), th^{5 5} (SSG); thā^{5 5} (N.PW), - (S.PW)

Note In N.Pwo, it is an irregular tone change, i.e. tone 55 (A23) instead of tone 35 (A1) as in Sgaw, NK and CK.

86. 'disappear, lost' PK: *ma *ma¹ (RB)

NK: ma^{5 3} (N.PO), ma^{5 5} (S.PO)

CK: ma¹¹ (KN); me¹¹ (KH); hē^{5 5} mē¹¹ (BW); ma¹¹ (KW)

SK: ma¹¹ (N.SG), ma³¹ (S.SG); mē¹¹ (N.PW), ma³³ (S.PW)

87. 'dog' PK: *thwi *thwi² (RB)

NK: thwi^{5 5} (N.PO), thwi³³ (S.PO)

CK: thwi¹¹ (KN), thwi¹¹ (KH); thwi³³ (BW); thi¹¹ (KW)

SK: thwi³¹ (N.SG), thwi^{4 5} (S.SG); thwi³³ (N.PW), thwi^{5 5} (S.PW)

88. 'do, make' PK: *ma *ma³ (RB)

NK: ma³³ (N.PO), ma^{5 3} (S.PO)

CK: ma³³ (KN); me¹¹ (KH); mē (BW); ma³³ (KW)

SK: ma³³ (N.SG), ma³³ (S.SG); mē^{5 5} (N.PW), ma³¹ (S.PW)

89. 'dove' PK: *(tho)lwi

NK: lwi^{5 3} (N.PO), lwi^{5 5} (S.PO)

CK: lwi¹¹ (KN); lwi¹¹ (KH); lwi³³ (BW); li¹¹ (KW)

SK: lwi¹¹ (N.SG), lwi³¹ (S.SG); lei¹¹ (N.PW), lwi³³ (S.PW)

90. 'dream (v.)' PK: *maŋ *maŋ¹ (RB)

NK: maŋ^{5 3} (N.PO), maŋ^{5 5} (S.PO)

CK: maŋ¹¹ (KN); mē¹¹ (KH); mi^{5 5} mā³³ (BW); m³¹ (KW)

SK: mi³³ mā³¹ (N.SG), m³¹ (S.SG); mē¹¹ (N.PW), m³³ (S.PW)

91. 'drink(v.)' PK: ɔ ɔ³ (RB)

NK: - (N.PO), - (S.PO)

- CK: ɔ̌au⁵³ (KN); ɔ̌o³³ (KH); ɔ̌⁵⁵ (BW); ɔ̌o⁵⁵ (KW)
- SK: ɔ̌ɔ³³ (N.SG), ɔ̌ɔ⁵⁵ (S.SG); ɔ̌ɔ⁵⁵ (N.PW), ɔ̌o³¹ (S.PW)
92. 'dry(v.)' PK: *hloɔ̌ *lo⁴ (RB)
- NK: lo³¹ (N.PO), lo³¹ (S.PO)
- CK: lau⁵³ (KN); - (KH); lo⁵⁵ (BW); lo⁵⁵ (KW)
- SK: lo³³ (N.SG), lo⁵⁵ (S.SG); lu³⁵ (N.PW), lu⁵³ (S.PW)
93. 'drunk' PK: *mwɪnɔ̌ *mun³ (RB)
- NK: mun³³ (N.PO), mu ɔ̌m⁵³ (S.PO)
- CK: mɪ³³ (KN); mɪ¹¹ (KH); mu³³ (BW); mɪ³³ (KW)
- SK: mu³³ (N.SG), mu³³ (S.SG); m̃a⁵⁵ (N.PW), m̃ɔ̌i³¹ (S.PW)
94. 'dung, excrement' PK: *ɔ̌eɔ̌ *ɔ̌e² (RB)
- NK: ɔ̌e⁵⁵ (N.PO), ɔ̌e³³ (S.PO)
- CK: ɔ̌ai¹¹ (KN); ɔ̌i¹¹ (KH); ɪ³³ (BW); ɔ̌i¹¹ (KW)
- SK: ɔ̌e³¹ (N.SG), ɔ̌e⁴⁵ (S.SG); ɔ̌i³³ (N.PW), ɔ̌i⁵⁵ (S.PW)
95. 'ear' PK: *naɔ̌ *na¹ (RB)
- NK: na⁵³ (N.PO), na⁵⁵ (S.PO)
- CK: na¹¹ (KN); ne¹¹ (KH); nɛ³³ (BW), na¹¹ (KW)
- SK: na¹¹ (N.SG), na³¹ (S.SG); nɛ¹¹ (N.PW), na³³ (S.PW)
96. 'earthworm' PK: *jɛɔ̌ɔ̌
- NK: jɛɔ̌⁴⁵ (N.PO), jɛɔ̌⁴⁵ (S.PO)
- CK: jɛɔ̌²¹ (KN); ja³³ (KH); - (BW); je³³ (KW)
- SK: - (N.SG), - (S.SG); - (N.PW), - (S.PW)
97. 'easy' PK: *hɲawɔ̌ *ɲɔ̌⁴ (RB)
- NK: - (N.PO), - (S.PO)
- CK: jau⁵³ (KN); ju³³ (KH); jo⁵⁵ (BW); ju⁵⁵ (KW)
- SK: ɲɔ̌³³ (N.SG), ɲɔ̌⁵⁵ (S.SG); jɔ̌³⁵ (N.PW), jo⁵³ (S.PW)
98. 'eat' PK: *ɔ̌amɔ̌ *ɔ̌am² (RB)
- NK: ɔ̌am⁵⁵ (N.PO), ɔ̌am³³ (S.PO)

CK: ɲaŋ¹¹ (KN); ɲɛ¹¹ (KH); a³³ (BW); ɲɔ¹¹ (KW)

SK: ɲɔ³¹ (N.SG), ɲɔ^{4 5} (S.SG); ɲɛ³³ (N.PW), ɲɔ^{5 5} (S.PW)

99. 'egg' PK: *ɲdej ɲ *di² (RB)

NK: di^{5 5} (N.PO), di³³ (S.PO)

CK: tɪ¹¹ (KN); dɔ¹¹ (KH); -dɪ³³ (BW); di¹¹ (KW)

SK: di³¹ (N.SG), di^{4 5} (S.SG); dei³³ (N.PW), dei^{5 5} (S.PW)

Note Kayan has an irregular onset, i.e. t- instead of d-.

100. 'eggplant' PK: *s-gaŋ ɲ

NK: - (N.PO), - (S.PO)

CK: kaŋ³³ (KN); kɛ¹¹ (KH); - (BW); -kɔ³³ (KW)

SK: sɔ¹¹kɔ³³ (N.SG), kɔ³³ (S.SG); - (N.PW), - (S.PW)

101. 'eight' PK: *khrɔt ɲ/*grɔt ɲ *xo¹ (RB)

NK: sɔ¹¹t²¹ (N.PO), sɔ¹¹t²¹ (S.PO)

CK: chau^{4 5} (KN); - (KH); - (BW); -xɔ¹¹ (KW)

SK: xɔ^{4 5} (N.SG), xo²¹ (S.SG); xo²¹ (N.PW), xu^{4 5} (S.PW)

Note The D3 tone in NK and Pwo suggests *gr- while the D12 tone in Kayan and Sgaw indicates *khr-.

102. 'elephant' PK: *k-chaŋ ɲ *chaŋ⁴ (RB)

NK: chaŋ³¹ (N.PO), chaŋ³¹ (S.PO)

CK: chaŋ^{5 3} (KN); chɛ³³ (KH); (gɔ) ɲa^{5 5} (BW); rɔ¹¹chɔ^{5 5} (KW)

SK: kɔ¹¹chɔ³³ (N.SG), kɔ¹¹chɔ^{5 5} (S.SG); ka¹¹chɛ^{3 5} (N.PW), kɔ¹¹chɔ^{5 3} (S.PW)

103. 'embrace, hug' PK: *phɔ ɲ ɲ *pho² (RB)

NK: phɔ²¹ (N.PO), - (S.PO)

CK: - (KN); phɛ^{5 5} (KH); phɔ³³ (BW); pho³³ (KW)

SK: phɔ^{4 5} (N.SG), pho²¹ (S.SG); pho^{4 5} (N.PW), phu²¹ (S.PW)

104. 'enter' PK: *n() ɲ ɲ *ny¹ (RB)

NK: - (N.PO), nɔ^{4 5} (S.PO)

CK: - (KN); no³³ (KH); nu¹¹ (BW); nu³³ (KW)

SK: nɪ²¹ (N.SG), nɔ¹¹ (S.SG); na²¹ (N.PW), - (S.PW)

105. 'exert forcing to expel' PK: *chuŋ ɲ

NK: chən³¹ (N.PO), chuə m³¹ (S.PO)

CK: chwaŋ^{5 3} (KN); chī³³ (KH); - (BW); chu^{5 5} (KW)

SK: chu³³ (N.SG), - (S.SG); - (N.PW), - (S.PW)

106. 'extinguish (t.v.)' PK: *p/bi□□ *phi□¹ (RB)

NK: pe□²¹ (N.PO), pi□²¹ (S.PO)

CK: pī□²¹ (KN); pi³³ (KH); bI¹¹ (BW); pi³³ (KW)

SK: phi□^{4 5} (N.SG), pai¹¹ (S.SG); phai□²¹ (N.PW), phai□^{4 5} (S.PW)

Note Pa-O and N.Sgaw have tone D12 which suggests *p- in PK.

107. 'eye, face' PK: *mɛ□□ *mɛ□¹ (RB)

NK: mɛ□^{4 5} (N.PO), mɛ□^{4 5} (S.PO)

CK: - (KN); - (KH); - (BW); - (KW)

SK: mɛ□²¹ (N.SG), mɛ¹¹ (S.SG); me□²¹ (N.PW), me□^{4 5} (S.PW)

108. 'fall down' PK: *dɛ□□

NK: thɛ□^{4 5} (N.PO), tɛ□^{4 5} (S.PO)

CK: tɛ□²¹ (KN); tɑ³³ (KH); de¹¹ (BW); te³³ (KW)

SK: tɛ□²¹ (N.SG), tɛ¹¹ (S.SG); the□²¹ (N.PW), the□^{4 5} (S.PW)

109. 'fat (adj)' PK: *bluŋ□

NK: pləŋ^{5 5} (N.PO), pwəŋ³³ (S.PO)

CK: bwaŋ¹¹ (KN); bī¹¹ (KH); bŭ³³ (BW); bu¹¹ (KW)

SK: bɔ³¹ (N.SG), bɔ^{4 5} (S.SG); bɑ̃³³ (N.PW), bõ^{5 5} (S.PW)

110. 'feather, fur, body hair' PK: *chuN□ *chɔn² (RB)

NK: chən^{5 5} (N.PO), chuə m³³ (S.PO)

CK: chwaŋ¹¹ (KN); ɕī¹¹ (KH); ɕŭ³³ (BW); ɕu¹¹ (KW)

SK: chu³¹ (N.SG), chu^{4 5} (S.SG); chɑ̃³³ (N.PW), chõ^{5 5} (S.PW)

111. 'feel full (after eating)' PK: *kə□□ *ka□² (RB)

NK: kə□²¹ (N.PO); kə□²¹ (S.PO)

CK: kə□^{4 5} (KN); kɔ^{5 5} (KH); - (BW); kə³³ (KW)

SK: - (N.SG), - (S.SG); kaə□^{4 5} (N.PW), - (S.PW)

112. 'female, woman' PK: *hmi□ *my² (RB)
 NK: mi^{5 3} (N.PO), mi³³ (S.PO)
 CK: mi¹¹ (KN); mɔ¹¹ (KH); mo³³ (BW); mi¹¹ (KW)
 SK: mi³¹~ (N.SG), mi^{4 5} (S.SG); mɔ³³ (N.PW), mi^{5 5} (S.PW)
113. 'female (humans), mother' PK: *mu□ *mo¹ (RB)
 NK: mu^{5 3} (N.PO), mu^{5 5} (S.PO)
 CK: mu¹¹ (KN); mu¹¹ (KH); bo¹¹ mu^{5 5} (BW); mu¹¹ (KW)
 SK: mo¹¹ (N.SG), mu³¹ (S.SG); mu¹¹ (N.PW), mu³³ (S.PW)
114. 'fire' PK: *hme□ *me² (RB)
 NK: me^{5 5} (N.PO), me³³ (S.PO)
 CK: mai¹¹ (KN); mi¹¹ (KH); mɪ³³ (BW); mi (KW)
 SK: me³¹~ (N.SG), me^{4 5} (S.SG); mi³³ (N.PW), mi^{5 5} (S.PW)
115. 'fish' PK: *da□□
 NK: tha□^{4 5} (N.PO), tha□^{4 5} (S.PO)
 CK: ta□²¹ (KN); te³³ (KH); da¹¹-pho³³ (BW); tɔ³³ (KW)
 SK: - (N.SG), - (S.SG); - (N.PW), - (S.PW)
116. 'five' PK: *ɲjat□ *ɲɛ¹ (RB)
 NK: ɲat^{4 5} (N.PO), ɲap^{4 5}~ɲat^{4 5} (S.PO)
 CK: ɲɛ□²¹ (KN); ɲɛ³³ (KH); jɛ¹¹ (BW); jɛ³³ (KW)
 SK: jɛ¹¹ (N.SG), jɛ³¹ (S.SG); jɛ¹¹ (N.PW), jɛ³³ (S.PW)
Note It is noticeable that Sgaw and Pwo (SK) have tone B3.
117. 'flank, rib' PK: *rə□□ *ʏy□¹ (RB)
 NK: rə□^{4 5} (N.PO), rə□^{4 5} (S.PO)
 CK: rə□²¹ (KN); rɔ³³ (KH); - (BW); rə³³ (KW)
 SK: ʏu□²¹ (N.SG), ʏəi¹¹ (S.SG); ʏaə□²¹ (N.PW), ʏaə□^{4 5} (S.PW)
118. 'flat' PK: *pja□
 NK: pja^{5 5} (N.PO), pja³³ (S.PO)
 CK: bja¹¹ (KN); - (KH); bə³³ bə³³ (BW); - (KW)

SK: bi³¹~ (N.SG), - (S.SG); pa³³ (N.PW), bai^{5 5} (S.PW)

119. 'flea' PK: *kli~/*kli~

NK: khli³¹ (N.PO), - (S.PO)

CK: khli^{4 5} (KN); kla^{5 5} (KH); kle³³ (BW); kle³³ (KW)

SK: kli^{5 5}, kli^{4 5} (N.SG), kli^{5 5} (S.SG); khlei^{5 5} (N.PW), khlei^{5 3} (S.PW)

Note S.Pa-O, Pwo and some varieties of Sgaw have tone A, while CK has tone D.

120. 'flesh, meat' PK: *hja~ *ja² (RB)

NK: ja^{5 5} (N.PO), ja³³ (S.PO)

CK: ja¹¹ (KN); ja¹¹ (KH); - (BW); ja¹¹ (KW)

SK: ja³¹~ (N.SG), ja^{4 5} (S.SG); ja³³ (N.PW), ja^{5 5} (S.PW)

121. 'flower' PK: *pho~ *pho⁴ (RB)

NK: - (N.PO), - (S.PO)

CK: phau^{5 3} (KN); pho³³ (KH); pho^{5 5} (BW); pho^{5 5} (KW)

SK: pho³³ (N.SG), pho^{5 5} (S.SG); pho^{3 5} (N.PW), pho^{5 3} (S.PW)

122. 'forget' PK: *s-be~

NK: phe³³ (N.PO), pe^{5 3} (S.PO)

CK: pe³³ (KN); pa¹¹ (KH); - (BW); pe³³ (KW)

SK: - (N.SG), pi³³ (S.SG); - (N.PW), - (S.PW)

123. 'forest, jungle' PK: *kla~

NK: khla³¹ (N.PO), kla³¹ (S.PO)

CK: - (KN); kle³³ (KH); - (BW), ka^{5 5} (KW)

SK: pja¹¹ la^{4 5} kla^{4 5} (N.SG), kla^{5 5} (S.SG); kla^{5 5} (N.PW), kla³¹ (S.PW)

124. 'four' PK: - (?) PNK: *lit~; PCK: *hlwi~; PSK: *lwi~ *lwi¹ (RB)

NK: lip^{4 5}~lit^{4 5} (N.PO), lip^{4 5}~lit^{4 5} (S.PO)

CK: lwi^{5 3} (KN); lwi³³ (KH); lu^{5 5} (BW); li^{5 5} (KW)

SK: lwi¹¹ (N.SG), lwi³¹ (S.SG); lwi¹¹ (N.PW), lei³³ (S.PW)

Note The PK form cannot be reconstructed since PNK, PCK and PSK have different onsets and tones.

125. 'fox, wolf' PK: *thwi~mi~

NK: thwi^{5 5} mi³³ (N.PO), thwi³³ mi^{5 3} (S.PO)

CK: thwi¹¹mi³³ (KN); thwi¹¹mi¹¹ (KH); - (BW); thi¹¹mi³³ (KW)

SK: thwi¹¹mi³³ (N.SG), thwi^{4 5}mi³³ (S.SG); - (N.PW), thwi^{5 5}mei³¹ (S.PW)

126. 'frog' PK: ***de** *de² (RB)

NK: de^{5 5} (N.PO), de³³ (S.PO)

CK: dai¹¹ (KN); di¹¹ (KH); di³³ (BW); di¹¹ (KW)

SK: de³¹ (N.SG), de^{4 5} (S.SG); di³³ (N.PW), di^{5 5} (S.PW)

127. 'front' PK: ***ha**

NK: ha³¹ (N.PO), ha³¹ (S.PO)

CK: ha^{5 3} (KN); he³³ (KH); - (BW); - (KW)

SK: ha³³ (N.SG), ha^{5 5} (S.SG); - (N.PW); - (S.PW)

128. 'fruit' PK: ***sa** *sa² (RB)

NK: - (N.PO), - (S.PO)

CK: sa¹¹ (KN); se¹¹ (KH); se³³ (BW); sa¹¹ (KW)

SK: sa³¹ (N.SG), sa^{4 5} (S.SG); sa³³ (N.PW), sa^{5 5} (S.PW)

129. 'full' PK: ***pw**

NK: pw³¹ (N.PO), pw³¹ (S.PO)

CK: pw^{5 3} (KN); ba³³ (KH); pwe^{5 5} (BW); b^{5 5} (KW)

SK: pw³³ (N.SG), - (S.SG); - (N.PW), - (S.PW)

130. 'ginger' PK: ***eh** *eh³ (RB)

NK: eh³¹ (N.PO), eh³¹ (S.PO)

CK: eh^{5 3} (KN); ea³³ (KH); e^{5 5} (BW); eh^{5 5} (KW)

SK: e³¹ (N.SG), e³³ (S.SG); i^{5 5} (S.SG); ai^{5 5} (N.PW), ei³¹ (S.PW)

131. 'go' PK: ***lw**

NK: lw³³ (N.PO), lw^{5 3} (S.PO)

CK: lw³³ (KN); - (KH); le³³ (BW); lw³³ (KW)

SK: lw³³ (N.SG), lw³³ (S.SG); le^{5 5} (N.PW), le³¹ (S.PW)

132. 'grasshopper' PK: ***dw**

NK: dw^{5 5} (N.PO), dw³³ (S.PO)

CK: dwɛ¹¹ (KN); - (KH); d̥ɛ³³ (BW); dɛ¹¹ (KW)

SK: dwɛ³¹̃ (N.SG), dwɛ^{4 5} (S.SG); thwɛ¹¹ (N.PW), thwe³³ (S.PW)

Note The initial th- and tone B3 in Pwo suggest a plain voiced *d- in PK.

Probably, *d̥- in PK became *d- in Pwo and then *d- became th-.

133. 'grandchild' PK: *lej̯̚ *li³ (RB)

NK: li³³ (N.PO), li^{5 3} (S.PO)

CK: - (KN); lɔ̯¹¹ (KH); li¹¹ (BW); li³³ (KW)

SK: li³³ (N.SG), li³³ (S.SG); lei^{5 5} (N.PW), lei³¹ (S.PW)

Note There is an irregular tone change in Bwe. The A3 tone in Bwe should be 33, not 11.

134. 'great-grandchild' PK: *lo̯̚ *lo³ (RB)

NK: - (N.PO), lo^{5 3} (S.PO)

CK: lau³³ (KN); lu¹¹ (KH); - (BW); lu³³ (KW)

SK: lo³³ (N.SG), lu³³ (S.SG); lu^{5 5} (N.PW), lu³¹ (S.PW)

135. 'great-great-gandchild' PK: *la̯̚ *la³ (RB)

NK: la³³ (N.PO), la^{5 3} (S.PO)

CK: la³³ (KN); la¹¹ (KH); - (BW); la³³ (KW)

SK: la³³ (N.SG), la³³ (S.SG); la^{5 5} (N.PW), la³¹ (S.PW)

136. 'hair (head~)' PK: *khu̯̚slu̯̚

NK: ka̯̚^{4 5}lu³³ (N.PO), ka³³lu^{5 3} (S.PO)

CK: ka̯̚²¹lu³³ (KN); kho^{5 5}lo¹¹ (KH); khə̯̚^{5 5}lu^{5 5} (BW); khu¹¹lu³³ (KW)

SK: kho³¹̃su³³ (N.SG), khu^{4 5}'su^{5 5} (S.SG); khu³³su^{3 5} (N.PW), - (S.PW)

137. 'handle' PK: *do̯̚

NK: tho^{5 3} (N.PO), to^{5 5} (S.PO)

CK: - (KN); tu¹¹ (KH); -do³³ (BW); tu¹¹ (KW)

SK: to¹¹ (N.SG), tu³¹ (S.SG); thu¹¹ (N.PW), thu³³ (S.PW)

138. 'hat' PK: *~kl()ŋ̯̚

NK: koŋ̯̚^{5 5} (N.PO), koŋ̯̚³³ (S.PO)

CK: klɔ̯̚¹¹ (KN); ko¹¹ (KH); k̥ɔ̯̚³³phla³³ (BW); klɔ̯̚¹¹ (KW)

SK: - (N.SG), - (S.SG); - (N.PW), - (S.PW)

139. 'hawk, eagle' PK: *lek
 NK: lek^{4 5} (N.PO), lek^{4 5}~lik^{4 5} (S.PO)
 CK: li²¹ (KN); la³³ (KH); le¹¹ (BW); le³³ (KW)
 SK: li²¹ (N.SG), lai¹¹ (S.SG.); lai²¹ (N.PW), lai^{4 5} (S.PW)
140. 'have, be, exist, stay, there is' PK: *ɔ
 NK: ɔ^{5 5} (N.PO), ɔ^{au33} (S.PO)
 CK: ɔ^{au11} (KN); ɔ^{o11} (KH); ɔ³³ (BW); ɔ^{o11} (KW)
 SK: ɔ^{o31}~ (N.SG), ɔ^{u4 5} (S.SG); ɔ³³ (N.PW), ɔ^{5 5} (S.PW)
141. 'head' PK: *klə
 NK: klə^{21tu5 5} (N.PO), klə^{21tu33} (S.PO)
 CK: klə^{4 5} (KN); klə^{5 5} (KH); ku³³klə³³ (BW); ki³³~klə³³ (KW)
 SK: - (N.SG), - (S.SG); - (N.PW); - (S.PW)
142. 'heart' PK: *sa *sa² (RB)
 NK: sa²¹ (N.PO), sa²¹ (S.PO)
 CK: θa^{4 5} (KN); se^{5 5} (KH); θa³³ (BW); sɔ³³ (KW)
 SK: sa^{4 5} (N.SG), sa²¹ (S.SG); sa^{4 5} (N.PW); θa²¹ (S.PW)
143. 'heavy' PK: *thə
 NK: thi^{5 5} (N.PO), thi³³ (S.PO)
 CK: thəⁱ¹¹ (KN); tho¹¹ (KH); tho³³ (BW); thə¹¹ (KW)
 SK: - (N.SG), - (S.SG); - (N.PW), - (S.PW)
144. 'hemp' PK: *srɔ
 NK: sau^{5 5} (N.PO), ɔ^{au33} (S.PO)
 CK: θau¹¹ (KN); so¹¹ (KH); θrɔ³³ (BW); so¹¹ (KW)
 SK: ɔ^{ɔ31}~ (N.SG), ɔ^{ɔ4 5} (S.SG); ɔ^{jɔ33} (N.PW), θo^{5 5} (S.PW)
145. 'hip' PK: *glam
 NK: - (N.PO), kwam^{5 5} (S.PO)
 CK: k^{a11}kⁱ¹¹ (KN); kla¹¹ (KH); - (BW); ka¹¹ (KW)
 SK: khi³³ko¹¹ (N.SG), kɔ³¹ (S.SG); kwã¹¹ (N.PW), khwɔ̃³³ (S.PW)
146. 'hit hard (with a fist)' PK: *doŋ

NK: doŋ³³ (N.PO), toŋ^{5 3} (S.PO)

CK: dəu¹¹ (KN); thɔ¹¹ (KH); to³³ (BW); thə¹¹ (KW)

SK: tɔ¹¹ (N.SG), - (S.SG); thə¹¹ (N.PW), - (S.PW)

Note NK has the A tone instead of the B tone as in CK and SK.

147. 'horn' PK: *noŋ□ *noŋ³ (RB)

NK: noŋ³³ (N.PO), noŋ^{5 3} (S.PO)

CK: no³³ (KN); nɔ¹¹ (KH); -no¹¹ (BW); nə (KW)

SK: nə³³ (N.SG), nə³³ (S.SG); nău^{5 5} (N.PW), nõu³¹ (S.PW)

Note Bwe has an rregular tone. Modern Bwe should have tone 33(A3) instead of Tone 11 (D3).

148. 'hornet' PK: *phrĩN□

NK: phrĩm³¹ (N.PO), phrĩn³¹ (S.PO)

CK: phrĩ^{5 3} (KN); phlĩ³³ (KH); phlu^{5 5} (BW); phrĩ^{5 5} (KW)

SK: phlɔ³³ (N.SG), phlĩ^{5 5} (S.SG); phlă^{3 5} (N.PW), phlăĩ^{5 3} (S.PW)

Note In SK *phr- has become phl-.

149. 'horse' PK: *k-sre□ *se² (RB)

NK: se³¹ (N.PO), se³¹ (S.PO)

CK: θi^{4 5} (KN), si^{5 5} (KH); θrɪ³³ (BW); si³³ri³³ (KW)

SK: kə³³se¹¹ (N.SG), kə¹¹se^{4 5} (S.SG); ka²¹si³³ (N.PW), kə¹¹θi^{5 5} (S.PW)

Note The PK tone cannot be reconstructed due to the fact that NK, CK and SK

have tones A, D and B, respectively. This etymon is an Austroasiatic loanword. The reconstructed form for 'horse' in Proto-Monic is *ksɛh

(Diffloth, 1984).

150. 'hundred' PK: *g-rja□ *ŋa³ (RB)

NK: rja³³ (N.PO), rja^{5 3} (S.PO)

CK: ja³³ (KN); je¹¹ (KH); gə³³jɛ³³ (BW); ja³³ (KW)

SK: kə¹¹ja³³ (N.SG), kə¹¹ja³³ (S.SG); ja^{5 5} (N.PW), ja³¹ (S.PW)

151. 'husband' PK: *gwa□

NK: wa³³ (N.PO), wa^{5 3} (S.PO)

CK: gwa³³ (KN); we¹¹ (KH); wa³³ (BW); wu³³ (KW)

- SK: wa³³ (N.SG), wa³³ (S.SG); wa^{5 5} (N.PW), wa³¹ (S.PW)
152. 'ill, painful, sick' PK: *cha¹ *cha⁴ (RB)
- NK: cha³¹ (N.PO), cha³¹ (S.PO)
- CK: cha^{5 3} (KN); che³³ (KH); ɛ^{5 5} (BW); cha^{5 5} (KW)
- SK: cha³³ (N.SG), cha^{5 5} (S.SG); cha^{3 5} (N.PW), cha^{5 3} (S.PW)
153. 'insect' PK: *gra¹ *xa¹ (RB)
- NK: sa^{5 3} (N.PO), ɬa^{5 5} (S.PO)
- CK: ca¹¹ (KN); khre¹¹ (KH); - (BW); ca¹¹ (KW)
- SK: xa¹¹ (N.SG), xa³¹ (S.SG); xa¹¹ (N.PW), xa³³ (S.PW)
154. 'intestines' PK: *bre¹ *phɣi¹ (RB)
- NK: phre^{4 5} (N.PO), phri^{4 5} (S.PO)
- CK: pri²¹ (KN); phra³³ (KH); - bwi¹¹ (BW); pre³³ (KW)
- SK: phɣi²¹ (N.SG), pɣai¹¹ (S.SG); xwai²¹ (N.PW), xwai^{4 5} (S.PW)
- Note* Pwo has an unusual onset, xw- .
155. 'iron' PK: *tha¹ *tha² (RB)
- NK: tha²¹ (N.PO), tha²¹ (S.PO)
- CK: tha^{4 5} (KN); the^{5 5} (KH); tha³³ (BW); thɔ³³ (KW)
- SK: tha^{4 5} (N.SG); tha²¹ (S.SG); tha^{4 5} (N.PW), tha²¹ (S.PW)
156. 'itchy' PK: *sa¹
- NK: sa²¹ (N.PO), sa²¹ (S.PO)
- CK: θa^{4 5} (KN); se^{5 5} (KH); θa³³ (BW); sɔ³³ (KW)
- SK: sa^{4 5} (N.SG), sa²¹ (S.SG); sa^{4 5} (N.PW), θa²¹ (S.PW)
157. 'kidney' PK: *k-le¹ *le¹ (RB)
- NK: le^{5 3} (N.PO), - (S.PO)
- CK: lai¹¹ (KN); li¹¹ (KH); -li¹¹θɛ³³ (BW); li¹¹ (KW)
- SK: kə³³le¹¹ (N.SG), kə¹¹le³¹ (S.SG); li¹¹ (N.PW), li³³ (S.PW)
158. 'leaf' PK: *hla¹ *la² (RB)
- NK: la^{5 5} (N.PO), la³³ (S.PO)
- CK: la¹¹ (KN); le¹¹ (KH); lɛ³³ (BW); la¹¹ (KW)

SK: la³¹~ (N.SG), la^{4 5} (S.SG); la³³ (N.PW), la^{5 5} (S.PW)

159. 'leech (land~)' PK: *s-wa [] [] *wa []¹ (RB)

NK: wa []²¹ (N.PO), wa []²¹ (S.PO)

CK: θu []^{4 5} (KN); swa^{5 5} (KH); - (BW); su³³ (KW)

SK: su []^{4 5} (N.SG), sɔu []²¹ (S.SG); wa []²¹ (N.PW), wa []^{4 5} (S.PW)

Note NK, CK and Sgaw (SK) have tone D12; however, it is D3 in both N.Pwo and S.Pwo.

160. 'leech (water~)' PK: *k/s-lej [] [] *li³ (RB)

NK: - (N.PO), - (S.PO)

CK: - (KN); - (KH); gɔ³³ I³³ (BW); su¹¹ li³³ (KW)

SK: sɔ¹¹ li³³ (N.SG), tɔ¹¹ li³³ (S.SG); lei^{5 5} (N.PW), lei³¹ (S.PW)

161. 'left (side)' PK: *ce [] [] *ce² (RB)

NK: ce^{5 5} (N.PO), ce³³ (S.PO)

CK: cai¹¹ (KN); ci¹¹ (KH); c I³³ (BW); ci¹¹ (KW)

SK: ce³¹~ (N.SG), ce^{4 5} (S.SG); ci³³ (N.PW), ci^{5 5} (S.PW)

162. 'leg' PK: *khaŋ [] []

NK: khaŋ^{5 5} (N.PO); khaŋ³³ (S.PO)

CK: haŋ¹¹ (KN); khɛ¹¹ (KH); kha³³ (BW); khɔ¹¹ (KW)

SK: khɔ³¹~ (N.SG), khɔ^{4 5} (S.SG); khɛ³³ (N.PW), khɔ^{5 5} (S.PW)

163. 'lemur (flying~), flying squirrel' PK: *ple [] []/ble [] [] *phli []¹ (RB)

NK: phli []²¹ (N.PO), - (S.PO)

CK: - (KN); pla³³ (KH); bli³³ (BW); ple³³ (KW)

SK: phli []^{4 5} (N.SG), plai¹¹ (S.SG); phlai []²¹ (N.PW), phlai []^{4 5} (S.PW)

Note Some Karenic languages have tone D12 which suggests *pl-, e.g. N.Pa-O, Bwe and some varieties of Sgaw.

However, both N.Pwo and S.Pwo have the D3 tone which indicates a plain voiced *bl-.

164. 'lick (v.)' PK: *hlaj [] [] *lɛŋ² (RB)

NK: - (N.PO), lai³³ (S.PO)

CK: lai¹¹ (KN); - (KH); l I³³ (BW); li¹¹ (KW)

SK: le³¹~ (N.SG), li^{4 5} (S.SG); - (N.PW), lai^{5 5} (S.PW)

165. 'lie down, sleep' PK: *hmej [] [] *mi⁴ (RB)

NK: - (N.PO), - (S.PO)

CK: mai^{5 3} (KN); mǎ³³ (KH); ɲǎ^{5 5} mi^{5 5} (BW); mi^{5 5} (KW)

SK: mi³³ (N.SG), mi^{5 5} (S.SG); mei^{3 5} (N.PW), mei^{5 3} (S.PW)

Note Kayan has an irregular vowel, ai; the regular one should be i̇.

166. 'light (~weight)' PK: *phwi

NK: phwi³¹ (N.PO), phwi³¹ (S.PO)

CK: phwi^{5 3} (KN); phwi³³ (KH); phwi^{5 5} ~ phi^{5 5} (BW); phi^{5 5} (KW)

SK: phɣi³³ (N.SG), phɣi^{5 5} (S.SG); khwi^{3 5} (N.PW), khwi^{5 3} (S.PW)

Note The initial khw- in Pwo is irregular.

167. 'linear measure of about 8-9 inches (a span)' PK: *tha *tha⁴ (RB)

NK: tha³¹ (N.PO), tha³¹ (S.PO)

CK: tha^{5 3} (KN); the³³ (KH); - (BW); - (KW)

SK: tha³³ (N.SG), tha^{5 5} (S.SG); tha^{3 5} (N.PW), tha^{5 3} (S.PW)

168. 'linear measure of about half a yard (closed-hand)' PK: *bla

NK: - (N.PO), - (S.PO)

CK: pla¹¹ (KN); ple¹¹ (KH); - (BW); pla¹¹ (KW)

SK: pla¹¹ (N.SG), pla³¹ (S.SG); phla¹¹ (N.PW), phla³³ (S.PW)

169. 'linear measure of about half a yard (open-hand)' PK: *dɪm

NK: thom³³ (N.PO), tǎm^{5 3} (S.PO)

CK: tɪ³³ (KN); tɪ¹¹ (KH); - (BW); tɪ³³ (KW)

SK: tǎ³³ (N.SG), tɪ³³ (S.SG); thǎ^{5 5} (N.PW); thǎ³¹ (S.PW)

170. 'linear measure of about two yards' PK: *khlɛN

NK: khlɛm³¹ (N.PO), khlɛŋ³¹ (S.PO)

CK: khlɛŋ^{5 3} (KN); khlɛ³³ (KH); - (BW), - (KW)

SK: khlɛ³³ (N.SG), khlɛ^{5 5} (S.SG); khlǎi^{3 5} (N.PW), khlǎi^{5 3} (S.PW)

171. 'listen to' PK: *-hna

NK: na^{5 5} (N.PO), na³³ (S.PO)

CK: na¹¹ (KN); - (KH); ɲǎ^{5 5} ne³³ (BW), na¹¹ (KW)

SK: kǎ³³ na³¹ ~ (N.SG), na^{4 5} (S.SG); na³³ (N.PW), na^{5 5} (S.PW)

172. 'live, stay' PK: ***ṽṽn** ***ṽo²** (RB)
 NK: **ṽoṽ³¹** (N.PO), **ṽoṽ³¹** (S.PO)
 CK: **ṽau¹¹** (KN); **ṽo¹¹** (KH); **ṽ³³⁻** (BW); **ṽo¹¹** (KW)
 SK: **ṽo³¹~** (N.SG), **ṽu^{4 5}** (S.SG); **ṽṽ³³** (N.PW), **ṽo^{5 5}** (S.PW)
173. 'liver' PK: ***swiṽ** ***sun²** (RB)
 NK: **siṽ^{5 5}** (N.PO), **su ṽn³³** (S.PO)
 CK: **ṽi¹¹** (KN); **so¹¹** (KH); **ṽ^{5 5}ṽa³³** (BW), **~βu¹¹** (KW)
 SK: **su³¹~** (N.SG), **su^{4 5},** (S.SG); **sṽ³³** (N.PW), **ṽṽi^{5 5}** (S.PW)
174. 'long (of time)' PK: ***ṽjaṽ** ***ṽi¹** (RB)
 NK: **ṽjaṽ^{4 5}** (N.PO), **ṽjaṽ^{4 5}** (S.PO)
 CK: **ṽi³³** (KN); **ṽja³³** (KH); **ji¹¹** (BW), **ji³³** (KW)
 SK: **jiṽ²¹** (N.SG); **ji¹¹** (S.SG); **jaiṽ²¹** (N.PW), **jaiṽ^{4 5}** (S.PW)
175. 'long (of length), tall, high' PK: ***ihaw**
 NK: **tho³¹** (N.PO), **tho³¹** (S.PO)
 CK: **thau^{5 3}** (KN); **thu³³** (KH); **tho^{5 5}** (BW); **thu^{5 5}** (KW)
 SK: **thṽ³³** (N.SG), **thṽ^{5 5}** (S.SG); **thṽ^{3 5}** (N.PW), **tho^{5 3}** (S.PW)
176. 'look after (things)' PK: ***pə**
 NK: **pə^{5 5}** (N.PO), **pə³³** (S.PO)
 CK: **pəi¹¹** (KN); **po¹¹** (KH); **pə³³i^{5 5}** (BW); **pə¹¹** (KW)
 SK: - (N.SG), - (S.SG); - (N.PW); - (S.PW)
177. 'loris' PK: ***k-ch()N**
 NK: - (N.PO), - (S.PO)
 CK: **chəṽ^{5 3}** (KN); **chṽ³³** (KH); - (BW); **chə^{5 5}** (KW)
 SK: **kə¹¹che³³** (N.SG), **ta¹¹chi^{5 5}** (S.SG); **chāi^{3 5}** (N.PW), **kə¹¹chēi^{5 3}** (S.PW)
178. 'louse (chicken~), insect' PK: ***(chjaN)gra**
 NK: **ṽja³¹ sa^{5 3}** (N.PO), **ṽja³¹ ṽa^{5 5}** (S.PO)
 CK: **ṽi^{5 3}ca¹¹** (KN); **cha³³khre¹¹** (KH); - (BW); **ṽi^{5 5}ca¹¹** (KW)
 SK: **chṽ³³xa¹¹** (N.SG), **chṽ^{5 5}xa³¹** (S.SG); **chē^{3 5}xa¹¹** (N.PW), **chṽ^{5 3}xa³³** (S.PW)

179. 'louse (head~)' PK: *swiN $\square$ *sun² (RB)
 NK: si^{5 5} (N.PO), si³³ (S.PO)
 CK: θi¹¹ (KN); so¹¹ (KH); θo¹¹ (BW); si¹¹ (KW)
 SK: su³¹~ (N.SG), su^{4 5} (S.SG); sã³³ (N.PW), θãi^{5 5} (S.PW)
180. 'lung' PK: *p-so $\square$
 NK: - (N.PO), - (S.PO)
 CK: θau¹¹ (KN); su¹¹ (KH); θə^{5 5}·o^{5 5} (BW); su¹¹ (KW)
 SK: pə³³so³¹~ (N.SG), su^{4 5} (S.SG); so $\square$ ^{4 5} (N.PW), bu³³θu $\square$ ²¹ (S.PW)
 Note Pwo has tone D while CK and Sgaw have tone B.
181. 'maggot' PK: *hloŋ $\square$ *loŋ² (RB)
 NK: loŋ^{5 5} (N.PO), loŋ³³ (S.PO)
 CK: lo¹¹ (KN); lo¹¹ (KH); lo³³ (BW); lo¹¹ (KW)
 SK: lo³¹~ (N.SG), lo^{4 5} (S.SG); lãu³³ (N.PW), lõu^{5 5} (S.PW)
182. 'male (of animals), father' PK: *pha $\square$ *pha⁴ (RB)
 NK: pha³¹ (N.PO), pha³¹ (S.PO)
 CK: pha^{5 3} (KN); phe³³ (KH); - (BW); pha^{5 5} (KW)
 SK: pha³³ (N.SG), pha^{5 5} (S.SG); pha^{3 5} (N.PW), pha^{5 3} (S.PW)
183. 'man, male' PK: *khwa $\square$ *khwa⁴ (RB)
 NK: kho³¹ (N.PO), kho³¹ (S.PO)
 CK: khau^{5 3} (KN); khu³³ (KH); - bo¹¹khə^{5 5} (BW); khu^{5 5} (KW)
 SK: khwa³³ (N.SG), khwa^{5 5} (S.SG); khwa^{3 5} (N.PW), khwa^{5 3} (S.PW)
184. 'mango' PK: *khɔ $\square$ $\square$
 NK: khɔ $\square$ ²¹ (N.PO), khɔ $\square$ ²¹ (S.PO)
 CK: khau $\square$ ^{4 5} (KN); khɛ^{5 5} (KH); - (BW); kho³³ (KW)
 SK: khɔ $\square$ ^{4 5} (N.SG), kho $\square$ ²¹ (S.SG); kho $\square$ ^{4 5} (N.PW), khu $\square$ ²¹ (S.PW)
185. 'many, much' PK: * $\square$ a $\square$ * $\square$ a⁴ (RB)
 NK: $\square$ a³¹ (N.PO), $\square$ a³¹ (S.PO)
 CK: $\square$ a^{5 3} (KN); $\square$ a³³ (KH); ɔ³³ɛ³³ (BW); - (KW)

SK: ɔ̃³³ (N.SG), ɔ̃^{5 5} (S.SG); ɔ̃³³ (N.PW), ɔ̃^{5 5} (S.PW)

Note Kayah has an irregular vowel, a instead of e. Bwe and Pwo have tone B while the rest have tone A.

186. 'mat' PK: *khɪɔ̃ ʔ *khɪɔ̃² (RB)

NK: khɪɔ̃^{5 5} (N.PO), khɪɔ̃³³ (S.PO)

CK: - (KN); - (KH); khɪɔ̃³³ (BW); - (KW)

SK: khɪɔ̃³¹ (N.SG), khɪɔ̃^{4 5} (S.SG); khɪɔ̃³³ (N.PW), khɪɔ̃^{5 5} (S.PW)

187. 'milipede' PK: *k/s-wɛj ʔ ʔ

NK: cɔ̃^{5 5} wɛ³³ (N.PO), cɔ̃¹¹ wɛ^{5 3} (S.PO)

CK: θã²¹ kə̃¹¹ wi³³ (KN); si^{5 5} kɔ̃^{5 5} wi¹¹ (KH); - (BW); te¹¹ we³³ (KW)

SK: si^{4 5} wɔ̃¹¹ de³³ (N.SG), si^{5 5} wa³¹ de^{5 5} (S.SG); wai¹¹ (N.PW), wai³¹ (S.PW)

Note NK, CK and S.Pwo have tone A while Sgaw and N.PWO have tone B.

188. 'mole' PK: *(jow ʔ) wi ʔ *wi³ (RB)

NK: ju^{5 3} wi³³ (N.PO), ju^{5 5} wi^{5 3} (S.PO)

CK: - (KN); jo¹¹ wi¹¹ (KH); wi³³ (BW), ju¹¹ wi³³ (KW)

SK: wi³³ (N.SG), wi³³ (S.SG); wei^{5 5} (N. PW), wei³¹ (S.PW)

189. 'money' PK: *hrun ʔ

NK: rən³¹ (N.PO), ruən³¹ (S.PO)

CK: rwaŋ^{5 3} (KN), ri³³ (KH); hŭ^{5 5} (BW); ru^{5 5} (KW)

SK: - (N.SG), - (S.SG); - (N.PW), - (S.PW)

190. 'monitor lizard (land~)' PK: *khwo ʔ ʔ

NK: - (N.PO), - (S.PO)

CK: re³³ khɔ̃^{4 5} (KN); khwɔ̃^{5 5} (KH); - (B); - (KW)

SK: khu^{4 5} (N.SG), khɔ̃²¹ (S.SG); khau^{4 5} (N.PW), khɔ̃³¹ (S.PW)

191. 'monitor lizard (water~)' PK: *tre ʔ *reŋ³ (RB)

NK: - (N.PO), - (S.PO)

CK: - (KN); re³³ (KH); tre^{5 5} (BW); tre^{5 5} ~ te^{5 5} re^{5 5} (KW)

SK: kre^{4 5} (N.SG), re^{5 5} (S.SG); - (N.PW), ɣei³¹ (S. PW)

Note Only N.Sgaw has tone D instead of A.

192. 'monkey' PK: *jo ʔ ʔ

NK: jo^{4 5} (N.PO), ju^{4 5} (S.PO)

CK: jɔ²¹ (KN); jɔ³³ (KH); jo¹¹ (BW); jə³³ (KW)

SK: - (N.SG), - (S.SG); - (N.PW), - (S.PW)

193. 'moon, month' PK: *la⁴ (RB)

NK: la³¹ (N.PO), la³¹ (S.PO)

CK: la^{5 3} (KN); le³³ (KH); lɛ^{5 5} (BW); la^{5 5} (KW)

SK: la³³ (N.SG), la^{5 5} (S.SG); la^{5 5} (N.PW), la³¹ (S.PW)

194. 'mother, female (of animals)' PK: *mə⁴

NK: mə^{5 3} (N.PO), mə^{5 5} (S.PO)

CK: mə¹¹ (KN); mo¹¹ (KH); - mo³³ (BW); mə¹¹ (KW)

SK: - (N.SG), - (S.SG); - (N.PW), - (S.PW)

195. 'mouse, rat' PK: *jow⁴ *ŋy¹ (RB)

NK: ju^{5 3} (N.PO), ju^{5 5} (S.PO)

CK: ɯ¹¹ (KN); jo¹¹ (KH); ju¹¹ (BW); ju¹¹ (KW)

SK: ji¹¹ (N.SG), ji³¹ (S.SG); jou¹¹ (N.PW), jou³³ (S.PW)

196. 'mynah' PK: *s-raN⁴

NK: ran^{5 5} (N.PO), ran³³ (S.PO)

CK: raŋ¹¹ (KN); - (KH); θə³³rɛ³³ (BW); sa¹¹ra¹¹ (KW)

SK: - (N.SG), - (S.SG); - (N.PW), - (S.PW)

Note Irregular vowels in CK, i.e. the vowel changes behave like the *-a rhyme

pattern, not the *-aŋ one.

197. 'nail' PK: *hmej⁴ *mi/ɛŋ/eŋ² (RB)

NK: mi^{5 5} (N.PO), mi³³ (S.PO)

CK: mi¹¹ (KN); mə¹¹ (KH); θə^{5 5}mi³³ (BW); mi¹¹ (KW)

SK: me³¹ (N.SG), mi^{4 5} (S.SG); mei³³ (N.PW), mei^{5 5} (S.PW)

198. 'name' PK: *min⁴ *min³ (RB)

NK: min³³ (N.PO), min^{5 3} (S.PO)

CK: mjəŋ³³ (KN); mi¹¹ (KH); - mi³³ (BW); mi³³ (KW)

SK: mi³³ (N.SG), mi³³ (S.SG); mǎi^{5 5} (N.PW); mǎi³¹ (S.PW)

199. 'narrow' PK: ***ɲɛn** *in² (RB)
 NK: ɲɛn^{5 5} (N.PO), ɲɛn³³ ~ ɲɛm³³ (S.PO)
 CK: ɲjaŋ¹¹ (KN); ɲi¹¹ (KH); - (BW); - ɲi¹¹ (KW)
 SK: ɲi³¹ (N.SG), i^{4 5} (S.SG); ǎi³³ (N.PW), ǎi^{5 5} (S.PW)
200. 'navel, umbilical cord' PK: ***de** *de³ (RB)
 NK: pa^{4 5} de³¹ (N.PO), de³¹ (S.PO)
 CK: dai^{5 3} (KN); di³³ (KH); -ɗi^{5 5} mU^{5 5} (BW); di³³ (KW)
 SK: de³³ (N.SG), de^{5 5} (S.SG); di^{5 5} (N.PW), di³¹ (S.PW)
201. 'neck, cheek' PK: ***bo** *bo² 'cheek' (RB)
 NK: - (N.PO), - (S.PO)
 CK: ka²¹ bɔ^{5 3} (KN); khɛ³³ bo³³ (KH); -ɓ³³ (BW); ko³³ bo^{5 5} (KW)
 SK: khɔ^{4 5} bo³³ (N.SG), - (S.SG); - (N.PW), kho²¹ bo³¹ (S.PW)
202. 'new' PK: ***saN** *saN⁴ (RB)
 NK: sa³¹ (N.PO), - (S.PO)
 CK: θi^{5 3} (KN); se³³ (KH); θɛ^{5 5} (BW); sa^{5 5} (KW)
 SK: sɔ³³ (N.SG), sɔ^{5 5} (S.SG); sɛ^{3 5} (N.PW), θɔ^{5 3} (S.PW)
203. 'nine' PNK: ***kot**; PCK, PSK : ***khwi** *khwi⁴ (RB)
 NK: kot^{4 5} (N.PO), kot^{4 5} (S.PO)
 CK: khwi^{5 3} (KN); - (KH); khwi^{5 5} (BW); sɔ³³ khi^{5 5} (KW)
 SK: khwi³³ (N.SG), khwi^{5 5} (S.SG); khwi^{5 5} (N.PW), khwi^{5 3} (S.PW)
204. 'obtain, receive, take, get' PK: ***ne** *ne¹ (RB)
 NK: ne^{5 3} (N.PO), - (S.PO)
 CK: nai¹¹ (KN); ni¹¹ (KH); ni^{5 5} (BW); ni¹¹ (KW)
 SK: ne¹¹ (N.SG), ne³¹ (S.SG); ni¹¹ (N.PW), ni³³ (S.PW)
205. 'old (of humans)' PK: ***bra** *phYɑ¹ (RB)
 NK: phra^{5 3} (N.PO), pra^{5 5} (S.PO)
 CK: pra¹¹ (KN); phre¹¹ (KH); θa³³ bwɛ³³ (BW); pra¹¹ (KW)
 SK: sa^{4 5} pYɑ¹¹ (N.SG), pYɑ³¹ (S.SG); - (N.PW), - (S.PW)

206. 'old (of things)' PK: *lej□ *laN¹ li³ (RB)

NK: li³³ (N.PO), - (S.PO)

CK: - (KN); lə¹¹ (KH); - (BW); li³³ (KW)

SK: lɔ¹¹ li³³ (N.SG), li³³ (S.SG); lei^{5 5} (N.PW); lei³¹ (S.PW)

207. 'one' PK: *t-la□

NK: ta²¹ (N.PO), ta²¹ (S.PO)

CK: la¹¹ ~ la²¹ (KN); t- (KH); to^{5 5}, tə³³⁻, də³³⁻ (BW); t- (KW)

SK: tə³³ (N.SG), tə²¹ (S.SG); la^{21~11} (N.PW), lə³¹ (S.PW)

Note The tone of this etymon cannot be reconstructed due to its regularity. Very often, the word 'one' is unstressed, so its vowel has been reduced. Moreover, in some CK languages there is also the harmony of vowels in

sesquisyllabic words, t 'C . □ □

208. 'open (the mouth)' PK: *□a□

NK: □aŋ³¹ (N.PO), □an³¹ (S.PO)

CK: □aŋ^{5 3} (KN); □ε³³ (KH); □^{5 5}-tha³³ (BW); □ɔ^{5 5} (KW)

SK: □ɔ³³ (N.SG), - (S.SG); - (N.PW), □³¹ (S.PW)

209. 'ox, bull, cow' PK: *bo□

NK: pho^{5 3} (N.PO), po^{5 5} (S.PO)

CK: pau¹¹ (KN); pu¹¹ (KH); bo³³ (BW); pu¹¹ (KW)

SK: - (N.SG), - (S.SG); - (N.PW), - (S.PW)

210. 'palm (of the hand)' PK: *ja□

NK: ja³³ (N.PO), ja^{5 3} (S.PO)

CK: □a³³ (KH); ja¹¹ (KH); - (BW); - (KW)

SK: ja³³ (N.SG), - (S.SG); jε^{3 5} (N.PW), ja^{5 3} (S.PW)

Note NK and CK have tone A3 which indicates a voice onset *j- but the tone A12 in SK suggests a voiceless onset *hj-.

211. 'pangolin, anteater' PK: *jo□

NK: ju³³ (N.PO), - (S.PO)

CK: □au³³ (KN); ju¹¹ (KH); j̥³³-θo³³ (BW); ju³³ (KW)

SK: j̥³³hɔ³¹ (N.SG), ju³³ (S.SG); ju¹¹ ~ (N.PW); j̥³¹ (S.PW)

Note S.Pwo has an irregular vowel, i.e. it should be u not i̥.

212. 'parrot, parakeet' PK: ***(thoꞑ)kiꞑ** *ki² (RB)
 NK: ki^{5 5} (N.PO), - (S.PO)
 CK: khi¹¹ (KN); khi¹¹ (KH); ki³³~khi³³ (BW); ki¹¹ (KW)
 SK: - (N.SG), ki^{4 5} (S.SG); kei³³ (N.PW), kei^{5 5} (S.PW)
213. 'path, trail, way' PK: ***glɛꞑ** *klɛ⁴ (RB)
 NK: - (N.PO), klai^{5 3} (S.PO)
 CK: khlɛ³³ (KN); - (KH); klɛ³³ (BW); klɛ³³ (KW)
 SK: klɛ³³~klɛꞑ^{4 5} (N.SG), klɛ³³ (S.SG); - (N.PW); - (S.PW)
214. 'peacock' PK: ***(thoꞑ)braꞑꞑ** *phYɑꞑ¹ (RB)
 NK: - (N.PO), - (S.PO)
 CK: praꞑ²¹ (KN); phre³³ (KH); ba¹¹ (BW); prɔꞑ³³ (KW)
 SK: bYɑꞑ²¹~phYɑꞑ²¹ (N.SG), pYɑ¹¹ (S.SG); sjaꞑ²¹ (N.PW); ɕɑꞑ^{4 5} (S.PW)
215. 'pen, coop' PK: ***k-bɔꞑꞑ**
 NK: phɔꞑ³³ (N.PO), pɔꞑ^{5 3} (S.PO)
 CK: pɔꞑ³³ (KN); po¹¹ (KH); - (BW); po³³ (KW)
 SK: kɔꞑ¹¹po³³ (N.SG), kɔꞑ²¹po³³ (S.SG); phɔꞑ^{5 5} (N.PW), phɔ̃³¹ (S.PW)
216. 'penis' PK: ***teꞑ** *the² (RB)
 NK: te^{5 5} (N.PO), te³³ (S.PO)
 CK: thai¹¹ (KN); thi¹¹ (KH); - (BW); thi¹¹ (KW)
 SK: the³¹ (N.SG), the^{4 5} (S.SG); thi³³ (N.PW), thi^{5 5} (S.PW)
217. 'pheasant' PK: ***(thoꞑ)reꞑꞑ**
 NK: - (N.PO), - (S.PO)
 CK: - (KN); ra³³ (KH); - (BW); - re³³ (KW)
 SK: Yɪꞑ²¹ (N.SG), Yai¹¹ (S.SG); Yaiꞑ²¹ (N.PW); Yaiꞑ^{4 5} (S.PW)
218. 'pig' PK: ***thɔꞑꞑꞑ** *thoꞑ² (RB)
 NK: thɔꞑ²¹ (N.PO), thɔꞑ²¹ (S.PO)
 CK: thauꞑ^{4 5} (KN); thɛ^{5 5} (KH); thɔꞑ³³ (BW); tho³³ (KW)
 SK: thɔꞑ^{4 5} (N.SG), thɔꞑ²¹ (S.SG); thoꞑ^{4 5} (N.PW); thuꞑ²¹ (S.PW)
219. 'pile' PK: ***plomꞑ**

NK: pom^{5 5} (N.PO), pom³³ (S.PO)

CK: blo¹¹ (KN); plu¹¹ (KH); plo³³ (BW); plə¹¹ (KW)

SK: pu³¹~ (N.SG), pu^{4 5}’ (S.SG); pãu³³ (N.PW), põu³¹ (S.PW)

Note Kayah and Sgaw have irregular vowels, i.e. u instead of ɔ̃ (Kayah) and u instead of ə (Sgaw).

220. ‘pointed (as of needles), sharp point’ PK: *cuŋ *cu³ (RB)

NK: cu³¹ (N.PO), kju³¹ (S.PO)

CK: cu^{5 3} (KN); cu³³ (KH); cu^{5 5} (BW); cu^{5 5} (KW)

SK: cuŋ^{4 5} (N.SG), cu^{5 5} (S.SG); cu^{5 5} (N.PW), cu³¹ (S.PW)

Note N.Sgaw has tone D while the rest have tone A.

221. ‘pole’ PK: *thuNŋ

NK: thun^{5 5} (N.PO), thuə m³³ (S.PO)

CK: thwaŋ¹¹ (KN); thĩ¹¹ (KH); - (BW); thu¹¹ (KW)

SK: thu³¹~ (N.SG), thu^{4 5}’ (S.SG); thã³³ (N.PW), thãĩ^{5 5} (S.PW)

222. ‘porcupine’ PK: *sunŋ

NK: sun^{5 5} (N.PO), suə m³³~suə n³³ (S.PO)

CK: θwaŋ¹¹ (KN); sĩ¹¹ (KH); θu³³ (BW); su¹¹ (KW)

SK: - (N.SG), su^{4 5}’ (S.SG); - (N.PW), - (S.PW)

223. ‘porcupine (brush-tailed)’ PK: *s-ŋbaŋ

NK: - (N.PO), - (S.PO)

CK: - (KN); se³³be³³ (KH); - (BW); sa^{5 5}ba^{5 5} (KW)

SK: ba³³ (N.SG), ba^{5 5} (S.SG); - (N.PW), θə^{5 5}ba³¹ (S.PW)

224. ‘pot’ PK: *bəŋŋ *phə³m³ (RB)

NK: phəŋ³³ (N.PO), pəŋ^{5 3} (S.PO)

CK: pə³³ (KN); pɔ̃¹¹ (KH); gə³³bo^{5 5} (BW); pə³³ (KW)

SK: sə¹¹pə³³ (N.SG), pə³³ (S.SG); phã^{5 5} (N.PW), phə³¹ (S.PW)

225. ‘pound (v.)’ PK: *dɔ̃ŋŋ

NK: thɔ̃ŋ^{5 3} (N.PO), tɔ̃ŋ^{5 5} (S.PO)

CK: tɔ̃¹¹ (KN); - (KH); dɔ̃³³ (BW); to¹¹ (KW)

- SK: to¹¹ (N.SG), tu³¹ (S.SG); - (N.PW), thō³³ (S.PW)
226. 'pull up' PK: *thɛ□□
 NK: - (N.PO), - (S.PO)
 CK: - (KN); - (KH); - (BW); the³³ (KW)
 SK: thɛ□^{4 5} (N.SG), thɛ□²¹ (S.SG); the□^{4 5} (N.PW); the□²¹ (S.PW)
227. 'pumpkin (ash~)' PK: *luN□
 NK: - (N.PO), - (S.PO)
 CK: lwaŋ¹¹ (KN); lī¹¹ (KH); lu¹¹ (BW); lu¹¹ (KW)
 SK: lu¹¹ (N.SG), lu³¹ (S.SG); lā¹¹ (N.PW), lō³³ (S.PW)
228. 'push' PK: *chaN□
 NK: - (N.PO), - (S.PO)
 CK: - (KN); chɛ¹¹ (KH); ɲa³³ (BW); chɔ¹¹ (KW)
 SK: chɔ³¹ (N.SG), chɔ^{4 5} (S.SG); chɛ³³ (N.PW), chō^{5 5} (S.PW)
229. 'put on, wear' PK: *sɔ□□ *so□² (RB)
 NK: sɔ□²¹ (N.PO), sɔ□²¹ (S.PO)
 CK: θau□^{4 5} (KN); sɛ^{5 5} (KH); θɔ³³ (BW); so³³ (KW)
 SK: sɔ□^{4 5} (N.SG), so□²¹ (S.SG); so□^{4 5} (N.PW); θu□²¹ (S.PW)
230. 'put on (a necklace), wear' PK: *le□□ *li□¹ (RB)
 NK: - (N.PO), - (S.PO)
 CK: li□²¹ (KN); la³³ (KH); li¹¹ (BW); le³³ (KW)
 SK: li□²¹ (N.SG), lai¹¹ (S.SG); lai□²¹ (N.PW); lai□^{4 5} (S.PW)
231. 'put out (the tongue)' PK: *hlɛm□
 NK: lɛm^{5 5} (N.PO), lɛm³³ (S.PO)
 CK: ljaŋ¹¹ (KN); - (KH); - (BW); li¹¹ (KW)
 SK: le³¹ (N.SG), li^{4 5} (S.SG); lāi³³ (N.PW), lāi^{5 5} (S.PW)
232. 'quail' PK: *(tho□)hrwi□□
 NK: ri□^{4 5} (N.PO), - (S.PO)
 CK: rwi□^{4 5} (KN); rwi^{5 5} (KH); - (BW); ri³³ (KW)
 SK: pə¹¹wi³³ (N.SG), pə¹¹wi^{5 5} (S.SG); phu³³wei³³ (N.PW); wei^{5 5} (S.PW)

Note The SK languages have irregular tone changes, i.e. Sgaw has tone A while Pwo has tone B.

