

Homeland of Karenic languages: From the perspective of plant names

Atsuhiko Kato

Keio University

Running head: Homeland of Karenic languages

Comparisons of plant names in Karenic languages reveal that names that can be traced back to Proto-Karen belong to plants that grow in temperate zones, such as bamboo, banyan, and mango. The names for coconut and palmyra palm, which are typical tropical plants, cannot be traced back to Proto-Karen and are borrowings. This suggests that Proto-Karen was spoken in a temperate zone. Meanwhile, the highest diversity of Karenic languages is observed in the area from southern Shan State to Kayah State and northern Karen State in Myanmar. Thus, as per linguistic migration theory, this area may have been the homeland of Karenic languages. Furthermore, the area largely has temperate zones. Hence, we can assume that the homeland of Karenic languages was in this area.

Keywords: Karenic languages, Tibeto-Burman languages, homeland, plant names, linguistic migration theory

1. Introduction

Karenic languages are a group of languages that form the Karenic branch of the Tibeto-Burman languages. The languages belonging to this branch have a subject-verb-object (SVO) basic word order, unlike most other Tibeto-Burman languages that generally have the SOV word order (see Kato, 2021b, for the typological characteristics of Karenic languages). The Karenic branch contains many languages, including Geba, Gek(h)o, Kayah, Kayaw, Kayan, Manu, Monebwa, Mopwa (Mobwa), Paku, Pa-O, Pwo Karen, Sgaw Karen, Thalebwa, Yinbaw, Yintale, and Zayein (Eberhard et al., 2019). In this paper, people who speak Karenic languages are referred to as “Karenic people”. Karenic people have a very wide range

of cultures and lifestyles. There are Buddhists, Christians, and animists. Some of them live in the plains, while others live in mountainous areas. Some of the groups, especially Sgaw Karen, Pwo Karen, and Kayah, are engaged in an armed struggle against the Myanmar government. Even a single ethnic group can have a high degree of diversity, for example, many Sgaw Karens practice Buddhism or Christianity and live in urban areas, while other worship spirits and live in mountainous areas.

Figure 1 presents a rough distribution of the Karenic languages. This map is based on the maps presented by Mitani (1984) and Asher and Moseley (2007), and the information provided in Bradley (2007), with additional information gathered in my field research carried out in Myanmar and Thailand.

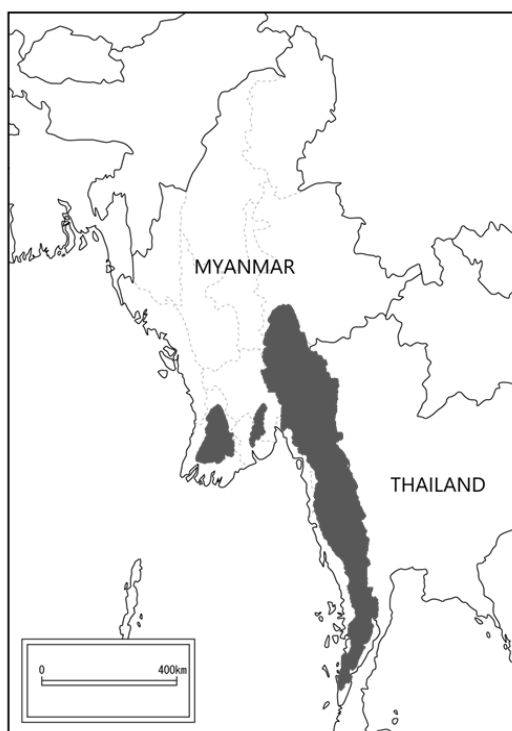


Figure 1. Distribution of Karenic languages

The purpose of this paper is to linguistically explore the homeland of Karenic languages through the plant names found in these languages. Historical studies of Karenic languages, such as the reconstruction of Proto-Karen or studies of their genealogical relationships, include Haudricourt (1946, 1953, 1972, 1975), Luce (1959), Jones (1961), Burling (1969), Peiros (1989), Solnit (2001, 2013), Manson (2002, 2009, 2019), Shintani (2003), and Luangthongkum (2013a, 2013b, 2014a, 2014b, 2019). However, there

have been no studies to estimate the homeland of Karenic languages in terms of linguistics.

Section 2 will list the Karenic languages included in this paper and present their genealogical positions within the Karenic branch. Section 3 will examine the correspondences between the plant names in the Karenic languages to determine whether the names can be traced back to Proto-Karen and determine the climate of the regions where it was spoken. Section 4 will examine the possibility of the area from southern Shan State to Kayah State and northern Karen State being the homeland of Karenic languages due to the highest diversity of Karenic languages in these regions, based on the linguistic migration theory. Section 5 will indicate that the areas discussed in Section 4 have a large portion of temperate climates. Section 6 will conclude the paper with relevant findings.

2. Karenic languages dealt with in this study

There are a total of 16 languages of the Karenic branch studied in this paper. Their names and data sources are listed below (see Section 3 for the order of the languages).

1. Northern Pa-O (Shintani, 2020b)
2. Southern Pa-O (Shintani, 2020b)
3. Eastern Kayah Li (Solnit, 1997)
4. Zayein (Shintani, 2014)
5. Nangki (Shintani, 2016)
6. Yingtalay (Shintani, 2018a)
7. Thaidai (Shintani, 2018b)
8. Sonkan Kayan (Shintani, 2018c)
9. Dosanbu Kayan (Shintani, 2018d)
10. Thamidai (Shintani, 2020a)
11. Pekon Kayan (Manson, 2010a)
12. Bwe (Blimaw) (Henderson, 1997)
13. Blimaw (Shintani, 2017)

14. Western Pwo Karen (Kyonpyaw dialect; the author's data)
15. Eastern Pwo Karen (Hpa-an dialect; the author's data)
16. Sgaw Karen (Hpa-an dialect; the author's data)

The word notations in each language follow the consulted source. The notation of Western Pwo Karen, Eastern Pwo Karen, and Sgaw Karen words, which are based on my data, follows Kato (2022), Kato (2021a), and Kato (forthcoming), respectively. These 16 languages were selected because reliable data were available for a certain number of plant names in each language. Note that Shintani's (2017) Blimaw is a different language from Henderson's (1997) Bwe, which, too, is called Blimaw. Shintani's Blimaw is closer to Mopwa than to Bwe.

