Register complex and tonogenesis in Khmu dialects

SUWILAI Premsrirat
Mahidol University, Thailand

The prosodic variation in different Khmu dialects demonstrates the general trend of Southeast Asian languages which is heading towards tonogenesis and at the same time reflects the register complex phenomenon which is a typical Mon-Khmer prosodic characteristic. In this paper general information on the phonological systems of Khmu dialects in northern Southeast Asia is provided, followed by a discussion on the register complex and tonogenesis in Khmu.

I.   General information about Khmu in Southeast Asia

Khmu is a minority language spoken in a wide area of northern Southeast Asia, covering parts of Thailand, Laos, northwestern Vietnam and in Yunnan province of southwestern China. Khmu belongs to the Mon-Khmer sub-family of the Austroasiatic language family which is believed to form a substratum in some of the present languages of Southeast Asia. There is probably a small number of Khmu people living in Burma but no investigation has been done there. Some Khmu people have also migrated to Europe, the United States, and Australia as refugees from Laos after the communist liberation. The total Khmu population is more than 500,000 speakers. In Thailand, the Khmu are found mainly in the border area between Thailand and Laos, especially in Nan and Chiengrai provinces in the north. There are also Khmu in Kanchanaburi and Uthaithani in the central part of Thailand. The population of the Khmu in Thailand is around 10,000 speakers. In Laos, where the majority of the Khmu live, the Khmu are found mainly in the eight provinces of Luang Namtha, Udomsaj, Bokaew, Sayaburi, Phongsali, Luang Prabang, Hua Phan, and Chiengkhwang, and the population would not be less than 500,000. In Vietnam the Khmu population is around 40,000 speakers and they are found in Điện Biên Phù and Sơn La of Lai Châu province and in Nghệ An. In China, the Khmu are found in Muang La of Sipsongpanna, Yunnan and the population is about 2,000 speakers. The Khmu have normally settled in the same areas as various Tai speaking groups. The Khmu and the Tai, therefore, have exerted a lot of linguistic and cultural influences on each other. The Khmu have been influenced by the Tai culture in many aspects, such as styles of dress and housing. The Khmu women normally dress themselves similarly

to the Tai women who live in the same area. The outside structure of their houses looks similar, although the use and function of various parts of the house may be different. Khmu people in all countries can normally speak the Tai varieties of the area. On the other hand, the Tai and other small ethnic groups in some areas can also speak some Khmu. Inter-marriage occurs among these groups. Apart from Tai languages, the Khmu can also speak the national language of the country in which they have settled. Because of the close contacts between Khmu and Tai and between Khmu and the national languages, the Khmu language reflects the influences of Tai and the national languages, especially in the lexicon and certain phonological and grammatical patterns. As for the life style, in general, the Khmu still maintain their Khmu-ness. Khmu people in all areas still maintain the hand-tying ceremony for good fortune, spirit sacrificing ceremonies for healing sick people, jar wine drinking, bronze drum beating, traditional Khmu singing *tɔ̀:m*, and telling folktales and stories.

II. Khmu phonology

In all Khmu areas, the Khmu people are found living together with Tai-speaking people among whom we also find some Mon-Khmer speakers. For example, the majority of the Khmu in northern Thailand live in the same area as the northern Thai speakers. In Laos, the Khmu live among the Tai Lue, Tai Daeng, Phuen and Lao people. The Khmu in Vietnam live in the same area as the Tai Dam and Tai Siang. The Khmu in China live with the Tai Lue. Because of the close contact with Tai languages which have been tonal for a long time, the Khmu have unavoidably been exposed to tone influence. Khmu dialects in various parts of SEA have a similar basic structure though vocabulary and pronunciation vary. The most outstanding differences of Khmu dialects are the variation among the phonological systems. Khmu dialects in SEA can be phonologically divided into two main dialects: 'Eastern' and 'Western' Khmu. The Eastern Khmu (EK) consists of the Khmu dialects spoken in Hua Phan, Phongsali, Luang Prabang and Chiengkhwang in Laos, in Điện Biên Phủ, Sơn La, Nghệ An in North Vietnam, some villages in Sipsongpanna in China and some villages in Thailand. The Western Khmu (WK) consists of the Khmu dialects spoken in Udomsaj, Bokaew and Luang Namtha in Laos, Chiengrai and Nan in Thailand, and some villages in Sipsongpanna.
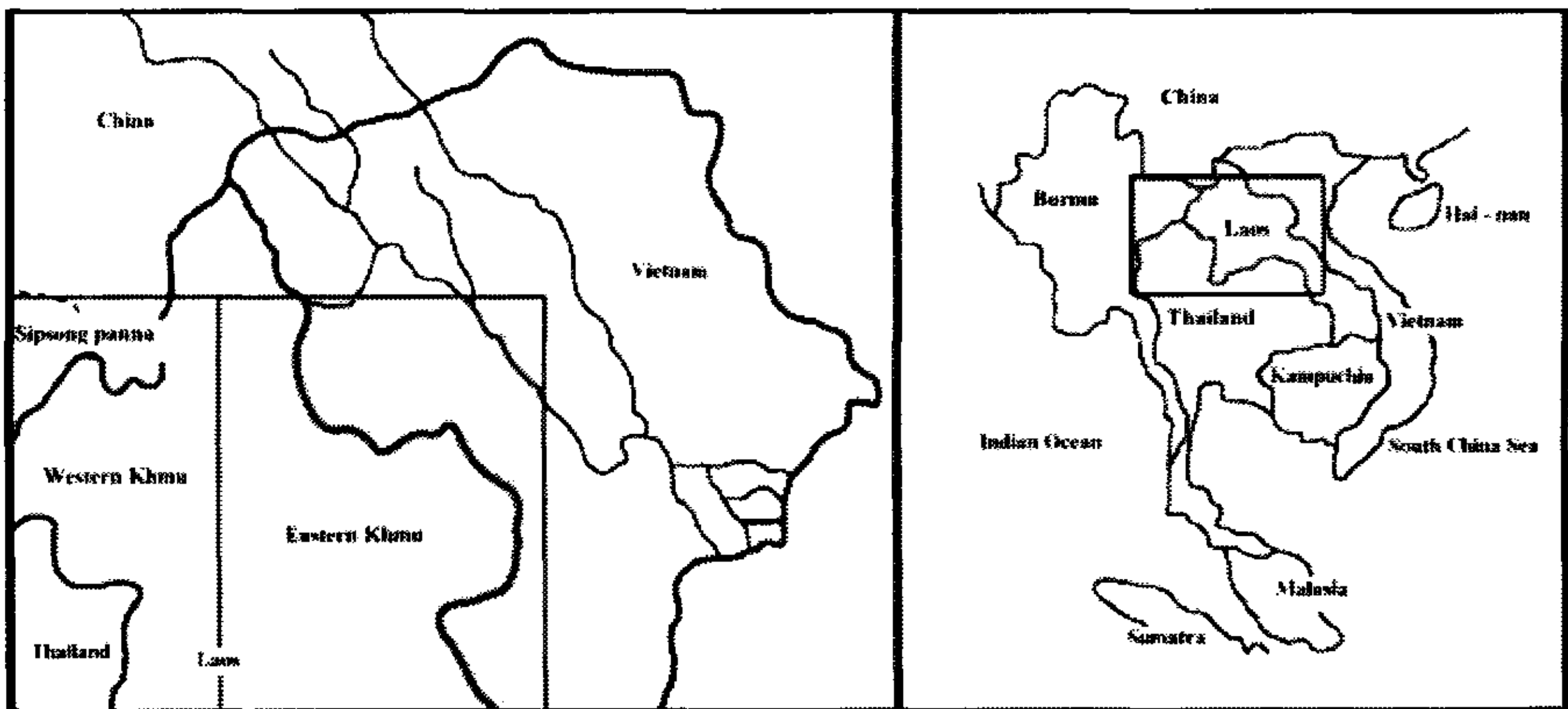
Map 1. Khmu speaking area showing the Western/Eastern division.