233. 'rabbit, hare' PK: *p- $\square$ d ε $\square$

NK: - (N.PO), - (S.PO)

CK: d ε ^{5 3} (KN); d ε ³³ (KH); p $\mathfrak{a}$ ³³d ε ³³ (BW); d ε ^{5 5} (KW)

SK: p $\mathfrak{a}$ ¹¹d ε ³³ (N.SG), - (S.SG); p $\mathfrak{a}$ ³³d ε ^{5 5} (N.PW), p $\mathfrak{a}$ ³¹ d ε ³³ (S.PW)

234. 'raise, feed (animals)' PK: * $\square$ b $\mathfrak{a}$ $\square$

NK: - (N.PO), - (S.PO)

CK: b $\mathfrak{a}$ i¹¹ (KN); bo¹¹ (KH); - (BW); b $\mathfrak{a}$ ¹¹ (KW)

SK: b i³¹ $\sim$ (N.SG), b i^{4 5} (S.SG); - (N.PW), b i^{5 5} l $\mathfrak{a}$ ³¹ (S.PW)

235. 'rattan' PK: *re $\square$ * γ e¹ (RB)

NK: re^{5 3} (N.PO), re^{5 5} (S.PO)

CK: rai¹¹ (KN); ri¹¹ (KH); - (BW); ri¹¹ (KW)

SK: γ e¹¹ (N.SG), re³¹ (S.SG); γ i¹¹ (N.PW), γ i³³ (S.PW)

236. 'reach into' PK: * $\square$ bo $\square$ $\square$

NK: bo $\square$ ²¹ (N.PO), bu $\square$ ²¹ (S.PO)

CK: b $\mathfrak{a}$ $\square$ ^{4 5} (KN); b $\mathfrak{a}$ $\square$ ^{5 5} (KH); - (BW); b $\mathfrak{a}$ ³³ (KW)

SK: b i $\square$ ^{4 5} (N.SG), b $\mathfrak{a}$ u $\square$ ²¹ (S.SG); bau $\square$ ^{4 5} (N.PW); b $\mathfrak{a}$ u $\square$ ³¹ (S.PW)

237. 'resin' PK: *thro $\square$ $\square$

NK: - (N.PO), - (S.PO)

CK: th $\mathfrak{a}$ $\square$ ^{4 5} (KN); th $\mathfrak{a}$ $\square$ ^{5 5} (KH); thro³³ (BW); th $\mathfrak{a}$ ³³ (KW)

SK: th i $\square$ ^{4 5} (N.SG), th $\mathfrak{a}$ u $\square$ ²¹ (S.SG); thau $\square$ ^{4 5} (N.PW); th $\mathfrak{a}$ u $\square$ ³¹ (S.PW)

238. 'right (side)' PK: *thw ε $\square$ *thwe⁴ (RB)

NK: thw ε ³¹ (N.PO), thw ε ³¹ (S.PO)

CK: thw ε ^{5 3} (KN); thwa³³ (KH); thw ε ³³ (BW); th ε ^{5 5} (KW)

SK: chw ε ³³ (N.SG), thw ε ^{5 5} (S.SG); - (N.PW), thw ε ^{5 5} (S.PW)

Note S.Sgaw has tone B. Perhaps, the word thw ε ^{5 5} in S.Pwo was borrowed from S.Sgaw, thw ε ^{5 5}. There is no cognate in N.Pwo.

239. 'ripe, cooked' PK: *hmin $\square$ *min⁴ (RB)

NK: min³¹ (N.PO), min³¹ (S.PO)

CK: mjəŋ^{5 3} (KN); mi³³ (KH); mi^{5 5} (BW); mi^{5 5} (KW)

SK: mi³³ (N.SG), mi^{5 5} (S.SG); mǎi^{5 5} (N.PW), mǎi^{5 5} (S.PW)

240. 'roast (by stuffing into fresh bamboo sections)' PK: *deN

NK: then³³ (N.PO), - (S.PO)

CK: tɛ³³ (KN); tai¹¹ (KH); de³³ (BW); tɛ³³ (KW)

SK: tɛ³³ (N.SG), tɛ³³ (S.SG); - (N.PW), thɛi³¹ (S.PW)

241. 'root' PK: *rwi

NK: rwi^{5 3} (N.PO), rwi^{5 5} (S.PO)

CK: rwi¹¹ (KN); rwi¹¹ (KH); wi³³ (BW); ri¹¹ (KW)

SK: ʔi¹¹ (N.SG), ʔi³¹ (S.SG); ʔei¹¹ (N.PW), wei³³ (S.PW)

242. 'rotten, bad smell' PK: *səŋ

NK: səŋ³¹ (N.PO), səŋ³¹ (S.PO)

CK: θə^{5 3} (KN); sɔ³³ (KH); θə^{5 5} (BW); sə^{5 5} (KW)

SK: - (N.SG), - (S.SG); - (N.PW), - (S.PW)

243. 'rub, clean, polish, scrub' PK: *thu *thu² (RB)

NK: - (N.PO), - (S.PO)

CK: - (KN); thi^{5 5} (KH); - (BW); - thi³³ (KW)

SK: thu^{4 5} (N.SG), thɔu²¹ (S.SG); thu^{4 5} (N.PW); thɔu³¹ (S.PW)

244. 'salt' PK: *sa *sa² (RB)

NK: sa^{5 5} (N.PO), sa³³ (S.PO)

CK: θa¹¹ (KN); se¹¹ (KH); θɛ³³ (BW); sa¹¹ (KW)

SK: ʔi^{4 5} sa³¹ (N.SG), sa^{4 5} (S.SG); - (N.PW), - (S.PW)

245. 'salty, sweet' PK: *chim *chəm⁴ (RB)

NK: chom³¹ (N.PO), chəm³¹ (S.PO)

CK: chi^{5 3} (KN); chi³³ (KH); - (BW); chi^{5 5} (KW)

SK: chə³³ (N.SG), chi^{5 5} (S.SG); chǎ³⁵ (N.PW), chǎi^{5 3} (S.PW)

246. 'sand' PK: *hmɛ

NK: - (N.PO), - (S.PO)

CK: mɛ^{4 5} (KN); - (KH); - (BW); me³³ (KW)

SK: mɛ̃^{4 5} (N.SG), mɛ̃²¹ (S.SG); mẽ^{4 5} (N.PW); mẽ²¹ (S.PW)

247. 'scale (of fish)' PK: *s-beŋ

NK: beŋ^{5 5} (N.PO), beŋ³³ (S.PO)

CK: be¹¹ (KN); ba¹¹ (KH); - (BW); be¹¹ (KW)

SK: sɔ̃¹¹ bi³³ (N.SG), - (S.SG); - (N.PW), bɛ̃^{5 5} (S.PW)

Note Only Sgaw has tone A while Pwo (SK), NK and CK have tone B.

248. 'sell' PK: *ca *cha⁴ (RB)

NK: ca³¹ (N.PO), ca³¹ (S.PO)

CK: cha^{5 3} (KN); che³³ (KH); ɛ̃^{5 5} (BW); cha^{5 5} (KW)

SK: cha³³ (N.SG), cha^{5 5} (S.SG); cha^{3 5} (N.PW), cha^{5 3} (S.PW)

249. 'serow, mountain goat' PK: *jaj

NK: jaj³³ (N.PO), - (S.PO)

CK: jaj³³ (KN); je¹¹ (KH); ji³³ (BW); ji³³ (KW)

SK: - (N.SG), - (S.SG); - (N.PW), - (S.PW)

250. 'seven' PK: *nwet *nwe

NK: ni^{1 21} (N.PO), nɔ̃^{4 5} (S.PO)

CK: nwɛ̃^{5 3} (KN); - (KH); nwe^{5 5}~nwi^{5 5} (BW); sɔ̃³³nẽ^{5 5} (KW)

SK: nwi^{4 5} (N.SG), nwi^{5 5} (S.SG); nwɛ̃^{5 5} (N.PW), nwɛ̃³¹ (S.PW)

Note Two protoforms which have tone A (CK, S.Sgaw, Pwo) and tone D (NK, N.Sgaw) can be reconstructed.

However, the correspondence of vowels is problematical.

251. 'sew' PK: *cha *cha² (RB)

NK: cha²¹ (N.PO), cha²¹ (S.PO)

CK: cha^{4 5} (KN); che^{5 5} (KH); ɛ̃³³ (BW); chɔ̃³³ (KW)

SK: cha^{4 5} (N.SG), cha²¹ (S.SG); cha^{4 5} (N.PW); cha²¹ (S.PW)

252. 'sharp' PK: *thaŋ

NK: thaŋ³¹ (N.PO), thaŋ³¹ (S.PO)

CK: - (KN); thɛ̃³³ (KH); tha^{5 5} (BW); thɔ̃^{5 5} (KW)

SK: - (N.SG), - (S.SG); - (N.PW), - (S.PW)

253. 'sheep' PK: *so

NK: sɔ̃^{5 5} (N.PO), so³³ (S.PO)

CK: ʰɔ³³ (KN); - (KH); ʰu⁵⁵ (BW); so⁵⁵ (KW)

SK: so³³ (N.SG), su⁵⁵ (S.SG); - (N.PW); -ʰu⁵⁵ (S.PW)

Note It is noticeable that the Christian Karen seem to remember the word ‘sheep’ quite well. However, this word has tone A in CK and Sgaw while it has tone B in NK and Pwo.

254. ‘shivering’ PK: *t/k-hnɔ́ɔ́

NK: ta²¹nɔ́²¹ (N.PO), ta¹¹nɔ́²¹ (S.PO)

CK: ka²¹nau⁴⁵ (KN); tɛ¹¹nɛ⁵⁵ (KH); - (BW); kɔ¹¹no³³ (KW)

SK: tɛ¹¹nɔ́⁴⁵ (N.SG), kɛ¹¹no²¹ (S.SG); - (N.PW); - (S.PW)

255. ‘shoot (v.)’ PK: *khaɔ́ɔ́ *khaɔ² (RB)

NK: kha²¹ (N.PO), kha²¹ (S.PO)

CK: kha⁴⁵ (KN); khe⁵⁵ (KH); khɛ³³ (BW); kha³³ (KW)

SK: kha⁴⁵ (N.SG), kha²¹ (S.SG); kha⁴⁵ (N.PW); kha²¹ (S.PW)

256. ‘short, low’ PK: *phə́ɔ́ *phy² (RB)

NK: phə⁵⁵ (N.PO), phə³³ (S.PO)

CK: phəi¹¹ (KN); pho¹¹ (KH); - (BW); - phə¹¹ (KW)

SK: phi³¹ (N.SG), phi⁴⁵ (S.SG); phəi³³ (N.PW), phi⁵⁵ (S.PW)

257. ‘shoulder’ PK: *pleŋɔ́

NK: pleŋ³¹ (N.PO), pleŋ³¹ (S.PO)

CK: ble⁵³ (KN); pla³³ (KH); ple⁵⁵ (BW); ple⁵⁵ (KW)

SK: - (N.SG), - (S.SG); - (N.PW), -(S.PW)

258. ‘shove (with the foot)’ PK: *thwɪ́nɔ́

NK: thɪn³¹ (N.PO), thuɔm³¹ (S.PO)

CK: - (KN); - (KH); - (BW); tht⁵⁵ (KW)

SK: jɔ¹¹ thɔ³³ (N.SG), thu⁵⁵ (S.SG); thɔ̃³⁵ (N.PW), thɔ̃i⁵³ (S.PW)

259. ‘six’ PK: *kxrowɔ́ *xy⁴ (RB)

NK: su³¹ (N.PO), su³¹ (S.PO)

CK: chu⁵³ (KN); - (KH); xu⁵⁵ (BW); - (KW)

SK: xi³³ (N.SG), xi⁵⁵ (S.SG); xou³⁵ (N.PW), xou⁵³ (S.PW)

260. ‘sink’ PK: *blɛɔ́ɔ́ *phlɛɔ¹ (RB)

NK: plɛ̃^{4 5} (N.PO), plɛ̃^{4 5} (S.PO)

CK: plɛ̃²¹ (KN); plɔ³³ (KH); bli¹¹ca^{5 5} (BW); ple³³ (KW)

SK: phlɛ̃²¹ (N.SG), plɛ̃¹¹ (S.SG); phlẽ²¹ (N.PW); phlẽ^{4 5} (S.PW)

261. 'sleeping' PK: *hmejnej̃ *mi⁴ (RB)

NK: - (N.PO), - (S.PO)

CK: mai^{5 3}nai¹¹ (KN); - (KH); - (BW); mi^{5 5}ni¹¹ (KW)

SK: mi³³ne¹¹ (N.SG), mi^{5 5}ni³¹ (S.SG); - (N.PW), - (S.PW)

262. 'slippery' PK: *plẽ

NK: ple³¹ (N.PO), ple³¹ (S.PO)

CK: blai^{5 3} (KN); pli³³ (KH); pli^{5 5} (BW); pli^{5 5} (KW)

SK: ble³³ (N.SG), ble^{5 5} (S.SG); phli^{5 5} (N.PW), phli^{5 3} (S.PW)

263. 'smell (v.)' PK: *(h)n̩im̩

NK: num³³ (N.PO), nu̯m̩^{5 3} (S.PO)

CK: n̩i³³ (KN); - (KH); nu^{5 5} (BW); n̩i³³ (KW)

SK: n̩ə³³ (N.SG), n̩i^{5 5} (S.SG); ñ̩^{3 5} (N.PW), ñ̩i^{5 3} (S.PW)

Note Tone A3 in NK and CK suggests a voiceless onset (*hn-) but tone A12 in

Sgaw and tone A1 in Pwo indicate a voiced onset *n-.

264. 'snail (land~)' PK: *khlõ

NK: khlo³¹ (N.PO), - (S.PO)

CK: - (KN); - (KH); khlo³³ (BW); - (KW)

SK: khlo³¹̃ (N.SG), khlu^{4 5} (S.SG); khlu³³ (N.PW), khlu^{5 5} (S.PW)

265. 'snail (water~)' PK: *s - ŋwĩ

NK: ŋwi^{5 5} (N.PO), ŋwi³³ (S.PO)

CK: ŋwi¹¹ (KN); - (KH); θə^{5 5}mi³³ (BW); si¹¹mi¹¹ (KW)

SK: - (N.SG), - (S.SG); - (N.PW), - (S.PW)

266. 'snake' PK: *row̩ *ʏy¹ (RB)

NK: ru^{5 3} (N.PO), ru^{5 5} (S.PO)

CK: rəu¹¹ (KN); ro¹¹ (KH); Ru³³ (BW); ru¹¹ (KW)

SK: ʏi¹¹ (N.SG), ʏi³¹ (S.SG); ʏou¹¹ (N.PW), βou³³ (S.PW)

267. 'son-in-law' PK: *ma□□ *ma□¹ (RB)
 NK: ma□^{4 5} (N.PO), ma□^{4 5} (S.PO)
 CK: ma□²¹ (KN); me³³ (KH); ma¹¹ (BW); mɔ³³ (KW)
 SK: ma□²¹ (N.SG), ma¹¹ (S.SG); - (N.PW); ma□^{4 5} (S.PW)
268. 'sour' PK: *chjaN□ *cheN² (RB)
 NK: ɕja^{5 5} (N.PO), ɕja³³ (S.PO)
 CK: ɕi¹¹ (KN); cha¹¹ (KH); ∫ I³³ (BW); ɕi¹¹ (KW)
 SK: chi³¹~ (N.SG), chi^{4 5}, (S.SG); chã³³ (N.PW), chãi^{5 5} (S.PW)
Note The correspondence pattern of vowels in SK suggests *-ɛN
269. 'spider' PK: *gaŋ□
 NK: kuŋ³³kaŋ³³ (N.PO), jɔŋ³¹kaŋ^{5 3} (S.PO)
 CK: kaŋ³³ (KN); - (KH); ɡa¹¹-ɡu¹¹ (BW); kɔ³³ (KW)
 SK: - (N.SG), - (S.SG); pu¹¹khɛ̃^{5 5} (N.PW), khɔ̃³¹ (S.PW)
Note Bwe has an irregular tone, i.e. tone 11 (D3) instead of tone 33 (A3) as in the other Karenic languages.
270. 'spit (v.)' PK: *thu□ *tho² (RB)
 NK: - (N.PO), thu³¹ (S.PO)
 CK: thu^{5 3} (KN); thu³³ (KH); - (BW); - (KW)
 SK: thu□^{4 5} (N.SG), thu^{5 5} (S.SG); thu^{5 5} (N.PW), thu^{5 3} (S.PW)
Note N. Sgaw has an irregular tone, i.e. tone D, while the rest have tone A.
271. 'spleen' PK: *g- hmaN□
 NK: - (N.PO), - (S.PO)
 CK: - (KN); mɛ³³ (KH); ɡɔ³³ma³³ (BW); mɔ^{5 5} (KW)
 SK: kɔ¹¹mɔ³³ (N.SG), mɔ^{5 5} (S.SG); mɛ̃^{5 5} (N.PW), mɔ̃^{5 3} (S.PW)
272. 'split (bamboo strip)' PK: *pla□ *bla² (RB)
 NK: - (N.PO), - (S.PO)
 CK: - (KN); - (KH); plɛ³³ (BW); pla¹¹ (KW)
 SK: bla³¹~ (N.SG), bla^{4 5}, (S.SG); bla³³ (N.PW), bla^{5 5} (S.PW)
273. 'spray (v.)' PK: *pru□

NK: phru³¹ (N.PO), pru³¹ (S.PO)

CK: phru^{5 3} (KN); - (KH); phru³³ (BW); phru^{5 5} (KW)

SK: - (N.SG), pru^{5 5} (S.SG); phju^{5 5} (N.PW), pru³¹ (S.PW)

274. 'spread out (a mat)' PK: ***da** *da³ (RB)

NK: da³¹ (N.PO), da³¹ (S.PO)

CK: da^{5 3} (KN); de³³ (KH); dɛ^{5 5} (BW); da^{5 5} (KW)

SK: da³³ (N.SG), da^{5 5} (S.SG); da^{5 5} (N.PW), da³¹ (S.PW)

275. 'squeeze, massage' PK: ***im**

NK: cim^{5 5} (N.PO), cim³³ (S.PO)

CK: ci¹¹ (KN); ci¹¹ (KH); ci³³ (BW); ci¹¹ (KW)

SK: ci¹¹ (N.SG), ci³¹ (S.SG); chāi¹¹ (N.PW), chɛ̃i³³ (S.PW)

276. 'squirrel' PK: ***li** *li² (RB)

NK: li^{5 5} (N.PO), li³³ (S.PO)

CK: - (KN); - (KH); li¹¹ (BW); - (KW)

SK: li^{4 5} lu¹¹ (N.SG), li^{4 5} (S.SG); lei³³ (N.PW), lei^{5 5} (S.PW)

Note N.Sgaw has tone D while the other Karenic languages have tone B.

277. 'stand (v.)' PK: ***thəŋ**

NK: thəŋ^{5 5} (N.PO), thəŋ³³ (S.PO)

CK: thə¹¹ (KN); thə¹¹ (KH); tɔ̃^{5 5} thə³³ (BW); tə¹¹ (KW)

SK: thə³¹ (N.SG), thə^{4 5} (S.SG); thā³³ (N.PW), thə̃^{5 5} (S.PW)

278. 'star' PK: ***cha** *cha² (RB)

NK: cha^{5 5} (N.PO), cha³³ (S.PO)

CK: cha¹¹ (KN); che¹¹ (KH); tɛ³³ (BW); cha¹¹ (KW)

SK: cha³¹ (N.SG), cha^{4 5} (S.SG); sja³³ (N.PW), ɕa^{5 5} (S.PW)

279. 'steal (v.)' PK: ***how**

NK: - (N.PO), - (S.PO)

CK: hu¹¹ (KN); ho¹¹ (KH); a^{5 5} hu³³ (BW); hu¹¹ (KW)

SK: hī³¹ (N.SG), hī^{4 5} (S.SG); - (N.PW), - (S.PW)

280. 'stingy' PK: ***i**

- NK: ci^{5 5} (N.PO), ci^{5 3} (S.PO)
- CK: ci¹¹ (KN); - (KH); I³³ (BW); ci¹¹ (KW)
- SK: - (N.SG), - (S.SG); - (N.PW), - (S.PW)
281. 'stomach' PK: *g-phowꪀ *phoŋ⁴ (RB)
- NK: - (N.PO), - (S.PO)
- CK: phu^{5 3} (KN); pho³³ (KH); gə¹¹ phu^{5 5} (BW); phu^{5 5} (KW)
- SK: kə¹¹phĩ³³ (N.SG), phĩ^{5 5} (S.SG); phou^{3 5} (N.PW), phou^{5 3} (S.PW)
282. 'stone, rock' PK: *loŋꪀ *loŋ¹ (RB)
- NK: loŋ^{5 3} (N.PO), loŋ^{5 5} (S.PO)
- CK: ləu¹¹ (KN); lɔ¹¹ (KH); lo³³ (BW); lə¹¹ (KW)
- SK: lə¹¹ (N.SG), lĩ³¹ (S.SG); lău¹¹ (N.PW), lõu³³ (S.PW)
283. 'string, rope' PK: *blejꪀ *phli³ (RB)
- NK: phli³³ (N.PO), pli^{5 3} (S.PO)
- CK: pĩ³³ (KN); plə¹¹ (KH); bli³³ (BW); pli³³ (KW)
- SK: pli³³ (N.SG), pli³³ (S.SG); phlei^{5 5} (N.PW), phlei³¹ (S.PW)
284. 'suck' PK: *cwiꪀ
- NK: cuk²¹ (N.PO), cuꪀ²¹ (S.PO)
- CK: cwiꪀ^{4 5} (KN); cwi^{5 5} (KH); - (BW); ci³³ (KW)
- SK: sə¹¹ wiꪀ^{4 5} (N.SG), - (S.SG); - (N.PW); θə¹¹ waiꪀ³¹ (S.PW)
285. 'sun' PK: *mĩꪀ *my¹ (RB)
- NK: mĩ^{5 3} (N.PO), mĩ^{5 5} (S.PO)
- CK: mĩ¹¹ (KN); mɔ¹¹ (KH); mu³³ (BW); mĩ¹¹ (KW)
- SK: mĩ¹¹ (N.SG), mĩ³¹ (S.SG); mə¹¹ (N.PW), mĩ³³ (S.PW)
286. 'sunlight, hot' PK: *goꪀ *kho¹ (RB)
- NK: ja³¹ kħə^{5 3} (N.PO), ko^{5 5} (S.PO)
- CK: kau¹¹ (KN); ku¹¹ (KH); - (BW); ku¹¹ (KW)
- SK: tɑ¹¹kħə¹¹ (N.SG), ku³¹ (S.SG); khu¹¹ (N.PW), khu³³ (S.PW)
287. 'sway, swing' PK: *thuꪀ
- NK: - (N.PO), - (S.PO)

CK: - (KN); thu¹¹ (KH); - (BW); thu¹¹ (KW)

SK: tho³¹ (N.SG), thu^{4 5} (S.SG); thu³³ (N.PW), thu^{5 5} (S.PW)

288. 'swollen, swelling' PK: *hɲɔɲɔ *ɲo² (RB)

NK: jɔɲ²¹ (N.PO), jɔɲ²¹ (S.PO)

CK: ɲau^{4 5} (KN); jɛ^{5 5} (KH); jɔ³³ (BW); jo³³ (KW)

SK: ɲɔɲ^{4 5} (N.SG), ɲɔɲ²¹ (S.SG); jo^{4 5} (N.PW); ju²¹ (S.PW)

289. 'tail' PK: *meɲ *me¹ (RB)

NK: me^{5 3} (N.PO), me^{5 5} (S.PO)

CK: mai¹¹ (KN); mi¹¹ (KH); me³³ (BW); mi¹¹ (KW)

SK: mɛ¹¹ (N.SG), mɛ³¹ (S.SG); mɛ¹¹ (N.PW), mi³³ (S.PW)

Note Sgaw and N.Pwo have irregular vowel, i.e. ɛ instead of e.

290. 'tasteless, faded' PK: *plaɲ

NK: pla³¹ (N.PO), - (S.PO)

CK: bla^{5 3} (KN); - (KH); plɛ^{5 5} (BW); - (KW)

SK: bla³³ (N.SG), bla^{5 5} (S.SG); bla^{5 5} (N.PW), bla³¹ (S.PW)

291. 'ten' PK: *chejɲ *chi⁴ (RB)

NK: ɕi³¹ (N.PO), ɕi³¹ (S.PO)

CK: thi^{5 3} (KN); chɔ³³ (KH); ɕi^{5 5} (BW); ɕi^{5 5} (KW)

SK: chi³³ (N.SG), chi^{5 5} (S.SG); chei^{5 5} (N.PW), chei^{5 3} (S.PW)

292. 'tendon' PK: *thwaNɲ

NK: - (N.PO), - (S.PO)

CK: thu¹¹ (KN); thwa¹¹ (KH); - (BW); - (KW)

SK: thu³¹ (N.SG), thu^{4 5} (S.SG); thɔ³³ (N.PW), thɔ^{5 5} (S.PW)

293. 'termite' PK: *b-ɲwaNɲ

NK: pi^{5 5} ɲwa³¹ (N.PO), pi³³ ɲwa³¹ (S.PO)

CK: baɲ²¹ ɲu^{5 3} (KN); phu^{5 5} wa³³ (KH); - (BW); - (KW)

SK: pɔ¹¹ ɲu^{4 5} (N.SG), pɔ¹¹ ɲu²¹ (S.SG); paɲ²¹ ɲu^{5 5} (N.PW), pu³³ ɲu³¹ (S.PW)

294. 'termite (winged~)' PK: *biɲ *phin¹ (RB)

NK: phi^{5 3} (N.PO), pi^{5 5} (S.PO)

CK: pi¹¹ (KN); pi¹¹ (KH); pə³³i³³ (BW); pi¹¹i¹¹ (KW)

SK: pi¹¹ (N.SG), pi³¹ (S.SG); phai¹¹ (N.PW), phɛi³³ (S.PW)

Note Pwo vowels are irregular, i.e. ai and ɛi instead of ei.

295. 'termite mound' PK: *ploŋ

NK: phloŋ^{5 5} (N.PO), ploŋ³³ (S.PO)

CK: bləu¹¹ (KN); plɔ¹¹ (KH); plo³³ (BW); plə¹¹ (KW)

SK: blə³¹ (N.SG), - (S.SG); phlău¹¹ (N.PW), phlôu³³ (S.PW)

Note Pwo has tone B3 which suggests a voiced onset *bl-.

296. 'thick' PK: *ɗwɪn

NK: dɪn³¹ (N.PO), duəm³¹ (S.PO)

CK: dɪ^{5 3} (KN); do³³ (KH); - (BW); dɪ^{5 5} (KW)

SK: - (N.SG), - (S.SG); - (N.PW), - (S.PW)

297. 'thin (as of cloth)' PK: *ɓ() bun³ (RB)

NK: bə³¹ (N.PO), bə³¹ (S.PO)

CK: bɪ^{5 3} (KN); bɪ³³ (KH); - (BW); - (KW)

SK: bu³³ (N.SG), bu^{5 5} (S.SG); - (N.PW), baə³¹ (S.PW)

298. 'thousand' PK: *hreŋ, *g-thɔŋ thɔŋ⁴ (RB)

NK: reŋ³¹ (N.PO), reŋ³¹ (S.PO)

CK: re^{5 3} (KN); ri³³ (KH); gə³³thɔ^{5 5} (BW); re^{5 5}, tho³³ (KW)

SK: kə¹¹tho³³ (N.SG), thu^{5 5} (S.SG); - (N.PW), thɔ^{5 3} (S.PW)

Note *hreŋ = thousand (amount of money), *g-thɔŋ = thousand (people)

299. 'thread (a needle)' PK: *dow

NK: thu^{5 3} (N.PO), tu^{5 5} (S.PO)

CK: tu¹¹ (KN); to¹¹ (KH); du³³ (BW); tu¹¹ (KW)

SK: tɪ¹¹ (N.SG), tɪ³¹ (S.SG); thou¹¹ (N.PW), thou³³ (S.PW)

300. 'three' PK: *səm səm⁴ (RB)

NK: səm³¹ (N.PO), səm³¹ (S.PO)

CK: θə^{5 3} (KN); sɔ³³ (KH); θo^{5 5} (BW); sə^{5 5} (KW)

SK: sə³³ (N.SG), sə^{5 5} (S.SG); sǎ^{3 5} (N.PW), sǎ^{5 3} (S.PW)

301. 'tick' PK: *khejꠞ *khi² (RB)

NK: - (N.PO), - (S.PO)

CK: khi¹¹ (KN); khə¹¹ (KH); khi³³ (BW); khi¹¹ (KW)

SK: khi³¹ (N.SG), khi^{4 5} (S.SG); khej³³ (N.PW), khei^{5 5} (S.PW)

302. 'tie (a string)' PK: *cəŋꠞ *cəm³ (RB)

NK: cəŋ³¹ (N.PO), cəŋ³¹ (S.PO)

CK: cə^{5 3} (KN); cə³³ (KH); co^{5 5} (BW); cə^{5 5} (KW)

SK: cə³³ (N.SG), cə^{5 5} (S.SG); cā^{5 5} (N.PW), cə³¹ (S.PW)

303. 'tiger' PK: *keꠞ *khe⁴ (RB)

NK: ke³¹ (N.PO), ke³¹ (S.PO)

CK: kai^{5 3} (KN); khi³³ (KH); khi^{5 5} (BW); khi^{5 5} (KW)

SK: khe³³ji¹¹ 'lion' (N.SG), - (S.SG); khi^{5 5} (N.PW), khi^{5 3} (S.PW)

304. 'tilted' PK: *khiŋꠞ

NK: khiŋ^{5 5} (N.PO), khiŋ³³ (S.PO)

CK: khi^{5 3} (KN); khi³³ (KH); - (BW); khi^{5 5} (KW)

SK: khi³³ (N.SG), - (S.SG); khāi^{5 5} (N.PW), khēi^{5 3} (S.PW)

Note NK has the B tone while CK and SK have tone A.

305. 'toad' PK: *deꠞsowꠞ

NK: de^{5 5}su^{5 5} (N.PO), de³³su³³ (S.PO)

CK: dai¹¹θu¹¹ (KN); di¹¹so¹¹ (KH); di³³θu³³ (BW); di¹¹su¹¹ (KW)

SK: de³¹si³¹ (N.SG), di^{4 5}si^{4 5} (S.SG); di³³sou³³ (N.PW), di^{5 5}sou^{5 5} (S.PW)

306. 'tomorrow' PK: *~rɔꠞ

NK: mi^{5 3}rau³³ (N.PO), mi^{5 5}rau^{5 3} (S.PO)

CK: - (KN); pa^{5 5}ro¹¹ (KH); - (BW); mə¹¹ro³³ (KW)

SK: iə^{4 5}ɣɔ³³ (N.SG), ɣɔ³³ (S.SG); khə¹¹ɣɔ^{5 5} (N.PW), - (S.PW)

307. 'tongue' PK: *bleꠞ *phle³ (RB)

NK: phle³³ (N.PO), ple^{5 3} (S.PO)

CK: plai³³ (KN); pli¹¹ (KH); - bli¹¹ (BW); pli³³ (KW)

SK: ple³³ (N.SG), ple³³ (S.SG); pli^{5 5} (N.PW), pli³¹ (S.PW)

308. 'torn' PK: *dɛ́ꠞꠞ/*dwɛ́ꠞꠞ

NK: thɛ́ꠞꠞ^{4 5} (N.PO), tɛ́ꠞꠞ^{4 5} (S.PO)

CK: tuꠞ²¹ (KN); tĩ³³ (KH); thɛ³³ (BW); tu³³ (KW)

SK: tɛ́ꠞꠞ²¹ (N.SG), tɛ¹¹ (S.SG); theꠞ²¹ (N.PW); theꠞ^{4 5} (S.PW)

Note CK has irregular vowels.

309. 'tortoise' PK: *kliꠞꠞ *khliꠞ² (RB)

NK: kleꠞ²¹ (N.PO), kliꠞ²¹ (S.PO)

CK: khliĩ^{4 5} (KN); khli^{5 5} (KH); khli³³ (BW); khli³³ (KW)

SK: khliꠞ^{4 5} (N.SG), khlaiꠞ²¹ (S.SG); khlaiꠞ^{4 5} (N.PW); khlaiꠞ³¹ (S.PW)

310. 'tread' PK: *jamꠞ

NK: jam^{5 3} (N.PO), jam^{5 5} (S.PO)

CK: ꠞaŋ¹¹ (KN); ji¹¹ (KH); ja³³ (BW); - (KW)

SK: jɔ¹¹ (N.SG), jɔ³¹ (S.SG); jǎi¹¹ (N.PW), jǎi³³ (S.PW)

Note Pwo has irregular vowels, i.e. ǎi instead of ẽ (N.PW) and ɔ̃ (S.PN).

311. 'trunk (of a tree), firewood' PK: *məŋꠞ

NK: miŋ^{5 3} (N.PO), məŋ^{5 5} (S.PO)

CK: mə¹¹ (KN); mɔ¹¹ (KH); - (BW); mə¹¹ (KW)

SK: mi¹¹ (N.SG), - (S.SG); mǎ¹¹ (N.PW), - (S.PW)

312. 'trunk (of an elephant)' PK: *k-mlɔ́ꠞꠞ *mlɔ́ꠞ³ (RB)

NK: - (N.PO), mɔ́ꠞ^{5 3} (S.PO)

CK: mɔ³³ (KN); mo¹¹ (KH); -blɔ¹¹ (BW); - (KW)

SK: kə¹¹mlɔ³³ (N.SG), mlu³³ (S.SG); mlɔ́^{5 5} (N.PW), mlǔ³¹ (S.PW)

313. 'two' PK: *k-hnejꠞ *ni⁴ (RB)

NK: ni³¹ (N.PO), ni³¹ (S.PO)

CK: ŋĩ^{5 3} (KN); nə³³ (KH); ki^{5 5} (BW); ki^{5 5} (KW)

SK: khi³³ (N.SG), khi^{5 5} (S.SG); khi^{3 5} (N.PW), nei^{5 3} (S.PW)

314. 'urine' PK: *chejꠞ *chi² (RB)

NK: ɕi^{5 5} (N.PO), ɕi³³ (S.PO)

CK: - (KN); chə¹¹ (KH); ʃi³³ (BW); chi¹¹ (KW)

SK: chi³¹ (N.SG), chi^{4 5} (S.SG); chei³³ (N.PW), chei^{5 5} (S.PW)

315. 'vagina' PK: *lin (RB)

NK: lin^{5 3} (N.PO), lim^{5 5} ~ lin^{5 5} (S.PO)

CK: - (KN); li¹¹ (KH); - (BW); li¹¹ (KW)

SK: li¹¹ (N.SG), li³¹ (S.SG); lǎi¹¹ (N.PW), iǝi³³ (S.PW)

316. 'vomit (v.)' PK: *prɔ (RB)

NK: prɔ²¹ (N.PO), prɔ²¹ (S.PO)

CK: phrau^{4 5} (KN); phrɛ^{5 5} (KH); pɔ³³ (BW); pro³³ (KW)

SK: bɣɔ^{4 5} (N.SG), bɣo²¹ (S.SG); pjo^{4 5} (N.PW); pju²¹ (S.PW)

317. 'vulture' PK: *hlaŋ / k-da (RB)

NK: lɛŋ^{5 5} ta^{4 5} (N.PO), lɔŋ³³ ta^{4 5} (S.PO)

CK: - (KN); lɛ¹¹ ta³³ (KH); lɔ³³ da^{5 5} (BW); la¹¹ ta³³ (KW)

SK: lɔ³³ kə¹¹ ta²¹ (N.SG), lɔ³³ kə¹¹ ta¹¹ (S.SG); lǎ^{5 5} ka²¹ tha²¹ (N.PW); lǎ^{5 5} ka²¹ tha^{4 5} (S.PW)

318. 'waist' PK: *j-de (RB)

NK: ~de³¹ (N.PO), ~de³¹ (S.PO)

CK: - (KN); - (KH); ja³³ de¹¹ (BW); - (KW)

SK: jɔ¹¹ de³³ (N.SG), jɔ³¹ de^{5 5} (S.SG); ja¹¹ di^{5 5} (N.PW), jǎ³³ di³¹ (S.PW)

319. 'wart' PK: *thu (RB)

NK: - (N.PO), - (S.PO)

CK: - (KN); thi^{5 5} (KH); - (BW); thi³³ (KW)

SK: thu^{4 5} (N.SG), thɔu²¹ (S.SG); thau^{4 5} (N.PW); thɔu³¹ (S.PW)

320. 'warm' PK: *lam (RB)

NK: - (N.PO), lam^{5 3} (S.PO)

CK: la³³ (KN); le¹¹ (KH); lɛ³³ (BW); la³³ (KW)

SK: kə¹¹ lɔ³³ (N.SG), lɔ³³ (S.SG); la^{5 5} (N.PW), lə³¹ (S.PW)

Notelt is not certain if the *-a or *-aN rhyme should be reconstructed. S.Pa-O, Sgaw and S.Pwo suggest *-aN while the rest indicate *-a.

321. 'wash (the face)' PK: *bla (RB)

NK: phla^{5 3} (N.PO), pla^{5 5} (S.PO)

CK: pla¹¹ (KN); ple¹¹ (KH); bla³³ (BW); pla¹¹ (KW)

SK: pla¹¹ (N.SG), pla³¹ (S.SG); phla¹¹ (N.PW), phla³³ (S.PW)

Note Bwe has an irregular vowel, i.e. a instead of ɛ.

322. 'wash (clothes)' PK: *chrow 𑜋𑜧𑜨𑜩 *chɣy² (RB)

NK: pa 𑜋𑜧²¹ chu^{5 5} (N.PO), chu³³ (S.PO)

CK: chu¹¹ (KN); cho¹¹ (KH); - (BW); chu¹¹ (KW)

SK: chɣi³¹ 𑜋𑜧𑜨𑜩 (N.SG), chɣi^{4 5} 𑜋𑜧𑜨𑜩 (S.SG); - (N.PW), - (S.PW)

323. 'wasp' PK: *(dəŋ 𑜋𑜧) 𑜋𑜧de 𑜋𑜧

NK: ka 𑜋𑜧²¹ de³¹ (N.PO), - (S.PO)

CK: dau 𑜋𑜧^{4 5} dai^{5 3} (KN); di³³ 𑜋𑜧 (KH); - (BW); di^{5 5} 𑜋𑜧 (KW)

SK: tɔ¹¹ de³³ (N.SG), ti³¹ de^{5 5} (S.SG); thɔ¹¹ di^{5 5} (N.PW), thɔ³³ di³¹ (S.PW)

324. 'water' PK: *thej 𑜋𑜧𑜨𑜩 *thi⁴ (RB)

NK: thi³¹ (N.PO), thi³¹ (S.PO)

CK: thi^{5 3} (KN); thɔ³³ (KH); chi^{5 5} (BW); thi^{5 5} (KW)

SK: thi³³ (N.SG), thi^{5 5} (S.SG); thei^{3 5} (N.PW), thei^{5 3} (S.PW)

325. 'wear (a shoulder bag)' PK: *ch(w)ɛŋ 𑜋𑜧

NK: chɛŋ³¹ (N.PO), chɛŋ³¹ (S.PO)

CK: ɕi^{5 3} (KN); cha³³ (KH); - (BW); ɕi^{5 5} (KW)

SK: che³³ thɔ³¹ 𑜋𑜧 (N.SG), chi^{5 5} (S.SG); chɛi^{3 5} (N.PW), chwɛi^{5 3} (S.PW)

326. 'weave (cloth), loom' PK: *tha 𑜋𑜧𑜨𑜩 *tha² (RB)

NK: - (N.PO), cha³³ (S.PO)

CK: tha¹¹ (KN); the³³ (KH); - (BW); tha¹¹ (KW)

SK: tha³¹ 𑜋𑜧 (N.SG), tha^{4 5} (S.SG); tha³³ (N.PW), tha^{5 5} (S.PW)

327. 'weave (baskets)' PK: *thɛŋ 𑜋𑜧𑜨𑜩 *thɛŋ¹ (RB)

NK: thɛŋ³¹ (N.PO), thɛŋ³¹ (S.PO)

CK: thi^{5 3} (KN); tha³³ (KH); the^{5 5} (BW); thi^{5 5} (KW)

SK: the³³ (N.SG), thi^{5 5} (S.SG); thɛi^{3 5} (N.PW), thɛi^{5 3} (S.PW)

328. 'weevil' PK: *roŋ 𑜋𑜧

- NK: roŋ⁵³ (N.PO), roŋ⁵⁵ (S.PO)
- CK: ro¹¹ (KN); rɔ¹¹ (KH); - (BW); rə¹¹ (KW)
- SK: ʔə¹¹ (N.SG), ʔi³¹ (S.SG); ʔǎu¹¹ (N.PW), βǝu³³ (S.PW)
329. 'wet' PK: *cɔŋ *cɔ² (RB)
- NK: cau⁵⁵ (N.PO), cau³³ (S.PO)
- CK: cau¹¹ (KN); co¹¹ (KH); cɔ³³ (BW); co¹¹ (KW)
- SK: ba³¹ ɔ³¹ (N.SG), cɔ⁴⁵ (S.SG); cɔ³³ (N.PW), co⁵⁵ (S.PW)
330. 'white' PK: *ɓwaŋ *ɓwa³ (RB)
- NK: bwa³¹ (N.PO), bwa³¹ (S.PO)
- CK: bau⁵³ (KN); bu³³ (KH); ɓu⁵⁵ (BW); bu⁵⁵ (KW)
- SK: kwa³³ (N.SG), wa⁵⁵ (S.SG); - (N.PW), kwa³¹ (S.PW)
331. 'wide' PK: *lɛŋ *lɛ¹ (RB)
- NK: lai⁵³ (N.PO), lai⁵⁵ (S.PO)
- CK: lɛ¹¹ (KN); lɛ¹¹ (KH); la³³ lɛ³³ (BW); lɛ¹¹ (KW)
- SK: lɛ¹¹ (N.SG), lɛ³¹ (S.SG); lɛ¹¹ (N.PW), lɛ³³ (S.PW)
332. 'wife' PK: *hmaŋ *ma⁴ (RB)
- NK: ma³¹ (N.PO), ma³¹ (S.PO)
- CK: ma⁵³ (KN); me³³ (KH); mɛ⁵⁵ (BW); ma⁵⁵ (KW)
- SK: ma³³ (N.SG), ma⁵⁵ (S.SG); ma⁵⁵ (N.PW), ma⁵³ (S.PW)
333. 'wild' PK: *miŋ *min³ (RB)
- NK: mi³³ (N.PO), mi⁵³ (S.PO)
- CK: mi³³ (KN); mi¹¹ (KH); - (BW); mi³³ (KW)
- SK: mi³³ (N.SG), mi³³ (S.SG); mei⁵⁵ (N.PW), mei³¹ (S.PW)
334. 'win' PK: *niŋŋ
- NK: nəŋ³³ (N.PO), nəŋ⁵³ (S.PO)
- CK: ni³³ (KN); ni¹¹ (KH); - (BW); - (KW)
- SK: nə³³ (N.SG), ni³³ (S.SG); nǎ⁵⁵ (N.PW), nǎi³¹ (S.PW)
335. 'wing' PK: *deŋŋ *di² (RB)

NK: de²¹ (N.PO), di²¹ (S.PO)

CK: - (KN); da^{5 5} (KH); d³³e³³ (BW); de³³ (KW)

SK: di^{4 5}che³³ (N.SG), dai²¹ (S.SG); dai^{4 5} (N.PW); dai³¹ (S.PW)

336. 'wrong, mistake, error' PK: *g/k-hma² *ma² (RB)

NK: ma^{5 5} (N.PO), ma³³ (S.PO)

CK: - (KN); - (KH); g³³ma³³ (BW); ma¹¹ (KW)

SK: k³¹ma³¹ (N.SG), ma^{4 5} (S.SG); - (N.PW), ma^{5 5} (S.PW)

337. 'yam, potato' PK: *hnwε² *mwε² (RB)

NK: nwε^{5 5} (N.PO), nwε³³ (S.PO)

CK: nwε¹¹ (KN); - (KH); nwε³³ (BW); nε¹¹ (KW)

SK: nwε¹¹ (N.SG), nwε^{4 5} (S.SG); nε³³ (N.PW), nε^{5 5} (S.PW)

338. 'year' PK: *hneŋ² *nin² (RB)

NK: neŋ^{5 5} (N.PO), neŋ³³ (S.PO)

CK: ne¹¹ (KN); na¹¹ (KH); d³³e³³ (BW); de¹¹ (KW)

SK: ni³¹ (N.SG), ni^{4 5} (S.SG); năi³³ (N.PW), n^{5 5}i^{5 5} (S.PW)

339. 'yellow' PK: *baŋ³ *baŋ³ (RB)

NK: - (N.PO), - (S.PO)

CK: baŋ^{5 3} (KN); bε³³ (KH); - (BW); b^{5 5} (KW)

SK: b³³ (N.SG), b^{5 5} (S.SG); b^{5 5} (N.PW), b³¹ (S.PW)

340. 'yes' PK: *mwε²

NK: mwε^{5 3} (N.PO), mwε^{5 5} (S.PO)

CK: mwε¹¹ (KN); ma¹¹ (KH); - (BW); mε¹¹ (KW)

SK: me¹¹ (N.SG), mε³¹ (S.SG); mε¹¹ (N.PW), mwε³³ (S.PW)

341. 'young, soft, tender' PK: *ca¹ *cha¹ (RB)

NK: cha^{5 3} (N.PO), ca^{5 5} (S.PO)

CK: ca¹¹ (KN); ce¹¹ (KH); jε³³ (BW); ca¹¹ (KW)

SK: - (N.SG), ca³¹ (S.SG); - (N.PW), - (S.PW)

There are five more PK roots where I am not so sure how they should be reconstructed: *phow^A 'grandfather', *hl^B 'house', *ch()m^B 'mortar', *moɿ^B 'sky', *(h)me^B khlow^B 'smoke' (n.) and *k/g-li^A 'wind'.

7. Remarks on the internal and external classifications

Internal classification

With regard to the classification and subgrouping of Karenic languages, Manson (2009,2011) gives a good overview of what has been proposed by Jone (1961), Burling (1969), Kauffman (1993), Bradley (1997) and Shintani (2003). He comments: "...all lacked in comprehensiveness and/or have not been based on the comparative method..." and then, concludes his recent findings: "Based on a database of over 130 language varieties, Karen languages are divided into at least 20 low level clusters based on

shared phonological developments." (Manson, 2011:1)²²

We can see from the existing classifications that when different criteria are used different results are achieved. Since the previous classifications were based on geographical, lexical or phonological criteria, I would like to propose here a new look at the classification of Karenic languages based on the semantic development of six basic noun classifiers. The use of basic noun classifiers in the present reflects not only their modern worldview but also helps the reconstruction of the past one.

To investigate the use of noun classifiers in modern Karenic languages and to reconstruct the cognitive or conceptual system of the basic noun classification in Proto-Karen, a word list comprising 175 items was devised. The phrase frame used for eliciting noun classifiers during the interviews with native speakers was : N-NUM-(CLF), e.g. "chickens two..., sugarcanes two ..., plates two....". To obtain the cognate sets of basic noun classifiers, and the data on nine Karen language varieties, i.e. N. Pa-O, S.Pa-O, Kayan, Kayah, Kayaw, N. Sgaw, S.Sgaw, N. Pwo and S. Pwo, were analysed and compared. The results can be summarised as in (1) – (10)

(1) The cognate sets and the PK reconstructed forms :

Gloss	PK	N.Pa-O	S.Pa-O	Kayan	Kayah	Kayaw	N.Sgaw	S.Sgaw	N.Pwo	S.Pwo
[+human]	*bra ^A	phra ³³	pra ⁵³	pra ³³	phre ¹¹	ra ³³	pYa ³³	Ya ³³	Ya ⁵⁵	Ya ³¹
[+4-legged]	*ɿdə ^A dɿ ³¹	-	-	dəɿ ⁵³	do ³³	də ⁵⁵	dɿ ³³	dɿ ⁵⁵		dəɿ ⁵⁵
[+long]	*ɿbɔN ^A	-	-	bɔ ⁵³	bo ³³	bo ⁵⁵	bo ³³	bu ⁵⁵	bɔ̃ ⁵⁵	bɔ̃ ³¹
[+flat]	*ɿba ^B	ba ⁵⁵	ba ³³	ba ¹¹	be ¹¹	ba ¹¹	be ³¹ ~	bi ⁴⁵	bai ³³	bai ⁵⁵
[+round]	*phlot ^B	phlot ⁵⁵	phlot ³³	phəu ¹¹	phlɔ̃ ¹¹	phlɔ̃ ¹¹	phlɔ̃ ³¹ ~	phlɿ ⁴⁵	phlāu ³³	phlōu ⁵⁵
[+round + big]	*hma ^A	-	-	ma ⁵³	me ³³	ma ⁵⁵	-	-	-	-

(2) The PK conceptual system of basic noun classification based on the six noun classifiers is presented in Figure 5.

²² In his earlier version (2009:1), Manson says that Karen languages are divided into 7 low level clusters. This analysis is based on a database of over 100 language varieties. It is interesting to note that the increase in the number of languages (from 100-130) is also an increase of low level clusters (from 7-20).

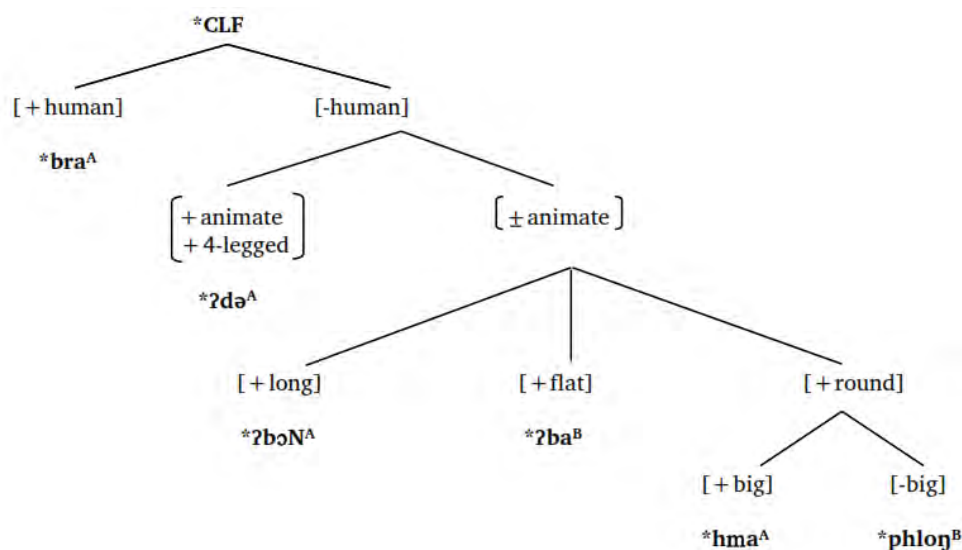


Figure 5 Conceptual system of basic noun classification in Proto-Karen

(3) The conceptual system as revealed by the use of basic noun classifiers shows how Karenic people construe their world of experience.

(4) Karen speakers primarily make a distinction between [+human] and [-human]. The [-human] category can be [+animate] or [±animate]. The [+animate] must be [+4-legged], while the [±animate] (if [+animate] must be [4-legged]) can be [+long],[+flat] or [+round]. The [+round] can be [+big] or [-big].

(5) From three to six basic noun classifiers were found in the nine modern Karenic language varieties studied, i.e. three in N. Pa-O and S.Pa-O; five in N.Sgaw, S.Sgaw, N. Pwo and S. Pwo; and six in Kayan, Kayah and Kayaw.

(6) "Shape" is more important than "size" in the Karen worldview.

(7) Based on the use of noun classifiers, the nine Karenic language varieties investigated can be classified into three groups : a. Pa-O; b. Kayan, Kayah and Kayaw; and c. Sgaw and Pwo.

(8) *hma^A [+round,+big] has been lost in Pa-O, Sgaw and Pwo. This means that the distinctive semantic values of *hma^A [+round,+big] and *phloŋ^B [+round, -big] have merged. Since *phloŋ^B has only been retained in Pa-O, Sgaw and Pwo, the meaning of *phloŋ^B has been changed to [+round], i.e. the distinction between the sizes [+big,-big] of round objects is no longer significant.

(9) The Pa-O seem to have a somewhat different conceptual system of noun classification,²³ for example, the classifier [+flat], which is ba⁵⁵ in N. Pa-O and ba³³ in

S. Pa-O, is used for most animals, except for snails and tortoises.

(10) Tai words, non-classifiers and classifiers, have been borrowed into N. Pa-O and used as classifiers, e.g. law³³ 'clf. for flutes', li⁵³ 'clf. for long objects' (from ri A4 'long'), phi³³ 'clf. for mats, blankets, etc.' (from phi in A1), phoŋ⁵⁵ 'flock, herd' (from fuuŋ A1) and so on. In the other Karenic language varieties studied, the classifier *ʔdə^A is only used for 4-legged animals, whereas the other animal species are classified according to their shape as the shapes of objects. The Pa-O use *ba³³ [+flat] as the classifier for all animal species without considering their shape, the same way as the use of the classifier tuəA3/toA3 in Tai languages.

²³ The Pa-O do not think that they are Karen but that they are a separate ethnic group. The Pa-O living and/or working in Thailand even have their own "National Day" which they celebrate annually. They can be found in Bangkok and in factories located in the vicinity of Bangkok. They regularly gather on Sundays for chatting and special events.

The above findings suggest the fact that the Pa-O, especially the northern group settling in the Shan State of Myanmar, have been in contact with Tai speaking peoples for a long time, long enough to have their worldview changed. A large number of Austroasiatic loans can also be found in S. Pa-O and Pwo as has been pointed out by Luce (1959) and Bauer (1992), due to their contact relationship with Mon, Palaung and Wa speakers. Most of the Pa-O men migrated to Thailand when they were quite young. The easiest way to get into the country was to be ordained as novices and come under the protection of famous and powerful Mon monks. A few years later, they left the temples to look for available jobs to earn their living.

In conclusion, a cognitive or conceptual system of noun classification can be used as additional evidence for language classification. The different conceptual system of noun classification of the Pa-O gives a good hint that the Pa-O separated from the other Karenic speaking groups quite earlier in time. Generally, my finding more or less supports the classifications proposed by Kauffman (1993) and Shintani (2003), even though their classifications are based on different criteria.

External classification

It has widely been accepted that Karenic is a branch of Tibeto-Burman. Weidert, one of the Tibeto-Burman specialists, states this quite explicitly in his monograph *Tibeto-Burman tonology* (1987). He provides a long list of both the present-day Karen roots and the reconstructed PK ones from many sources and compares them with the cognates, or in some cases lookalike words, found in the other branches of Tibeto-Burman with emphasis on the TB languages spoken in areas distant from where the Karen live, to avoid cases of contact relationship. About 168 etyma are given in the monograph, the detail of which are to be found on pages 334-367 of the monograph.

To support the view that Karenic is a branch of Tibeto-Burman, I compared my reconstructed PK roots (see the PK lexicon part) with the PTB roots reconstructed by Matisoff (2003) and/or Benedict (1972) which can be accessed from the STEDT database. The following is a summary of my findings with regard to the retentions and some sound changes from PTB to PK.

(1) Retention of PTB *a in PK

PTB *g/m/b-la-y > PK *ɭbla ^B 'arrow'	PTB *tsa-t, *tsa > PK *cha ^A 'ill'
PTB *g-p ^w a, *r-p ^w a, *r-wa > PK *hwa ^B 'bamboo'	PTB *m-twa > PK *tha ^A 'handspan'
PTB *ka > PK *kha ^B 'bitter'	PTB *r/g-na > PK *hna ^B 'listen to'
PTB *s-ka > *m-ka-y > PK *ka ^B 'chin, jaw'	PTB *pa > PK *pha ^A 'male, father'
PTB *ra > PK *bra ^A 'person, people'	PTB *s/g-la > PK *ɭla ^A 'moon, month'
PTB *ma-t > PK *ma ^B 'lost, disappear'	PTB *tsa > PK *sa ^B 'salt'
PTB *r/g-na, *s-na > PK *na ^B 'ear'	PTB *b-wa > PK *ɭbwa ^A 'white'
PTB *r-gya > PK *g-rja ^A 'hundred'	PTB *ŋya > PK *hɭa ^B 'flesh, meat'
PTB *pwa, *wa > PK *gwa ^A 'husband'	PTB *m-twa, *twa > PK *tha ^A 'handspan'

(2) Retention of PTB *-N in PK

PTB *d-wam > PK *tham ^A 'bear (animal)'	PTB *mi:n > PK *min ^A 'name'
PTB *ɔ̃am > PK *ɔ̃am ^B 'drink'	PTB *s-min > PK *hmin ^A 'ripe, cooked'
PTB *r/s-maŋ > PK *maŋ ^B 'dream'	PTB *dzyim > PK *chim ^A 'salty, sweet'
PTB *r-kaŋ > PK *khaŋ ^B 'leg'	PTB *m-nam > PK *(h)nim ^A 'smell'
PTB *kaŋ > PK *gaŋ ^A 'spider'	PTB *r-luŋ > PK *loŋ ^B 'stone, rock'
PTB *g-laŋ > PK *hlaŋ ^{A/B} 'hawk, vulture'	PTB *s-toŋ > PK *g-thɔŋ ^A 'thousand'
PTB *m-sin > PK *swin ^B 'liver'	PTB *g-sum > PK *səm ^A 'three'
PTB *s-luŋ > PK *hloŋ ^B 'maggot'	PTB *s-ni (:) ŋ > PK *hneŋ ^B 'year'

(3) PTB *vd. C- > PK *vl. C-/ *glottalised C-

PTB *d-wam > PK *tham 'bear (animal)'	PTB *d-yuk > PK *t-khroɽ ^D 'sambha deer'
PTB *dow > PK *tho ^B 'bird'	PTB *dzyim > PK *chim ^A 'salty, sweet'
PTB *mi > PK *hmī ^B 'femal, woman'	PTB *b-wa > PK *ɔ̃bwa ^A 'white'
PTB *mey > PK *hme ^B 'fire'	PTB *ga:p > PK *khaɽ ^D 'shoot (v.)'
PTB *ba:r > PK *phɔ ^A 'flower'	PTB *gwa-n > PK *chwɛŋ ^A 'wear (shoulder bag)'
PTB *lap > PK *hla ^B 'leaf'	PTB *dzo:p > PK *cwik ^D 'suck'
PTB *r-na > PK *hna ^B 'listen to'	PTB *g-nis > PK *k-hnej ^A 'two'
PTB *r-wa > PK *hwa ^B 'bamboo'	PTB *g-laŋ > PK *hlaŋ ^{A/B} 'hawk'
PTB *dwəy > PK *ɔ̃dej ^B 'egg'	PTB *d-ləy > PK *klej ^B 'crossbow'

(4) PTB *s-Vd. Sonorant > PK *vl. sonorant

PTB *s-min > PK *hmin ^A 'ripe, cooked'	PTB *s-luŋ > PK *hloŋ ^B 'maggot'
PTB *s-ley > PK *hli ^B 'squirrel'	PTB *s-mwəy > PK *hmej ^A 'sleep'
PTB *s-ni(:)ŋ > PK *hneŋ ^B 'year'	PTB *s-lay > PK *hlaŋ ^B 'lick'

(5) PTB *vl. unasp. Stop > PK *vl. asp. Stop

PTB *tak > PK *tha ^B 'weave (cloth)'	PTB *kroy > PK *khlo ^B 'snail'
---	---

PTB *r-kaŋ > PK *khaŋ^B 'leg'

PTB *pa > PK *pha^A 'male, father'

PTB *s-kew > PK *khwi^A 'nine'

PTB *d-kruk > PK *khrow^A 'six'

PTB *s-toŋ > PK *g-thɔN^A 'thousand'

PTB *krəw-t > PK *chrow^B 'wash
(clothes)'

PTB *pwa:y > PK *phe^A 'bran, chaff'

(6) PTB *-r > PK *-

PTB *s(y)ar > PK *gra^B 'insect'

PTB *swa:r > PK *chja^B 'sour'

PTB *s-kar > PK *cha^B 'star'

(7) PTB *-s > PK *-t

PTB *was > PK *kwat^D 'bee (Apis cerana)' PTB *s-nis > PK *ɲnwet^D 'seven'

PTB *rus > PK *khrwit^D 'bone'

(8) PTB *-stop = *-stop or > *-ɻ in PK

PTB *nuk > PK *nok^D 'brain'

PTB *k-r-pwat > PK *s-waɻ^D 'land
leech'

PTB *b-r-gyat & *b-g-ryat > PK *khr/grɔt^D

PTB *ɲap & *ga:p > PK *khaɻ^D 'shoot (v.)'

eight

PTB *smik & *s-myak > PK *mɛɻ^D 'eye, face' PTB *ma:k > PK *maɻ^D 'son-in-law'

PTB *l-tsyak > PK *thaɻ^D 'iron, needle'

PTB *dzo:p > PK *cwik^D 'suck'

PTB *m-sak > PK *saɻ^D 'itchy'

PTB *dyuk > PK *t-khroɻ^D 'sambha
deer'

(9) PTB *u > PK *o (vowel lowering)

PTB *nuk > PK *nok^D 'brain'

PTB *d-kruk > PK *khrow^A 'six'

PTB *d-yuk > PK *t-khro^{2D} 'sambha deer'

PTB *b-ru:l > PK *row^B 'snake'

PTB *ruŋ > PK *noŋ^A 'horn'

PTB *r-luŋ > PK *loŋ^B 'stone'

PTB *s-luŋ > PK *hloŋ^B 'maggot'

PTB *b-yuw > PK *jow^B 'mouse, rat'

(10) PTB *i_y > PK *i

PTB *s-hwi_y > PK *swi^B 'blood'

PTB *r-mi_y > PK *rwi^B 'root'

PTB *kwi_y > PK *thwi^B 'dog'

PTB *bwi_y > PK *wi^A 'bamboo rat'

PTB *s/g-li_y > PK *kli^A 'flea'

(11) PTB *e_y > PK *e

PTB *b-re_y > PK *bre^A 'buy'

PTB *re_y > PK *re^B 'rattan'

PTB *me_y > PK *hme^B 'fire'

PTB *d-key > PK *khe^A 'tiger'

PTB *r-ne_y-t > PK *ne^B 'get, obtain'

(12) PTB *a_y > PK *e

PTB *pwa:_y > PK *phe^A 'bran, chaff'

PTB *s/m-lay > PK *ble^A 'tongue'

PTB *r-may > PK *me^B 'tail'

(13) PTB *ə_y > PK *ej

PTB *d-lə_y > PK *klej^B 'crossbow'

PTB *b-lə_y > PK *lej^A 'grandchild'

PTB *mlə_y > PK *khlej^A 'boat'

PTB *r/s-mwə_y > PK *hmej^A 'lie

down, sleep'

PTB *twə_y 𑜄𑜤𑜂𑜫 *dwə_y > PK *ɽdej^B 'egg'

PTB *d-kə_y > PK *d-khej^A 'barking

deer'

PTB *nə_y > PK *nej^A 'sun, day'

8. Discussion

It is noticeable that the tones of some protoforms cannot be reconstructed because of their unusual patterns of correspondences; for example, the ones in simple words used in everyday life, 'child', 'pungent', 'take', 'pus', 'paddy', 'blow', 'hand', 'breathe', etc. To cope with this problem, the *B ✓ tone equivalent to Proto-Tai and Proto-Mienic tone *C was reconstructed by Haudricourt (1975), Manson (2009) and the others. Based on my field notes, the tones of the words 'paddy', 'blow', 'hand', and 'child' in modern Karenic languages are tone A in NK, tone D in CK and tone A/B in SK. Besides the addition of tone *B ✓, could there be another solution to the problem?

The *B ✓ tone does not seem to behave like the *C tone in the Proto-Tai and Proto-Mienic with which I am familiar. The *C tone has healthily existed in modern Tai and modern Mienic languages. It has never completely merged with tones A, B, C or D as found by Shintani (2003)²⁴ and Manson (2009) for Karen. How could the PK *B ✓ tone have had such a short life? Immediately after the Great Tone Split, it merged with tone *D, especially in Proto - CK.²⁵ Since there is no clear trace of its existence, except the unusual patterns of tone correspondences of some sets of words, would it make more sense to guess or hypothesise that once upon a time, the words 'paddy', 'blow', 'hand' and 'child' had tone *A? Actually, Pa-O and Sgaw, which can represent NK and SK languages, still have tone A in these words. As time went by, resulting from a heavy contact with the other tonal Tibeto-Burman and Tai languages spoken in the area, the original PK tone *A had independently developed in different sub-groups. In comparison with the Chinese, Tai and Hmong-Mien tones, the Karen tones are quite recent. Since PTB is non-tonal, its branches or clusters and the sub-groups within each branch have their own history of tonal developments.

