Atsuhiko Kato

18 Typological profile of Karenic languages

18.1 Introduction

Karenic languages constitute the Karenic branch of the Tibeto-Burman family. The ethnic groups that speak Karenic languages (called "Karenic people" here) include several groups, such as the Bwe, Geba, Gek(h)o, Kayah, Kayaw, Kayan, Manu, Monebwa, Mopwa (Mobwa), Paku, Pa-O, Pwo Karen, Sgaw Karen, Thalebwa, Yeinbaw, Yintale, and others (listed in alphabetical order). Today, these ethnic groups live in the eastern, southeastern, and southwestern parts of Myanmar and the northern and western parts of Thailand.

As pointed out by Kato (2016), the range of people who consider themselves to be ethnic "Karen" can vary depending on various contexts, such as political, ethnic, and linguistic. The Geba people, for example, have their own identity as a separate group from the Karen or Kayah; however, sometimes they may associate themselves as belonging to the Karen and, less often, to the Kayah. Despite this complicated situation, Pwo Karen and Sgaw Karen are always recognized as Karen, and, in the narrowest sense, can be considered to constitute the ethnic "Karen". In Myanmar, the Pwo Karen and Sgaw Karen jointly hold many events frequently, including traditional festivals such as the Karen New Year Festival. These two groups also often act together in political movements; the Karen National Union (KNU), one of the largest and keenest anti-government armed groups in Myanmar, are mainly comprised of Pwo and Sgaw. Intermarriage between Pwo and Sgaw is also fairly common. In addition to Pwo Karen and Sgaw Karen, the ethnic groups that speak languages close to Sgaw, i.e., Monebwa, Paku, and Thalebwa, often consider themselves Karen. Moreover, the Bwe, Geba, and Mopwa, whose languages cannot be genealogically said to be very close to Pwo or Sgaw but who have strong contact with the Sgaw people, may also consider themselves Karen. Other groups such as the Kayah, Kayan, and Pa-O usually see themselves as belonging to ethnic groups that are separate from Karen. However, all ethnic groups speaking Karenic languages may be considered as a unified ethnic unit that constitutes "Karen"; for example, a radical Karen nationalist would prefer this because the concept of "Karen" comprises more groups and has a larger population when all Karenic people are treated under the same rubric. The English word "Karen" probably originated from an old form of the Burmese exonym /kăyìɴ/ (cf. written form: <karaṅ>)[1], which can be further traced back to a Proto-Karen autonym. Some scholars have attempted to reconstruct it, e.g., *k-ɲaŋ^A (Solnit 2001), *bra-ka-louŋ (Shintani 2003a), and *k-rjaŋ^A (Luangthongkum 2012). Autonyms in contemporary Karenic languages include /phlòʊɴ/ (Eastern Pwo), /phlóuɴ/ (Western Pwo), /pɣākəɲɔ́/ (Sgaw), and /gèɓá/ (Geba), /kəjɛ̄/ (Kayah Li; Solnit 1997: xviii), and /kayân/ (Kayan; Manson 2010: xxi)

The *Ethnologue* (22nd edition; Eberhard et al. 2019) lists 21 Karenic languages: Bwe, Geba, Geko, Mobwa, Paku, Phrae Pwo, Eastern Pwo, Northern Pwo, Western Pwo, Sgaw, Kawyaw, Eastern Kayah, Western Kayah, Kayan, Kayaw, Lahta, Pa-O, Wewaw, Yinbaw, Yintale, and Zayein. Based on their estimated population data, the Karenic languages that have a relatively large population of speakers are: Sgaw (2,250,000), Pwo (1,326,000), Pa-O (858,740), Kayah (178,000), Kayan (133,180), Geba (40,000), Kayaw (20,100), and Bwe (17,200) (population counts in parentheses). The Genealogical relationships among the Karenic languages have been discussed by scholars, including Jones (1961), Manson (2002, 2017a), and Shintani (2003b), but opinions vary regarding the subgrouping of the languages.

The Union of Myanmar has two states for Karenic people: Kayin State (Karen State) and Kayah State. These are the only administrative units in the world that are legitimately

---

[1] The Mon exonym for Karen is /kəriəŋ/ (written form: <kareṅ>) (Mathias Jenny, p.c.). This Mon form could also be the origin of the word "Karen".

established for the Karenic people. People speaking Karenic languages are mostly farmers, who make swiddens in mountainous areas and wet rice fields in the plains. The majority of the Pwo Karen and Sgaw Karen population in Myanmar live in the plains of Kayin State, Bago Region, Yangon Region, Mon State, Tanintharyi Region, and Ayeyarwady Region, and some of the population live in urban areas including Yangon and Mawlamyine. However, Pwo Karen and Sgaw Karen in Thailand mainly live in the mountainous areas. Karenic people other than Pwo and Sgaw usually live in the mountainous areas of the southern Shan State, Kayah State, Kayin State, and Bago Region, with the exception of Pa-O, some of whom live in the plains of Kayin State and Mon State. In the plains, people that speak the Karenic languages are usually Buddhists or Christians, whereas in the mountainous areas, they are usually Christians or animists. The majority of Pa-O and Pwo Karen are Buddhists, Sgaw Karen also has a large population of Buddhists, and the rest are mostly Christians or animists.

Some of the Karenic languages have writing systems. Pwo Karen in Kayin State has several writing systems, including the Buddhist Pwo Karen Script (called /láithûlì/ in Pwo; I assume that it was created by Pwo Karen Buddhist monks in the late 18th or early 19th century based on the Mon script, though legend has it that it dates back to the Pagan Period), Christian Pwo Karen Script (created by American Baptist missionaries in the middle of 19th century based on the undermentioned Christian Sgaw Karen Script), and Leke Script (with original new shapes, which may be created in the middle of 19th century). For the Pwo Karen scripts, see Stern (1968) and Kato (2001a, 2001b, 2006). Different religious followers of Pwo Karen tend to use different writing systems (Leke is the religion specific to Pwo Karen in Kayin State). Many Sgaw Karen use the Christian Sgaw Karen Script; this was created in the 1830s by Wade, an American missionary, based on the Burmese script with a considerable modification. Some Buddhist Sgaw also use this, whereas some Buddhist Sgaw use the Buddhist Sgaw Karen Script (called /lìʔtəláʔɲá/ in Sgaw), which was probably created after the 1960s based on the Burmese script; moreover, some Buddhist Sgaw use the Myaing Gyi Ngu Script with original new shapes that were created probably in the 1990s. According to Solnit (1997: 304–308), Kayah has a roman-letter orthography developed by Catholic missionaries and an Indic-style orthography created by indigenous people. Geba and Gekho also have roman-letter writing systems created by Catholic missionaries. The Pa-O living in Shan State have a Burmese-based writing system that reportedly has a history of hundreds of years, and the Pa-O living in Mon and Kayin States use a modified version of this. There are several other scripts that have been developed recently: for example, the roman-letter Kayan orthography (Manson 2010: 67–69) and the Burmese-based Manu orthography (Wai Lin Aung 2013: 14)

All Karenic languages have an SVO word order, which is aberrant among the Tibeto-Burman languages, the majority of which are of the SOV-type. Most likely, the ancestor of the Karenic languages originally had an SOV word order, which changed to SVO at the Proto-Karen stage. This change was likely due to contact with some Mon-Khmer language(s). Matisoff (2000) suggests heavy contact with Mon in the late first millennium AD. Manson (2009) suggests that Mon-Khmer loanwords in Karenic languages imply a greater connection with the Palaungic branch of Mon-Khmer than the Monic branch. Kato (in print) argues that Pwo and Sgaw have loanwords from Mon that can be considered to be borrowed before Pwo and Sgaw split because they match the regular historical tonal changes observed in pure Karen words; he concludes that Pwo and Sgaw have a long history of contact with Mon, although it is not obvious whether the contact can be traced back to the Proto-Karen stage. Below are some examples. Mon forms shown in angle brackets "< >" indicate Shorto's (1962) literary forms.

(1) Eastern Pwo /pàɪntərân/, Western Pwo /təràɴ/, Sgaw /pɛ́trɔ́/ 'window, door', cf. Mon <pāṅtaraṅ>
     Eastern Pwo /phjâ/, Western Pwo /phjà/, Sgaw /phjá/ 'market', cf. Mon <phyā>
     Eastern Pwo /kəmâ/, Western Pwo /kəmà/, Sgaw /kəmá/ 'pond', cf. Mon <kamā>

Kato also notes in the same paper that Burmese and Tai loanwords in Pwo and Sgaw do not match the regular patterns of tonal changes, and concludes that the contact of Pwo and Sgaw with Burmese or Tai languages, including Thai and Shan, is relatively recent. At present, Karenic languages generally have great interaction with the Burmese or Thai, and have many Burmese loanwords in Myanmar and Thai loanwords in Thailand.

Dialectal variety within some Karenic languages is rather large. For example, Kato (2019) lists four dialectal groups of Pwo Karen that are not mutually intelligible: Western Pwo Karen, Htoklibang Pwo Karen, Eastern Pwo Karen, and Northern Pwo Karen. For more details regarding Pwo Karen dialects, see Kato (1995, 2009b), Dawkins and Phillips (2009a, b), and Phillips (2017). According to Manson (2017a, 2019), Kayan seems to have an even wider variety of dialects. He says that the "Kayan cluster" is highly diverse both phonologically and lexically, and that several speech varieties are mutually unintelligible (Manson 2019: 5).

In this chapter, to describe the general typological profile of Karenic languages, I will mainly use samples of the Hpa-an dialect of Eastern Pwo Karen (hereafter, Eastern Pwo Karen is abbreviated as EPK, and Pwo Karen as PK), as I am the most familiar with this dialect among the Karenic languages.[2] I will also use many samples of the Hpa-an dialect of Sgaw Karen (hereafter, abbreviated as SK), with which I am familiar to a certain extent. As stated by Manson (2017a), the modern comprehensive grammatical studies available in the present for Karenic languages are Solnit (1997) for Kayah Li, Manson (2010) for Kayan, and my grammar for Pwo Karen (Kato 2004); thus, in this chapter, Solnit (1997) and Manson (2010) are often referenced for comparison.

18.2 Phonology

18.2.1 Syllable structure and segmental phonemes

The syllable structure of EPK can be represented as C1(C2)V1(V2)(C3)/(T). "C" stands for a consonant, "V" for a vowel, and "T" for a tone. C1 is an initial consonant, C2 is a medial consonant, and C3 is a final consonant. One or two vowels may occur, and are represented with V1 and V2. Bracketed elements may or may not occur. The sequence of V1, V2, and C3 has limited combinations; thus, it is better to treat them as a unit, that is, a rhyme. Many EPK basic words consist of one syllable; thus, we can consider EPK to be a monosyllabic language. This monosyllabicity is also true of the other Karenic languages.

Table 1 shows all the EPK simple onsets and rhymes. The phonetic values that should be noticed are: /θ/[t̪~t̪θ~θ], /c/[tɕ], /ph/[pʰ], /th/[tʰ], /ch/[tɕʰ], /kh/[kʰ], /b/[ɓ], /d/[ɗ], /j/[j~ɟ], /r/[r~ɹ], /i/[ᵊi], /i̠/[ɪ], and /ʁ/[ʁ~ɦ]. The uvular nasal /ɴ/ can only occur as C3, and all the other consonants can occur as C1. The consonants that can occur as C2 are: /w, l, r, j/. The consonants /ɲ/, /ŋ/, and /r/ occur mostly in loanwords from Mon or Burmese. Rhymes are formed with the following four combinations: V1, V1V2, V1ɴ, and V1V2ɴ. There are 21 rhymes. The final uvular nasal /ɴ/ often functions as an element that only nasalizes the last part of the preceding vowel, especially in rapid speech. /ɴ/ of the rhymes /eiɴ/, /əɯɴ/, and /oʊɴ/ is often entirely dropped; in such cases, no nasalization is observed in the vowels, either. /i̠ɴ/ is only found in Burmese loanwords. The vowel of the rhyme /aɴ/ is phonetically realized as a diphthong [ăɔɴ]. This diphthong differs from the diphthongs of /eiɴ/, /əɯɴ/, /oʊɴ/, and /aiɴ/ because it is a rising diphthong, while the first and second elements of the diphthongs in /eiɴ/ [ei(ɴ)], /əɯɴ/ [əɯ(ɴ)], /oʊɴ/ [oʊ(ɴ)], and /aiɴ/ [aiɴ] are of equal prominence.