This paper is mainly based on the research done in seven Khmu speaking areas in Thailand, Laos, Vietnam and Yunnan (China) during 1996-2002[1]. Other researches on Khmu by Smalley[2] (1961), Svantesson[3] (1983), Suwilai[4] (1987), Cooper and Cooper[5] (1999), Mayuree (2003)[6] have been included where necessary. The main Khmu dialects mentioned here consist of: three Eastern dialects which are represented by speakers *ls*, *vt*, and *cl*; four Western dialects which are represented by speakers *lk*, *tw*, *ct* and *lb*. The dialects of these speakers are as follows:

ls –	the Khmu dialect spoken in Hua Phan, Laos
vt –	the Khmu dialect spoken in Nghệ-An, Vietnam
cl –	the Khmu dialect spoken in Pung Soa village, Sipsongpanna, Yunnan, China
lk –	the Khmu dialect spoken in Nalae, Udomsaj, Laos
tw –	the Khmu dialect spoken in Huay Yen village, Chiengkhong district, Chiengrai province, Thailand
ct –	the Khmu dialect spoken in Om Kae village, Sipsongpanna, Yunnan, China
lb –	the Khmu dialect spoken in Muang Hun, Udomsaj, Laos

---

[1]This research was granted by SEASREP. The result was published as a Khmu Series of Thesaurus and Dictionaries, namely Thesaurus of Khmu Dialects in SEA, Dictionary of Khmu in China, Dictionary of Khmu in Laos, Dictionary of Khmu in Vietnam and Dictionary of Khmu in Thailand.

[2]Smalley (1961) studied a Khmu dialect spoken in the Luang Prabang area. This dialect belongs to the *EK*.

[3]Svantesson (1983) studied a Khmu dialect spoken in the Luang Namtha area. This dialect belongs to the *WK*.

[4]The author first started working on a Khmu dialect in north Thailand in 1977. This dialect belongs to the *WK*.

[5]Cooper and Cooper (1999) studied a Khmu dialect spoken in Huay Co village, Wiengkaen district, Chiengrai province, Thailand. This dialect belongs to the *WK*.

[6]Mayuree (2003) studied a Khmu dialect (Khmu Kwaen) spoken in the Luang Namtha area. This dialect belongs to *WK*.

1) Segmental and suprasegmental inventories in Khmu

The EK dialects of *ls, vt, cl* are phonologically uniform. They have no tone and no register but have a rich consonant inventory. They have a full set of contrastive voiced and voiceless initial consonants. The WK dialects vary in their prosodic features. Some WK dialects have a lax-tense register contrast *(lk, tw)*, whereas some have tonal contrast *(ct, lb)*. In the following section the phonological systems of the WK and EK dialects are presented.

1.1 Initial consonants

The WK have 22 initial consonants, whereas the EK have 36 initial consonants. The difference between the two sets of initial consonants is related to the suprasegmental system of the WK dialects to be discussed in section 1.5.

A: Eastern Khmu consonant phonemes						B: Western Khmu consonant phonemes				
p	t	c	k	ʔ		p	t	c	k	ʔ
p^h	t^h	c^h	k^h			p^h	t^h	c^h	k^h	
b	d	ɟ	g			(f)*	s			h
(f)*	s			h		m	n	ɲ	ŋ	
m	n	ɲ	ŋ			ʔm/ʔb	ʔn/ʔd			
^{h}m	^{h}n	^{h}ɲ	^{h}ŋ			l				
ʔm	ʔn	ʔɲ				r				
l						w		j		
^{h}l										
r										
^{h}r										
w		j								
^{h}w		^{h}j								
		ʔj								

f is a borrowed phoneme from Tai-Lao or Vietnamese

1.2 Final consonants

All Khmu dialects have the same system of 15 final consonants as indicated below. However, the phonetic realization of the final alveolar fricative in different dialects can vary between [s], [ç], and [h]. The *lb* dialect and some *ls* still keep the final fricative [s]. A detailed description of this variation is given in the description of the consonants below.