With regard to Benedict's reconstruction of the glottalised voiceless stop series, i.e. *ʔp-, *ʔt-, *ʔc- and *ʔk- (Benedict, 1975: 12), I hesitate to accept his claim. Phonetically, these voiceless glottalised sounds are difficult to produce. "Voiceless" is a kind of phonation type. When producing voiceless sounds, the vocal folds are open. Thus, from a stage of complete closure of the vocal folds as when a glottal stop is made to an open stage of the glottis so that the airstream can easily pass through during the voiceless stage, is harder than producing voiced implosives or voiced glottalised plosives, such as ɓ/ɓb, ɗ/ɗd, etc. When a voiced sound is produced, the vocal folds come very close to each other to increase subglottal air-pressure; and after, to decrease the high air-pressure below, they start to blow open for the opening phase so that the air from the lungs can pass through, and then, the vocal folds begin to move towards each other again for the closing phase. The vibratory cycle or the glottal cycle ends as soon as the glottis is nearly closed. This mechanism causes the vocal folds to vibrate. As for the production of voiced implosives, the larynx with a closed glottis moves downwards, while the pulmonic egressive airstream causes a vibration of the glottis. In short, changing from a complete-closure stage to an open stage of the glottis, as when a glottalised voiceless stop (e.g. ʔp) is produced, is more difficult than changing from a complete-closure stage to a nearly-closed stage, as when a glottalised voiced stop (e.g. ɓb) is made.²⁶ More information on the stages of the glottis or phonation types is to be found in Ladefoged (1971), Catford (1977) and Laver (1994).

Benedict (1979:13) reconstructs eight PK roots with final *-s, i.e. *khrus 'bone', *ʔbūs 'carry with headstrap', *khwis 'comb', *ŋas 'five', *lis 'four', *ʔkus 'nine', *[hy]as 'pungent', *hnəs 'seven' when I reconstruct *-t in most of them, i.e. no. 31. *'bone'* *khrwit; 48. *'chili pepper, pungent'* *hat; 61. *'comb (v.)'* *khwit; 116. *'five'* *ŋjat; 124. *'four'* PNK: *lit, PCK: *hlwi, PSK: *lwi; 203. *'nine'* PNK: *kot, PCK, and PSK: *khwit and 250. *'seven'* PNK: *nwet, PCK and PSK: *nwe. In modern Karenic languages, the correspondences of tones in these words are not neat, for example, the NK varieties have tone D while the CK and SK ones tend to have tone A or B. If we agree with Benedict's reconstruction of the *-s, then, we have to say that the PK *-s > *-t in NK (Pa-O) which has been retained until today. This *-s was completely dropped in CK and SK and as a result, these roots have their own history of tonal developments.

The reconstruction of the *-s seems to help solve some problems of the irregular tone correspondences, even though there is no /-s/ in modern Karenic languages. Between the two solutions, i.e. the *B ✓ tone and the *-s, if I have to choose one, my preference will be the *-s. However, I wonder why the *-s has not been kept in any Pa-O or Pwo language

²⁴ Instead of tones A,B,C and D, Shintani uses tones 1,2,2 ✓ and 3.

²⁵ This is my interpretation based on Supplement 1 : the development of tone in Karen languages in Manson (2011).

²⁶ In the year 1984, I was invited by the Institute of Asian Studies in Copenhagen to analyse our Mlabri field notes. Søren Egerod, Jørgen Rischel and I did some fiber-optic experiments as I was producing the Thai voiceless plosives, p,t and k, which some linguists, e.g. Jimmy Harris, believe are glottalised sounds. My glottis could not be viewed because there was an interference of the epiglottis. We suspected that perhaps they were epiglottalised plosives.

varieties, since the speakers of Pa-O and Pwo have been in close contact with Austroasiatic speaking peoples both in the past and the present. It is known that Austroasiatic languages have final fricatives.

In conclusion, more fresh data collected in Myanmar by competent field linguists is badly needed before definite answers can be reached.

Acknowledgements

I would like to express my appreciation and gratitude to the Thailand Research Fund (TRF) for funding the "Karen Linguistics Project" from 2009-2012. Many thanks go to my Thai and international colleagues, especially James A. Matisoff, for their support and contribution of ideas to my work and also to my research assistants, Sujinat Jitwiriyanont, Karnthida Kerdpol and Siwaporn Tuanthaisong for their help in many different ways. Last but not least, I would like to thank my Karen language consultants and the local authorities in the many research sites for their kind co-operation, assistance and hospitality.

References

- Abramson, Arthur. 1995. Laryngeal timing in Karen obstruents. In *F. Bell-Berti and L.J. Raphael (Eds.), Producing Speech : Contemporary Issue. For Katherine Safford Harris*. New York :AIP Press.
- Bauer, Christian. 1992. Mon-Karen contacts. 27th International Conference of Sino-Tibetan Languages and Linguistics, Paris.
- Benedict, Paul K. 1972. *Sino-Tibetan: a conspectus*. New York: Cambridge University Press.
- Benedict, Paul K. 1979. Four forays into Karen linguistic history. *Linguistics of the Tibeto-Burman Area* 5.1:1-35
- Benedict, Paul K. 1983. Proto-Karen final stops. *Linguistics of the Tibeto-Burman Area* 7.2: 112-123.
- Bennett, J. Fraser. 1992. *Phonetics and phonology of Karen languages*. (ms.)
- Bradley, David. 1997. Tibeto-Berman languages and classification. In *Papers in Southeast Asian linguistics 14*, ed. David Bradley, 1-71. Canberra: Pacific Linguistics.
- Brunelle, M & Finkeldey, J (2011) Tone perception in Sgaw Karen, Proceedings of the 16th International Congress of Phonetic Sciences, 372-375.
- Burling, Robbins. 1969. Proto-Karen: A reanalysis. *Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics vol 1*. Ann Arbor, MI: Department of Linguistics, University of Michigan.
- Catford, John Cunnison. 1977. *Fundamental Problems in Phonetics*. Edinburgh: Edinburgh University Press.
- Enfield, N.J. 2005. Areal linguistics and mainland Southeast Asia. In *The annual review of anthropology*. 34 : 181-206
- Finkeldey, Joshua. 2011. Tone and phonation types in Sgaw Karen. In *Sgaw Karen papers presented to Nimrod Andrew, Marc Brunelle (eds.)*, 25-35. University of Ottawa. (online).
- Haudricourt, André-Georges. 1946. Restitution du Karen commun. *Bulletin de Société Linguistique de Paris* 42.1.103-111.
- Haudricourt, André-Georges. 1953. À propos de la restitution du Karen commun. *Bulletin de la Société Linguistique de Paris*. 49.1.129-132
- Haudricourt, André-Georges. 1975. Le système des tons du Karen commun. *Bulletin de la Société Linguistique de Paris*. 70:339-343.

Henderson, Eggénie J. A. 1979. Bwe Karen as a two-tone language? An inquiry into the Interrelation of pitch, tone and initial consonant. In *Southeast Asian linguistic studies* 3, ed.

Nguyen Dang Liem, 301-326. Canberra: Pacific Linguistics.

Henderson, Eggénie J. A. 1997. *Bwe Karen dictionary: with text and English-Karen word list (2 vols)*. London: School of Oriental and African Studies.

Intajamornrak, Chommanad. (in press). Variation and change of Phraw Pwo Karen vowels and tones induced by language contact with the Tai languages. *Manusya* 15.2.

Jitwiriyant, Sujinat. (in press). Ban Pa La-u Sgaw Karen tones: an analysis of semitones, quadratic trendlines and coefficients. *Manusya* 15.2.

Jones, Robert B. 1961. *Karen linguistic studies: Description, comparison and texts*. Berkeley: University of California Press.

Jones, Robert B. 1971. Some problems in reconstructing Proto-Karen tones. 4th International Conference of Sino-Tibetan Languages and Linguistics, Indiana.

Kato, Atsuhiko. 2009. A basic vocabulary of Htoklibang Pwo Karen with Hpa-an, Kyonbyaw, and Proto-Pwo Karen forms. *Asian and African languages and linguistics* no. 4, 169-217.

Kauffman, William G. 1993. *The great tones split and Central Karen*. University of North Dakota: M.A. Thesis.

Kerdpol, Karntida. (in press). Formant transitions as effective cues to differentiate the places of articulation of Ban Pa La-u Sgaw Karen nasals. *Manusya* 15.2.

Kullavanijaya, Pranee and Theraphan L-Thongkum. 2000. Linguistic criteria for determining Tai ethnic groups: case studies on central and Southwestern Tais. In *Proceedings of the International Conference on Tai Studies*, July 29-31, 1998, pp. 273-297. Bangkok: Institute of Language and Culture for Rural Development, Mahidol University.

Ladefoged, Peter. 1971. *Preliminaries to Linguistic Phonetics*. Chicago: The University of Chicago Press.

Langella, François. (2012). Polyfunctionality in Pwo Karen: the case of ʔàʔ- (<T-B pronominal prefix *ʔa-) in Tadao Miyamoto, Naoyuki Ono, Kingkarn Thepkanjana and Satoshi Uehara (eds.) *Typological Studies on Languages in Thailand and Japan*. Tokyo: Hituzi Syobo Publishing. 41-55.

LaPolla, Randy J. 2001. The role of migration and language contact in the development of the Sino-Tibetan language family. In *Areal diffusion and genetic inheritance: case studies in language change*, eds. R. M. W. Dixon & A. Y. Aikhenvald, 225-254. Oxford: Oxford University Press.

Laver, John. 1994. *Principles of Phonetics*. Cambridge: Cambridge University Press.

Lewis, M. Paul (ed.). 2009. *Ethnologue: Languages of the World*, Sixteenth edition. Dallas, Tex.: SIL International. (online)

L-Thongkum, Theraphan. 1991. A preliminary reconstruction of Proto-Lakkja (Cha Shan Yao). *Mon-Khmer Studies* 20, 57-89.

L-Thongkum, Theraphan. 1993. A view on Proto-Mjuenic (Proto-Yao). *Mon-Khmer Studies* 22, 163-230.

L-Thongkum, Theraphan. 1997. Implications of the retention of proto-voiced plosives and fricatives in the Dai Tho language of Yunnan province for a theory of tonal development and Tai language classification. In *Jerold A. Edmondson and David B. Solnit* (Eds.),

Publication 124, 193-221. Comparative Kradai: The Tai branch, J. Dallas: SIL and U. of Texas at Arlington. Publications in Linguistics.

L-Thongkum, Theraphan. (2011). Numerals and classifiers in Modern Karenic languages and Proto-Karen. Paper presented at the 44th International on Sino-Tibetan Languages and

Linguistics, organised by Central Institute of Indian Languages, Mysore, India.

L-Thongkum, Theraphan. (2012). Proto-Karen (*k-rjaŋ^h) fauna. Paper presented at SEALS 22, co-organised by French laboratories and institutions, Agay, France.

- Luce, Gordon H. 1959. Introduction to the comparative study of Karen languages. *Journal of the Burma research society* 42.1:1-18.
- Luce, Gordon H. 1985. *Phases of pre-Pagan Burma: Languages and history*. Oxford: Oxford University Press (2 volumes).
- Manson, Ken. 2002. *Karenic language relationships: a lexical and phonological analysis*. Payap University.
- Manson, Ken. 2004. Bibliography of Karen linguistics. Department of Linguistics, Graduate School, Payap University.
- Manson, Ken. 2009. Prolegomena to reconstructing Proto-Karen. *LaTrobe working papers in linguistics* 12.
- Manson, Ken. 2010. Bibliography of Karen linguistics. Payap University. (online)
- Manson, Ken. 2011. The subgrouping of Karen. Paper presented at SEALS 21, Bangkok. (online)
- Matisoff, James A. 1973. Tonogenesis in Southeast Asia. In *Consonant types and tone*, ed. Larry Hyman, 73-95. Los Angeles: University of Southern California.
- Matisoff, James A. 1983. Linguistic diversity and language contact. In *Highlanders of Thailand*, eds. J. MacKinnon & Wanat Bhruksasri, 56-86. Oxford: Oxford University Press.
- Matisoff, James A. 1991. Sino-Tibetan linguistics: present state and future prospects, *Annual review of anthropology* 20: 469-504.
- Matisoff, James A. 2003. *Handbook of Proto-Tibetan-Burman: System and philosophy of Sino-Tibetan reconstruction*. Berkeley: University of California Press.
- Mazaudon, Martine. 1977. Tibeto-Burman tonogenetics. *Linguistics of the Tibeto-Burman area* 3.2:1-123.
- Pittayaporn, Pittayawat. 2013. Sonorancy of the dorsal rhotics /ʁ/ in Sgaw Karen. Unpublished manuscript.
- Ratanakul, Suriya. 1986. *Thai-Sgaw Karen dictionary*. Institute of Language and Culture for Rural Development, Mahidol University.
- Schliesinger, Joachim. 2000. *Ethnic groups of Thailand : non-Tai-speaking peoples*. Bangkok, Thailand : White Lotus Press.
- Shintani, Tadahiko. 2003. Classification of Brakaloungic (Karenic) languages, in relation to their tonal evolution. In *Proceedings of the symposium cross-linguistic studies of tonal phenomena: historical development, phonetics of tone, and descriptive studies*, ed. Shigeki Kaji, 37-54. Tokyo: Tokyo University of Foreign Studies.
- Solnit, David B. 2001. Another look at Proto-Karen. Paper presented at the 34th International Conference of Sino-Tibetan languages and linguistics, Kunming.
- Teeranon, Phanintra. (in press). Initial consonant voicing perturbation of fundamental frequency of clear vowels: a controversial case from Ban Doi Pwo Karen. *Manusya* 15.2.
- Teerarajanarat, Sirivilai. (in press). Using GIS for exploring Karen settlements: a case of the Western and Northern Thailand along the Thai-Burmese border. *Manusya* 15.2.
- Thomas, Dominique and Leonardo Alves-Soares (2011). On the three-way plosive contrast of Sgaw Karen. In *Sgaw Karen papers presented to Nimrod Andrew, Marc Brunelle, (eds.)*, 5-
24. University of Ottawa. (online)
- Van Driem, George. 2001. *Languages of the Himalayas: An ethnolinguistic handbook of the greater Himalayan region*. Leiden: Brill.
- Weidert, Alfons. 1987. *Tibeto- Burman tonology*. Amsterdam: John Benjamins.

ผศ.ดร. พณินทรา ธีรานนท์

Initial consonant voicing perturbation of the fundamental

frequency of oral vowels and nasal vowels:

A controversial case from Ban Doi Pwo Karen

Initial consonant voicing perturbation of the fundamental frequency of oral vowels and nasal vowels:

A controversial case from Ban Doi Pwo Karen

Phanintra Teeranon

School of Liberal Arts, Mae Fah Luang University

0. Abstract

This paper aims to analyze the acoustic characteristics of initial consonant voicing perturbation of the fundamental frequency of oral vowels and nasal vowels of Ban Doi Pwo Karen. Three age groups of informants were selected: over-sixty years old (>60), middle aged (35-45), and under-twenty years old (<20). The acoustic analysis method was employed to analyze the mean vowel duration (msec), mean vowel amplitude (dB), and mean vowel fundamental frequency (Hz). The results showed that voiceless initial consonants tend to cause a lower fundamental frequency than that of the voiced initial consonants. This has excited controversy concerning the tonogenesis theory of initial voicing perturbation on vowels. However, it was later found that the vowels following by voiceless initial consonants were breathy and it was the voice register of vowels that caused the low fundamental frequency values. In contrast to other studies, the nasal vowels were not always higher in fundamental frequency when compared to oral vowels. In all age groups, nasal vowels following either voiceless or voiced consonants were found to be higher in fundamental frequency than oral vowels, except in the younger age group where the fundamental frequency of nasal vowels following voiced consonants was lower than that of the oral vowels following voiced consonants.

1. Introduction

According to linguistic classification (Grimes, 1988), the languages spoken in Southeast Asia (SEA) can be grouped into five language families: Tai-Kadai (e.g. Thai); Miao-Yao (e.g. Miao, White Hmong); Sino-Tibetan (e.g. Chinese, Burmese); Austroasiatic or Mon-Khmer (e.g. Vietnamese, Mon, Khmer); and Austronesian (e.g. Malay, Javanese). The languages in these five families contain 3 main variations, and are classified as the following: tonal languages (e.g. Thai, Chinese); non-tonal languages (e.g. Malay, Indonesian); and register languages (e.g. Chong, Mon). The distinctive feature of Tai-Kadai, Miao-Yao and Sino-Tibetan is tone. In contrast, most of the Mon-Khmer and Austronesian languages are

non-tonal languages. At present, however, a number of non-tonal languages and register languages, such as some of the languages of New Caledonia in the Austronesian family and some Khmer dialects in the Mon-Khmer family, have been found to have changed from a stage of tonelessness to that of a tonal language. Some of the present tonal languages have also been found to have shifted in their number of tones, that is to say, increasing in the number of tones or decreasing in the number of tones. This process of transformation is widely discussed as tonal evolution or the tonogenesis theory (Matisoff, 1973).

In the year 1954, Haudricourt (1954) claimed that the Vietnamese language, which is tonal, had derived its tones from its non-tonal ancestral language. Therefore, he proposed that Vietnamese should be classified as a language in the Mon-Khmer family, not as it had previously classified in Sino-Tibetan. He argued that the main influences causing tone birth in the Vietnamese language had been the loss of initial and final consonants (See Table 1). Since then, Haudricourt's proposal has been broadly verified by linguists (Matisoff, 1973; Maddieson, 1984; L-Thongkum, 1988; L-Thongkum et al., 2007; Thurgood, 2007, etc.). Following this, others have proposed new influences that give birth to tones, for instance, the influence of the high-low dimension of vowels. Up to now, linguists have discovered the various influences of consonants and vowel quality on the tonal evolution of the five Southeast Asian language families. These influences are; 1. Internal factors, (i) monosyllabification (ii) initial consonants (iii) vowel quality (iv) final consonants, and 2. external factors. Both internal and external factors have been claimed seemingly, as universal phenomena.

Data from published studies proves SEA languages to have developed tones by the similar influences of consonants and vowels. The most established factor causing tone birth and development is initial consonant voicing; the voiced consonant changing to a voiceless consonant causing a low pitch or low tone in the later state. This is accepted to be a universal phenomenon by linguists studying SEA languages. However, an uncommon phenomenon has emerged from a dialect of Cantonese, T'ientsin, in which a high tone was found to have developed from an initial voiced consonant. This was reported in 1977 (Li, 1977). In the Pwo Karen of Ban Doi, Chiangrai, the identical phenomenon seems to be appearing in some words. It has been observed that an initial voiceless consonant has initiated a lower tone, while a voiced initial consonant has initiated a higher tone. This corresponds to what had been found in T'ientsin. Therefore, the objective of this study is to attest to the fundamental frequency behaviour of vowels influenced by initial consonant voicing in Ban Doi Pwo Karen. This study intends to investigate pitch behaviour in the vowels following voiceless

and voiced initial consonants, i.e to investigate whether or not the voiceless consonant induces a high pitch, while the voiced consonant induces a low pitch.

2. Literature review

Tonogenesis is the study of tone evolution or tone development in languages, especially in Southeast Asian languages. Prior to linguistics publishing linguistic results, many comparativists tried to reconstruct a proto-language to demonstrate that languages may form more contrastive pitches—that is, tones—through changes in the initial and final consonant features in their parent languages. This concept was first introduced by Przyluski (1924). A classical model of tonal evolution has been proven in the Vietnamese language by Haudricourt (1954) and has later been clarified by Matisoff (1973).

Table 1 Tonal development of the Vietnamese language in the early 6th century (non-tonal)

		Final consonants		
		/*-Ø/, /*-N/	/*-h/	/*- /
Initial consonants	Voiceless /*p-/	pa, paN	pah	pa
	Voiced /*b-/	ba, baN	bah	ba

(adapted from Haudricourt, 1954 and Matisoff, 1973)

According to Table 1, the Vietnamese language in the early 6th century was a non-tonal language. Its syllables during this period were of (i) the open type, ending in a vowel /*-Ø/ or a nasal consonant /*-N/, and (ii) the closed type, ending in a fricative /*-h/ or a glottal stop /*-/. There were, in addition, two types of voicing distinction for initial consonants: voiceless /p-/ and voiced /b-/. The loss of final consonants, /*-N/, /*-h/, /*- /, caused phonologically distinctive pitches or tones to emerge, namely level, falling and rising tones, as shown in Table 2.

Table 2 Three tones of the Vietnamese language around the 12th century

		Tones		
		Level	Falling	Rising
Initial consonants	Voiceless /*p-/	pa	pa	pa
	Voiced /*b-/	ba	ba	ba

(adapted from Haudricourt, 1954 and Matisoff, 1973)

By the end of the 12th century, the number of tones had increased to six through the loss of initial consonant voicing, as shown in Table 3. As the voiced initial consonants lost their voicing, /*b-/ > /*p-/, low and high tones emerged to avoid homophones. Tones emerged to replace consonant voicing and to differentiate word meanings. That is to say, when /*p-/ became /p/, a high tone resulted, and when /*b-/ became /p/, a low tone resulted.

Table 3 Six tones in the Vietnamese language by the end of the 12th century

Tones	Level	Falling	Rising
/*p-/ > /*p-/ > High tone	pa	pa	pa
/*b-/ > /*p-/ > Low tone	pa	pa	pa

(adapted from Haudricourt, 1954 and Matisoff, 1973)

The study of tonal evolution or tonogenesis has developed into a long progression, following the statement of the theory of consonant effect on the development of contrastive tones in the Vietnamese language.

Recently, Thurgood (2007) revised the model of Haudricourt (1954); explaining that pitch assignment occurs not from the consonant-based account but from the laryngeal-based or voice quality. This is due to instability in assigning pitch to the syllable from initial consonant types and final consonants.

In the case of initial consonants, it is argued that it is not Haudricourt's initial consonant types but the voice quality, breathy voice, tense voice, and clear voice, that assigns pitch to the syllable. Mostly, a breathy voice induces low pitch, a tense voice high pitch and a clear voice mid pitch.

Considering final consonants, laryngeal features and final consonants are considered to give birth to pitch contour. Glottal fricatives, /*-h/ or glottal stops, /*- / alone are not the

primary cause of falling or rising pitch but abrupt glottal stops [- abrupt] and [-h nonbreathy] relate to the raising of pitch or to a high pitch and a creaky glottal stop [- creaky] and a breathy final -h [- breathy] relate to the lowering of pitch or a low pitch (Thurgood, 2007).

In short, linguists have become aware of many universal phenomena which are thought to give birth to tones: 1. internal factors, (i) monosyllabicization (ii) initial consonants (iii) vowel quality (iv) final consonants, and 2. external factors.

2.1 Internal factors

2.1.1 Monosyllabicization

According to Matisoff (1973), languages whose basic syllable structures are in monosyllabic form are likely to develop tones. Most of the words in a language consist of stressed and unstressed syllables. An unstressed syllable tends to drop more than a stressed one (Thach, 1999). This may possibly result in the creation of monosyllabic words and even bring homophones into a language. However, languages create tones to distinguish word meanings and also to avoid having homophones. Several studies have regarded monosyllabicization as the source of tone birth in the five language families of SEA (Henderson, 1982; L-Thongkum, 1984; Matisoff, 1973; Teeranon, 2008; Thach, 1999; Thurgood, 1999). Table 4 examines monosyllabicization in the Kiengiang Dialect of the Khmer language. In the case of the unstressed syllable, the first syllable of each word tends to drop and may result in monosyllabic words.

Table 4 Monosyllabicization in the Kiengiang Dialect of the Khmer language

Standard Khmer		Kiengiang Dialect	Meaning
ci cra:m	>	cra m	‘to chop’
c ngkra:n	>	kra n	‘kitchen’

(adapted from Thach, 1999, 84)

Linguists such as Abramson (2004), Brunelle (2005), Matisoff (1973), and Michaud (2012) have also revealed that monosyllabicization results in tonal contrast across Southeast Asian language families.

2.1.2 Initial consonants

One of the most documented studies of tonal evolution concerns the influence of initial consonants on the development of tones. An illustration of the influence of initial consonants on the development of tones in SEA can be found in Table 5. Symbol ‘p’ represents a voiceless initial consonant, ‘b’ represents a voiced initial consonant, ‘v’ represents a vowel, / ` / is a low tone, and / ´ / is a high tone.

Table 5 A model of the influence of initial consonants on the development of tone

Proto-language, Non-tonal language		Non-tonal language		Tonal language
*pv	>	pv	>	p´
*bv	>	bv	>	p`

Table 5 shows a loss of initial contrast and refers to the original initial consonants in the ancestral language (Proto-language). When the voicing distinction of the initial consonant in the Proto-language is lost, a higher tone / ´ / seemingly appears in words formerly beginning with Proto-voiceless initial consonants as compared with words beginning with proto-voiced initial consonants accepting the result of the lower tone / ` /.

Matisoff (1973) often refers to examples from the Tai-Kadai language family. One of these, the Thai language, is depicted in Table 6:

Table 6 The influence of initial consonants on tones in the Thai language.

Proto-Tai		Ancient Thai		Modern Thai	Meaning
*hmaa	>	hmaa	>	ma a	‘a dog’
*maa	>	maa	>	maa	‘to come’

(adapted from Robinson, 1994: 16)

Another interesting case of the influence of initial consonants on tone birth is found in the Khmer language. The Phnom Penh Khmer was discovered for the loss of its cluster /-r-/ in /Cr-/, which also causes a falling-rising tone -- as in Table 7:

Table 7 The lost of cluster in the Phnom Penh Khmer

Standard Khmer	Phnom Penh Khmer	Meaning
pram	p eam	‘five’
triw	t ĩw	‘correct’

(adapted from Guion and Wayland, 2004: 1)

Hence, the influence of initial consonants causing tone birth has been proven in the five language families of SEA (Diffloth, 1980; Erickson, 1975; Haudricourt, 1954; Li, 1966; L-Thongkum, 1984; Maddieson, 1984; Maran, 1973; Sun, 2003; Teeranon, 2008; Thurgood, 1999).

2.1.3 Vowel quality

2.1.3.1 Voice register of vowels

Register is phonologically defined as a prosodic system (Henderson, 1952), a contrast between phonation types. For example, the modal voice vowel and the breathy voice vowel in the Suai language spoken in Thailand:

Modal voice vowel		Breathy voice vowel	
/lu /	‘to howl’	/lu /	‘thigh; the lap’
/lu m/	‘a mouthful’	/lu m/	‘to gobble chunks of food’

(adapted from Abramson, L-Thongkum, Nye, 2004: 148)

Phonetically, register is called phonation type; it describes a cluster of laryngeal and supralaryngeal activities, one property of which may be dominant and the rest secondary. The term register is generally understood to mean a ‘register complex’, one property of which may be dominant and the rest secondary. The complex of phonetic characteristics typically includes such features as phonation type, pitch, vowel quality, vowel length, loudness and perhaps others (Abramson, Thongkum, and Nye 2004: 147).

The register of vowels is regarded as the transcending stage (or middle point) for SEA non-tonal languages changing to tonal ones. Apparently, an example is derived from the Mon-Khmer language family dialects -- Khmu (Premssirat, 2003).

Table 8 The loss and the replacement of register distinction in Khmu dialects

Non-tonal language	Register language	Tonal language		Meaning
Eastern Khmu	Western Khmu Dialect 1	Western Khmu Dialect 2	Western Khmu Dialect 3	
bu c	pu c	p ù c	pù c	‘fermented rice’
bok	po k	p òk	pòk	‘to cut the trees’
bu m	pu m	p ùm	pùm	‘to chew’
aŋ	ca ŋ	c àŋ	càŋ	‘to weigh’
la ŋ	kla ŋ	k là ŋ	klâ ŋ	‘stone’

(adapted from Premsrirat, 2003: 25)

According to Table 8, the Eastern Khmu dialect retains its voicing distinction of initial consonants, whereas dialect 1 of Western Khmu shows the register distinction of vowels. However, dialect 2 and dialect 3 of Western Khmu drop both initial voicing distinction and the register of vowels. With the loss of these two features, dialect 2 has now substituted low tones.

Until recently, only two language families -- Mon-Khmer and Austronesian -- have reportedly given birth to tones from the influence of vowel registers (Edmondson and Gregerson, 1993; Henderson, 1982; L-Thongkum, 1988, 1989, 1990; Thurgood, 1993, 1996, 1999).

2.1.3.2 High and low vowel dimension

In Southeast Asian languages, high vowels, e.g. /i/, have a higher pitch than low vowels, e.g. /a/, in non-tonal languages such as Malagasy (Whalen and Levitt, 1995), an Austronesian language; in register languages such as Paroak (Watkins, 2002), a Mon-Khmer language; and in tonal languages such as Thai (Mohr, 1971; Zee, 1980; Bunphan et al., 1982; Svantesson, 1988; Rose, 1997). All of the findings conclude that high vowels cause a higher fundamental frequency than that of low vowels.

However, Lehiste (1970) and Teeranon (2008) have revealed that the influence of initial consonants on the following vowels is much greater than the effect of the voice register of the vowel itself. It can be inferred that after the monosyllabicization process

or the change of syllabic structure to the monosyllable (Matisoff 1973), tones occurred as the voicing states of initial consonants influenced the pitch of the following vowels, not higher and lower pitches within the vowels.

2.1.3.3 Vowel length

Hu, a language of the Mon-Khmer family, has been confirmed to have 2 tones, i.e. high tone vs. low tone which developed from the loss of vowel length. According to Table 9, the ancestral language of Hu (Proto-Palaungic) had both short vowels and long vowels. In the middle stage of Hu tonal evolution, the short vowels remained short but the long vowels began to lose their length. Following this, the new short vowels merged with the original short vowels. To compensate for the loss of vowel length, two tones have developed in Hu as shown in Table 9. (Diffloth, 1980; Svantessen, 1991).

Table 9 The loss of vowel length in the Hu language

Proto-Palaungic				Hu	Meaning
*yam	>	*yam	>	yám	‘to die’
*yaam	>	*yam	>	yàm	‘to cry’

Over the past 20 years, the birth of tones in many minority languages of the Mon-Khmer (Diffloth, 1980; Svantessen, 1991) and Sino-Tibetan language families (Sun, 2003) have been found to indicate this type of influence. Recently, SEA languages such as Tai-Kadai, Miao-Yao, and other unmentioned categories, have developed tones through the loss of vowel length (L-Thongkum et al., 2007).

2.1.4 Final consonants

The influence of final consonants on the development of tones is widely observed in vowels preceded by a glottal stop consonant /*- /, including those headed by a glottal fricative consonant /*-h/. The loss of /*- / tends to develop into a higher tone than those preceded by /*-h/. This phenomenon is often found in Middle Chinese (Hombert et al, 1979; Sagart, 1993). Jingpaw spoken in Myanmar (Maran, 1973) and Usat which is an Austronesian language spoken in Hainan (Thurgood, 1996), and so forth.

However, not all languages conform to this hypothesis, for instance, Eastern Chamic languages (Phu Van Han et al., 1992; Thurgood, 1993) and some dialects of Tibetan (Mazaudon 1977; Sun, 2003), since the loss of the final /- / causes a low tone. As stated in Thurgood (2007) laryngeal features and final consonants are considered to give birth to pitch contour, so it is not only the final consonants that play a major role in pitch but also the phonation types of the vowels. This corresponds to the results that have been found in L-Thongkum (1989, 1990) where it was pinpointed that the voice register of vowels plays a major role in causing tone birth.

Above all, the influence of final consonants, proven and verified, appears universally as a common phenomenon in SEA languages (Lohde, 2003; Thavisak, 2001; Watkins, 2002).

2.2 External factors

The external factor, or language contact, is the only factor unlikely to be attested to within the frame of the tonogenesis theory.

According to Matisoff (1973), the Chinese language is the only true tonal language in the whole of SEA. The Chinese language, through its own internal factors, has developed its tones. In other words, tones did not emerge out of language contact. On the other hand, Tai-Kadai, Miao-Yao and the Austronesian language families have been claimed to have developed tones by interweaving parts from internal factors with language contact.

A number of observations have also been made by Thurgood (1996, 1999). In Western Cham, an Austronesian Language, tones have developed following influence from the Khmer language and Eastern Cham or Phan Rang Cham is transforming into a tonal language through the influence of Vietnamese. Additionally, the Chamic language which is known as Usat, a language of Hainan, has given birth to tones with the presence of the Chinese language.

3. Methodology

3.1 Context and language

This research was conducted at Ban Doi (Doi Village), Tambon Chokchai (Chokchai Sub-district), Amphoe Doi Luang (Doi Luang District), Chiang Rai Province, Thailand. Pwo Karen is the main language spoken by the villagers. The approximate population is 693 (356 males and 377 females). According to the phonological study, the sound system of Pwo Karen of Ban Doi is as follows (See Table 10 And Table 11):

Table 10 Pwo Karen of Ban Doi consonants

	Places of articulation	Bilabial	Labio- dental	Alveolar	Palatal	Velar	Glottal
Manners of articulation							
Stop	voiceless	p		t	c	k	
	unaspirated						
	voiceless	ph		th	ch	kh	
	aspirated						
Fricative	voiced	b		d			
Nasal	voiceless		f			x	h
	voiced						
Approximant		w			j		
Lateral				l			

Table 11 Pwo Karen of Ban Doi vowels

	Front	Central	Back
High	i	, ̃	u, ũ
Mid	e, ẽ	ə, ẽ	o, õ
Low	, ̃	a, ã	, ̃

/ai, au, ui, ə, ia, əu, i, ã, əi/

Synchronic analysis reveals that there are 4 tones in the smooth syllables, high, mid, rising and low. The other 2 allotones are in the checked syllables, low , and high .

3.2 Wordlist

The wordlist used to attest to the influence of voiceless and voiced initial consonants followed by oral vowels contains 16 words; that is 8 words for voiceless initial consonants and 8 words for voiced initial consonants. Voiceless and voiced initial consonants followed by nasal vowels are represented by 8 words, i.e. 4 words contain voiceless initial consonants and 4 words voiced initial consonants. In the case of voiceless consonant onset, the glottal stop / / has been regarded as a voiceless stop like /p/, /t/, /c/, /k/. This is due to / / being classified as the same phonation type as the remaining sounds, voiceless, in that there is

no vibration of the vocal folds (Ladefoged, 1999: 609). Almost all of the vowels in the wordlist were low vowels to diminish vowel height effect on the pitch. Tones on the vowels are mid tone (no tone mark) to avoid pitch perturbation. Moreover, it appears that in the wordlist not only monosyllabic words were selected but also the last syllables of disyllabic and sesquisyllabic words because the last syllables receive a prominent sound or stress as in the monosyllabic words. The wordlist is as follows:

Oral vowels		Nasal vowels	
Voiceless	Voiced	Voiceless	Voiced
k ‘to be’	m ‘birth scar’	natã ‘naughty’	sanã ‘to forget’
‘to love’	thà ‘chilli’	mi ã ‘burn’	cham ~ ‘to think’
a ‘much’	ba ‘cheap’	sathã ‘angry’	lal ~ ‘piece’
aka ‘to grill’	lola ‘a kind of tree’	kaik ~ ‘crooked’	kh b ~ ‘hip’
mapha ‘to rent’	à a ‘meat’		
sa ‘to breathe’	phai a ‘skin’		
‘to have’	ásaba ‘pity’		
bac ‘wet’	kàd ‘dull’		

3.3 Language consultants and data collection

For the Pwo Karen of Ban Doi, twelve language consultants of both sexes, equally, who were >60 years of age, 35-45 years of age, and <20 years of age were chosen. The language consultants were asked to pronounce each word 5 times. The number of test tokens for oral vowels in each age group was 960 (4 language consultants x 3 age group x 16 words x 5 times). The overall number of test tokens for nasal vowels was 480 (4 language consultants x 3 age groups x 8 words x 5 times). Regarding the nasal vowels, the words were pronounced twice to produce an identical number with the oral vowels above. Therefore, the test tokens used were 960 for both oral vowels and nasal vowels. The recording was done with a SONY IC Recorder ICD-MS515.

3.4 Data analysis

The entire vowel (vocalic portion) of each token was measured in respect of vowel duration, vowel amplitude and fundamental frequency. In each vowel, the time was normalized at the following points: 0%, 25%, 50%, 75% and 100%, a total of 5 measurement

points. The Praat program version 4.2.09 was used for the analysis. The duration of vowels was measured in milliseconds (m sec) in the range of the beginning of the vowel onset (0%) to the vowel offset (100%). The amplitude values were measured at each peak of the vowel. The fundamental frequency values were measured at 0-100 msec as it has been reported that the initial consonant voicing influences the following vowels from 0-100 msec (Hombert, Ohala, and Ewan, 1979). This range was normalized starting from the vowel onset or 0 msec to 100 msec into the following points: 0%, 50%, and 100%. Each point was averaged. Statistical analysis, Mean, Standard Deviation (SD) and *t-test* were used. Then, bar and line graphs were drawn.

4. Results

The results detail mean vowel duration, mean vowel amplitude and mean fundamental frequency values comparing the three age groups: the over-sixty group (>60), the middle group (35-45) and the under-twenty group (<20).

4.1 Duration

Table 12 Mean duration values, Standard deviation (SD) and *t-test* of oral vowels and nasal vowels in the three age groups: the over-sixty group (>60), the middle group (35-45) and the under-twenty group (<20)

Age group	Oral vowels (msec)			Nasal vowels (msec)		
	Voiceless onset	Voiced onset	<i>t-test</i> result (p)	Voiceless onset	Voiced onset	<i>t-test</i> result (p)
>60	224.86	273.89	.002	287.96	453.35	.000
	SD = .73	SD = .77		SD = .99	SD = .94	
35-45	288.50	329.50	.002	271.11	309.98	.002
	SD = .67	SD = .89		SD = .87	SD = .75	
<20	327.48	343.34	.045	231.73	232.05	.200
	SD = .92	SD = .71		SD = .90	SD = .87	

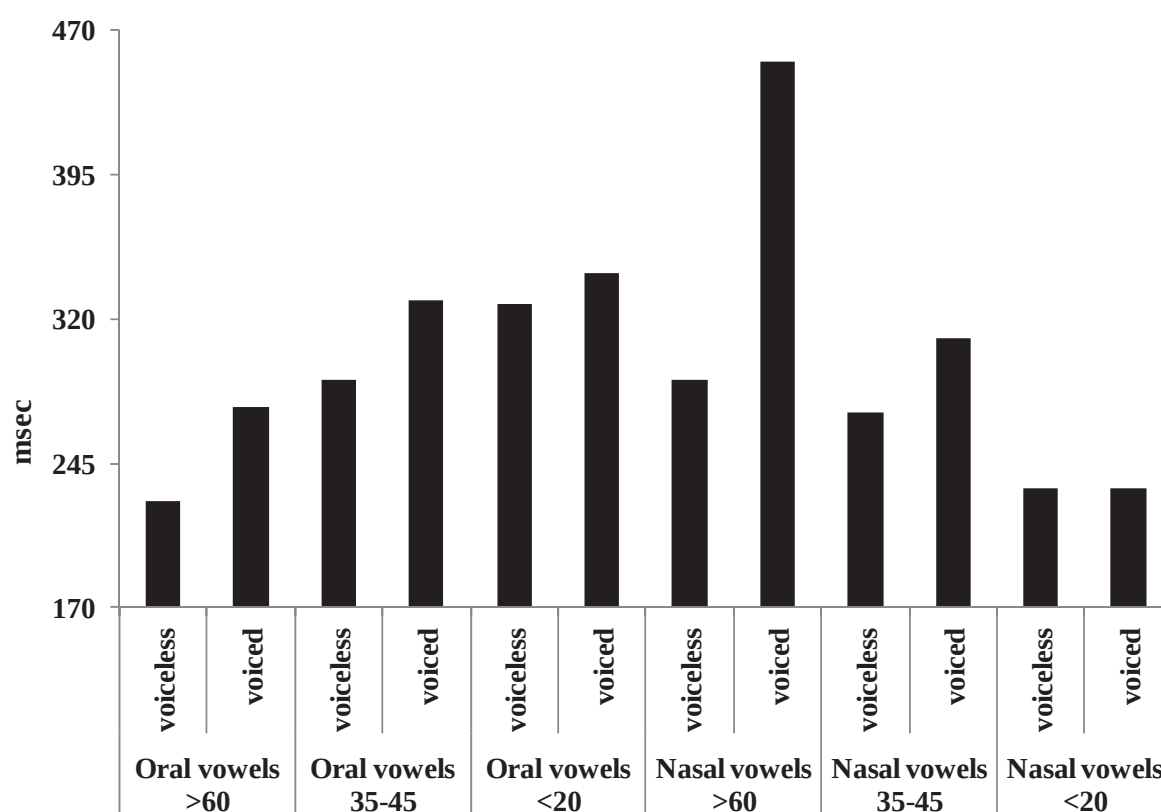


Figure 12 Mean duration values of oral vowels and nasal vowels in the three age groups: the over-sixty group (>60), the middle group (35-45) and the under-twenty group (<20)

Table 12 and Figure 12 show the mean duration values of oral and nasal vowels following voiceless and voiced initial consonants in the Pwo Karen of Ban Doi. The mean duration values of both types of oral and nasal vowels following voiceless initial consonant was significantly lower ($p < .05$) than that of voiced initial consonants for the over-sixty group (>60), the middle group (35-45), and the under-twenty group (<20), except that the nasal vowels following voiceless consonants were found to be insignificantly lower than that of the voiced consonant ($p = .200$). The mean duration values of nasal vowel were found to be inconsistently higher than those of the oral vowels. The size of the differences between the mean duration of the vowel following voiceless and following voiced consonant in each age group was not significantly different ($p > .05$).

4.2 Amplitude

Table 13 Mean amplitude values, Standard deviation (SD), and *t-test* of oral vowels and nasal vowels in the three age groups: the over-sixty group (>60), the middle group (35-45) and the under-twenty group (<20)

Age group	Oral vowels (dB)			Nasal vowels (dB)		
	Voiceless onset	Voiced onset	<i>t-test</i> result (p)	Voiceless onset	Voiced onset	<i>t-test</i> result (p)
>60	77.19	79.83	.56	78.64	79.99	.67
	SD = .99	SD = 1.24		SD = 1.79	SD = .90	
35-45	76.27	77.49	.62	79.86	80.08	.93
	SD = .82	SD = 1.79		SD = .78	SD = .97	
<20	80.95	81.02	.71	81.25	81.87	.89
	SD = 1.05	SD = 1.63		SD = 1.22	SD = 1.45	

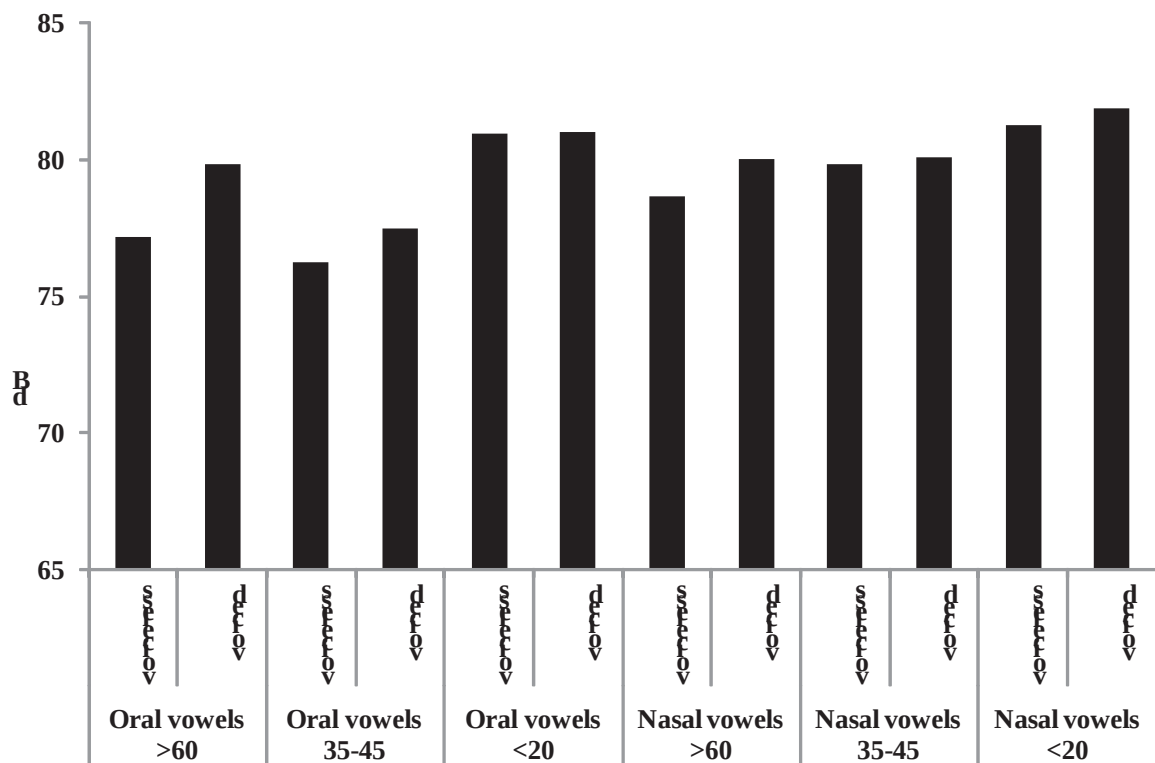


Figure 13 Mean amplitude values of oral vowels and nasal vowels in the three age groups: the over-sixty group (>60), the middle group (35-45) and the under-twenty group (<20)

Table 13 and Figure 13 show the mean amplitude values of oral and nasal vowels following voiceless and voiced initial consonants in the Pwo Karen of Ban Doi. The mean amplitude values of both oral and nasal vowels following voiceless initial consonants were insignificantly lower ($p > .05$) than those of voiced initial consonants for the over-sixty group (>60), the middle group (35-45) and the under-twenty group (<20). The mean amplitude values of nasal vowels were found to be higher than those of the oral vowels. The differences between the mean durations in each age group were not significantly different ($p > .05$). It was found that the size of differences between the mean amplitude of the vowel following voiceless and following voiced consonants in each age group was not significantly different ($p > .05$).

4.3 Fundamental frequency

Table 14 Mean fundamental frequency values, Standard deviation (SD) and *t*-test of oral vowels and nasal vowels in the three age groups: the over-sixty group (>60), the middle group (35-45) and the under-twenty group (<20)

	Age group								
	0%	>60 50%	100%	0%	35-45 50%	100%	0%	<20 50%	100%
Fundamental frequency of oral vowels following voiceless initial consonant (pa)	177.79	178.03	178.86	203.78	199.45	192.88	221.99	219.50	214.65
SD	2.44	3.12	2.89	3.05	1.26	2.75	2.47	2.50	2.36
Fundamental frequency of oral vowels following voiced initial consonant (ba)	183.09	189.21	194.22	203.92	199.33	193.67	225.88	219.12	213.35
SD	1.48	1.95	2.34	2.52	5.77	3.84	4.38	4.51	4.20
<i>t</i> -test result (p)	.089	.045	.042	.35	.39	.18	.09	.15	.22
Fundamental frequency of	198.65	197.52	197.20	213.35	211.52	209.62	222.25	218.90	215.52

nasal vowels following voiceless initial consonant (pã)									
SD	5.65	3.36	4.12	5.73	5.39	4.92	3.55	2.52	4.75
Fundamental frequency of nasal vowel following voiced initial consonant (bã)	210.91	204.89	200.63	205.12	200.10	197.88	224.44	220.57	213.79
SD	2.67	4.97	4.21	2.62	2.52	3.79	4.11	4.17	4.58
<i>t-test</i> result (p)	.038	.077	.095	.049	.043	.041	.093	.13	.11

Table 14 shows mean fundamental frequency values at the 3 measurement points: 0%, 50%, and 100%. Standard deviation (SD), and *t-test* result or p-values of both oral vowels and nasal vowels following voiceless initial consonants and voiced initial consonants in the three aged group: the over-sixty group (>60), the middle group (35-45) and the under-twenty group (<20) were also shown. In Table 14 and Figures 14-16, “pa” represents oral vowels following voiceless initial consonants, while “ba” represents oral vowels following voiced initial consonants. As for “pã”, it represents nasal vowels following voiceless initial consonants, while “bã” represents nasal vowels following voiced initial consonants.

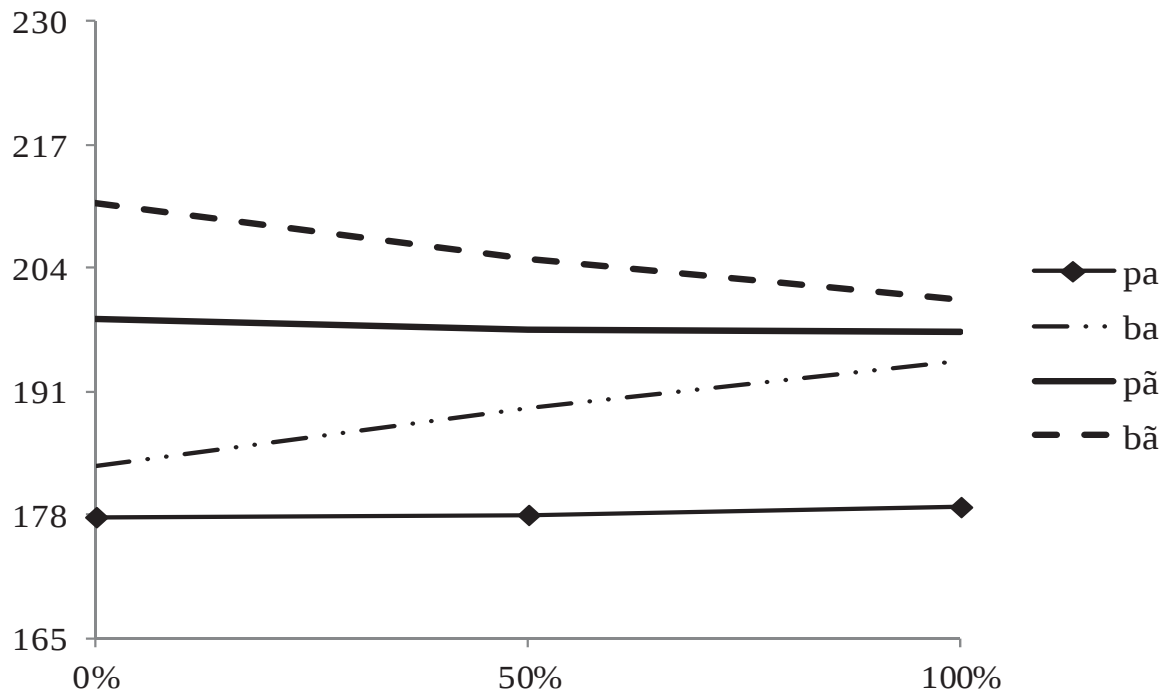


Figure 14 Mean fundamental frequency values of oral vowels and nasal vowels in the >60 years of age

From Table 14 and Figure 14, all lines were level and both oral and nasal vowels following voiced were clearly higher in pitch from 0%-100%. Regarding oral vowels, the vowels following voiceless initial consonants at 0% were found to be insignificantly higher in mean fundamental frequency values compared to those following voiced initial consonant but they were significantly higher at 50% and 100%. Regarding mean fundamental frequency values, those of the nasal vowels were higher than that of the oral vowels.

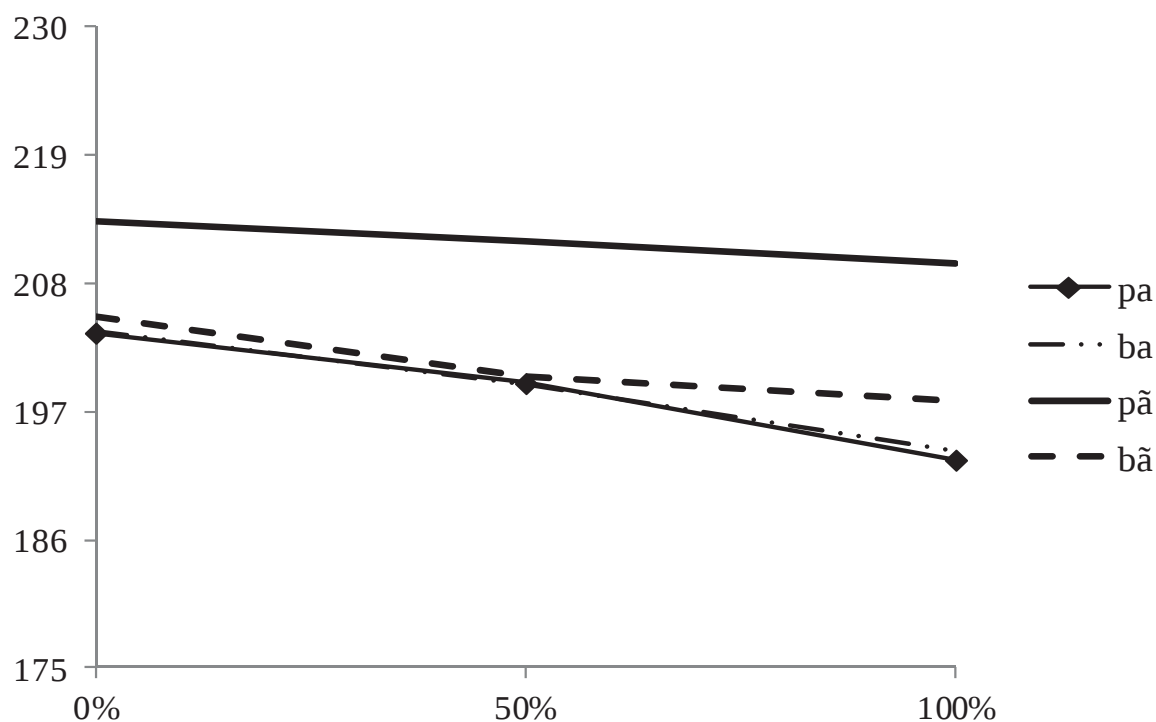


Figure 15 Mean fundamental frequency values of oral vowels and nasal vowels in the 35-45 years of age

From Table 14 and Figure 15 in the 35-45 age group all lines were found to be level. Oral vowels following voiced initial consonants were insignificantly higher in mean fundamental frequency values than that of the voiceless vowels, while the nasal vowels following voiced initial consonants were significantly lower in mean fundamental frequency values comparing to those following voiceless initial consonants. Regarding mean fundamental frequency values, those of the nasal vowels were higher than that of the oral vowels.

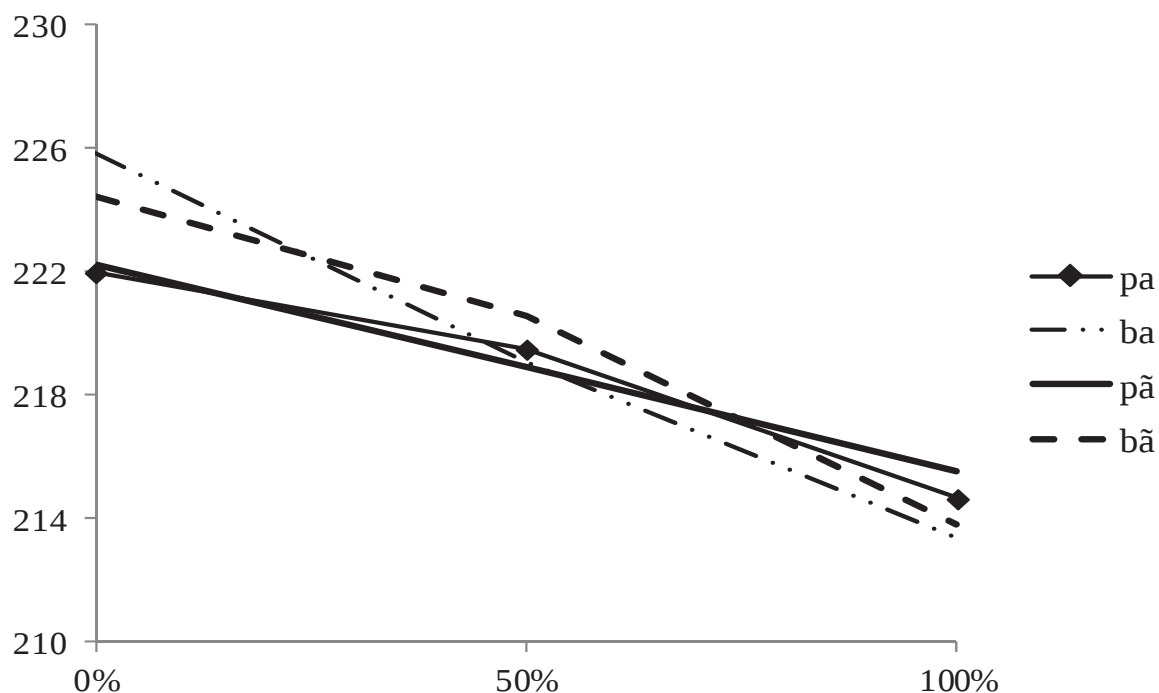


Figure 16 Mean fundamental frequency values of oral vowels and nasal vowels in the <20 years of age

From Table 14 and Figure 16 mean fundamental frequency values of clear vowels following voiced initial consonants were insignificantly higher in pitch at the onset compared to those following voiceless initial consonants. However, at the midpoint and the offset, it was found to be insignificantly lower. In the case of nasal vowels, the vowels following voiceless initial consonants were found to be insignificantly higher in mean fundamental frequency values than those following voiced initial consonants at the mid point. All lines seem to be similar in pitch contour but different in pitch height.

The size of the pitch differences seems to be larger in the >60 years of age and smaller in the 35-45 years of age, and <20 years of age, respectively. Moreover, for all vowel types, e.g. oral and nasal vowels, pitch contours were found to change from level to falling.

5. Discussion and conclusion

Linguists have claimed many universal factors (Erickson, 1975; Gandour, 1974; Hombert et al., 1979; House and Fairbanks, 1953; Lehiste, 1970; Lehiste and Peterson, 1961; L-Thongkum, 1990; Maddieson, 1984; Watkins, 2002) in the causes of tone development—initial consonant voicing, voiceless and voiced. These factors have caused the tones in all Tai languages (Li, 1966), some of the Mon-Khmer languages such as Vietnamese and Plang

(Diffloth, 1980; Haudricourt, 1954), Sgaw Karen and Cantonese in Sino-Tibetan (Haudricourt, 1954; Li, 1977; Sun, 2003), and Cham in Austronesian (Thurgood, 1999). However, there is one language reported not to have behaved in the same way, T'iensin, a dialect of Peking (Li, 1977). Ban Doi Pwo Karen seems to be another language that presents the uncommon phenomenon of the fundamental frequency behaviour of vowels influenced by voiced initial consonant being higher than the voiceless initial consonant. All of the three age groups showed identical behaviour. The size of the fundamental frequency differences seems to be larger in the >60 years of age and smaller in the <20 years of age. Regarding all vowel types, e.g. oral and nasal vowels, the fundamental frequency contours were found to change in the same direction, level to falling. This means that pitch height may have been used as a cue to perceive tone differences. It can be inferred from this that voiced initial consonant tends to cause a high tone, while voiceless initial consonants tend to cause a low tone. The phenomenon is an unusual case. However, when searching into the voice register of the vowels following some voiceless consonants, it was found that those vowels in the three age groups were pronounced with breathy voice vowels. As the voice register of vowels was found to play a major role in causing tone birth (L-Thongkum, 1989, 1990; Thurgood 2007), so breathy vowels have prominently affected the mean fundamental frequency values causing low pitch in vowels following voiceless consonants.

Therefore, as well as the process of stiffening (for voiceless) and slackening (for voiced) the cricothyroid muscles, there might be low muscular tension with weak medial compression and medium longitudinal tension of the vocal folds causing the vibrations' frequency to be just below the value typical of the modal voice (Eckert and Laver, 1994).

The mean amplitude of both types of vowel (oral vowel and nasal vowel) following voiceless initial consonants is lower than that of voiced initial consonants for the over-sixty group (>60), the middle group (35-45), and the under-twenty group (<20). The amplitude of the nasal vowel was found to be higher than that of the oral vowel. This is in line with previous research (Amelot and Rossatto, 2007; Ladefoged, 2003; Picket, 1998, Whalen and Beddor, 1989). However, the duration of the nasal vowel was found to be inconsistently higher than that of the oral vowel. This does not correspond with the other research in which the duration of the oral vowels was found to be less than that of the nasal vowels (Amelot and Rossatto, 2007; Ladefoged, 2003; Picket, 1998, Whalen and Beddor, 1989).

In conclusion, the initial consonant voicing perturbation of the fundamental frequency of oral vowels and nasal vowels was in the same direction, that is, a voiceless initial consonant causing a lowering of mean fundamental frequency and a voiced initial consonant

causing an increase in of mean fundamental frequency in Pwo Karen of Ban Doi. This challenges the tonogenesis universality claimed in previous research. The three age groups were reliable in showing a tone development tendency. However, physiological analysis should be initiated for further explanation.

Acknowledgements

I would like to express my deepest gratitude to Prof. Dr. Theraphan L-Thongkum for all her valuable support as a part of the Karen project funded by the Thailand Research Fund (TRF) throughout the year. I am also thankful to the anonymous readers of the Journal for all their fruitful comments.

References

- Abramson, A. S. 2004. Toward prosodic contrast: Suai and Pattani Malay. In B. Bel and I. Marlien (eds.), **Proceedings of the International Symposium on Tonal Aspects of Languages: Emphasis on Tone Social Sciences**. Beijing: The Institute of Linguistics, Chinese Academy of Social Sciences.
- Abramson, A. S., T. L-Thongkum, and P. W. Nye. 2004. Voice register in Suai (Kuai): An analysis of perceptual and acoustic data. **Phonetica** 61:147-171.
- Amelot, A. and S. Rossato. 2007. Velar movements for two French speakers. **Proceedings of the 16th International Congress of Phonetic Sciences** (August 6-10): 489-492.
- Brunelle, M. 2005. **Register in Eastern Cham: phonological, phonetic and sociolinguistic approaches**. Unpublished Ph.D. dissertation, Faculty of the Graduate School, Cornell University.
- Bunphan, J., P. Deepuengton, and S. Savetamalya. 1982. Fundamental frequency of [i], [u] and [a] of Standard Thai. **Science of Language** 2: 7-20. (in Thai)
- Diffloth, G. 1980. The Wa languages. **Linguistics of the Tibeto-Burman Area** 5.2: 1-182.
- Edmondson, J. A. and K. J. Gregerson. 1993. Western Cham as a register language. In J. A. Edmondson and K. J. Gregerson (eds.), **Tonality in Austronesian Languages**, 61-74. Honolulu: University of Hawaii Press.
- Eckert, H. and J. Laver. 1994. **Menschen und ihre Stimmen: Aspekte der vokalen Kommunikation**. Weinheim: Psychologie Verlags Union.

