The subgrouping of Karen

Ken Manson

Payap University & SIL International

ken.grammar@gmail.com

Several classifications of Karen languages have been published (e.g. Jones 1961; Burling 1969; Kauffman 1993; Bradley 1997; Manson 2002; Shintani 2003). However they have all lacked in comprehensiveness and/or have not been based on the comparative method. 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.

Introduction

Karen languages are unusual among Tibeto-Burma as being a distinct branch with no members of uncertain status. Internal relationships and reconstruction has been limited to either a few well known groups (normally the three largest groups: Sgaw, Pwo, and Pa'O, e.g. Haudricourt 1946, 1953; Jones 1961; Burling 1969) or only one aspect is compared (Manson 2002; Shintani 2003). This presentation, building on the lexical and phonological comparison of Manson (2002), considers tonal development, reflexes of proto-initials and proto-rhymes to identify higher-order subgroupings.

Previous reconstructions of Karen include Haudricourt (1946, 1953), Jones (1961) and Burling (1969). The reconstruction by Haudricourt was based on the orthographies of two languages (Sgaw & Pwo). Jones & Burling's reconstructions are limited to four languages (Sgaw, Pwo, Pa'O & Palaychi) and are phonologically unmotivated with respect to tonal development.

Overview of Karen languages

Speakers of Karen languages are located primarily in eastern Burma from the southernmost tip of Burma to southern Shan State. There are also populations of Sgaw and Pwo in Thailand in the provinces along the western border with Burma. Bradley (1997:46) suggests a total population of 3.9 million, but notes that this is "substantially under enumerated". The total population of ethnic Karen is somewhere between 6 and 10 million; however, not all ethnic Karen still speak Karen languages. Many now speak only Burmese, especially those living on the plains.

The actual number of Karen languages is unknown as there has never been a comprehensive survey of Karen. It would appear from the literature that there are between 20-30 Karen languages. Eighteen languages whose existence has been reasonably documented are shown in Figure 1, where they are listed in relative geographical position to each other. Karen languages found in the mountains of eastern Burma usually have numerous dialects, some often difficult to understand to other speakers of the same ethnicity (see Bennett 1991 for Kayah and Manson & Chou 2008 for Kayan).

N. Pa'O

Pwo

Lahta

E. Kayah

Kayan

Yinbaw

W. Kayah

Geker

Kayaw   Yintale

Gekho   Bre

Manu

Geba

Pwo

Bwe   Palaychi

Sgaw

Paku   Dermuha

Sgaw

Pwo

S. Pa'O

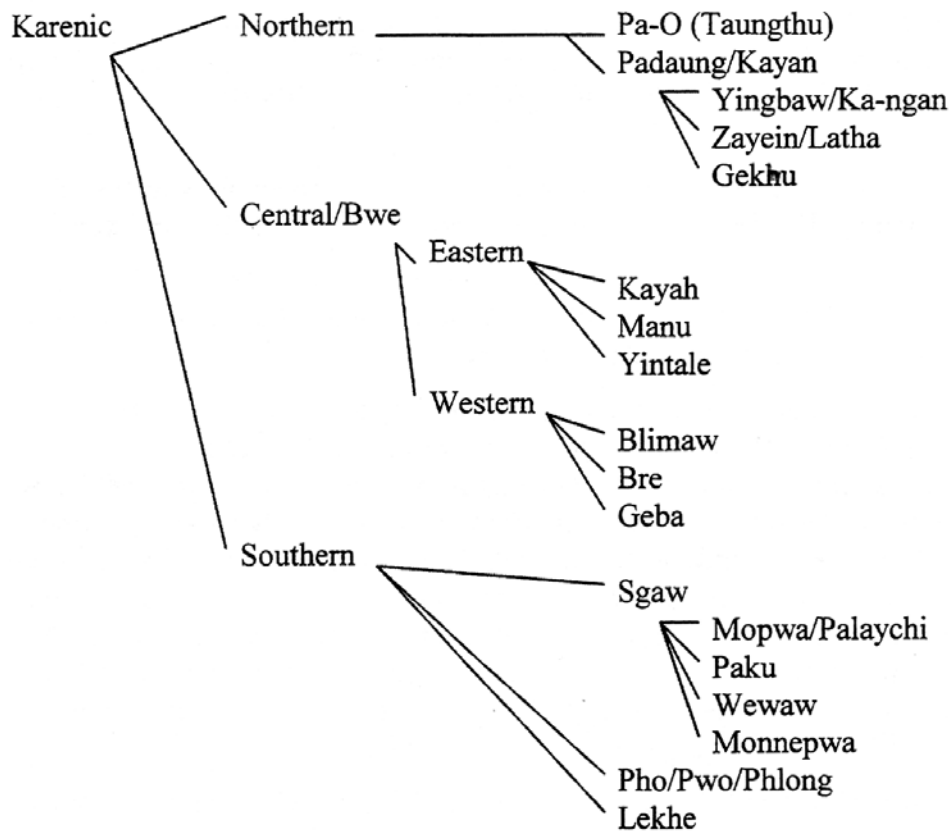Figure 1: Geographic distribution of Karen language clusters