Back:mid, Text:dark :: mid-W:0.2, B-Peak:35 :: gamma:1, B:130, W:200

	-p	-t	-c	-k	-ʔ
		-s [s, ç, h]			-h
	-m	-n	-ɲ	-ŋ	
		-l			
		-r			
	-w		-j		

Examples

-p	kɛːp	[kɛːp]	'to pinch a thing (such as charcoal with tongs)'
	grɨp	[grɨp]	'to seize'
-t	snaːt	[sᵊnaːt]	'gun'
	glaːt	[glaːt]	'stop, already passed (storm)'
-c	pic	[pic]	'to throw away'
	rwəːc	[rəwəːc]	'woman's head scarf'
-k	pklaːk	[pⁱklaːk]	'to tell a lie'
	caːk	[tɕaːk]	'to tear'
-ʔ	tɔʔ	[tɔʔ]	'to dip in (curry or chili paste)'
	kmbraʔ	[kᵊmbraʔ]	'wife'
-s	khras	[khraç]	'to laugh'
-h	jɔh	[jɔh]	'to go'
	smŋəh	[sᵊmŋəh]	'to look up'
-m	ʔom	[ʔom]	'water'
	maːm	[maːm]	'blood'
-n	klɨn	[klɨn]	'smell'
	laːn	[laːn]	'to open (eyes)'
-ɲ	piɲ	[piɲ]	'to shoot'
	taɲ	[taɲ]	'to hit down (using big stick)'
-ŋ	kʰɨəŋ	[kʰɨəŋ]	'to dig'
	tmʰmɔːŋ	[tᵊmʰmɔːŋ]	'nail'
-l	kaːl	[kaːl]	'in front of'
	tmʰmɨl	[tⁱmʰmɨl]	'garbage'
-r	mar	[mar]	'snake'
	smpɔːr	[sᵊmpɔːr]	'tamarind'
-w	liw	[liw]	'to make (big rope)'
	pɛːw	[pɛːw]	'undeveloped (limbs)'
-j	taj	[taj]	'bag'
	knkuːj	[kᵊnkuːj]	'to be sad and subdued'

Suwilai Premsrirat 2004, "Register complex and tonogenesis in Khmu dialects", in *The Mon-Khmer Studies Journal*, vol. 34, pp. 1-17. (purl.org/sealang/suwilai2004register.pdf)

Back:mid, Text:dark :: mid-W:0.11, B-Peak:32 :: gamma:1, B:130, W:200

1.3 Consonant clusters

Many consonant clusters are found in Khmu. All consist of a stop consonant followed by a liquid *l* or *r*, or a labial approximant *w*:

bl-	blik	'chili'
br-	briʔ	'jungle'
pl-	plɔːŋ	'leg'
pʰr-	pʰriːt	'to shift slightly'
dr-	druːɲ	'termite'
ɟr-	ɟriəh	'to wash rice'
cʰr-	cʰraːp	'astride, straddle'
gl-	gleʔ	'husband'
gr-	grip	'to seize'
gw-	gwaj	'to call'
kr-	kruə	'to burn an animal skin'
kl-	kləʔ	'hair (on the head)'
kʰr-	kʰruʔ	'to order somebody to work'
kw-	kwir	'to stir'
kʰw-	kʰwɛːn	'to recover'

Apart from these examples, there are many disyllabic words whose presyllable has an initial stop consonant and *l* or *r* as a final consonant. The presyllable in disyllabic words is unstressed, and its vowel is just a transitional one. In natural speech it is deleted. The syllable final consonant *l* or *r* then is able to form consonant clusters as in the following examples.

cr-	cᵊrguːl ~ crguːl	'finger'		cᵊrkɔːj ~ crkɔːj	'a kind of turmeric'		
sr-	sᵊrmaʔ ~ srmaʔ	'sickness'		sᵊrwaːk ~ srwaːk	'to mix'		
kl-	kᵊlmeʔ ~ klmeʔ	'sugar cane'		kᵊljɔːŋ ~ kljɔːŋ	'to swim'		
kr-	kᵊrlɛh ~ krlɛh	'smart, clever'		kᵊrŋɛːc ~ krŋɛːc	'to poke (a finger)'		
	kᵊrwaʔ ~ krwaʔ	'sheath'		kᵊrɲaʔ ~ krɲaʔ	'resin used to catch animals'		
pr-	pᵊrɟol ~ prɟol	'heel'		pᵊrdin ~ prdin	'vertical line'		
	pᵊrnaːj ~ prnaːj	'skewer'		pᵊrloŋ ~ prloŋ	'door'		
tr-	tᵊrlic ~ trlic	'sprained'		tᵊrhuət ~ trhuət	'wrinkle'		

The presyllable can be heard even when the final consonant is not *l* or *r* but is any consonant which agrees in place of articulation with the final consonant of the main syllable, or is an open syllable for which the reduced vowel has not been completely deleted, and the initial consonant of the main syllable is a stop. Presyllables with final nasal consonants in which the reduced vowel is very short or even deleted have prominent syllabicity.

cᵊ-	cᵊnəːj ~ cnəːj	'soft back fin of fish'
cⁱN-	cⁱmpiəŋ ~ cmpiəŋ	'straw'
kᵃ-	kᵃdɔŋ ~ kdɔŋ	'jar'
kᵊN-	kᵊmbraʔ ~ kmbraʔ	'wife'
lᵃ-	lᵃwaːŋ ~ lwaːŋ	'sky'
lᵊN-	lᵊmdaːŋ ~ lmdaːŋ	'eggplant'
pᵊ-	pᵊnir ~ pnir	'wing'
pⁱN-	pⁱŋ moːj ~ pŋmoːj	'the day after tomorrow'
tᵊ-	tᵊlɔːj ~ tlɔːj	'banana'
tᵊN-	tᵊmbraʔ ~ tmbraʔ	'cooking stove'
hⁱN-	hⁱmɲɛŋ ~ hmɲɛŋ	'to listen'
hᵊr-	hᵊrlɔʔ ~ hrlɔʔ	'language, story'
sᵊ-	sᵊʔɔːŋ ~ sʔɔːŋ	'wood'
sⁱN-	sⁱmpɔːr ~ smpɔːr	'tamarind'
rᵊ-	rᵊbaːŋ ~ rbaːŋ	'gong'
rⁱN-	rⁱmheʔ ~ rmheʔ	'tendon'
rⁱt-	rⁱtmuːt ~ rtmuːt	'flat, wooden container for cooling cooked sticky rice'

1.4 Vowels

All Khmu dialects have similar vowel phonemes. There are 19 monophthongs and three diphthongs. Even though some Khmu dialects have developed contrastive voice qualities, the vowels are still the same. The mid-low central vowel also occurs in Khmu but only as long vowel ʌː.

Tongue position / Tongue level	Unrounded		Rounded			
	Front		Central		Back	
	short	long	short	long	short	long
High	i	iː	ɨ	ɨː	u	uː
Mid	e	eː	ə	əː	o	oː
Mid-low				ʌː		
Low	ɛ	ɛː	a	aː	ɔ	ɔː

The diphthongs iə, ɨə, and uə glide from high vowels to central vowels. There is no distinction between short and long diphthongs.

Back:mid, Text:dark :: mid-W:0.14, B-Peak:35 :: gamma:1, B:130, W:200

1.5 Suprasegmental phonemes

The EK dialects *(ls, vt, cl)* have no tone or register, but have a consonant voicing distinction in the initial consonants. Some WK dialects have registers (*lk, tw*) and some have tones *(ct, lb)*.