The tree diagrams in Figures 2 to 4 show the genealogical relationships between Karenic languages. Figures 2, 3, and 4 were proposed by Manson (2002), Shintani (2003), and Manson (2017), respectively. The capital letters A to G, indicating the subgroups in each tree, are added by me for convenience. The same alphabet does not mean the same genealogical group. According to the respective diagrams, each of the 16 languages belongs to the subgroups presented as follows (the three capital letters in parentheses after each language name indicate the subgroups in Figures 2, 3, and 4, respectively):

1. Northern Pa-O (D, G, A)
2. Southern Pa-O (D, G, A)
3. Eastern Kayah Li (F, F, C)
4. Zayein (E, E, B)
5. Nangki (E, E, B)
6. Yingtalay (E, E, B)
7. Thaidai (E, E, B)
8. Sonkan Kayan (E, E, B)
9. Dosanbu Kayan (E, E, B)
10. Thamidai (E, E, B)
11. Pekon Kayan (E, E, B)
12. Bwe (G, D, C)

- 13. Blimaw (? , C, D)
- 14. Western Pwo Karen (B, A, A)
- 15. Eastern Pwo Karen (B, A, A)
- 16. Sgaw Karen (A, B, D)

Blimaw's subgroup is uncertain in Figure 2 and is indicated by a "?". The genealogical diversity exhibited by these languages suggests that they are sufficient to serve as samples of Karenic languages, regardless of the classification.

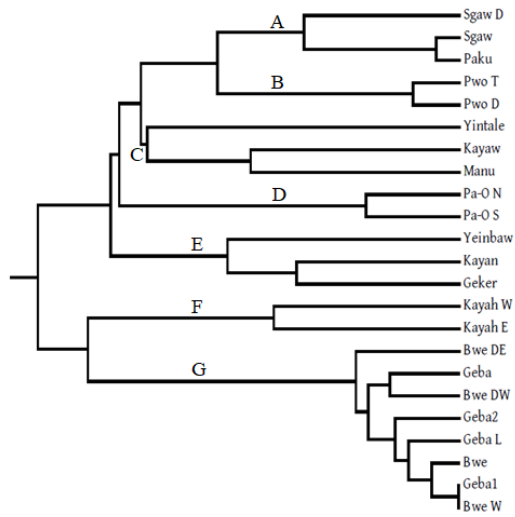


Figure 2. Manson (2002)

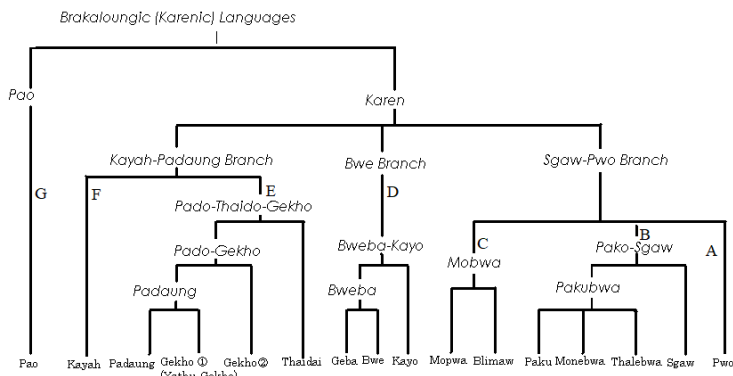


Figure 3. Shintani (2003)

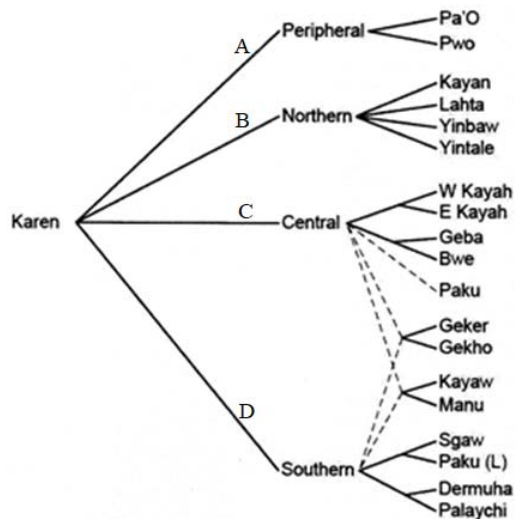


Figure 4. Manson (2017)

3. Comparison of plant names

This section compares the plant names in Karenic languages. Thirteen plants included in this paper are: (1) bamboo, (2) banana, (3) banyan, (4) betel palm, (5) coconut, (6) jackfruit, (7) mango, (8) palmyra palm, (9) pine, (10) rattan, (11) sugarcane, (12) tamarind, and (13) teak. These plants were chosen because their names are observed in a wide range of Karenic languages and they have been mentioned in many of the literature consulted. In Sections 3.1 to 3.13, for each plant name, we will observe the

correspondences in the Karenic branch and consider whether a proto-form can be reconstructed at the Proto-Karen stage. In Section 3.14, a discussion would be presented based on the considerations made in Sections 3.1 to 3.13.

Tables 1 to 13 list the nouns for each plant in Karenic languages. The left column of each table lists the language name and the right column lists the equivalent plant name. In the tables, the languages are divided into six groups for ease of understanding. These six groups are arranged with the more leftward groups in Figure 3 higher. The order of languages within a group is random.

1. Northern Pa-O and Southern Pa-O
2. Eastern Kayah Li
3. Zayein, Nangki, Yingtalay, Thaidai, Sonkan Kayan, Dosanbu Kayan, Thamidai, and Pekon Kayan
4. Bwe
5. Blimaw
6. Western Pwo Karen, Eastern Pwo Karen, and Sgaw Karen

This grouping was made because, as far as the plant names are concerned, the words seem to correspond well with each other within each group in terms of comparative linguistics. Based on Shintani's (2003) classification in Figure 3, which best fits my image of the genealogical relations of Karenic languages among the three classifications shown in Section 2, Group 1 here corresponds with Group G, Group 2 with Group F, Group 3 with Group E, Group 4 with Group D, Group 5 with Group C, and Group 6 with both Groups A and B.

The symbol “—” in the tables indicates that the equivalent word was not provided in the data source. For Western Pwo Karen, Eastern Pwo Karen, and Sgaw Karen, which are based on my data, the same symbol indicates that an equivalent word was not found in the survey.

Data sources are presented in Section 2; however, Eastern Kayah Li forms may be taken from Solnit (2013) or personal communications with him (June 2022). Such cases are noted as “(Solnit, 2013)” and “(Solnit, p.c., 2022)”. Some of the Pekon Kayan forms may be taken from Manson (2010b), in which cases they are noted as “(Manson, 2010b)”.

3.1 Bamboo

Table 1 lists the words for bamboo.

Table 1. Words denoting bamboo in Karenic languages

Languages	Words
Northern Pa-O	wa34
Southern Pa-O	wa44
Eastern Kayah Li	ve
Zayein	va42
Nangki	va31
Yingtalay	va55
Thaidai	hua45
Sonkan Kayan	hua55
Dosanbu Kayan	hua55
Thamidai	hua42
Pekon Kayan	hwâ
Bwe	hũ
Blimaw	wạ33
Western Pwo Karen	wà
Eastern Pwo Karen	wá
Sgaw Karen	wâ

All of these forms can be traced back to a single form of Proto-Karen. Matisoff (2003, p. 305; 2015,

p. 630) reconstructed *hwa as the Proto-Karen form and Luangthongkum (2019) reconstructed *hwa^{B.1}. Similarly, I assume *hwa² to be a Proto-Karen form (*hw is a voiceless labial-velar approximant). The forms of all 16 languages can be explained as having evolved from the proto-form *hwa². Thus, the form denoting bamboo is believed to have existed at the Proto-Karen stage.

