

'Word Combination' in Lai¹

George Bedell
International Christian University

The Lai language is written with the Roman alphabet according to a system originally introduced by British colonial officers in the late nineteenth century and considerably modified by American missionaries in the first half of the twentieth. According to CACC Adult Literacy Project (1998), it may be summarized as follows².

Syllable Initial Stops and Affricates³:

p	t	c	tl	t̚	k
ph	th	ch	thl	t̚h	kh
b	d				

Syllable Initial Fricatives:

f	s				h
v	z				

Syllable Initial Sonorants:

m	n		l	r	ng
hm	hn		hl	hr	hng

Vowels:

a	e	i	o	u	aw
---	---	---	---	---	----

Diphthongs:

	ai	ao	au	awi
	ei	eo	eu	
ia		io	iu	
	oi			
ua	ui	uo		

Triphthongs:

iai	iao	iau
iei		
uai	uao	uau
uei		

Syllable Final Stops:

p	t		k	h
---	---	--	---	---

Syllable Final Sonorants:

m	n	l	r	ng
mh	nh	lh	rh	ngh

This orthographic system provides a way to represent most of the distinctive phonological contrasts in Lai. It is likely that it is in some respects exuberant (in particular in vowel combinations), and does not incorporate some restrictions on distribution. It does not represent two features of Lai: vowel length and tone. I have not been able to study these matters adequately and will make no claims about them here; but it seems to me that these weaknesses of the orthography are intimately related to variation within the Lai speaking population. Some people advocate representing vowel length by doubling a vowel letter. But it is not clear (at least to me) how large a functional load the distinction carries for those who have it, though there are minimal pairs. And there seem to be speakers who do not have it. There is a keen interest among Lais to standardize the orthography but no consensus as to exactly how the variability should be resolved. The most salient controversy is not any of the matters just referred to, but rather what Lais call *biafang komh*, usually rendered into English as 'word combination'.

Lai belongs to the type of language (classically exemplified by Chinese) in which each separately meaningful linguistic element (often termed 'word', but corresponding more closely to 'morpheme' as used by linguists) is a phonological syllable⁴. The orthography as presented above makes some use of syllable structure: the letter 'h' represents aspiration (either alone or combined with a voiceless obstruent) or voicelessness (when combined with a sonorant) in syllable initial position, but glottalization (either alone or combined with a sonorant) in syllable final position. Together with the non-distinctive phonetic differences among other consonants in syllable initial versus syllable final position, a certain amount of ambiguity will be present in the orthography unless syllable boundaries are indicated. Not indicating syllable boundaries per se would probably not constitute a major impediment for Lai speakers, but it does highlight these properties of the orthography. 'Word combination' refers to writing syllables together without leaving a space between them, and results in the disappearance of syllable boundaries. Virtually all users of written Lai have practised 'word combination' to some degree, but there is wide variation in where it is, or should

be, applied.

To get a more concrete idea of what is involved here, consider two Lai translations of a Bible verse (Matthew 2:16). Each of (a) and (b) is followed by a fairly literal retranslation to compare with the English RSV given first.

Then Herod, when he saw that he had been tricked by the wise men, was in a furious rage, and he sent and killed all the male children in Bethlehem and in all that region who were two years old or under, according to the time which he had ascertained from the wise men.
(Revised Standard Version)

(a) Khi can ah Herad nih arfi a thiam mi hlennak kha a hngal tik ah amah nih a thin-a-hung ngaingai i Betlehem khua le a pawng kam hrawng khua chung, a ngakchia pa kum hni tang deu mi kha a that ter hna. Khi kum ngakchia pa kong kha arfi a thiam mi sin in a hngal. (Strait, 1950, pp. 3-4)

'At that time, when Herod realized the trick of the astrologers, he was very angry and had the male children under two years in Bethlehem and the nearby towns killed. He knew about male children of that age from the astrologers.'

(b) Herod nih, nichuahlei mifimhna nih an ka hlen, ti kha ahngalh tikah, a thin a hung ngaingai. Arfi a chuahnak kong nichuahlei mifim hna nih an chim ning khan a caan cu a tuak i Bethlehem khua le a pawngkam i a um mi khua hna chung i ngakchia-pa kum hnih ri in a tanglei paoh kha thahdih hna awk ah nawlbia a pek hna.

(Van Bik, 1978, pp. 4-5^s)

'When Herod realized, 'the wise men from the East have tricked me,' he was very angry. He calculated the time the wise men from the East had told him concerning the appearance of the star, and he ordered that boys under two years in Bethlehem and the nearby towns be all killed.'

In the English version, there are 54 words and 66 syllables. The Lai versions are not significantly different: (a) has 57 words and 69 syllables; (b) has 65 words and 86 syllables. Unlike Lai, English has morphemes which do not comprise separate syllables: noun plurals like *men* or *years*, or verb past tenses like *saw* or *tricked*. If syllables per morpheme rather than per word were calculated, English might in fact be less compact than Lai.