Table 1. Voiced/voiceless contrast in Eastern Khmu consonants.

1. Stops		2. Sonorants	
bak	'to ride'	ŋɔʔ	'to fear'
pak	'to break (wood)'	ʰŋɔʔ	'paddy rice'
buːc	'rice wine'	məh	'to be'
puːc	'to take off clothes'	ʰməh	'what'
bok	'to cut down a tree'	raːŋ	'flower'
pok	'to take a bite'	ʰraːŋ	'tooth'
buːm	'to chew'	laʔ	'to go out'
puːm	'to fart'	ʰlaʔ	'ear'
glaːŋ	'stone'	waʔ	'to chase'
klaːŋ	'eagle'	ʰwaʔ	'monkey'
ɟaŋ	'to weigh'	ɟaːŋ	'animal basket'
caŋ	'astringent'	ʰɟaːŋ	'female animal'

In the WK dialects the prosodic features are significant instead of the voicing contrast in the initial consonants of EK. The WK dialects can be divided into three types on the basis of voice quality and tone as follows:

1.   Voice quality contrast: *lk, tw*
2.   Tone contrast and aspirated initial stop consonant: *lb*
3.   Tone contrast: *ct*

The WK dialects with register voice quality contrast (*lk, tw*) are spoken in both Laos and Thailand. In Laos, in the northern part of Udomsaj, the dialect is called Khmu Lue and in the southern part of Udomsaj it is called Khmu Khroang. The Khmu varieties with register in northern Thailand are spoken in Chiengrai (cf. Suwilai 1993, Cooper 1999) and Nan (cf. Cholthisa 1988). In the WK dialects of *lk* and *tw*, lower pitch and lax/breathy voice quality correspond to voiced initial consonants in EK. In addition, in WK the higher pitch and clear tense voice quality correspond to voiceless initial consonants in EK. Even though the lax register in most WK dialects does not have the strong breathy voice quality that is found in other Mon-Khmer languages, such as Mon, Chong, Kuy, or Iduh, nonetheless, it is quite distinctive for the native speakers. The Khmu people explain the differences between lax and tense registers as differences between "big voice" and "small voice". When the initial consonant is a stop consonant the breathy voice

quality is quite clear and obvious. The tense register has clear voice and is loud and tense[7]. However, as for the words with sonorant initials, the voiced consonants *m, n, ŋ, r, l, w, j* in words with low tone correspond to voiced consonants in EK and the same set of voiceless consonants in words with high tone correspond to the voiceless consonants ^{h}m, ^{h}n, ^{h}ŋ, ^{h}r, ^{h}l, ^{h}w, ^{h}j in EK. The tone contrast paired with voiceless aspirated stop initials has been observed in the *lb* dialect. In all these dialects the contrast is between low and high tones. For the *ct* dialect, even though it is considered to be tonal, phonetically the lax voice quality is clearly heard in words with low pitch that occurs with initial stops. It is also noticeable that the *ct* and *lb* Khmu dialects which have a tone contrast have more monosyllabic words than the other Khmu dialects.

Table 2. Voice quality and tonal contrasts in Western Khmu dialects.

English gloss	Register contrast	Tonal contrast	
	lk, tw Khmu	*lb* Khmu	*ct* Khmu
'rice wine' 'to take off clothes'	puːc [pu̯ːc] púːc [pûːc]	p^{h}uːc [p^{h}ùːc] púːc [púːc]	puːc [pùːc] púːc [pûːc]
'to cut down a tree' 'to take a bite'	pok [po̯k] pók [pók]	p^{h}ok [p^{h}òk] pók [pók]	pok [pòk] pók [pók]
'to chew' 'to fart'	puːm [pu̯ːm] puːm [pûːm]	p^{h}uːm [p^{h}ùːm] púːm [púːm]	puːm [pùːm] púːm [pûːm]
'stone' 'eagle'	klaːŋ [kla̯ːŋ] kláːŋ [klâːŋ]	k^{h}laːŋ [k^{h}làːŋ] kláːŋ [kláːŋ]	klaːŋ [klàːŋ] kláːŋ [klâːŋ]
'to weigh' 'astringent'	caŋ [ca̯ŋ] cáŋ [câŋ]	c^{h}aŋ [c^{h}àŋ] cáŋ [cáŋ]	caŋ [càŋ] cáŋ [câŋ]
'to fear' 'paddy rice'	ŋoʔ [ŋo̯ʔ] ŋóʔ [ŋóʔ]	ŋoʔ [ŋòʔ] ŋóʔ [ŋóʔ]	ŋoʔ [ŋòʔ] ŋóʔ [ŋóʔ]
'to chase' 'monkey'	waʔ [wa̯ʔ] wáʔ [wáʔ]	waʔ [wàʔ] wáʔ [wáʔ]	waʔ [wàʔ] wáʔ [wáʔ]
'flower' 'tooth'	raːŋ [ra̯ːŋ] ráːŋ [râːŋ]	raːŋ [ràːŋ] ráːŋ [ráːŋ]	raːŋ [ràːŋ] râːŋ [râːŋ]
'to be' 'what'	məh [mə̯h] mə́h [mə̂h]	məh [mə̀h] mə́h [mə́h]	məh [mə̀h] mə́h [mə̂h]

Even though it is obvious that the *ct* and *lb* dialects have only two phonemic tones, words of five and six phonetic tones are actually found in *ct* and *lb*, respectively. In *ct* Khmu a word with a final stop or a fricative and low tone on the short vowel has three variants: level (unmarked), low (v̀), and half high (v̭) as in *k^{h}rip* [k^{h}rip ~ k^{h}rìp ~ k^{h}rip̭] 'to seize', *k^{h}leʔ* [k^{h}leʔ ~ k^{h}lèʔ ~

---

[7]According to Cooper and Cooper 1999 (unpublished paper) the Huay Jo Khmu spoken in Chiangrai demonstrates very strong breathy voice.