Table 1: Onsets and rhymes in EPK.

Onsets						Rhymes									
p	θ	t	c	k	ʔ	i	ɨ	ɯ	ai	aʊ	i̠N	əN	eiN	əɯN	oʊN
ph		th	ch	kh		i̠		ʊ			aN	ON	aiN		
b		d				e	ə	o							
			ɕ	x		h	ɛ	a	ɔ						
				ɣ	ʁ										
m		n	ɲ	ŋ											
w			j												
		l													
		r													

Next, let us look at SK simple onsets and rhymes for comparison. Table 2 shows all the SK onsets and rhymes. The phonemic transcription of SK follows Kato (2002). SK /θ/ is pronounced [t̪~t̪θ~θ], the same as for Eastern Pwo /θ/. Both come from Proto-Karen *s, which was reconstructed by Haudricourt (1946) and is still preserved in many SK dialects spoken in Thailand. /s/ [s] and /sh/ [sʰ] come from *c and *ch (see Manson 2009), which are also preserved in many SK dialects spoken in Thailand. The SK phonemes /b/ and /d/ are implosives, similar to Eastern Pwo /b/ and /d/. These can be considered as reflects of Proto-Karen *ɓ (or *ʔb) and *ɗ (or *ʔd). In many Karenic languages, these Proto-Karen sounds are preserved as implosives or preglottalized voiced stops. Note that Haudricourt's (1946) Proto-Karen plain voiced stops *b, *d, and *g have changed to /ph/, /th/, and /kh/ in PK, and /p/, /t/, and /k/ in SK, as will be mentioned again in 18.2.2.

Table 2: Onsets and rhymes in SK.

Onsets						Rhymes					
p	θ	t	c	k	ʔ	i	ɯ	u	iʔ	ɯʔ	uʔ
ph		th	ch	kh		e	ə	o	eʔ	əʔ	oʔ
b		d				ɛ	a	ɔ	ɛʔ	aʔ	ɔʔ
		s	ɕ	x	h						
		sh									
		z		ɣ	ɦ						
m		n	ɲ	ŋ							
w			j								
		l									
		r									

It is certain that Proto-Karen had series of syllable-final nasals and stops, as assumed by scholars including Manson (2009), Solnit (2013), and Luangthongkum (2019). Proto-Karen probably had final nasals such as *-m, *-n, and *-ŋ because these are still present in Pa-O (see Jones 1961). However, regarding the final stops, it is difficult to reconstruct them with indisputable certainty because all Karenic languages have reduced them to a glottal stop or completely dropped them. Pa-O has many rhymes with final stops /-p, -t, -k, -ʔ/; however, as Solnit (2013) and Manson (2017a: 154) state, most of these with /-p, -t, -k/ are borrowed from Tai (Shan) or Pali. SK still preserves a final glottal stop, but final nasals have entirely disappeared. In this regard, Pwo is more conservative than Sgaw. In the Hpa-an dialect of Pwo, Proto-Karen final nasals are preserved as a uvular nasal. Final stops have completely disappeared in Hpa-an; however, a final glottal stop is still preserved in the Kyonbyaw dialect of Western PK and the Tavoy dialect of EPK (see Kato, 1995). For example, compare Hpa-an /thò/ 'pig' with Kyonbyaw /thoʔ/ 'pig' and Tavoy /thòʔ/ 'pig'. Generally, Proto-Karen rhymes

have been highly simplified in every Karenic language. Such simplification has occurred due to the dropping and merging of final consonants, as mentioned above, as well as the monophthongization of diphthongs. Examples of monophthongization would be EPK /ʔɛ́/ 'to love' and SK /ʔɛ̂/ 'to love', whose vowels probably used to be *ai, which is preserved in Western PK, e.g. /ʔài/ 'to love'. The simplification of rhymes is an on-going phenomenon. For example, as mentioned above, the final /ɴ/ of /eiɴ/, /əɯɴ/, and /oʊɴ/ in the Hpa-an dialect of EPK is often dropped. Moreover, the EPK final /-aɴ/ can also be pronounced as /-ɔ/ in some words, such as /náɴ ~ nɔ́/ 'to wake' and /màɴ ~ mɔ̀/ 'to bark (as a dog)'.

The possible combinations of C1 and C2 in EPK are shown in Table 3. The medial /r/ in EPK generally occurs only in Mon or Burmese loanwords. Table 4 shows the possible combinations of C1 and C2 in SK, which appears in Wade (1849).

Table 3: Possible combinations of C1 and C2 in EPK.

		C1																		
		p	θ	t	c	k	ʔ	ph	th	ch	kh	b	d	x	h	m	n	ɲ	j	l
C2	w	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+
	l	+			+		+				+	+				+				
	r	+			+						+									
	j	+								+		+				+				+

Table 4: Possible combinations of C1 and C2 in SK.

		C1																		
		p	θ	t	k	ph	th	kh	b	d	s	sh	x	m	n	ɲ	ŋ	j	l	r
C2	w	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+
	l	+		+	+				+	+				+						
	r	+	+	+	+	+	+	+			+	+		+						
	j	+				+			+					+						
	ɣ	+				+			+		+	+		+						

The medials of Karenic languages are generally semivowels or liquids. Even the SK fricative medial /-ɣ-/ [ɣ] can be pronounced [ɰ], which is also noted by Manson (2017a: 154), e.g. /pɣā/[pɣā~pɰā] 'person'. In dialects of the Ayeyarwady delta, it has generally merged with /-w-/ when it occurs after a bilabial, e.g. /pwā/ 'person'. The SK medial /-ɣ-/ can be traced back in many cases to the Proto-Karen medial *-r-. The Proto-Karen medial *-l- is relatively well preserved in many Karenic languages. An example would be words that mean 'arrow': EPK /phlā/, SK /plà/, Bwe /blɛ/ (Henderson 1997: vol. 2, 7), Eastern Kayah Li /plè/ (Solnit 1997: 354), and Kayan /plā/ (Manson 2010: 234).

18.2.2   Suprasegmental phonemes

EPK (the Hpa-an dialect) has four tones, as shown below with examples:

High-level	á	[a55]	/má/ 'son-in-law'
Mid-level	ā	[a̤33 ~ 334]	/mā/ 'very'
Low-level	à	[a11]	/mà/ 'to do'
Falling	â	[a51]	/mâ/ 'wife'

The mid-level tone is pronounced with a breathy phonation and may be pronounced with a rising pitch in utterance-final position and before a pause.

SK (the Hpa-an dialect) has four plain tones and two checked tones:

High-level   á   [a55]   /má/ 'wife'   High-checked   áʔ [aʔ44]/náʔ/ 'sword'

Mid-level	ā	[a33]	/mā/ 'to do'	
Low-level	à	[a11]	/nà/ 'ear'	Low-checked à? [a?11]/mà?/ 'son-in-law'
Falling	â	[a51]	/ɲâ/ 'fish'	

The high- and low-checked tones can be interpreted either as the high- (or possibly mid-) and low-level tones occurring in a syllable with the final glottal stop or as tones that are independent tones from the plain ones. It is uncertain which interpretation is appropriate.

Table 5 shows how the Proto-Karen tones (represented as 1, 2, 2′, and 3)[3] reconstructed by Haudricourt (1946, 1953, 1975) have changed in EPK and SK. Tone 3 is the tone with final stops. Haudricourt (1946) only reconstructed Tones 1, 2, and 3, and added one more proto-tone, i.e. Tone 2′, in his 1975 paper.[4] B, M, and H represent the initial consonant groups of Proto-Karen. B is the group of voiced consonants, M is the group of voiceless non-aspirated stops including glottalized/implosive stops, and H is the group of voiceless aspirated stops as well as voiceless fricatives and sonorants. The Proto-Karen tones are split in this way based on the types of initial consonants. The tonal split in Karen resembles that of the Tai languages (see Li 1977) in that tones split based on the three types of initial consonants. For details regarding Haudricourt's reconstruction of Proto-Karen tones, see Kato (2018).

Table 5: Proto-Karen tones and the tones of EPK (left) and SK (right).

	1 (plain)	2 (plain)	2′ (plain)	3 (checked)
B	/à/[11] : /ā/[33]	/ā/[33] : /à/[11]		/á/[55] : /à?/[11]
M	/à/[11] : /á/[55]	/á/[55] : /â/[51]	/á/[55] : /á/[55]	/à/[11] : /à?/[11]
H	/â/[51] : /á/[55]	/á/[55] : /â/[51]	/á/[55] : /á/[55]	/à/[11] : /á?/[44]

Haudricourt (1946) reconstructed the Proto-Karen tones using the PK dictionary by Purser and Saya Tun Aung (1922), which also shows corresponding SK forms in brackets. Haudricourt (1975) assumes that Tone 2′ merged with Tone 2 in syllables beginning with B-series consonants at the Proto-Karen stage; thus, there are no contemporary examples of B2′. All Karenic languages still have tones. The tonal correspondences of 16 Karenic languages with Proto-Karen are shown in Shintani (2003b). If we consider his description and if tones with a final glottal stop (i.e. checked tones) are treated as independent tones, Karenic languages have either two, three, four, five, or six tones.

As has been already mentioned in 18.2.1, the Proto-Karen initials *b, *d, and *g have changed to /ph/, /th/, and /kh/ in PK, and /p/, /t/, and /k/ in SK. This devoicing happened in the process of the tonal split. When Haudricourt (1946) reconstructed Proto-Karen *b, *d, and *g, no Karenic language that has [b], [d], or [g] had been found. However, his assumption that the Proto-Karen had voiced stops was attested by the reports of Luce (1959) and Henderson (1961, 1979) on Bwe Karen, which preserves plain voiced stops. According to Shintani (2003b), Proto-Karen *b, *d, and *g are still preserved intact as [b], [d], and [g] in Bwe, Geba, Paku, and Monebwa, and these languages have a four-way contrast of homorganic stops, e.g. /p/, /ph/, /b/, and /ɓ/, which is a contrast that Proto-Karen probably used to have. In the other Karenic languages, *b, *d, and *g have changed to [p], [t], and [k], or [ph], [th], and [kh]. For example, Proto-Karen *go[2] 'hot' became /khŭ/ in EPK and /kò/ in SK. The Geba form corresponding to them is /gō/ (Kato 2008a: 191). Similarly, in all Karenic languages except Geba, Proto-Karen voiceless sonorants belonging to the H-series have been voiced in the process of tonal split. For example, Proto-Karen *m̥a[1] 'wife' became /mâ/

---

[3] Luangthongkum (2019) claims that Haudricourt's (1975) reconstruction of Tone 2′ is unnecessary. However, as I discussed in Kato (2018), reconstructing Tone 2′ can explain the tonal developments that occurred in the transition from Proto-Karen to modern Karenic languages. This hypothesis that assumes Tone 2′ is accepted by Solnit (2001, 2013), Shintani (2003b), and Manson (2009). Tones 1, 2, 2′, 3 may be represented as Tones A, B, B′, C or A, B, B′, D by different scholars.