Figure 2: Karen language relationships (Bradley 1997)

Tonal development

It is a well known fact that voiced initial consonants cause the syllable to be pronounced at a lower pitch than a voiceless initial. Most languages of Southeast Asia typically have a reduced first syllable, and where there have been phonological changes in the initial consonants, often pitch and/or phonation differences remain. From about the 10th-17thC there was a "Great Tone Split" that affected all the languages of the area (Weidert 1987). Karlgren (1915) was the first to deduce that the split in tones was conditioned by the initial consonant. Haudricourt (1961) (see Court (1972) for an English translation) expanded on Karlgren's work and showed that for all the language families of South-east Asia there was a split in the tones of the language. Usually there was a 2-way split, but in some languages there was a 3-way split, based on the manner of articulation of the syllable initial consonant.

The development of tone in Karen can be traced back to a three-way distinction in open syllables (*A, *B, *B´) plus closed syllables (*C). There were at least six steps in the development of modern-day tonal systems in Karen. No all Karen languages have undergone all the steps.

1. Initial proto-voiced consonants caused the following vowel to be pronounced with breathy phonation
2. This breathiness caused a lowering of the pitch (Thurgood 2002, 2006)
3. Proto-voiced stops became voiceless unaspirated
4. Proto-voiceless became voiced (or no change)
5. (Breathy phonation caused the initial consonant to be pronounced aspirated)
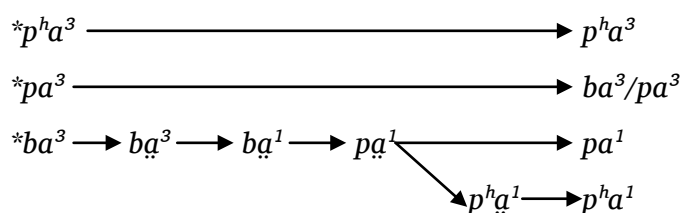6. Breathy phonation lost, leaving only a tonal difference

$$*p^ha^3 \longrightarrow p^ha^3$$
$$*pa^3 \longrightarrow ba^3/pa^3$$
$$*ba^3 \longrightarrow ba̤^3 \longrightarrow ba̤^1 \longrightarrow pa̤^1 \longrightarrow pa^1$$
$$\searrow p^ha̤^1 \longrightarrow p^ha^1$$

Figure 3: The development of tone and initial consonant

Karen Tone Box

Manson (2009) proposed a Tone Box for Karen languages in a similar vein to the Gedney Tone Box for Tai languages (Gedney 1972):