kʰlɛ̰́ʔ], 'husband', *jɔh* [jɔh ~ jɔ̰́h], 'to go'. A word with high tone in *lb* has the same variation as in the *ct* dialect. A word with a final sonorant and low tone normally has two variants: level and low pitch as in *ma:m* [maːm ~ màːm] 'blood', and *caŋ* [caŋ ~ càŋ] 'to weigh' *rwa:j* [rəwaːj ~ rəwàːj] 'tiger'. A word with a high tone on a short vowel and a final stop or a fricative has two variants: high (v́) and rising-falling (v̂) as in *sóʔ* [sóʔ] 'Khmu axe', *tíh ~ t̂ih* 'mushroom'. A word with a final sonorant normally has either high or rising-falling pitch, as in *kɔ́:n* [kɔ́ːn ~ kɔ̂ːn] 'child', *tʰmíl* [tʰmíl ~ tʰm̂il] 'garbage' *kntúər* [kntúər ~ kntûər] 'neck'. In *lb* Khmu, a word with a low tone and a final stop or a fricative has three variants: level (unmarked), low (v̀), half high (v̠), as in *kʰrip* [kʰrip ~ kʰr̀ip ~ kʰr̠ip] 'to seize', *jɛʔ* [jɛʔ ~ jɔ̀h ~ jɛ̰́ʔ] 'you (male)', *kʰleʔ* [kʰleʔ ~ kʰlèʔ ~ kʰlɛ̰́ʔ] 'husband', *mec* [mec ~ mèc ~ mḛ́c] 'to fear, feel', *jɔh* [jɔh ~ jèʔ ~ jɔ̰́h] 'to go'. The high tone has two variants: high and rising-falling as in *tíʔ* 'hard' and *t̂is ~ tís* 'mushroom'. A word with a final stop after long vowels has a falling-rising tone (v̌) as in: *kʰu:t* [kʰŭːt] 'to enter', *pʰle:k* [pʰlɛ̌ːk] 'eel', and *sna:t* [snǎːt] 'gun'. A word with final sonorant normally has level or low pitch, as in *ma:m* [maːm ~ màːm] 'blood', *jɔ:j* [jɔːj ~ jɔ̀ːj] 'balance'.

2) Syllables and word structure of Khmu

The phonological word in Khmu consists of one, two or three syllables. The stress occurs on the main final syllable of the word. The majority of words are monosyllabic or disyllabic (sesquisyllabic). The presyllable C(C)^V(C) of disyllabic words is unstressed and the vowel is just a transitional or reduced vowel. Examples from an EK dialect is provided.

Monosyllable	'C(C)V(C)	ʰreʔ	'plantation'
		gaːŋ	'house'
		tʰraːk	'water buffalo'
Disyllable	C(C)^V(C)' C(C)V(C)	lᵊ'ŋiɲ	'dark'
		kᵊr'huʔ	'to shake'
		hⁱm'ɲɛŋ	'to listen'
		hⁱm'maːl	'soul'
Trisyllable	C(C)^V(C) C^V(C)' C(C)V(C)	tᵊlaːm'paːm	'butterfly'
		kᵊldɔːɲ'deːɲ	'firefly'
		tɕᵊlɔːŋ'dɔːŋ	'grasshopper'

The syllable structure in Khmu is clearly related to the prosodic feature. In EK *(ls, vt, cl)* both initial and final consonants in the presyllable are still preserved, whereas in the WK dialects that developed a register contrast such as *lk, tw* the presyllable sometimes became syllabic and the WK dialects that developed a tone contrast especially in *ct* and *lb,* the entire presyllable or the initial consonant of the presyllable is deleted. The *ct* and *lb* Khmu dialects, therefore, have more monosyllabic words than the others.

EK		WK *lk, tw*		WK *ct, lb*		Gloss
hᶦmmaːl	~	rᵊmaːl	~	maːl		'soul'
kᵊndrɔːŋ	~	hᵊntrɔːŋ	~	ntrɔːŋ/trɔːŋ		'back part of the body'
hᵊrlɔʔ	~	hᵊrlɔ́ʔ	~	rlɔ́ʔ/lɔʔ		'language'
hᵊʔeʔ	~	hᵊʔéʔ	~	ʔéʔ		'firewood'
hᶦʔiər	~	hᵊʔiər	~	ʔiər		'chicken'
hᵘʔuʔ	~	hᵘʔúʔ	~	ʔúʔ		'bad smell'
hᶦʔɨr	~	hᶦʔɨr	~	ʔɨr		'fragrant'
rᶦŋkoʔ	~	ŋkó ʔ	~	kóʔ		'husked rice'
kᵊmʰmuʔ	~	kəmmúʔ	~	kəmúʔ		'Khmu language'

However there are some WK dialects that have developed a tone contrast but still keep presyllable (*sesquisyllabic*) as in Khmu Yuan (Svantesson, 1983, 1994) and Khmu Kwaen (Mayuree, 2003) spoken in the Luang Namtha area in Laos. Those examples appear in Khmu Kwaen as *hîmmaːl, k̂intrɔːŋ, hîrlɔʔ, hîʔéʔ, hîʔiər, hîʔúʔ, hîʔɨr, r̂iŋkoʔ, k̂immuʔ* respectively.
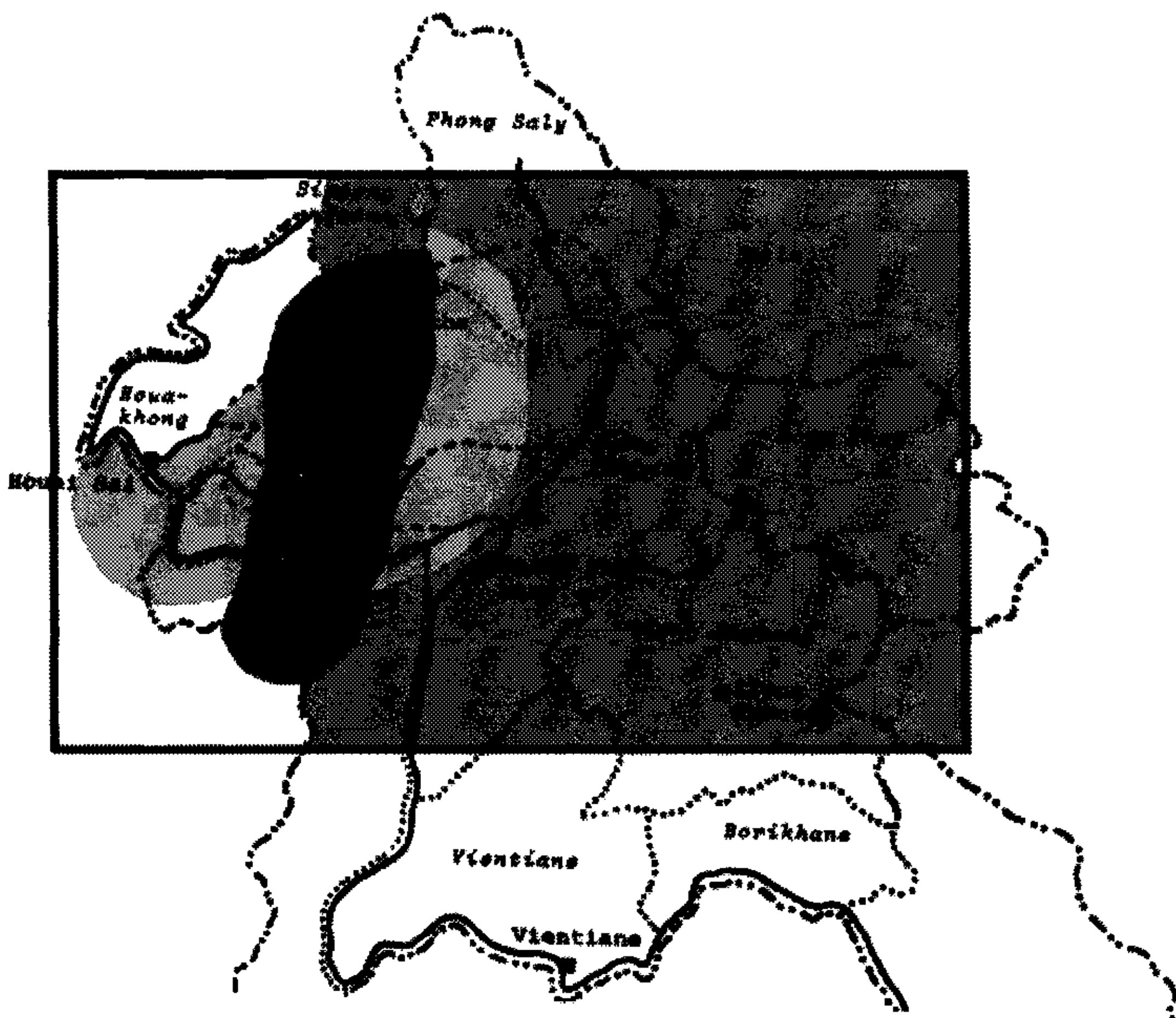
Map 2. Different prosodic features in different Khmu dialects.