3.2 Banana

Table 2 lists the words for banana.

Table 2. Words denoting banana in Karenic languages

Languages	Words
Northern Pa-O	tha [?] 34 thi ³ 4
Southern Pa-O	θa ⁴ 2 thi ⁴ 4
Eastern Kayah Li	dīklwí
Zayein	le ⁴ 2 ja ⁴ 2 le ⁴ 2
Nangki	—
Yingtalay	θa ⁵ 5 klo ⁵ 3
Thaidai	cə ³ 3 kuai ⁴ 2
Sonkan Kayan	khlu ³ 1
Dosanbu Kayan	khlu ⁴ 2
Thamidai	cu ⁴ 2 kui ⁵ 5
Pekon Kayan	k ^h lwí
Bwe	yà

1. The capital letters A, B, and C in the reconstructed Proto-Karen forms represent the tones reconstructed by Haudricourt (1946). A and B are plain tones and C is a stopped tone. Some researchers refer to these as 1, 2, and 3. Haudricourt himself used 1, 2, and 3, hence, I follow this method. For the process of reconstruction of proto-tones by Haudricourt (1946) and modification in Haudricourt (1975), see Kato (2018).

Blimaw	la11 kwe11
Western Pwo Karen	θaʔklú
Eastern Pwo Karen	θàkwì
Sgaw Karen	təkwi

Manson (2019) reconstructed **k^hlwi^A* for the Proto-Kayan stage, which is the proto-language of languages including Zayein, Nangki, Yingtalay, Thaidai, Sonkan Kayan, Dosanbu Kayan, Thamidai, and Pekon Kayan in the table. Similar forms are found in Eastern Kayah Li, Blimaw, Western Pwo Karen, Eastern Pwo Karen, and Sgaw Karen. These resemble the Proto-Mon-Khmer form, **t₁luuy[]* or **t₁luəyʔ*, which Shorto (2006, p. 408) reconstructed. Thus, it is possible that Proto-Karen borrowed a word for a banana from a certain Mon-Khmer language and that the forms in Table 2 evolved from that word. However, this needs further consideration because the forms do not always show regular phonological correspondences. For example, the forms of Western Pwo Karen, Eastern Pwo Karen, and Sgaw Karen are similar; however, they do not show regular correspondence. That is, Sgaw Karen *tə* usually does not correspond with Western Pwo Karen *θaʔ* or Eastern Pwo Karen *θà* and the Western Pwo Karen *kl-* usually does not correspond with Eastern Pwo Karen *kw-* or Sgaw Karen *kw-*. Therefore, each Karenic language may have borrowed Mon-Khmer words for banana independently.

Many Karenic languages, however, have different words for banana. These words are related to the form of Bwe shown in Table 2. Solnit (2013) provided the following forms:

1. Pa-O *ηàʔ*
2. Eastern Kayah Li *ηē*
3. Pekon Kayan *ηá*
4. Kayaw *jə̄.sà*
5. Blimaw (Henderson's 1997 Bwe) *yà*
6. Pwo Karen *jāʔ*
7. Sgaw Karen *jàʔ*

My Pwo Karen and Sgaw Karen data have the following forms and represent a type of banana that

grows in the jungle: Western Pwo Karen *jaʔ*, Eastern Pwo Karen *jà*, and Sgaw Karen *jàʔ*. These forms, phonologically, correspond well among Karenic languages. Solnit (2013) reconstructed **ŋakD* in Proto-Karen. Similarly, I reconstruct **ŋak³* in Proto-Karen. All the forms listed above can be explained as having evolved from this reconstructed form. Thus, a form for banana is believed to have existed at the Proto-Karen stage. I assume that the forms that evolved from **ŋak³* were the original words for banana in Karenic languages; however, later, for reasons unknown, words of Mon-Khmer origin became dominant. A possible reason may have been that the original words represented “cooking bananas”, while the Mon-Khmer words represented “dessert bananas”.

3.3 Banyan

Table 3 lists the words for banyan.

Table 3. Words denoting banyan in Karenic languages

Languages	Words
Northern Pa-O	ŋɔŋ34
Southern Pa-O	kloŋ44
Eastern Kayah Li	—
Zayein	ɤɔŋ42
Nangki	sa31 klɔ31
Yingtalay	klou55
Thaidai	khlaɔ45
Sonkan Kayan	khlou55
Dosanbu Kayan	khlou55
Thamidai	khlaɔ42
Pekon Kayan	k ^h lɔn (Manson, 2010b)
Bwe	klo
Blimaw	xlɔ55

Western Pwo Karen	khlòuN
Eastern Pwo Karen	khlóuN
Sgaw Karen	khlô

All these words, except the Northern Pa-O form, can be assumed to be related to each other and a Proto-Karen form can be reconstructed. Manson (2019) reconstructed *k^hlɔN^B for Proto-Kayan. Although no previous studies have proposed a reconstructed form at the level of Proto-Karen, I assume *khlɔŋ² for Proto-Karen. The language forms, except Northern Pa-O, can be explained as having evolved from this reconstructed form. Loss of syllable-final consonants is common in Karenic languages; thus, the loss of syllable-final nasals in the nine languages, e.g. Sgaw Karen *khlô*, in the table is not strange. Thus, a form for banyan is believed to have existed at the Proto-Karen stage. Note that the first syllable, *sa31*, of the Nangki form is a morpheme representing fruit, the reasons for which are unknown.

3.4 Betel palm

Table 4 lists the words for betel palm (areca).

Table 4. Words denoting betel palm in Karenic languages

Languages	Words
Northern Pa-O	mok34 mu53
Southern Pa-O	maʔ32 mu44
Eastern Kayah Li	mū (Solnit, p.c., 2022)
Zayein	məŋ42 mɔ33
Nangki	mouʔ53
Yingtalay	ma53
Thaidai	kun42
Sonkan Kayan	ni31 mɔu31
Dosanbu Kayan	mɔu31

Thamidai	kwen55
Pekon Kayan	bamò
Bwe	màmú
Blimaw	la11 ma11
Western Pwo Karen	θeʔ
Eastern Pwo Karen	θè
Sgaw Karen	θēʔ

Manson (2019) reconstructed *bəmō^A for Proto-Kayan. For the Proto-Karen level, Peiros (1989) reconstructed *mō^B; however, I do not believe that a Proto-Karen form can be reconstructed since Blimaw, Western Pwo Karen, Eastern Pwo Karen, and Sgaw Karen have different forms. Therefore, it cannot be demonstrated that a form denoting betel palm existed at the Proto-Karen stage. Note that the Thaidai and Thamidai forms were probably borrowed from Burmese *kúN* <kvam“>. Angle bracketed forms of Burmese and Mon represent the written forms (Burmese and Mon transliterations follow the systems of the Library of Congress and Shorto [1962], respectively).