	A	B	B′	C
Proto-aspirated	**1 (III)** Water [*tʰi] Branch [*pʰaŋ] Flower [*pʰɔ] Chicken [*sʰan] Sleep [*m̥i] Die [*t̥ʰi]	**4 (VI)** Star [*sʰa] Leaf [*l̥a] Fingernail [*m̥i] Fire [*m̥e] Give [*p̥ʰe] Bitter [*kʰa]	**7 (Va)** Bone [*kʰri] Child [*pʰo] Right [*tʰwe] Spicy [*hɛ] Take [*pʰi] Pus [*pʰi/mi]	**10 (VIII)** Sky [*m̥oʔ] Iron [*t̥ʰaʔ] Pig [*tʰɔʔ] Skin/bark [*pʰeʔ] Shoot [v] [*kʰaʔ] Dark [*kʰeʔ/kʰuʔ]
Proto-voiceless	**2 (II)** Silver [*r̥ɔn] Ginger [*ʔeŋ] Rabbit [*tɛ] Navel [*te] Spear [*pan] White [*pwa]	**5 (VIa)** Egg [*ti] Cheek [*pu] Liver [*sin] Eat [*ʔam] Left [*se] Be at, exist [*ʔɔ]	**8 (V)** Paddy [*pɨ] Blow/howl [*ʔu] Head [*klo] Hand [*su] Breathe [*sa] Many [*ʔa]	**11 (VIIIa)** Alcohol [*siʔ] Wing [*teʔ] Heart [*saʔ] Call/shout [*kaʔ] Near [*pɔʔ]
Proto-voiced	**3 (I)** Nest [*bwe] Tongue [*ble] Person [*bra] Name [*min] Drunk [*mun] Red [*le]	**6 (IV)** Sun [*mɤ] Stone [*loŋ] Snake [*ru] Arrow [*bla] Old [humans] [*bra] Hot [*go]		**12 (VII)** Monkey [*zoʔ] Eye/face [*meʔ] Brain [*nɔʔ] Intestines [*breʔ] Rib [*rɤʔ] Deep [*jɔʔ]

Table 1: Karen Tone Box (Manson 2009)

The words in each cell have been retained in all/most modern-day Karen languages. The phonological elements in square brackets in Table 1 are my reconstructed form for the lexical item; however there will be some variation in manner of articulation depending on the language variety under investigation, yet the position of articulation should not change. Luce's (1985) tone patterns (Roman numerals in brackets) are included.

Tone patterns

Taking the Karen Tone Box and applying it to Karen language varieties shows a number of striking patterns (Table 2):

- the top two rows prototone *B′ merged with prototone *C
- all languages except Pwo had a 2-way split[1]
- breathy phonation is associated with protovoiced initial consonants
- closed syllables with non-voiced initials are either high (5) or mid (3) tone

---

[1] Palaychi also had a 3-way split, but has yet to be included in the analysis.

5	1	3	3
5	1	3	3
4̤	1̤		2̤

Kayaw

5	1	3	3
5	1	3	3
2̤	3̤		5̤

Kayan

5	3	3	3
5	3	3	3
3	3		1

Bwe-Geba

3	1	5	5
3	1	5	5
1̤	4̤		3̤

Kayah

4	3	5	5
4	3	5	5
1̤	1̤		3̤

Yintale

5	1	5	5
5	1	5	5
1̤	3̤		5̤

Yinbaw

5	2	5	5
5	2	5	5
3	2		3

Paku

1	5	5	5
1	5	5	5
1̤	3̤		4̤

Manu

5	2	5	3
5	2	5	3
3	1		1

Sgaw

2	4	2	3
2	4	2	3
1	5		1

Pa'O

2	5	5	3
1	5	5	3
1	3		4

Pwo

Table 2: Tone Boxes for Karen languages

Taking Kauffman's (1993:45-55) proposal for the development of tone in "Central" Karen as a starting point I proposed a possible scenario for the development of tone in Karen languages (Manson 2009). This proposal has been revised and expanded (see Supplement 1: The development of tone in Karen languages). The development in tone cannot be separated from the development of consonants. The four high-level clusters (Peripheral; Northern; Central; Southern) each share significant consonant developments that align with tonal development.