III.   Register complex in Khmu dialects

The prosodic situation in Khmu can actually be described as "register complex" because it consists of phonetic features and their interrelationship.

Back:mid, Text:dark :: mid-W:0.2, B-Peak:44 :: gamma:1, B:130, W:200

They are consonant voicing, voice quality (phonation), aspiration, pitch level and contour, etc. The complex syllable prosody in Khmu dialects can be illustrated by an analogy with an amoeba as explained by Thomas (personal discussion).

An Amoeba is a one celled animal that has a changeable body outline and usually moves about by forming an extension of its body, a temporary projection called a pseudopod, or false foot, by which it moves about and captures food.

The changeable structure of Khmu syllable prosody among all dialects is similar to the structure of an amoeba. Different Khmu dialects select different phonetic features to highlight their prominent prosodic characteristics. The EK such as *vt, ls, cl* select the contrast in consonant voicing whereas the WK such as *lk* and *tw* select the register contrast and WK dialects such as *ct, lb* select the tone contrast. The *lb* dialect also selects consonant aspiration as part of its tone contrast. Interesting phonetic details of Khmu dialects with register and tone contrast were provided in section 1.5. The syllable structure of different Khmu dialects is changed according to the prosodic feature selected. The disyllabic structure is well preserved in EK whereas in the WK dialects that select phonation type (register contrast) keep the disyllabic word with syllabic nasal, and some of the dialects that select pitch level (tone contrast) as a distinctive feature have more monosyllabic words.

In the Khmu dialects, register complex, tone and phonation are equal partners and any one of these features can develop into a contrastive system. Tone is one part of the register complex, and phonation is another part of the register complex. The contrastive tones in some Khmu dialects have developed from a register complex (*lb, ct*). In other dialects a register complex has developed into a contrastive voice quality (*lk, tw*), or contrastive consonant voicing (*ls, vt, cl*), etc. So a register complex is sometimes a source for tonogenesis, as is often found in SEA. The contrastive consonant voicing (voiced/voiceless) in Khmu dialects is probably conservative and contrastive register/tone is innovative. However they all belong to the Khmu register complex.

On a larger scale, the amoeba analogy of the Khmu register complex is actually a replica of the structure within the Austroasiatic family. Different Mon-Khmer languages select different prominent prosodic features. Languages like Mlabri and Lavue have consonant voicing contrast; Nyah Kur, Chong, Mon and Kui have register contrast; Vietnamese, Samre and some Khmu dialects have tone contrast; and Kasong and So (Thavung) have both contrastive tones and phonation types.

At the same time on a smaller scale even some Khmu dialects in Thailand also demonstrate the register complex. Since most of the Khmu in Thailand moved from their original homes in various parts of northern Laos, their languages and cultures are different from one another. They are remnants of almost all major Khmu dialects in Laos. The variation of different Khmu villages is clear and can be seen in the suprasegmental features, initial and final consonants. Different suprasegmental

Suwilai Premsrirat 2004, "Register complex and tonogenesis in Khmu dialects", in *The Mon-Khmer Studies Journal*, vol. 34, pp. 1-17. (purl.org/sealang/suwilai2004register.pdf)

Back:dark, Text:dark :: mid-W:0.31, B-Peak:37 :: gamma:1.4, B:130, W:200

features, consonant voicing contrast, voice quality contrast and high and low pitch contrast have been employed by different Khmu dialects in Thailand. Apart from that, in the villages where there is a lot of contact with Thai speaking people the Khmu language spoken by the young (under 40 years of age) is also different from the language spoken by the old (over 40 years of age). The situation in Thailand is different from the Khmu language situation in other countries in terms of fast changing and shifting process. Examples of variation are given below. It is also noticeable that this type of Khmu dialect is also developing a system of more than two tones.

Variation of Khmu among the old and the young in Thailand.