3.5 Coconut

Table 5 lists the words for coconut.

Table 5. Words denoting coconut in Karenic languages

Languages	Words
Northern Pa-O	mok34 ʔun31
Southern Pa-O	moʔ32 ʔon42
Eastern Kayah Li	míʔuse (Solnit, p.c., 2022)
Zayein	mak33 qa33 ʔoan42
Nangki	ʔou31
Yingtalay	ma53 ʔouj55

Thaidai	ma33 ʔon42
Sonkan Kayan	ʔouŋ55
Dosanbu Kayan	ʔouŋ55
Thamidai	ʔoun42
Pekon Kayan	ʔõŋ
Bwe	u
Blimaw	ʔou55
Western Pwo Karen	phláu
Eastern Pwo Karen	phlò
Sgaw Karen	pl̄5 (x̄5)

The Northern Pa-O, Southern Pa-O, Zayein, Yingtalay, and Thaidai forms are borrowed from Shan *maak² ʔun¹* (Hudak, 2000). The Eastern Kayah Li (the second syllable), Nangki, Sonkan Kayan, Dosanbu Kayan, Thamidai, Pekon Kayan, Bwe, and Blimaw forms are borrowed from the above-mentioned Shan form or Burmese *ʔoun* <'un">. The Western Pwo Karen, Eastern Pwo Karen, and Sgaw Karen forms are most likely borrowed from Mon *prèa* <brau> (Shorto, 1962, p. 154). The Hpa-an dialect of Sgaw Karen calls coconut *pl̄5*; however, certain dialects use the bracketed form *x̄5*. Both forms can be considered to have emerged from the form **br̄5*, which I assume for the Proto-Sgaw Karen level.

These words are borrowed from different origins and a Proto-Karen form cannot be considered to be reconstructed. Thus, it cannot be demonstrated that a form denoting coconut existed at the Proto-Karen stage.

3.6 Jackfruit

Table 6 lists the words for jackfruit.

Table 6. Words denoting jackfruit in Karenic languages

Languages	Words
-----------	-------

Northern Pa-O	mok34 klaŋ22
Southern Pa-O	maʔ32 klaŋ53
Eastern Kayah Li	—
Zayein	laŋ42
Nangki	ʔa232 laŋʔ53
Yingtalay	—
Thaidai	di42 ne33
Sonkan Kayan	ba31 laŋ31
Dosanbu Kayan	ba42 laŋ31
Thamidai	pe55 nen55
Pekon Kayan	balāŋ
Bwe	təbəne, thəbənı
Blimaw	təl1 ɲil1
Western Pwo Karen	nwê
Eastern Pwo Karen	nwē
Sgaw Karen	pənwè

Manson (2019) reconstructed *bəlaŋ^B for the Proto-Kayan stage. However, for the Proto-Karen stage, it seems difficult to reconstruct a proto-form. Thus, it cannot be demonstrated that a form denoting jackfruit existed at the Proto-Karen stage. Note that Thamidai, Western Pwo Karen, Eastern Pwo Karen, and Sgaw Karen forms are probably borrowed from Sanskrit/Pali *panasa* via Mon *pənh* <pnah> (Shorto, 1962, p. 145) or Burmese *péinné* <pinnai>.

3.7 Mango

Table 7 lists the words for mango.

Table 7. Words denoting mango in Karenic languages

Languages	Words
Northern Pa-O	ta31 khoʔ34
Southern Pa-O	θa42 khoʔ32
Eastern Kayah Li	təkhé
Zayein	θa33 khwa42
Nangki	khu31
Yingtalay	ta33 khə53
Thaidai	kho45
Sonkan Kayan	khə53
Dosanbu Kayan	khaoʔ53
Thamidai	khaoʔ53
Pekon Kayan	k ^h àu
Bwe	—
Blimaw	la11 hu11
Western Pwo Karen	khoʔ
Eastern Pwo Karen	khò
Sgaw Karen	təkhōʔ, θəkhōʔ

All the words listed in Table 7 can be traced back to Proto-Karen. Solnit (2013) reconstructed *khok and Luangthongkum (2019) reconstructed *khəʔ^D for Proto-Karen. I assume the Proto-Karen form is *səkhok³ and all the forms in Table 7 can be regarded as reflections of this form. The vowel *ɛ* in the Eastern Kayah Li form may seem odd, but this shows a regular correspondence. For example, Eastern Kayah Li *thé* ‘pig’ (Solnit, 1997, p. 371) corresponds to Eastern Pwo Karen *thò* ‘pig’ and Sgaw Karen *thōʔ* ‘pig’. The consonant *h* in the second syllable of the Blimaw form does not show a regular correspondence to voiceless aspirated velar stops in the other languages, but presumably there was a sporadic change of *kh* to *h* in this word. Thus, a form representing mango is believed to have existed in Proto-Karen.

3.8 Palmyra palm

Table 8 lists the words for palmyra palm.

Table 8. Words denoting palmyra palm in Karenic languages

Languages	Words
Northern Pa-O	than ³⁴
Southern Pa-O	thæn ⁴²
Eastern Kayah Li	—
Zayein	—
Nangki	—
Yingtalay	—
Thaidai	than ⁴⁵
Sonkan Kayan	—
Dosanbu Kayan	thaŋ ⁵⁵
Thamidai	na ⁴² then ⁴²
Pekon Kayan	—
Bwe	thá
Blimaw	tho ⁵⁵
Western Pwo Karen	thàn
Eastern Pwo Karen	tà
Sgaw Karen	tá

Northern Pa-O, Southern Pa-O, Thaidai, Dosanbu Kayan, Thamidai, Bwe, Blimaw, and Western Pwo Karen forms probably borrowed from Burmese *thán* <than'>. Eastern Pwo Karen and Sgaw Karen forms were probably borrowed from Mon *ta* <tā> (Shorto, 1962, p. 108). Thus, it cannot be demonstrated that a form denoting palmyra palm existed at the Proto-Karen stage. Note that both Burmese and Mon forms

might possibly reflect Sanskrit *tāla*.

3.9 Pine

Table 9 lists the words for pine.

Table 9. Words denoting pine in Karenic languages

Languages	Words
Northern Pa-O	ŋo31
Southern Pa-O	—
Eastern Kayah Li	ŋū (Solnit, p.c., 2022)
Zayein	ju55
Nangki	mei31 ŋou232
Yingtalay	ta33 ŋo53
Thaidai	thin33 ju33
Sonkan Kayan	ŋou31
Dosanbu Kayan	ŋou42
Thamidai	—
Pekon Kayan	θêŋnó
Bwe	—
Blimaw	thi55 ju44
Western Pwo Karen	—
Eastern Pwo Karen	chàʊpənâin
Sgaw Karen	chyó (Drum Publication Group, 2000)

Some of these words begin with *ŋ*; however, their etymology is unknown. The Thaidai and Blimaw forms are borrowed from Burmese *thínyú* <thañ“ rū”>. My Sgaw Karen consultants from Hpa-an mentioned that they do not know the Sgaw Karen word for pine because pines do not grow in their

villages where the climate is too hot for pines; however, Drum Publication Group (2000) contains the form *chyó*, thus, it has been to the list. Nevertheless, I do not believe that a Proto-Karen form for pine can be reconstructed. Thus, it cannot be demonstrated that a form denoting pine existed at the Proto-Karen stage.