Proto-Karen initial consonants and clusters

Proto initial consonants are fairly well preserved throughout Karen languages. The main variation in reflexes is found in proto initial consonant clusters, especially for Cr clusters. Consider bilabial initials:

	*pʰ	*p	*b	*pʰl	*pl	*bl	*pʰr	*pr	*br
Manu	pʰ	b	p	pʰl	pl	pl	?	pʰ	pj
Kayaw	pʰ	b	p	pl	pl	pl	?	pr	pr
Yinbaw	pʰ	b	p	pʰl	bl	pl	pʰr	pr	pr
Pwo	pʰ	ɓ	pʰ	pʰl	bl	pʰl	pʰl	pj	pj
Pa'O	pʰ	b	pʰ	pʰr	pʰr	pʰr/pl	pʰr	pʰr	pʰr
Kayan	pʰ	b	p	pʰl	pl	pl	pʰr	pr	pr
Lahta	pʰ	b	p	pl	?	p(l)	?	pl	pl
Paku (L)	pʰ	ɓ	p	pʰl	ɓl	pl	pʰɣ	ɓ	bʀ
Sgaw	pʰ	b	p	pʰl	bl	pl	pʰɣ	bɣ	pɣ
Yintale	pʰ	b	p	pʰl	pl	pl	?	pw	pj
Palaychi	pʰ	b	p	pʰl	bl	pl	?	p	ɣ
Dermuha	pʰ	ɓ	p	pʰl	?	ɓl	pʰl	p	p
Geker	pʰ	ɓ	p	?	ɓl	pl	pʰr	pr	pr
Gekho	pʰ	b	b	?	bl	bl	pr	br	br
E Kayah	pʰ	b	p	pl	?	pl	pʰr	pʰr	pʰr
W Kayah	pʰ	b	p	pl	pl	pl	pr	pr	pr
Bwe	pʰ	ɓ	b	pʰl	pl	pl	pw	p	bj
Geba	pʰ	b	p	pʰl	pl	pl	pl	p	pl
Paku	pʰ	ɓ	p	pʰl	ɓl	pɣ	pʰɣ	p(ɣ)	bj

Table 3: Reflexes of proto-bilabial initial consonants

The pronunciation of consonant clusters in Karen languages ranges from co-articulation to the clear insertion of an epithetic schwa. For example the word for 'sea, river' is *pəle* in Pwo and *plḙ* in Kayan.

Proto-rhymes

Vowels show the greatest variation in Karen languages and even within a language there can be divergent dialects which look more like another Karen language. In many Karen languages, especially from the southern region, the simple vowel rhymes have not changed.

Nasal final rhymes are key to identifying historical developments within Karen. For example, Sgaw, Palaychi and Paku have merged **am*, **an*[2], and **aŋ* to ɔ. Geker and Gekho show this change only for **aŋ*, Geker and Gekho **am* is denasalised and fronted, a pattern seen also in Kayah, Bwe and Geba. **aŋ* in Manu and Kayaw has merged with **aʔ* and then denasalised and backed to ɔ.

	*a	*aʔ	*aŋ	*am
Manu	a	ɔ	ɔ	a
Kayaw	a	ɔ	ɔ	a
Pwo	a	aʔ	ã	ã
Pa'O	a	aʔ	aŋ	am
Kayan	a	a	aɴ	aɴ
Yintale	a	a	aɴ	aɴ
Lahta	a	a	aɴ/a	aɴ/e
Yinbaw	a	a	ã	ã
Paku (L)	a	a	ɔ	ɔ
Sgaw	a	a	ɔ	ɔ
Palaychi	a	a	ɔ	ɔ
Dermuha	a	a	o	o
Geker	a	a	ɔ	ɛ
Gekho	a	a	ɔ	ei
E Kayah	e	e	ɛ	e
W Kayah	ɛ	ɛ	ja	e
Bwe	ɛ	a	a	ɛ
Geba	ɛ	a	a	ɛ
Paku	i/e	a	a	a

Table 4: Reflexes of selected proto-rhymes

Proto-Karen had lost its distinction between syllable final stops */p, t, k/* at the time Proto-Karen speakers came into contact with Mon-Khmer. Only Pa'O has syllable final consonants, and they only occur in apparent loans from Mon-Khmer, see Luce (1985 Charts E-J, especially Chart E 'Austric connections'). Further Southern

---

[2] *an is not a strongly supoorted reconstruction for Proto-Karen.