- Erickson, D. 1975. Phonetic implications for a historical account of tonogenesis in Thai. In J. G. Harris and J. R. Chamberlain (eds.), **Studies in Tai Linguistics in Honor of W. J. Gedney**, 100-111. Bangkok: Central Institute of English Language.
- Gandour, J. T. 1974. Consonant types and tone in Siamese. **Journal of Phonetics** 2: 337-350.
- Grimes, J. E. 1988. Correlations between vocabulary similarity and intelligibility. **Notes on Linguistics** 41: 19-33.
- Guion, S. G. and R. P. Wayland. 2004. Aerodynamic coarticulation in sound change or how onset trills can condition a falling tone. In A. Agwuele, W. Warren, and S-H Park (eds.), **Proceedings of the 2003 Texas Linguistics Society Conference**, 107-115. Somerville: Cascadilla Proceedings Project.
- Haudricourt, A-G. 1954. De l'origine des tons en vietnamien. **Journal Asiatique** 242: 69-82.
- Henderson, E. J. A. 1982. Tonogenesis: some recent speculations on the development of tone. **Philological Society**, 1-24. Oxford: Oxford University Press.
- Hombert, J-M., J. Ohala, and W. Ewan. 1979. Phonetic explanations for the development of tones. **Language** 55: 37-58.
- House, A. S. and G. Fairbanks. 1953. The influence of consonant environment upon the secondary acoustical characteristic of vowels. **Journal of Acoustical Society of America** 25: 105-113.
- Ladefoged, P. 1999. Linguistic phonetic descriptions. In W. J. Hardcastle and J. Laver (eds.), **The Handbook of Phonetic Sciences**, 589-618. Massachusetts: Blackwell.
- Ladefoged, P. 2003. **Phonetic data analysis: an introduction to fieldwork and instrumental technique**. Oxford: Blackwell.
- Lehiste, I. 1970. **Suprasegmentals**. Cambridge: MIT Press.
- Lehiste, I. and G. Peterson. 1961. Some basic considerations in the analysis of intonation. **Journal of Acoustical Society of America** 33: 419-425.
- Li, F-K. 1966. The relationship between tones and initials in Tai. In N. Zide (ed.), **Studies in Comparative Austroasiatic Linguistics**, 82-86. The Hague: Mouton.
- Li, F-K. 1977. **A handbook of Comparative Tai**. Manoa: The University of Hawaii Press.
- L-Thongkum, T. 1984. **Nyah-Kur (Chaobon)-Thai-English Dictionary**. Bangkok: Chulalongkorn University Printing House.
- L-Thongkum, T. 1988. Phonation types in Mon-Khmer languages. In O. Fujimura (eds.), **Vocal fold physiology: voice production, mechanisms and functions**, 319-333. New York: Raven Press.

- L-Thongkum, T. 1989. An acoustic study of the register complex in Kui (Suai). **Mon-Khmer Studies 15**: 1-19.
- L-Thongkum, T. 1990. The interaction between pitch and phonation type in Mon: phonetic implication for a theory of tonogenesis. **Mon-Khmer Studies 16-17**: 11-24.
- L-Thongkum, T., P. Teeranon, and C. Intajamornrak. 2007. The interaction between vowel length and pitch in SEA languages: an implication for tonal evolution. In J. G. Harris, S. Burusphat, J. E. Harris (eds.), **Studies in Tai and Southeast Asian Linguistics**, 225-240. Bangkok: Ekphimthai.
- Lohde, K. 2003. **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, Department of Linguistics, Chulalongkorn University.
- Maddieson, I. 1984. The effects on f_0 of a voicing distinction in sonorants and their implication for a theory of tonogenesis. **Journal of Phonetics 12**: 9-15.
- Maran, L. R. 1973. On becoming a tone language: a Tibeto-Burman model of tonogenesis. In L. M. Hyman (ed.), **Consonant Types and Tones (Southern California Occasional Papers in Linguistics 1)**, 97-114. Los Angeles: University of Southern California.
- Matisoff, J. A. 1973. Tonogenesis in Southeast Asia. In L. M. Hyman (ed.), **Consonant Types and Tones (Southern California Occasional Papers in Linguistics 1)**, 71-95. Los Angeles: University of Southern California.
- Mazaudon, M. 1977. Tibeto-Burman tonogenetics. **Linguistics of the Tibeto-Burman Area 3.2**: 1-123.
- Michaud, A. 2012. Monosyllabicization: Patterns of evolution in Asian languages. In N. Nau, T. Stolz, and C. Stroh (eds), **Monosyllables: From Phonology to Typology**, 115-130. Berlin: Akademie Verlag.
- Mohr, B. 1971. Intrinsic variations in the speech signal. **Phonetica 23**: 65-93.
- Phu Van Han, J. A. Edmondson and K. Gregerson. 1992. Eastern Cham as a tone language. **Mon-Khmer Studies 20**: 31-44.
- Pickett, J. M. 1998. **The acoustics of speech communication**. MA: Allyn and Bacon.
- Premssirrat, S. 2003. Khmu dialects: a case of register complex and tonogenesis. In E. Shigeki (ed.), **Proceedings of the Symposium Cross-Linguistic Studies of Tonal Phenomena: Historical Development, Phonetics of Tone and Descriptive Studies**, 13-28. Tokyo: ILCAA, University of Foreign Studies.

- Przyluski, J. 1924. Les langues Munda. In A. Meillet and M. Cohen (eds.), **Les Langues du Monde**, 385-403. Paris: Librairie Ancienne Édouard Champion.
- Robinson, E. R. 1994. **Further classification of Southwestern Tai “P” group languages**. Unpublished M.A Thesis, Department of Linguistics, Chulalongkorn University.
- Rose, P. 1997. Seven-tone dialect in Southern Thai with super-high: Pakphanang tonal acoustics and physiological inferences. In A. S. Abramson (ed.), **Southeast Asian Linguistic Studies in Honour of Vichin Panupong**, 191-208. Bangkok: Chulalongkorn Printing House.
- Sagart, L. 1993. Chinese and Austronesian: evidence for a genetic relationship. **Journal of Chinese Linguistics** 21.1: 1-64.
- Sun, J. T.-S. 2003. Variegated tonal developments in Tibetan. In D. Bradley, R. LaPolla, B. Michailovsky and G. Thurgood (eds.), **Language Variation: Papers on Variation and Change in the Sinosphere and in the Indosphere in Honour of James A. Matisoff**, 35-52. Honolulu: University of Hawaii Press.
- Svantesson, J-O. 1988. U. **Linguistics of the Tibeto-Burman Area** 11: 64–133.
- Svantesson, J-O. 1991. Hu – a language with unorthodox tonogenesis. In J. Davidson (ed.), **Austroasiatic Languages: Essays in Honour of H. L. Shorto**, 67-79. London: SOAS.
- Teeranon, P. 2008. Tonal evolution from the influence of initial consonant clusters: another look from Lavua’, Wa, and Plang. In N. Ronnakiat and A. Saengmanee (eds.), **A Collection of Essays on Southeast Asian Linguistics: Festschrift for Professor Theraphan Luangthongkum’s 60 Birthday**, 30-42. Bangkok: Phikkhanet Printing Center. (in Thai)
- Thach, N. M. 1999. Monosyllabization in Kiengiang Khmer. **Mon-Khmer Studies** 29: 81-95.
- Thavisak, A. 2001. The effects of glottal finals on pitch in Southeast Asian languages. **Mon-Khmer Studies** 31: 57-64.
- Thurgood, G. W. 1993. Phan Rang Cham and Utsat: tonogenetic themes and variants. In J. A. Edmondson and K. J. Gregerson (eds.), **Tonality in Austronesian Languages, Oceanic Linguistics Special Publication** 24, 91-106. Honolulu: University of Hawaii Press.
- Thurgood, G. W. 1996. Language contact and the directionality of internal drift: The development of tones and registers in Chamic. **Language** 72.1: 1-31.

- Thurgood, G. W. 1999. **From ancient Cham to modern dialects: two thousand years of language contact and change: Oceanic linguistics special publication no.28.** Honolulu: The University of Hawaii Press.
- Thurgood, G. W. 2007. Tonogenesis revisited: revising the model and the analysis. In J. G. Harris, S. Burusphat, and J. E. Harris (eds.), **Studies in Tai and Southeast Asian Linguistics**, 263-291. Bangkok: Ekphimthai.
- Watkins, J. 2002. **The phonetics of Wa: experimental phonetics, phonology, orthography and sociolinguistics.** Canberra: Research School of Pacific and Asian Studies, The Australian National University.
- Whalen, D. H. and A. G. Levitt. 1995. The universality of intrinsic F0 of vowels. **Journal of Phonetics** 23: 349-366.
- Whalen, D. H. and P. S. Beddor. 1989. Connections between nasality and vowel duration and height: Elucidation of the Eastern Algonquian intrusive nasal. **Language** 65: 457-486.
- Zee, E. 1980. Tone and vowel quality. **Journal of Phonetics** 8: 247-258.

อ. ดร. ชมนาด อินทจามรรักษ์

**Variation and change of the Phrae Pwo Karen vowels and
tones induced by language contact with the Tai languages**

VARIATION AND CHANGE OF THE PHRAE PWO KAREN VOWELS AND TONES INDUCED BY LANGUAGE CONTACT WITH THE TAI LANGUAGES

Chommanad Intajamornrak¹

Abstract

This paper aims to analyze and compare the acoustic characteristics of the vowels and tones in the Phrae Pwo Karen spoken by three generations. The data was collected at Khangchai Village in Wang Chin District, Phrae Province. A wordlist of Pwo Karen vowels and tones was recorded directly on to computer using Adobe Audition version 2. Fifteen female informants were divided into three groups: those over 60 years old, those 35-50 years old and those under 25 years old. The total number of test tokens was 405 for vowel analysis, and 810 for tone analysis. The fundamental frequencies and formant frequencies were measured using Praat version 5.1.43.

The results show that there are nine monophthongs in Phrae Pwo Karen, namely /i, e, ε, i, ə, a, u, o, ɔ/. Considering the vowel spaces, it is noticeable in the over-60 group that front vowels /i, e, ε/ occur very close to each other, i.e., with only a little difference in the tongue height position (F1). Whereas the back vowels /u, o/ occur close to each other, the vowel /ɔ/ appears close to the vowel /a/. In the 35-50

group and the under-25 group, the vowel /ε/ moves downward and the vowel /ɔ/ moves upward. There are 4 tones, namely, the high tone, mid tone, low tone and falling tone. In the over-60 group, the high tone begins at a high pitch and stays level until the end. The mid tone starts at a mid pitch and stays level until the end. The low tone starts at a mid pitch and falls to a low pitch. The falling tone begins at a high pitch and then rises slightly before sharply falling to a low pitch. For the 35-50 group, the acoustic characteristics of the 4 tones are similar to those of the older group; however, the high tone behaves differently. The onset of the high tone is lower and rises slightly until the end. The change of the high tone is clearer in the under-25 group. It starts from a mid pitch and rises sharply to a high pitch.

In conclusion, the acoustic characteristics of the vowels and tones as spoken by three generations suggest that Phrae Pwo Karen is changing because of the variation among the three groups. The variation and change seems to be caused by language contact with Tai Yuan and Standard Thai.

1. Introduction

Pwo Karen is one of the six Karen languages spoken in Thailand. The other five are Sgaw, Pa-O, Kayan, Kayah, and Kayaw. Pwo Karen and Sgaw Karen are the two languages with the greatest number of speakers. Pwo Karen can further be subdivided into dialects on the basis of classical criteria such as their degree of lexical similarity (or mutual intelligibility) and geographic distribution. The three examples below provide a glimpse into cross-dialect differences which including the presence of some non-cognate words and differences in consonants, vowels and tones between

¹ Lecturer, Faculty of Humanities, Naresuan University

cognates. ‘Sangkhla’ refers to the dialect of Sangkhla District, Kanchanaburi Province, ‘Omkoï’ for that of Omkoï District, Chiangmai Province (Phillips 2009), and ‘Khangchai’ for the variety spoken in Khangchai Village, Wangchin District, Phrae Province.

Sangkhla	Omkoï	Khangchai	Gloss
Tα? ³¹	Tα? ³²	sa ³³	‘heart’
na ²¹ kuŋ ⁵¹	nɛ ³⁴	na ²¹	‘ear’
phlɔ ⁵¹	ke ³¹ ma ⁴⁵	ʔaŋ ³³ klaŋ ⁴⁵	‘to give’

The present study focuses on the third of these dialects: Khangchai. It belongs to the Phrae Pwo Karen dialect of Pwo Karen, which is also known as North-eastern Pwo (Lewis 2009). The speakers of Pwo Karen in Phrae Province have been settled in this area for at least 150 years (National Archives of Thailand 1999). Pwo Karen speaking people outnumber those of the other non-Tai ethnicities, e.g. Hmong, Akha, Lisu, and Mlabri. Khangchai Village is one of fourteen Pwo Karen speaking villages (in two districts) and almost all the speakers are bilingual. They can speak Phrae Tai Yuan, with the exception of some elderly speakers who are passive bilinguals. The young generation can also speak Standard Thai very well, since they have been educated in school.

The factors which bring about language variation and change can be distinguished into two groups, internal and external. Internal factors deal with features of a language itself, for example, vowel shifting in English, whereas external factors usually deal with language contact and social factors, e.g. age, gender, social class, etc. Weinreich (1968: 3) explains that among the extra-linguistic factors which must be considered in a language contact situation, some are inherent in the

bilingual speakers’ relation to the languages they come into contact with. These include, for example, the speakers’ facility of verbal expression in general and their ability to keep the two languages apart, their relative proficiency in each language and specialization in the use of each language by topic and interlocutor. In addition, there are certain features of bilingual groups such as size, demographic facts, and political relations.

Further, Thomason (2001:60) also concluded that the results of language contact can be classified into three types: contact-induced change, extreme language mixture, and language death. Contact-induced change is varied in the kind and degree of change by two predictors: social factors and linguistic factors. Social factors include intensity of contact, presence vs. absence of imperfect learning and speakers’ attitude. Linguistic factors include universal markedness, the degree to which features are integrated into the linguistic system, and typological distance between source and recipient languages. The next type of result of language contact is extreme language mixture. This level leads to pidgins, creoles, and bilingual mixed languages. The last type is language death. This occurs through attrition or the loss of linguistic material and grammatical replacement.

Many scholars have studied tonal variation and change in Southeast Asian languages (Akharawatthanakun 2002 and 2009, Intajamornrak 2011, L-Thongkum 1994, Teeranon 2002) and found that acoustic characteristics of tones are influenced both by internal factors and by external factors. There are also generational differences in acoustic characteristics of vowels (Cox 1999), Decker and Mackenzie 2000,

Jacewicz, Fox and Salmons 2011). As the language used by different generations can provide insights into variation and change, this paper aims to analyze and compare the acoustic characteristics of the vowels and tones of Phrae Pwo Karen as spoken by three generations.

2. Phonological system of Phrae Pwo Karen (Khangchai Village)

Fieldwork data collected in 2010-2011 show that the Pwo Karen dialect spoken at Khangchai Village has 21 consonant phonemes. Only /ʔ/ and /ŋ/ can occur in final position. There are 9 monophthongs and 4 diphthongs, without vowel length distinction. The tonal system comprises 4 tones in non-checked syllables and 2 tones in checked syllables.

Consonants

	Bilabial	Alveolar	Palatal	Velar	Glottal
Stop	p p ^h b	t t ^h d	c c ^h	k k ^h	ʔ
Nasal	m	n		ŋ	
Fricative		s		x ɣ	h
Approximant	w		j		
Lateral		l			

Vowels

Monophthongs	Front	Central	Back
High	i	ɨ	u
Mid	e	ə	o
Low	ɛ	a	ɔ
Diphthongs	ai	ɔi	ai au

Tones

Non-checked syllable (CV/CVN): /high/ [44 ~ 45] /mid/ [33] /low/ [21]

/falling/ [41]

Checked syllable (CVʔ): /midʔ/ [332 ~ 33] /fallingʔ/ [442 ~ 42]²

² /ʔ/ is marked for tones that occur in checked syllables. It does not mean a glottalized tone

3. Methodology

The data was collected at Khangchai Village in Wangchin District, Phrae Province. Because of the variation among Pwo Karen dialects, the 1,000 vocabulary items were collected in the first fieldwork. The vocabulary included action verbs, stative verbs, body parts and secretion, natural objects and phenomena, manmade objects and constructions, kinship terms, numerals, colors, time, direction and location, pronouns, and questions³. After that, a wordlist for acoustical measurement of vowels and consonants was chosen from among these vocabulary items.

The words for acoustical measurement of the nine monophthongs are all mid tone and occur in non-checked syllables. The initials are voiceless aspirated stops which prevent the acoustic analysis from being affected by phonation. A wordlist for vowels is shown below.

Vowels

/i/	khi ³³	‘tick’
/e/	kheŋ ³³	‘cricket’
/ɛ/	thɛ ³³	‘to crack’
/i/	phi ³³	‘short’
/ə/	khəŋ ³³	‘to dig’
/a/	khaŋ ³³	‘calf’
/u/	khuf ³³	‘smoke’
/o/	thoŋ ³³	‘bag’
/ɔ/	thɔŋ ³³	‘blood vain’

The words for acoustical measurement of tones occur both in non-checked syllables and checked syllables and also have voiceless aspirated stops as initial consonants. A wordlist for tones is shown below.

³ The Thai and English gloss for collecting the vocabulary was conducted by Prof. Dr. Theraphan L-Thongkum, the head of the “Karen Linguistics” project.

Tones

Non-checked syllables⁴

/high/	khəŋ ⁴⁵	‘chopping board’
/mid/	khəŋ ³³	‘to dig’
/low/	thəŋ ²¹	‘comb’
/falling/	phəŋ ⁴¹	‘hole’
/high/	thaŋ ⁴⁵	‘gizzard’
/mid/	thaŋ ³³	‘exit’
/low/	thaŋ ²¹	‘to hit’
/falling/	phaŋ ⁴¹	‘lance’
/high/	phoŋ ⁴⁵	‘abdomen’
/mid/	thoŋ ³³	‘bag (shoulder~)’
/low/	thoŋ ²¹	‘turn upside down’
/falling/	thəŋ ⁴¹	‘to turn back’

Checked syllables

/midɿ/	thauɿ ³³	‘to polish’
/fallingɿ/	theɿ ⁴¹	‘to fall’
/midɿ/	thaiɿ ³³	‘thigh’
/fallingɿ/	thaɿ ⁴¹	‘needle’

⁴ The word khəŋ⁴⁵ ‘chopping board’ is a loanword from Thai, but it has been used for a long time. The older speakers also use this word. Even though it is a loanword, it doesn’t affect the tone shape since it is integrated in the native lexicon. The word thoŋ³³ ‘bag (shoulder~)’ is not borrowed from Thai word thun²⁴ ‘plastic or paper bag’ because Pwo Karen also has the word ta²¹ soŋ³³, which means ‘plastic bag’.

/midɿ/ phaiɿ³³ 'skin'
 /fallingɿ/ thuɿ⁴¹ 'pig'

The wordlist of Pwo Karen vowels and tones was recorded directly on to computer using Adobe Audition version 2. Fifteen female⁵ informants were divided into three groups: those over 60 years old, those 35-50 years old, and those under 25 years old. The informants were asked to pronounce each test-word three times randomly for each list, with a three-to-five second break between each word. The total number of test tokens was 405 (9 words x 3 times x 15 informants) for vowel analysis, and 810 (18 words x 3 times x 15 informants) for tone analysis.

The formant frequencies were measured at every 10 millisecond interval between 25% - 75% using Praat version 5.1.43. Then, the variation of each vowel within its space was plotted by vowel plot program⁶ on a graph for each speaker.

The fundamental frequencies were measured at every 10% of normalized duration using Praat version 5.1.43. The measured fundamental frequencies in Hertz were converted into semitone values. The formula was semitones = 3.32 x 12 x Log (Hz to be translated / Hz reference level). This is to help minimize the variation among the pronunciation of the five female speakers in each group. Microsoft Excel 2007 was used to analyze and plot graphs of the semitone values.

⁵ Only female informants were selected because different gender affects acoustical measurement.

⁶ Created by Mr. Patavee Chanvivit and thanks to Ms. Supaporn Phalipat for her help in drawing the graphs from the vowel plot program.

4. Results

4.1 Vowels

The results show the vowel spaces of the nine monophthongs in Phrae Pwo, Karen namely /i e ɛ i ə ɑ u o ɔ/. The formant frequencies of the Phrae Pwo Karen vowels were divided into 3 groups, the over-60 group, the 35-50 group, and the under-25 group, as shown in Figures 1, 2, and 3.⁷

Figure 1 shows that the overall space of the nine monophthongs as spoken by five speakers of the over-60 group is similar. The front vowels /i e ɛ/ occur very close to each other, i.e., with only a little difference in the tongue height position or the first formant frequency (F1) especially for speakers 1, 2 and 4. The central vowel /i ə/ and the back vowels /u ɑ/ occur close to each other, whereas the vowels /a/ appears close to the vowel /ɔ/. Concerning the tongue height position, there are only two low vowels, which are /a/ and /ɔ/.

In terms of the tongue advancement or the second formant frequency (F2), there are front vowels /i e ɛ/, central vowels /i ə ɑ/, and back vowels /u o ɔ/. The front vowels occur quite separately from the central and back vowels. It also appears that the vowels /a/ and /ɔ/ occur close to each other, with only a little difference in the tongue advancement position.

⁷ See the formant frequencies of each vowel in Appendix 1.

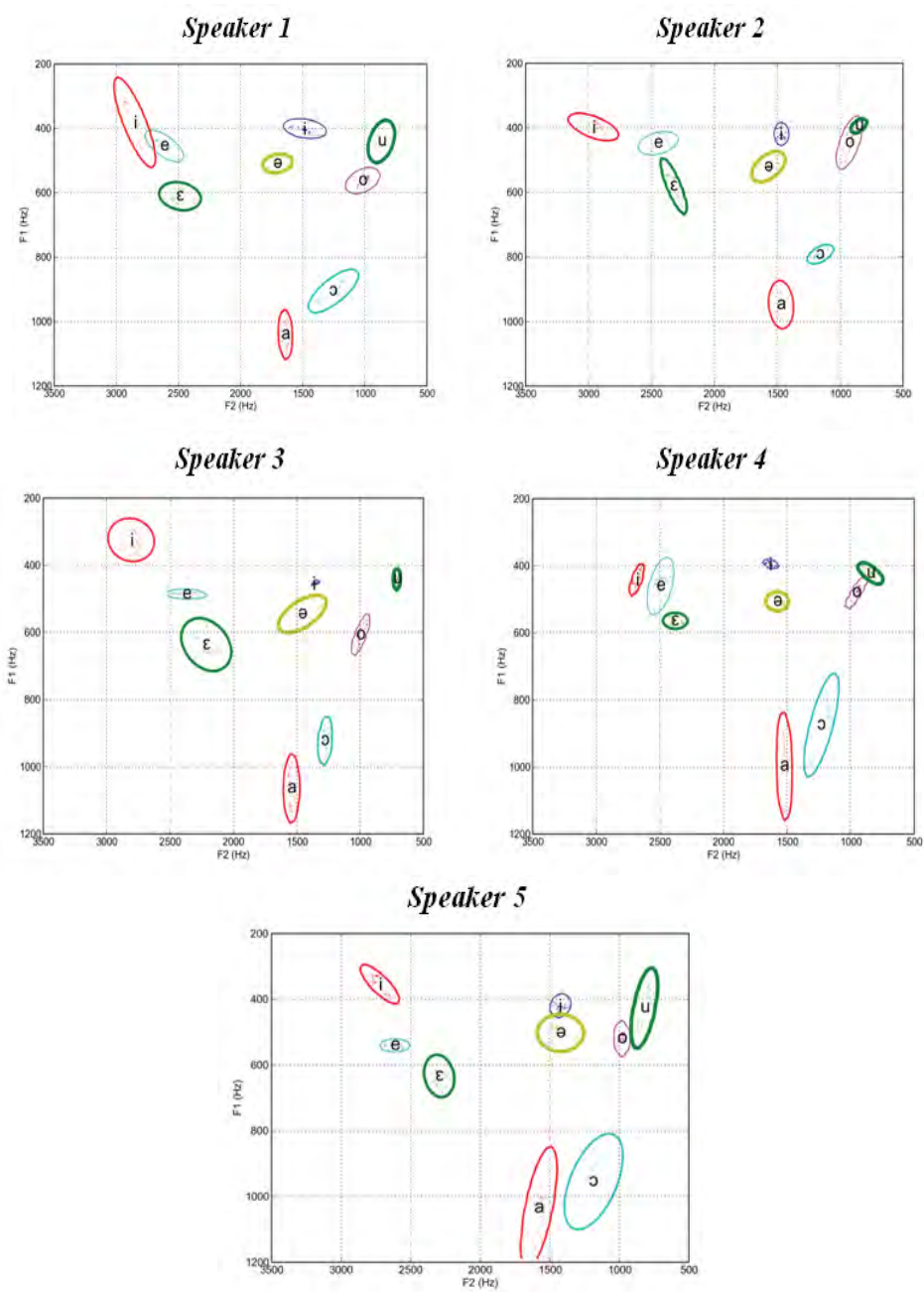


Figure 1 Vowel spaces of Phrae Pwo Karen vowels as spoken by five speakers of the over-60 group

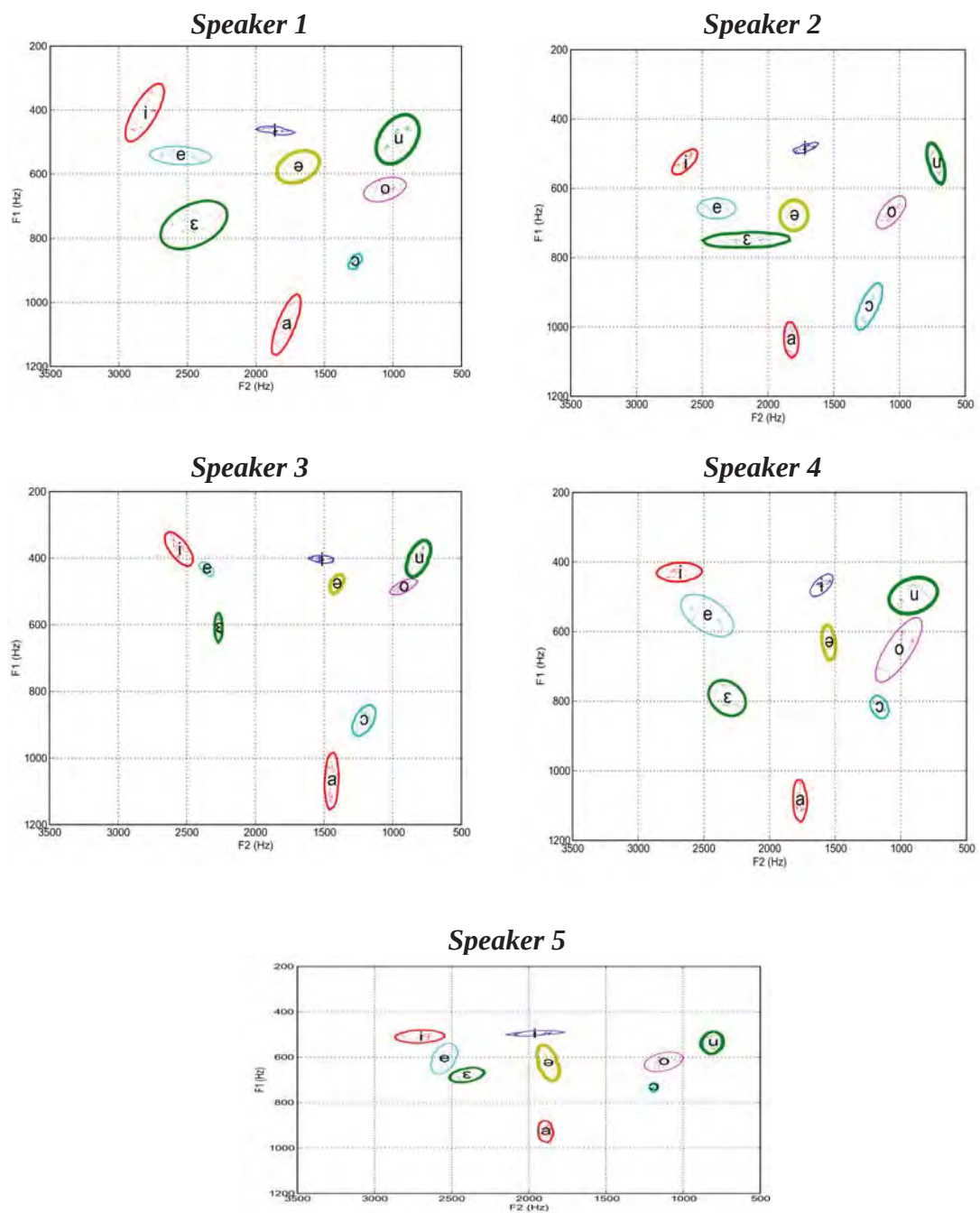


Figure 2 Vowel spaces of Phrae Pwo Karen vowels as spoken by five speakers of the 35-50 group

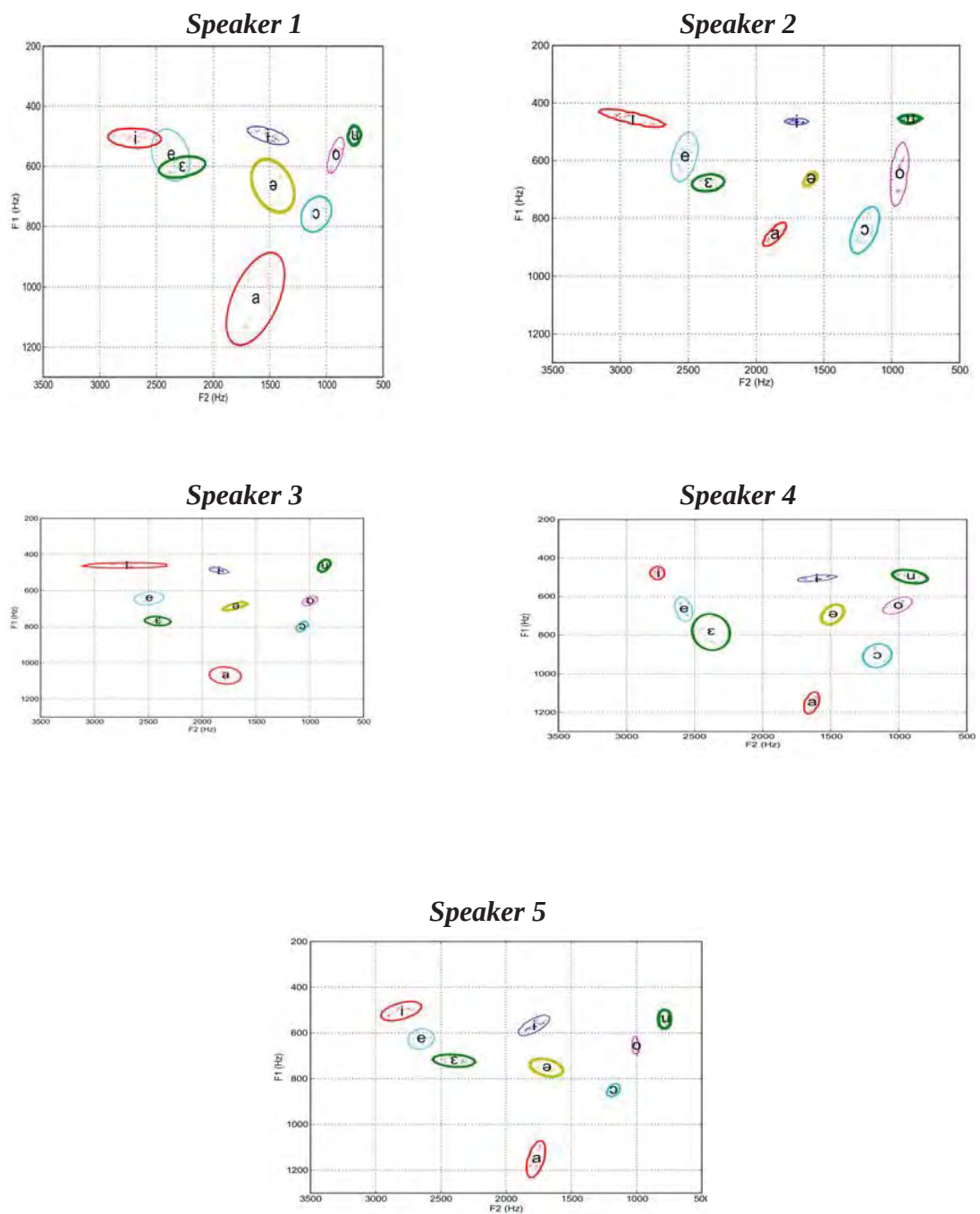


Figure 3 Vowel spaces of Phrae Pwo Karen vowels as spoken by five speakers of the under-25 group

Figure 2 shows that the overall space of the nine monophthongs as spoken by five speakers of the 35-50 group seems to be different from that spoken by the over 60 group. Even though the front vowels /i e ε/ occur close to one another, it can be clearly seen in speakers 1, 3, and 4 that the vowel /ε/ moves downward separately from the vowels /i e/. This means that the difference of tongue height position or the first formant frequency (F1) is wider in these three speakers. The central vowels /i ə/ occur close to each other separately from the vowel /a/. It is also noticeable that the back vowel /ɔ/ moves upward in the data from speakers 4 and 5. It can be said that these speakers move the tongue height position higher than the other speakers.

Considering the tongue advancement or the second formant frequency (F2), the tongue position of the vowels /a/ and /ɔ/ occurs separately from each in all of the speakers except for speaker 3. It clearly shows that for the vowel /ɔ/, the back of the tongue moves to a similar position as the vowels /u o/. It is noticeable that the vowels /a/ and /ɔ/ of speaker 3 occur close to each other, with only a little difference in the tongue advancement position which is similar to the over-60 group.

In Figure 3, the overall space of the nine monophthongs as spoken by five speakers of the under-25 group looks similar to that spoken by the 35-50 group. Considering the front vowels, the vowel /ε/ moves downward, whereas the vowel /ɔ/ moves upward in the vowel area. The difference of the tongue height position of the vowels /i e ε/ is clearer, except for speaker 1. As for the vowel /ɔ/, the tongue moves to the higher position in all of the speakers.

In conclusion, the vowel /ε/ moves downward in speakers 3 and 5 of the over-60 group (see Figure 1), in speakers 1, 2, and 4 of the 35-50 group (see Figure 2), and in speakers 2, 3, 4, and 5 of the under-25 group (see Figure 3). The vowel /ɔ/ moves upward in the vowel area as seen in Figure 2 in which the vowel /ɔ/ moves upward in speakers 1, 2, 4 and 5. In Figure 3, the vowel /ɔ/ moves upward in all of the speakers. It looks as if the variation and change occur in Phrae Pwo Karen vowels as clearly seen in the 35-50 group and the under-25 group.

4.2 Tones

There are 4 tones in Phrae Pwo Karen, namely, the high tone, mid tone, low tone and falling tone with an allotone in the mid tone and falling tone. Phonetically, the high tone is high level and high rising [44 ~ 45], the mid tone is mid level [33], the low tone is low falling [21], and the falling tone is high falling [41]. The mid tone in checked syllables is [332 ~ 32] and the falling tone is [442 ~ 42].

The semitones converted from fundamental frequencies of each tone as spoken by five speakers of each group were plotted using Microsoft Excel 2007 as shown in Figures 4, 5, and 6.⁸

⁸ See the fundamental frequencies of Phrae Pwo Karen tones of each speaker in each group in Appendix 1 and Appendix 2.

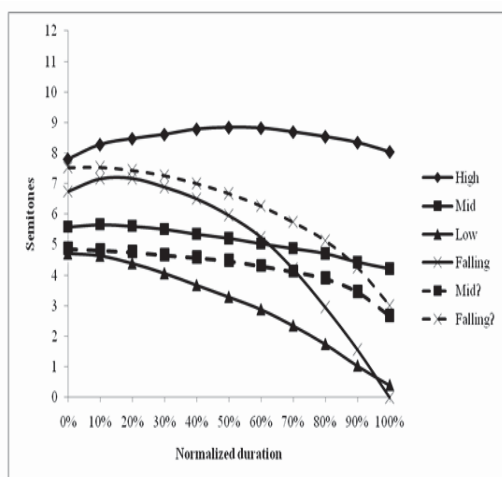


Figure 4 Semitones of Phrae Pwo Karen tones as spoken by the over-60 group⁹

Figure 4 shows that in the over-60 group, in non-checked syllables, the high tone begins at a high pitch and stays level until the end. The mid tone starts at a mid pitch and also stays level until the end. The low tone starts at a mid/low¹⁰ pitch and falls to a low pitch. The falling tone begins at a high pitch and then rises slightly before sharply falling to a low pitch.

In checked syllables, the mid tone starts at a mid/low pitch and stays level before falling slightly at 80% of the duration. The falling tone begins at a high pitch and then falls to a mid/low pitch.

For the 35-50 group as shown in Figure 5, the acoustic characteristics of the 4 tones are similar to those of the over-60 group. In non-checked syllables, the high tone begins at a high pitch and rises slightly until the end. The mid tone starts at mid pitch and also stays level until the end. The low tone starts at a mid pitch and falls sharply to a low pitch. The falling tone begins at a high pitch and then sharply falls to a low pitch at 50% of the duration.

However, it is noticeable that the high tone behaves differently from that of the over-60 group.

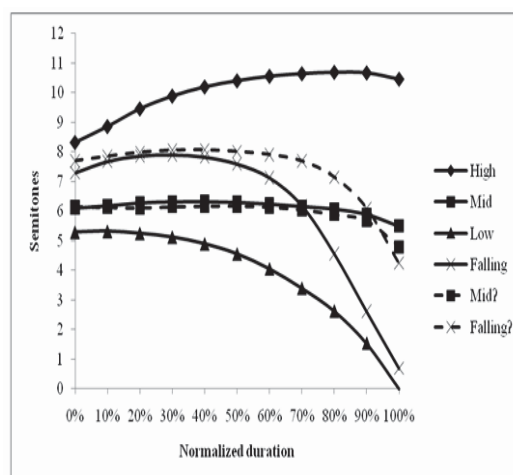


Figure 5 Semitones of Phrae Pwo Karen tones as spoken by the 35-50 group

In checked syllables, the mid tone has a similar contour as in non-checked syllables. It starts at mid pitch and also stays level until the end. The falling tone begins at a high pitch and stays level until 60% of the duration before falling to a mid pitch.

In Figure 6, in non-checked syllables, the change of the high tone is clearer in the under-25 group. It starts from a mid pitch and rises sharply to the highest pitch of the scale. The acoustic characteristic of the mid tone is mid level starting at a mid pitch and staying level to the end of the duration. The low tone is low-falling starting at a mid pitch and falling sharply to a low pitch. The falling tone is high-falling starting at a high pitch and slightly rising before sharply falling at 60% of the duration.

⁹ Solid line represents a non-checked syllable. Dotted line represents a checked syllable.

¹⁰ The mid/low pitch means a mid pitch which moves to a low pitch.

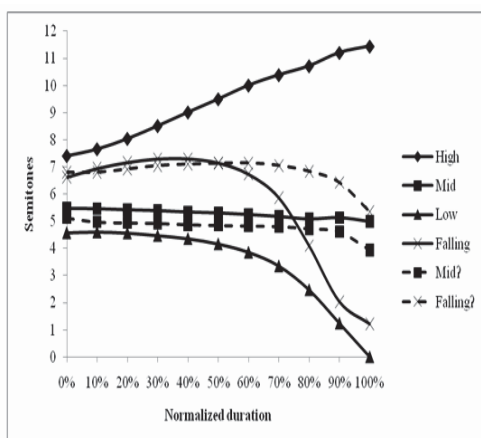


Figure 6 Semitones of Phrae Pwo Karen tones as spoken by the under-25 group

In checked syllables, the mid tone starts at a mid pitch and stays level before falling slightly at the last 10% of the duration. The falling tone begins at a high pitch and stays level until 70% of the duration before falling to a mid pitch.

It looks as if the acoustic characteristics of Phrae Pwo Karen tones as spoken by the 35-50 group and the under-25 group have changed, i.e., the contour of the high tone in non-checked syllables has changed from high level to mid-rising.

5. Conclusion and discussion

The acoustic characteristics of the vowels and tones in the Phrae Pwo Karen spoken by three generations are shown below;

5.1 Vowels

As mentioned in the Introduction, there are nine monophthongs in Phrae Pwo Karen, without length distinction: /i e ɛ i ə a u o ɔ/. The front vowels /i e ɛ/ occur very

close to one another, i.e. with only a little difference in tongue height position (F1).

The central vowels /i ə/ and the back vowels /u o/ occur close to each other, whereas the vowel /a/ appears close to the vowel /ɔ/. See Figure 7.

In Figure 7, it is noticeable that the vowel /ɔ/ starts to move upward in the 35-50 group and obviously changes its position in the under-25 group. This situation might be explained in terms of external factors. If the overall space of Standard Thai vowels (Intajamornrak 2002)¹¹ and of Tai Yuan vowels is considered¹², the tongue height position (F1) of the vowel /ɔ/ is similar to the low front vowel /ɛ/ (see Figure 8). Contact with Standard Thai or Tai Yuan might cause the vowel /ɔ/ to shift upward. This evolution is at its clearest in data from speakers 4 and 5 of the 35-50 group (see Figure 2) and also in speakers 3, 4, and 5 of the under-25 group (see Figure 3). A noticeable evolution also affects the front vowel /ɛ/, which moves downward in speakers 1 and 4 of the 35-50 group (see Figure 2) and also in speakers 3 and 4 of the under-25 group (see Figure 3). Speakers 1 and 2 in the under-25 group are not included because their vowels /ɛ/ and /ɔ/ do not have the same tongue height position, and the vowel /ɔ/ also appears close to the vowel /a/.

¹¹ Standard Thai speakers are 45-60 years old. See the word list in Intajamornrak 2002.

¹² Only long vowels are considered because the duration of vowels in Phrae Pwo Karen is phonetically similar to the duration of long vowels in Standard Thai and Tai Yuan.

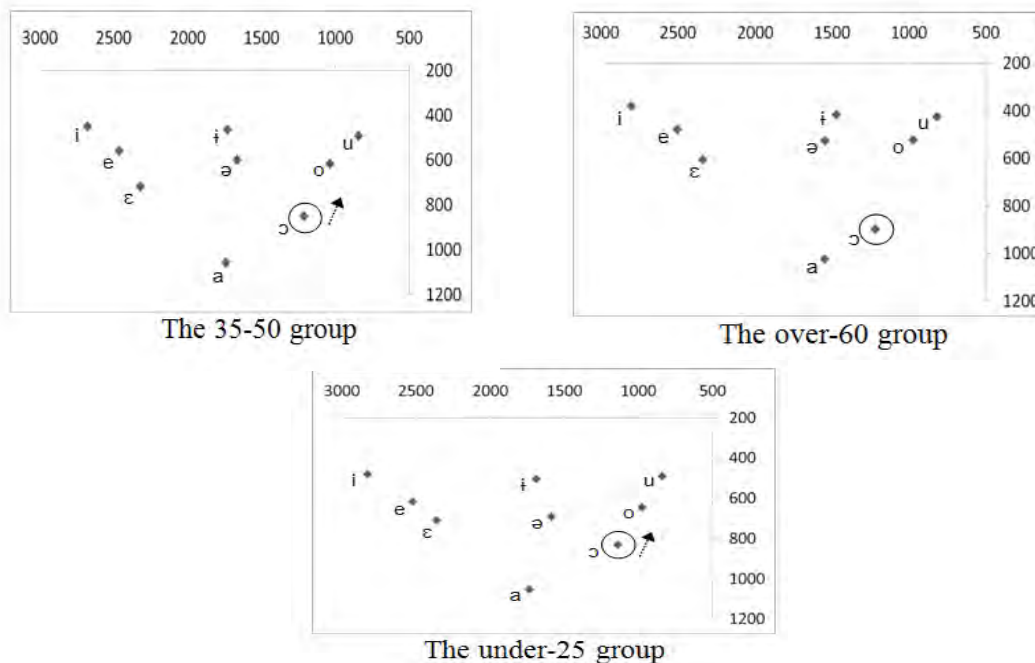


Figure 7 The overall space of Phrae Pwo Karen vowels as spoken by three groups of speakers

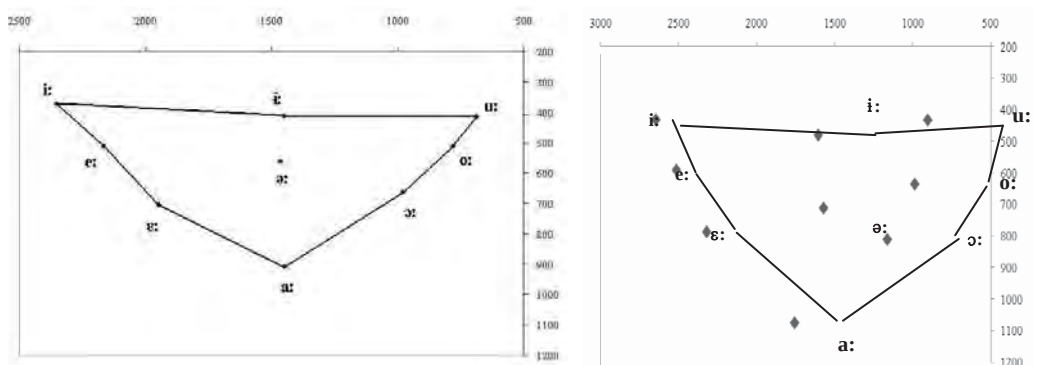


Figure 8 The overall space of Standard Thai long vowels (left) and of Tai Yuan long vowels¹³ (right)

¹³ The Tai Yuan long vowels were spoken by three female Tai Yuan (Phrae) native speakers between 35-50 years old which can be compared with Standard Thai from Intajamornrak 2002 and this age represents language in the present. The methodology for acoustical measurement of Tai Yuan vowels is the same as for Phrae Pwo Karen vowels. See the wordlist below.

/ii/ pii	‘year’	/ii/ khii	‘to be’	/uu/ khuu	‘ditch’
/ee/ thee	‘to pour’	/əə/ thəə	‘you’	/aa/ tao	‘big’
/εε/ phεε	‘raft’	/aa/ thaa	‘to paint’	/ɔɔ/ khɔɔ	‘neck’

* For the words ‘thee’ ‘thəə’ ‘too’, the informants use these words sometimes and pronounce them with Tai Yuan tones.

Figure 8 shows the overall space of Standard Thai long vowels and Tai Yuan long vowels. The position of the vowel /ɔ:/ for both Standard Thai and Tai Yuan is closer to the other back vowels /u: o:/ than the vowel /a:/. Considering the F1, the vowels /ɛ:/ and /ɔ:/ appear in the same tongue height position. Additionally, with regard to the tongue advancement, the vowel /a:/ is a central vowel, while the /ɔ:/ is a back vowel. In language-contact situations, the difference between a language spoken by a majority group and a language spoken by a minority group is one of the factors which bring about language variation and change. Karen is said to be a minority language whereas Tai Yuan is a majority language in Phrae as is Standard Thai, which is a majority language of the country. Thomason (2001: 66) explains that if one of two groups in contact is much larger than the other, the smaller group's language tends to acquire features from the larger group's language.

5.2 Tones

Recall that there are four tones in Phrae Pwo Karen, namely the high tone, mid tone, low tone and falling tone. However, based on the over-60 group, the acoustic characteristics of each tone reveal that: *the high tone* is high level starting at a high pitch and staying level to the end of the normalized duration; *the mid tone* is mid level starting at a mid pitch and then slightly falling over 75% of the duration; *the low tone* is mid-falling starting at a mid pitch and sharply falling to a low pitch; *the falling tone* is high-falling starting at a high pitch and sharply falling to a low pitch. In checked syllables, the mid tone is mid level with the same contour as those in non-checked syllables; the falling tone is high-falling with less

degree of pitch change than for those occurring in non-checked syllables.

However, the shape or contour of the four tones as spoken by each group of speakers is very similar except for the high tone. Figure 9 shows that the tonal contour changes from level to rising. In the over-60 group, it begins at a high pitch and stays level to the end point. The high tone of the 35-50 group also starts at a high pitch but continually rises which can be seen in the difference from the onset to the offset. The change is clear in the under-25 group, where it becomes a rising tone with an obvious degree of pitch change or pitch contour. Moreover, the tonal onset is lowered from a high pitch to a mid-high pitch. The acoustic characteristics of the high tone in the under-25 group and some speakers of the 35-50 group seem to be very similar to the high tone of Standard Thai. The Tai Yuan high tone is high-level (Intajamornrak 2011) which is similar to the high tone of the over-60 group. Therefore, in this situation it might not influence the Pwo Karen high tone.

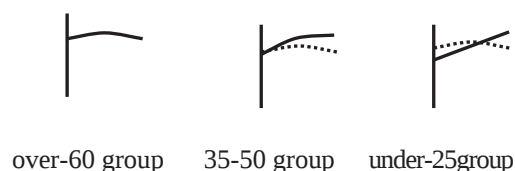


Figure 9 The high tone as spoken by three groups of speakers

Teeranon and Rungrojsuwan (2009) analyzed the high tone of Standard Thai and found that there are three variants of high tone as spoken by the under-20 group: high level [34], rising with slightly falling [322] and rising [334]. The variant that has the highest frequency distribution is rising [334] (78%) (see Figure 10). Moreover, the

pitch height of the high tone as spoken by the under-20 group moves downward from high to mid. Since the young generation studies Standard Thai in school, the high tone of Standard Thai may influence the native high tone.



Figure 10 The high tone variants in the under-20 group (adapted from Teeranon and Rungrojsuwan 2009: 41)

However, if the fundamental frequencies of Phrae Pwo Karen high tones as spoken by each speaker of the group are considered (see Appendix 2), the variation seems to occur in middle age as seen in speakers 1 and 2. It is possible that some Karen speakers at this age went to school and also work in the Khangchai Women's Weaving Group, so they have a chance to use Tai Yuan and Standard Thai frequently.

In conclusion, the acoustic characteristics of the vowels and tones as spoken by three generations suggest that Phrae Pwo Karen is in the process of changing because of the variation among the three groups. The variation and change seems to be caused by language contact with Tai Yuan and Standard Thai. As Weinreich (1968) said, the bilingual speaker's relation to the languages that come into contact must be considered in language contact situations. Here, contact with Tai Yuan and Standard Thai is a factor that brings about variation and change in Pwo Karen because of being

the majority language of Phrae and the national language. However, as the high tone shows, the situation occurs differently among the different groups of speakers. The over-60 group seems to be more familiar with Tai Yuan than Standard Thai, whereas the speakers of the 35-50 group have a chance to use Standard Thai, so tonal variation can be seen in the latter group. In the under-25 group, they are definitely familiar with Standard Thai because they use it in school and they are also very good at Tai Yuan. Therefore, their language seems to be influenced by both Standard Thai and Tai Yuan.

Acknowledgements

I would like to express my gratitude to the Thailand Research Fund (TRF) for financial support of my research through the research project on "Karen Linguistics" and Naresuan University for my travel grant. I would also like to thank Professor Dr. Theraphan Luangthongkum, the head of the project, for her valuable suggestion of my research, Dr. Pittayawat Pittayaporn and Dr. Alexis Michaud for their very useful comments. Last but not least, many thanks go to my Karen informants for their kindness and cooperation.

References

- Akharawatthanakun, P. 2002. Tonal variations and changes in a language mixture area: a case study of northeastern Thailand. *Manusya* 5.2: 30-51.
- Akharawatthanakun, P. 2009. Linguistic hybridization: a case study of Khün spoken in Nan Province, a language mixture area in northern Thailand. Paper presented at the 42nd

- International Conference on Sino Tibetan Language and Linguistics*, 2-4 November 2009, Payap University, Chiangmai, Thailand.
- Cox, F. 1999. Vowel change in Australian English. *Phonetica* 56: 1-27.
- Decker, P. and S. Mackenzie. 2000. Slept through the ice: A further look at lax vowel lowering in Canadian English. In *Sociolinguistic Dialectology*, edited by Gordon J. Easson. Special Issue, Toronto Working Papers in Linguistics Vol. 18: 1-11.
- Intajamornrak, C. 2002. The acoustic characteristics of vowels produced by Thai tracheoesophageal and normal speakers, and the perception of tracheoesophageal vowels (ลักษณะทางกลศาสตร์ของเสียงสระภาษาไทยที่ออกเสียงโดยผู้พูดที่ใช้หลอดลม-หลอดอาหาร และผู้พูดปกติ และการรับรู้เสียงสระที่ออกเสียงโดยผู้พูดที่ใช้หลอดลม-หลอดอาหาร). M.A. Thesis, Department of Linguistics, Faculty of Arts, Chulalongkorn University. (in Thai)
- Intajamornrak, C. 2011. Tonal variation and change in Tai Lue spoken in Phrae Province. Paper presented at *the Thammasat International Symposium on Language and Linguistics 2011 (TISLL 2011)*, September 23, 2001, Thammasat University, Bangkok, Thailand.
- Jacewicz, E., R. A. Fox. and J. Salmons. 2011. Vowel change across three age groups of speakers in three regional varieties of American English. *Journal of Phonetics* 39: 683-693.
- L-Thongkum, T. 1994. Phonological variation and phonological change in Mien-Yao: a result of language contact with Thai. In *Current Issues in Sino Tibetan Linguistics*, edited by Hajime Kitamura et al, pp. 915 – 927. Osaka: Museum of Ethnology.
- Lewis, M. P. (ed.). 2009. *Ethnologue: Languages of the world*. Sixteenth edition. Dallas, Tex.: SIL International. Online.
- National Archives of Thailand. 1999. *Culture, Historical development, Identity and Local wisdom of Phrae Province* (วัฒนธรรม พัฒนาการทางประวัติศาสตร์ เอกลักษณ์ และภูมิปัญญา จังหวัดแพร่). Bangkok: The organizers of His Majesty the King's 6th Cycle (72nd) Birthday Anniversary. (in Thai)
- Phillips, A. 1996. *Dialect comparison among the Pwo Karen of central Thailand*, Report II, Thammasat University and Summer Institute of Linguistics.
- Phillips, A. 2009. Lexical similarity in Pwo Karen. *PYU Working Papers in Linguistics*.
<http://ic.payap.ac.th/graduate/linguistics/papers/Pwo_Karen_Lexical_Similarity.pdf>
- Teeranon, P. 2002. Changes in phonetic characteristics of Thai Tone 4 (การเปลี่ยนแปลงสัทลักษณะของวรรณยุกต์เสียงตรีในภาษาไทย). *Journal of Arts, Silpakorn University* (วารสารอักษรศาสตร์ มหาวิทยาลัยศิลปากร) 24 (1-2): 188-209. (in Thai)

- Teeranon, P. and R. Rungrojsuwan. 2009.
Change in the Standard Thai high tone:
An Acoustic Study. *Manusya Special
Issue 17*: 34-44.
- Thomason, S. G. 2001. *Language Contact*.
Washington: Georgetown University
Press.
- Weinreich, U. 1968. *Languages in Contact*.
New York: Linguistic Circle of New
York.

Appendix 1

Table 1 The formant frequencies of Phrae Pwo Karen vowels: the over-60 group

	i	e	ɛ	ĩ	ə	ɑ	u	o	ɔ
F1	2772.81	2542.18	2354.57	1465.39	1534.90	1545.20	820.64	965.56	1223.36
F2	376.79	466.02	601.80	423.40	483.95	1022.35	423.17	491.65	905.56

Table 2 The formant frequencies of Phrae Pwo Karen vowels: the 35-50 group

	i	e	ɛ	ĩ	ə	ɑ	u	o	ɔ
F1	2688.82	2475.75	2341.73	1731.33	1675.31	1736.77	832.06	1021.55	1213.64
F2	457.02	545.79	705.07	464.44	579.72	1048.48	501.55	603.83	842.69

Table 3 The formant frequencies of Phrae Pwo Karen vowels: the under-25 group

	i	e	ɛ	ĩ	ə	ɑ	u	o	ɔ
F1	2774.39	2550.49	2529.50	1677.74	1577.62	1728.39	838.92	973.99	1126.47
F2	503.80	606.70	699.18	506.70	682.67	1069.53	496.25	626.10	818.85

Table 4 The fundamental frequencies (Hz) of Phrae Pwo Karen tones: the over-60 group

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
H	183.33	182.52	179.94	176.57	172.65	168.82	164.89	159.88	154.40	148.18	142.82
M	192.72	193.59	193.01	191.93	190.13	188.66	186.74	184.99	183.24	180.28	178.03
L	206.05	211.10	211.28	207.83	203.27	196.97	189.01	177.98	165.52	152.79	139.55
F	218.96	225.13	227.67	229.50	231.84	232.59	232.32	230.60	228.54	226.00	222.03
M?	184.98	184.45	183.77	182.76	181.86	180.74	179.01	177.11	174.80	170.44	162.85
F?	215.64	215.77	214.55	212.41	209.26	205.33	200.55	194.49	187.81	178.47	166.06

Table 5 The fundamental frequencies (Hz) of Phrae Pwo Karen tones: the 35-50 group

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
H	215.31	215.69	214.77	213.24	210.46	206.34	200.39	192.93	184.53	173.22	158.63
M	225.69	226.65	227.92	228.43	228.48	228.16	227.43	226.38	225.27	222.95	217.94
L	241.77	247.21	249.89	250.48	249.34	246.00	239.49	226.86	206.54	184.47	165.05
F	256.56	264.68	273.99	280.90	285.96	289.41	291.99	293.46	294.24	293.90	290.31
M?	226.46	226.08	225.84	226.40	226.65	226.71	226.19	225.01	222.93	220.22	209.17
F?	247.52	249.69	251.60	252.62	252.82	252.02	250.55	247.43	239.78	225.69	202.68

Table 6 The fundamental frequencies (Hz) of Phrae Pwo Karen tones: the under-25 group

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
H	230.64	231.02	230.49	229.26	227.75	225.21	221.25	214.88	204.20	190.16	176.98
M	242.78	242.63	242.06	241.63	240.90	240.46	239.62	238.66	237.43	238.08	236.13
L	259.57	264.40	267.88	269.94	269.97	267.62	261.28	248.52	224.35	199.11	189.81
F	271.48	275.41	281.58	289.39	297.91	306.40	315.50	322.58	328.79	338.24	342.89
M?	238.00	235.95	235.51	235.28	234.59	234.22	233.94	233.68	232.59	231.02	222.07
F?	262.15	262.19	263.93	265.92	266.74	267.20	267.41	265.99	262.73	256.55	241.09

H = high tone, M = mid tone, L = low tone, F = falling tone

M? = mid tone in checked syllable, F? = falling tone in check syllable

Appendix 2

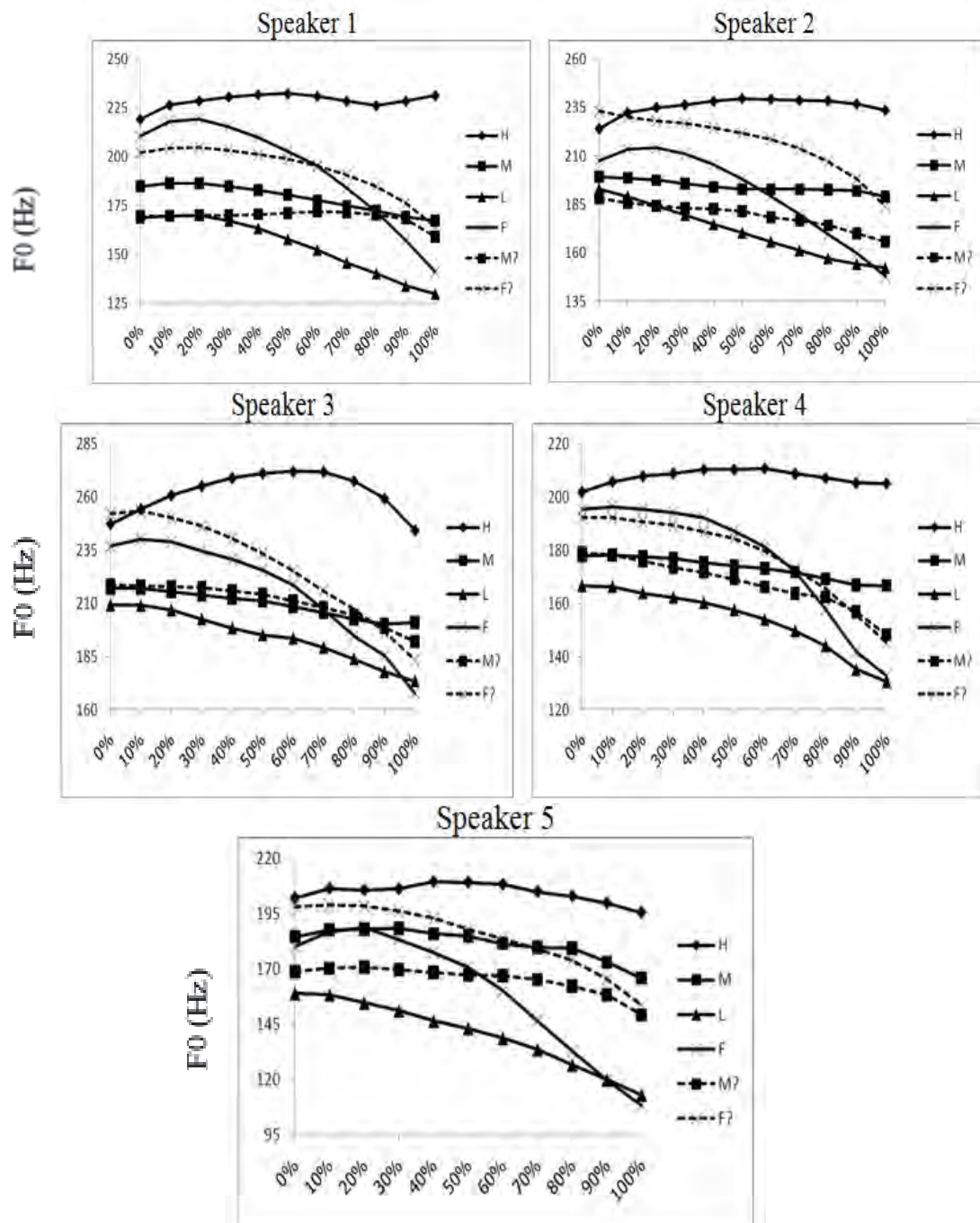


Figure 1 Fundamental frequencies of Phrae Pwo Karen tones as spoken by the over-60 group

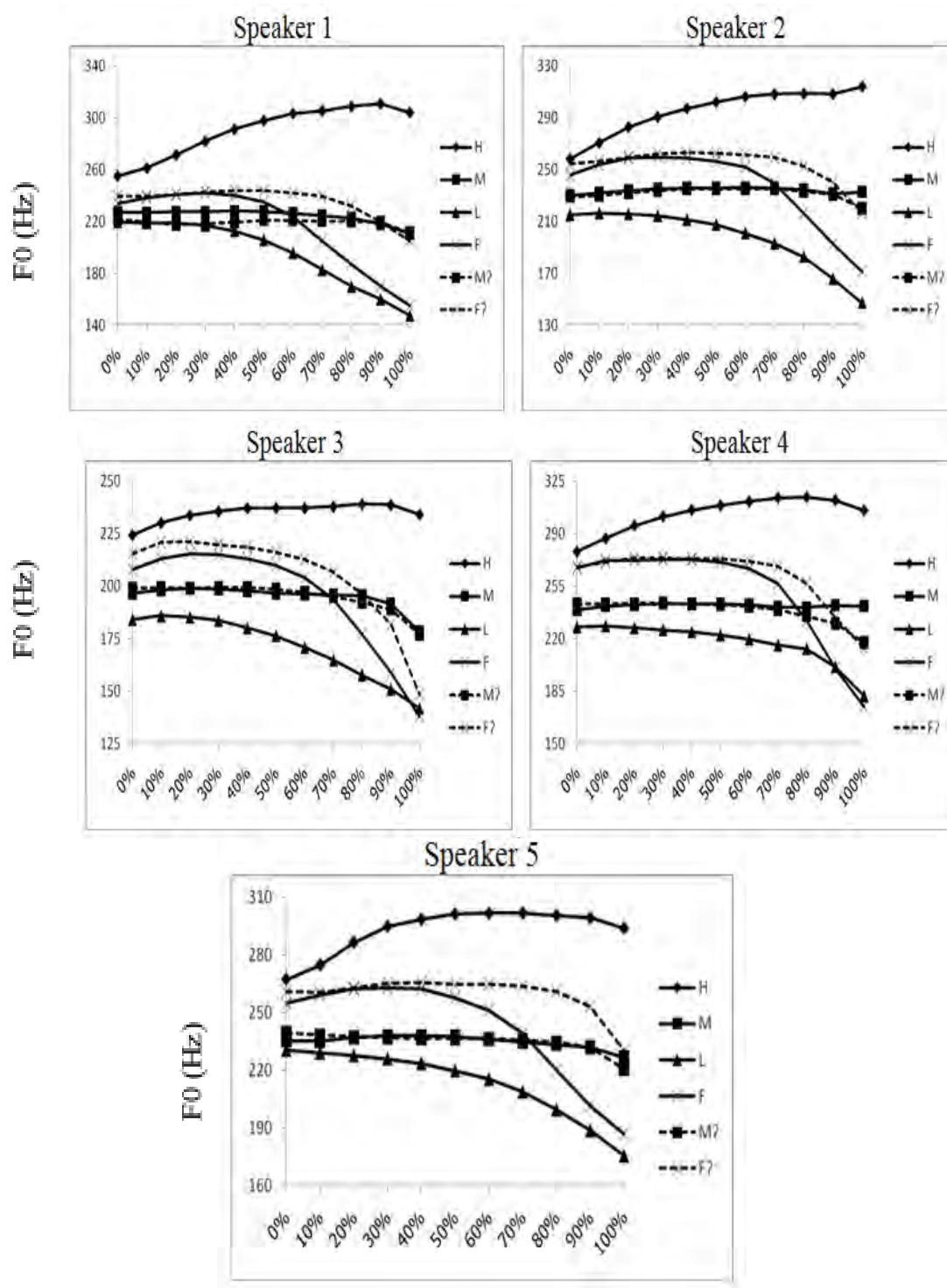


Figure 2 Fundamental frequencies of Phrae Pwo Karen tones as spoken by the 35-50 group

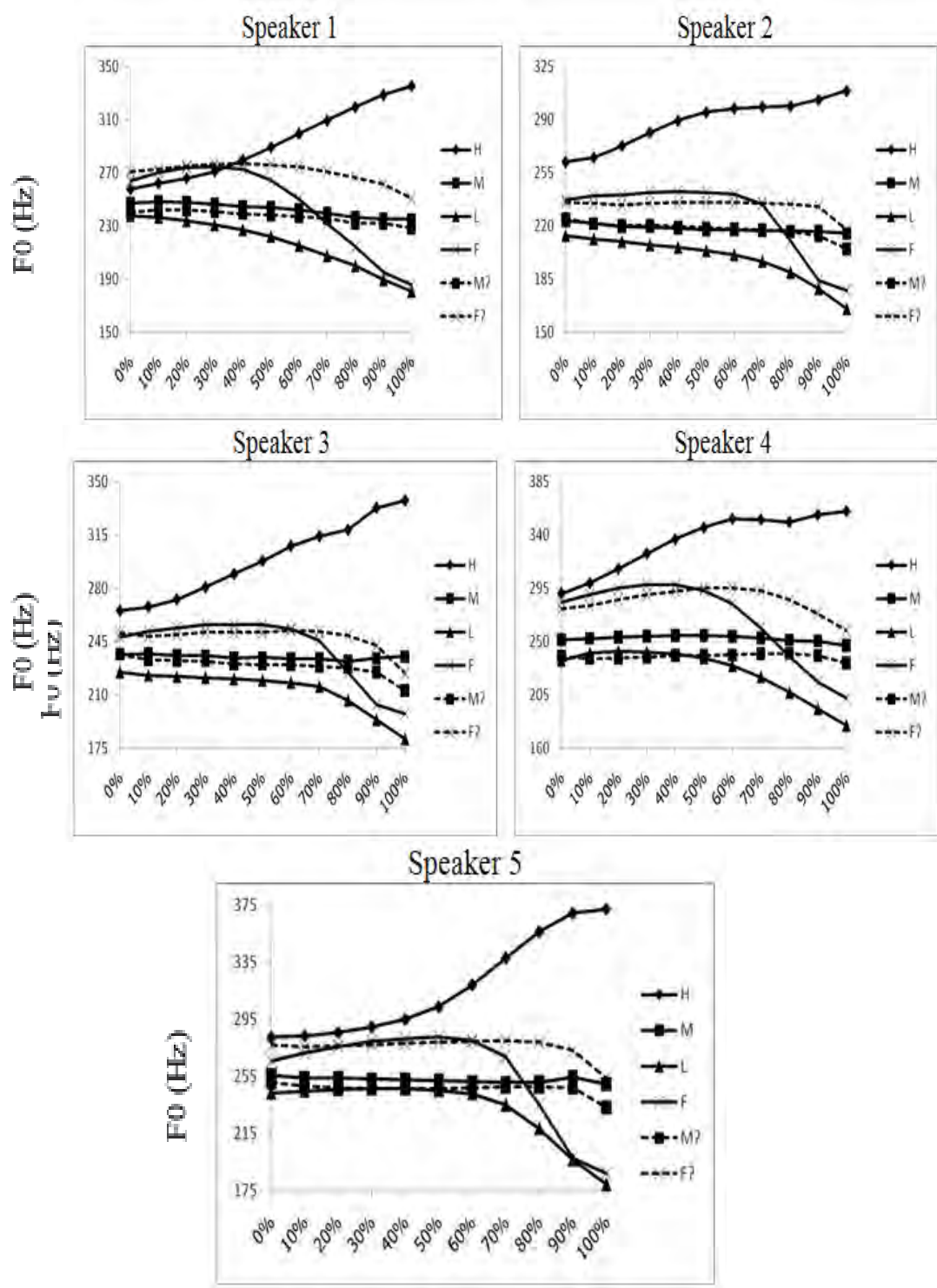


Figure 3 Fundamental frequencies of Phrae Pwo Karen tones as spoken by the under-25 group

อ. ดร. พิทยาวัฒน์ พิทยาภรณ์

Rhoticity and sonorancy of /ʁ/ in Sgaw Karen

Rhoticity and sonorancy of /ɣ/ in Sgaw Karen

Author Pittayawat Pittayaporn
Affiliation Department of Linguistics, Faculty of Arts
 Chulalongkorn University
Address 12th floor, Barommarajakumari Bldg.
 Phayathai Rd., Pathumwan, Bangkok 10330
Email Pittayawat.P@chula.ac.th

Abstract

The status of rhotic segments as a phonological class is controversial. Lindau (1985) as well as Ladefoged and Maddieson (1996), among others, argues that they do not share any phonetic properties in spite of many robust phonological similarities (Lindau 1985; Walsh Dickey 1997; and Wiese 2001). With its two contrastive rhotic sounds /r/ and /ɣ/, Sgaw Karen of the Tibeto-Burman language family is a good case study. This study examines the dorsal fricative /ɣ/ in Sgaw Karen, focusing on its phonological and phonetic patterning within the sound system. It first argues that this central fricative sound is best characterized as a rhotic sonorant since it patterns phonologically with the liquids /r/ and /l/. It then shows that acoustically /ɣ/ can be grouped with /r/ and the other sonorants as a class in respect of the amplitude of the first harmonic (H1), the amplitude of the second harmonic (H2), intensity (INT), fundamental frequency (F0) and harmonic-to-noise ratio (HNR). More specifically, /ɣ/ patterns with the sonorants in having relatively high values for all the phonetic properties. This paper thus argues that the phonological and phonetic properties of Sgaw Karen /ɣ/ are compatible with the traditional classification of rhotics as sonorants. Furthermore, it discusses the ambiguity of /ɣ/ as phonetically a fricative but phonologically a sonorant.