[4] Haudricourt himself did not number the added tone as 2′. I am using this numbering for convenience.

in EPK and /má/ in SK. However, only Geba preserves voiceless sonorants (see Kato 2008a and Naw Hsar Shee 2008), and the Geba word corresponding to 'wife' is /m̥ɛ́/ (Kato 2008a: 184).

EPK has atonal syllables, which are phonetically unstressed and short, and cannot occur utterance-finally. They must be followed by a syllable with a tone when they occur in an utterance. The only vowel that can occur in an atonal syllable is /ə/, and it is transcribed as /Cə/ without a tonal notation. In an atonal syllable, C2 never occurs. EPK has many disyllabic words that have an atonal first syllable. These are "sesquisyllabic" in Matisoff's (1973: 84ff) terms, in that they are composed of the unstressed first syllable and the stressed second syllable. Examples are: /kəchā/ 'lord, master', /kəchân/ 'elephant', /cəxwà/ 'king', /chərâ/ 'teacher', /chəná/ 'evil-spirit', /təkhwâ/ 'cousin', /təɕā/ 'surely', /təwân/ 'village', /thərài/ 'expense', /pənā/ 'buffalo', /phəjā/ 'to release', and /θədán/ 'shrimp'. SK also has many sesquisyllabic words, such as /kəsà/ 'lord, master', /kəshó/ 'elephant', /θərâ/ 'teacher', /təkhwá/ 'cousin', /θəwɔ́/ 'village', /pənà/ 'buffalo', and /ləpɔ́/ 'wave'.

Karenic languages generally have sesquisyllabic words. Solnit (1997: 25) presents Kayah Li examples, such as /kədā/ 'door' and /təmɔ̀/ 'sun'. Manson (2010: 57) presents Kayan examples, such as /na.là/ 'ear', /ka.nɔ̰́/ 'brain', and /ʔa.cʰû/ 'thorn'. In PK and SK, the minor syllable in a sesquisyllabic word never has a tone; however, in some other Karenic languages, it may bear tones. For example, according to Solnit (1997: 25), the Kayah Li prefix /ʔi-/, which appears as the minor syllable in a sesquisyllabic word, may have two tones (low or high), whereas full syllables may have five tones.

In some Karenic languages, some prefixes or proclitic morphemes with atonal or unstressed syllables show vowel harmony. For example, according to Kato (2008a), in Geba, the vowel of the personal pronouns /jV/ (1SG), /wV/ (1PL), /nV/ (2SG), and /sV/ (3SG) show vowel harmony with the vowel of the following syllable. 'V' here represents an atonal vowel that is the same as that of the following syllable. Examples are: /ja ʔā/ (1SG - to.eat) 'I ate', /je lē/ (1SG - to.go) 'I went', /jɛ mɛ̄/ (1SG - to.do) 'I did', and /jɔ ɗɔ/ (1SG - to.speak) 'I spoke'. Apart from these personal pronouns that occur before a verb or a noun, vowel harmony in Geba is also observed in the negative marker /tV/ that occurs before a verb, in the realis marker /kV/ that occurs before a verb, and in the numeral /tV/ 'one' that occurs before a numeral classifier. Not all atonal morphemes show vowel harmony. An example would be the nominalizing prefix /ʔa-/, e.g. /ʔaθɛ̄/ 'fruit' (< /ʔa-/ + /θɛ̄/ 'to bear fruit'). Solnit (1997: 26) reports that vowel harmony is observed in Kayah Li, Kayaw, and Bwe (see also the "Introduction" of Henderson [1997] written by Solnit). Manson (2010: 152) also reports that a Kayan adverbializer /θa-/ shows vowel harmony. In PK and SK, vowel harmony is not observable.

In EPK, the first syllables of some disyllabic words can be pronounced either as a full syllable with a tone or as an atonal syllable. Examples are: /khʉ́lòɴ ~ khəlòɴ/ 'mountain', /phɨ́bàɪɴ ~ phəbàɪɴ/ 'blanket', and /khánthài ~ khəthài/ 'base, foundation'. These forms tend to appear in rapid speech with an atonal syllable. Manson (2017a: 153) mentions that the Kayan word /kʰaŋ⁴².du⁴²/ [kʰan⁴².du⁴²] 'thigh' is also pronounced with forms that have weakened first syllables including [kʰə.du⁴²] and [kə.du⁴²]. These examples imply that one of the sources of sesquisyllabic words in Karenic languages is the weakening of the full first syllables of disyllabic words.

18.3 Word classes

Kato (2004, 2008b) groups EPK words into five word classes: nouns, verbs, adverbs, particles, and interjections. It is unnecessary to set up the category of adjectives, and words that denote states are considered a subgroup of verbs, i.e., stative verbs. These word classes can be defined in terms of the bundle of syntactic features that each word class has. Interjections are considered as a special word class that cannot have a syntactic relationship with other elements. For the other four word classes, Kato (2004, 2008b) proposes three tests to distinguish them: Test 1 is whether a word can constitute an utterance on its own;

Test 2 is whether a word allows a verb particle to occur; and Test 3 is whether a word can be an argument of a verb. See Table 6.

Table 6: Tests to distinguish EPK word classes.

	Test 1	Test 2	Test 3
Nouns	Yes	No	Yes
Verbs	Yes	Yes	Yes
Adverbs	Yes	No	No
Particles	No	No	No

Nouns, verbs, and adverbs can constitute an utterance on their own, but particles cannot. Only verbs can allow a verb particle to occur. Verb particles are one of the seven kinds of particles. Kato (2004: 370–372) lists over sixty verb particles, including the negative marker /lə/ and the irrealis marker /mə/. For example, in /ʔəwê mə ɣɛ̂/ (3SG - IRR - to.come) 'S/he will come', the irrealis marker /mə/ cannot occur without the presence of a verb /ɣɛ̂/; thus, a verb allows a verb particle to occur. Nouns and verbs can function as the subject argument or object argument of a verb. For example, in /jə θànán láiʔàʊ/ (1SG - to.forget - book) 'I forgot a book', the noun /láiʔàʊ/ 'book' is the object argument of the verb /θànán/ 'to forget', and in /jə θànán ɣɛ̂/ (1SG - to.forget - come) 'I forgot to come', the verb /ɣɛ̂/ 'to come' functions as the object argument of the same verb, /θànán/ 'to forget'. The grammars of Karenic languages by Solnit (1997) for Kayah Li, Manson (2010) for Kayan, and Kato (2004) for Pwo Karen, establish different word classes; however, all these grammars commonly recognize nouns, verbs, and adverbs, and consider the class of adjectives as unnecessary.

18.4 Word formation

EPK is an isolating type of language; its words do not inflect at all. However, there are three productive word-formation processes. They are compounding, affixation, and reduplication.

Compounding is a highly productive process of word formation in EPK. Nouns and verbs are involved in compounding and the resultant words are also nouns or verbs. There are four patterns in the formation of compound nouns: N+N > N, N+V > N, V+V > N, and V+N > N. In the formation of compound verbs, there are three patterns: V+V > V, N+V > V, and V+N > V. Here, we will look at four patterns from the patterns mentioned above: N+N > N, N+V > N, V+V > V, and V+N > V.

- N+N > N. In this pattern, the second element is usually the semantic head:

    (2) /mé/ 'eye' + /thî/ 'water' > /méthî/ 'tear'

    (3) /ɣéiɴ/ 'house' + /khʊ́/ 'head' > /ɣéiɴkhʊ́/ 'roof of a house'

    (4) /phɔ̂/ 'flower' + /phɜ̀ɴ/ 'pot' > /phɔ̂phɜ̀ɴ/ 'vase'

    (5) /déiɴ/ 'sesame' + /θʊ́/ 'oil' > /déiɴθʊ́/ 'sesame oil'

However, in some words, the former element is the semantic head, as in (6) and (7), and there are also cases in which both elements can be considered the semantic heads, as in (8) and (9):

    (6) /láiɴ/ 'cart' + /mí/ 'fire' > /láiɴmí/ 'train'

    (7) /phɜ̀ɴ/ 'pot' + /thà/ 'iron' > /phɜ̀ɴthà/ 'Chinese pot made of iron'

    (8) /mʊ̄/ 'mother' + /phā/ 'father' > /mʊ̄phā/ 'parents'

(9)   /chái/ 'rice field' + /xàʊ/ 'swidden' > /cháixàʊ/ 'agricultural land'

In "N+N > N" pattern, the second element is sometimes a numeral classifier, as in (10) and (11). Numeral classifiers constitute a subgroup of nouns.

(10)   /lái/ 'writing' + /béiɴ/ 'ɴC[flat thing]' > /láibéiɴ/ 'booklet'

(11)   /lái/ 'writing' + /phlóʊɴ/ 'ɴC[round thing]' > /láiphlóʊɴ/ 'alphabet'

● N+V > N. This pattern can be considered in terms of syntactic relationship of N and V; 'N' may be the subject of the 'V' as in (12), the object as in (13), and the adjunct as in (14) and (15). In many cases, 'N' denotes a generic concept of the resultant noun but sometimes, it is not the case; In (15), a 'handle' is not a kind of 'hand'.

(12)   /lōʊɴ/ 'stone' + /jî̀/ 'be green' >/lōʊɴjî̀/ 'jade'

(13)   /lái/ 'writing' + /pɔ̄/ 'to read' > /láipɔ̄/ 'textbook'

(14)   /thà/ 'iron, needle' + /chà/ 'to sow' > /thàchà/ 'sewing needle'

(15)   /cɯ́/ 'hand' + /phóɴ/ 'to hold' >/cɯ́phóɴ/ 'handle, grip'

● V+V > V. Examples are below:

(16)   /ʔáɴ/ 'to eat' + /ʔɔ̀/ 'to drink' > /ʔáɴʔɔ̀/ 'to eat and drink'

(17)   /thé/ 'to be cut' + /phà/ 'to be split' > /théphà/ 'to crack, explode'

(18)   /ɕɯ́/ 'be calm' + /máʊ/ 'be comfortable' > /ɕɯ́máʊ/ 'be peaceful'

● V+N > V. Examples are below. In (20), /pərêɴ/ is not the object noun of the verb /kè/ because /kèpərêɴ/ as a whole can take an object, as in /kèpərêɴ lái/ (to.write – letter) 'to write a letter'.

(19)   /bá/ 'to hit' + /θà/ 'heart' > /báθà/ 'to want (something)'

(20)   /kè/ 'to write' + /pərêɴ/ 'news' > /kèpərêɴ/ 'to write (as a letter)'

Compounding is highly productive in Karenic languages in general. Some of the compound words seem to be able to trace back to the Proto-Karen stage. An example would be words that mean 'tree leaf': EPK /θéiɴlá/ (< /θéiɴ/ 'tree' + /lá/ 'leaf'), SK /θêlâ/ (< /θê/ 'tree' + /lâ/ 'leaf'), Kayah Li /sɔle/ (< /sɔ/ 'tree' + /le/ 'leaf') (Solnit 1997: 44), and Kayan /θêŋlâ/ (< /θêŋ/ 'tree' + /lâ/ 'leaf') (Manson 2007: 58).