Linguists often assume that any language has both morphological and syntactic structure; the interface between these levels is the word. Morphology is concerned with how morphemes are combined to create words, and syntax with how words are combined to create sentences. There may be differences in the relative complexity of the two levels across languages, but the distinction between them is taken to be clear. An ideal orthography based on the Roman alphabet will represent the morphemes which make up a word contiguously, but a space will separate each word in a sentence. The problem of 'word combination' in Lai is of concern to Lais who wish to write their language. But it is of interest to linguists because of its bearing on the distinction between morphology and syntax. Lai verbs are accompanied by a rather large number of particles indicating (among other things) tense and aspect, subject and object agreement, directionality, valence, and nominalization. One can speak of a 'verb complex' to maintain neutrality as to which (if any) of these particles are affixes (belonging to morphology) and which (if any) are independent adverbs (belonging to syntax). To a lesser degree, there is also a 'noun complex' consisting of a noun accompanied by particles indicating (among other things) number, case, gender, demonstratives, quantifiers and postpositions.

As extreme positions in the 'word combination' debate, we can imagine treating each particle in the verb or noun complex as an independent word, or treating each verb or noun complex as a single word. The first position is illustrated by rewriting (b) as (b')⁶ and the second by rewriting it as (b''), though in this case there might be different views of where the 'complex boundaries' are.

(b') Herod nih, ni chuah lei mi fim hna nih an ka
hlen, ti kha a hngalh tik ah, a thin a hung ngai ngai.
Ar fi a chuah nak kong ni chuah lei mi fim hna nih
an chim ning khan a caan cu a tuak i Bethlehem

khua le a pawng kam i a um mi khua hna chung i
ngak chia pa kum hnih ri in a tang lei paoh kha thah
dih hna awk ah nawl bia a pek hna. (83 words)

(b") Herodnih nichuahlei mifimhnanih
ankahlentikha ahngalh tikah, athinahung ngaingai.
Arfi achuahnnakkong nichuahlei mifimhnanih
anchimningkhan acaancu atuaki Bethlehem khuale
apawngkami aummi khuahnachungi ngakchiapa
kumhnhiriin atangleipaohkha thahdihhnaawkah
nawlbua apekha. (26 words)

Lais often complain about reading difficulty when 'word combination' differs from what they advocate or are used to, but the extreme diffuse alternative (b') seems considerably more readable and natural than the extreme compact alternative (b").

Lai words like *arfi* 'star' or *ngakchia* 'child' in (a) and (b) are unusual from the point of view of the 'monosyllabic' character of Lai mentioned above. Whatever their etymology may be, they appear to be single morphemes (and ipso facto single words) in current Lai. In fact no one advocates not doing 'word combination' in such cases. This is in contrast to the larger *ngakchia pa* 'male child' or *pawng kam* 'near at' in (a) versus *ngakchiapa* 'boy' or *pawngkam* 'vicinity' in (b). In these latter cases there is variation; they differ from the previous ones in that the component syllables, have their own meanings and can be used independently of these combinations (and are therefore distinct morphemes). A short manual *Laica Trialning* (CCLS, no date) has been devoted to guidelines for when 'word combination' should be used and when not. It contains the following general statements.

Laica Trialning Tawhfun (pp. 33-4)

1. Sullam pakhat a ngeimi biafang paoh cu komh
ding. (One word, one meaning)
eg. *Sihmanhsehlaw, lunglawmnak*.
2. Cafangaw pahnih aa chan tikah desh (-) le
apostrophe (') te hna hman lo ding a si.
eg. *Ca-uk* cu *cauk* tiah
Ngandam te'n cu *ngandam tein* tiah
trial ding a si.

3. Biatlang aa thawknak cafang le min paoh cu cafang ngan in trial ding. Cun inverted commas (" ... ") aa thawknak cafang cu a ngan in trial ding a si.
eg. *Ka pa nih, "Rawl tampi in ei," tiah a ka ti.*
4. Hla cuangmi min paoh cu aa kom in trial i a thawknak lawng cafang ngan in trial a si lai.
eg. 1. *Vanteek khuari na tor hluarmar cu, Mangling le a nau Tinchum than tawn kha bang, An siang cung i thirthluan phir le cuanki aw hlei, An pauter tawn kan mual zung in.*
2. *Khua khum pau lai e, ra tim thai cel hlah, Mang cia ei Tinkhar lai ni thei lo, Bualkung keimei dawh ngun Lalling kha, Sangchum thlabang ceu ti lai ri lo.*

[Principles of Lai Orthography]

1. All syllables which express a single meaning should be combined.
eg. *sihmanhsehlaw* 'but', *lunglawmnak* 'happiness'
2. When two vowel letters come together no dash or apostrophe should be used.
eg. *ca-uk* 'book' should be written *cauk*, and
ngandam te'n 'in a healthy manner' should be written *ngandam tein*
3. A capital letter should be written at the beginning of each sentence and each name.
Also the first letter in a quotation should be a capital.
eg. *Ka pa nih, "Rawl tampi in ei," tiah a ka ti.*
'My father told me, "Eat a lot".'