Rhoticity and sonorancy of /ɣ/ in Sgaw Karen

Pittayawat Pittayaporn
Chulalongkorn University
Pittayawat.P@chula.ac.th

1. Introduction

Rhotics have traditionally been recognized as a distinct class of sounds in opposition to glides, laterals, nasals, and obstruents. More specifically, they are defined as non-lateral members of the liquid class, which form a subclass of sonorants. The existence of a class of rhotics is strongly supported by the fact that its members share many aspects of phonological behavior (Lindau 1985; Hall 1993; Hall 1997; Walsh Dickey 1997; Wiese 2001). For example, rhotics tend to be the closest consonantal sounds to the vowels within consonant clusters¹, e.g. freeze [CrVC] and cart [CVrC]. In synchronic variation, rhotics of different types have been observed to alternate with each other. For example, the rhotic sounds phonemicized as /r/ in Belgian French (Wiese 2001) may be realized as [ʀ], [ʁ], [ʁ̥], [χ], [ʁ̥], [ʁ̥], [ʁ̥], or as creakiness on the vowel in different phonological environments. German (Hall 1993; Wiese 2001), European Portuguese (Jesus and Shadle 2005), and the Veracruz dialect of Spanish (Bradley 2012) also show a similar pattern of variation. This strongly suggests that rhotic sounds form a phonological class. However, from a phonetic point of view, rhotic segments appear to be a wildly heterogeneous set of sounds from different places and manners of articulation (e.g. Lindau 1985; Ladefoged and Maddieson 1996; Magnuson 2007). With respect to place, members of the rhotic class range from coronal sounds like [r], [r], and [ɹ] to dorsal sounds like [ʀ], and [ʁ]. As for manner, sounds that have been labeled rhotics include trills like [r] and [ʀ], taps like [ɹ], approximants like [ɹ], and even fricatives like [ʁ]. This paradox has led many phoneticians to question the reality of rhotics as a unified class of sound. Lindau (1985), while recognizing their phonological similarities, claims that there is no single physical property that defines rhotics as a class. Ladefoged and Maddieson (1996: 245-245) go even further and claim that the overall unity of rhotics seems to rest mostly on their “historical connections” and “the choice of the letter ‘r’ to represent them all.”

This striking disconnect between phonology and phonetics raises questions about the relationship between the two. In particular, the class of rhotics poses a problem for the view that phonological categories are phonologically unified, cf. Jakobson, et al. (1963), Chomsky and Halle (1968), and Steven et al. (1986), among others. This articulatory diversity among reported rhotics is illustrated in **Table 1**, which lists segments labeled as rhotics in Lindau (1985), and Ladefoged and Maddieson (1996), the two most influential studies on rhotics.

¹Sometimes glides intervene between the rhotic and the vowel, e.g. Spanish /grweso/ ‘thick’, but these semi-vowels are non-consonantal segments (Chomsky and Halle 1968).

Table 1 Sounds labeled as rhotics

	Ladefoged and Maddieson(1996)		Lindau(1985)	
	coronal	dorsal	coronal	dorsal
Trill	R	ʀ	r	ʀ
Tap/flap	ɾ, ɽ		ɾ, ɽ	
Approximant	ɹ, ɻ, ɹ̥		ɹ	
Fricative		ʁ	ʒ	ʁ, ɣ

According to a cross-linguistic study by Maddieson (1984), of 317 languages surveyed, only 60 languages or 18.9% have more than one rhotic segment. In most of these languages, the contrast is between rhotics that differ in terms of manner, not place of articulation. This makes Sgaw Karen, a Tibeto-Burman language spoken in the Thailand-Myanmar borderlands, particularly interesting. Most Sgaw Karen dialects in Thailand tend to have, in addition to the alveolar trill /r/, a voiced velar fricative /ɣ/. While the coronal trill /r/ is a proto-typical rhotic segment, the dorsal fricative /ɣ/ is often treated in a great number of languages as a non-rhotic segment, more specifically as an obstruent. For example, Modern Greek has a voiced obstruent /ɣ/ in contrast with a voiceless /x/ (Holton, Mackridge et al. 1997). However, in Sgaw Karen /ɣ/ does not pattern with other obstruents but seems to behave rather like liquids /l/ and /r/. If this central sound is indeed a liquid, it would have to be considered a rhotic segment making the language unique among the world's languages in having two rhotics sounds which differ both in place and manner of articulation. Therefore, Sgaw Karen /ɣ/ is an important case study in understanding the nature of rhotics as a phonological class.

This study examines the dorsal fricative /ɣ/ in Sgaw Karen, focusing on its phonological and phonetic patterning within the sound system. It first argues that the central fricative /ɣ/ is best characterized as a rhotic sonorant since it patterns phonologically with other liquids in the language, namely /r/ and /l/. It then shows that acoustically /ɣ/ can be grouped with the coronal rhotic /r/ and other sonorants as a class with respect to the amplitude of the first harmonic (H1), the amplitude of the second harmonic (H2), intensity (INT), fundamental frequency (F0) and harmonic-to-noise ratio (HNR). This paper thus argues that treating the Sgaw Karen dorsal fricative /ɣ/ as a sonorant segment in spite of its usual fricative realization is justified. In other words, the phonological and phonetic properties of /ɣ/ in this language are compatible with the traditional classification of rhotics as sonorants despite the phonetic ambiguity of the segment.

2. Rhoticity of the voiced dorsal fricative /ɣ/

Traditionally, rhotics are grouped together with laterals as liquids, a subclass of sonorants. Other types of sonorants include vowels, glides, and nasals. These types of segments are said to be more sonorous than obstruent sounds, which include stops, affricates, and fricatives. This section is a brief review of the phonological treatments of rhotics in respect of their feature specification and sonority. It

then argues that distributional facts about Sgaw Karen /ɣ/ support treating it as belonging with /r/ in the class of rhotics.

In Feature Theory (Jacobson, Fant et al. 1963; Chomsky and Halle 1968; Fant 1969), rhotics are standardly assumed to be specified as [+sonorant, -continuant, -lateral, -nasal]. These features are defined articulatorily, and entail mutual exclusivity between two classes of opposing feature values. For example, the feature [+lateral] is defined as produced with the middle section of the tongue lowered so that the airflow passes through the sides of the tongue. The sound /n/, /m/, and /ŋ/ are specified as [+nasal] because it is produced by lowering the velum to allow the airflow to go through the nasal passage. All the other sounds of English, e.g. /ɹ/, /l/, /d/, /t/, /k/, etc., on the other hand, are specified as [-nasal] because they lack nasal airflow. A sound that belongs to the former group cannot simultaneously belong to the latter. However, the articulatory diversity of rhotics has led a number of researchers to believe that the traditional featural classification is not adequate. **Table 2** gives a standard feature specification of different types of consonantal segments.

Table 2 Featural specification for consonantal sounds

	stops	fricatives	nasals	laterals	rhotics
[sonorant]	-	-	+	+	+
[continuant]	-	+	-	- ²	+
[nasal]	-	-	+	-	-
[lateral]	-	-	-	+	-

As seen above, rhotics are defined featurally as [+consonantal, +sonorant, -nasal, -lateral]. In other words, rhotics are viewed as centraloral sonorants. However, this feature system seems to display two shortcomings. First of all, Wiese (2001) argues based on articulatory descriptions of rhotics in earlier studies that [+continuant] is the only feature shared by all reported rhotic sounds (see **Table 1**). Moreover, a discrepancy arises between the diversity of rhotics and the exclusivity of the classes. More specifically, rhotics and fricatives are two distinct classes of segments because the former is [+sonorant] but the latter is [-sonorants]. This feature system thus seems to predict that rhotics cannot be fricatives and vice versa. However, rhotic segments in many languages are phonetically realized as fricatives, e.g. French [ʁ], and generally assumed to be specified as [-sonorant], cf. Wiese (2001). It is therefore problematic how such consonants should be defined in terms of their distinctive features.

One proposal is to add a new feature [rhotic] to the inventory of phonological features (Bhat 1974; Lindau 1985; Hall 1997). However, this feature does not have any clear phonetic correlate and fails to distinguish between rhotic and non-rhotic sounds in a meaningful way (Wiese 2001: 343-344). A slightly different direction is to view rhotics as belonging to the laminal place of articulation and characterize them with standard place features (Walsh Dickey 1997). As Wiese pointed out (Wiese 2001:

² The value of the feature [continuant] for lateral segments is ambiguous and seems to differ from language to language. See Hall (2007) for an introduction to this issue.

344-346), this proposal is not adequate because it makes incorrect predictions about possible rhotic sounds and fails to explain the well documented phonological processes involving rhotics, such as pre-rhotics vowels lowering and backing. Another direction is to invoke underspecification. This option is also problematic as no featural representations allowed by current theories of underspecification can account for the diverse types of rhotic contrasts attested (Wiese 2001: 347-349). For the purpose of this paper, a precise featural specification of rhotics is not immediately relevant. However, the result of this study suggests that the standard feature system is not incompatible with the reported articulatory diversity of rhotics (see §5).

A fundamental problem in capturing the classhood of rhotics is exactly what sounds should be considered rhotics. Ladefoged and Maddieson (1996) defines rhotics orthographically and includes in the rhotic class all and only sound written or transcribed with symbols derived from Latin <r, R> or Greek <ρ, Ρ>. On the other hand, Dixon (1980) defines rhotics articulatorily as central approximant segments produced at the top of the mouth. In contrast, Lindau (1985) uses shared phonological characteristics as the criteria for the membership of this class of consonants. Among the three views, Lindau (1985) assumes the most inclusive definition. The sounds that are included by Lindau but excluded by the other two authors are the fricatives [χ] and [ʒ]. In addition to these consonants, any other sounds that display r-like phonological behavior would be considered rhotics by Lindau, e.g. [ɣ] in Tiv (Zec 1995), and [ʒ] in Mandarin (Duanmu 2000)³. Because most later acoustic and phonological investigations of rhotics (e.g. Inouye 1995; Walsh Dickey 1997; Wiese 2001; Webb 2002; Magnuson 2007) are built upon Lindau's study, this paper adopts her basic view that rhotic segments are identified by their phonological distribution.

Sgaw Karen is an ideal case study of rhotics as a phonological class because the shared phonological and phonetic properties between the two rhotics in this language are certainly not due to etymology or orthography. Firstly, /ɣ/ and /r/ in Sgaw Karen do not have the same etymological source. The velar fricative /ɣ/ occurs in both native as well as non-native morphemes, contrasting with the alveolar trill /r/, which is found only in a small set of non-native morphemes. In addition, morphemes having the two sounds are not historically connected. While the dorsal /ɣ/ is the regular reflex of *r- in Proto-Tibeto-Burman and Proto-Karen (Jones 1960; Matisoff 2003; L-Thongkum ms.), the coronal /r/ is found only in loanwords and other morphemes of unclear origin⁴. Secondly, only one of the two rhotics is represented by Latin <r>. The rhotics are represented with totally different letters both in Roman-based (Gilmore 1898; Jesuit Social Services Centre 2005) and Burmese-based scripts (Gilmore 1898). The coronal /r/ is expectedly written with the Roman <r>, and the Burmese <ရ> (originally representing /r/). In contrast, the dorsal /ɣ/ is represented by the Roman <g> or <gh>, and Burmese <ဂ> (for /g/). Given the etymological and orthographical differences, the shared phonological and phonetic properties of /ɣ/ and /r/ must be due to their status as rhotics.

³ In contrast to many researchers, e.g. Karlgren (1915-26) and Dong (1958), who describes this sound as [ʒ], Duanmu (2000) transcribes this sound as [ɹ] but mentions that it is pronounced with some frication.

⁴ Examples of borrowing with /r/ include /təru³/ 'Chinese' from Burmese /təjou[?]/ <tarut>, and /səra:¹/ 'teacher' from Burmese /s'həjə/ <ch'ara>, /ri³/ 'to iron' from Thai /ri:t³/, and /ro¹/ 'fence' from Thai /ruə⁴/ . Words with unclear etymology include /pri³/ 'small' and /ru¹/ 'mane'.

One of the most important diagnostics of rhoticity is phonotactics. Rhotic segments occur closest to the vocalic nucleus within a tautosyllabic consonant cluster, forming CrV- or -VrC sequences (Lindau 1985; Walsh Dickey 1997; Wiese 2001). However, this phonotactic property is in fact not unique to rhotics but also true of laterals. Cross-linguistically, both types of liquids “facilitate” the formation of consonant clusters (Walsh Dickey 1997; Proctor 2009). Therefore, to identify a given central segment as a rhotic sound is essentially showing its liquid status by examining its distribution with respect to consonant clusters. For example, the German uvular trill /ʀ/, often phonetically realized as [ʁ] or [χ], is considered rhotic not because it is produced as a trill but because it shows phonological properties characteristic of rhotics (cf. Hall 1993). For example, /ʀ/ can occur in a CrV- sequence as in *Drang* /dʁaŋ/ ‘impulse’ and *Attrappe* /a.tʁa.pə/, and in a -VrC sequence as in *Arm* /aʁm/ ‘arm’, and *Zorn* /t͡sɔʁn/ ‘anger’. Similarly, Spanish [r] is considered a rhotic sound because it can occur in a CrV- sequence as in *preso/preso* /ˈpɾeso/ ‘prisoner’, and a -VrC as in *arte/arte* /ˈarte/ ‘art’ (cf. Harris 1983; Inouye 1995). In contrast, American English [ɹ], though a flap, is not considered rhotic because it does not occur in clusters. Rather it is an allophone of /d/ and /t/ which are clearly not rhotic segments (cf. Inouye 1995).

Given that rhotics are identified by their position within the syllable, their phonological characterization must therefore be related to sonority, an abstract phonological notion used to characterize segment sequencing within syllables. As Wiese (2001) convincingly argues, rhotics can be characterized in terms of sonority. The sonority hierarchy defines a ranking of sound classes on the basis of their sequencing within the syllable (Clements 1990; Zec 1995; Parker 2002). While the most sonorous segments typically function as syllable peaks, the least sonorous ones are preferably confined to syllable margins. One standard version (Clements 1990) of the sonority hierarchy is given in **Table 3**.

Table 3 Sonority hierarchy according to Clements (1990)

	Sonority index
Vowels	5
Glides	4
Liquids	3
Nasals	2
Obstruents	1

According to the above scale, obstruents are the least sonorous types of segments with a sonority index of 1. Sonorants, on the other hand, are distributed along the rest of the scale. At the top end of the hierarchy, vowels constitute the most sonorous class of segments with a sonority index of 5. Most crucially, liquid segments including rhotics occupy the middle point, having a sonority index of 3. From this point of view, rhotics are associated with the CrV- and the -VrC patterns because their sonority index is higher than other consonantal sounds but lower than vowels and glides (Wiese 2001). The crucial distribution of rhotics thus follows directly from the fact that liquids are intermediate between the other types of consonantal segments (i.e. obstruents and nasals) and all the non-

consonantal segments (i.e. glides and vowels). The results of this study suggest that with respect to sonority, the Sgaw Karen /ɣ/ resembles sonorants rather than obstruents (see §4).

Having reviewed the phonological treatments of rhotics, it is now possible to demonstrate that the voiced dorsal fricative /ɣ/ in Sgaw Karen is indeed a rhotic sound. As the rhoticity of segments depends on their distribution within each language, it is important to first examine the sound inventory of Sgaw Karen. The description is based on the speech of a 26-year-old female native speaker. In spite of having lived in Bangkok for seven years, her sound system is identical to that of the other speakers in her native Sut Huay Na village, Mae Ho sub-district, Mae Sariang district, Mae Hong Son Province, Thailand. Although located in Mae Ho sub-district, this Sgaw Karen variety is quite different from the one spoken at the center of the sub-district but appears to be the same dialect as that spoken in Sop Moei district. In the Sut Huay Na dialect, there are altogether 24 consonants as summarized in the consonant inventory in **Table 4**. Among these, two central consonants display distribution expected of rhotics, namely the velar fricative /ɣ/ and its more proto-typical counterpart, the alveolar trill /r/.

Table 4 Inventory of Sgaw Karen consonants at Sut Huay Na

	labial	alveolar	palatal	Velar	Glottal
Voiceless stops	p	t	c	k	ʔ
Aspirated stops	p ^h	t ^h	c ^h	k ^h	
Voiced stops	b	d			
Fricatives	(v)	s z		x	h
Nasals	m	n	ɲ	ŋ	
Liquids		l r		ɣ	
Glides	w		(j)		

From the above inventory, three observations regarding the consonant system are noteworthy. Firstly, the two palatal obstruents /c/ and /c^h/ are best analyzed as stops but they are phonetically realized with a slight affrication, e.g. [tɕáʔ]⁵ ‘to ask’ and [tɕ^hā] ‘painful’. Phonologically, there is no evidence that they behave differently from other stops. Secondly, the voiced labial fricative [v] should perhaps be analyzed phonologically as a glide /w/ because [v] and [w] are in complementary distribution. While [v] occurs only initially, e.g. [vā] ‘husband’, [w] only occurs in Cw- clusters, e.g. [k^hwā] ‘male (for animal)’. Thirdly, the status of [j] is unclear as it is found in only one word [jàʔ] ‘blanket’. However, native speakers insist that substituting it with other sounds is not possible⁶. Notice that the

⁵ This dialect of Sgaw Karen has 4 lexical tones: 1) level e.g. /nɔ̃/ ‘sesame’, 2) low e.g. /nɔ̀/ ‘slow’, 3) high glottalized e.g. /nɔ̃ʔ/ ‘elder sister’ and 4) low glottalized, e.g. /nɔ̀ʔ/ ‘grass’.

⁶ Sut Huay Na /z/ corresponds to /j/ in some other dialects, e.g. Mae Chaem dialect studied by Ratanakul(1986). Native speakers often tolerate variation between [z] and [j], but curiously they do not accept *[za⁴] for ‘blanket’. This puzzle may indicate an on-going (re-)introduction of /j-/ into this dialect.

voiced velar fricative /ɣ/ is classified with the liquid /l/ and /r/ rather than with the voiced alveolar fricative /z/. Of crucial relevance is the status of the voiced velar fricative /ɣ/ as liquid. Phonetically, this dorsal segment is pronounced as [ɣ] or [ɣ̥] with varying degrees of frication. In word-initial position, the segment is often produced with a clearly audible friction represented by a period of aperiodic waves in Figure 1. The frication is particularly strong when /ɣ/ is the second member of a consonant cluster as seen in Figure 2. In contrast, in other positions the frication is not as robust but still observable as in Figure 3. Although the actual amount and intensity of frication seems to vary across speakers and across phonetic environments, it is clear that this dorsal segment is phonetically a fricative sound.

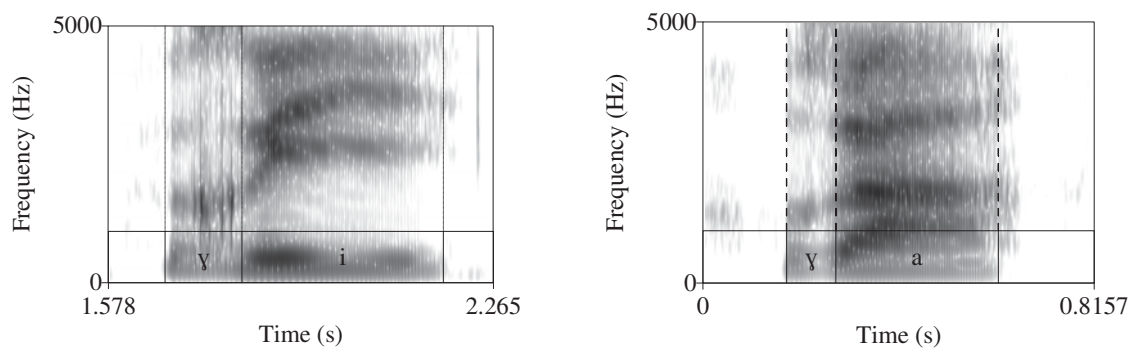


Figure 1 Spectrograms of /ɣi²/ 'root' and /ɣa²/ 'to hurry' pronounced in isolation by a female speaker

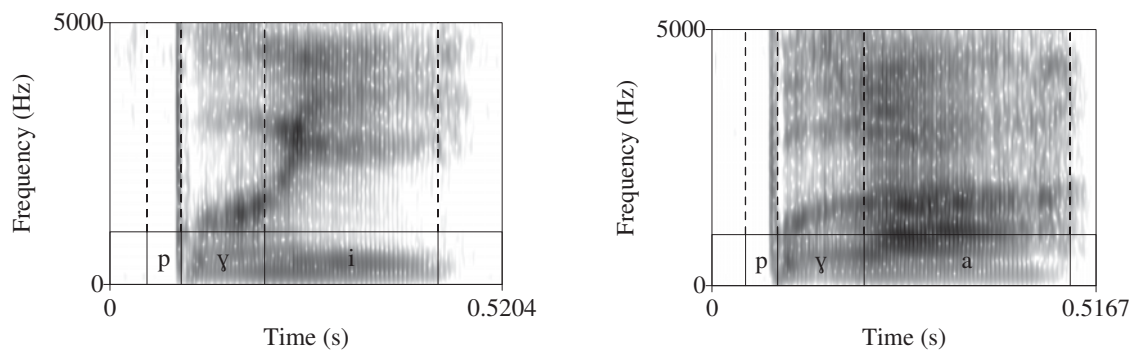


Figure 2 Spectrograms of /pɣi²/ 'to melt' and /pɣa²/ 'old' pronounced in isolation by a female speaker

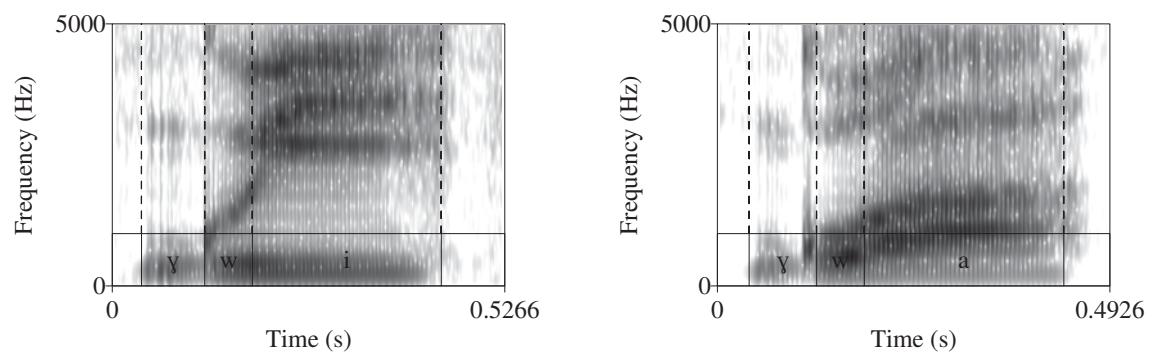


Figure 3 Spectrograms of /ɣwi¹/ ‘bag’ and /ɣwa¹/ ‘white’ pronounced in isolation by a female speaker

In spite of its phonetic realization, /ɣ/ phonologically however patterns with the prototypical liquids /l/ and /r/, occurring in onset consonant clusters. In Sgaw Karen, only open syllables of types CV and CCV are allowed. In addition to /ɣ/, only the glide /w/ and the liquid /l/ and /r/ can occur as the second member of an onset cluster. Attested clusters in this Karen dialect are exemplified in **Table 5**.

Table 5 Attested onset clusters in Sgaw Karen

	/-w-/	/-l-/	/-r-/	/-ɣ-/
/p-/		pla ² 'to let go' ple ³ 'to release' ple ¹ 'tongue'	pri ³ 'remnant' pru ³ 'to call(chicken)' pro ³ 'to gush out'	pɣo ¹ 'sacred' pɣe ¹ 'full'
/p ^h -/		p ^h ly ⁴ 'round' p ^h le ⁴ 'quick' p ^h lo ⁴ 'to wear'	p ^h ro ³ 'to come off'	p ^h yu ³ 'to slip' p ^h ya ³ 'fishnest' p ^h yi ¹ 'light'
/b-/		blɣ ⁴ 'ant hill' blo ¹ 'kind of bamboo'		bye ⁴ 'to broadcast' bya ¹ 'tame' byo ³ 'to vomit'
/t-/	twi ² 'to drag'			
/t ^h -/	t ^h wi ⁴ 'dog'			
/d-/	dwi ³ 'to point ^v ₃ '			
/c-/	cwa ³ 'to crawl' cwi ¹ 'horn beetle'			cɣo ¹ 'cymbal' cɣo ¹ 'to catch with fingertips'
/c ^h -/	c ^h we ⁴ 'crab' c ^h we ¹ 'to step aside'			c ^h yo ⁴ 'tight, secure' c ^h yo ³ 'to convulse'
/k-/	kwa ³ 'axe' kwi ³ 'pond' kwe ³ 'wasp'	kle ³ 'to mix' kla ² 'to bite' kle ⁴ 'unlucky'	kre ¹ 'door' krɔ ³ 'pod (of knife)' kre ³ 'kind of cricket'	
/k ^h -/	k ^h wi ³ 'to throw'	k ^h ly ⁴ 'fig' k ^h lo ¹ 'to boil' k ^h li ³ 'turtle'	k ^h ri ¹ 'boat' k ^h ri ³ 'to pile up' k ^h ro ⁴ 'shellfish'	
/s-/	swi ³ '(flower) to shut' swa ¹ '(stair) step' swe ⁴ 'to cut up'			
/l-/	lwi ³ 'cabbage' lwa ⁴ 'to saw'			
/m-/		mle ³ 'promiscuous' mlɔ ¹ '(elephant) trunk'		
/n-/	nwi ³ 'seven' nwe ⁴ 'taro' nwi ¹ 'tuft'			
/ɣ-/	ɣwa ¹ 'white' ɣwi ² 'delicious' ɣwa ⁴ 'to duck'			

To summarize, Sgaw Karen /ɣ/ patterns with the liquids /r/ and /l/, indicating that it is phonologically a liquid segment in spite of its fricative realization. As /ɣ/ clearly does not have a lateral articulation, it must be considered a central sound. This means that /ɣ/ must be considered a rhotic segment. Furthermore, as liquids form a subclass of sonorants, the rhoticity of Sgaw Karen /ɣ/ thus recalls the very discrepancy between the phonological status and the phonetic realization of rhotic segments. The following section reports on an acoustic study of the dorsal fricative /ɣ/ in Sgaw Karen, which shows that the segment does in fact show phonetic properties characteristic of sonorants in spite of its realization as a fricative.

3. Sonorancy of the dorsal fricative /ɣ/

Although most researchers define features in terms of production, the articulatory properties of the features are assumed to correspond to some acoustic correlates (Jacobson, Fant et al. 1963; Stevens and Blumstein 1981; Keating 1990). For sonorants, amplitude of the first harmonic (H1), amplitude of the second harmonic (H2), and intensity (INT) seem to be the acoustic properties that distinguish them from obstruents. According to Steven and Keyser (1989), [+sonorant] segments are acoustically characterized by amplitude at lower frequencies that are not substantially different from those of the surrounding vowels. In addition, Parker (2002)'s claim on the phonetic basis of sonority entails that sonorants, being higher in the sonority hierarchy, have a higher intensity than obstruents. An acoustic study was thus carried out to determine whether phonologically characterizing the dorsal fricative /ɣ/ as a sonorant segment is phonetically-grounded. The goal of this study was to test the hypothesis that one or more acoustic correlates group /ɣ/ with other sonorant segments. More specifically, it was hypothesized that /ɣ/ pattern with /r/ and other sonorants rather than obstruents with respect to amplitude of the first harmonic (H1), amplitude of the second harmonic (H2), and also intensity (INT).

3.1 Subjects and recordings

Ten native speakers (5 female, and 5 male) were recorded during a field trip to Mae Hong Son Province in October 2010. All of the subjects had been born and grown up in Sut Huay Na village. In addition to Sgaw Karen, all of them are fluent speakers of Central Thai, the national language. Some may also speak Northern Thai, the lingua franca of the area. The subjects all have basic familiarity with Romei, the roman-based script of Sgaw Karen used by Christian missionaries (Jesuit Social Services Centre 2005). Given the linguistic and literacy background, the recording sessions were carried out in Thai, but the word list presented to the speakers was written in a simplified Romei script⁷.

⁷ Because this dialect of Sgaw Karen has fewer tones than the dialect on which the standard Romei script is based, certain tone diacritics were not used.

The word list consisted of 112 monosyllables with simple alveolar and velar onsets. All content monosyllabic words whose initial consonant is from the following set of consonants: /s, z, n, l, r, x, ɲ, ʎ/ were included. These consonants are fricative, nasal, and liquid segments from the alveolar and the velar series. While the fricatives represent the class of obstruents, the nasals and the liquids represent the class of sonorants. The labial, the palatal, and the glottal series were excluded because Sgaw Karen does not have the liquid sounds from these places of articulation. **Table 6** gives a Romei transcription for the eight segments included in this study.

Table 6 Consonants included in this study in IPA and in Romei transcriptions

IPA	Romei
/r/	r
/ʎ/	g
/l/	l
/n/	n
/ɲ/	ng
/z/	y
/s/	s
/x/	q

The words were arranged in a randomized order and written in a simplified version of the Romei script (see Appendix 1). The subjects were asked to read each word in the frame sentence /c̣ɣ té ...kʰl̥ kʰl̥/ ‘I say ... quickly’ three times at a normal speech rate with short pauses between repetitions. The subjects were recorded at a sampling frequency of 48 kHz, using a Marantz PMD660 portable solid-state recorder with a head-mounted unidirectional microphone Shure SM10A.

3.2 Data processing

The recordings were analyzed using Praat 5.1.42 (Boersma and Weenink 2010). Abrupt changes in the spectrogram were used as cue to the onset and endpoint of the consonants. For the nasals /n/, /ɲ/ and the lateral /l/, the consonantal duration is signaled by markedly weaker energy from its surrounding vowels. The trill /r/, in contrast, is characterized by intermittent closures. For the fricatives /s/, /x/ and /z/, the consonantal duration is signaled by aperiodic waveforms. Similarly, the fricative /ʎ/ is also characterized by aperiodic waveforms but sometimes preceded by a period of silence. In some speakers, however, the voiced dorsal fricative appears less aperiodic but is still clearly distinguished from the surrounding vowels by its low energy similar to the liquid /l/. Average intensity (INT), fundamental frequency (F0), amplitude of the first harmonic (H1), amplitude of the second harmonic (H2), and harmonics-to-noise ratio (HNR) were measured. F0, H1, and H2 values of the voiceless

segments are assumed to be 0⁸. While INT is the acoustic correlate of sonority (Parker 2002), H1 and H2 are the correlates of sonorancy (Stevens and Keyser 1989). F0 and HNR have not been claimed to be directly related to either sonority or sonorancy but they are also included as potential candidates. Before the measurements were conducted, the peak intensity of the sound files was scaled to 0.99 in order to normalize for variation that may have arisen naturally in the production by the subjects.

3.3 Statistical analysis

To test the hypothesis that the dorsal group /ɣ/ pattern with other sonorant segments with respect to some phonetic properties, a series of discriminant analyses were performed on the results of the acoustic measurements. Discriminant analysis (DA) is a method used to model the linear relationship between a dependent variable and one or more predictor variables⁹. It can thus be employed to test predictions about grouping theories, and to investigate the differences between groups based on attributes of the cases. The method assumes that each predictor variable is normally distributed and that the samples are correctly assigned to the dependent categories at the initial classification. Furthermore, it assumes that the categories are natural, mutually exclusive, and collectively exhaustive.

A discriminant analysis was performed separately on the sound tokens from each individual subjects. The predictor variables in the discriminant analyses are the values of intensity (INT), first-harmonic amplitude (H1), second-harmonic amplitude (H2), and fundamental frequency (F0) of initial consonants obtained by the procedures described in §3.1. The analysis was designed to test whether classifying the two Sgaw Karen rhotics /ɣ/ as sonorants is phonetically supported. To test the hypothesis, two dependent categories were set up based on the sonorant/obstruent distinction. The sound tokens were initially classified into two dependent groups based on the sonorancy of their initial consonants. The first one consists of obstruent-initial cases including those with fricative onsets /s/, /x/ and /z/. In contrast, the second one is the class of sonorant-initial cases, which consists of sound tokens having initial nasals, laterals, and rhotics: /n/, /ŋ/, /l/, /r/, and crucially /ɣ/. The classification is summarized in **Table 7**.

Table 7 Classification of cases into 2 groups based on the sonorancy of their onsets

Group	Segments
1	/z/, /s/, /x/
2	/n/, /ŋ/, /l/, /r/, /ɣ/

⁸ The Praat script used to measure these segments returned F0, H1, and H2 values of the voiceless segments /x/ and /s/ but manual measurements shows that they are false. The values for the voiced segments were also randomly re-checked manually.

⁹ See Burns and Burns (2008) for introduction to this technique.

For the hypothesis to be verified, two predictions must be confirmed. First, it must be shown that the initial classification in which /ɣ/ is grouped with the sonorants is statistically supported. Second, some acoustic variables associated with the sonorant segments must be significant predictors in assigning the cases into the two dependent groups. The acoustic variables used as predictor variables in the discriminant analysis were the H1, H2, INT, and F0 values described in §3.2. The results of the study described in the following section confirm both predictions, thus supporting the hypothesis. In addition, two additional discriminant analyses were performed to test whether the sonorancy-based analyses is the only possible grouping of the consonants. The first alternative classification groups the consonants into three groups with the dorsal fricative /ɣ/ in its own groups. As for the second alternative, the consonants are divided into two groups according to their manner of articulation with /ɣ/ in the same group as the other fricatives. In contrast, the second alternative classification groups the consonants into three groups with the dorsal fricative /ɣ/ in its own group. The results indicate that only the manner-based alternative is also statistically supported, pointing to the phonetic ambiguity of /ɣ/.

4. Results

As the discriminant analysis was conducted on each of the ten subjects separately, the results show clearly that there is a certain degree of variation among speakers. However, all agree in having H1, H2, and INT as important predictor variables in discriminating among the consonants. For an overall picture of the phonetics of consonants, this section first gives the results from the acoustic measurements before proceeding to present the results from the sonorancy-based discriminant analysis compared with those from the alternative classifications.

4.1 Acoustic measurements

A quick examination of the mean INT, H1, H2, F0, INT, and HNR values of all the subjects combined is particularly revealing. Overall, the sonorants including /ɣ/ tend to show larger values than the obstruents for all acoustic properties. The only exception is /r/, which shows rather low HNR. The mean values of the acoustic properties for all speakers combined are given in Figure 4, Figure 5, Figure 6, Figure 7, and Figure 8. Acoustic results for each subject are given in Appendix 2.

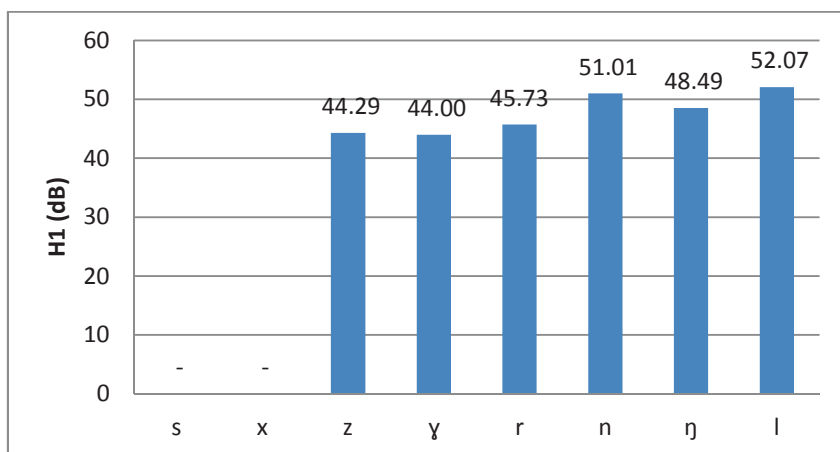


Figure 4 Mean first-harmonic amplitude (H1) for each consonant

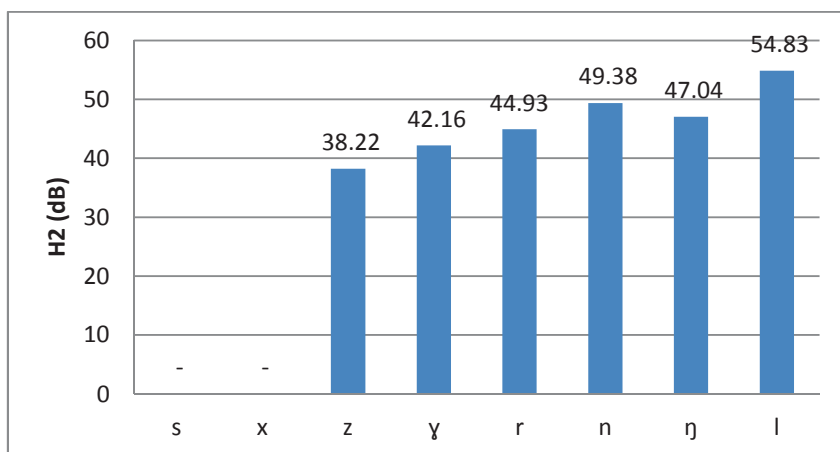


Figure 5 Mean second-harmonic amplitude (H2) for each consonant

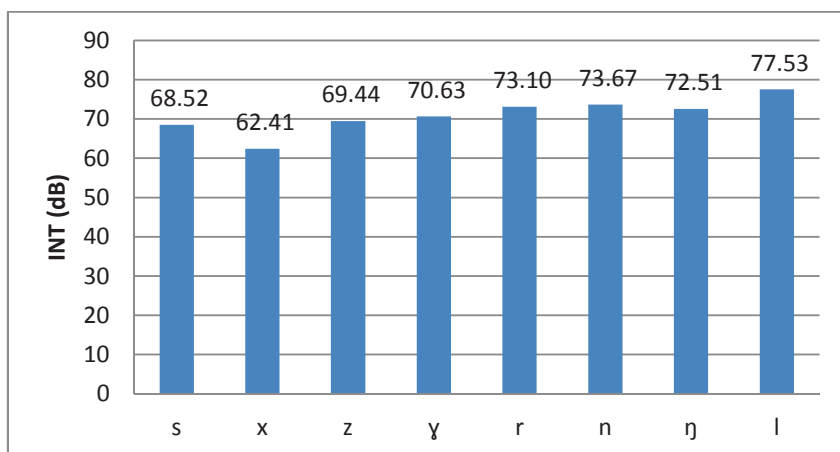


Figure 6 Mean intensity (INT) for each consonant

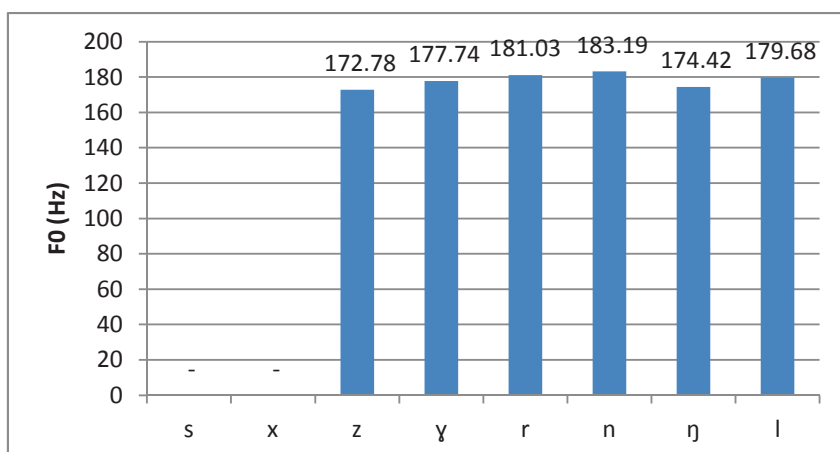


Figure 7 Mean fundamental frequency (F0) for each consonant

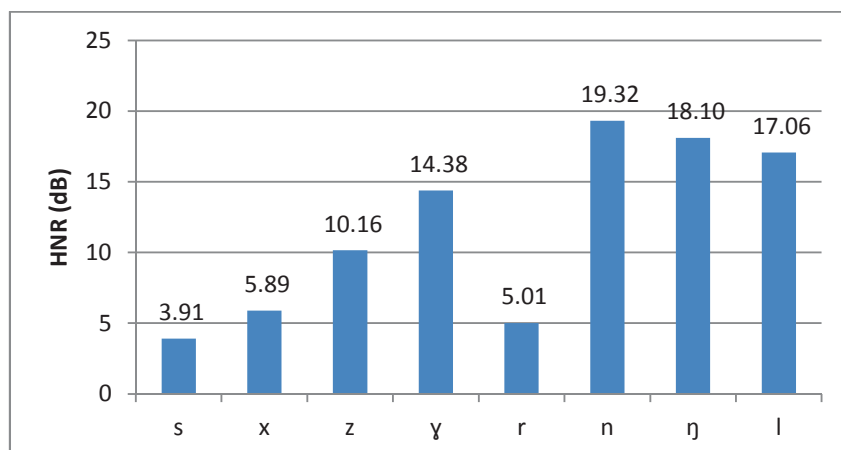


Figure 8 Mean harmonics-to-noise ratio for each consonant

Some important observations can be made regarding the acoustic measurements. Firstly, the two voiced fricatives /z/ and /ɣ/ and the trill /r/ seem to fall somewhere between the voiceless fricative /x/ and /s/ on the one hand, and the prototypical sonorants /n/, /ŋ/, and /l/ on the other, with respect to INT, H1, H2, and F0. Secondly, the fricative /ɣ/ is unsurprisingly highly similar to the voiced fricative /z/ with respect to most acoustic parameters. The only exception is HNR, for which /ɣ/ has noticeably higher values. As HNR correlates with the amount of noise, the relatively high value for /ɣ/ means that the segment is characterized with less friction than other fricative sounds. The obstruent /z/, on the other hand, is similar to the other fricatives, showing a relatively large amount of friction noise. These three observations suggest that /ɣ/ is fairly intermediate between prototypical sonorants and the prototypical obstruents.

4.2 Validity of the sonorancy-based classification

The hypothesis that /ɣ/ pattern with /r/ and other sonorants with respect to some acoustic properties predicts that the initial sonorancy-based classification is statistically supported. One way of testing whether the initial classification of cases into the sonorant and the obstruent groups is statistically supported is by using the group means of the predictor variables to describe each group. In the discriminant analysis technique, predictor variables are combined into a single composite function called the discriminant score. The mean of the discriminant score for a dependent category is in turn called the group centroid. The further apart the centroids, the more distinguishable the dependent categories. Therefore, for the initial classification to be confirmed a significant difference between the functions at group centroids is expected. **Table 8** displays the functions at group centroids for the sonorant and obstruent groups for each subject.

Table 8 Functions at group centroids for each subject (sonorancy-based classification)

Subjects	Functions at Group Centroids	
	Group 1 obstruents	Group 2 sonorants
BP	-1.940	1.827
CP	-1.597	1.304
KS	-1.647	1.464
OS	-1.506	1.402
PS	-2.330	2.206
PT	-1.687	1.481
SC	-1.545	1.215
ST	-2.211	1.892
TM	-2.227	1.702
WM	-2.023	1.686

Unstandardized canonical discriminant functions evaluated at group means

For all subjects, the function at the centroid for the first and the second groups are noticeably different. For example, the functions at group centroids for subject BP are -1.940 for the first group and 1.827 for the second group. Similarly, the functions at the centroid for subject PS are -1.687 and 1.481 for the first and second groups. The fact that the difference between the functions is very large indicates that the two groups are well-distinguished from each other. This means that the initial classification of cases in terms of their sonorancy is statistically supported, confirming the prediction that grouping the cases with initial /y/ with cases with /r/ and other sonorant onsets is statistically supported.

Another way of confirming the initial classification is by examining the accuracy in predicting the group membership of each case. The discriminant analysis technique posits a linear equation that predicts which group each case belongs to. If the two groups are well differentiated from each other, the predictive accuracy of classification, or the hit ratio, is relatively high. An acceptable hit ratio must be larger than that due to chance. Typically, a hit ratio 25% larger than that due to chance is considered acceptable. When the data from all the subjects are combined, 1491 and 1726 cases are initially assigned to group 1 and group 2 respectively. This means that the acceptable hit ratios are approximately 71.35% and 78.62% respectively. According to the cross-validated classification results, the two dependent groups both have high hit ratios for all subjects. **Table 9** gives the classification results for each subject.

Table 9 Classification result for each subject (sonorancy-based classification)

Subjects	Predictive accuracy (%)		
	Group 1 obstruents	Group 2 sonorants	Weighted average
BP	93.2 (N = 162)	91.9 (N = 172)	92.51 (N = 334)
CP	75.4 (N = 142)	93.1 (N = 174)	85.13 (N = 316)
KS	83.6 (N = 152)	97.1 (N = 171)	90.1 (N = 323)
OS	75.9 (N = 162)	98.9 (N = 174)	87.80 (N = 336)
PS	91.3 (N = 160)	100 (N = 169)	95.74 (N = 329)
PT	78.1 (N = 151)	93.6 (N = 172)	86.38 (N = 323)
SC	72.1 (N = 136)	96.5 (N = 173)	85.76 (N = 309)
ST	91.2 (N = 148)	97.1 (N = 173)	94.39 (N = 321)
TM	94.0 (N = 133)	94.8 (N = 174)	94.46 (N = 307)
WM	92.4 (N = 145)	97.1 (N = 174)	94.98 (N = 319)

To take subject ST as an example, the two dependent groups both have equally high hit ratios, 91.2% and 97.1% respectively. These two values are much larger than the acceptable hit ratios. The average accuracy, i.e. the weighted mean of the hit ratios of the two groups, is 94.36%, which is also very high. This indicates that the boundary between the two categories is quite sharp. Subjects TM and WM are the other subjects for whom the hit ratios are comparable between the two independent groups. In contrast, for subject CP, the hit ratios for the obstruent group and sonorant groups are 75.4% and 93.1% respectively, yielding an average accuracy of 85.13%. Overall, the hit ratio for the obstruent group is considerably lower than the sonorant group, suggesting that the boundary between the two is slightly fuzzy. In particular, the lower accuracy for group 1 means that a number of obstruent-initial cases are wrongly predicted to belong to group 2 or the sonorant group. However, the fact that the hit ratios of the two groups are still larger than the acceptable values suggests that overall the two groups are still clearly separated. In addition to CP, most subjects including KS, OS, PT, SC, and TM also show such overlap between the two groups. BP is the only subject for which the predictive accuracy is larger for the obstruent group than the sonorant group. Overall, all the hit ratios are larger than the acceptable ratios.

Therefore, the result of the statistical analysis shows that the classification of tokens into two groups with cases with initial /y/ in the same group with those with /r/ and the other sonorants is

statistically supported. Firstly, the validity of the sonorancy-based classification is indicated by the fact that the distance between the functions at group centroids for each speaker is large. Secondly, the predictive accuracy for each speaker is very high. These two results clearly demonstrate that /y/ can be grouped with cases with /r/ and other sonorant onsets.

4.3 Phonetic variables

Having established that /y/ can be put together with /r/ and the other sonorants into a category clearly distinguishable from the obstruents, it is now possible to specify what acoustic properties underlie the categorization. The second prediction of the hypothesis is that H1, H2, and INT variables are significant in the sonorancy-based classification. As seen in §4.1, /y/ is intermediate between the obstruents /z, x, s/ and the prototypical sonorants /r, l, n, ŋ/. A rough inspection at mean H1, H2, INT, F0, and HNR suggests that the sonorant groups have noticeably higher H1, H2, and INT than the obstruent groups. For example, the mean H1 for subject PS is 49.73 dB for the sonorant group but only 10.30 dB for the obstruent group. This comparison holds for all subjects as shown in Figure 9. It is important to note that the extremely low mean H1 of the obstruent group is due to the fact that voiceless segments do not have a first harmonic.

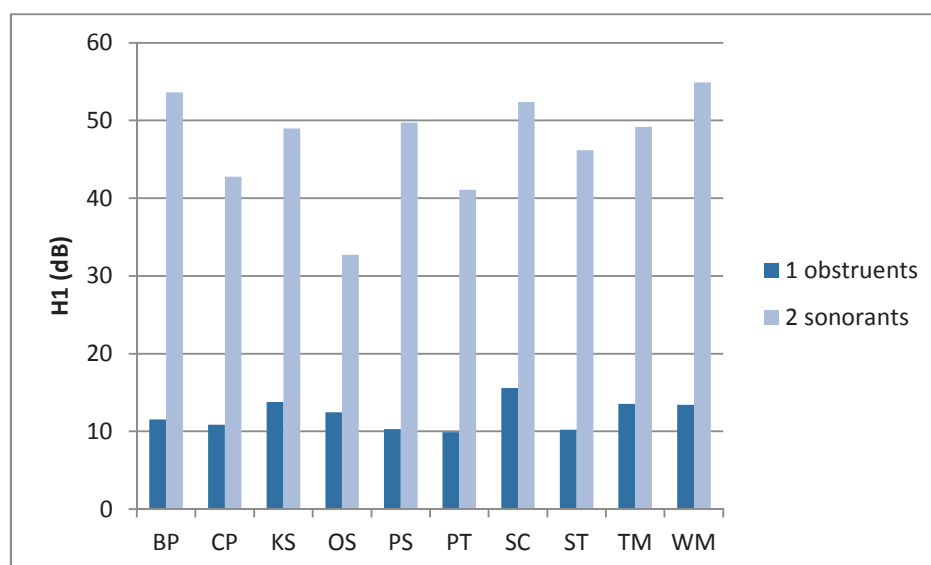


Figure 9 Mean first-harmonic amplitude (H1) for each subject (sonorancy-based classification)

Similar to H1, the mean H2 of sonorants is noticeably higher than obstruents for all speakers. For example, subject PS has a mean of 27.69 dB for the sonorant group compared to the 6.70 dB for the obstruent group. The comparison of H2 between the two categories is shown in Figure 10. Again, it is

important to note that the extremely low mean H2 of the obstruent group is due to the fact that voiceless segments lack a second harmonic.

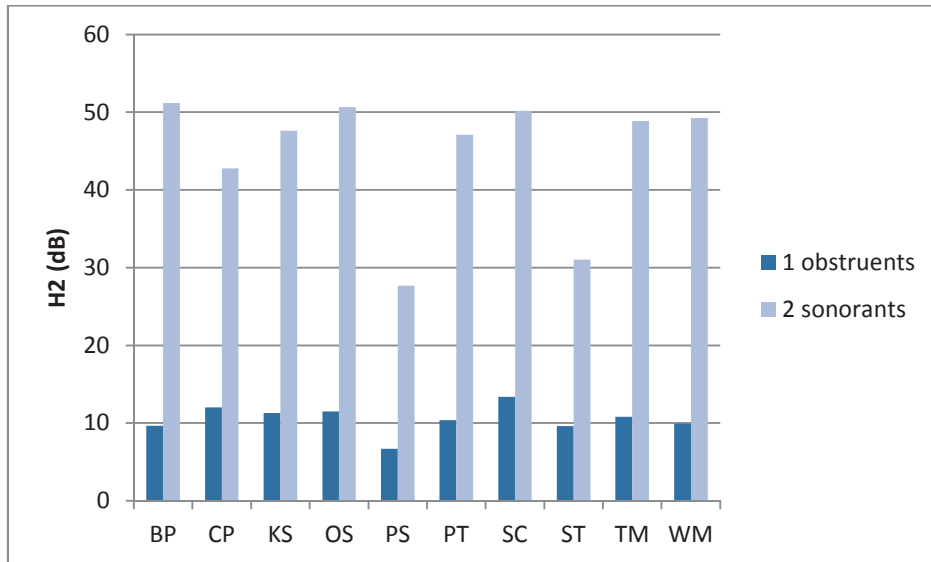


Figure 10 Mean second-harmonic amplitude (H2) for each subject (sonorancy-based classification)

In addition, the mean intensity is also higher for the sonorant group than for the obstruent group. For example, subject PS has a mean intensity of only 72.13 dB for the former but only 63.42 dB for the latter. Figure 11 gives a comparison of the mean intensity for each subject.

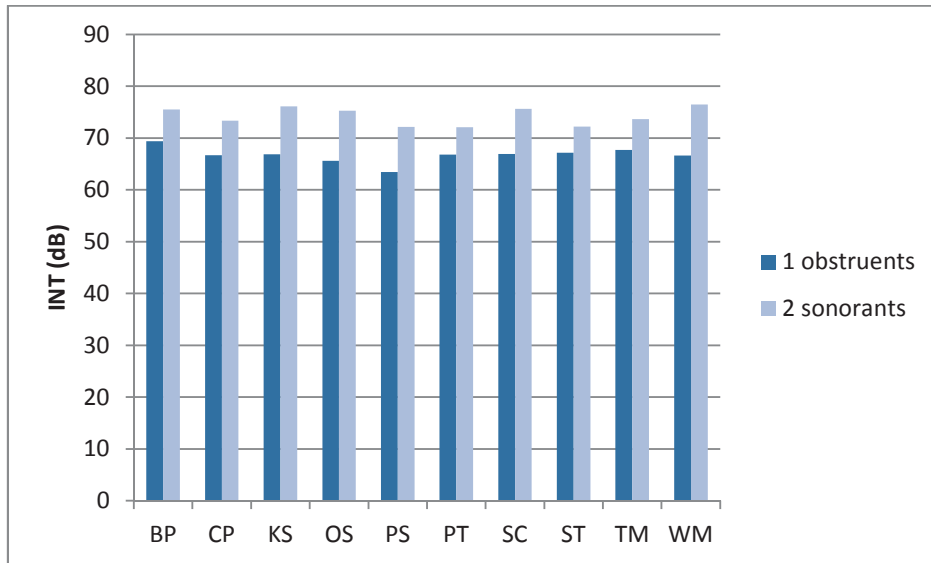


Figure 11 Mean intensity (INT) for each subject (sonorancy-based classification)

Similarly, the mean F0 also seem to distinguish between the two groups. The average F0 values for subject PS, for instance, are higher for the sonorants than the obstruents. The former has a mean of 255.98 Hz, compared to a mean of 62.84 Hz for the latter. The comparison of F0 for each subject is given in Figure 12. Again, the big difference between the two groups is due to the fact that voiceless segments do not have a fundamental frequency.