Reduplication in EPK only applies to verbs. Nouns are never reduplicated. Reduplication derives an adverb from a stative verb. For example, /phlɛ́phlɛ́/ 'fast' is a reduplicated form of the verb /phlɛ́/ 'be fast', and it functions as an adverb that occurs after the verb and modifies it. Other examples are: /xɛ̀xɛ̀/ 'slowly' (< /xɛ̀/ 'be slow'), /ʔáʔá/ 'much' (< /ʔá/ 'be many'), and /ɣ̀ìɣ̀ì/ 'well' (< /ɣ̀ì/ 'be good'). When the verb is disyllabic, it is usually reduplicated in the pattern of AABB, e.g. /thîthîchàchà/ 'exactly' (< /thîchà/ 'be exact') and /xûxûxàɴxàɴ/ 'unitedly' (< /xûxàɴ/ 'be united'). However, some disyllabic words are reduplicated in the form of ABAB, e.g. /bádàbádà/ 'moderately' (< /bádà/ 'be moderate'). In Kayan, reduplication also has a function of deriving adverbs from the stative verbs (Manson 2010: 151). Kayah Li reduplication does not seem to have such a function but has an interesting function of expressing the meaning of 'also, too, either' (Solnit 1997: 52–53). For example, in the

9

sentence /vɛ̄ cwá to to/ (I - to.go - NEG - RDP) 'I won't go either', the meaning of 'either' is expressed by reduplicating the last syllable of the clause. This process can be applied to any syntactic element as far as it is in the clause-final position.

EPK has a limited number of affixes. Kato (2004: 57–62) lists ten affixes; eight out of these are prefixes and two are suffixes. Here, we will see two important affixes.

● /ʔə-/ is a derivational prefix attached to a verb or noun. This morpheme is related to the Proto-Tibeto-Burman prefix ***a-** (Benedict 1972: 121–123; see also Matisoff 2003: 104–117; and Matisoff 2018), which has various functions in Tibeto-Burman languages, including nominalization. It is homophonous with the 3rd person singular pronoun /ʔə/ 's/he; his or her'. It is prefixed to verbs and derives nouns. Examples are: /ʔəkhʊ̄/ 'hot thing' (< /ʔə-/ + /khʊ̄/ 'be hot'), /ʔəchâ/ 'injury, wound' (< /ʔə-/ + /châ/ 'to ache'), /ʔədà/ 'rug' (< /ʔə-/ + /dà/ 'to spread'), /ʔəθá/ 'fruit' (< /ʔə-/ + /θá/ 'to bear fruit'), /ʔəθâɴ/ 'new one' (< /ʔə-/ + /θâɴ/ 'be new'), and /ʔəʔwà/ 'white' (< /ʔə-/ + /ʔwà/ 'be white'). This prefix is also prefixed to a noun and derives another noun. Examples are: /ʔəkhâɪɴ/ 'backside, rear' (< /ʔə-/ + /khâɪɴ/ 'buttock') and /ʔəkhʉ́/ 'roof' (< /ʔə-/ + /khʉ́/ 'head'). Moreover, this prefix is also utilized to change a dependent nominal morpheme into an independent noun. For example, in /ʔəchóɴ/ 'body hair', the part /chóɴ/ cannot occur as an independent noun but evidently indicates the meaning of 'body hair' because it is used in a compound noun; an example of this would be /kháchóɴ/ 'beard' (< /khá/ 'chin' + /chóɴ/ 'body hair'). The independent noun /ʔəchóɴ/ is formed by prefixing /ʔə-/ to the bound form /chóɴ/. Examples of nouns of this type are: /ʔənā/ 'blade', /ʔəphèɴ/ 'inside', /ʔəmèɪɴ/ 'name', /ʔəkhwâ/ 'female', /ʔəmʉ́/ 'male', and /ʔəlʊ̄/ 'voice'. Some words have both forms with and without this prefix. That is, both forms can be used as an independent noun. Examples of this are: /(ʔə)chʉ̋/ 'thorn', /(ʔə)khlî/ 'seed', and /(ʔə)nòʊɴ/ 'horn'. The prefix /ʔə-/ is evidently cognate with the Kayah Li prefix /ʔa-/ (Solnit 1997: 41–44), the Kayan prefix /a-/ (Manson 2010: 104–105), and the SK prefix /ʔə-/. Proto-Karen probably had a prefix that can be reconstructed as *(**ʔ)a-**, from which these prefixes have developed.

● /chə-/ is a derivational prefix related to the noun /chē/, meaning 'thing', which originated from Proto-Karen ***da²** 'thing'.[5] It is mainly prefixed to verbs and derives nouns. Examples are: /chəkhlàɪɴ/ 'language' (< /chə-/ + /khlàɪɴ/ 'to speak'), /chəkhʊ̄/ 'heat, hotness' (< /chə-/ + /khʊ̄/ 'be cold'), /chəchèɴ/ 'rain' (< /chə-/ + /chèɴ/ 'to rain'), /chədóɴ/ 'wall' (< /chə-/ + /dóɴ/ 'to fence'), /chəmà/ 'job' (< /chə-/ + /mà/ 'to do'), and /chəʔɛ́/ 'love' (< /chə-/ + /ʔɛ́/ 'to love'). This prefix may also occasionally be attached to a noun. Examples are: /chəpərâɴ/ 'information' (< /chə-/ + /pərâɴ/ 'news about somebody') and /chəɣàɴ/ 'picture' (< /chə-/ + /ɣàɴ/ 'appearance, figure'). Compared to /ʔə-/, nouns that are derived with /chə-/ tend to have more abstract meanings; compare /ʔəkhʊ̄/ 'hot thing' with /chəkhʊ̄/ 'heat, hotness'. This prefix is cognate with the Kayan prefix /ta-/ (Manson 2010: 105–107) and SK nominalizer /tà/. It is not certain whether the proto-form ***da²** 'thing' had already been grammaticalized as a prefix at the Proto-Karen stage.

18.5 Clause structure

The basic structure of the EPK verb-predicate clause can be schematized as presented in Figure 1. In the position of the verb, represented as 'V', concatenated verbs, which will be discussed later, may occur. I call the part consisting of the verb and the verb particle(s) as a verb complex. In addition to the elements shown in the schema, after the adverbial elements,

---

[5] The Proto-Karen ***d** generally became /th/ in Pwo Karen dialects. In this regard, the onset of the prefix /chə-/ shows an irregular correspondence with the Proto-Karen ***d**. In Northern Pwo Karen, spoken in northern Thailand, the form corresponding to the Eastern prefix /chə-/ is /thə-/ (see Phillips 2017: 70–80), whose onset shows a regular correspondence with Proto-Karen. Probably, the onset of the Eastern Pwo prefix /chə-/ used to be an alveolar stop and became /ch/ later for some reason.

another verb complex may occur, which is the V2 of a separated type serial verb construction. Furthermore, some adverbial elements may appear clause-initially.
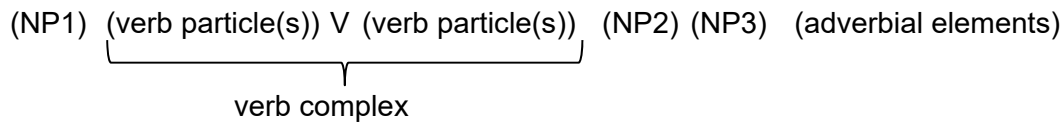
(NP1) (verb particle(s)) V (verb particle(s)) (NP2) (NP3)   (adverbial elements)

verb complex

Figure 1: Basic structure of the EPK clause.

NP1 is the subject, and NP2 and NP3 are the objects of the verb. When the verb is monotransitive, NP2 can appear, and when the verb is ditransitive, both NP2 and NP3 can appear. When the verb is intransitive, only NP1 can appear. (21), (22), and (23) are examples of intransitive, monotransitive, and ditransitive sentences of EPK, respectively. In many Southeast Asian and East Asian SVO languages, existential and phenomenon sentences utilize a VS order; however, EPK does not use that order in such sentences, as seen in (21). In a ditransitive clause, the noun denoting Recipient occurs immediately after the verb, and the noun denoting Theme follows it, as seen in (23). All Karenic languages have basic SVO word order, and the positions of the arguments of verbs are basically the same as EPK.

(21) *phlòʊɴmwì*   *ɣɛ̂*   [EPK]
      guest   to.come
      'A guest came; There comes a guest.'

(22) *ɕáphàɴ*   *dʊ́*   *θàkhléiɴ*   [EPK]
      PN   to.hit   PN
      'Shapan hit Thakhlein.'

(23) *ɕáphàɴ*   *phi̠lâɴ*   *θàkhléiɴ*   *khòθá*   [EPK]
      PN   to.give   PN   mango
      'Shapan gave Thakhlein a mango.'

To see elements other than NPs, let us take sentence (24) as an example:

(24) *nə*   *mə*   *ʔáɴ*   *bá*   *kʊ́*   *ʔáʔá*   *lə́*   *jə*   *ɣéiɴ*   *phə̀ɴ*   *ɕī̠*   [EPK]
      2SG   IRR   to.eat   OPP   cake   much   LOC   1SG   house   inside   too
      'You will also get a chance to eat a lot of cake inside my house.'

/mə/ and /bá/ are verb particles. /mə/ is the irrealis marker, and /bá/ denotes opportunity. In this example, /mə ʔáɴ bá/ is the verb complex. Kato (2004) eleven verb particles that appear before the verb and fifty verb particles that appear after the verb.

In the position of "adverbial elements", adverbs, adpositional phrases, adverbial particles, and numeral-classifier phrases may occur. In (24), /ʔáʔá/ is an adverb, /lə́ jə ɣéiɴ phə̀ɴ/ is an adpositional phrase, and /ɕī̠/ is an adverbial particle. These adverbial elements occur after the verb. Adpositional phrases are formed by adpositional particles including /lə́ ~ lé/ 'at; from; to', /dē/ 'with', /thōɴ/ 'around', /ni̠/ 'as much as', and /bê ... θò/ 'like'. Adpositional particles generally precede the noun phrase, with the exception of /bê ... θò/, which is a circumposition. Out of the constituents that occur in the position of adverbial elements, adpositional phrases may be topicalized and placed in the sentence-initial position (see Section 13 for topicalization).

The form /phə̀ɴ/ 'inside' in (24) is a locational noun. Its independent form is /ʔəphə̀ɴ/ 'inside' with the prefix /ʔə-/, which is dropped in most cases when it modifies a noun. Locational nouns occur after the noun that they modify. They are a subgroup of nouns;

11

however, they are special because they can function like an adposition. For example, in /ʔəwê ʔɔ́ (lə́) ɣéiɴ phə̀ɴ/ (3SG - be - LOC - house - inside) 'He is in the house', the adpositional particle /lə́/ can be omitted because /phə̀ɴ/ can make the preceding noun function as an adjunct. (However the omission seems to occur usually when the adpositional phrase appears immediately after the verb.) Other locational nouns include /ʔəklà/ 'between', /ʔəméjâ/ 'front', /ʔəlāɴkhâiɴ/ 'back', /ʔəphâɴkhʉ́/ 'on, above', and /ʔəphâɴlá/ 'under'.

It is difficult whether we should recognize noun-predicate clauses in EPK. In a sentence such as /ʔəwê mwɛ̄ phlòʊɴ/ (3SG - COP - Karen) 'He is a Karen', the copular verb /mwɛ̄/ is occasionally dropped. However, some native speakers say that it makes the expression sound rude or sloppy. Therefore, it cannot be said with certainty that the expression /ʔəwê phlòʊɴ/ 'He is a Karen' is completely acceptable.

Karenic languages generally do not have a tense marking. However, in EPK, the irrealis marker /mə/ (also pronounced as /mʊ̄/) helps to indicate the future, as is seen in (24) above. A more simple EPK example is /jə mə ɣɛ̂/ (1SG - IRR - to.come) 'I will come'. The SK irrealis marker /kə/ also helps to indicate the future, as in: /jə kə hɛ́/ (1SG - IRR - to.come) 'I will come'. These markers should not be considered as a future marker because they can be used to denote a surmise about past time. For example, in EPK, /mʊ̄ɣá ʔəwê mə ʔɔ́ lə́ jò/ (yesterday - 3SG - IRR - to.be - LOC - here) 'Yesterday, he was probably here'. However, it is certain that EPK /mə/ and SK /kə/ are usually used in a clause depicting an event in the future. Kayan has an irrealis prefix /kà-/ (Manson 2010: 140–141), and Geba also has a preverbal irrealis particle /kV/ (Kato 2008a: 215–216). The forms of SK, Kayan, and Geba are probably cognates. EPK /mə/ is related to the Western Pwo Karen verb particle /mô/, which means 'want to'. The Western Pwo irrealis marker is /kə/, and this is probably cognate with the SK, Kayan, and Geba forms. Thus, it seems that EPK used to have an irrealis marker that is cognate with Western /kə/, but it was replaced by /mə/ later.