- Change from disyllabic (sesquisyllabic) to monosyllabic word:
  cmrɔʔ → mrɔʔ → rɔʔ   'man'
  cntáh → tah   'prawn'

- Simplification of consonant clusters:
  pléc/péc → pét   'to clean the internal organs of animal'
  pʰríʔ/pʰlíʔ → pʰíʔ 'peppery hot'
  kləʔ → kəʔ   'hair'

- Loss of lateral articulation leaving only voicelessness: $^h l \sim h$:
  ʰráːŋ → ʰláːŋ → háːŋ   'tooth'
  ʰréʔ → ʰléʔ → héʔ   'field, farm'

- Final –r varies between final –l and –j:
  mar → mal → maj   'snake'
  hʔiər → ʔiəl → ʔiəj   'chicken'
  kntúər → ntúəl → ntúəj   'neck'
  húːr → húːl → húːj   'rotten'

IV.   Tonogenesis in Khmu dialects

From the point of view of tonogenesis, the synchronic differences among the Khmu dialects in Southeast Asia seem to reflect different stages of tonogenesis. As a living language, Khmu shows the different diachronic stages of tone development, from the beginning to the innovative stages. The EK dialects (ls, vt, cl)[8] represent the beginning stage. It has the voiced/voiceless distinction in the syllable-initial consonants. The development of tones in WK originated when the voiced/voiceless contrast in the consonantal system in EK broke down and the languages were forced to exploit prosodic differences for

---

[8]This type of Eastern Khmu dialect includes the Khmu variety of Luangprabang area described by Smalley (1961) and also the Khmu variety spoken in some villages in Nan, Lampang and Uthaithani of Thailand described by Suwilai (2002).

contrastive purposes. The initial consonant of the syllable conditions the pitch height of the syllable. However, in the case of Khmu the voiced initial consonant not only provokes low pitch in some dialects, but it also causes lax/breathy voice quality. The voiceless initial consonant provokes tenseness and higher pitch of the syllable as seen in *lk* and *tw*[9]. WK *lb* also represents the stage where the historically voiced initial has lost its voicing and has become a voiceless initial with aspiration and low pitch when the initial consonant is a stop consonant. This dialect has distinctive tones[10]. In the same situation, WK *ct*[11] has the voiceless initial and distinctively lower pitch. On the other hand, the voiceless initial consonant co-occurs with higher pitch (high rising). However, it is noticeable that even the *ct* dialect can be described as having a tone contrast, but phonetically when the word with low pitch occurs with initial stop the lax voice quality is still slightly heard. The following tables show how the different Khmu dialects are at different stages of tonogenesis.

The lax/breathy register in *lk* and *tw* is rather weak especially when it occurs in a word with initial sonorants but the tense register is rather strong (to be a good Khmu speaker). In the phonemic transcriptions used here words with tense register are therefore symbolized as (v́), while the low lax breathy register is unmarked.

Table 3. Stages of Tonogenesis: words with initial stop.

English gloss	1. Beginning stage: *ls, vt, cl* Voicing contrast	2. Intermediate stage: *lk, tw* Register contrast	3. Final stage:	
			lb Tone contrast with asp.initial	*ct* Tone contrast
'rice wine' 'to take off clothes'	buːc [buːc] puːc [puːc]	puːc [pṳːc] púːc [pûːc]	pʰuːc [pʰùːc] púːc [púːc]	puːc [pùːc] púːc [pûːc]
'to cut down a tree' 'to take a bite'	bok [bók] pok [pók]	pok [po̤k] pók [pók]	pʰok [pʰòk] pók [pók]	pok [pòk] pók [pók]
'to chew' 'to fart'	buːm [buːm] puːm [puːm]	puːm [pṳːm] púːm [pûːm]	pʰuːm [pʰùːm] púːm [púːm]	puːm [pùːm] púːm [pûːm]
'to weigh' 'astringent'	ɟaŋ [ɟáŋ] caŋ [cáŋ]	caŋ [ca̤ŋ] cáŋ [câŋ]	cʰaŋ [cʰàŋ] cáŋ [cáŋ]	caŋ [càŋ] caŋ [câŋ]
'stone' 'eagle'	glaːŋ [glaːŋ] klaːŋ [glaːŋ]	klaːŋ [kla̤ːŋ] kláːŋ [klâːŋ]	kʰlaːŋ [kʰlàːŋ] kláːŋ [kláːŋ]	klaːŋ [klàːŋ] kláːŋ [klâːŋ]

---

[9]This type of Khmu dialect is found in villages in Chiengrai province of Thailand along the Mekhong River.

[10]This type of Khmu dialect is found in some villages in Nan province of Thailand where the development of monosyllabic word structure is more radical.

[11]This type of Khmu dialect would include the Khmu Yuan which is the Khmu dialect described by Svantesson (1983) though there are some differences.

Table 4. Stages of Tonogenesis: Words with initial sonorant.

English gloss	1. Beginning stage: *ls, vt, cl* Voicing contrast	2. Intermediate stage: *lk, tw* Register contrast	3. Final stage: *ct, lb* Tone contrast
'to fear'	ŋɔʔ [ŋɔʔ]	ŋɔʔ [ŋɔ̰ʔ]	ŋɔʔ [ŋɔ̀ʔ]
'paddy rice'	ʰŋɔʔ[ʰŋɔʔ]	ŋɔ́ʔ [ŋɔ́ʔ]	ŋɔ́ʔ [ŋɔ́ʔ]
'to chase'	waʔ [waʔ]	waʔ [wa̰ʔ]	waʔ [wàʔ]
'monkey'	ʰwaʔ[ʰwaʔ]	wáʔ [wáʔ]	wáʔ [wáʔ]
'flower'	raːŋ [raːŋ]	raːŋ [ra̰ːŋ]	raːŋ [ràːŋ]
'tooth'	ʰraːŋ [ʰraːŋ]	ráːŋ [râːŋ]	ráːŋ [ráːŋ]

In summary, the combination of the distribution of prosodic differences and the proposed development of tone in some Khmu dialects supports Haudricourt's hypothesis on the development of Vietnamese tones (1954) and Matisoff's tonogenesis (1998). At the same time, it reveals the important stage prior to the emergence of tone, namely, the register complex, which is characteristic of most Mon-Khmer languages today. Tones developed in Khmu when the voicing contrast in the consonantal system (as in EK: *ls, vt, cl*) decayed and the languages were forced to exploit prosodic differences for contrastive purposes (as in WK: *lk, tw, ct, lb*). The initial consonant conditions the pitch height. However, in the case of Khmu not only does the voiced initial consonant provoke a low pitch, in some dialects it also causes a lax/breathy voice quality (*lk, tw*). Some dialects such as *lb* and *ct* dialects have clear voice quality. However, in the *lb* dialect aspiration also occurs with the stop consonant which is believed to be a stage that developed after the register contrast. Instead of having voice quality, it has aspiration. There is research evidence that a Khmu dialect of this type in Thailand had the voice quality before (Preedaporn, 1987), but now the voice quality has disappeared and has become a pure tonal contrast. For the *ct* dialect which is considered to have a tone contrast, it still has the phonetic characteristic of the slightly lax voice quality in stop consonants in words with low pitch while the voiceless initial consonants provoke higher pitch. Apart from that, the *lb* and *ct* dialects seem to be developing tone systems with five or six tones as mentioned in section II. It is obvious that these two dialects with tone contrast have a more monosyllabic word structure than other dialects. The evidence of the pure tone *lb* dialect that is spoken in some villages in Nan (Thailand) even demonstrates that disyllabic words can all become monosyllabic by deleting the presyllable. From my own observation, this phenomenon first started with young speakers and now most speakers of all age groups use mainly monosyllabic words.