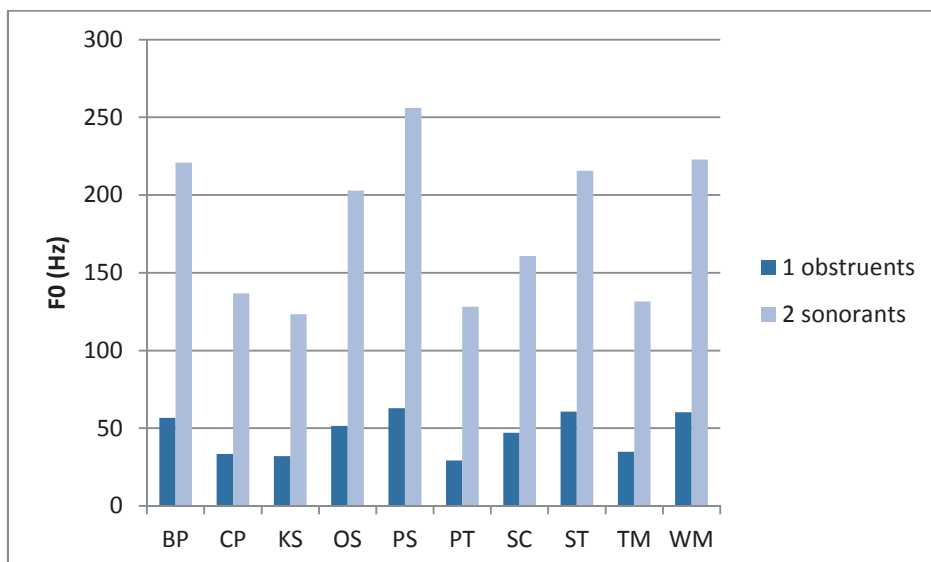


Figure 12 Mean fundamental frequency F0 for each subject (sonorancy-based classification)

Similar to the other properties, HNR is higher for the sonorants than the obstruents. For example, the mean HNR of the sonorant group for subject PS is 16.52 dB, higher than the mean 8.38 dB of the obstruent group. The comparison of mean HNR between the two categories is shown in **Figure 13**.

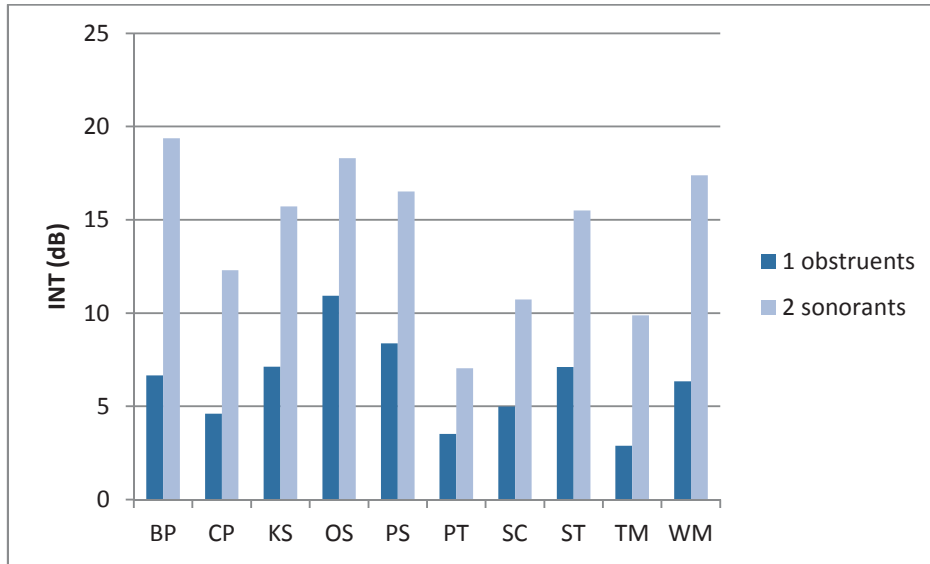


Figure 13 Mean harmonics-to-noise ratio (HNR) for each subject (sonorancy-based classification)

To precisely identify the acoustic properties that characterize the categories, the structure matrix needs to be examined. The structure matrix table indicates the significance of predictor variables in discriminating between the dependent groups. The level of importance of the predictor variables is represented by the size of the loading coefficients given in the matrix. The larger the size of the coefficients, the more important the variable is. The cut-off between important and less important predictor variables is usually taken to be 0.30 disregarding the +/- signs. The loading coefficients are summarized in **Table 10**.

Table 10 Loading coefficients for each subject

	INT	H1	H2	F0	HNR
BP	.351	.773	.860	.648	.762
CP	.479	.811	.785	.667	.547
KS	.742	.722	.874	.792	.593
OS	.874	.900	.931	.847	.516
PS	.450	.662	.893	.561	.384
PT	.455	.798	.880	.778	.619
SC	.728	.828	.930	.847	.770
ST	.280	.735	.819	.554	.581
TM	.396	.611	.771	.520	.606
WM	.827	.743	.881	.669	.533

The best predictor variable for each subject is given in bold. Shaded boxes are values smaller than 0.30.

Table 10 above shows that all the acoustic properties included are important predictors in classifying the segments into two groups based on their sonorancy. Furthermore, a closer examination reveals that H2 is the best predictor variable for most subjects. Only CP have H1 as the best predictor respectively. Overall, H1 and H2 have the largest coefficients among the five properties included. Recalling that amplitudes at lower frequencies are the acoustic properties claimed by Steven and Keyser (Stevens, Keyser et al. 1986; Stevens and Keyser 1989) to be correlates of sonorancy (see §3), INT interestingly appears to be the least reliable predictor of sonorancy as it shows relatively small values for most subjects, especially ST whose loading coefficient is .280 only. This is surprising as INT is claimed to be an acoustic correlate of sonority, which is closely related to sonorancy. It seems even less reliable than F0 and HNR, which are not directly related to sonorancy.

In conclusion, the discriminant analysis shows that it is phonetically appropriate to classify the dorsal fricative /ɣ/ in the same category as the more prototypically rhotic /r/ and other sonorants, in opposition to the obstruent group. Moreover, it shows that H1, H2, INT, F0, and HNR underlie the distinction between the two groups, with H2 as the best predictor. The sonorants typically have higher values than the obstruents for all the acoustic properties. Crucially, the analysis reveals that /ɣ/ patterns with other sonorants in terms of H1, H2, and INT as well as F0 and HNR, confirming the prediction that some acoustic properties associated with sonorant segments group /ɣ/ in the class of sonorants.

4.4 Alternative classifications

Even though the results of the discriminant analysis in §4.2 and §4.3 supports treating /ɣ/ as a sonorant, it is possible that the grouping is not the only classification that is statistically supported. In other words, there may be more than one way to group the consonants into clearly distinguished categories. This section provides the results from two additional analyses and shows that the dorsal

fricative /ɣ/ unsurprisingly can also be grouped with the obstruents, contrasting with the initial classification just discussed. However, this paper argues that the sonorancy-based classification above is the still the best classification.

The first alternative classification is based on the manner of articulation of the segments. It puts the consonants into two groups according to their manner of articulation with /ɣ/ in the same group as the other fricatives, as summarized in Table 11. This classification is designed to test whether a classification based on the actual phonetic realization that groups the dorsal /ɣ/ with the prototypical obstruents rather than the sonorants is also possible.

Table 11 Classification of cases into two groups based on the manner of articulation of the onset

Group	Segments
1	/s/, /x/, /z/, /ɣ/
2	/n/, /ŋ/, /l/, /r/

The discriminant analysis reveals that this classification is also statistically justified. The functions at the centroids of the two groups are quite far apart for all subjects, as shown in Table 12. In addition, the hit ratios for all subjects are also very high, roughly equal to the predictive accuracy of the sonorant-based classification, as shown in **Table 13**.

Table 12 Functions at group centroids for each subject (manner-based classification)

Subjects	Functions at Group Centroids	
	Group 1 fricatives	Group 2 nasals and liquids
BP	-.994	1.559
CP	-1.123	1.506
KS	-1.382	2.078
OS	-.802	1.240
PS	-.869	1.382
PT	-1.083	1.608
SC	-.929	1.246
ST	-1.163	1.687
TM	-1.718	2.278
WM	-1.421	2.014

Table 13 Classification result for each subject (manner-based classification)

Subjects	Predictive accuracy (%)		
	Group 1 fricatives	Group 2 nasals and liquids	Weighted average
BP	82.4 (N = 204)	86.9 (N = 130)	84.13 (N = 334)
CP	82.9 (N = 181)	89.6 (N = 135)	85.76 (N = 316)
KS	94.8 (N = 194)	92.2 (N = 129)	93.81 (N = 323)
OS	69.1 (N = 204)	95.5 (N = 132)	79.46 (N = 336)
PS	74.3 (N = 202)	98.4 (N = 127)	83.89 (N = 329)
PT	85.0 (N = 193)	91.5 (N = 130)	87.62 (N = 323)
SC	80.8 (N = 177)	87.1 (N = 132)	83.50 (N = 309)
ST	85.8 (N = 190)	93.1 (N = 131)	88.79 (N = 321)
TM	93.1 (N = 175)	97.7 (N = 132)	95.11 (N = 307)
WM	92.0 (N = 187)	97.0 (N = 132)	94.04 (N = 319)

These two facts indicate that grouping /ɣ/ in the obstruent group is also statistically supported. This result is clearly related to the observation in §4.1 that /ɣ/ is intermediate between the obstruents and the sonorants, and indicates a phonetic ambiguity of the segment with regards to sonorancy. In other words, the discriminant analyses performed reveals that phonetically the dorsal fricative /ɣ/ patterns with both the sonorants and the obstruents at the same time. It is not possible to definitely decide which grouping is better because comparing the average predictive accuracy of the two competing classifications yields different results for different speakers. However, the sonorancy-based classification seems to fare better overall. The average accuracy for the sonorancy-based classification is noticeably higher than that of the manner-based classification for subjects BP, OS, PS, PT, SC, ST and WM but only slightly lower for subjects CP, KS, and TM. Furthermore, the functions at the group centroids in the sonorancy-based classification are further apart than in the manner-based classification for all speakers but KS and TM. The comparison between the two analyses thus seems to suggest that the sonorancy-based classification is a better one.

The second alternative is to put the consonants into three groups with /ɣ/ on its own, as summarized in **Table 14**. This classification is designed to test whether /ɣ/ in fact does not belong to either the sonorant or the obstruent categories.

Table 14 Classification of cases into 3 groups

Group	Segments
1	/s/, /x/, /z/
2	/ɣ/
3	/n/, /ŋ/, /l/, /r/

The result of this discriminant analysis shows that it is not statistically grounded to have /ɣ/ alone in its own group. The functions at the centroids for all subjects are not very far apart, as shown in **Table 15**. In addition, the classification results for many subjects are also relatively low, as shown in **Table 16**. For example, subject PS has a rather low hit ratio of 78.6% for the second group, and a very low hit ratio of 63.8% for the third group. The average predictive accuracy is only 79.03%. This means that for this subject it is not appropriate to treat /ɣ/ and the other obstruents as two distinct groups. Another example is subject OS whose hit ratios for the three groups are 74.1%, 69.0%, and 66.7% respectively. The average accuracy for this subject is only 70.54%. These hit ratios are similarly low, indicating that treating /ɣ/ as a separate class is not statistically supported. The result of this second alternative analysis suggests that /ɣ/ must be treated as part of either the sonorant or obstruent class, highlighting the phonetic ambiguity of the segment.

Table 15 Functions at group centroids for each subject (ternary classification)

Subjects	Functions at Group Centroids		
	Group 1 obstruents	Group 2 dorsal fricative	Group 3 sonorants
BP	-1.962	1.371	2.002
CP	-1.645	.387	1.618
KS	-1.846	-.144	2.222
OS	-1.518	1.047	1.529
PS	-2.334	2.016	2.273
PT	-1.794	.558	1.904
SC	-1.594	.575	1.463
ST	-2.370	1.010	2.354
TM	-2.685	-.036	2.717
WM	-2.301	.316	2.427

Table 16 Classification result for each subject (ternary classification)

Subjects	Predictive accuracy (%)			
	Group 1	Group 2	Group 3	Average
BP	89.5 (N = 162)	78.6 (N = 42)	80.8 (N = 130)	84.73 (N = 334)
CP	73.2 (N = 142)	66.7 (N = 39)	85.9 (N = 135)	77.85 (N = 316)
KS	73.0 (N = 152)	95.2 (N = 42)	87.6 (N = 129)	81.73 (N = 323)
OS	74.1 (N = 162)	69.0 (N = 42)	66.7 (N = 132)	70.54 (N = 336)
PS	91.3 (N = 160)	78.6 (N = 42)	63.8 (N = 127)	79.03 (N = 329)
PT	74.8 (N = 151)	88.1 (N = 42)	82.3 (N = 130)	79.57 (N = 323)
SC	69.1 (N = 136)	75.6 (N = 41)	71.2 (N = 132)	70.87 (N = 309)
ST	85.1 (N = 148)	73.8 (N = 42)	82.4 (N = 131)	82.55 (N = 321)
TM	85.7 (N = 133)	81.0 (N = 42)	95.5 (N = 132)	89.25 (N = 307)
WM	80.7 (N = 145)	81.0 (N = 42)	92.4 (N = 132)	85.58 (N = 319)

In summary, two additional discriminant analyses reveal that grouping the dorsal fricative /ɣ/ with the prototypical sonorants is not the only possible classification. It is also possible to treat it as belonging to the same class as the prototypical obstruents. Despite this ambiguity, it seems that treating /ɣ/ as member of the sonorant class might still be the best.

5. Conclusion

In conclusion, this paper argues that the dorsal fricative /ɣ/ in Sgaw Karen is phonologically a sonorant. Phonologically, its sonorant status is evidenced by the fact that it occurs as the second member of consonant clusters just like other liquids (see §2). Given its central and oral articulation, this means that the segment is phonologically a rhotic sound. Phonetically, the segment shows intermediate amplitude of the first harmonic, amplitude of the second harmonic, intensity, fundamental frequency, and harmonics-to-noise ratio that allow it to be grouped with either the sonorants or the obstruents. However, treating /ɣ/ seems to offer the best classification result (see §3 and §4). Therefore, the phonological and phonetic evidence together provides support for the claim that the dorsal fricative /ɣ/ in Sgaw Karen is a sonorant despite its fricative realization. In other words, the phonological and phonetic properties of /ɣ/ in Sgaw Karen are compatible with the traditional classification of rhotics as sonorants.

6. Discussion

The results of this study are revealing with respect to the phonological status of the dorsal fricative /ɣ/ in Sgaw Karen. It demonstrates that the Sgaw Karen dorsal fricative /ɣ/ can be characterized as a rhotic sonorant from the point of view of both phonology and phonetics. However, an apparent problem is that labeling the dorsal /ɣ/ as [+sonorant] is problematic as it contradicts the common assumption that fricative segments are assumed to be obstruents, and thus featurally specified as [-sonorants]. This issue disappears when the definition of the feature [sonorant] is carefully examined. The label “fricative” is a phonetic term used to describe sounds made by forcing the air through a constricted passage thus creating a turbulent noise (Ladefoged 2006). However, there is no phonological feature that alone groups fricative sounds as a distinct phonological class. Even though fricative segments are typically specified as [-sonorant] because of their observable phonetic realization, the feature [sonorant] does not make reference to the passage or the turbulent noise.

In *SPE* (Chomsky and Halle 1968: 302), sonorants, which are specified as [+sonorant], are defined as segments produced “with a vocal tract configuration in which spontaneous voicing is possible.” In contrast, obstruents, or [-sonorant] segments, are defined as produced “with a cavity configuration that makes spontaneous voicing impossible.” This explicit definition of [sonorants] means that friction is not directly relevant to the sonorancy of a segment. Whether a given sound is a sonorant or an obstruent depends on whether it is compatible with spontaneous voicing. In other words, the fricative [s] is an obstruent not because of its characteristic turbulent noise but because it is produced with a vocal configuration that does not allow spontaneous voicing. On the other hand, this strict interpretation of the term “fricative” allows for some fricatives to be classified as [+sonorant] phonologically. This means that the dorsal fricative /ɣ/ in Sgaw Karen can be characterized as a sonorant fricative, in contrast to the more canonical obstruent fricatives like /s/ or /x/. Examples of fricatives that display the phonological properties characteristic of sonorants include [ʋ] in French (cf. Tranel 1987; Demolin 2001), European Portuguese [χ] in European Portuguese (cf. Jesus and Shadle 2005), [ʒ] in Mandarin Chinese (Duanmu 2000), and [ð] Wood Cree (Starks and Ballard 2005).

Lastly, the sonorancy of the dorsal fricative /ɣ/ raises an interesting theoretical issue about the relationship between phonetic realization and phonological categorization. If fricatives can be either sonorants or obstruents, it is puzzling whether the two types are realized differently depending on the languages. To put in more concretely, it is unclear whether the rhotic /ɣ/ as in Sgaw Karen and Tiv has a higher H2 than the non-rhotic /ɣ/ in Vietnamese and Modern Greek. The first possibility is that sonorant fricatives are phonetically different from obstruent fricatives. In this case, rhotic /ɣ/ would be realized with higher second-harmonic amplitude than their non-rhotic counterparts. This possibility would suggest that the phonetic realization of segments varies according to their roles in the phonological systems. The second possibility is that the two types of fricatives are similar in terms of the relevant phonetic parameters, but languages set up different sonorancy thresholds to distinguish between sonorants and obstruents. In this scenario, rhotic and non-rhotic /ɣ/ would be realized as roughly the same, but the minimum values of the phonetic properties required of sonorant segments are higher in

languages where /ɣ/ is non-rhotic. This possibility would suggest that the phonetic realization of segments is fairly constant across languages but it is the categorization that shifts.

The phonetic ambiguity of /ɣ/ revealed by this study seems to suggest the second possibility. Being one of the most “phonetically sonorous” segments among stops and fricatives, it may fall on either side of the “sonorancy threshold” imposed by each individual language. This would also not only capture the ambiguity of /ɣ/ but also explain why voiced dorsal fricatives /ɣ/ and /ʁ/ are the fricatives that most frequently show rhotic behaviors in languages of the world. However, future cross-linguistic study on the acoustic differences of rhotic and non-rhotic fricatives, especially in languages that have both, will shed light on this aspect of phonetic-phonology connection.

References

- Bhat, D. N. S. (1974). "The phonology of liquid consonants." Working papers on language universals **16**: 73-104.
- Boersma, P. and D. Weenink (2010). Praat: doing phonetics by computer.
- Bradley, T. G. (2012). "Rhotic variation and contrast in Veracruz Mexican Spanish." Estudios de Fonética Experimental **21**: 43-74.
- Burns, R. and R. Burns (2008). Business research methods and statistics using SPSS (additional advanced chapters online at <http://www.uk.sagepub.com/burns/chapters.htm>). London, SAGE Publications Ltd. .
- Chomsky, N. and M. Halle (1968). The sound pattern of English. New York, Harper and Row.
- Clements, G. N. (1990). The role of syllabification cycle in core syllabification. Papers in Laboratory Phonology I: Between the grammar and physics of speech. J. Kingston and M. E. Beckman. Cambridge; New York, Port Chester, Melbourne, Sydney, Cambridge University Press: 283-333.
- Demolin, D. (2001). Some phonetic and phonological observation concerning /R/ in Belgian French 'r-atics: Sociolinguistic, phonetic and phonological characteristics of /r/. Special issue of Études & Travaux 4. . H. van de Velde and R. van Hout: 63-74.
- Dixon, R. M. W. (1980). The Australian languages. Cambridge, Cambridge University Press.
- Dong, S. (1958). Yuyin changtan [Introduction to phonetics]. Beijing, Wenhua Jiaoyu Chubanshe.
- Duanmu, S. (2000). The phology of Standard Chinese. Oxford, Oxford University Press.
- Fant, C. G. M. (1969). "Distinctive features and phonetic dimensions." Speech Transmission Laboratory. Quarterly Progress and Status Reports **10**(2-3): 1-18.
- Gilmore, D. (1898). A grammar of the Sgaw Karen. Rangoon, American Baptist Mission Press.
- Hall, T. A. (1993). "The phonology of German /R/." Phonology **10**(1): 83-105.
- Hall, T. A. (1997). The phonology of coronals. Berlin, John Benjamins Publishing Company.
- Hall, T. A. (2007). Segmental features. (ed.) The Cambridge handbook of phonology. P. de Lacy. Cambridge, Cambridge University Press: 311-333.
- Harris, J. W. (1983). Syllable structure and stress in Spanish: a nonlinear analysis. Cambridge, MA, MIT Press.
- Holton, D., P. Mackridge, et al. (1997). Greek grammar: a comprehensive grammar of the modern language. London, Routledge.
- Inouye, S. B. (1995). Trills, taps, and stops in contrast and variation. Department of Linguistics. Los Angeles, UCLA. **Ph.D.**
- Jacobson, R., C. G. M. Fant, et al. (1963). Preliminaries to Speech Analysis: The distinctive features and their correlates. Cambridge, MA, MIT Press.
- Jesuit Social Services Centre (2005). Lix Pgaz K'nyau Av Hta. Bangkok, Jesuit Social Services Centre.
- Jesus, L. M. T. and C. H. Shadle (2005). "Acoustic analysis of European Portuguese uvular [ʁ, χ] and voiceless tapped alveolar [ɾ] fricatives." Journal of the International Phonetic Association **35**(1): 27-44.
- Jesus, L. M. T. and C. H. Shadle (2005). "Acoustic analysis of European Portuguese uvular [χ, ʁ] and voiceless tapped alveolar [ɾ] fricatives." Journal of the International Phonetic Association **35**(1): 27-44.
- Jones, R. B. (1960). Karen Linguistic Studies: Description, comparison, and texts. Berkeley, University of California Press.
- Karlgren, B. (1915-26). Étude sur la phonologie chinoise. Uppsala, K. W. Appelberg.
- Keating, P. A. (1990). "Phonetic representations in a generative grammar." Journal of Phonetics **18**: 321-334.
- L-Thongkum, T. (ms.). Proto-Karen Phonology and Lexicon, Chulalongkorn University.

- Ladefoged, P. (2006). A course in phonetics. Boston, Thomson-Wadsworth.
- Ladefoged, P. and I. Maddieson (1996). The sounds of the world's languages. Oxford, UK ; Cambridge, MA, Blackwell Publishing.
- Lindau, M. (1985). The story of r. Phonetic Linguistics. V. A. Fromkin. Orlando, FL, Academic Press: 157-168.
- Maddieson, I. (1984). Patterns of sounds. Cambridge; New York, Cambridge University Press.
- Magnuson, T. J. (2007). The story of /r/ is two vocal tracts. 16th International Congress of Phonetic Sciences, Saarbrücken, Germany.
- Matisoff, J. A. (2003). Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction. Berkeley, University of California Press.
- Parker, S. (2002). Quantifying the Sonority Hierarchy. Department of Linguistics. Amherst, University of Massachusetts Amherst. **Ph.D.**
- Proctor, M. I. (2009). Gestural characterization of a phonological class: the liquids. New Haven, Yale University. **Ph.D.**
- Ratanakul, S. (1986). Thai-Sgaw Karen dictionary. S. Srichampa and W. Niyomtham. Nakhon Pathom, Graduate School and Institute of Language and Culture for Rural Development, Mahidol University.
- Starks, D. and E. Ballard (2005). "Woods Cree /ð/: an unusual type of sonorant." International Journal of American Linguistics **71**: 102-115.
- Stevens, K. N. and S. E. Blumstein (1981). The search for invariant acoustic correlates of phonetic features. Perspectives on the study of speech. P. D. Eimas and J. Miller. Hillsdale, NJ, L. Erlbaum Associates: 1-38.
- Stevens, K. N. and S. J. Keyser (1989). "Primary features and their enhancement in consonants." Language **65**(1): 81-106.
- Stevens, K. N., S. J. Keyser, et al. (1986). Toward a phonetic and phonological theory of redundant features. Invariance and variability in speech processes. J. S. Perkell and D. H. Klatt. Hillsdale, NJ, Lawrence Erlbaum Associates: 426-449.
- Tranel, B. (1987). The sounds of French. Cambridge, Cambridge University Press.
- Walsh Dickey, L. (1997). The phonology of liquids. Department of Linguistics. Amherst, MA, UMass, Amherst. **Ph.D.**
- Webb, E. R. (2002). The relational /r/: three case studies in rhotic integrity and variation, The University of Texas at Austin. **Ph.D.**
- Wiese, R. (2001). The phonology of /r/. Distinctive Feature Theory. T. A. Hall. Berlin, Mouton de Gruyter: 335-368.
- Wiese, R. (2001). The unity and variation of (German) /r/. 'r-atics: Sociolinguistic, phonetic and phonological characteristics of /r/. Special issue of Études & Travaux **4**. H. van de Velde and R. van Hout: 27-44.
- Zec, D. (1995). "Sonority constraints on syllable structure." Phonology **12**: 85-129.

Appendix 1

Wordlist for recording

	IPA	Simplified Romei	Gloss
1.	zɔ ¹	yau	to subside
2.	ɣa ¹	ga	clf.for human
3.	nuw ³	nuv	to enter
4.	na ³	nav	to believe
5.	xo ¹	qo	to burn
6.	ɣa ²	gaj	to hurry
7.	ɣe ²	geij	to study
8.	xu ¹	qoo	even, flat
9.	lɔ ³	lauv	to play
10.	ɣe ¹	gai	to play
11.	xa ³	qav	dirty
12.	le ⁴	leif	to lick
13.	luw ²	luj	to feed
14.	re ⁴	reif	bit
15.	lo ²	loj	to use magic
16.	se ¹	sai	sound
17.	ɣɔ ³	gauv	to insert
18.	zɔ ²	yauj	to step on
19.	si ³	siv	liquor
20.	su ⁴	soof	louse
21.	ɣɔ ¹	gau	red
22.	ruw ¹	ru	mane
23.	ŋɔ ¹	ngau	foolish
24.	le ¹	lei	cleared, bare
25.	la ⁴	laf	leaf
26.	sɔ ³	sauv	to put
27.	sɔ ¹	sau	new
28.	zi ¹	yi	far
29.	ne ²	neij	to acquire
30.	se ⁴	seif	wood
31.	xa ⁴	qaf	chicken louse
32.	xuw ³	quv	dry field
33.	xo ⁴	qof	to chase
34.	xi ¹	qi	bone
35.	zu ²	yooj	to swallow
36.	ro ²	roj	pile (of wood)

	IPA	Simplified Romei	Gloss
37.	lɔ ²	lauj	haystack
38.	lɤ ²	lej	rock
39.	xɛ ³	qeiv	to beg
40.	nɤ ⁴	nɛf	edge
41.	xɤ ⁴	qɛf	roasted rice
42.	nɛ ⁴	naif	to point
43.	xi ⁴	qif	scurf
44.	xɔ ³	qauv	bee wax
45.	zo ¹	yo	Yuan, Northern Thai
46.	xɔ ⁴	qauf	jail
47.	nɔ ¹	nau	black sesame
48.	ɣu ²	guj	snake
49.	ze ¹	jei	cicada
50.	ŋa ¹	nga	to hire
51.	nɔ ²	nauj	slow
52.	nɤ ¹	ne	to smell
53.	su ¹	soo	black
54.	xɛ ¹	qei	grill
55.	xɑ ¹	qa	to yell at
56.	si ¹	si	to die
57.	xɛ ⁴	qeif	to stew
58.	ɣo ²	goj	to bare
59.	lu ¹	loo	diarrhea
60.	su ³	soov	land leech
61.	nu ²	nooj	milk
62.	zi ³	yiv	long time
63.	so ¹	so	oil
64.	ɣi ³	giv	to grind
65.	xɔ ¹	qau	stairs
66.	zu ¹	yoo	to fly
67.	ni ⁴	nif	year
68.	sa ⁴	saf	fruit
69.	xɛ ¹	qai	lean
70.	za ⁴	yaf	blanket
71.	so ⁴	sof	to wilt
72.	ɣɛ ³	gaiv	to slice
73.	lu ⁴	luf	cotton, thread
74.	la ¹	la	moon
75.	nɔ ³	nauv	elder sister
76.	xɤ ³	qev	konjac
77.	ɣɤ ²	gej	weevil

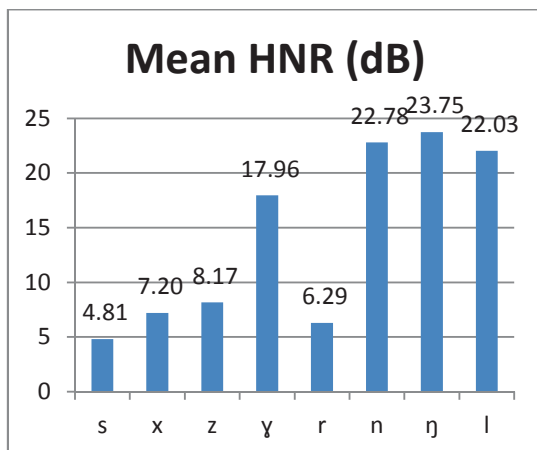
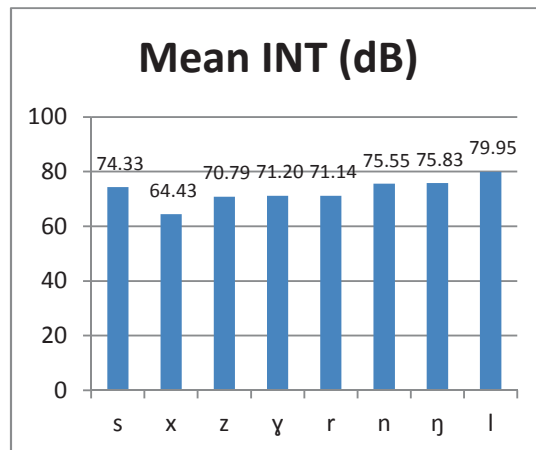
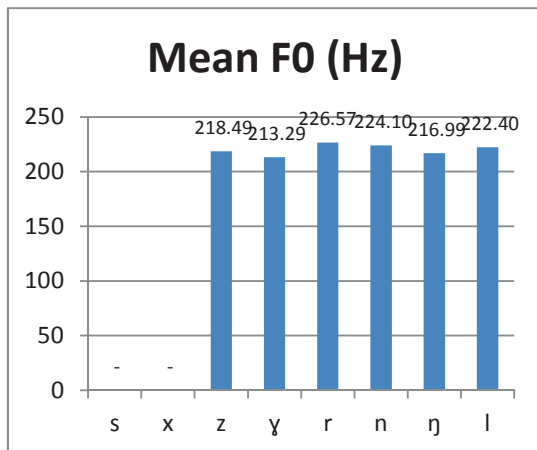
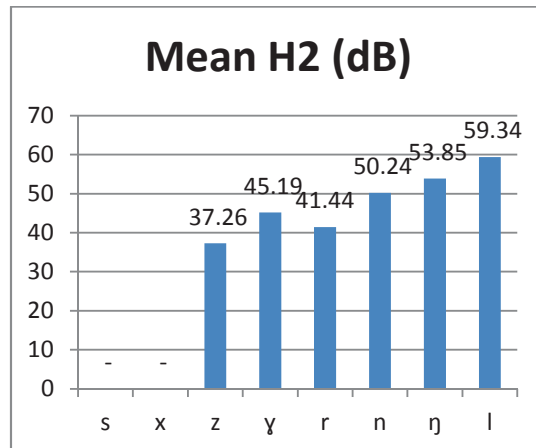
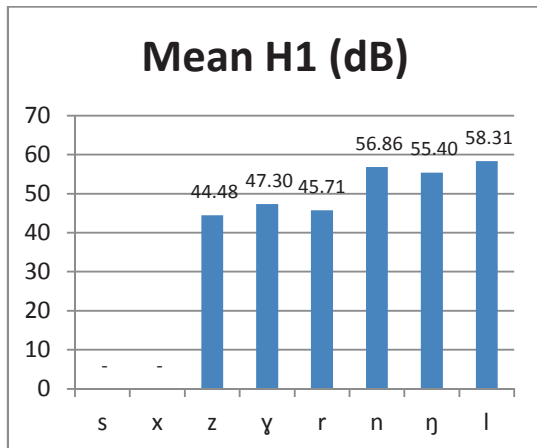
[removed because subjects uniformly have /j-/]

	IPA	Simplified Romei	Gloss
78.	se ¹	sei	hog plum
79.	ru ³	roov	to tie
80.	ra ³	rav	many, much
81.	xe ³	qaiv	knife
82.	zu ³	yoov	to raise
83.	ni ¹	ni	to smile
84.	lo ³	lov	car
85.	sw ³	suv	soldier
86.	yw ¹	gu	to sieve
87.	ro ¹	ro	fence
88.	yi ²	gij	root
89.	lo ¹	lau	to go down
90.	zo ³	yauv	deep
91.	ye ¹	gei	beautiful, good
92.	ze ³	yaiv	to chop
93.	ne ³	naiv	pineapple
94.	sy ³	sev	three
95.	lo ¹	lo	graveyard
96.	no ⁴	nauf	grass
97.	no ⁴	nof	pond, lake
98.	ly ⁴	lef	maggot
99.	xy ¹	qe	heavy
100.	li ³	liv	letter (of alphabet)
101.	zw ²	yuj	mouse
102.	se ³	saiv	areca nut
103.	ri ³	riv	to roll
104.	xi ³	qiv	to pick up
105.	li ¹	Li	grandchild
106.	si ⁴	sif	to comb
107.	le ²	leij	cowrie
108.	sa ³	sav	heart
109.	xw ¹	qu	to search
110.	ze ²	yaij	five
111.	li ²	lij	vagina
112.	ly ¹	le	warm

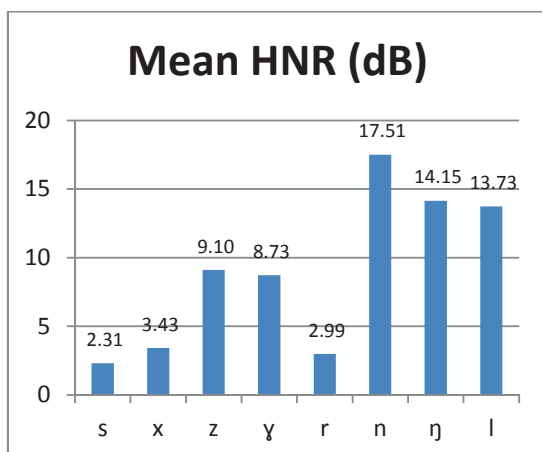
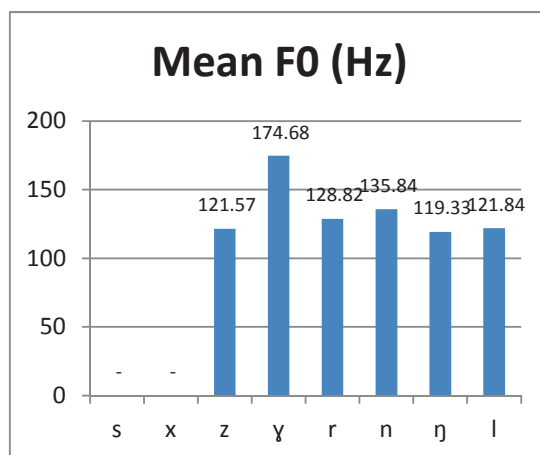
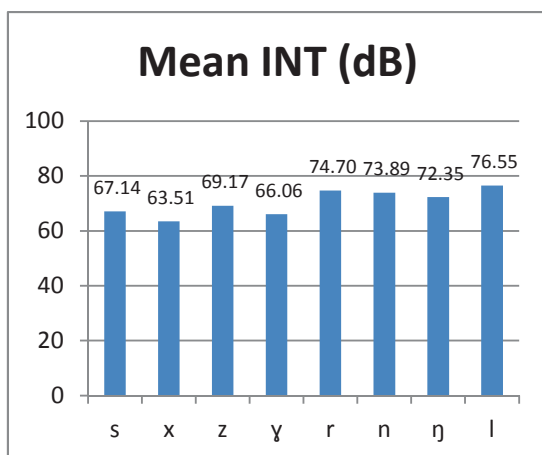
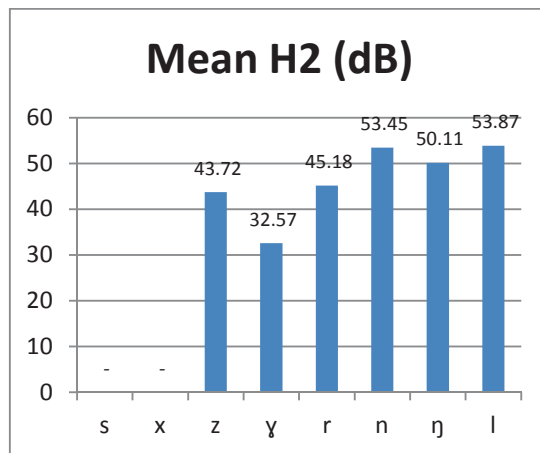
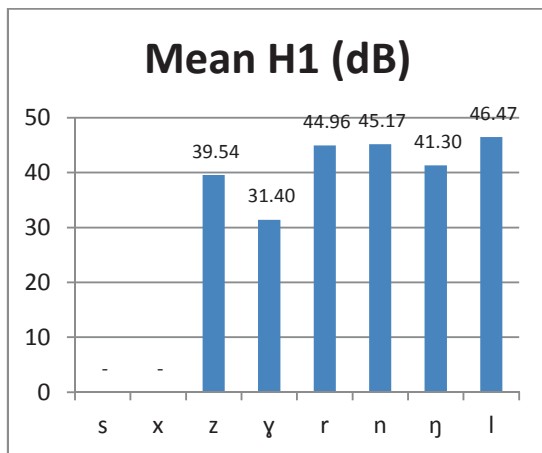
Appendix 2

Acoustic measurements

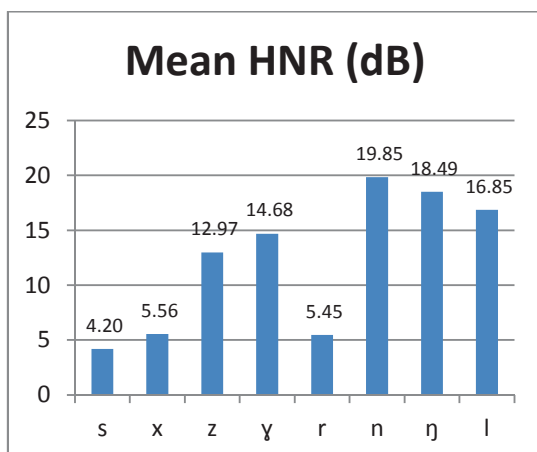
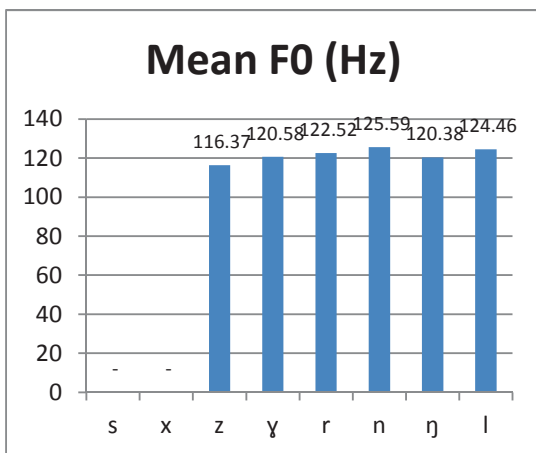
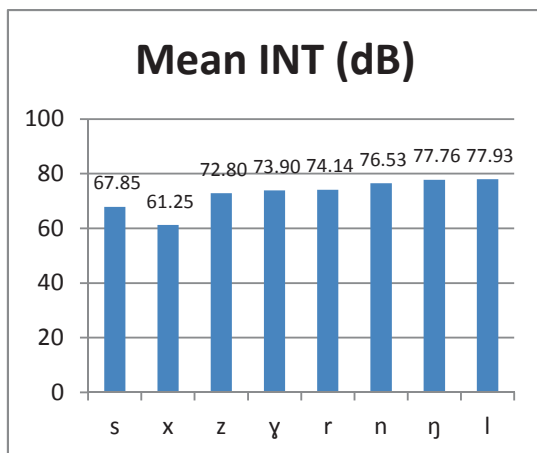
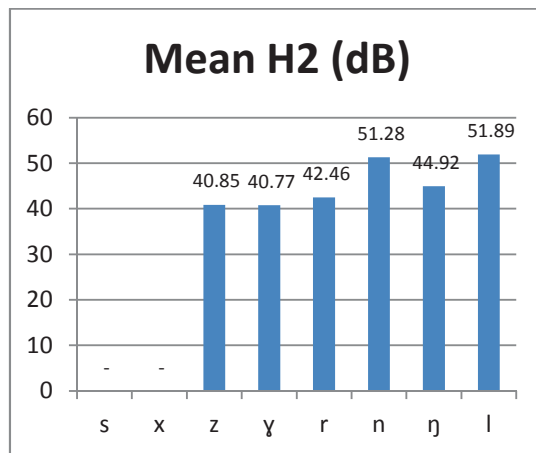
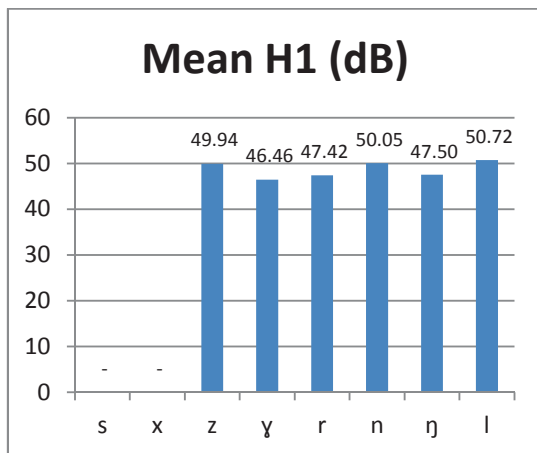
Subject BP



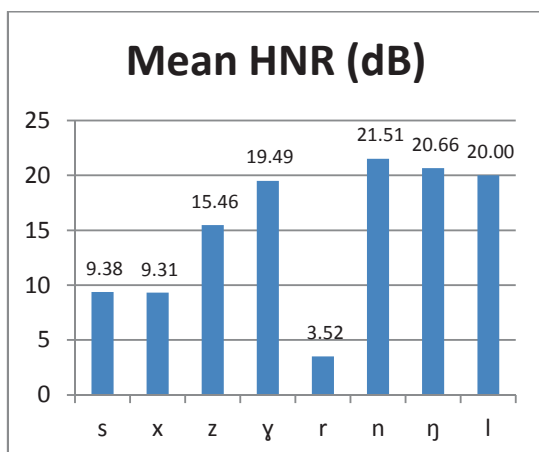
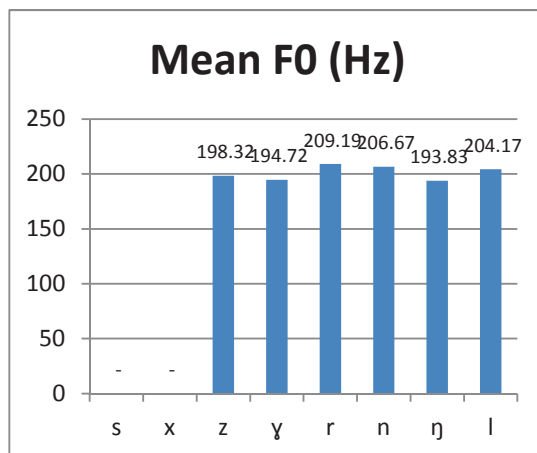
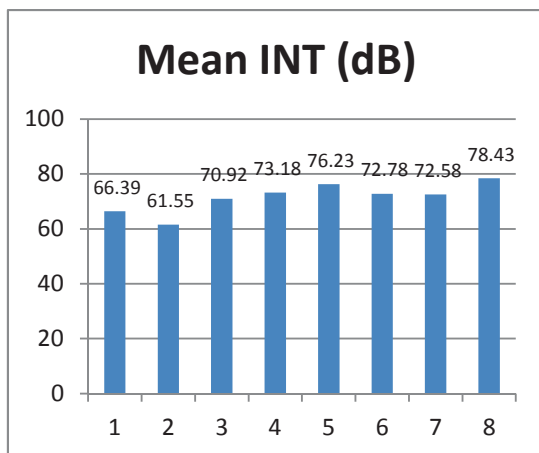
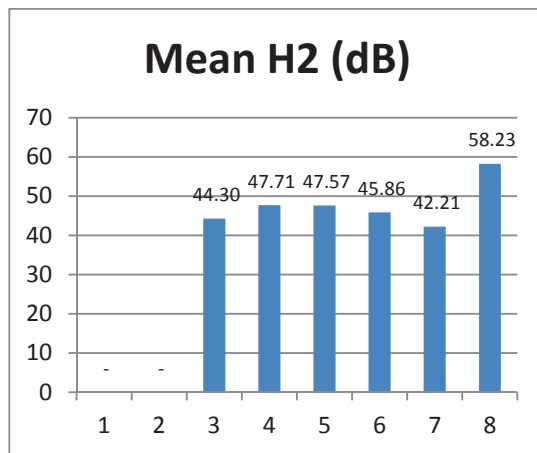
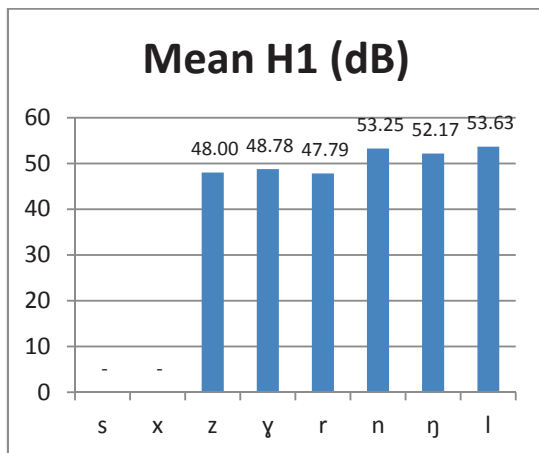
Subject CP



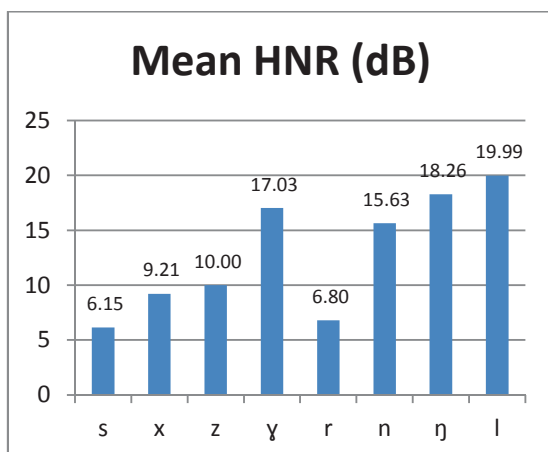
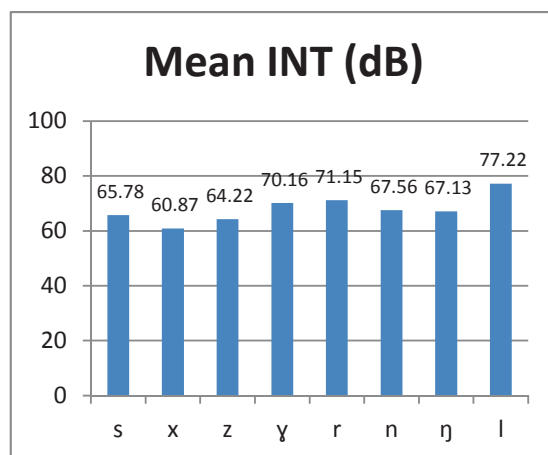
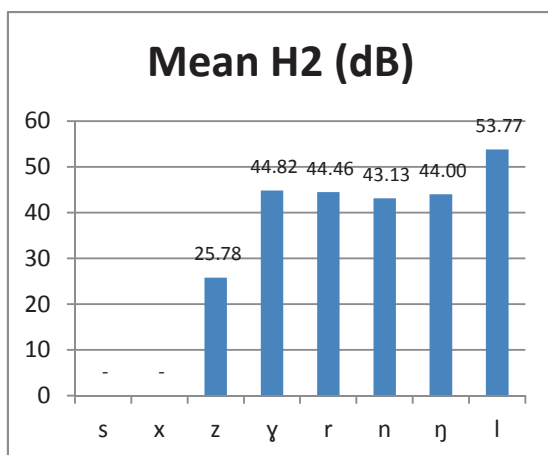
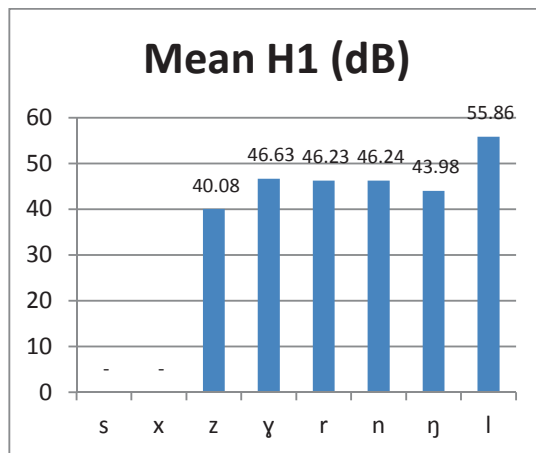
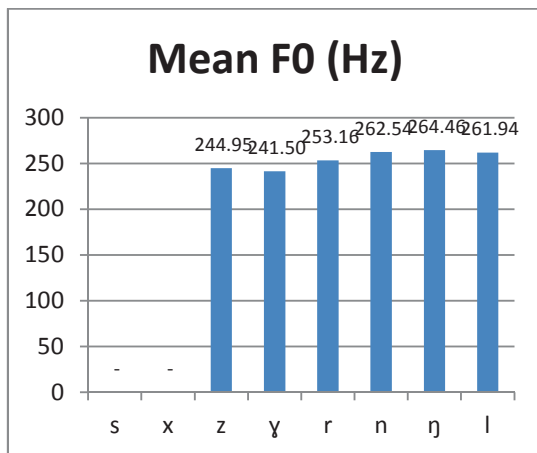
Subject KS



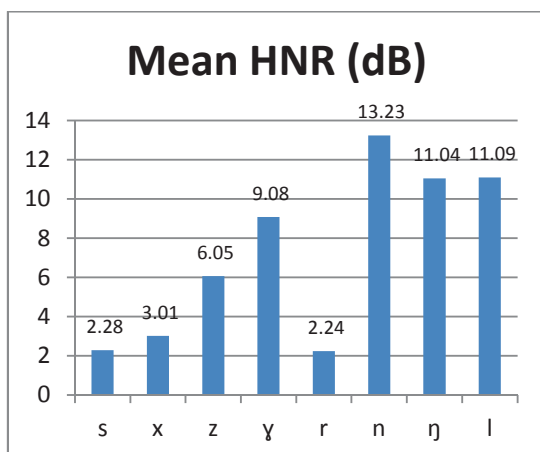
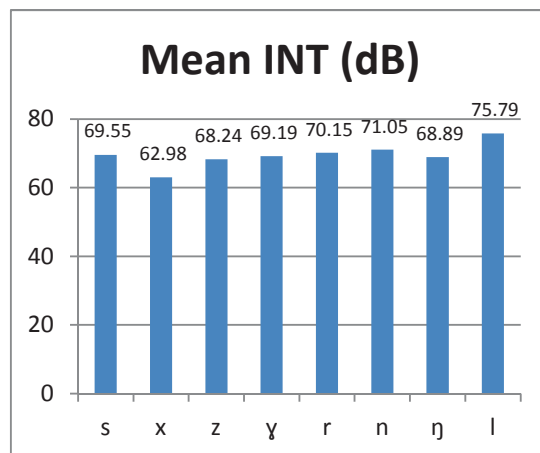
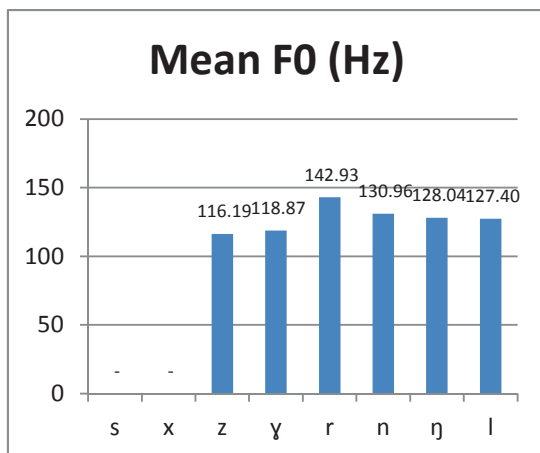
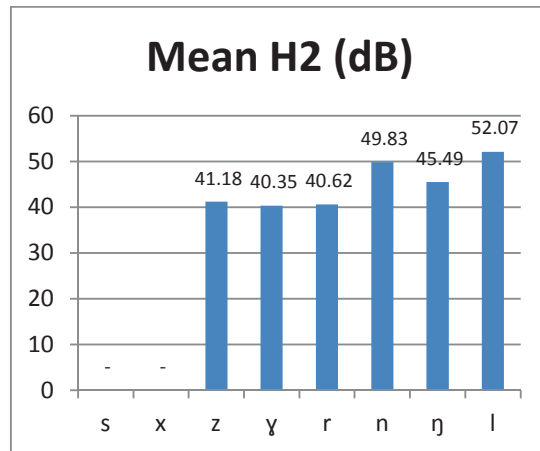
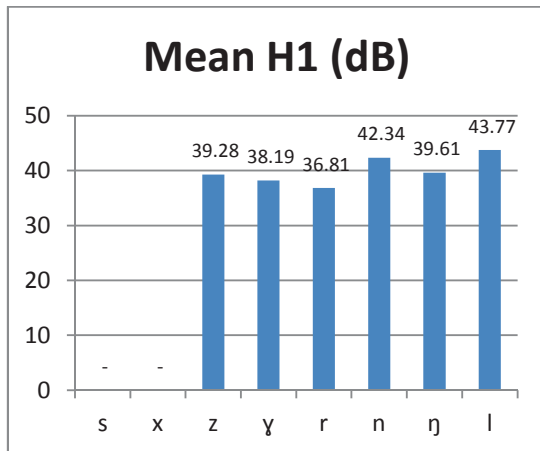
Subject OS



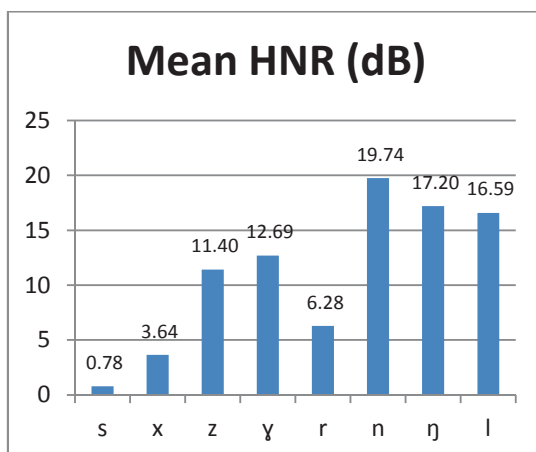
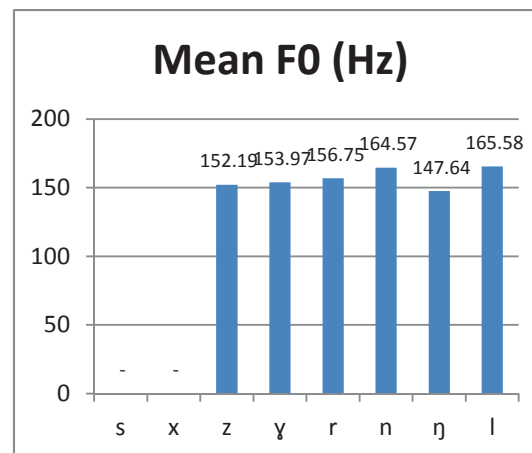
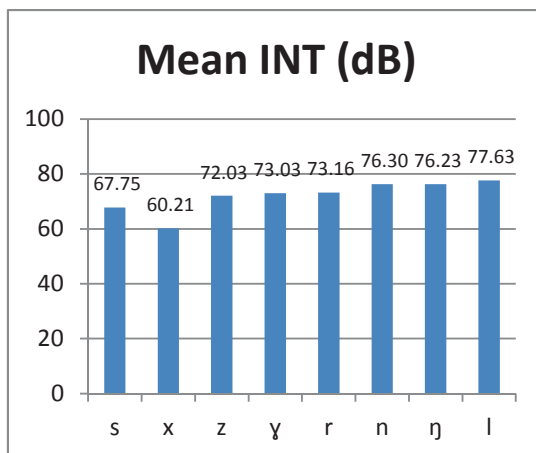
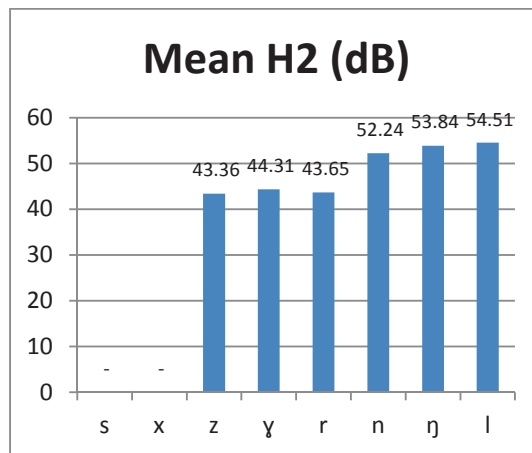
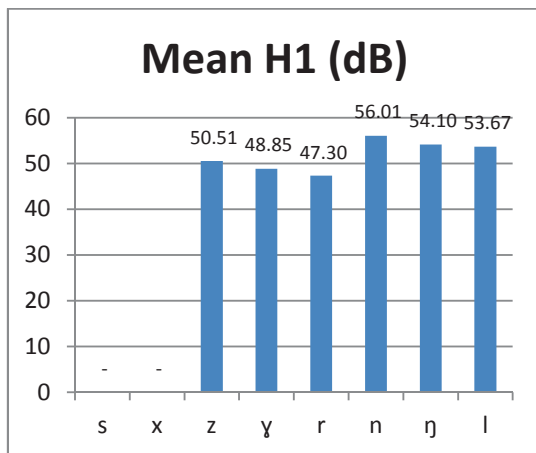
Subject PS



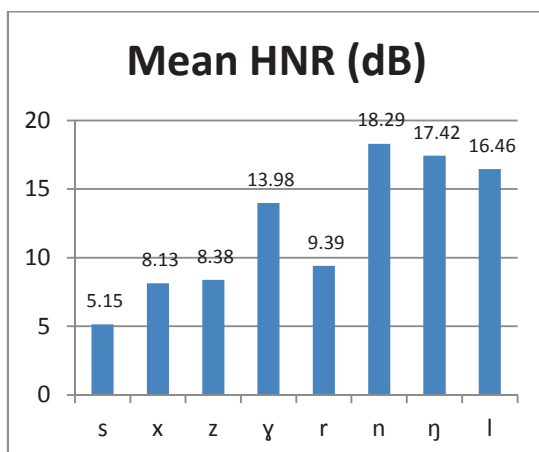
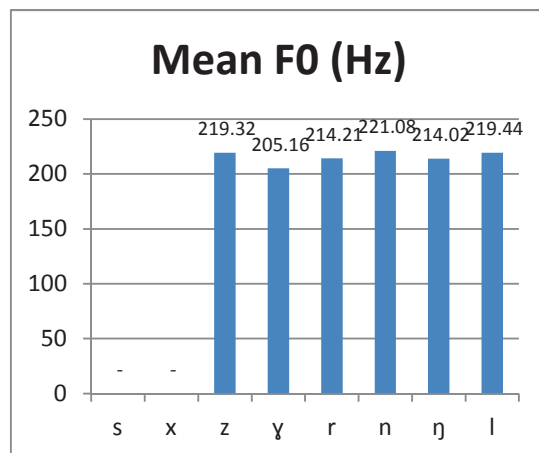
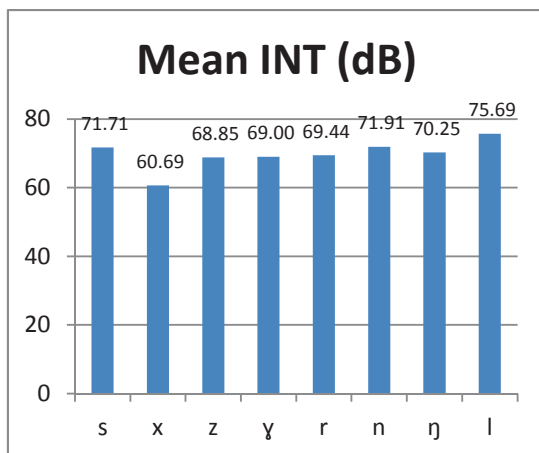
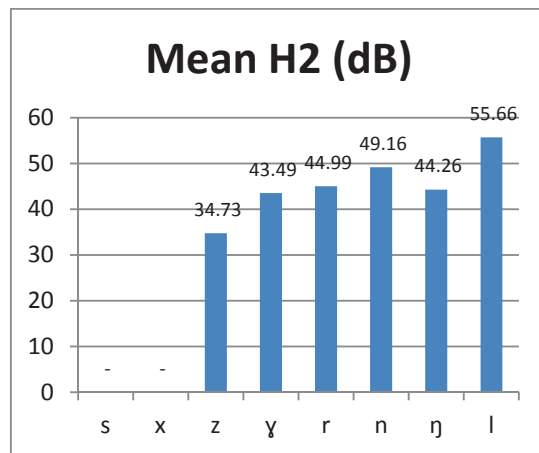
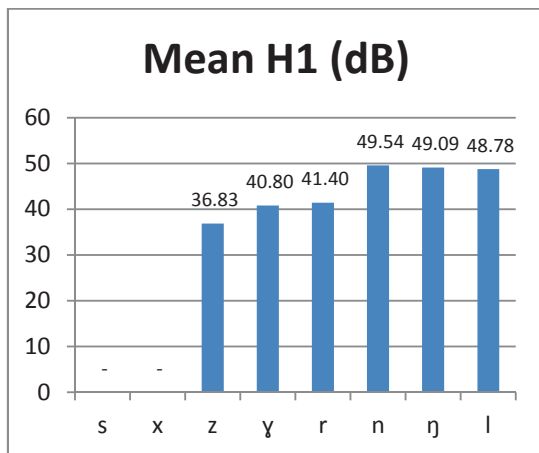
Subject PT



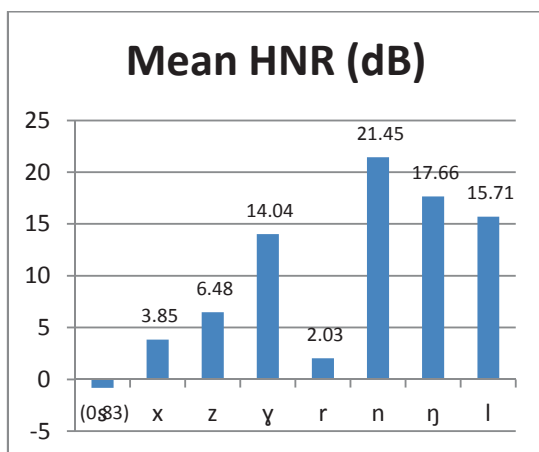
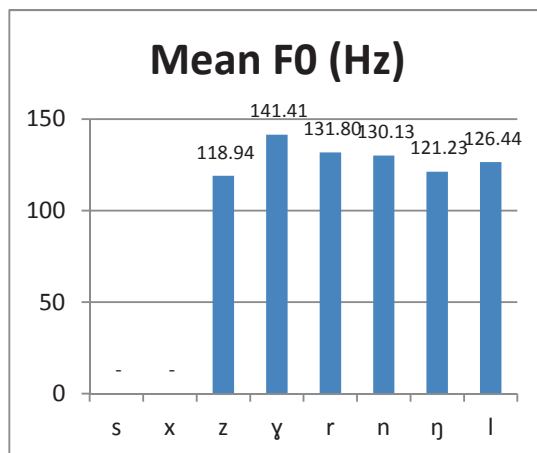
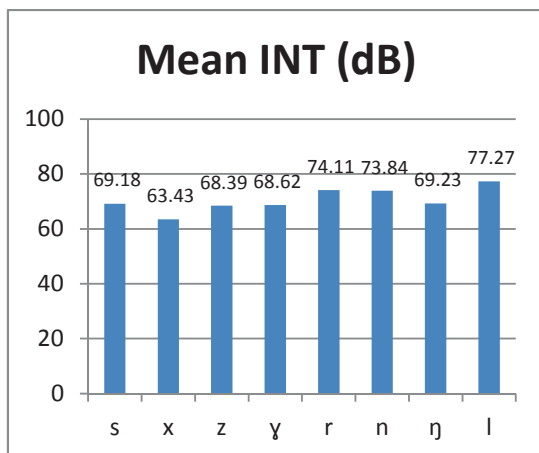
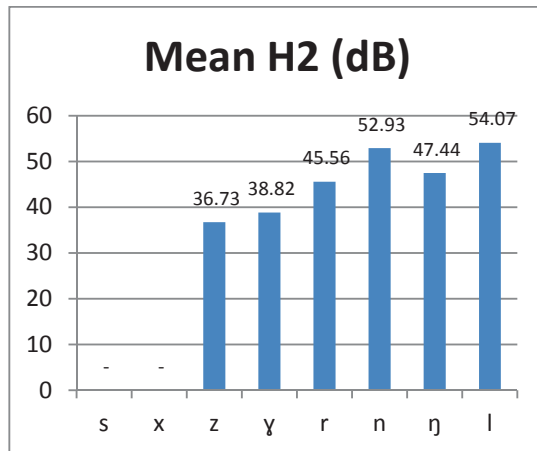
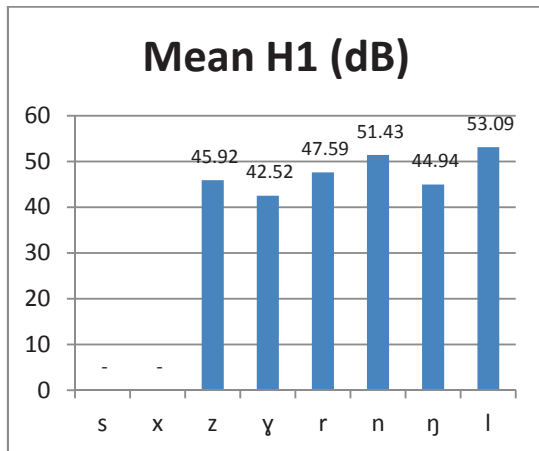
Subject SC



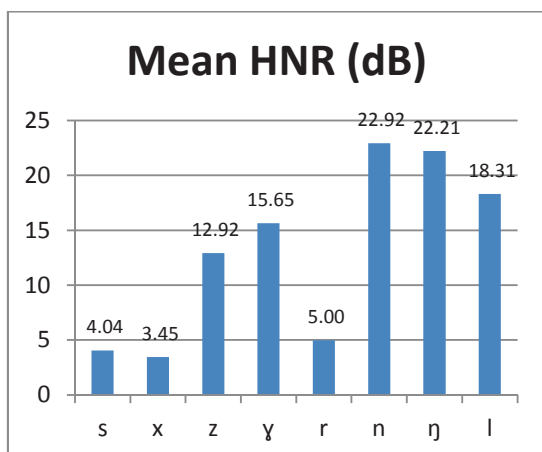
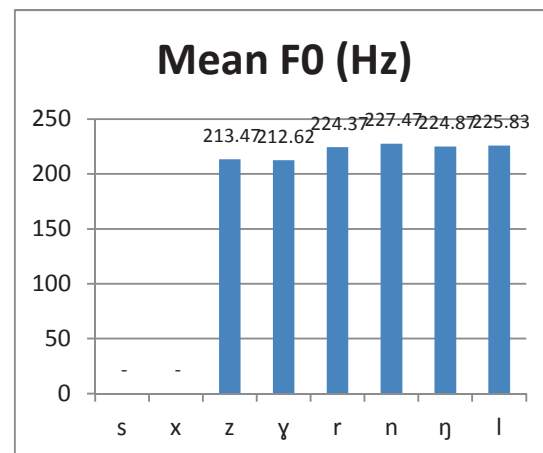
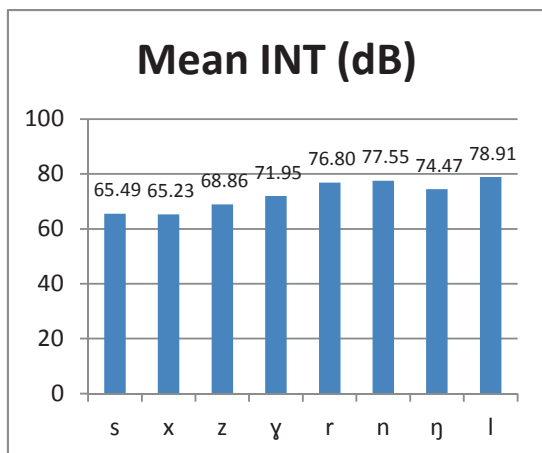
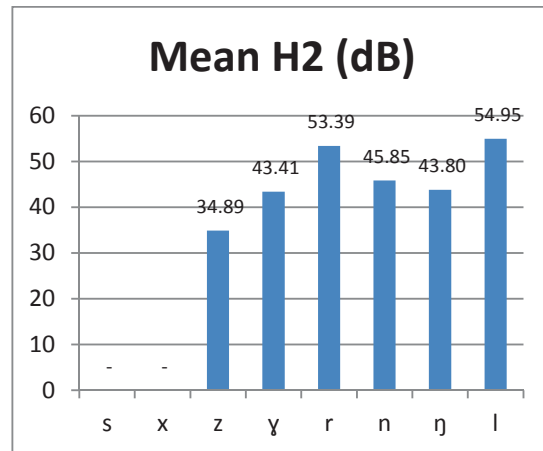
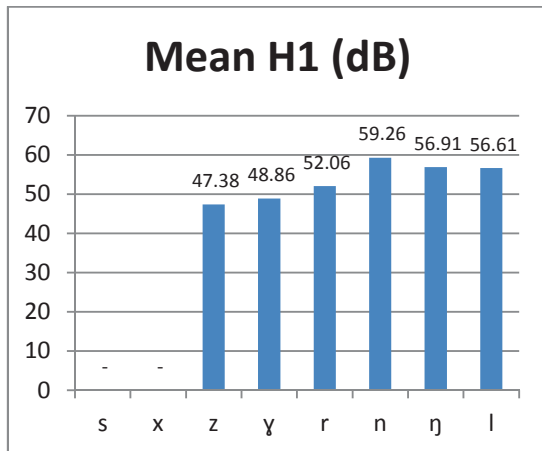
Subject ST



Subject TM



Subject WM



นาย François Langella

Polyfunctionality in Pwo Karen: The case of ʔàʔ-

(<T-B pronominal prefix *ʔa-)

Hituzi Linguistics in English No. 19

Typological Studies on Languages in Thailand and Japan

Edited by:
Tadao Miyamoto, Naoyuki Ono,
Kingkarn Thepkanjana, Satoshi Uehara

Hituzi Syobo Publishing



Contents

Contents	v
About the Contributors	ix
Preface	xiii
 Chapter Introductions	 1
 Chapter 1 Syntactically naughty? : Prosody of final particles in Thai	 13
Pittayawat Pittayaporn and Pirachula Chulanon	
 Chapter 2 Etymological inquiry in a quest for the universality of human cognition: Focusing on word formation by sound symbolism	 29
Hiroyuki Eto	
 Chapter 3 Polyfunctionality in Pwo Karen: The case of ʔàʔ- (<T-B pronominal prefix *ʔa-)	 41
François Langella	
 Chapter 4 Global distribution of nominal plural reduplication	 57
Vipas Pothipath	

Chapter 3

Polyfunctionality in Pwo Karen: The case of ʔàʔ- (<T-B pronominal prefix *ʔa-)

François Langella

1. Introduction

Karenic languages are spoken in an area straddling the Thai-Burmese border. Following Thurgood's classification (2003), Pwo Karen belongs to the Southern group of the Karenic family. The dialect studied in this paper is that spoken in Dong Dam village, a village of about 400 inhabitants located in Lamphun province, Northern Thailand. I shall refer to this dialect as Dong Dam Pwo Karen.