Although Karenic languages do not mark tense, they have markers denoting various aspect categories. EPK has a perfective marker, /jàʊ/, which is an adverbial particle. An example is /ʔəwê ɣɛ̂ jàʊ/ (3SG - to.come - PFV) 'He has come'. The SK equivalent is /lí/, as in /ʔəwɛ́ hɛ́ lí/ (3SG - come - PFV) 'He has come'. Both EPK and SK do not have a marker that denotes continuous aspect. Thus, EPK /jə ʔáɴ mị̀/ (1SG - to.eat - rice) and SK /jə ʔɔ̂ mē/ (1SG - to.eat - rice) can both express meanings, such as 'I ate rice', 'I was eating rice', 'I am eating rice', and 'I eat rice (every day)'. Kayan (Manson 2010: 137) has a continuous aspect prefix /âu-/, for example, /âu-mé/ (CONT-to.sleep) 'to be sleeping', which originated from the verb meaning 'to exist, live'. However, in EPK and SK, the EPK verb /ʔɔ́/ and SK verb /ʔɔ̂/, both of which mean 'to exist, live' have not been grammaticalized to a form denoting aspect.

18.6 Pronouns

The EPK pronouns have three forms: form I, form II, and the emphatic form, as is shown in Table 7. Form I is used for the subject of a verb, e.g., /**jə** ʔáɴ khòθá/ (1SG - to.eat - mango) 'I ate a mango'. It is also used in Slot 3 (see Section 7) of the noun phrase to denote a possessor, e.g., /**jə** khòθá/ (1SG - mango) 'my mango'. Form II is mainly used for the object of a verb, e.g., /jə dá **nè̠**/ (1SG - to.see - 2SG) 'I saw you', or the object of an adpositional particle, e.g., /jə ʔáɴ dē **ʔè̠**/ (1SG - to.eat - with - 3SG) 'I ate with him/her'. It is also used for the pronouns that are topicalized (see Section 13), e.g., /**jè̠** nɔ́ ʔáɴ khòθá/ (1SG - TOP - to.eat - mango) 'As for me, I ate a mango'. The emphatic form is typically used when the pronoun is emphasized, whatever syntactic role the emphasized pronoun has, e.g., /**jəwê** ʔáɴ khòθá/ (1SG - to.eat - mango) '*I* ate a mango'. When the subject of a sentence is third person singular, the emphatic form /ʔəwê/ is usually used instead of the form I /ʔə/, e.g., /**ʔəwê** ʔáɴ/ (3SG - to.eat) 'He ate'.

Table 7: EPK pronouns.

	Form I	Form II	Emphatic form
1SG	*jə*	*jə̀*	*jəwê, jəwêdá*
1PL	*hə (pə)*	*hə̀ (pə̀)*	*həwê (pəwê), həwêdá (pəwêdá)*
2SG	*nə*	*nə̀*	*nəwê, nəwêdá*
2PL	*nəθí*	*nəθí*	*nəθíwê, nəθíwêdá*
3SG	*ʔə*	*ʔə̀*	*ʔəwê, ʔəwêdá*
3PL	*ʔəθí*	*ʔəθí*	*ʔəθíwê, ʔəθíwêdá*

In many Karenic languages, pronouns placed before verbs and nouns are phonologically weak, as is the case with the EPK pronouns /jə/ (1SG), /nə/ (2SG), and /ʔə/ (3SG), which have atonic syllables. In EPK, a topicalized pronoun in the form II sometimes appears together with the same pronoun in form I in the subject slot in a sentence, such as in (25).

(25) **jə̀**   **jə**   *mə*   *lɔ́*   *nə̀*   [EPK]
    1SG[TOP]   1SG[SBJ]   IRR   to.tell   2SG[OBJ]
    'As for me, I will tell you.'

The weak forms /jə/, /nə/, and /ʔə/ can be interpreted as proclitics, and these forms used before verbs could be taken as an example of the pronominalization that is widely observed in Tibeto-Burman languages (for pronominalization, see, e.g., Nishi 1995). LaPolla (1994: 74) discusses a similar phenomenon in SK in the context of pronominalization.

18.7 Noun phrases

The structure of the Pwo Karen noun phrase can be schematized as in Figure 2. This is a revised version of the schema shown by Kato (2019: 144). Brackets denote optional items.

Slot 1	Slot 2	Slot 3	Slot 4	Slot 5	Slot 6	Slot 7	Slot 8	Slot 9
(RC)	(NM)	(PRON)	HEAD NOUN	(RC)	(ADP)	(NUM + NC)	(PL)	(DEM)

Figure 2: The order of components within the EPK noun phrase

Slot 1 is for a pre-head relative clause. Slot 2 is for a noun modifier, a noun that modifies the head noun. Slot 3 is for a pronoun, which denotes a possessor. Slot 4 is for the head noun. Slot 5 is for a post-head relative clause. Slot 6 is for an adpositional phrase. Slot 7 is for "numeral + numeral classifier". Slot 8 is for a particle that denotes plurality. Slot 9 is for a demonstrative, more precisely, a particle that has a demonstrative function. Below are examples:

(26)   Slot 2 | Slot 3 | Slot 4   [EPK]
       phā | *(ʔə)* | *chái*
       father | 3SG | rice.field
       'Father's rice field'

(27)   Slot 3 | Slot 4 | Slot 6 | Slot 8 | Slot 9   [EPK]
       jə | *θò* | *dē khúláʋ* | *θὲ* | *nɔ́*
       1SG | friend | with hat | PL | that
       'those friends of mine with hats'

13

(28)

Slot 1	Slot 4	Slot 5	Slot 6			Slot 7	Slot 8	Slot 9	[EPK]
ʔəwêxwè	já	phàdʊ́	lə́	cəpwɛ̄ phânkhʊ́	ʔò	θə̄n béiɴ	θɛ̀	nɔ́	
3SG to.buy	fish	be.big	LOC	table above	away	three NC	PL	that	

'those three big fish on that table, which he bought'

When a possessor noun modifies the head noun as in (26), the third-person pronoun may appear before the possessed noun. /ʔə chái/ (3SG - rice.field) alone can mean 'his rice field'.

Slot 7 in Figure 2 is the slot for "numeral (NUM) + numeral classifier (NC)", which I call an "NC phrase" here. When a noun is modified by a numeral, the numeral must be followed by a numeral classifier corresponding to the noun. Examples of head nouns with an NC phrase are: /thwí lə dùʊ/ (dog - one - NC[animal]) 'one dog', /phlòʊɴ nī ɣà/ (person - two - NC[person]) 'two people', /châiɴ θə̄ɴ béiɴ/ (shirt - three - NC[flat thing]) 'three pieces of shirt', /khòθá lī phlóʊɴ/ (mango - four - NC[round thing]) 'four mangos', and /lé jɛ̄ bòɴ/ (stick - five - NC[long thing]) 'five sticks'. /dùʊ/ is usually used for mammals. Birds, fish, and insects are counted with /béiɴ/ (flat thing), and snakes and lizards are counted with /bòɴ/ (long thing). EPK numeral classifiers are a type of dependent nouns. When they occur in Slot 7, they need to be preceded by a numeral. Thus, /thwí lə dùʊ jò/ (dog - one - NC[animal] - this) 'this one dog' and the zero head /lə dùʊ jò/ (one - NC[animal] - this) 'this one (animal)' are fine but */thwí dùʊ jò/ (dog - NC[animal] - this) and */dùʊ jò/ (NC[animal] - this) are ungrammatical. It would be worth noting that an NC phrase can be floated into the position of the "adverbial elements" in Figure 1, if the entire noun phrase is the subject or object. Thus, /təwâphjā nī ɣà ɣɛ̂ lə́ jò/ (student - two - NC[person] - to.come - LOC - here) 'Two students came here' can be changed to /təwâphjā ɣɛ̂ lə́ jò nī ɣà/ (student - to.come - LOC - here - two - NC[person]).

According to Manson's (2017: 160) generalization about the structure of the noun phrase in Karenic languages, the NC phrase follows the demonstratives. He cites Naw Hsa Eh Ywar's (2013: 62) Kayan Lahta example:

(29) ɲa˧ ʃa˧ pi˥ do˩ **shu˩ ba˩** [Kayan Lahta]
1SG chicken be.small that six CLF
'those six small chickens of mine'

In EPK, however, the NC phrase is placed before the demonstrative, as in (28). A simpler EPK example than (28) would be: /já phàdʊ́ **nī béiɴ** jò/ (fish - be.big - two - NC[flat thing] - this). The order in SK is the same as EPK: /ɲâ pháʔdô **khí bê** ʔī̄/ (fish - be.big - two - NC[flat thing] - this). In this way, the order of elements in the noun phrase varies from one Karenic language to another.

Relative clauses are also formed in various ways in different Karenic languages. In EPK, there are two ways of forming a relative clause: the first uses the relative marker /lə́/ and the second does not. We will first observe the latter case, which is the more colloquial. In this case, when the relativized noun is the subject of the relative clause, the relative clause is placed in Slot 5. In the relative clause, the subject noun is gapped:

(30) phlòʊɴ [ɣɛ̂ lə́ phlòʊɴ thîkhā̄ɴ] [EPK]
person to.come LOC Karen country
'the person who came to Kayin State'

The same phonological string as (30) can be used as a full sentence that means 'People (a person) came to Kayin State'. When the relativized noun is a non-subject noun, the relative clause is placed in Slot 1. The relativized non-subject noun is gapped in the relative clause:

(31) [jə   tháʊ   lə́   dàʊ   phèɴ]   khánphài   nɔ́   [EPK]
     1SG   to.ride   LOC   room   inside   shoes   that
     'those shoes which I wear in the room'

In the case in which the relative marker /lə́/ is used, relative clauses are placed in Slot 5 regardless of the syntactic role of the relativized noun. The resumptive pronoun corresponding to the head noun occurs in relative clauses with /lə́/, as seen in (32) and (33). (When the relativized noun is a non-subject noun, the resumptive pronoun occurs only if the noun is an animate noun.)

(32) phlòʊɴ   [lə́   ʔə   γɛ̂   lə́   phlòʊɴ   thîkhāɴ]   [EPK]
     person   REL   3SG   to.come LOC   Karen   country
     'the person who came to Kayin State'

(33) phlòʊɴ   [lə́   jə   dʊ́   ʔə̀]   nɔ́   [EPK]
     person   REL   1SG   to.hit 3SG   that
     'that person whom I hit'

SK only has the relative clauses with the relative marker /lə́/, and its usage is very similar to that of EPK:

(34) pɣākəɲɔ́   [lə́   ʔə   hɛ́   lə́   pɣākəɲɔ́   kɔ̀]   [SK]
     person   REL   3SG   to.come LOC   Karen   country
     'the person who came to Kayin State'

(35) pɣākəɲɔ́   [lə́   jə   tɔ̀   ʔɔ̄]   nê   [SK]
     person   REL   1SG   to.hit 3SG   that
     'that person whom I hit'

Kayan has internally headed relative clauses as well as externally-headed relative clauses that are formed with relative markers. Manson (2010: 325) argues that the noun meaning 'woman' in (36) is the head of the internally headed relative clause that is enclosed in square brackets:

(36) [prə̀mû   wân-bâ   písāpʰò   sʰɨ́   nṳ̀   prɛ̀]   ŋá   mè   mwē   kʰí   pʰò   [Kayan]
     woman   to.wash-BEN   child   water   that   CLF   there TOP   to.be   1SG   child
     'The woman over there washing the baby is my daughter.'