V. Conclusion

From the study of the complex prosodic situation in Khmu dialects described above, it would be more appropriate to call the prosody of Khmu dialects in SEA as "register complex". However, the study also demonstrates that Khmu conforms to the general trend of SEA languages that are heading toward tonogenesis.

REFERENCES

Cholthisa Ong-arj. 1988. "A phonology of Khmu Lue at Nan province." M.A. thesis, Institute of Language and Culture for Rural Development, Mahidol University.

Cooper, Michael and Susan Cooper. 1999. "A preliminary phonological analysis of Huay Jo Khmu." A paper presented at *SEALS IX Conference*, Berkeley, USA.

Diffloth, Gérard. 1974. "Austro-Asiatic languages." *New Encyclopaedia Britannica* 2:480-484.

Egerod, Søren. 1971. "Phonation types in Chinese and South East Asian languages." *ALH* XIII 2:159-171.

Haudricourt, A.G. 1954. "De l'origine des tons en vietnamien." *JA* 242:69-82.

Henderson, E.J.A. 1982. "Tonogenesis: some recent speculations on the development of tone." *TPS* 1-24.

Huffman, Franklin E. 1976. "The register problem in fifteen Mon-Khmer languages." *AS* Part I. Honolulu: University of Hawaii Press.

Kaji, Shigeki (ed.). 1999. *Proceedings of the Symposium Cross-Linguistic Studies of Tonal Phenomena: Tonogenesis, Typology, and Related Topics*. Institute for the Study of Languages and Cultures of Asia and Africa (ILCAA), Tokyo University of foreign Studies.

Kaji, Shigeki (ed.). 2001. *Proceedings of the Symposium Cross-Linguistic Studies of Tonal Phenomena: Tonogenesis, Japanese Accentology, and Other Topics*. Institute for the Study of Languages and Cultures of Asia and Africa (ILCAA), Tokyo University of Foreign Studies.

Lindell, Kristina, Jan-Olof Svantesson and Damrong Tayanin. 1981. "Phonology of Kammu dialects." *CLAO* 9:45-71.

Matisoff, James A. 1998. "Tibeto-Burman tonology in an areal context." In Shigeki Kaji, ed., *Proceedings of the Symposium 'Cross-Linguistic Studies of Tonal Phenomena: Tonogenesis, Typology, and Related Topics'*, pp. 3-32, Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Matisoff, James A. 1973. "Tonogenesis in Southeast Asia." In Larry M. Hyman ed., *CTT*, Southern California Papers in Linguistics, No.1.

Mayuree Thawornpat. 2003. *Khmu Kwaen.* (unpublished manuscript)

Phu Van Han, Edmondson, J. and Kenneth Gregerson. 1997. "Eastern Cham as a tone language." *MKS* 20:31-43.

Pornsawan Ploykaew. 2001. "Samre grammar" Ph.D. dissertation, Institute of Language and Culture for Rural Development, Mahidol University.

Preedaporn Sisakorn. 1987. "The sound system of Khmu at Nampan in Nan." M.A. thesis, Institute of Language and Culture for Rural Development, Mahidol University.

Proschan, Frank. 1989. "Khmu verbal art in America: The poetics of Khmu verse." Ph.D. dissertation, The University of Texas at Austin.

Smalley, William. 1961. *Outline of Khmu Structure.* New Haven: American Oriental Society.

Suksawang Simana, Somsaeng Sayavong and Elizabeth Preisig. 1994. *Kmhmu-Lao-French-English Dictionary.* State Printing Enterprise, Lao P.D.R.

Back:dark, Text:dark :: mid-W:0.27, B-Peak:35 :: gamma:1.4, B:130, W:200

Sunee Kamnuansin 2002. "Kasong syntax: The study at Danchumphon sub-district, Borai district, Trat province." M.A. thesis, Institute of Language and Culture for Rural Development, Mahidol University.

Suwilai Premsrirat. 2002. *The Thesaurus and Dictionary Series of Khmu Dialects in Southeast Asia.* Mon-Khmer Studies-Mahidol. University Special Publication No. 1

Suwilai Premsrirat. 2001. "Tonogenesis in Khmu dialects of Southeast Asia." *MKS* 31:47-56.

Suwilai Premsrirat. 1997. "Phonological variation in Khmu of Northern Thailand." *MKS* 29:57-69.

Suwilai Premsrirat. 1996. "Phonological characteristics of So (Thavung), a Vietic language of Thailand." *MKS* 26:161-178.

Suwilai Premsrirat. 1993. *Thai-Khmu-English Dictionary.* Institute of Language and Culture for Rural Development, Mahidol University.

Suwilai Premsrirat. 1987. *Khmu, A Minority Language of Thailand.* (A Khmu Grammar and A Study of Thai and Khmu Cutting Words) Papers in South-East Asian Linguistics No. 10, Pacific Linguistics Series A-No. 75.

Suwilai Premsrirat. 1986. "Aspect of inter-clausal relations in Khmu." In Davidson ed, *Essays in Honour of H.L Shorto,* London, School of Oriental and African Studies.

Suwilai Premsrirat. 1984. "World view in Khmu." *JTU* 15.1:94-113.

Svantesson, Jan-Olof. 1983. *Kammu Phonology and Morphology.* Travaux de l'Institut de Linguistique de Lund XVIII.

Svantesson, Jan-Olof et al. 1994. *Khmu-Lao Dictionary.* Ministry of Information and Culture Vientiane.

Institute of Language and Culture
for Rural Development
Mahidol University
Phuttha Monthon, Salaya
Nakhon Pathom 73170
THAILAND
<lcspm@mahidol.ac.th>

Back:mid, Text:dark :: mid-W:0.21, B-Peak:32 :: gamma:1, B:130, W:200