The morpheme that is the topic of this paper has been a subject of continual interest in the field of Sino-Tibetan linguistics. The main reason for this is probably that reflexes of Proto-Tibeto-Burman *ʔa- are found in a wide array of Tibeto-Burman languages, and that they occur in various environments in these languages. This ubiquity and versatility have initially led Wolfenden (1929) to reconstruct two distinct proto-elements to account for this phenomenon. Such analysis has fallen out of favor, and empirical research based on a larger sample of languages has recognized that the various semantic functions of *ʔa- are all "outgrowths of one and the same proto-element" (Matisoff 2003: 104, whose reconstruction I follow. See also Benedict 1972, as well as Lehman 1975 who reaches similar conclusions).

In his *Handbook of Proto-Tibeto-Burman* (2003), Matisoff spells out six semantic functions attested across the Tibeto-Burman family, which I have used as a point of departure to investigate the range of functions covered by Dong Dam Pwo Karen ʔàʔ-. Of these six functions, three are not found in Dong Dam

Pwo Karen, of which two will be dealt with in Section 2. Drawing on Solnit's analysis of Eastern Kayah Li (1997), I have divided the remaining four functions into two groups: pronominal functions (Section 3) and non-pronominal functions (Section 4). Finally, in Section 5, I discuss whether this division makes sense in the case of Dong Dam Pwo Karen. Section 6 summarizes the findings presented in the paper.

Before starting this survey of the functions of Dong Dam Pwo Karen ʔàʔ-, some remarks on selected aspects of the language's morphosyntax are in order. Dong Dam Pwo Karen exhibits morphosyntactic features that are quite typical of Mainland Southeast Asian languages (Enfield 2005): it is largely isolating and analytic, has lexical tones and a numeral classifier system and makes extensive use of verb serialization.¹ More directly relevant to this study, all noun modifiers follow the head noun, except for possessive modifiers, which occur in phrase-initial position. Furthermore, nouns can be divided into free and bound nouns. A free noun can stand alone, meaning that it can head a noun phrase and fill an argument slot in a clause, as opposed to bound nouns, which need some kind of morphological prop to behave as free nouns. I have used a numeral classifier construction in order to distinguish between the two types of nouns. A free noun must be able to fill the empty slot in the following frame: _____ + Numeral + Classifier. For example, *dāuʔ* 'house' is a free noun since it can directly enter in construction with *khéi phlau* [two CLF], as in *dāuʔ khéi phlau* 'two houses'. On the other hand, *-kjùʔ* 'sheath' is a bound noun, since **kjùʔ khéi bei* (intended meaning: 'two sheaths') is syntactically ill-formed.

2. Functions not attested in Dong Dam Pwo Karen

2.1 With kinship terms

Matisoff (2003) reports that **ʔa-* is often found prefixed to kinship terms, both in their vocative (1) and referential (2) uses.

(1) *a-pa* 'Father!' (Lahu)

(2) *wa* or *əwa* 'father' (Kachin) (Benedict 1972: 121)

Dong Dam Pwo Karen ʔàʔ- does not occur with kinship terms in their vocative use. A child from Dong Dam village simply uses the prefix-less form

phá 'Father!' to call his father. ʔà- does occur with kinship terms in their referential uses but always anaphorically refers to a third person possessor: ʔà- phá thus means 'his/her/their father', not 'father'. Note that in Kachin, ə- alone does not carry this possessive meaning, and a distinct morpheme, kə-, is used if such a meaning is intended: kəwa 'his father' (Benedict 1972: 121).

2.2 "Aspectual" verbal prefix

In the "aspectual verbal prefix" function, *ʔa- is prefixed to a verbal root and involves valency-changing operations such as transitivity/intransitivity or causativizing/decausativizing processes (Matisoff 2003: 122). For example, the Jingpho verb rái 'to be' becomes ʔarái 'to arrange, to prepare', literally 'to make (something) come into being'. This function goes unreported for all the Karenic languages formerly described (Benjakul 1997; Jones 1961; Kaewsilpa 1982; Kato 2003; Solnit 1997), and Dong Dam Pwo Karen is no exception.

3. Pronominal functions

This section deals with two separate functions in Matisoff's classification. In both these functions, *ʔa- anaphorically refers to a third person entity, attached to a noun in one and to a verb in the other. Matisoff (2003: 121) calls these two functions "third person possessive" and "verbal prefix showing agreement to a third person subject", respectively. The third person possessive function is further divided into two subfunctions. In (3a), Lai Chin ʔa- is prefixed to the noun referring to the possessed entity, rool 'food', standing in for an already established possessor.² In (3b), Lahu ə- is prefixed to the possessed entity and preceded by the lexical possessor vâ? 'pig'.

(3) Third person possessive (examples from Matisoff 2003: 121)

- a. "Pronominal possessor": ʔa-rool 'his/her food' (Lai Chin)
- b. "Prefixed to the thing possessed": vâ? ə-šā 'the flesh/meat of a pig' (Lahu)

In (4), the Lai Chin prefix ʔa- occurs prefixed to a verb, anaphorically referring to a third person subject.

(4) ʔa-kal 'he/she goes' (Lai Chin, Matisoff 2003: 121)

3.1 Dong Dam Pwo Karen pronominal system

We now turn to Dong Dam Pwo Karen. The pronominal system of Dong Dam Pwo Karen is summarized in Table 1.

Table 1. Dong Dam Pwo Karen Pronominal System

Person	Possessor-N	Subject-V	V Object
1SG	(jā) cāʔ-N	(jā) cāʔ-V	V jā
1PL	(xuu) ʔāʔ-N	(xuu) ø-V	V xuu
2	(nē) nāʔ-N	(nē) nāʔ-V	V nē
3	ʔāʔ-N	ø-V	

A noticeable feature of the Dong Dam Pwo Karen pronominal system is the presence of two sets of pronouns, free pronouns *jā/xuu/nē* and bound pronouns *cāʔ-/nāʔ-/ʔāʔ-*. The frame test introduced in Section 1 equally applies to these pronouns. In the possessive constructions in (5), the second person free pronoun *nē* may be directly quantified by *khéi yá* (5a), while this cannot be so for the second person bound pronoun *nāʔ-* in (5b), which, indeed, forms an inseparable whole with the head noun *dāuʔ* ‘house’.

- (5) a. *nē khéi yā nāʔ-dāuʔ*
 2 two CLF 2-house
 ‘the house of you two’
- b. **nāʔ-khéi yā dāuʔ*
 2-two CLF house
 (Intended: ‘the house of you two’)

Bound and free pronouns thus contrast in their distribution: only free pronouns may occur as object arguments, while they are optional, and indeed often absent, when referring to the subject or possessor entity. Bound pronouns, in contrast, do not occur in object function, while they are obligatory in subject function and in possessive constructions, as shown in (6).

- (6) a. (jā) cāʔ-erj phāʔla ʔa
 1SG 1SG-eat betel a.lot
 ‘I eat a lot of betel.’

- b. *jā ēŋ phàʔla ʔa
 1sg eat betel a.lot
 (Intended: "I eat a lot of betel.")
- c. (jā) càʔ-dāuʔ
 1sg 1sg-house
 'my house'
- d. *jā dāuʔ
 1sg house
 (Intended: 'my house')

These distributional facts suggest that bound pronouns form a person-marking paradigm, close to what Bresnan and Mchombo (1987) call a system of "anaphoric agreement", as opposed to a "grammatical agreement" system. In grammatical agreement, found for example in French, a noun phrase bears the argument relation with the verb and an inflectional suffix marks agreement with this noun phrase on the verb root, in a somehow redundant manner. In anaphoric agreement, the noun phrase that may at first sight appears to function as the clausal subject "has only a non-argument function — either as an adjunct or as a topic or focus of the clause or discourse structure", while "the verbal affix functions as an incorporated pronominal argument of the verb" (Bresnan & Mchombo 1987, cited in Bhat 2004: 17). In a similar fashion, the bound pronoun in possessive constructions can be regarded as an "incorporated" possessor modifying the head noun, which can be optionally preceded by a noun or free pronoun specifying the identity of the possessor, presumably for pragmatic effects. A possessive construction can thus be described as the following schema (R) Pro-D, in which R stands for the possessor and D for possessed entity.

Another remarkable particularity of Dong Dam Pwo Karen is that it is asymmetrical in two respects. Firstly, free pronouns are asymmetrically distributed across the three persons. Free pronouns are only found to express the first and second persons: Dong Dam Pwo Karen does not have a third person free pronoun. Such a gap is not unusual in typological terms (Bhat 2004; Dixon 2010). Secondly, the third person bound pronoun ʔàʔ- is not used as an "incorporated pronominal argument to the verb" (7a), but only to signal a third person possessor (7b).

- (7) a. (phó i) ø-ɽeŋphàɽla ɽa
 grandma 3-eat betel a.lot
 “Grandma eats a lot of betel.”
- b. phlāu làɽ-ɽā no ɽàɽ-dāuɽ
 person one-CLF that 3-house
 ‘that person’s house’

Finally, a few words on number marking. In accordance with Corbett’s Animacy Hierarchy (Corbett 2000, cited in Bhat 2004), number marking in Dong Dam Pwo Karen is restricted to the existence of two first person pronouns, marked for number, singular *jā* and plural *xui*. Although the first person singular has a dedicated bound pronoun, *càɽ-*, there is no dedicated first person plural bound pronoun. Rather, it seems that the first person plural *xui* is treated as a third person pronoun, since it pairs up with the third person *ɽàɽ-* in possessive constructions, and the first person plural is zero-marked on the verb. Alternatively, one could argue that third person marking has been extended to the marking of the first person plural. It remains unclear at this stage which analysis should be preferred over the other.

3.2 Pronominal *ɽàɽ-* in Dong Dam Pwo Karen and other Karenic languages

The examination of the pronominal system reveals that only the third person possessive function is available to Dong Dam Pwo Karen *ɽàɽ-*. In this function *ɽàɽ-* is a member of the pronominal paradigm {*càɽ-*, *nàɽ-*, *ɽàɽ-*} marking reference to a third person possessor, which can optionally be specified before the prefixed head noun.

Table 2 recapitulates the facts presented in the previous sub-section and contrasts them with two other Karenic languages: Hpa-An Pwo Karen (Kato 2003) and Eastern Kayah Li (Solnit 1997).

Table 2. Comparison of Reflexes of TB *ɽa- in Three Karenic Languages

	Dong Dam PK <i>ɽàɽ-</i>	Hpa-An PK <i>ɽə-</i>	Eastern Kayah Li <i>ɽa</i>
Morphological status	bound form	bound form	free form
3 rd p. subject	no	in sub. clauses only	yes
In possessive constructions	R <i>ɽàɽ-</i> D	R (<i>ɽə-</i>)D	R <i>ɽa</i> D

Table 2 shows that the reflexes of Proto Tibeto-Burman *ʔa- in the two Pwo Karen dialects and Eastern Kayah Li contrast with each other in terms of their morphological status. Interestingly, this seems to correlate with their functional range, since only in Eastern Kayah Li, which only have free pronouns, is ʔa found to function in subject position without restrictions. Furthermore, Hpa-An Pwo Karen sits halfway between Dong Dam Pwo Karen and Eastern Kayah Li, since, on the one hand, ʔə- is found prefixed to the thing possessed (like Dong Dam Pwo Karen), but not compulsorily (unlike Dong Dam Pwo Karen), and, on the other hand, it may function as a subject (like Eastern Kayah Li), although only in subordinate clauses (unlike Eastern Kayah Li).

4. Non-pronominal functions

4.1 Verb nominalizer

In Tibeto-Burman languages, reflexes of *ʔa are frequently used to derive nouns from verbs, such as in Jingpho wák ‘to notch’ → ʔəwák ‘a notch’. In Dong Dam Pwo Karen, ʔàʔ- is hardly ever found in this use. The sole example I could find that resembles the Jingpho example is sa ‘to bear fruits’ → ʔàʔ-sa ‘the fruits (of a tree)’, illustrated in (8).

- (8) a. khu.sa ləʔ-thau jō sa jəuʔ
 mango one-CLF this bear.fruit already
 “This mango tree is already giving fruits.”
- b. khu.sa ləʔ-thau jō ʔàʔ-sa ʔo jəuʔ
 mango one-CLF this -fruit exist already
 “This mango tree, it already has fruits.”
 (Literally: “This mango tree, its fruits already are/exist.”)

However, (8) remains an isolated example in my data, which casts doubt on the nominalizing function of ʔàʔ- in this example. A safer option at this stage would be to consider that the morpheme *sa* can function both as a verb and as a noun. Consequently, I would rather treat ʔàʔ- in (8a) as a third person possessive pronoun prefixed to the head noun *sa* ‘fruits’, as the literal translation between parentheses suggests.

Much more widespread is the use of ʔàʔ- to derive free nouns from adject-

tival roots, with the meaning 'the adjV one', for example *du* 'big' → *ɔ̀à-ɔ̀du* 'the big one'.³ Evidence that the derived forms are nouns can be found in their ability to be quantified (9a), and indeed to take noun modifiers (9b) or function as core arguments (9c).

- (9) a. *ɔ̀à-ɔ̀du khéi phlau*
 -big two CLF
 'the two big ones (i.e. fish baskets)'
- b. *ɔ̀à-ɔ̀pi jō*
 -small this
 'this small one (i.e. fish basket)'
- c. *cà-ɔ̀-thái ɔ̀à-ɔ̀sɛŋ jəu*
 1sg-weave -new already
 'I've already woven a new one (i.e. lid of a fish basket)."

An important property of the derived nominals in (9) is that all refers to entities that the speaker assumes to be identifiable by the addressee, as a result of being possibly pointed at at the time of the utterance, such as in (9a) and (9b), or of having been previously mentioned, such as (9c), that answers the addressee's initial enquiry about the missing lid of a fish basket. This equally applies whether *ɔ̀à-* attaches to an adjectival root or to a bound noun (see Section 4.2).

4.2 Phonological "bulk-provider"

Before getting into the details of this function, let me first go back to Matisoff's (2003: 117) definition of phonological "bulk-provider", for the sake of clarity. Very frequently this prefix is added to roots that are already nouns, merely to give them a bit more of phonological bulk, providing them with the salience to serve as constituents in larger constructions.

According to Matisoff, this use is attested both with nouns that are already free nouns and with bound nouns. Furthermore, it may or may not involve "semantic specialization", meaning that it may or may not involve some sort of meaning increment to the noun it is attached to.⁴ When no semantic specialization is involved, the main function of **ɔ̀à-* may be to preserve the integrity of the

rythmic pattern of the utterance. An example provided by Matisoff is the Lahu pair *bo ~ ð-bo*, which both mean 'favor, grace, advantage'. In contrast, semantic specialization is at work in Burmese *ṛəmyak* 'knot in timber', derived from the noun *myak* 'eye'.

My data does not show any cases not involving semantic specialization, and all instances of *ṛàṛ-* do bring some kind of meaning increment to the affixed noun. However, this may merely reflect the fact that my data comes, for the most part, from elicitation sessions, which carries the risk that informants will focus on giving to each word its own semantic contribution to the sentence, thus downplaying the role of prosody.

Keeping this caveat in mind, uses of *ṛàṛ-* with both free and bound nouns always involve in my data a relational interpretation, meaning that the addressee is invited to construe the referent of the affixed noun in relation to some other entity. This typically involves a possessive relationship. To further the discussion, it is worth noting that the morphological status of a noun (free vs. bound) is not without consequences on the type of possessive relationship that it can enter into with a modifier noun. This is illustrated in example (10).

- (10) a. *pàṛnè ṛàṛ-rōŋ*
 buffalo -pen
 'the buffalo's pen'
- b. *ṛ sàuṛ ṛàṛ -rōŋ*
 Sau -pen
 (Intended meaning: 'Sau's pen', i.e. the pen that belongs to Sau.)
 (Actual meaning: 'Sau's pen', i.e. ṛ the pen to keep Sau in.)

(10a) and (10b) are two examples of possessive constructions following the schema [(possesso)R *ṛàṛ*-(possesse)D] presented in Section 3.1. The oddity of (10b) shows that when the D slot is filled by a bound noun such as *-rōŋ* 'pen', the range of semantic relationships holding between the possessor and the possessed entity is restricted. In other words, only a noun specifying the kind of possessed entity being talked about can fit the R slot. Therefore, when *ṛàṛ-* occurs prefixed to a bound noun it can only refer to a specifying noun other than a third person owner. Accordingly, *càṛ-rōŋ* [1sg-pen] and *nàṛ-rōŋ* [2-pen] are at best very odd, since they respectively mean that the speaker and the addressee

are the tenants of the pen.

To summarize, when ʔàʔ- is prefixed to a free noun, it can straightforwardly be interpreted as a third-person possessive pronoun, and a member of the three-member paradigm $\{\text{cà-}, \text{nà-}, \text{ʔàʔ-}\}$. Cases when it is affixed to a bound noun are more complex insofar as it is not anymore aligned with the first and second personal pronouns cà- and nà- . If ever a phonological “bulk-provider” function has to be distinguished, it is possibly in this use with bound nouns, which, as seen above, need some kind of prop, or “phonological bulk” to occur as autonomous constituents. A related question is whether ʔàʔ- remains pronominal in this function. This forms the topic of the next section.

5. Discussion: Pronominal versus non-pronominal ʔàʔ-

5.1 Eastern Kayah Li: pronominal ʔa and formative prefix ʔa-

In both Benedict’s *Sino-Tibetan Conspectus* (1972) and Matisoff’s *Handbook of Proto-Tibeto-Burman* (2003), ʔàʔ- and its cognate forms in other Tibeto-Burman languages are taken to be modern reflexes of a sole Tibeto-Burman pronominal element: $*\text{ʔa-}$. Lehman (1975) follows this view, but leaves open the possibility that two prefixes may be distinguished in modern Tibeto-Burman languages, on the basis of the functions they fulfill in these languages. Solnit’s analysis of Eastern Kayah Li (1997: 41–46) precisely echoes Lehman’s point of view, insofar as he distinguishes the “unmarked third person pronoun” ʔa from the formative prefix ʔa- , “which allows the [bound] Noun to occur as a free Form without adding any other semantic coloration”. Solnit (1997: 43) adduces the following example as evidence for the non-pronominal nature of ʔa- in Eastern Kayah Li.

- (11) $\text{vē} \quad \text{sīnē} \quad \text{ʔo}, \quad \text{manē} \quad \text{ʔaplo} \quad \text{ʔo} \quad \text{to}$
 1s gun exist but exist NEG
 ‘I have a gun, but no bullets.’

In (11), Solnit argues, “it would be inaccurate to analyze ʔaplo [...] as a possessive construction ‘its small-round-thing’”. Since -plo ‘small round thing’ is a bound noun and cannot function as a main constituent on its own, and since ʔaplo does not refer to the bullets of a particular gun such as the one referred to in the first clause but “simply names the general category ‘bullet’”, it follows

that ʔa- is no more than a “colorless” morphological device used to derive a free noun from a bound noun. The resulting form is thus best analyzed as “a sort of abbreviation of the compound noun *sine plɔ* ‘bullet’”.

Although the uses of reflexes of *ʔa- in Eastern Kayah Li and in Dong Dam Pwo Karen show similarities, the genetic relation between the two languages is of course no sufficient reason to replicate Solnit’s analysis with the Dong Dam Pwo Karen data. Whereas Solnit is able to differentiate two morphemes on the basis of their morphological status, free pronoun ʔa and prefix ʔa-, there is only one bound form ʔàʔ- in Dong Dam Pwo Karen.⁵ Beyond morphology, the pragmatic properties of free nouns derived from bound nominal roots provide evidence against distinguishing between a pronominal and a non-pronominal ʔàʔ-.

5.2 Pragmatic properties of constructions involving the prefix ʔàʔ-

The use of derived nouns involving the prefix ʔàʔ- is subject to a major discourse-level restriction, namely that the speaker assumes that his interlocutor will be able to identify which kind of entity he is referring to. I will refer to this constraint as the *identifiability constraint*.⁶ The following exchange, from which (9c) was extracted, took place between a basket weaver and a villager inquiring about the missing lid of a fish basket. The use of both ʔàʔ-phlī and ʔàʔ-séŋ is licensed by the shared knowledge of what kind of lid and which new thing, respectively, are being talked about.

- (12) A: ʔàʔ-phlī ʔo phàʔlē
 -lid be.at where
 B: yéyōŋ, càʔ-thái ʔàʔ-séŋ jāuʔ
 broken 1SG-weave -new already
 A: “Where is the/its lid?”
 B: “(It’s) broken. I have already woven a new one.”

Although I would cautiously reckon that more naturalistic data are needed to confirm my hypothesis, I am inclined to consider that, in utterances such as (12) ʔàʔ- retains its pronominal function, and that ‘its lid’ is the appropriate gloss for ʔàʔ-phlī.

Furthermore, it is not obvious that compounds of the form [free specifying noun-bound noun] can be analyzed as the full-fledged expressions of nouns

sharing the structure [ʔàʔ-bound noun], as Solnit proposes for Eastern Kayah Li. As illustrated in (13), such compounds contrast with possessive constructions, usually, but not systematically, involving a lesser degree of specificity: (14a) does so, (13b) may do so, but (13c) does not.⁷

- (13) a. chén-ja 'chicken' vs. chén ʔàʔ-ja 'the flesh, meat of a chicken'
 b. mi-kōŋ «a rice kratip» vs. mi ʔàʔ-kōŋ «a rice kratip, a kratip of rice»
 c. dāuʔ-khlāuʔ or dāu ʔàʔ-klāuʔ 'behind the house, the back of the house'

Finally, some high-frequency bound nouns are found to occur with other prefixes, such as tàʔ- or thàʔ-:

- (14) a. ʔàʔ-sa làʔ ʔo dāiʔ ba
 -fruits NEG exist yet NEG
 'It (i.e. the tree) is not giving fruits yet.'
 (Literally: "Its fruits do not exist yet.")

- b. tàʔ-sa làʔ ʔo ba
 fruits NEG exist NEG
 "There are no fruits."
 (Literally: "Fruits do not exist.")

- (15) a. ʔàʔ-ja 'its flesh, meat' ʔàʔ-dāi 'its eggs'
 b. thàʔ-ja 'meat' thàʔ-dāi 'eggs'

The contrast between the forms referenced under (a) and those referenced under (b) in (14) and (15) parallels that opposing the possessive constructions from the compounds in (13). For example, ʔàʔ-ja 'the meat, its meat' could refer to the meat or flesh of a particular animal, whereas thàʔ-ja 'meat' refers to meat in general.⁸ The origins of the prefixes tàʔ- and thàʔ- remains unsure, although it can be hypothesized that the latter thàʔ- comes from the coalescence of the free noun thə 'thing, stuff' and the pronominal prefix ʔàʔ-. Indeed, my informant once tried to make the meaning of thà-dāi 'eggs' more transparent to me by spelling it out as thə ʔàʔ-dāi 'eggs of something'.

6. Closing remarks

To summarize, I wish to present the findings of this paper in the form of a semantic map, following Haspelmath's notational convention (Haspelmath 2003). The square area in Figure 1 delimits the functional range of Dong Dam Pwo Karen ꠌꠌꠌ-.

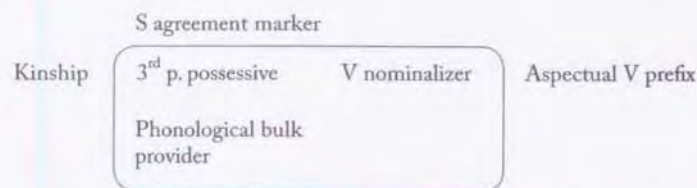


Figure 1. Semantic map showing the functional range of Dong Dam Pwo Karen ꠌꠌꠌ-.

Of the six functions attributed to modern reflexes of *ꠌa- across Tibeto-Burman languages, three go unattested in the case of Dong Dam Pwo Karen ꠌꠌꠌ-. These are represented outside of the square area: with kinship terms, as an aspectual verb prefix and as a verb prefix referring to a third person subject. As for the three remaining functions, ꠌꠌꠌ- appears primarily as a third person possessive pronoun, and I have attempted to show that it retains its possessive interpretation even with bound nouns, in constructions that resemble what Matisoff describe as involving the use of *ꠌa- as a “phonological bulk provider”. Finally, the function of verb nominalizer is represented within the functional range of ꠌꠌꠌ-, although a proviso is necessary: it was only found to occur with adjectival verbs, in a rather different configuration than in Matisoff's examples.

Acknowledgements

This research is part of the Karen Linguistics Project directed by Prof. Theraphan Luangthongkum. I am grateful to the Thailand Research Fund and to the Faculty of Arts, Chulalongkorn University, for providing financial support for my field trip in March 2010.

Notes

- 1 There are four lexical tones on non-checked syllables, while checked syllables (final glottal stop -ʔ) only take two tones, low and high. Transcription is as follows:

Mid tone	a (no diacritic)
Low tone	à
High tone	ā
High rising tone	á
- 2 Following Dixon's use (2010), the term "possession", as well as its derivatives "possessor" and "possessed entity", spans the range of semantic relationships cross-linguistically observed to be expressed through a single construction, such as kinship, whole-part relationships, hyponymy, etc.
- 3 Following Solnit (1997) on Eastern Kayah Li, I assume that words denoting properties, usually corresponding to the category "adjective" in English, form a subcategory of verbs in Dong Dam Pwo Karen, which I will refer to as "adjectival verbs", "adjV" in short.
- 4 I borrow the term "semantic specialization" from Matisoff (2003: 109), who doesn't use it in a technical sense. As suggested by an anonymous reviewer, the term "metaphorization" might be more adequate, i.e. a knot in timber is conceptualized as an eye.
- 5 However, Solnit (1997) notes that ʔa- is an unusual prefix, insofar as it is not unstressed, and does not undergo vowel harmony.
- 6 I voluntarily avoid the term "definiteness". I prefer the notion of "identifiability" which, following Lambrecht (1994), is assumed to be a universal cognitive category, which closely correlates to the language-specific syntactic category of "definiteness", in languages that have it.
- 7 In a similar fashion, Matisoff observes that in Lahu "free-headed compounds" (what I call here possessive constructions) of the type vāʔ ð-šā [pig ð-flesh] 'the pig's flesh, the flesh of a pig' are "more specific and definite in meaning [...] while the bound headed ones are vaguer and more general", e.g. vāʔ-šā 'pork' (Matisoff 1973: 71).
- 8 Interestingly, the form thāʔ- is also found attached to verbs: ʔe 'spicy' gives thāʔ-ʔe 'chilli' and ʔeŋ 'eat' gives thāʔ-ʔeŋ 'food'. The former example would contrast with ʔāʔ-ʔe 'the spicy one'.

References

- Benedict, Paul. 1972. *Sino-Tibetan: A conspectus*. Cambridge: Cambridge University Press.
- Benjakul, Lalin. 1997. *A syntactical study of the Pwo Karen dialect in Huay-Hom-Nok village, Tambon Tha Mae Lob, Mae Tha district, Lamphun province*. Bangkok: Mahidol University MA Thesis.
- Bhat, D. N. S. 2004. *Pronouns*. New York: Oxford University Press.
- Bresnan, Joan & Sam Mchombo. 1987. Topic, pronoun and agreement in Chichewa. *Language* 63. 741–782.
- Corbett, Greville. 2000. *Number*. Cambridge: Cambridge University Press.
- Dixon, Robert M. W. 2010. *Basic linguistic theory*. Vol.1 & 2. Oxford: Oxford University Press.
- Enfield, Nicholas J. 2005. Areal linguistics and mainland Southeast Asia. *Annual Review of Anthropology* 34. 181–206.

- Jones, Robert B. 1961. *Karen linguistic studies: Description, comparison, and texts*. (UC Publication in Linguistics 25). Berkeley & Los Angeles: University of California Press.
- Kaewsilpa, Chutima. 1982. *A description of Pho Karen: A Tibeto-Burman language in Thailand*. Bangkok: Mahidol University MA thesis.
- Kato, Atsuhiko. 2003. Pwo Karen. In Graham Thurgood & Randy LaPolla (eds), *The Sino-Tibetan languages*, 632–648. London: Routledge.
- Haspelmath, Martin. 2003. The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. In Michael Tomasello (ed.), *The new psychology of language*, vol. 2, 211–243. New-York: Erlbaum.
- Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.
- Lehman, Frederic K. 1975. Wolfenden's non-pronominal a- prefix in Tibeto-Burman. *Linguistics of the Tibeto-Burman Area* 2(1). 19–44.
- Matisoff, James. 1973. *The grammar of Lahu*. UC Publications in Linguistics 75. Berkeley & Los Angeles: University of California Press.
- Matisoff, James. 2003. *Handbook of Proto-Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction*. UC Publications in Linguistics 135. Berkeley & Los Angeles: University of California Press.
- Peansiri, Eknijom. 1982. *A study of informational structuring in Thai sentences*. Honolulu: University of Hawai'i PhD dissertation.
- Solnit, David. 1997. *Eastern Kayah Li: Grammar, texts, glossary*. Honolulu: University of Hawai'i Press.
- Thurgood, Graham. 2003. A subgrouping of the Sino-Tibetan languages: The interaction between language contact, change and inheritance. In Graham Thurgood and Randy LaPolla (eds), *The Sino-Tibetan languages*, 632–648. London: Routledge.
- Wolfenden, Stuart. 1929. *Outline of Tibeto-Burman linguistic morphology*. London: Royal Asiatic Society.

นายศุภณัฐ จิตวิริยพันธ์

**Ban Pa La-U Sgaw Karen tones: An analysis of semitones,
quadratic treadlines and coefficients**

BAN PA LA-U SGAW KAREN TONES: AN ANALYSIS OF SEMITONES, QUADRATIC TRENDLINES AND COEFFICIENTS¹

Sujinat Jitwiriyanont²

Abstract

Each dialect of the Sgaw Karen language has a different tonal system. Despite the different number of tones, all of the previous studies agree on the fact that all tones of Sgaw Karen are level tones. However, according to my phonological analysis of the tonal system of Ban Pa La-u Sgaw Karen, this dialect has a contour tone. The tonal system comprises four tones, i.e. /low/, /mid/, /high/ and /falling/ occurring in non-checked syllables. The high and low tones also have allotones in checked syllables. To confirm my analysis of the tonal system and the new finding of a contour tone in this Sgaw Karen dialect, an acoustic analysis of the tones occurring in citation forms was attempted. The analysis included three parts: (1) semitones to illustrate the pitch shape and height; (2) quadratic trendlines to indicate the direction and degree of pitch change; and (3) coefficients to show the

generalization of each tone and the discrimination among tones.

The fundamental frequencies in hertz of the four tones were measured and then converted into semitone values to help minimize the variations in the pronunciation of the six female speakers. In addition, time and semitone values were used to generate 2nd degree polynomial or quadratic equations and trendlines for which coefficients were plotted to model the pitch contour shapes.

The results revealed that: the low tone was mid-falling or low-falling with an obvious degree of pitch change and curved or linear pitch movement; the mid tone was mid level with a less obvious degree of pitch change and wide curved or linear pitch movement; the high tone was mid-high level with a lesser degree of pitch change and wide curved or linear pitch movement; and the falling tone was high-falling with an obvious degree of pitch change and curved pitch movement. In checked syllables, the low and high tones were realized as low-falling and high-falling respectively. The low tone had a greater degree of pitch change than the other; however, both had curved pitch movement.

1. Introduction

Sgaw Karen is a tonal language belonging to the Karenic branch, which is an affiliation within the Tibeto-Burman languages. Sgaw Karen speaking people outnumber those of the other Karenic languages spoken in Thailand, i.e. Pwo, Pa-O, Kayan, Kayah, and Kayaw. Sgaw Karen speaking people are widely distributed throughout the country. The Karen in Thailand are generally called “Kariang” which roughly refers to all

¹ This paper was presented at RGJ Seminar Series LXXXII on Southeast Asian Linguistics organized by the Research Institute for Languages and Cultures of Asia at Mahidol University. I am grateful to Dr. Chutamanee Onsuwan, the discussant of the session, for her comments and suggestions which have been useful for the revision.

² Ph.D. Candidate, Department of Linguistics, Faculty of Arts, Chulalongkorn University

Karenic groups. At Ban Pa La-u, Tambon Huay Sat Yai, Amphoe Hua Hin, Changwat Prachuap Khiri Khan (a province in the western part of Thailand), some Karen identify themselves as “Kariang”, and others as “Karang”. Following the linguistic fieldwork, we found that Kariang is Pwo Karen while Karang is Sgaw Karen.

Each dialect of the Sgaw Karen language has a different number of tones. Previous studies have reported three to four tones in Sgaw Karen, i.e. /high/, /mid/ and /low/ (Jones 1961a, Ratanakul 1986b) or /mid/, /breathy high/, /breathy low/ and /creaky low/ (Dhananjayananda 1983³) or /mid high/, /mid/, /mid-low/ and /low/ (Lar Baa 2001) in non-checked syllables and two in checked syllables, i.e. /high/ and /low/. In addition we observed in a Linguistic Field Methods class (November 2010 – February 2011) that the Sgaw Karen dialect spoken at Ban Huay Mi in Amphoe Pay, Changwat Mae Hong Son (a province in the northern part of Thailand) has two tones, i.e. /low/ and /high/. Despite the different number of tones, all of the previous studies agree on the fact that all tones in Sgaw Karen are level tones. However, according to my phonological analysis of the tonal system of

Ban Pa La-u Sgaw Karen, this dialect of Sgaw Karen has a contour tone, i.e. a falling tone. The objectives of this research are to attest by acoustic study that the Sgaw Karen dialect spoken at Ban Pa La-u has four tones in non-checked syllables and two allotones in checked syllables and to investigate the acoustic cues to tonal discrimination. This research will provide a new finding about a contour tone in Sgaw Karen and prove the effectiveness of applying quadratic equations to analyze the acoustic characteristics of tones.

2. Phonological sketch of Ban Pa La-u Sgaw Karen

The Sgaw Karen dialect spoken at Ban Pa La-u has 23 consonant phonemes which occur in the initial position. Only / ʔ / can occur in the final position. There are 11 vowel phonemes: 9 monophthongs and 2 diphthongs. The vowel length is not contrastive. The tonal system comprises 4 tones, i.e. /low/, /mid/, /high/ and /falling/ occurring in non-checked syllables. There are allotones of the high and low tones in checked syllables. The high and low tones occurring in checked syllables are analyzed as allotones of those in non-checked syllables instead of tonemes, for they are conditioned by syllable type.

³ The Sgaw Karen dialect spoken at Ban Huay Tom was originally reported by Dhananjayananda (1983) to have a tonal system comprising six contrastive tones. In this paper, I re-analyze the tonal system of Ban Huay Tom Karen to define four contrastive tones in non-checked syllables and two in checked syllables, i.e. the glottalized high and the glottalized low tones occurring in checked syllables have been re-analyzed as allotones of the high level and low level tones in non-checked syllables respectively, since they are conditioned by syllable type.

Consonants

	<i>Bilabial</i>	<i>Alveolar</i>	<i>Palatal</i>	<i>Velar</i>	<i>Glottal</i>
<i>Plosive</i>	p b ph	t d th	c ch	k kh	ʔ
<i>Nasal</i>	m	n	ɲ	ŋ	
<i>Fricative</i>		s		x ɣ	h
<i>Trill</i>		r			
<i>Approximant</i>	w	l	j		

Vowels

<i>Monophthongs</i>	Front	Central	Back
high	i	ɨ	u
mid	e	ə	o
low	ɛ	a	ɔ
<i>Diphthongs</i>	ai au		

Tones /low/⁴ /mid/ /high/ /falling/

⁴ The low tone is produced with a breathy voice, but the breathy voice is not marked because it can be predicted and phonation type is not contrastive in this language.

3. Methodology

I collected the data for phonological analysis in order to find a minimal set of tones to use as test words for acoustic study at Ban Pa La-u, Tambon Huay Sat Yai, Amphoe Hua Hin, Changwat Prachuap Khiri Khan. The test words were pronounced three times by six female native speakers of Sgaw Karen, ranging in age from 23 to 43. The total number of test tokens was 108 (6 speakers x 6 words⁵ x 3 times). The speakers were asked to say the test words in the sentence frame⁶:

/tɛ̃ wɛ̃ _____ tɔ̃? bɔ̃/
say COMP one time
'say the word _____ once'

The following are the test words of four lexical tones in non-checked syllables and two allotones in checked syllables:

Non-checked syllables

Low tone (Low) /lɔ̃/ 'play'
Mid tone (Mid) /lɔ/ 'down'
High tone (High) /lɔ́/ 'deceive'
Falling tone (Falling) /lɔ̌/ 'straw'

⁵ To avoid the influence of initial and final consonants on pitch, I used a minimal set, of which I could find only one set due to the time limitations of the fieldtrip, for non-checked syllables which have four tones, for acoustic analysis. Therefore, only six words including two analogous pairs in checked syllables were used.

⁶ There is no tone sandhi in this dialect of Sgaw Karen so this sentence frame can be used for eliciting the data.

Checked syllables⁷

Low tone (Low?⁸) /nòʔ/ 'Ms (title)'

High tone (High?) /nóʔ/ 'grass'

The Praat program version 5.2.11 was used for recording and F0 measurement. A high-quality desktop microphone was placed approximately 30 cm from the speakers. Mono recordings and a default sampling frequency of 44,100 hertz were selected. Regarding pitch analysis, the standard pitch range setting, which was from 75 to 500 hertz, was adapted as all of the speakers were female⁹. The floor was set to 100 hertz and the ceiling to 500 hertz.

Two F0 measurement tasks were done. The first was to analyze semitone values and the other to generate quadratic trendlines. Regarding the first F0 measurement task, the fundamental frequencies (F0) at 5 points of time for each vowel were measured at 0%, 25%, 50%, 75% and 100% from vowel onset to offset. Only 5 points of time, instead of 11 points, were measured because they did not show significant differences. Moreover, the elaborated pitch measurement was further done for the second F0 measurement task for which the fundamental frequency was measured every 0.01 second from vowel onset to offset.

⁷ Checked syllables with the initial consonant /l/ could not be found for the minimal set. Therefore /n/ was chosen because they are both voiced alveolar sounds.

⁸ The glottal stops (ʔ) are used to mark the allotones of the high and the low tones in checked syllables.

⁹ It was difficult to find male informants because they went out to work.

The measured fundamental frequencies in hertz were converted into semitone values. In fact, there are a number of psycho-acoustic pitch scales, apart from semitones, such as mels, Bark and ERB-rate. Semitones were used in this study owing to their effectiveness. Nolan (2007) has done an experimental evaluation of pitch scales and the results reveal that semitones most accurately reflect intuitions about the intonational equivalence of the subjects.

The formula was $\text{semitones} = 3.32 \times 12 \times \text{LOG}(\text{Hz to be translated} / \text{Hz reference level})$. On account of the different purposes of converting hertz into semitones, the reference levels used for plotting the semitone-value line graph and the scatter plot for generating quadratic trendlines were different. The line graph of semitones was to illustrate the overview of the tone shapes and the comparison of the pitch height of all tones, so the reference level in the case of the semitone line graph was the lowest pitch point among all tones which, in this study, was always that of the low tone in the non-checked syllable (as an example of Speaker 4 in Table 1). The fundamental frequencies

were converted into semitones and the lowest pitch was subtracted to zero to help minimize the variation in the pronunciation of the six female speakers, showing the different pitch height and range between each tone produced by the individual speaker and among the other speakers.

The other purpose of converting hertz into semitones was to show the degree of pitch change in each tone. Thus the reference level was the lowest pitch point of each tone (as an example of Speaker 4 in Table 2). When the lowest pitch point from each pitch measurement of each tone was normalized to zero, the maximum pitch became equal to the degree of pitch change. The normalized semitones showed the total amount of pitch change in each contour compared across tones and speakers. The scatter plots and the quadratic trendlines that best fitted the F0 measurements were generated by Microsoft Excel 2010. The 2nd degree polynomial or quadratic equation used to generate the trendline was $y = ax^2 + bx + c$. The a- and b-coefficients for each trendline were saved to create a model based on the relationship of the a- and b-coefficients from the equations.

Table 1 The semitone values of each tone produced by Speaker 4, prepared for a line graph. Zero was the reference level.

Semitones: Speaker 4						
Time	Low	Low?	Mid	High	High?	Falling
0%	4.782749	7.144141	2.657708	4.861412	7.349272	6.520224
25%	4.133223	6.862081	2.048872	4.651888	7.379556	6.263299
50%	3.209767	6.529904	1.560691	4.668621	7.272468	5.546492
75%	1.818102	5.560063	1.442503	4.473226	6.549445	3.430742
100%	0	3.208046	1.30989	4.140227	5.402263	2.37995

Table 2 The semitone values of each tone produced by Speaker 4, prepared for a scatter plot to generate the quadratic trendline. The zeroes were the reference level of each tone.

Normalized Semitones : Speaker 4						
Time (sec)	Low	Low?	Mid	High	High?	Falling
0	3.905162	2.086838	1.061225	0.273398	1.858174	4.078882
0.01	3.819686	2.071617	0.911586	0.258577	1.736927	4.077479
0.02	3.632387	2.054405	0.514968	0.310423	1.617278	3.997847
0.03	3.477174	2.049013	0.336621	0.313052	1.514942	3.908419
0.04	3.337588	2.024117	0.363141	0.347427	1.408199	3.78989
0.05	3.043458	1.957536	0.267551	0.329549	1.291427	3.658584
0.06	2.814755	1.784974	0.232249	0.273675	1.124818	3.539144
0.07	2.635281	1.538286	0.121453	0.25557	0.790055	3.423321
0.08	2.554128	1.205144	0.104036	0.199834	0	3.261531
0.09	2.094947	0.717159	0.06911	0.125465		3.075494
0.1	1.914202	0	0	0.086418		2.781599
0.11	1.514439			0		2.41639
0.12	0.994252					1.9926
0.13	0.630345					1.560399
0.14	0					0.929184
0.15						0.588183
0.16						0.325644
0.17						0

4. Results

4.1 Semitones

The semitones converted from the average fundamental frequencies of each tone produced by each Pa La-u Sgaw Karen speaker were plotted using line graphs

illustrating the phonetic realization of the four lexical tones.

The six line graphs show the overview of the four tones in non-checked syllables and two in checked syllables. In spite of some differences in details among speakers, they reflect a similar pattern (see Figure 1).

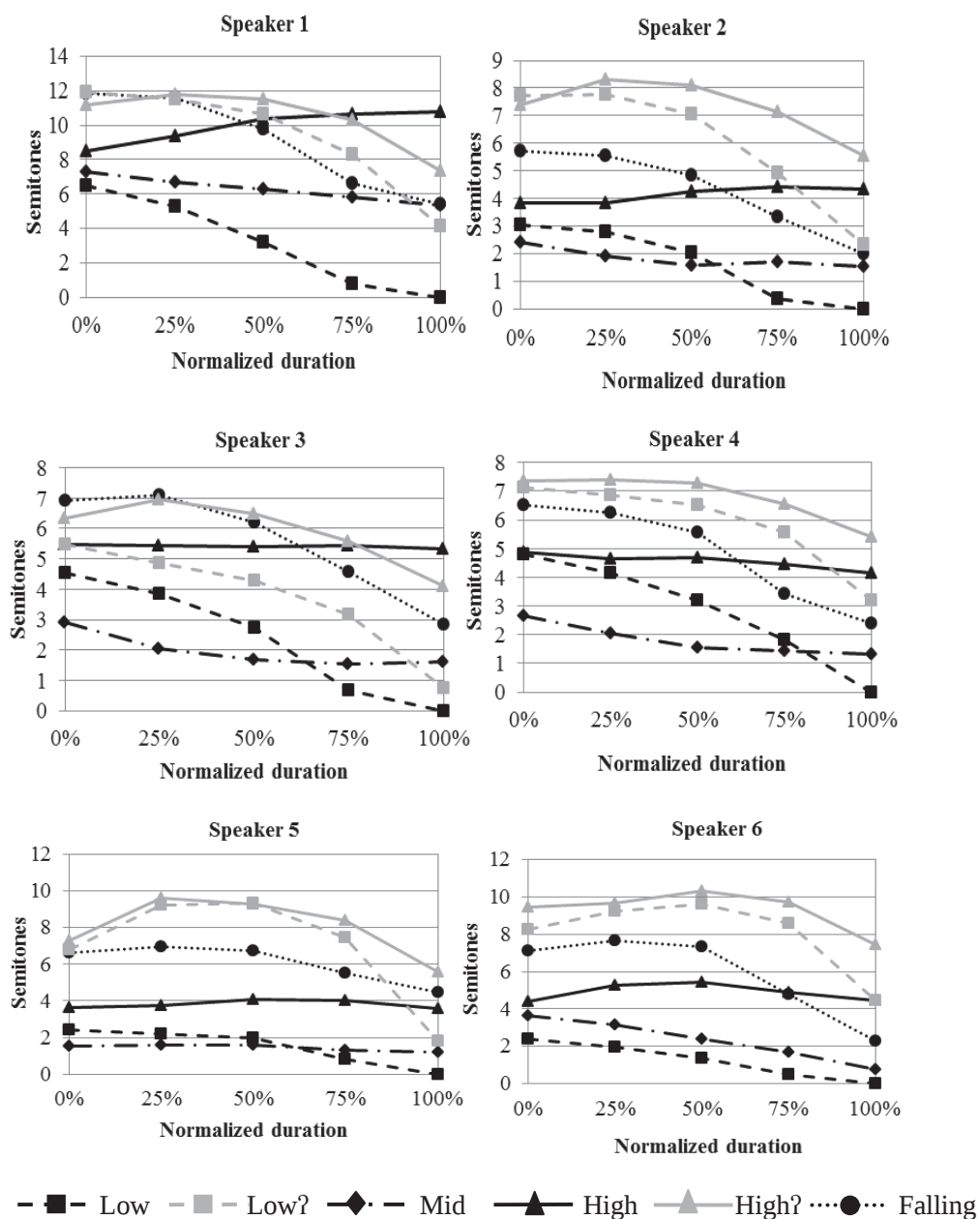


Figure 1 Semitones of the Pa La-u four lexical tones

In Figure 1, the result of the F0 measurements indicates that the four lexical tones in Pa La-u Sgaw Karen can be categorized by pitch height and contour. To begin with the non-checked syllables, the generalization was that the *low tone* was mid-falling or low-falling. It began at mid or low range, and fell steadily to the lowest point of the semitone scale. The *mid tone* was mid level or low level beginning at the mid or low range and staying level or falling slightly. The *high tone* was mid-high level. It began with mid or mid-high pitch and stayed relatively level in that it rose or fell for not more than one semitone scale except for that of Speaker 1. Despite three semitone scales of pitch contour, the high tone of Speaker 1 could be concluded to have a small amount of contour similar to the other five speakers due to the widest pitch range. This is supported by the finding of Dilley (2005), suggesting that tone is relative in a dualistic way, one part of which is that tones are scaled relative to an individual's pitch range. Moreover, although the high tone of some speakers began at the mid range similar to the low and the mid tone, it was noted that the pitch of the high tone from 25% to 100% was always higher than the other two tones. The *falling tone* was high-falling. It began at a high range then either fell to mid range (Speaker 1, 2 and 4) or rose to the highest point at 25% before falling to mid range (Speaker 3, 5 and 6). Furthermore, the pitch shapes of the allotones of the high and the low tones in the checked syllables were either falling or rising-falling. Although they shared a similar pitch contour, they were differentiated by the endpoint. The high tone always ended higher than the low tone.

According to the semitones of the four lexical tones in both non-checked and checked syllables, the phonetic characteristics of many tones, i.e. the low tones in both non-checked and checked syllables, the high tone in checked syllables, and the falling tone were of a falling or convex shape. Moreover, it was speculated that duration was relevant to tonal identity. Among the tones with phonetically falling shape, the tone with shorter duration would fall more steeply than the longer counterpart. Nevertheless, the different durations of the tones were ignored in the semitone line graph owing to the normalized duration represented by the x-axis. Therefore, it was interesting to use the method proposed by Andruski and Costello (2004) to investigate the pitch contour cues for tone discrimination in Pa La-U Sgaw Karen as the method revealed the details of pitch contour and the durations of tone in real time. Andruski and Costello (2004) noted that "in languages with a crowded tonal space, multiple tones can have similar contours. Even in languages with few tones, details of contour shape may be used by listeners for tone identification." They proposed the method of using polynomial equations to model pitch contour shape in lexical tones. Green Mong was used as a test case. In the paper, two kinds of polynomial equations, namely linear equations (1st degree polynomial equations) and quadratic equations (2nd degree polynomial equations) were used to test their hypothesis. It was found that discriminant analysis using the coefficients as predictor variables showed a greater effectiveness of using quadratic coefficients for tone classification than linear coefficients.

The method of using quadratic equations and trendlines to analyze lexical tones proposed by Andruski and Costello (2004) was adopted in this study in order to investigate the pitch contour cues for tone discrimination in Pa La-u Sgaw Karen.

4.2 Quadratic trendlines

A quadratic equation can be applied in order to analyze lexical tones, especially contour tones, in languages. The output of the quadratic equation is the trendline which can be used to describe and compare the details of contour shape across tones. The quadratic trendline shows the direction and the degree of pitch change, which pinpoints some interesting pitch contour cues compared with the semitone-value line graph. The information carried by the trendlines elaborates the phonetic realization of the lexical tones.

A quadratic equation is a polynomial equation of the second degree. The form of quadratic equation used in the study of Andruski and Costello (2004) was $y = a + bx + cx^2$. However, this current study followed the form of quadratic equation generally used in mathematic literature i.e. $y = ax^2 + bx + c$ where a and b are coefficients. Therefore, the a -coefficient in this study was the c -coefficient in Andruski and Costello's study. The trendline generated from a quadratic equation is a graph in the form of a parabola. The parabola can open either upward or downward as in Figure 2.

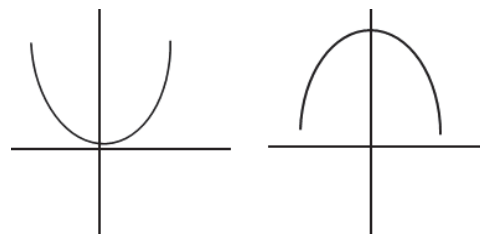


Figure 2 A parabola opening upward (left) and a parabola opening downward (right)

However in the case of analyzing tone or lexical pitch in languages, the quadratic trendline generated from the value of time and semitones is always half of the parabola. This is because the x -axis which represents time value is never negative.

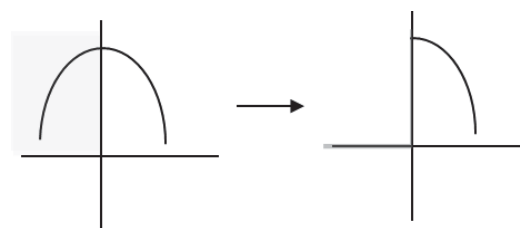


Figure 3 A half parabola due to the positive value of the x -axis representing duration

In this paper, the values of time and semitones were used to generate 2nd degree polynomial or quadratic equations and trendlines for each actual pitch contour. As in the example of Speaker 1 in Figure 4, the y -axis shows the semitone scale for interpreting the direction and degree of pitch change while the x -axis shows the duration in seconds.

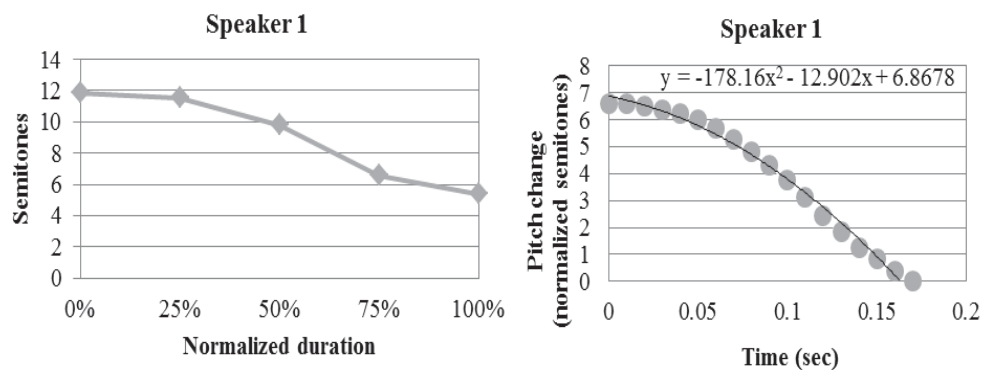


Figure 4 The comparison of the semitone-value line graph (left) and a quadratic trendline fitted to average fundamental frequencies of the falling tone (right) produced by Speaker 1

In the case of Speaker 1's falling tone, the quadratic equation for generating this trendline was $y = -178.1x^2 - 12.90x + 6.867$ in the form $y = ax^2 + bx + c$. The y-axis of the line graph (left) represents the level of pitch height whereas the y-axis of the scatter plot (right) represents the degree of pitch change. The line graph indicates that the pitch height of the starting point of the falling tone produced by Speaker 1 was approximately 12 semitones and the lowest point at the end was approximately 6 semitones. This conforms to the information conveyed by the trendline of the scatter plot in that the y-axis indicated an amount of pitch change of approximately 6 semitones. While the x-axis of the line graph represents the normalized duration, the other represents the real duration indicating a duration of approximately 0.17 second. The semitone-value line graph and the trendline both suggest a similar phonetic realization in terms of the pitch shape of the falling tone produced by Speaker 1 in that the line graph illustrates the overall pitch shape of the falling tone as high falling and the trendline also shows the direction of the pitch

movement as a convex or falling contour.

As in Figures 5-10, it is important to note that the y-axis of the quadratic trendline graph on the right does not represent the pitch height as it traditionally does in the line graph of hertz or semitone values but the degree of pitch change or contour. Besides, the x-axis of the quadratic trendline graph on the right represents real duration (seconds) instead of normalized duration (5 equidistant timepoints) as in the line graph of semitone values on the left.

The quadratic trendlines of the four lexical tones produced by 6 speakers can be classified into 3 groups based on the patterns. The main difference is the pitch movement of the high, mid and low tones in non-checked syllables. The other tones, i.e. the falling tone in non-checked and the high and low tones in checked syllables, all had a convex pitch contour.

Pattern 1: Linear pitch movement (S1)

The uniqueness of the first pattern was that, except for the falling tone, all tones in non-

checked syllables had a linear pitch movement. In other words, the pitch movement of the low, mid, and high tones in non-checked syllables was a straight line (see Figure 5).

As in Figure 5, the trendline on the right shows that the falling tone had the largest amount of pitch change and the mid tone had the smallest. Although the low tone in non-checked syllables had the lowest starting point, illustrated by the semitone-value line graph on the left, the amount of pitch change was not the smallest. According to the trendlines, the pitch movement of the falling tone, the low, and the high tones in checked syllables was convex or falling. The pitch movement of the low tone, mid tone, and high tone in non-checked syllables was a straight line. This suggests that in spite of the contour shape of the low tone resembling the falling

tone in the semitone-value line graph, the overall pitch movement was similar to that of the level tone. This was due to the different duration of the low and falling tones. The low tone had approximately 0.1 seconds of duration more than the falling tone. Longer duration caused the gradual pitch fall; consequently, the overall pitch movement was more like a level tone than a contour tone.