3.10 Rattan

Table 10 lists the words for rattan.

Table 10. Words denoting rattan in Karenic languages

Languages	Words
Northern Pa-O	rei53
Southern Pa-O	—
Eastern Kayah Li	ri (Solnit, 2013)
Zayein	ꞑei42
Nangki	ʔað232
Yingtalay	rui31
Thaidai	rei33
Sonkan Kayan	rei33
Dosanbu Kayan	rei31
Thamidai	ri55
Pekon Kayan	rē
Bwe	wɪ
Blimaw	ɣi33
Western Pwo Karen	ɣê
Eastern Pwo Karen	ɣī
Sgaw Karen	ɣè

All the words, except for the Nangki form, are related to each other and a proto-form can be reconstructed. Solnit (2013) reconstructed *reB for Proto-Karen and Luangthongkum (2019) reconstructed *re^B. Similarly, I assume *re² is the Proto-Karen form and all the forms, except the Nangki form, can be explained as having developed from this form. Thus, a form for rattan is believed to have existed in Proto-Karen.

3.11 Sugarcane

Table 11 lists the words for sugarcane.

Table 11. Words denoting sugarcane in Karenic languages

Languages	Words
Northern Pa-O	toŋ ³⁴ lai ³¹
Southern Pa-O	toŋ ⁴⁴ lai ⁴²
Eastern Kayah Li	dīkle bō
Zayein	qiŋ ⁴² kwai ⁵⁵ boŋ ⁵⁵
Nangki	he ³¹ kli ²³²
Yingtalay	khi ⁵³
Thaidai	din ⁴² khle ⁴²
Sonkan Kayan	khle ³¹ boŋ ³¹
Dosanbu Kayan	khlai ⁴² boŋ ⁴²
Thamidai	ceŋ ⁵⁵ khlae ⁵⁵
Pekon Kayan	dèŋk ^h lái ^{bú}
Bwe	dăkhlé
Blimaw	chi ¹¹ dōu ⁵⁵
Western Pwo Karen	shipho?
Eastern Pwo Karen	ciphó
Sgaw Karen	thípò?

The words denoting sugarcane vary from one language to another, making it impossible to reconstruct a proto-form of these words in Proto-Karen, though the forms with the cluster /kl-/ or /khl-/ in Eastern Kayah Li, Nangki, Thaidai, Sonkan Kayan, Dosanbu Kayan, Thamidai, Pekon Kayan, and Bwe could reflect some relationship. Thus, it cannot be demonstrated that a form for sugarcane existed at the Proto-Karen stage.

3.12 Tamarind

Table 12 lists the words for tamarind.

Table 12. Words denoting tamarind in Karenic languages

Languages	Words
Northern Pa-O	mok ³⁴ kreŋ ³⁴
Southern Pa-O	moŋ ⁴² kreŋ ⁴⁴
Eastern Kayah Li	—
Zayein	mə ³³ ɤəŋ ⁴²
Nangki	ʔa ²³² klai ⁵³
Yingtalay	mi ⁵⁵ piaŋ ⁵³
Thaidai	bun ⁴² klein ³³
Sonkan Kayan	ma ³¹ klaiŋ ⁵³
Dosanbu Kayan	ŋklaiŋ ³¹
Thamidai	ma ⁴² klen ⁵⁵
Pekon Kayan	—
Bwe	màklé
Blimaw	mɯ ¹¹ kle ⁵⁵
Western Pwo Karen	mánkhláun
Eastern Pwo Karen	mànkhlōn

Manson (2019) reconstructed *mak^hlɛŋ^B for Proto-Karen. However, the forms shown in Table 12 seem to have different etymologies. According to the SEALang Mon-Khmer Etymological Dictionary, made available on the Internet by Paul Sidwell, the forms for tamarind in Riang-Lang and Riang-Sak are *maŋ² kleəŋ²* and *maŋ² kleŋ²*, respectively. The forms of Northern Pa-O, Southern Pa-O, Zayein, Nangki, Thaidai, Sonkan Kayan, Dosanbu Kayan, Thamidai, Bwe, and Blimaw seem to have a certain relation with the Riang form (Burmese *mǎjī* <man‘ kyaññ‘>, too, may be related to this form). In contrast, the Western Pwo Karen, Eastern Pwo Karen, and Sgaw Karen forms for tamarind seem to be related to the Mon form *mèaŋ klòn* <mañ glon> (Shorto, 1962, p. 168). One of my Sgaw Karen consultants mentioned that in Sgaw Karen villages in the northern Karen state, tamarind is called *mǎklē* instead of *mòklò*. The form *mǎklē* seems to have some relation to the Riang form mentioned above. It is possible that after the Proto-Karen stage, the northern Karenic languages, such as Pa-O, Kayan, Bwe, and the northern Sgaw Karen dialects, adopted the words for tamarind from Palaungic to which Riang belongs, while the southern Karenic languages, such as Western Pwo Karen, Eastern Pwo Karen, and the southern Sgaw Karen dialects, adopted the words from Mon. Thus, it is difficult to demonstrate that a word for tamarind existed at the Proto-Karen stage.

3.13 Teak

Table 13 lists the words for teak.

Table 13. Words denoting teak in Karenic languages

Languages	Words
Northern Pa-O	mai22 sak34
Southern Pa-O	—
Eastern Kayah Li	leha
Zayein	mai42 sak42

Nangki	—
Yingtalay	la55 hai55
Thaidai	cu42
Sonkan Kayan	kiao31
Dosanbu Kayan	kla55
Thamidai	cun55
Pekon Kayan	cú
<hr/>	
Bwe	bahi
<hr/>	
Blimaw	la11 hei55
<hr/>	
Western Pwo Karen	pə̀jì
Eastern Pwo Karen	pə̀jí
Sgaw Karen	pə̀hî
<hr/>	

The Northern Pa-O and Zayein forms are borrowed from Shan *may⁵ shak⁴* (Hudak, 2000). The forms of Thaidai, Thamidai, and Pekon Kayan are probably borrowed from Burmese *cún* <kyvan'>. The origins of the other forms are unknown. Looking at these correspondences, it is not possible to reconstruct a proto-form for teak in Proto-Karen. Thus, it is difficult to demonstrate that a form denoting teak existed at the Proto-Karen stage.