18.8 Negation

In EPK, when the main clause is negated, the adverbial particle /ʔé/ is placed in the clause-final position:

(37) ʔəwê   ʔáɴ   bá   mḭ̀   dài   ʔé   [EPK]
     3SG   to.eat OPP   rice   still   NEG
     'He has not managed to eat rice yet.'

When a subordinate clause is negated, the verb particle /lə/, which originated from the Proto-Karen negative marker *ta (Manson 2017a: 157)[6] and is probably related to the Proto-Tibeto-

---

[6] Forms corresponding to EPK /lə/ in many other Karenic languages still preserve the onset of the Proto-Karen negative marker *ta, e.g., SK /tə/ and Geba /tV/. The Proto-Karen onset *t became /l/ in EPK in two morphemes: /lə/ 'negative marker' and the numeral /lôɴ ~ lə/ 'one' (see Matisoff's [2003: 262] Proto-Tibeto-

Burman negative imperative marker *ta (Benedict 1972: 97; Matisoff 2003: 162), is placed immediately before the verb; simultaneously, the verb particle /bá/, whose origin is unknown, is placed in the final position of the subordinate clause, as in (38). That is, "double negation" (Dryer 2005) is employed in a subordinate clause.

(38) ʔəwê **lə** ɣɛ̂ lɛ́ jò **bá** ʔəkhú̀còn, jə bá mà [EPK]
3SG NEG to.come LOC here NEG because 1SG must to.do
'Because he did not come here, I have to do.'

In (38), /bá/ may also be placed immediately after the verb, i.e. /ʔəwê **lə** ɣɛ̂ **bá** lɛ́ jò ʔəkhú̀còn .../. Sometimes, this negative form /lə V bá/ is used in the main clause, as seen in (39). In this case, the sentence typically presupposes that the listener wants to know the reason for something and the sentence shows the reason. Thus, (39) can be translated into English as 'it is that he doesn't remember me anymore', or 'it is because he doesn't remember me anymore'. In the case that /lə V bá/ is used in the main clause, /bá/ can be omitted.

(39) ʔəwê **lə** θíjâ lɛ̀n **(bá)** [EPK]
3SG NEG to.know anymore NEG
'(It is that) he doesn't remember (me) anymore.'

Manson (2017b) concisely summarizes the patterns of negation in Karenic languages. He groups them into five types: (I) the negative marker is placed immediately before the verb, (II) the negative marker is placed immediately before the verb and a second marker is placed immediately after the verb, (III) the negative marker is placed immediately before the verb and a second marker is placed in the clause-final position, (IV) the negative marker is placed immediately after the verb, and (V) the negative marker is placed in the clause-final position. Manson assumes that Type (I) is the original pattern. (37) is an example of (V), (38) is an example of (III), the case in which /bá/ in (38) is placed immediately after the verb is an example of (II), and (39) without /bá/ is an example of (I). Thus, in Pwo Karen, the only Type (IV) is not observed. A Pa-O example of Type (IV) from Cooper (2018: 29), which is not observable in EPK, is presented in (40). According to Manson, Type (I) is observed in Kayan, Lahta, Gekho, and Paku, (II) in Sgaw, (III) in Bwe, Geba, and Sgaw, (IV) in Pa-O, and (V) in Monu (Manu), Kayaw, Kayah, and Palaychi.

(40) khwè phrɛ́ lɛ̀n phé bá **tâw** na mɔ́k.cɔ́k [Pa-O]
1SG to.buy to.come to.give to.hit NEG 2SG orange
'I didn't buy you oranges.'

18.9 Interrogative sentences

In Karenic languages, polar questions (yes-no questions) are generally indicated by placing a question marker in the clause-final position. In EPK, the polar question marker /ʁâ/, a kind of particle, is used. Below is an example:

(41) nə mə ɣɛ̂ ʁâ [EPK]
2SG IRR to.come Q
'Will you come?'

In SK, /ɦá/ is used for making a polar question, such as in /nə kə hɛ́ ɦá/ (2SG - IRR - to.come -

---

Burman form *tan 'one'). The corresponding Western Pwo Karen forms are /lə/ 'negative marker' and /kə/ 'one'.

Q) 'Will you come?'. In Kayah Li, the polar question marker is /ɛ̄/ (Solnit 1997: 233), and in Kayan, it is /yá/ (Manson 2010: 271). All of these forms are probably cognates.

In the case of a content question, EPK uses the content question marker /lɛ̂/ in the clause-final position, such as in the example below:

(42) *nə   mə   ʔáN   chənɔ́   lɛ̂*   [EPK]
       2SG IRR to.eat what   Q
       'What will you eat?'

In this sentence, /lɛ̂/ is not omittable. In SK, the form /lɛ̂/, which is also not omittable, is used in a content interrogative sentence, as in /nə kə ʔɔ̂ mənɯ̄ lɛ̂/ (2SG - IRR - to.eat - what - Q) 'What will you eat?'. In Kayah, /tɛ̄/ is generally used (Solnit 1997: 244). In Kayan, /lé/ is used, but it is omittable. For example, /tará̰ (lé)/ (what - Q) 'What?' (Manson 2010: 276). All these forms are probably cognates and probably also cognate with the Burmese content question marker /lɛ́/.

18.10 Verb serialization

EPK has two types of serial verb constructions: the concatenated type and separated type. Here, we will discuss serial verb constructions that contains minimum of two verbs, represented as V1 (the first one) and V2 (the second one). The concatenated type does not allow other elements including a noun phrase to occur between the two verbs, while the separated type allows. The concatenated type and separated type in Pwo Karen respectively correspond to Aikhenvald and Dixon's (2006) "contiguous serial verb construction" and "non-contiguous serial verb construction". Using the terms defined by Role and Reference Grammar (see Van Valin and LaPolla 1997), the concatenated type corresponds to the "nuclear juncture", and the separated type to "core juncture". (43) and (44) below are examples of the concatenated type:

(43) *jə   **xwè**   **ʔáN**   kʊ́*   [EPK]
       1SG to.buy   to.eat   cake
       'I bought and ate cake.'

(44) *jə   **dʊ́**   **θî**   thò*   [EPK]
       1SG to.hit   to.die   pig
       'I hit the pig intending to kill it.'

When the combinations of verbs are Vi+Vi, Vi+Vt, and Vt+Vt, the subject arguments of them are coreferential, and both verbs are volitional, as in (43). However, in the case of Vt+Vi, the object argument of V1 and the subject argument of V2 is coreferential, and V1 and V2 must be volitional and non-volitional respectively, as in (44). The serialization of Vt+Vi type has a causative meaning.

In the separated type, V2 denotes a result of V1, as in (45) and (46), or denotes potentiality including possibility/ability/permission, as in (47). In the separated type, there is no restriction on the combination of shared arguments, but V2 must be non-volitional.

(45) *jə   **ʔáN**   mì̠   **blɛ̀**   jàʊ*   [EPK]
       1SG to.eat   rice be.full   PFV
       'I ate rice and got full.'

(46) *jə   **dʊ́**   thò   **θî**   mèiN*   [EPK]
       1SG to.hit   pig   to.die   naturally
       'When I hit the pig, it died.'

(47) jə **nâN** kā **θɪ́**   [EPK]
      1SG to.drive car be.capable
      'I can drive a car.'

The semantic difference between (44) and (46) is noteworthy. In (44), the death of the pig was intended by the agent of V1, whereas it was not intended but happened accidentally in (46). Such a difference between the two types is also observed in Kayan (Manson 2010: 301–302).

Solnit (1997: 56–57) points out that Karenic languages show a preference for immediate concatenation of verbs. This is also true of Pwo Karen. In the serialization /**xwè ʔáN** kʉ́/ of (43), if the noun /kʉ́/ is put between the verbs, it will be ungrammatical */**xwè** kʉ́ **ʔáN**/ (to.buy - cake - to.eat). In Mainland Southeast Asian languages that are of SVO-type, this type of serialization is often fine, such as in Thai /**sɯ́ɯ** khənǒm **kiN**/ (to.buy - cake - to.eat) 'buy cake and eat it'. Solnit states that Kayah Li's preference for immediate concatenation is high even in comparison to what is known of the syntax of other Karenic languages, such as Sgaw and Pa-O. Kayah Li lacks the separate type serialization. According to my research, Geba also lacks the separated type serialization. In Geba, ability, which is expressed by using the separate type in EPK as in (47) is expressed with concatenation (Kato 2008a: 213):

(48) jɔ **ɗɔ́** **zā** gadā ʔalē   [Geba]
      1SG to.speak be.capable Burma language
      'I can speak Burmese.'

In the concatenated type serialization in EPK, the order of verbs follows the chronological order of the events. In (43), for example, the verb /xwè/ (V1) 'to buy' precedes /ʔáN/ (V2) 'to eat' because the actions of buying and eating happen in this chronological order. However, this principle may be broken when one of the serialized verbs is a motion verb. Typical motion verbs are /lɪ̀/ 'to go' and /ɣɛ̂/ 'to come'. EPK has a verb order rule that requires a motion verb to occur before a non-motion verb (Kato 2004: 222–228). For example, if the verbs /ɣɛ̂/ 'to come' and /xwè/ 'to buy' are concatenated, they must be arranged in the order of /ɣɛ̂ xwè/ whatever the chronological order may be and */xwè ɣɛ̂/ is ungrammatical. Thus, /jə ɣɛ̂ xwè já/ (1SG - to.come - to.buy - fish) has two readings: (a) 'I came to buy a fish', and (b) 'I bought a fish and came with it'. In the case of (b), the serialization does not follow the chronological order. SK has the same rule, and the sentence /jə hɛ́ pɣē ɲâ/ (1SG - to.come - to.buy - fish) has the same two readings. However, Geba does not seem to have such a rule about motion verbs because the following example with a motion verb occurring after a non-motion verb is found in my Geba data: /zɛ swē lē/ (1SG - to.run - to.go) 'I went running'. To express the same meaning, EPK and SK have to use the sentences /jə lɪ̀ klí/ (1SG - to.go - to.run) and /jə lɛ̄ xè/ (1SG - to.go - to.run) respectively, in both of which a motion verb occurs before a non-motion verb.

In EPK, some verbs in concatenated serial verb constructions have changed their meanings into more abstract ones and are thus more flexible in terms of co-occurrence with other verbs. These can be called "versatile verbs", following Matisoff's (1969) terminology. Kato (2004) treats them as verb particles because each has been grammaticalized in certain points. Aikhenvald and Dixon's (2006) "asymmetrical serial verb constructions" are serial verb constructions that contain this type of verbs. In the following example, the verb particle /jʉ̄/, originating from the verb /jʉ̄/ meaning 'to look at', expresses the meaning 'to try to V'.

(49) jə **ʔáN** **jʉ̄** phlòʊN chəʔáNchəʔɔ̀   [EPK]
      1sg to.eat to.look Karen   food
      'I tried to eat Karen food.'

For more details on serial verb constructions in EPK, see Kato (1998, 2004: 207–275). Kato (2017) and Kato (2019a) also discuss them to some extent. Descriptive studies on serial verb constructions in other Karenic languages include Kato (1992) and Weinhold

(2011; in Brunelle's [2011] Sgaw Karen papers) on Sgaw Karen, Solnit (1997: Chapter 4) and Sonit (2006) on Kayah, Manson (2010: 287–302) on Kayan, Swanson (2011) on Bwe, and Cooper (2017) on Pa-O. Kato (in print) compares EPK and SK in terms of word order in serial verb constructions.