Pattern 2: Curved pitch movement (S2-S5)

All of the tones in the second pattern had a non-linear pitch movement. Like the other patterns, the falling tone and the two tones in checked syllables were convex. The low, mid, and high tones, in contrast to Pattern 1, had a curved pitch movement – both convex and concave. The different shape of curvature brought about an insignificant difference due to the small degree of pitch change (see Figures 6-9).

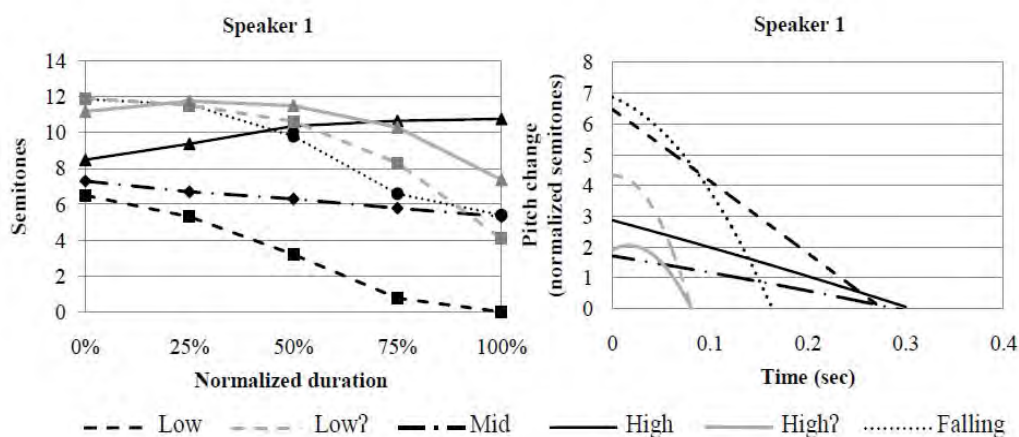


Figure 5 A comparison of semitones of the Pa La-u four lexical tones (left) and quadratic trendlines generated from semitones (right) produced by Speaker 1

As illustrated in Figures 6-9, the tones produced by Speakers 2 – 5 had nearly the same pitch-contour pattern. The trendlines illustrate that no tones had a pitch movement in an exact straight line. The semitone-value line graph shows that the overall tone shapes of the low tone and the falling tone in non-checked syllables were of a similar contour shape as for Speaker 1 in Figure 5. Nonetheless, unlike the linear trendline of Speaker 1's low tone, the trendlines showing the pitch movement of the low tones of this pattern were convex. With regard to the other tones in non-checked syllables, the pitch movement of the mid tone was a wide concave (S2,S3,S4) or a wide convex (S5); and that of the high tone was a wide convex (S2,4,5) or an extremely wide concave

which resembled a straight line but not exactly. (S3). However, the degree of the pitch change was less than 1 semitone, leading to an insignificant contour, and they were reasonably categorized as level tones conforming to the overall pitch shape in the semitone-value line graph. The falling tone was obviously realized as a falling contour. The pitch-change range was approximately 3-4 semitones except for Speaker 5. What differentiated Speaker 5's falling tone from the low tone, which had nearly the same amount of pitch change, was a rise-fall pitch contour and shorter duration causing a steeper contour. In checked syllables, both the semitone-value line graph and trendline showed that the low and high tones were falling and the low tone had greater degrees of pitch excursion.

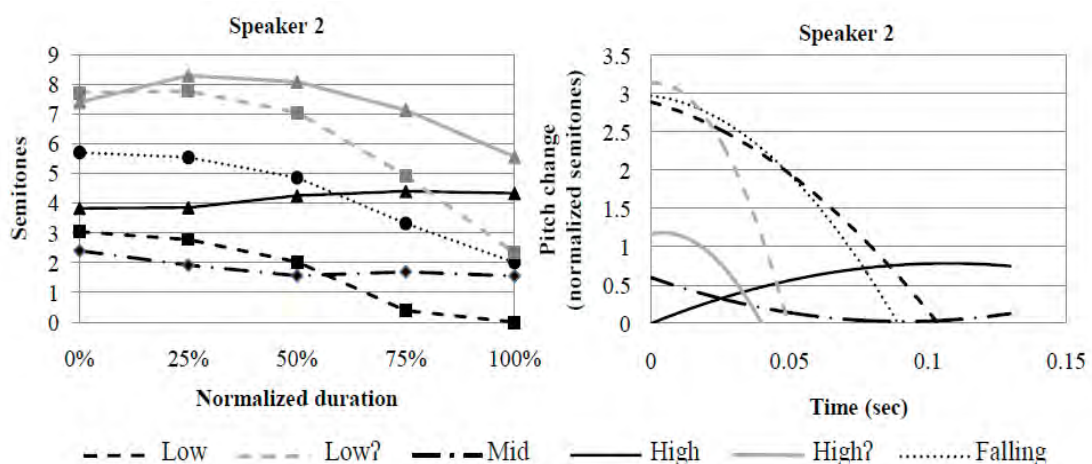


Figure 6 A comparison of semitones of the lexical tones in Pa La-u four (left) and quadratic trendlines generated from semitones (right) produced by Speaker 2

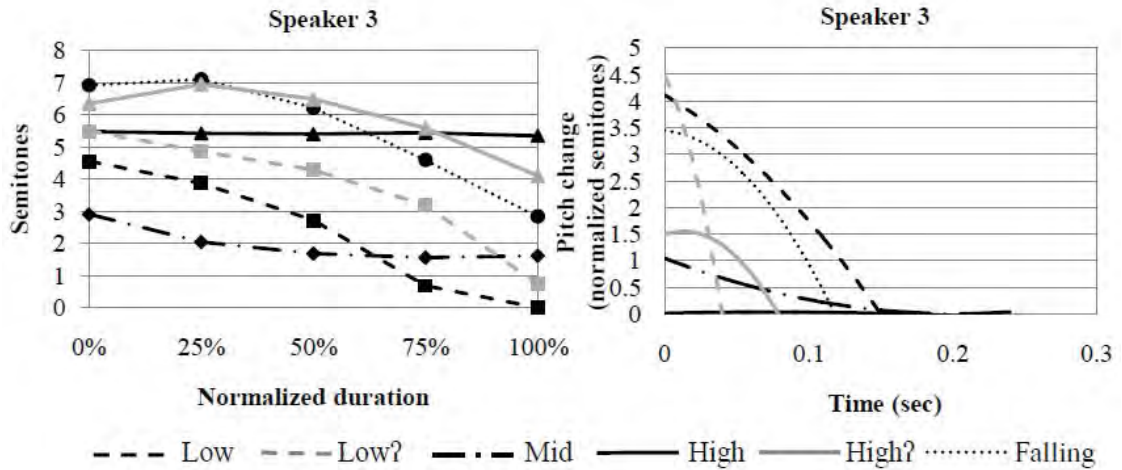


Figure 7 A comparison of semitones of the lexical tones in Pa La-u four (left) and quadratic trendlines generated from semitones (right) produced by Speaker 3

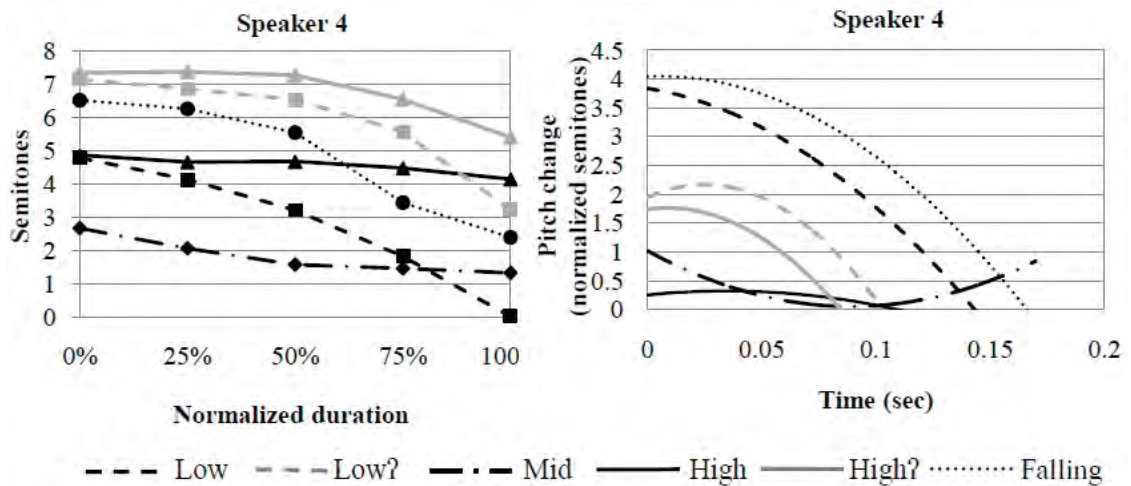


Figure 8 A comparison of semitones of the lexical tones in Pa La-u four (left) and quadratic trendlines generated from semitones (right) produced by Speaker 4

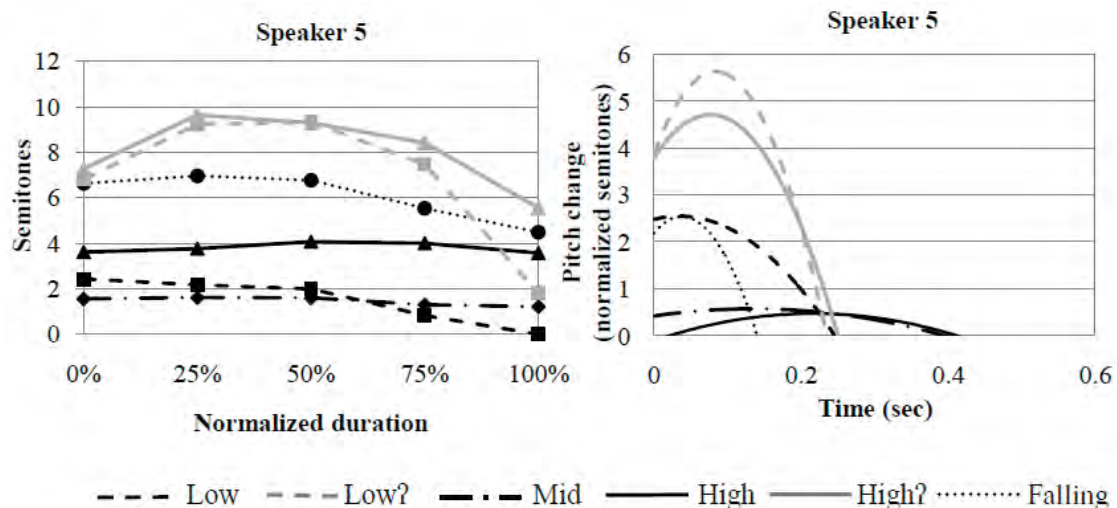


Figure 9 A comparison of semitones of the lexical tones in Pa La-u four (left) and quadratic trendlines generated from semitones (right) produced by Speaker 5

Pattern 3: Mixed type (S6)

The last pattern combined Patterns 1 and 2 in that the mid and low tones in non-checked syllables had a straight-line pitch movement but the high tone in non-checked syllables had a curved pitch movement as shown in Figure 10.

As seen in Figure 10, the trendline of the falling tone shows the largest amount of pitch change and the pitch contour shape as high falling. Due to the small pitch change below 1 semitone and the relatively long duration compared to the other tones, the high tone was categorized as a level tone, conforming to the phonetic realization illustrated by the semitone-value line graph. The pitch movement of the low and mid tones in non-checked syllables was a straight line, suggesting the characteristics of a level tone. In checked syllables, the low tone had

much more pitch excursion than the high tone; both were realized as falling.

To sum up, the generalization is that in non-checked syllables, the low tones whose phonetic realizations were mid falling or low falling, as the line graphs showing the semitones on the left of Figures 5-10 indicate, had three characteristics of trendlines: (1) the pitch movement was a straight line similar to the level tone; (2) the pitch movement was convex with wide curvature nearly similar to a straight line; (3) the pitch movement was convex but less obvious than the falling tone, which was relevant to the longer duration of the low tone. This suggests that the characteristics of the low tone were fairly level, similar to the impressionistic description.

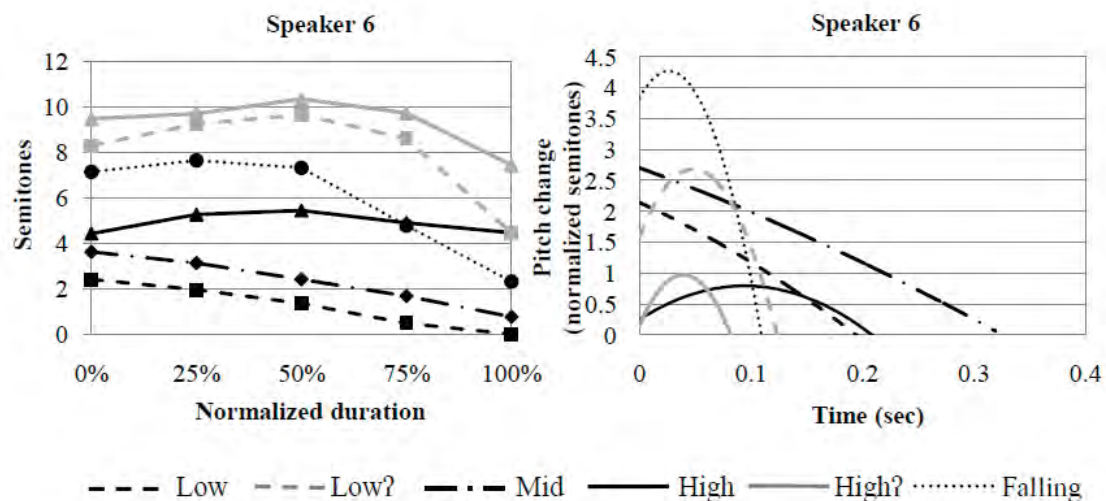


Figure 10 A comparison of semitones of the Pa La-u four lexical tones (left) and quadratic trendlines generated from semitones (right) produced by Speaker 6

The trendlines of the mid tone and the high tone, which were level tones, in non-checked syllables were not always a straight line. In spite of some curved trendlines, the range of the pitch change showing in semitones was narrow, resulting in a less obvious pitch contour. The contour was not significant when the degree of pitch movement was very little (not more than 1 semitone). Likewise, the results emphasize their realization as level tones. With regard to the falling tone, the semitone-value line graph together with the trendline unanimously confirmed the high-falling pitch contour.

In checked syllables the high tone and the low tone both had a falling pitch contour indicated by the direction and shape of the trendlines. The two tones were distinguished by the amount of pitch change. The pitch-change range of the low tones was larger

than the other. The results support the conclusion that the trendlines illustrate more detailed information about contour shape.

Table 3 shows the phonetic realization of the four lexical tones by synthesizing the information conveyed by both semitone-value line graph and quadratic trendline.

4.3 Coefficients

As the trendlines of tones across speakers were not identical, the line graphs cannot show the generalization and discrimination of tones well. Therefore, the a- and b-coefficients of the quadratic equations were further used to generate a scatter plot in order to create a model for tonal identity. The model reflects the generalization of each tone and discrimination among tones.

Table 3 Phonetic realization of the four lexical tones

Lexical tones	Overall shape	Starting point	Endpoint	Degree of pitch change	Duration	Pitch movement
Low	falling	mid/low	lowest	Obvious ¹⁰	long	linear / curved
Mid	level	mid	mid	less obvious	long	linear / wide curved
High	level	mid-high	high	less obvious	long	linear / wide curved
Falling	falling	high	low	obvious	short	curved
Low?	falling	high	higher than high?	more than high?	short	curved
High?	falling	high	lower than low?	less than low?	short	curved

This model based on the relationship of the a-and b-coefficients from the quadratic equations reflects the aspects of pitch contour shape distinguishing one tone category from another. The b-coefficient indicates the initial slope at the intercept. With the positive b-coefficient, the initial slope of the contour is upward, and the negative b-coefficient causes a downward initial slope (see Figure 11).



Figure 11 Information carried by the b-coefficients, the dotted circle indicating the initial slope

The a-coefficient indicates whether the pitch contour is concave or convex. A positive a-coefficient brings about a concave contour and a negative result in a convex contour. In addition, when the value of the a-coefficient is zero, its contour is a straight line. The amount of the absolute value of a-coefficients also indicates the size of the curvature. The larger the absolute value of a-coefficient, the steeper the pitch contour becomes (see Figure 12).

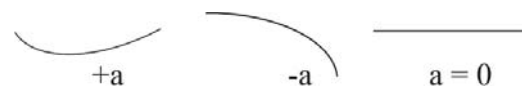


Figure 12 Information carried by the a-coefficients

¹⁰ The criterion for identifying the “obvious” pitch change was a pitch-change degree of more than 3 semitones for linear pitch movement and more than 1 semitone for curved pitch movement

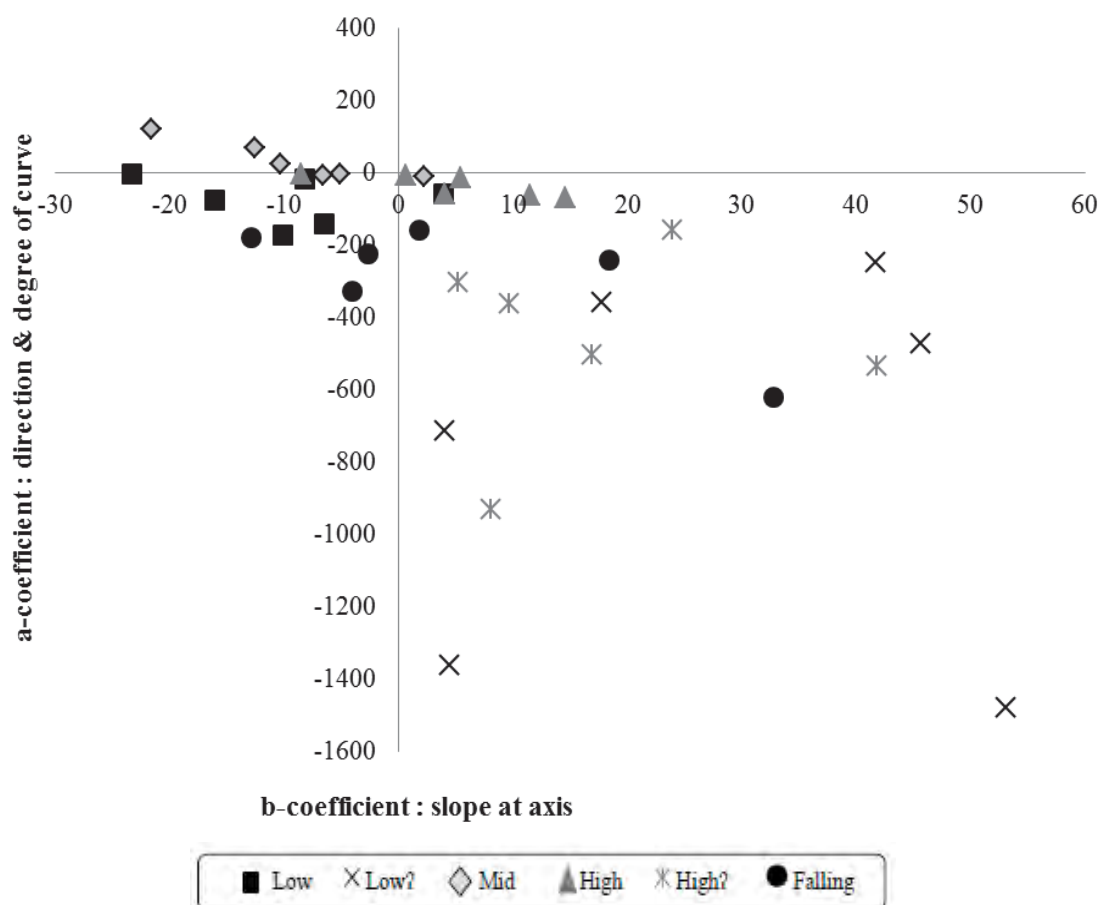


Figure 13 Relationship of the a- and b-coefficients from the quadratic equations

The a- and b- coefficients from all trendlines were plotted in the scatter plot to show the clustering and scattering of the tones suggesting the generalization of each tone and discrimination across tones (see Figure 13).

In Figure 13, the x-axis represents the b-coefficient, and the y-axis represents the a-coefficient. There are no tones in the upper right quadrant because positive a- and b-coefficients indicate a rising contour. The

rising tone is not a phoneme in the tonal system of Pa La-u Sgaw Karen. Only the mid and high tones in non-checked syllables are level owing to the value of a-coefficients (y-axis) near zero. Although the high tone has a slightly upward initial slope whereas the mid tone has a slightly downward initial slope due to the different positive/negative values of b-coefficient, they both are categorized as level tones since the contour is slight and their overall shapes are relatively straight lines.

The model reflects the different falling or convex shapes of the low (in non-checked syllables) and falling tones. The low tones have a progressive fall indicated by negative b-coefficient values. Although one token of the low tones has a positive b-coefficient value, the absolute value is so small that it does not show a significant upward initial slope.

The falling tones have a rise-fall pattern and a steep fall. The rise-fall pattern is specified by a positive b-coefficient value. Compared to the low tones, all of the falling tones have larger absolute values of a-coefficients, pointing out the steeper contours. In addition, having smaller absolute values of a-coefficients, the majority of the low tones fall very close to the same space occupied by the mid and high tones in non-checked syllables. Therefore, some of the low tones have a contour shape resembling the level tone. These values reflect the fact that the falling tone has a more obvious contour pattern than the low tone in non-checked syllables.

Of all the tones, the low and high tones in checked syllables show the most obvious rise-fall pattern. All have positive b-coefficient values and most have extreme negative a-coefficient values. However, the low tones have steeper contours as the absolute values of a-coefficients are larger.

5. Conclusion and discussion

Acoustic analysis confirmed the characteristics of Pa La-u Sgaw Karen's four lexical tones consisting of three level tones and one contour tone. The line graph of semitones, the quadratic trendlines, and the model based on the

relationships of coefficients indicate the complex of acoustic cues to tonal identity. Tones in the Pa La-u Sgaw dialect are well-defined by pitch shape, starting point, endpoint, degree of pitch change, duration, pitch movement, and slope.

This research supports the effectiveness of applying quadratic equations to analyze the acoustic characteristics of tones in language as suggested by Andruski and Costello (2004). In addition, the values of quadratic coefficients can further be plotted in a scatter plot to reflect the generalization of the individual tone together with the discrimination among tones by showing clustering and scattering in a particular area.

In fact, an acoustic analysis of the tones in connected speech was also performed but is not included in this paper because it did not show any significant pattern. The variability of the tones in connected speech could have resulted from different phonetic environments. Moreover, the data is too small and thus, inadequate for statistical analysis.

Acknowledgements

I would like to express my gratitude to Prof. Dr. Theraphan Luangthongkum for her supervision and comments. Financial support from the Thailand Research Fund through the Royal Golden Jubilee Ph.D. Program and the Faculty of Arts, Chulalongkorn University (Grant No. PHD/0201/2552) to Prof. Dr. Theraphan Luangthongkum and Sujinat Jitwiriyanont is acknowledged. My special thanks go to my friends in the Department of Linguistics and the Sgaw Karen at Ban Pa La-u.

References

- Andruski, Jean E. and James Costello. 2004. Using polynomial equations to model pitch contour shape in lexical tones. *Journal of the International Phonetic Association*, 34, 125–140.
- Dhananjayananda, Puttachart. 1983. *The phonology of Sgaw Karen, with comparisons with Thai*. MA thesis. Mahidol University, Nakhon Pathom
- Dilley, Laura C. 2005. *The phonetics and phonology of tonal systems*. Ph.D. Dissertation, MIT,
- Jones, Robert B. 1961a. *Karen Linguistic Studies: Description, Comparison and Texts*. Berkeley: University of California Press.
- Lar Baa, Saw. 2001. *The phonological basis of a Northwest Karenic orthography*. M.A. thesis. Payap University, Chang Mai
- Nolan, F. 2007. Intonational equivalence: an experimental evaluation of pitch scales. In *Proceedings of the 15th International Congress of Phonetic Sciences*; Barcelona, Spain, August 3-9, 2003.
- Ratanakul, Suriya. 1986b. *Thai-Sgaw Karen dictionary*. Bangkok: Institute of Language and Culture for Rural Development.

นางสาวกานต์ธิดา เกิดผล

**Formant transitions as effective cues to differentiate the
places of articulation of Ban Pa La-u Sgaw Karen nasals**

**Formant transitions as effective cues to differentiate the places of articulation of
Ban Pa La-u Sgaw Karen nasals¹**

Karnthida Kerdpol

Department of Linguistics, Faculty of Arts, Chulalongkorn University

Abstract

The Sgaw Karen dialect of Ban Pa La-u, Amphoe Hua Hin, Thailand, has four nasals: /m/, /n/, /ɲ/ and /ŋ/, that appear in syllable-initial position. Review of the relevant literature indicates that initial /ɲ/ has been less studied acoustically due to the lack of palatal nasals in the consonant systems of most languages. Thus, this Sgaw Karen dialect is suitable for investigating the place of articulation of nasals.

The acoustic characteristics examined include the duration, intensity, and frequencies of the formants (resonances) of the nasal murmurs as well as the frequencies of the formant transitions into the following vowels /a/ and /ɔ/. The significance of each acoustic characteristic as a place cue has been statistically tested with ANOVA and Tukey's HSD.

The results confirm the previous findings (Liberman, Delattre, Cooper, and Gerstman, 1954; Malécot, 1956; Recasens, 1983; Harding and Meyer, 2003) that transitions provide better cues for differentiating place of articulation for nasals. Furthermore, this study found that among the formant frequencies in a formant transition, the F2 transition provides the most effective cue to identifying the places of nasal articulation, i.e. /ɲ/ has the highest F2 frequency value at the nasal-vowel juncture, followed by /n/, /ŋ/ and /m/, respectively. The relational pattern between F2-F3 transitional directions can also aid in differentiating nasal articulation places; however, the pattern of transitional direction depends on vocalic context. The F2-F3 transitional patterns among the places of articulation clearly differ in the /ɲ/ context. In the case of a

¹ This paper was presented at the RGJ Seminar Series LXXXII: Southeast Asian Linguistics organized by The Royal Golden Jubilee Ph.D. Program (RGJ), Thailand Research Fund and Research Institute for Languages and Cultures of Asia, Mahidol University on August 5, 2011. I would like to express my gratitude to Dr. Chutamanee Onsuwan who was the discussant of the session for her valuable comments and suggestions.

following /a/, the F2–F3 transitional patterns for /n-/ and /ŋ-/ are very similar and do not act as a place cue. Although the second nasal formant (NF2) evinces consistent relational patterns, differences among /m/, /n/, and /ŋ/ are not statistically significant, implying their similarity. Likewise, neither intensity nor duration of nasal murmurs can be used as cues to differentiate place of articulation for nasals.

1. Introduction

In the production of nasal sounds, the oral and nasal cavities are coupled by the lowering of the velum. As a result, the acoustic characteristics of nasal sounds are more complex than those of purely oral sounds. In the case of nasal consonants, there is an obstruction in the oral cavity, but the velum is lowered, allowing air to flow through the nasal cavity. The nasal cavity acts as the main resonator, while the oral cavity acts as a side branch, absorbing sound energy. According to Ohala (1975), since the nasal cavity is fixed in shape and size, the nasal formants caused by resonance in the nasal cavity are similar for the various places of articulation. The first nasal formant (NF1) is very low in frequency. Formants above the first nasal formant are low in energy. On the other hand, the antiformant corresponding to sound energy being absorbed in the oral cavity varies from one place to another. The shorter the oral tract is, the higher the antiformant frequency value becomes. The antiformant can, then, serve to differentiate place of articulation for the various nasal stops.

Place contrast for nasals has been acoustically and perceptually studied to find what the best cue to place of articulation is. Some studies have found that nasal murmurs consisting of nasal formants and the antiformant act as place cues. Others have argued that transitions provide better cues. However, due to the limited number of languages being examined, only three places of articulation have commonly been investigated, i.e., bilabial, alveolar, and velar. The database compiled by Maddieson and Precoda (1984) reveals that the bilabial nasals are found most frequently in world languages, followed by velar, alveolar, and palatal, in descending order of frequency. Study of palatal nasals has been rare. Fortunately, nasal sounds in the Sgaw Karen of Ban Pa La–u in Hua Hin District are articulated in four places of articulation, i.e., bilabial, alveolar, palatal, and velar, thus offering an excellent opportunity to examine the rare nasal /ɲ/.

According to Matisoff (2008), the Sgaw Karen language belongs to the Karenic branch of the Tibeto–Burman language family. In Thailand, Sgaw Karen people are mostly found in the northern and western provinces. In this study, the Sgaw Karen dialect spoken at Ban Pa La–u, located in Hua Hin District, Prachuap Khiri Khan Province, has been examined. The population of Ban Pa La–u numbers one thousand one hundred and is made up of Sgaw, Pwo, and Thais. Most of the Karen people are Christian, although

some are Buddhist. Their birth places vary; some were born in the village, some in Myanmar and some around the Myanmar–Thailand border. My Sgaw Karen informants speak both Thai and Sgaw. The sound inventory of the Ban Pa La–u Sgaw Karen is as follows.

Table 1 Sgaw Karen Consonant System

Place Manner	bilabial	alveolar	palatal	velar	glottal
stop	ph p b	th t d	ch c	kh k	
nasal	m	n		ŋ	
fricative		s		x	h
trill		r			
approximant	w	l	j		

Only /h/ can occur in final position.

Table 2 Sgaw Karen Vowel System

	front	central	back
high	i		u
mid	e	ə	o
low		a	

Two diphthongs, /ai/ and /au/, are found.

There are four tones in Sgaw Karen:

Tone 1 is a mid tone with the phonetic realization [33] in non-checked syllables.

Tone 2 is a low tone with the phonetic realization [21] or [21] in non-checked syllables.

Tone 3 is a high tone with the phonetic realization [44] in non-checked syllables.

Tone 4 is a falling tone with the phonetic realization [452] in non-checked syllables.

To find nasal place cues, both nasal murmurs and formant transitions have been widely examined. Some studies have examined data from natural speech, while others have conducted experiments using synthetic speech. The results of these previous studies suggest that formant transitions provide better cues for differentiating place of articulation; however, nasal murmurs also play a place-contrastive role. The relational patterns among the nasal formants and antiformants of nasal murmurs and the formant frequencies of formant transitions reported in these various studies have been fairly consistent.

1.1 Nasal murmurs

Nasal murmurs have been found to be potential cues in differentiating the places of articulation in some studies (Malécot, 1956; Recasens, 1983). Nasal murmurs occur during the closure phase of nasal stop production and consist of nasal formants (NF) and an antiformant (NZ). Nasal formants arise from resonance in the nasal cavity, which functions as the main resonator. Ohala (1975) has argued that nasal formants tend to be stable across different nasal places of articulation due to the fixed size and volume of the nasal cavity. Ohala (1975), Recasens (1983), and Harding and Meyer (2003) state that the first nasal formant (NF1) occurs at about 200–300 Hz and has more energy than other nasal formants, which occur at higher frequencies. Furthermore, the disappearance of energy at certain frequencies is the result of energy absorption in the oral cavity. According to Ohala (1975), the antiformant frequency is inversely proportional to the length of the oral cavity. A longer oral cavity produces a lower antiformant frequency. Comparisons of the nasal murmurs for different places of articulation have shown that the highest to lowest frequency values for NF1 run from /ŋ/ through / / and /n/ to /m/. As for the antiformant, /ŋ/ likewise has the highest frequency value, followed by / /, /n/, and /m/. Moreover, the antiformant lies close to a particular nasal formant at each place of articulation, i.e., NZ is close to NF2 of /m/, NF3 of /n/, NF4 of / / and NF4 or higher of /ŋ/. Furthermore, perceptual studies such as Malécot (1956); Dukiewicz (1967), House (1957), Nakata (1959), Henderson (1978) (as cited in Recasens 1983: 1346) have found that the murmurs of /m/ and /n/ were quite effective in allowing identification of place of articulation, with /m/ receiving the highest correct score due to its having the lowest nasal formant and antiformant of all the nasal stops, and Ohala (1975) and Recasens (1983) have claimed that /ŋ/ is distinguishable from /n/ and / / due to its higher NF1 value and the lack of an antiformant in the middle of the nasal spectrum.

1.2 Formant transitions

Formant transitions have been proven to be effective cues in distinguishing place of articulation for nasals, especially with respect to the second formant (F2). Formant

transitions start from the release of the consonant and move toward the more-or-less steady state of the vowel. F2 has been the focus of place cue studies since its transitional direction and frequency value at the nasal-vowel juncture has proven to be a cue in many acoustic-perceptual studies. Studies on the contrastive role of transitions have shown that transitions following nasal initials differ by places: a rising transition follows /m/; a flat or falling transition depending on vowel type follows /n/ (Liberman et al., 1954; Recasens, 1983) a falling (Liberman et al., 1954) or a slightly rising, falling, or flat transition follows /ŋ/ and a falling direction after / / (Recasens, 1983). Additionally, some studies have found that the first formant (F1) and third formant (F3) aid in differentiating place of articulation (Recasens, 1983; Narayan, 2008).

Although F1 is not usually examined because all final nasals show the same falling transition, various studies (Recasens, 1983) have found that the degree of fall differs among the various place of articulation. F1 transitions fall the farthest for / / and the least for /ŋ/, with /m/ and /n/ lying in between. This means that /ŋ/ has the highest F1 value at the nasal-vowel juncture, followed by /m/-/n/ and / /, in that order.

Narayan (2008) has found that the F3 value at the nasal-vowel juncture helps to distinguish between /n/ and /ŋ/ which have similar F2s. Furthermore, Recasens (1983) has found that F3 falls between /n/ and a vowel, while it rises after /ŋ/. However, results for F3 transitions vary from study to study. Some have found that the F3 transition falls for /m/, /n/ and /ŋ/ after vowels and rises for / / (Magdics, 1969; Vaggies, Ferrero, Caldognetto-Magno, and Lavagnoli, 1978; Dukiewicz, 1967; and Fant, 1960, all cited in Recasens, 1983: 1347), but Recasens (1983) has found that the F3 transition after vowels falls for /m/ and /ŋ/ but rises for /n/ and / /.

This study aims to find effective cues for differentiating place of articulation of four initial nasals: /m/, /n/, / /, and /ŋ/, in the Sgaw Karen dialect of Ban Pa La-u, with the primary focus being the palatal nasal / /, which has been less studied acoustically. The acoustic parameters that were taken into account were the intensity, duration, and formant frequency of the nasal murmurs and the formant frequency of the formant transitions of /a/ and / /. Although, a review of the relevant literature shows intensity and duration to be considered generally poor place cues, they were examined in this study to verify the previous claims. The hypotheses of this study are (1) that formant transitions, especially F2, provide better cues for distinguishing place of articulation for nasal stops and (2) that of the four nasals, /n/ and /ŋ/ share the most similar acoustic characteristics.

2. Methods

2.1 Participants

Seven Sgaw Karen females aged between 19 and 43 were recorded. All speakers had Sgaw parents; however, their birth places varied. Some had been born on the Thai–Myanmar border, some had been born in Prachuap Khiri Khan Province, and one had been born in Myanmar. Females were chosen to avoid variation between genders; however, the participants' ages varied because of the difficulty in finding participants and the limited working time. Given the eight-day time limit, it was hard to find ideal participants since most Sgaw Karen people had to go to work. Therefore, in order to complete the field record on time, I controlled age range as much as possible, and most participants were between 30 and 43 years old, with one speaker of 19 and another of 24. Despite the wide age range, analysis showed that age had not affected the acoustic results.

2.2 Corpus and setup

The corpus consisted of two word lists, /a/ and /ɔ/. These were chosen because /a/ and /ɔ/ were the only two vowels that co-occurred with all 4 initial nasal stops. In both lists, the four nasals /m/, /n/, /ɲ/,² and /ŋ/ were in syllable-initial position. Here are the test words with their meanings: in /a/ context, *ma1* 'do' or *ma3* 'wife', *na1* 'you' or *na3* 'ghost', *me2 a1* 'front', *ŋa1/do2ŋa1* 'hire'; in the /ɔ/ context, *m 3* 'bite (classifier)', *n 3* 'older sister', *ɔ 3* 'easy' and *ŋ 3ŋa 2* 'dumb'.³ In the /a/ context, *ma3* and *na3* were substituted for *ma1* and *na1* in case the recorded sound of words with a mid level tone was not of good quality⁴. Although three test words were disyllabic, the syllable under examination in all these words, except *ŋ 3ŋa 2*, was stressed, which was assumed to have characteristics similar to those of monosyllabic words. Word lists were recorded in mono through a Sony ECM-719 microphone via a Creative USB Sound Blaster Play into a Lenovo notebook at a sampling rate of 16 kHz using Praat version 5.2.2.6. The Praat program was suitable for recording in this study due to the limited time

² /ɲ/ indicates a palatal nasal with an offglide.

³ Tone numbers 1–4 used here match those given earlier in this paper.

⁴ This happened with only one informant. It was not clear why she uttered words containing /a/ and a mid tone with more breath. Therefore, words were recorded that had the same structure except for the tone, resulting in test tokens suitable for further acoustic analysis.

and the surrounding environment. The data were collected over eight days of fieldwork, and the short time and noisy environment, where uncontrolled sounds from rain, animals, passing vehicles, etc. could not be avoided, were not ideal for recording. Therefore, sounds recorded with Praat had to be checked immediately to see if they were clear enough for further measurement and analysis. The speakers were asked to hold the microphone about 5" from their lips, which pretesting revealed to be the optimal distance, and to pronounce each word in the list at least 5 times. The total number of recorded test words was 280 (8 words $\times$ 5 times $\times$ 7 participants).

2.3 Acoustic analysis

Two characteristic parts of the acoustic signal were examined: nasal murmurs and formant transitions of the following vowels. Segmentation was based on waveform and wide-band spectrogram. Nasal murmurs were measured from the first periodic pulse to the beginning of the subsequent vowel which was signalled by a high-amplitude periodic pulse. Formant transitions were measured from the nasal murmur offset to the beginning of the vocalic steady state, signalled by the end of F2 distortion. See Figure 1 for an example of segmentation.

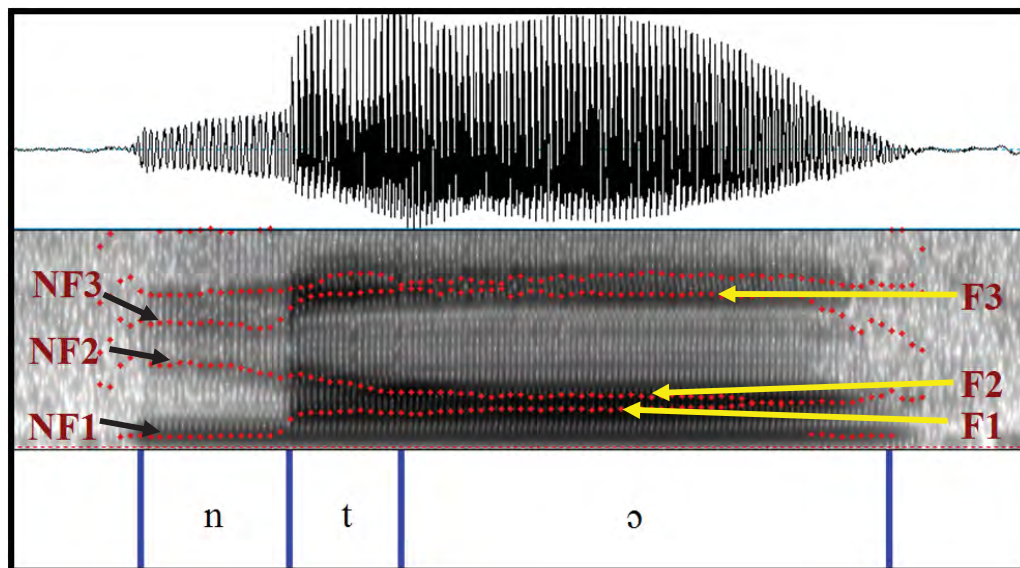


Figure 1 Segmentation of nasal murmurs /n/, transition signalled with 't' and /ʊ/ vowel

In the nasal murmur phase measurement were taken of acoustic characteristics including intensity, duration, and nasal formant frequency. Murmur intensity was measured at three positions: 0%, 50%, and 100%. These positions were chosen because the transitional directions in all nasal contexts were the same; therefore, it was not necessary to measure at additional positions. Murmur duration, corresponding to the period from the point where nasal murmurs began with the first periodic pulse to the beginning of the vowel, was measured. Only nasal formants in the / / word list were examined due to sound quality. The first, second, and third nasal formants (NF1, NF2, NF3) were measured at 25%, 50%, and 75% of the total nasal murmur duration to ensure that the frequency values truly belonged to the nasal formants. In the formant transition phase, the first, second, and third formants (F1, F2, F3) were each measured at five positions: 0%, 25%, 50%, 75%, and 100% of the total formant transition duration. These positions were chosen to enable transitional direction graph plotting.⁵ The frequencies of the nasal murmur formants and of the formant transitions were obtained automatically using Praat's formant tracker (LPC analysis).

2.4 Statistics

The significance of each acoustic characteristic as a cue in differentiating place of articulation for nasal stops was statistically tested with ANOVA (analysis of variance) and Tukey's HSD, a multiple comparison test. ANOVA provides the means to test whether several groups differ significantly. It is suitable only for comparing more than 2 groups. The T-test was not chosen because it is suitable for comparing 2 groups, resulting in a higher chance of committing an error. However, ANOVA only tells us whether there are statistically significant differences among groups. It cannot identify which pair is significantly different. Therefore, Tukey's HSD is required to find out which pairs differs significantly.

3. Results

3.1 Intensity

Intensity was not a cue for differentiating place of articulation since it showed no regular pattern either within a vowel or between two vowels. Below is the intensity in dB

⁵ Nasal formants can be studied by measuring spectra based on observation of nasal formant frequencies or relative spectral energy changes in low- and high-frequency ranges at the nasal release; however, due to time consumption in obtaining the values, spectra measurements were not attempted in this study.

measured at three different positions during nasal murmurs. See the average and SD values of intensity (dB) measured from 0%, 50% and 100% of nasal murmurs in Table 3.

Table 3 Mean ($\bar{x}$) and Standard Deviation (SD) Values for Intensity (dB) during Murmurs

vocalic context	position	value	place of articulation			
			m	n		ŋ
a	0%	x	53.07	51.07	50.43	49.05
		SD	6.69	5.32	4.70	7.34
	50%	x	58.78	56.96	53.41	54.32
		SD	6.20	5.71	5.26	6.94
	100%	x	62.20	61.42	58.90	54.84
		SD	8.83	6.19	6.16	8.09
	0%	x	47.69	48.10	47.13	46.53
		SD	4.88	3.81	3.16	3.64
	50%	x	54.19	54.42	54.59	52.97
		SD	4.54	3.50	3.25	3.13
	100%	x	59.84	59.27	58.39	57.93
		SD	7.20	4.75	4.49	4.08

The measurements were submitted to ANOVA and Tukey's HSD. ANOVA revealed no statistically significant difference among nasal places of articulation in the / / context, but there was a statistically significant difference among nasal places at the 50% and 100% position in the /a/ context with $p < 0.01$. The results of Tukey's HSD revealed that the intensity values of /m/ were significantly greater than those of / / with $p < 0.05$ at the 50% and 100% position. The intensity value of /m/ was significantly greater than that of /ŋ/ with $p < 0.05$ at 50%, and the intensity value of /n/ was significantly greater than that of / / with $p < 0.01$ at 100%.

Although some statistically significant differences were found in the /a/ context, they were not consistent across every position in both vocalic contexts. Moreover, when compared with studies examining the acoustic characteristics of stops and nasals with different places of articulation (Trongdee, 1987; Tarnsakun, 1988), the intensity results

showed no regular relational pattern of intensity in different places. Therefore, intensity could not distinguish nasal places of articulation. This result accords with intensity not usually being examined as a potential place cue. This may be due to the great variation in speakers' speech volume, which is very hard to control.

3.2 Duration

The results show that duration was not a place cue. There was no regular pattern of real-time duration across four nasals in two different vocalic contexts. Durations in milliseconds (msec) of nasal murmurs for nasal stops at four different of places of articulation are shown in Table 4.

Table 4 Duration of Murmurs in Milliseconds

vocalic context	value	place of articulation			
		m	n		ŋ
a	x	121	119	175	132
	SD	67	60	108	38
	x	117	124	140	121
	SD	37	48	44	35

In both vocalic contexts, / / had the highest duration value. In the /a/ context, the / / duration value was followed by those for /ŋ/, /m/, and /n/, in descending order. In the / / context, on the other hand, the duration of / / was followed by /n/, /ŋ/, and /m/, in that order. ANOVA revealed no statistically significant difference in the / / context; however, the /a/ context produced a statistically significant difference among nasal places of articulation with $p < 0.05$. Tukey's HSD then showed that, in the /a/ context, the duration value for / / was significantly higher than those for /m/ and /n/, with $p < 0.05$. These inconsistent results of relational patterns and statistical differences suggest that duration does not distinguish place of articulation for nasal stops. Furthermore, comparison with other works (Trongdee, 1987; Tarnsakun, 1988; Narayan, 2008) did not show a consistent pattern. This confirms the fact that duration is not usually investigated as a potential cue to place differentiation. The variation found across vowels may have been affected by speech rate. The difference in speech rate may lengthen or shorten the duration of nasals in a non-systematic way.

3.3 Nasal formants

Due to sound quality, only nasal formants during nasal murmurs from the / / word list were examined. However, the NF3 of nasals preceding / / could possibly be noise formants. Of the three nasal formants, only NF2 produced a consistent relational pattern: at all three measurement positions /n/ > / / > /m/ > /ŋ/. The mean values for NF2 at different points in the nasal murmur are shown in Table 5.

Table 5 Mean and Standard Deviation Values for NF2 during Murmurs

vocalic context	position	value	place of articulation			
			m	n		ŋ
	25%	x	1157.85	1376.11	1338.03	1101.53
		SD	227.77	422.39	280.42	146.40
	50%	x	1146.38	1324.61	1303.28	1089.84
		SD	181.57	376.50	288.17	190.72
	75%	x	1108.95	1345.61	1300.70	1046.07
		SD	148.18	329.01	303.41	142.13

ANOVA showed some statistically significant difference among nasal places, with $p < 0.001$ for NF1 and $p < 0.01$ for NF2, but there was no statistically significant difference among nasal places for NF3. Furthermore, Tukey's HSD revealed that only the NF2 values for /n/ and / / were significantly higher than those of /ŋ/, with $p < 0.05$ for every measurement position. At 75%, the NF2 value for /n/ was significantly higher than that of /m/, with $p < 0.01$. As for the other two nasal formants, NF1 and NF3 did not produce regular relational patterns across three different positions. However, Tukey's HSD revealed some patterns for NF1. The NF1 values for /ŋ/ were significantly greater than those for /m/ and /n/, with $p < 0.05$ for every measurement position. Moreover, at 50% and 75%, the NF1 values for /ŋ/ were significantly higher than those for / /, with $p < 0.05$. The mean values for NF1 at different points are shown in Table 6.

Table 6 Mean and Standard Deviation Values for NF1 during Murmurs

vocalic context	position	value	place of articulation			
			m	n		ŋ
	25%	x	274.61	278.98	306.37	352.46
		SD	78.66	57.86	55.26	93.25
	50%	x	292.52	285.07	302.89	363.37
		SD	108.79	57.76	54.09	83.11
	75%	x	317.78	293.41	297.24	386.95
		SD	121.21	59.14	53.03	90.05

3.4 Formant transitions

Measurements at 0% and 25% were subjected to ANOVA and Tukey's HSD since the values at both positions were the two closest to the initial nasals and, therefore, had the better potential for displaying the different acoustic characteristics of the four nasals. Consistent relational patterns were found for F2 across nasal places. The F2 values for / / were the highest, followed by /n/, /ŋ/, and /m/. Mean values for F2 in all four places of articulation are shown in Table 7.

Table 7 Mean and Standard Deviation Values for F2 (Hz) during Formant Transitions

vocalic context	position	value	place of articulation			
			m	n		ŋ
a	0%	x	1533.37	1987.82	2620.41	1986.56
		SD	285.93	169.46	90.55	292.30
	25%	x	1590.03	1977.05	2514.19	1975.71
		SD	254.77	180.35	85.27	215.93
	0%	x	1006.97	1608.91	2697.74	1134.62
		SD	152.14	165.58	111.03	176.50
	25%	x	1032.57	1532.44	2464.24	1150.81
		SD	146.12	165.51	150.66	159.86

ANOVA revealed that all four places differed significantly in both vocalic contexts with $p < 0.001$. However, statistical results for / / and /a/ contexts differ according to Tukey's HSD. In the / / context, almost every nasal pair differed significantly, with $p < 0.001$, except for the /ŋ/-/m/ pair, where /ŋ/ was significantly greater than /m/, with $p < 0.05$. The statistically significant difference among nasal places in the / / context reflects the potential for F2 being a place cue. However, while almost every nasal pair differed significantly in the /a/ context, with $p < 0.001$, /n/ and /ŋ/ did not differ significantly, reflecting their similarity. Scatter plots in Figure 2 and Figure 3 show F2 and F3 at the 0% position in both vocalic contexts.

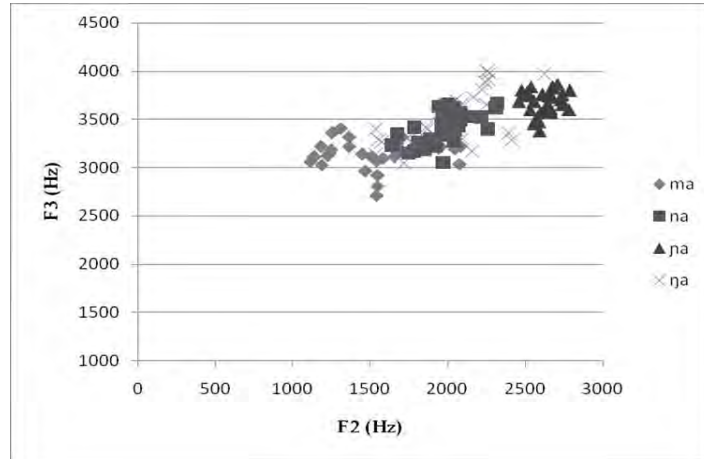


Figure 2 F2 and F3 at 0% of formant transitions in the /a/ context

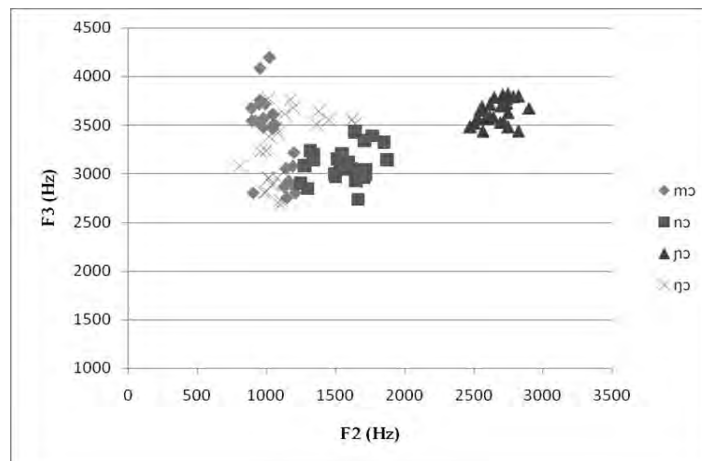


Figure 3 F2 and F3 at 0% of formant transitions in the / / context

In Figure 2, the F2 of /n/ and /ŋ/ overlap greatly. In Figure 3, there was a clearer grouping of F2s for each place of articulation; however, although /ŋ/ somewhat overlapped with /m/, this did not have a statistical significance.

As for F1, consistent relational patterns were found at 0% for both vowels and at 25% in the /a/ context. The F1 values for /m/ were higher than those for /n/, /ŋ/, and / /, in descending order. At 25% in the / / context, /n/ values were higher than those for /m/, /ŋ/, and / /, once again in descending order. However, this did not necessarily lead to

significant differences since /m/ and /n/ did not differ significantly. ANOVA revealed a statistically significant difference among nasal places in both vocalic contexts, with $p < 0.01$. Furthermore, Tukey's HSD revealed that the F1 values for /m/ and /n/ were significantly higher than those for /ŋ/ at both positions in both vocalic contexts, with $p < 0.05$. Moreover, the F1 values for /ŋ/ were significantly higher than those for / / in the /a/ context, with $p < 0.01$. At 25% in the / / context, the F1 values for /n/ were significantly higher than those for /ŋ/, with $p < 0.001$. However, the fact that the F1 values for /m/, /n/, and /ŋ/ did not differ significantly, except for one significant difference between /n/ and /ŋ/ at 25% in the / / context, shows that F1 cannot be a cue for distinguishing places of articulation.

Furthermore, the F3 values did not show regular relational patterns across nasal places in either vowel context. ANOVA indicated a statistically significant difference among nasal places in the /a/ context and at 0% in the / / context, with $p < 0.001$. Also, Tukey's HSD revealed that the F3 values of / / were higher than those of /ŋ/, /n/, and /m/ at 0% in the /a/ context, where / / was significantly higher than the others, with $p < 0.001$, and /ŋ/ and /n/ were significantly higher than /m/, with $p < 0.001$. At 25%, the / / values were higher than those for /n/, /ŋ/, and /m/; every sound was significantly higher than /m/, with $p < 0.01$, and / / was significantly higher than /ŋ/, with $p < 0.05$. The relational patterns at both positions in the / / context did not show the same pattern within vowels or between two vowels. Tukey's HSD revealed that at 0%, the / / value was significantly higher than those /m/, /ŋ/, and /n/, with $p < 0.001$, and that /m/ was significantly higher than /n/, with $p < 0.05$. At 25%, the /m/ value was higher than those for /ŋ/, /n/, and / /, but the differences were not statistically significant. The inconsistent relational patterns in both positions, with /m/ having a high F3 value, differing from the patterns in the /a/ context, resulted from high F3 values for /m/ obtained from two informants. If those values were discounted, the relational patterns at each position would be similar to those found in the /a/ context. Additionally, Figures 2 and 3 above show that the F3 values at 0% of the formant transitions in both vocalic contexts are very close. The irregular patterns of the F3 frequency do not make F3 a good place cue.

Apart from formant frequency comparisons, F2 and F3 transitional directions were compared across different nasals in both vocalic contexts. The results show that the relation between F2 and F3 transitional directions in the / / context helped in distinguishing place of articulation better than in the /a/ context, where F2 and F3 had the same falling transitional directions for /n/ and /ŋ/. Graphs showing F2 and F3 transitional directions for each vowel can be found in Table 8.

Table 8 Mean Values of F2 and F3 during Formant Transitions and F2 and F3 Transitional Directions Following Four Initial Nasals in the /a/ and / / Contexts

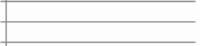
place of articulation	vocalic context	formant frequency (Hz)	position of formant transition					Graph
			0%	25%	50%	75%	100%	
m	a	F3	3136.61	3223.08	3256.90	3280.74	3282.79	
		F2	1533.37	1590.03	1638.7	1673.85	1720.81	
		F3	3319.62	3402.49	3456.24	3475.71	3507.56	
		F2	1006.97	1032.57	1052.77	1076.73	1108.10	
n	a	F3	3407.83	3400.96	3378.32	3374.83	3352.67	
		F2	1987.82	1977.05	1936.4	1892.36	1841.99	
		F3	3121.43	3326.46	3399.61	3468.77	3474.17	
		F2	1608.90	1532.44	1419.62	1332.78	1226.54	
	a	F3	3669.07	3498.27	3367.47	3317.05	3312.09	
		F2	2620.41	2514.19	2326.26	2133.68	1978.39	
		F3	3651.34	3312.07	3279.88	3321.12	3365.44	
		F2	2697.74	2464.24	1963.57	1495.95	1289.14	
ŋ	a	F3	3465.31	3386.07	3336.20	3308.43	3291.56	
		F2	1986.07	1975.36	1919.81	1879.68	1831.17	
		F3	3385.39	3388.64	3380.63	3345.89	3385.39	
		F2	1150.80	1161.45	1161.33	1136.37	1150.80	

Table 8 shows outstanding patterns for /m/ and / /, namely, F2 and F3 transitions rise for /m/ and fall significantly for / /. The F2 and F3 transitional directions for /n/ and /ŋ/ are very similar in the /a/ context; however, in the / / context, F2 falls in formant transition after /n/, while it rises only a little after /ŋ/, making it look rather flat. F3 rises after /n/, while it rises only a little after /ŋ/, once again making it look rather flat. From these results, it can be concluded that F2 and F3 transitional directions depend on vocalic context.

4. Discussion

The results appear to confirm previous studies' claims that formant transitions are better cues (Malécot, 1956; Delattre, Liberman and Cooper, 1955). Comparisons with other studies reveal that F2 and F3 formant transitional directions after different nasal places were found to have a regular pattern in most studies (Liberman et al., 1954; Delattre et al., 1955; Recasens, 1983). On the other hand, formant patterns within the nasal murmur were inconsistent both within this study at different positions and between this study and others (Recasens, 1983; Trongdee, 1987). In addition, even though the relational patterns for NF2 were regular in all three measurement positions and the NF2s for /n/ and / / were significantly higher than those for /ŋ/, the lack of statistically significant difference among /m/, /n/, and / / reflects the similarity of NF2 for these nasals. Therefore, NF2 is not considered a good cue to differentiate the places of articulation. This might be explained by the anatomical characteristics of the nasal cavity. According to Ohala (1975), the nasal formants correspond to resonance in the nasal cavity, which is fixed in size and volume; hence, the frequencies resonating in the nasal cavity tend to be very close regardless of place of articulation. Therefore, nasal murmurs probably do not provide place cues. Apart from the results themselves, another difficulty in the attempt to use nasal formants as place cues is that it is complicated to measure nasal formants in natural speech. The nasal formant frequencies appearing on the spectrum were neither clear nor consistent throughout the nasal murmur phase, and it was hard to separate noise frequencies from those of nasal formants.

F2 being an effective cue is consistent with other works (Delattre et al., 1955; Recasens, 1983; Harding and Meyer, 2003). Acoustic-perceptual studies have found the presence of certain F2 transitional directions at each place of articulation to be essential for correct place identification. The transitional direction rises after /m/, falls after /n/ and / /, and either falls or lies rather flat after /ŋ/, depending on which vowel follows. The F2 transitional directions of labial, alveolar, and velar nasal consonants arrived at in my research agree with previous studies on transitional directions of corresponding stops

(Delattre et al., 1955; Pickett, 1980). The frequency values for F2 at the nasal–vowel juncture in the /a/ context in the present study were similar to the results achieved by Recasens (1983), in which / / had the highest value, /m/ had the lowest value, and the values for /n/ and /ŋ/ fell in between. The difference is that, in Recasens’s study, /n/ and /ŋ/ were differentiated because the F3 of /ŋ/ rose before vowels. In contrast, I found the F3 of /ŋ/ to fall in a manner similar to /n/. However, the F2 values for /n/ and /ŋ/ in the / / context differed significantly. Both the F2 values and transitional directions, then, aid in contrasting places of articulation. F3 showed more than one pattern. This could be seen in the case of the F3 transitional directions of /ŋ/ which differed in the /a/ and / / contexts. Moreover, the F3 values of /n/ and /ŋ/ which helped distinguish both nasal places in Narayan (2008) did not differ significantly in the present study. Therefore, it could not help differentiate alveolar and velar places of articulation.

F1 is not an effective cue because the patterns are not consistent across studies. A comparison of F1 relational patterns between the present study and the literature review in Recasens’s study (1983) shows different patterns. In Recasens’s paper, /ŋ/ had the highest F1 value, followed by /m/ or /n/ and / /. In the present study, in contrast, /m/ had the highest F1 value, followed by /n/, /ŋ/, and / / in descending order. These inconsistent patterns may reflect language-specific traits or the ineffectiveness of F1 as a place cue. However, the lowest F1 value of / / can be explained by the oral constriction/F1 rule which says that “the frequency of F1 is lowered by any constriction in the front half of the oral part of the vocal tract, and the greater the constriction, the more the F1 is lowered.” (Pickett, 1980). In producing a palatal nasal, / / the tongue moves towards the hard palate, forming a constriction at the palate; hence, the F1 value decreases. Furthermore, the rising of the F1 during formant transitions after nasals likewise accords with the oral constriction/F1 rule, which also applies to constriction at the lips or teeth. In producing nasal consonants, oral constrictions occur in the front half of the mouth or at the lips, whereas there is no constriction in the oral cavity when producing vowels. Therefore, F1 is lower in nasals than in vowels, and the transitional direction is upward.

The transitional directions of F2 can be explained by the relationship between F2 and vocal–tract length. A sound produced in a longer vocal tract has a lower formant frequency. In producing /n/, / /, and /ŋ/, there is an obstruction within the oral cavity, causing the resonating chamber to be shorter; thus, the F2 rises. On the other hand, during vowel production, there is no obstruction, so the vocal tract is longer. Therefore, the F2 of a vowel is lower than that of nasal consonants. Consequently, when a vowel follows a nasal stop, F2 falls during the transition phase. This is true in the /a/ context but not in the / / context. In the / / context, the F2 rises only a little after /ŋ/, making it look rather flat. This might be because in producing both / / and /ŋ/, the tongue moves toward the

velum, even touching it for /ŋ/, so their places of articulation are rather similar. This may result in only minor changes of the vocal tract, so the F2 values for / / and /ŋ/ are close, producing a flat transition. In the case of /m/, the resonating chambers of /m/ and vowels seem to have close to the same vocal-tract length, so the length of the vocal tract may not be the reason why the F2 of /m/ is lower than that of the following vowel. This may, instead, be explained by the lip closure which results in a lowering of formant frequencies (Ladefoged, 1993). Hence, the F2 of /m/ is lower than that of the following vowel. Consequently, the formant rises from low to high, during the nasal-to-vowel transition phase

To obtain more concrete results, future research should study both formant transitions in more vocalic contexts and stops at corresponding places of articulation. This would help confirm the general characteristics of each place of articulation found in nasal place cue studies. Additionally, formant transitions are noticeably easier to analyze due to the clarity of formant frequencies during formant transitions. In contrast, nasal murmurs are much less clear due to sound quality and the characteristics of the nasal murmurs themselves. Nasal formants are hard to determine since the formant frequencies shown on the spectrogram are not continuous and they can be altered by noise formants. Moreover, the location of antiformants cannot be specified due to sound quality. This situation can probably be overcome by using extra-high-quality recording equipment; however, recording during fieldwork makes it hard to avoid surrounding noise.

Acknowledgements

This paper is an output of the Linguistics Field Methods class. This research was funded by the Faculty of Arts, Chulalongkorn University and the Thailand Research Fund (TRF) through the Royal Golden Jubilee Ph.D. Program (Grant No. PHD/0002/2552) to Prof. Dr. Theraphan Luangthongkum and Karnthida Kerdpol for the thesis called “Nasality in Karenic languages: An acoustic study”. I am very grateful to Prof. Dr. Theraphan Luangthongkum for her supervision and comments. My special thanks go to the Sgaw Karen people of Ban Pa La-u for their kind cooperation. Finally, I would like to thank all my friends in the Linguistics Department for their help and support in different ways.