3.14 Discussion on Karenic plant names

Words for bamboo, banana, banyan, mango, and rattan may have been present in Proto-Karen. In contrast, words for betel palm, coconut, jackfruit, palmyra palm, pine, sugarcane, tamarind, and teak cannot be demonstrated to have been present at the Proto-Karen stage. In particular, the words for coconut and palmyra palm in the languages included in this paper are likely to have originated from the borrowings of a later period than Proto-Karen. The words for coconut are borrowed from Shan, Mon, and Burmese and the words for palmyra palm are borrowed from Mon and Burmese. Therefore, it is reasonable to assume that words representing these two plants did not exist in Proto-Karen.

Coconut (*Cocos nucifera*) and palmyra palm (*Borassus flabellifer*) are typical tropical plants, which

require high temperatures to grow. According to Blombery and Rodd (1982, p. 185), the “extreme minimum temperature” required for both these plants to grow is 10°C or higher. In contrast, bamboo, banana, banyan, mango, and rattan, plants for which Proto-Karen is believed to have had names, have many species that can grow in temperate zones, that is, they do not require temperatures as high as coconut palm and palmyra palm do. For example, rattan (*Calamus*) has about 370 species (Blombery and Rodd, 1982, p. 68), some of which require an extreme minimum temperature of 10°C or higher, while others require a lower extreme minimum temperature of 3°C to 10°C (pp. 185–186). For information about the environments in which bamboo, banana, banyan (especially, *Ficus microcarpa*), and mango grow, see Kigomo (2007, pp. 51–52), Turner (2003), Sakai (1979, p. 5), and Rajan (2012, pp. 72–74), respectively. The temperatures given in these references suggest that while these plants prefer warmer temperatures, they can also grow in temperate zones. Hence, it can be assumed that the speakers of Proto-Karen lived in a temperate zone.

4. Homeland of Karenic languages as per the linguistic migration theory

It is important to explore the regions where the largest number of Karenic languages are spoken. According to Asher and Moseley (2007) and my fieldwork, the largest number of Karenic languages are spoken throughout southern Shan State, Kayah State, and northern Karen State in Myanmar. This is approximately the area encircled in Figure 5. Within this relatively small area, numerous Karenic languages, including Bwe, Geba, Gekho, Kayah, Kayaw, Kayan, Manu, Monebwa, Mopwa, Paku, Pa-O, Sgaw Karen, Thalebwa, Yinbaw, Yintale, and Zayein, are spoken. There are only three Karenic languages spoken outside the circle, i.e., Pwo Karen, Sgaw Karen, and Pa-O. Furthermore, the distribution of Pa-O outside the circle is limited to the northern part of Mon State and the central part of Karen State. Thus, primarily Pwo Karen and Sgaw Karen are the Karenic languages spoken outside the circle.

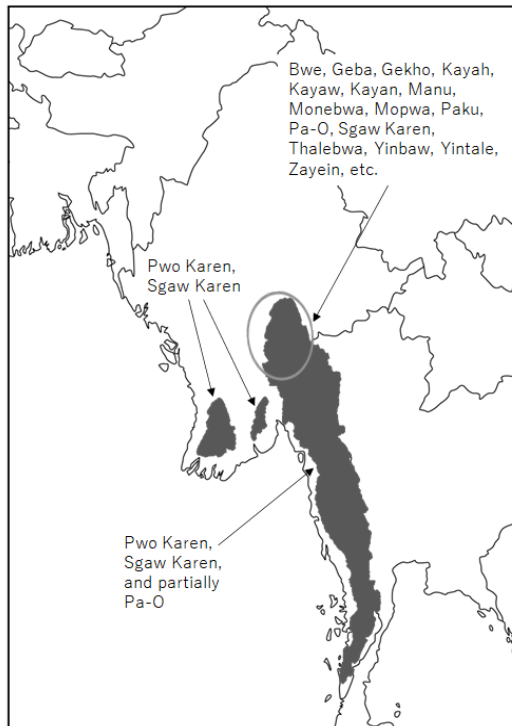


Figure 5. Distribution of Karenic languages (cf. Figure 1) and the area where the largest number of them are spoken

According to the linguistic migration theory, first suggested by Sapir (1916) and refined by Dyen (1956), the area where the largest number of related languages are spoken is the “homeland” of the languages (see also Blust, 1984; Campbell and Mixco, 2007, pp. 106–108). Therefore, the homeland of the Karenic languages can be presumed to have been located within the encircled area in Figure 5.

5. Homeland of Karenic languages and its climate

In Section 4, we discussed that, in light of the linguistic migration theory, the homeland of Karenic languages can be assumed to have been located within the area from southern Shan State to Kayah State and northern Karen State. In Section 3, we discussed that the speakers of Proto-Karen lived in a temperate zone. If the area presumed in Section 4 had a temperate climate, it would be even more likely that the homeland was located in this area.

Figure 6 represents a climate map of Myanmar that was created by the author based on the world map of the Köppen-Geiger climate classification by Kotték et al. (2006). The encircled area of Figure 5 overlaps with this map. It is clear that there is a large area of Cwa or Cwb, i.e., a temperate zone, in the circle. The encircled area is located at the southwestern edge of the Shan Plateau and, despite being south of the Tropic of Cancer, a significant portion of the area has a temperate climate due to its high altitude, ranging from 500 to 2000 meters above sea level. Therefore, it is reasonable to assume that the last homeland of Karenic languages was located within the encircled area.

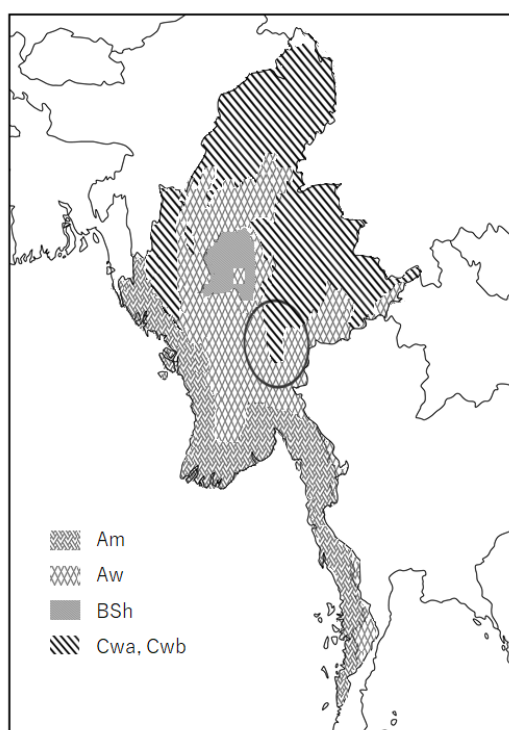


Figure 6. The climates of Myanmar and the area where the largest number of Karenic languages are spoken (cf. Figure 5)

6. Conclusion

Proto-forms for bamboo, banana, banyan, mango, and rattan can be reconstructed in Proto-Karen. All these plants can grow in temperate zones. However, the words for coconut and palmyra palm, both of

which are typical tropical plants, cannot be reconstructed at the Proto-Karen stage. The words representing coconut and palmyra palm in the modern Karenic languages are most likely borrowed from other languages post-Proto-Karen. Therefore, it can be assumed that the Proto-Karen-speaking area had a temperate climate. Meanwhile, looking at the distribution of Karenic languages, the area from southern Shan State to Kayah State and northern Karen State shows the highest diversity. Therefore, as per the linguistic migration theory, this area could be the homeland of Karenic languages. Moreover, this area has a large distribution of temperate climate. This supports the belief that the last homeland of Karenic languages was most likely located in this area (encircled in Figure 6).