18.11 Clause linkage

In EPK, clauses can be embedded as arguments without any special marking. The following example (50) is interpreted as a construction in which the part /ʔəwê ʔán mì/ is a complement clause embedded as the subject of /phlɛ́/, because the part /ʔəwê ʔán mì/ has some features of nounhood, including the fact that it can be clefted, such as /chə phlɛ́ nɔ́ mwē ʔəwê ʔán mì/ (thing - be.fast - TOP - COP - 3SG - to.eat - rice) 'Literal translation: It is his eating that is fast'. (50) can more properly be translated as 'his eating rice is fast', if we make much of the syntactic construction of this sentence. An example of a clause embedded as the object is (51).

(50) [ʔəwê ʔán   mì]   phlɛ́   [EPK]
      3SG   to.eat rice   be.fast
      'He eats (rice) fast.'

(51) jə   dá   [ʔəwê ʔán   mì]   [EPK]
      1SG   to.see 3SG   to.eat rice
      'I saw that he was eating (rice).'

Interestingly, in Kayah Li, the same meaning as in (50) is expressed by using verb concatenation. (52) is an example from Solnit (1997: 91). Kayah Li's strong preference for immediate concatenation is observed in this point as well:

(52) ʔa ʔe   phrɛ̄   dī   [Kayah Li]
      3   to.eat be.fast rice
      'He eats (rice) fast.'

Solnit says that syntactically there is no embedding in the verb concatenation of (52). Geba also uses verb concatenation to express situations similar to (52), such as /ja ʔā plá dĭ/ (1SG - to.eat - be.fast - rice) 'I eat rice fast' (Kato 2008a: 174). Kayan also seems to have the same characteristic; see example (53) from Manson (2010: 151).

(53) phón   phrái dyán ká   [Kayan]
      to.cook be.fast rice   IMP
      'Quickly cook the rice.'

Adverbial clauses in EPK are usually placed before the main clause. They are introduced in the sentence by using subordinate clause particles whose occurring positions differ from each other. Examples are: /ʔəwê ʔè lì, jə lì ʔé/ (3SG - if - to.go | 1SG - to.go - NEG) 'If he goes, I won't go'; /ʔəwê lì ʔəkhúcòN, jə lì/ (3SG - to.go - because | 1SG - to.go) 'Because he went, I went'; /ʔəwê lì lānâN, jə lì ʔé/ (3SG - to.go - although | 1SG - to.go - NEG) 'Although he went, I didn't go'; and /kəlā ʔəwê lə lì dài bá, jə mə lì/ (while - 3SG - NEG - to.go - still - NEG | 1SG - IRR - to.go) 'Before he goes, I will go'.

18.12 Voice

As Manson (2017: 159) states, some Karenic languages have a passive construction but its use is infrequent. SK is one of the Karenic languages that have passive usages. It uses the verb /bâ/ 'to hit' to make a passive clause. (54) is an example cited from Kan Gyi (1915: 19).

This example corresponds to the active clause /jə pà thú jā lɔ̄/ (1SG - father - to.kick - 1SG - EMP) 'My father kicked me'. This construction is a very formal expression and rarely used in a daily conversation.

(54)   *jə*   *bâ*   *tà*   *thú*   *jā*   *lə́*   *jə*   *pà*   *lɔ̄*   [SK]
       1SG   to.hit   NMLZ   to.kick   1SG   LOC   1SG   father   EMP
       'I was kicked by my father.'

EPK does not have a passive construction. However, agent-defocusing effect (see Myhill 1997), which is a significant functional role of the passive voice in many languages, is fulfilled by the noun meaning 'thing' that occurs in the subject slot. (55) is an example. SK also has the same construction, e.g. /tà tò jā/ (thing - to.hit - 1SG) 'I was hit (by someone).'

(55)   *chə*   *dʊ́*   *jə̀*   [EPK]
       thing   to.hit   1SG
       'I was hit (by someone).' (Literally: 'A thing hit me.')

As Kato (1999, 2009a) describes, EPK has several causative markers, which belong to the category of verb particles. Out of them, /dàʊ/ (also pronounced /dà/ and /dài/) is a "genuine" causative marker because it has no corresponding homophonous verb in EPK. In a clause with /dàʊ/, the causee occurs as NP2 (see Figure 1) and the object argument of the verb occurs as NP3, as seen in (56):

(56)   *jə*   *dàʊ*   *ʔáɴ*   *ʔəwê*   *kʊ́*   [EPK]
       1SG   CAUS   to.eat   3SG   cake
       'I let him eat cake.'

/dàʊ/ is cognate with the Kayah Li verb /dʌ́/ 'to let; to give' (Solnit 1997: 65), which can be used as a causative element in serialized verbs. /dàʊ/ is also cognate with the Kayan causative prefix /dè-/, which is derived from the verb /dè/ 'order' (Manson 2010: 262–263). It is also cognate with the SK causative marker /dɯ́ʔ/. Thus, its original form at the Proto-Karen stage might have already had a causative use. Jenny (2015: 170) assumes that Kayah Li /dʌ́/ and SK causative marker /dɯ́ʔ/ originated from Proto-Tibeto-Burman **ter**/**s-ter** 'give, CAUSATIVE' (Matisoff 2003: 399, 615).

It would be worth noting that Karenic languages have applicative constructions. Kato (2009a), in discussing valence-changing verb particles in EPK, pointed out that EPK has several applicative markers including the benefactive applicative, comitative applicative, prioritive applicative, assistive applicative, and substitutive applicative. See (57) below, an example of comitative applicative:

(57)   *jə*   [*ʔáɴ*   *ɣə̀ɴ*]   *ʔəwê*   *mì̠*   [EPK]
       1SG   to.eat   APPL   3SG   rice
       'I ate rice with him.'

The bracketed part is a verb complex. The verb particle /ɣə̀ɴ/ is placed after the verb and the applied noun appears as NP2. A similar meaning can be expressed by using the adposition /dē/ 'with' as in /jə ʔáɴ mì̠ dē ʔəwê/ (1SG - to.eat - rice - with - 3SG). However, in (57), in terms of semantics, the referent of the applied noun is more actively engaged in the event denoted by the sentence. Such a meaning is also observed in applicative constructions of other languages (Shibatani 2006: 244). (58) is an example of comitative applicative in Kayah Li (Solnit 1997: 104), and (59) and (60) are examples of instrumental applicative in Geba (Kato 2008b: 174) and Palaychi (the author's field data). Palaychi is one of the languages of the Mopwa group (for more details on Mopwa, see Naw Veronica 2011).

(58) *cwá   kʌ̄   vɛ*   [Kayah Li]
    to.go APPL 1SG
    'Go with me.'

(59) *ja   ʔā   ʔī   zwī*   [Geba]
    1SG   to.eat APPL spoon
    'I ate with a spoon.'

(60) *zà   ʔ̀ɔ   zr̀ʔ   jō*   [Palaychi]
    1SG   to.eat APPL spoon
    'I ate with a spoon.'

Furthermore, EPK has a middle marker, /θà/, which originated from the noun meaning 'heart', as is discussed in detail by Kato (2019b). (61) and (62) are examples of the marker that is used for an anticausative construction and reflexive construction, respectively. SK also has a cognate middle marker /θáʔ/, whose uses are quite similar to EPK /θà/.

(61) *ʔəwê   ʔánlɛ̀   θà   jàʊ*   [EPK]
    3SG   to.change(tr.) MID   PFV
    'He has changed.'

(62) *ʔəwê   chè   làɴ   θà*   [EPK]
    3SG   to.stab down MID
    'He stabbed himself.'

18.13 Pragmatics

Karenic languages use topicalization frequently. Topicalization in EPK is a left-dislocation of an element and the element is often followed by a topic marker. EPK has several topic markers. Mostly /nɔ́/ is used. In the sentence /ɕáphàɴ dʊ́ θàkhléiɴ/ (Shapan - to.hit - Thakhlein) 'Shapan hit Thakhlein' (=22), the subject can be topicalized, such as /ɕáphàɴ nɔ́ dʊ́ θàkhléiɴ/ (Shapan - TOP - to.hit - Thakhlein) 'As for Shapan, he hit Thakhlein', and the object can also be topicalized, such as /θàkhléiɴ nɔ́ ɕáphàɴ dʊ́/ (Thakhlein - TOP - Shapan - to.hit) 'As for Thakhlein, Shapan hit him'. Topicalization can be applied to various syntactic elements including the subject noun, object noun, adpositional phrase, and complement clause.

In EPK, ellipses of arguments that are recoverable from the discourse context are quite frequent. For example, when one is asked with the sentence /nə mə lḭ̀ ʁâ/ (2SG - IRR - to.go - Q) 'Will you go?', the answer can be either with or without the subject, /(jə) mə lḭ̀/ (1SG - IRR - to.go) 'Yes, (I) will go'. However, the frequency of ellipsis can vary from language to language within the Karenic branch. I have the feeling that in SK, subject pronouns are retained more often than EPK in daily conversation, but it needs further detailed investigation.

Abbreviations

1,2,3	first, second, and third person	COP	copular verb
ADP	adpositional phrase	DEM	demonstrative
APPL	applicative	EMP	emphatic
BEN	benefactive	EPK	Eastern Pwo Karen
CAUS	causative	IMP	imperative
CLF	classifier	IRR	irrealis
CONT	continuous	LOC	location particle

MID	middle marker		PRON	pronoun
NC	numeral classifier		Q	question
NEG	negation		RC	relative clause
NM	noun modifier		RDP	reduplicated syllable
NMLZ	nominalizer		REL	relative marker
NP	noun phrase		SBJ	subject
NUM	numeral		SG	singular
OBJ	object		SK	Sgaw Karen
OPP	opportunity		TOP	topic marker
PFV	perfective		V	verb
PK	Pwo Karen		Vi	intransitive verb
PL	plural		Vt	transitive verb
PN	personal noun			

References

Aikhenvald, Alexandra Y. & R. M. W. Dixon (eds.). 2006. *Serial verb constructions: A cross-linguistic typology*. Oxford: Oxford University Press.

Benedict, Paul K. 1972. *Sino-Tibetan: A conspectus*. Cambridge: Cambridge University Press.

Brunelle, Marc (ed.). 2011. *Sgaw Karen papers, presented to Nimrod Andrew*. Ottawa: University of Ottawa.

Cooper, Alys Boote. 2018. Secondary verbs in Pa-O: A preliminary study. In Pittayawat Pittayaporn et al. (eds.), *Papers from the Chulalongkorn International Student Symposium on Southeast Asian linguistics 2017*, 21–31. Honolulu: University of Hawai'i Press.

Dawkins, Erin & Audra Phillips. 2009a. *A sociolinguistic survey of Pwo Karen in Northern Thailand*. Chiang Mai: Linguistic Department, Payap University.

Dawkins, Erin & Audra Phillips. 2009b. *An investigation of intelligibility between West-Central Thailand Pwo Karen and Northern Pwo Karen*. Chiang Mai: Linguistic Department, Payap University.

Dryer, Matthew S. 2005. Negative morphemes. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*, 454–457. Oxford: Oxford University Press.

Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2019. *Ethnologue: Languages of the world. Twenty-second edition*. Dallas, TX: SIL International.

Haudricourt, André-Georges. 1946. Restitution du karen commun. *Bulletin de la Société de Linguistique de Paris* 42(1). 103–111 [Reprinted: Haudricourt 1972, 131–140].

Haudricourt, André-Georges. 1953. A propos de la restitution du karen commun. *Bulletin de la Société de Linguistique de Paris* 49(1). 129–132. (Reprinted: Haudricourt 1972, 141–145.)

Haudricourt, André-Georges. 1972. *Problèmes de phonologie diachronique.* Paris: SELAF.

Haudricourt, André-Georges. 1975. Le système des tons du karen commun. *Bulletin de la Société de Linguistique de Paris* 70(1). 339–343.