Pa'O and Northern Pa'O often show variable final consonants: 'spicy' S.Pa'O hap^5, N.Pa'O hat^5; 'weather' S.Pa'O k^han^5, N.Pa'O k^ham^5; 'slow' S.Pa'O $ju\textipa{P}^5$, N.Pa'O ʑik^5.

Synthesis

Creating a classification of Karen based on shared innovations, rather than shared retentions has been done emphasising rhyme development over consonant changes or tonal splits.
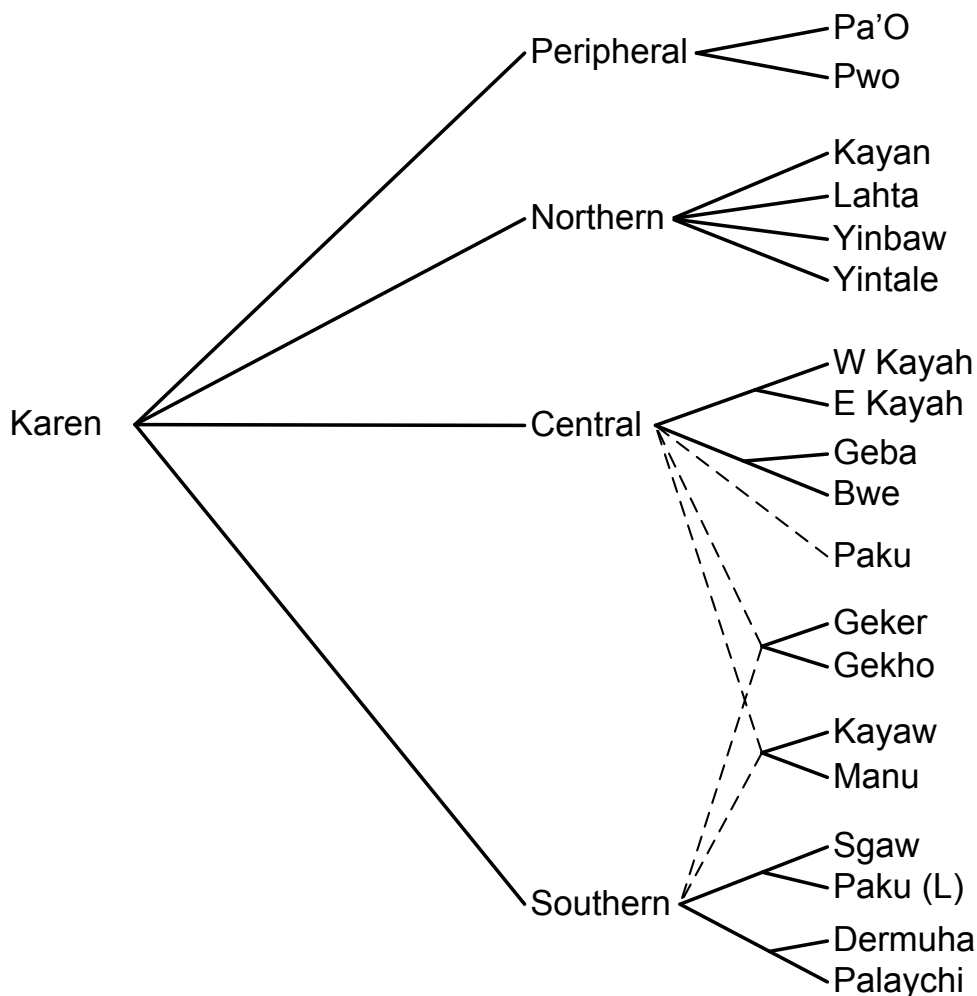

Figure 4: The Classification of Karen languages

Each branch has been labelled based on geographic location. The Peripheral branch (Pa'O and Pwo) are found to the north, east and south of the Karen languages. This branch is identified by proto voiced stop initials appearing as aspirated stops (e.g. *$p > p^h$). The Northern branch (Kayan, Lahta, Yinbaw and Yintale) are identified by the merging of nasal finals (e.g. *am, *an > aɴ) and the merging of stop final rhymes with the open rhyme equivalent (e.g. *aʔ, *a > a). The Central Karen languages (Kayah, Bwe, Geba) are identified by vowel raising (e.g. *a > ɛ). In the Southern branch (Sgaw, Luce's Paku, Palaychi, Dermuha) nasal final rhymes have merged and the rhyme has then been raised (e.g. *am, *aŋ > ɔ).