Figure 7 is a photograph of the village of Pa-O near Taunggyi, Shan State. This village is located within the circle in Figure 6 and its climate is probably the Cwa of Köppen-Geiger. Coconut and palmyra palm trees cannot be found in this village. The landscape of the homeland of Karenic languages would have been similar to Figure 7. In contrast, Figure 8 is a photograph of a Pwo Karen village near Hpa-an, Karen State and Figure 9 is of a Pwo Karen village near Kyonpyaw, Ayeyarwady Region. Both villages are located outside the circle in Figure 6 and their climate is probably the Am of Köppen-Geiger. Both villages are filled with coconut trees, evident in the photographs.



Figure 7. A Pa-O village near Taunggyi, Shan State (photograph taken by the author in 1994)



Figure 8. A Pwo Karen village near Hpa-an, Karen State (photograph taken by the author in 2019)



Figure 9. A Pwo Karen village near Kyonpyaw, Ayeyarwady Region (photograph taken by the author in 1993)

This paper demonstrates the possible homeland of Karenic languages. However, there are certain limitations. This study does not include the etymology of certain words, those of which proto-forms could not be reconstructed. The words for coconut and palmyra palm are borrowed from Shan, Mon, or Burmese. Future research must examine the etymology of the words for betel palm, jackfruit, pine, sugarcane, tamarind, and teak to determine whether each form is a native Karenic word or has been borrowed.

In addition, a precise estimation of the time of arrival of the speakers of Proto-Karen in the homeland is required. Kato (2019) estimates the split of Western Pwo Karen and Eastern Pwo Karen to have occurred in the seventeenth century based on their lexical similarity. Similarly, the split of Pwo Karen and Sgaw Karen is estimated to have occurred in the eleventh century. However, the time for Proto-Karen further goes back. I presume that the speakers of Proto-Karen would have already arrived in the homeland shown in Figure 6 between the first century to the fifth century A.D. Presumably, this migration would have taken place from north to south along the Salween River. According to Ge et al. (2013), who estimated temperature changes in China over the past 2000 years, although there were several warm intervals, they exhibited similar or lower levels than contemporary temperatures. It is quite possible that the Shan Plateau adjacent to China has experienced similar temperature changes. Thus, when Proto-Karen

was spoken, a large portion of the area shown in Figure 6 probably belonged the temperate zone, just as it does today.

Acknowledgements

This study is based on the paper read at the 2nd Workshop on Linguistic and Cultural Diversity in the Northeast India – Myanmar – Southwest China region, JSPS-NTU/NUS Joint Research Project. I express my gratitude to all those who gave their valuable comments, especially Prof. Hideo Sawada. I am deeply indebted to my Karen friends including, Saw Hla Chit, Nant Dahlia Win, Saw Khin Maung Aye, and Naw Snow Paw, who provided the information about the plant names in Pwo Karen and Sgaw Karen.

References

- Asher R. E. & Moseley, Christopher (Eds.). (2007). *Atlas of the world's languages (second edition)*. Routledge.
- Blombery, Alec & Rodd, Tony (1982). *Palms*. Angus & Robertson Publishers.
- Blust, Robert (1984). The Austronesian homeland: A linguistic perspective. *Asian Perspectives*, 26, 45–67.
- Bradley, David (2007). East and Southeast Asia. In Asher R. E. & Moseley, Christopher (Eds.), *Atlas of the world's languages (second edition)* (pp. 159–197). Routledge.
- Burling, Robbins (1969). Proto-Karen: A reanalysis. *Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics*, 1–116.
- Campbell, Lyle, & Mixco, Mauricio J. (2007). *A glossary of historical linguistics*. Edinburgh University Press.
- Drum Publication Group (2000). *Karen English dictionary*. Drum Publication Group.
- Dyen, Isidore (1956). Language distribution and migration theory. *Language*, 32, 611–626.
- Eberhard, David M., Gary F. Simons, & Charles D. Fennig (Eds.). (2019). *Ethnologue: Languages of the world*. Twenty-second edition. SIL International.
- Ge, Q., Z. Hao, J. Zheng, and X. Shao (2013). Temperature changes over the past 2000 yr in China and

comparison with the Northern Hemisphere. *Climate of the Past*, 9, 1153–1160.

<https://doi.org/10.5194/cp-9-1153-2013>

Haudricourt, André-Georges (1946). Restitution du karen commun. *Bulletin de la Société de Linguistique de Paris*, 42(1), 103–111. (Reprinted: Haudricourt 1972, pp. 131–40.)

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, pp. 141–145.)

Haudricourt, André-Georges (1972). *Problèmes de phonologie diachronique*. SELAF.

Haudricourt, André-Georges (1975). Le système des tons du karen commun. *Bulletin de la Société de Linguistique de Paris*, 70(1), 339–43.

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

Hudak, Thomas John (2000). *Cushing's Shan-English dictionary: A phonetic version*. Arizona State University.

Jones, Robert. B. (1961). *Karen linguistic studies: Description, comparison, and texts*. University of California Publications in Linguistics, No. 25. University of California Press.

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 (2019). Karen and surrounding languages. In Norihiko Hayashi (Ed.), *Topics in middle Mekong linguistics* (Journal of Research Institute 60) (pp. 123–150). Kobe City University of Foreign Studies.

Kato, Atsuhiko (2021a). Pwo Karen writing systems. *Reports of the Keio Institute of Cultural and Linguistic Studies*, 52, 23–55.

Kato, Atsuhiko (2021b). Typological profile of Karenic languages. In Paul Sidwell and Mathias Jenny (Eds.), *The languages and linguistics of Mainland Southeast Asia* (pp. 337–367). Mouton de Gruyter.

<https://doi.org/10.1515/9783110558142-018>

Kato, Atsuhiko (2022). Pwo Karen writing systems 2: Western Pwo Karen. *Reports of the Keio Institute of Cultural and Linguistic Studies*, 53, 23–57.

Kato, Atsuhiko (2023). Letalanyah: A Buddhist writing system of Sgaw Karen. *Reports of the Keio*

Institute of Cultural and Linguistic Studies, 54, 27–52.