Henderson, Eugénie. J. A. 1961. Tone and intonation in Western Bwe Karen. *Burma Research Society Fiftieth Anniversary Publication* 1. 59–69.

Henderson, Eugénie. J. A. 1979. Bwe Karen as a two-tone language? *Pacific Linguistics,* Series C, 45. 301–326.

Henderson, Eugénie. J. A. 1997. *Bwe Karen dictionary: With texts and English-Karen word list*, 2 vols. London: School of Oriental and African Studies University of London.

Jenny, Mathias. 2015. The far west of Southeast Asia: "Give" and "get" in the languages of Myanmar. In N. J. Enfield & Bernard Comrie (eds.), *Languages of Mainland Southeast Asia*, 156–208. Berlin & Boston: Mouton de Gruyter.

Jones, Robert B. 1961. *Karen linguistic studies*. Berkeley & Los Angeles: University of California Press.

Kan Gyi. 1915. *Introduction to the study of Sgaw Karen*. Rangoon: American Baptist Mission Press.

Kato, Atsuhiko [加藤昌彦]. 1992. Verb serialization in Sgaw Karen [スゴー・カレン語の動詞連続]. *Journal of Asian and African Studies* [アジア・アフリカ言語文化研究] 45. 177–204. (In Japanese)

Kato, Atsuhiko. 1995. The phonological systems of three Pwo Karen dialects. *Linguistics of the Tibeto-Burman Area* 18(1). 63–103.

Kato, Atsuhiko [加藤昌彦]. 1998. On head verbs of serial verb constructions in Pwo Karen [ポー・カレン語の動詞連続における主動詞について]. *Journal of the Linguistic Society of Japan* [言語研究] 113. 31–61. (In Japanese)

Kato, Atsuhiko. 1999. Two types of causative construction in Pwo Karen. In Tadahiko L. A. Shintani (ed.), *Linguistic & anthropological study on the Shan culture area*, 55–93. Tokyo: Research Institute for Languages and Cultures of Asia and Africa.

Kato, Atsuhiko [加藤昌彦]. 2001a. Buddhist Pwo Karen script [仏教ポー・カレン文字]. In *The Sanseido encyclopaedia of linguistics* [言語学大辞典], vol. 7, 847–851. Tokyo: Sanseido [三省堂]. (In Japanese)

Kato, Atsuhiko [加藤昌彦]. 2001b. Christian Pwo Karen script [キリスト教ポー・カレン文字]. In *The Sanseido encyclopaedia of linguistics* [言語学大辞典], vol. 7, 333–337. Tokyo: Sanseido [三省堂]. (In Japanese)

Kato, Atsuhiko [加藤昌彦]. 2002. A contrastive basic vocabulary of Eastern and Western Pwo Karen in Myanmar [ビルマにおける東部および西部ポー・カレン語の対照基礎語彙]. *Southeast Asian Studies, Tokyo University of Foreign Studies* [東京外大　東南アジア学] 7. 212–249. (In Japanese)

Kato, Atsuhiko [加藤昌彦]. 2004. A grammar of Pwo Karen [ポー・カレン語文法]. Tokyo: University of Tokyo PhD dissertation. (In Japanese)

Kato, Atsuhiko [加藤昌彦]. 2006. Difference in prevalence of writing systems in the same language: A case of Pwo Karen [同一言語内における文字普及状況の差異について: ポー・カレン語の事例]. In Asako Shiohara [塩原朝子] & Shigeaki Kodama [児玉茂昭] (eds.), *Writing unwritten languages* [表記の習慣のない言語の表記], 89–110. Tokyo: Research Institute for Languages and Cultures of Asia and Africa. (In Japanese)

Kato, Atsuhiko [加藤昌彦]. 2008a. Basic materials in Geba [ゲーバー語基礎資料]. *Asian and African Languages and Linguistics* [アジア・アフリカの言語と言語学] 3. 169–219. (In Japanese)

Kato, Atsuhiko [加藤昌彦]. 2008b. Is the category "adjective" necessary in Pwo Karen? *Asian and African Languages and Linguistics* [アジア・アフリカの言語と言語学] 3. 77–95. (In Japanese)

Kato, Atsuhiko. 2009a. Valence-changing particles in Pwo Karen. *Linguistics of the Tibeto-Burman Area* 32(2). 71–102.

Kato, Atsuhiko. 2009b. A basic vocabulary of Htoklibang Pwo Karen with Hpa-an, Kyonbyaw, and Proto-Pwo Karen forms. *Asian and African Languages and Linguistics* 4.169–218.

Kato, Atsuhiko [加藤昌彦]. 2016. Ethnic groups and languages in Myanmar [ミャンマーの諸民族と諸言語]. *ICD NEWS* [法務省法務総合研究所国際協力部報] 69. 8–26. (In Japanese)

Kato, Atsuhiko. 2017. Pwo Karen. In Graham Thurgood & Randy LaPolla (eds.), *The Sino-Tibetan languages, 2nd edition*, 942–958. London & New York: Routledge.

Kato, Atsuhiko. 2018. How did Haudricourt reconstruct Proto-Karen tones? *Reports of the Keio Institute of Cultural and Linguistic Studies* 49. 21–44.

Kato, Atsuhiko. 2019a. Pwo Karen. In Alice Vittrant & Justin Watkins (eds.), *The Mainland Southeast Asia linguistic area*, 131–175. Berlin & Boston: Mouton de Gruyter.

Kato, Atsuhiko. 2019b. The middle marker in Pwo Karen. *Reports of the Keio Institute of Cultural and Linguistic Studies* 50. 21–62.

Kato, Atsuhiko. In print. Karen and surrounding languages. In Hayashi Norihiko (ed.). *Topics in Middle Mekong linguistics*.

LaPolla, Randy J. 1994. Parallel grammaticalizations in Tibeto-Burman languages: Evidence of Sapir's "drift". *Linguistics of the Tibeto-Burman Area* 17 (1). 61–80.

Li, Fang-Kuei. 1977. *A handbook of comparative Tai*. Honolulu: The University Press of Hawaii.

Luangthongkum, Theraphan. 2012. Proto-Karen (*k-rjaŋᴬ) fauna. Paper presented at SEALS 22, Agay.

Luangthongkum, Theraphan. 2019. A view on Proto-Karen phonology and lexicon. *Journal of the Southeast Asian Linguistics Society* 12(1). i–lii.

Luce, Gordon. H. 1959. Introduction to the comparative study of Karen languages. *Journal of Burma Research Society* 42(1). 1–18.

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

Manson, Ken. 2007. *Pekon Kayan phonology*. Chiang Mai: Department of Linguistics, Payap University.

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

Manson, Ken. 2010. *A grammar of Kayan, a Tibeto-Burman language*. Bundoora: La Trobe University PhD dissertation.

Manson, Ken. 2017a. The characteristics of the Karen branch of Tibeto-Burman. In Picus Sizhi Ding & Jamin Pelkey (eds.), *Sociohistorical linguistics in Southeast Asia: New horizons for Tibeto-Burman studies in honor of David Bradley*, 149–168. Leiden & Boston: Brill.

Manson, Ken. 2017b. From right to wrong: Negation in the Karen languages. Paper read at the meeting of Australian Linguistic Society.

Manson, Ken. 2019. Proto-Kayan. ms.

Matisoff, James A. 1969. Verb concatenation in Lahu: The syntax and semantics of "simple" juxtaposition. *Acta Linguistica Hafniensia* XII(1). 69–120.

Matisoff, James A. 1973. Tonogenesis in Southeast Asia. In Larry M. Hyman (ed.), *Consonant types and tone* (Southern California Occasional Papers in Linguistics 1), 71–96. Los Angeles: University of California Press.

Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction*. Berkeley, Los Angeles & London: University of California Press.

Matisoff, James A. 2018. Rethinking the Proto-Tibeto-Burman *a- prefix: Glottal and nasal complications. *Journal of Asian and African Studies* 96. 29–69.

Myhill, John. 1997. Toward a functional typology of agent defocusing. *Linguistics* 35. 799–844.

Naw Hsa Eh Ywa. 2013. *A grammar of Kayan Lahta*. Chiang Mai: Payap University M.A. thesis.

Naw Hsar Shee. 2008. *A descriptive grammar of Geba Karen*. Chiang Maiɛ: Payap University MA thesis.

Naw Veronica. 2011. *The phonology of Dermuha and a phonological and lexical comparison between Dermuha, Sgaw Karen and Pwo Karen*. Chiang Mai: Payap University MA thesis.

Nishi, Yoshio. 1995. A brief survey of the controversy in verb pronominalization in Tibeto-Burman. In Yoshio Nishi, James A. Matisoff & Yasuhiko Nagano (eds.), *New horizons in Tibeto-Burman morphosyntax* (Senri Ethnological Studies 41), 1–16. Osaka: National Museum of Ethnology.

Phillips, Audra. 2000. West-Central Thailand Pwo Karen phonology. *33rd ICSTLL Papers*, 99–110. Bangkok: Ramkhamhaeng University.

Phillips, Audra. 2017. *Entities and the expression of grounding and referential coherence in Northern Pwo Karen narrative discourse*. Edmonton: University of Alberta PhD dissertation.

Purser, W. C. B. & Saya Tun Aung. 1922. *A comparative dictionary of the Pwo-Karen dialect*. Rangoon: American Baptist Mission Press.

Shibatani, Masayoshi. 2006. On the conceptual framework for voice phenomena. *Linguistics* 44 (2). 217–269.

Shintani, Tadahiko. 2003a. Notes à propos de l'étymologie du mot karen. *Linguistics of the Tibeto-Burman Area* 26 (1). 15–21.

Shintani, Tadahiko. 2003b. Classification of Brakaloungic (Karenic) languages in relation to their tonal evolution. In Shigeki Kaji (ed.), *Proceedings of the Symposium Cross-linguistic Studies of Tonal Phenomena: Historical Development, Phonetics of Tone, and Descriptive Studies*, 37–54. Tokyo: Research Institute for Languages and Cultures of Asia and Africa.

Solnit, David B. 1997. *Eastern Kayah Li: Grammar, texts, glossary*. Honolulu: University of Hawai`i Press.

Solnit, David B. 2001. Another look at Proto-Karen. Paper presented at the 34th Internatinal Conference of Sino-Tibetan Languages and Linguistics, Kunming.

Solnit, David B. 2006. Verb serialization in Eastern Kayah Li. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *Serial verb constructions: A cross-linguistic typology*, 144–159. Oxford: Oxford University Press.

Solnit, David B. 2013. Proto-Karen rhymes. Paper read at the 46th International Conference on Sino-Tibetan Languages and Linguistics, Dartmouth University.

Solnit, David B. 2017. Eastern Kayah Li. In Graham Thurgood & Randy LaPolla (eds.), *The Sino-Tibetan languages, 2nd edition*, 932–941. London & New York: Routledge.

Stern, Theodore. 1968. Three Pwo Karen scripts: A study of alphabet formations. *Anthropological Linguistics* 10(1). 1–39.

Swanson, Kirstie. 2011. *Serial verb constructions in Bwe Karen*. Chiang Mai: Payap University MA thesis.

Van Valin, Robert D., Jr. & Randy J. LaPolla. 1997. *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press.

Wade, J. 1849. *Vocabulary of the Sgau Karen language*. Tavoy: Karen Mission Press.

Wai Lin Aung. 2013. *A descriptive grammar of Kayah Monu*. Chiang Mai: Payap University MA thesis.

Weinhold, Meredith Lucey. 2011. Serial verb constructions in Sgaw Karen: A comparison of Karenic VP structures. In Marc Brunelle (ed.), *Sgaw Karen papers, presented to Nimrod Andrew*, 76–84. Ottawa: University of Ottawa.