The final two clusters show developments found in both the Southern and Central branches and so are tentatively associated with both until more thorough research can be done. Geker and Gekho both show *aŋ > ɔ and *am > ɛ. Kayaw and Manu form a cluster based on the merging of velar consonant final rhymes (e.g. *aʔ, *aŋ > ɔ) and bilabial consonant final rhymes merging with the simple rhyme (e.g. *am > a).

Bibliography

Bennett, J. Fraser. 1991. Two more Kayah Li dialects: A comparison of *dɔ tà má* and *dɔ shò pía* dialects with Western and Eastern Kayah. presented at 24th ICSTLL, Bangkok

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 1

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: Mahidol University. [Translation and expansion of Haudricourt 1961]

Gedney, William J. 1972. A checklist for determining tones in Tai dialects. In *Studies in linguistics in honor of George L. Trager*. ed. M. Estellie Smith, 423-437. The Hague: Mouton

Haudricourt, André-Georges. 1946. Restitution du Karen commun. *Bulletin de la Societe de Linguistique de Paris* 42:103-111.

Haudricourt, André-Georges. 1953. A propos de la restitution du Karen commun. *Bulletin de la Societe de Linguistique de Paris* 49:129-132.

Haudricourt, André-Georges. 1961. Bipartition et tripartition des systemes de tons dans quelques langues d'extreme-orient. *Bulletin de la Societe de Linguistique de Paris* 56.1:163-180

Haudricourt, André-Georges. 1975. Le systeme de tons du Karen commun. *Bulletin de la Societe de Linguistique de Paris* 70:339-343.

Jones, Robert B. 1961. *Karen linguistic studies: Description, comparison and texts*. Berkeley: University of California Press.

Karlgren, Bernhard. 1915. *Études sur la phonologie chinoise*. Archives d'Études Orientales.

Kauffman, William. 1993. *The great tone split and Central Karen,* University of North Dakota: MA.

Luce, Gordon H. 1959. Introduction to the comparative study of Karen languages. *Journal of the Burma Research Society* 42:32-51.

Luce, Gordon H. 1985. *Phases of Pre-Pagan Burma: Languages and history*. Oxford: Oxford University Press.

Manson, Ken. 2002. *Karen language relationships: A lexical and phonological analysis*. Chiang Mai: Dept of Linguistics, Payap University.

Manson, Ken. 2009. A prolegomena to reconstructing Proto-Karen. *La Trobe Working Papers in Linguistics* 12.

Manson, Ken. 2011a. *Karen tone boxes*. Karen Linguistic Discussion Group, Research paper 1. Chiang Mai: Dept of Linguistics, Payap University

Manson, Ken. 2011b. *Proto-Karen bilabials*. Karen Linguistic Discussion Group, Research paper 2. Chiang Mai: Dept of Linguistics, Payap University

Manson, Ken. 2011c. *Proto-Karen velars*. Karen Linguistic Discussion Group, Research paper 3. Chiang Mai: Dept of Linguistics, Payap University

Manson, Ken. 2011c. *Proto-Karen vowels*. Karen Linguistic Discussion Group, Research paper 9. Chiang Mai: Dept of Linguistics, Payap University

Manson, Ken & Grace Chou. 2008. *A survey of Kayan*. ms.

Shintani, Tadahiko. 2003. Classification of Brakaloungic (Karenic) languages, in relation to their tonal evolution. *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. 2001. *Another look at Proto-Karen*. Presented at 34th ICSTLL, Kunming

Thurgood, Graham. 2002. Vietnamese and tonogenesis: Revising the model and the analysis. *Diachronica* 19:333-363.

Thurgood, Graham. 2006. *Tonogenesis revisited: Revising the model and the analysis*. Ms

Weidert, Alfons K. 1987. *Tibeto-Burman tonology: A comparative account*. Amsterdam: John Benjamins