- Kigomo, Bernard N. (2007). *Guidelines for growing bamboo*. KEFRI guideline series, No. 4. Kenya Forestry Research Institute.
- Kottek, Markus., Grieser, Jürgen, Beck, Christoph, Rudolf, Brund, & Rubel, Franz (2006). World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift, 15*(3), 259–263.
<https://doi.org/10.1127/0941-2948/2006/0130> (Maps are available at: <http://koeppen-geiger.vu-wien.ac.at/present.htm>)
- Luangthongkum, Theraphan (2013a). A view on Proto-Karen phonology and lexicon. (unpublished ms. contributed to STEDT).
- Luangthongkum, Theraphan (2013b). Problems of the B' tone in Proto-Karen (PK). Paper presented at SEALS 23, Bangkok.
- Luangthongkum, Theraphan (2014a). Karenic as a branch of Tibeto-Burman: More evidence from Proto-Karen. Paper presented at SEALS 24, Yangon.
- Luangthongkum, Theraphan (2014b). Proto-Karen (*k-rjaŋA) Fauna. *MANUSYA: Journal of Humanities, Special Issue, 20*, 86–123.
- Luangthongkum, Theraphan (2019). A view on Proto-Karen phonology and lexicon. *Journal of the Southeast Asian Linguistics Society, 12*(1), i–lii.
<https://doi.org/10524/52441>
- Luce, Gordon. H. (1959). Introduction to the comparative study of Karen languages. *Journal of Burma Research Society 42*(1), 1–18.
- Manson, Ken (2002). *Karenic language relationships: A lexical and phonological analysis*. Department of Linguistics, Payap University.
- Manson, Ken (2009). *Prolegomena to reconstructing Proto-Karen*. La Trobe Working Papers in Linguistics 12.
- Manson, Ken (2010a). A grammar of Kayan, a Tibeto-Burman language. Ph.D. dissertation at La Trobe University.
- Manson, Ken (2010b). Kayan-English/English-Kayan dictionary. (unpublished ms. contributed to STEDT).
- Manson, Ken (2017). The characteristics of the Karen branch of Tibeto-Burman. In Ding, Picus Sizhi, &

Pelkey, Jamin (Eds.), *Sociohistorical linguistics in Southeast Asia: New horizons for Tibeto-Burman studies in honor of David Bradley* (pp. 149–168). Leiden/Boston: Brill.

https://doi.org/10.1163/9789004350519_010

Manson, Ken (2019). A reconstruction of Proto-Kayan. (unpublished ms.)

Matisoff, James A. (2003). *Handbook of Proto-Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction*. University of California Press.

Matisoff, James A. (2015). *The Sino-Tibetan etymological dictionary and thesaurus*. Berkeley: The Sino-Tibetan Etymological Dictionary and Thesaurus Project.

Mitani, Yasuyuki (1984). Tōnan Ajia no gengo bunpu [Linguistic distribution in Southeast Asia]. In Taryo Obayashi (Ed.), *Minzoku no sekaishi 6: Tōnan Ajia no minzoku to rekisi* [History of ethnic groups: Ethnicity and history of Southeast Asia]. Yamakawa Shuppansha. (in Japanese)

Peiros, Iliá (1989). Некоторые уточнения протокаренской реконструкции [Some enhancements for the reconstruction of Proto-Karen]. In Историческая акцентология и сравнительно-исторический метод [Historical accentology and the comparative method] (pp. 225–254). Nauka. (in Russian)

Rajan, Shailendra (2012). Phenological responses to temperature and rainfall: A case study of mango. In Sthapit, Bhuwon, Rao, V. Ramanatha, & Sthapit, Sajal (Eds.), *Tropical fruit tree species and climate change* (pp. 71–96). Bioversity International.

Sakai, Akira (1979). Freezing tolerance of evergreen and deciduous broad-leaved trees in Japan with reference to tree regions. *Low Temperature Science. Ser. B, Biological Sciences*, 36, 1–19.

Sapir, Edward (1916). *Time Perspective in Aboriginal American Culture, A Study in Method*. Ottawa Government Printing Bureau.

Shintani, Tadahiko L. A. (2003). 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* (pp. 37–54). Research Institute for Languages and Cultures of Asia and Africa.

Shintani, Tadahiko L. A. (2014). *The Zayein language*. Linguistic survey of Tay cultural area (LSTCA) No. 102. Research Institute for Languages and Cultures of Asia and Africa.

Shintani, Tadahiko L. A. (2016). *The Nangki language*. Linguistic survey of Tay cultural area (LSTCA)

- No. 109. Research Institute for Languages and Cultures of Asia and Africa.
- Shintani, Tadahiko L. A. (2017). *The Blimaw language*. Linguistic survey of Tay cultural area (LSTCA) No. 112. Research Institute for Languages and Cultures of Asia and Africa.
- Shintani, Tadahiko L. A. (2018a). *The Yingtalay language*. Linguistic survey of Tay cultural area (LSTCA) No. 115. Research Institute for Languages and Cultures of Asia and Africa.
- Shintani, Tadahiko L. A. (2018b). *The Thaidai language*. Linguistic survey of Tay cultural area (LSTCA) No. 116. Research Institute for Languages and Cultures of Asia and Africa.
- Shintani, Tadahiko L. A. (2018c). *The Sonkan language*. Linguistic survey of Tay cultural area (LSTCA) No. 118. Research Institute for Languages and Cultures of Asia and Africa.
- Shintani, Tadahiko L. A. (2018d). *The Dosanbu Kayan language*. Linguistic survey of Tay cultural area (LSTCA) No. 120. Research Institute for Languages and Cultures of Asia and Africa.
- Shintani, Tadahiko L. A. (2020a). *The Thamidai language*. Linguistic survey of Tay cultural area (LSTCA) No. 126. Research Institute for Languages and Cultures of Asia and Africa.
- Shintani, Tadahiko L. A. (2020b). *The Pao language*. Linguistic survey of Tay cultural area (LSTCA) No. 131. Tokyo: Research Institute for Languages and Cultures of Asia and Africa.
- Shorto, H. L. (1962). *A dictionary of Modern Spoken Mon*. Oxford University Press.
- Shorto, H. L. (2006). *A Mon-Khmer comparative dictionary*. (Main editor: Paul Sidwell; Assisting editors: Doug Cooper & Christian Bauer) Pacific Linguistics.
- Sidwell, Paul (director) & Cooper, Doug (co-director). *SEalang Mon-Khmer etymological dictionary*. <http://sealang.net/monkhmer/dictionary/>; conferred on November 10th, 2022.
- Solnit, David B. (1997). *Eastern Kayah Li: Grammar, texts, glossary*. University of Hawai'i Press.
- Solnit, David B. (2001). Another look at Proto-Karen. Paper presented at the 34th International Conference of Sino-Tibetan Languages and Linguistics, Kunming.
- Solnit, David B. (2013). Proto-Karen Rhymes. Paper read at the 46th International Conference on Sino-Tibetan Languages and Linguistics, Dartmouth University.
- Turner, D. W. (2003). *Bananas: Response to temperature*. Agfact (NSW Agriculture, Australia), H6.2.6.

Author: Atsuhiko Kato

Postal address:

Institute of Cultural and Linguistic Studies

Keio University

2-15-45 Mita, Minato-ku, Tokyo 108-8345 Japan