4. Names in songs are to be combined, with a capital letter at the beginning only.⁷]

The major principle bearing on 'word combination' is the first, which specifies that morphemes which have a unified meaning should be written together (i. e. constitute a single word). The principle is vague and subject to diverse interpretations. Thus the example *sihmanhsehlaw* is a composite of *si* 'be', *hmanh* 'even', *seh* 'let it' and *law* 'and'. In that it corresponds to English *but*, we might say that it has a unified meaning; but it might also be taken as corresponding to something like *be that as it may*. There is no objective criterion for unified meaning (certainly *Laica Trialning* does not attempt to formulate one), and the temptation to compare Lai and English is sometimes overwhelming. To understand better both the problem and the proposed solutions, we need to look into the detailed cases considered.

A plausible candidate for morphological status is the agreement system of Lai, which consists of person and number particles which come both before and after the verb, as described in Bedell (1995). The first entry in *Laica Trialning* is *a*, which marks third person singular subject or possessor agreement.

- 1.1 *a* timi biafang cu amah lawngte a dir i
mincan sullam a ngeih ahcun komh lo
siseh.
eg. *a si*, *a tupa*.
- 1.2 *a* timi biafang cu amah lawngte a dir i
mincan sullam a ngeih lo ahcun komh
siseh.
eg. *asinain*, *asiloah*, *azeipaoh*, *abikin*,
apakhatnak. (p. 1)

[If *a* stands by itself with the meaning of a noun, it should not be combined.
eg. *a si* 'it is'; *a tupa* 'his grandson']

If *a* does not stand by itself with the meaning of a noun, it should be combined.
eg. *asinain* 'nevertheless'; *asiloah* 'otherwise'; *azeipaoh* 'whatever';

abikin 'particularly'; *apakhatnak*
'firstly']

In general this agreement marker is a separate word on semantic grounds: it is taken to be a pronoun, either subject or possessor. Only in frozen idioms (of which the corresponding English adverb or conjunction does not contain any pronoun) is it to be written as part of a larger word. The third person plural *an* is also mentioned.

- 1.8 *an* timi biafang cu thil pakhatkhat can i hman
tikah sullam dangte a ngeih
ahcun komh lo siseh.
eg. *an thra*, *an pa*.

- 1.9 *an* timi biafang cu thil pakhatkhat can i hman
tikah sullam dangte a ngeih lo ahcun komh
siseh.
eg. *anmah*, *annih*. (p. 2)

[If *an* has a distinct meaning in reference to several things, it should not be combined.

- eg. *an thra* 'they are good'⁸, *an pa* 'their father'

If *an* does not have a distinct meaning in reference to several things, it should be combined.

- eg. *anmah* 'they'; *annih* 'they']

Here, the agreement marker *an* is taken as a pronoun and a separate word. Since it does not occur in frozen idioms parallel to the cases cited for *a*, the contrast is with the third person plural independent pronouns. The parallel *amah* 'he/she' and *anih* 'he/she' could have been cited for the singular case as well. The independent pronouns represent a different type of frozen form.

The first person singular agreement marker *ka* is also mentioned.

- 8.1 *ka* timi biafang cu keimah tinak a si tikah
komh lo siseh.
eg. *Ka tha a ba*.

- 8.2 *ka* timi biafang cu amah lawngte a dir i sullam
a ngeih lo ahcun komh siseh.
eg. *innka*, *kutka*, *mahka*, *hika*, *khika*,
deika, *chuahka*, etc. (pp. 10-11)

[When *ka* is equivalent to *keimah* 'I', it should not be combined.

eg. *Ka tha a ba*. 'I am tired.'

If *ka* does not have meaning by itself, it should be combined.

eg. *innka* 'door'; *kutka* 'gate'; *mahka* 'here'; *hika* 'here'; *khika* 'there'; *deika* 'dawn'; *chuahka* 'just after birth', etc.]

The decision with *ka* is the same as with *a* or *an* above: if it is an agreement marker, it is a separate word. The examples given for combining it are of two types: there is a *ka* whose basic meaning is 'mouth', seen in the first two examples; another with locative meaning (either spatial or temporal), seen in the others. These may be related, but they are clearly unrelated to the first person singular agreement marker. With the plural *kan*, the discussion mirrors that of *an*, though sentence examples are given.

- 8.5 *kan* timi biafang cu kanmah timi sullam a ngeih
ahcun komh lo siseh.
eg. *Laimi kan si*. *Kan pa a ra*.

- 8.6 *kan* timi biafang cu amah lawngte a dir i sullam
a ngeih lo ahcun komh siseh.
eg. *Kanmah ta a si*. *Kannih cu kan*
kal cang lai. (p. 11)

[If *kan* has the meaning 'we', it should not be combined.

eg. *Laimi kan si*. 'We are Lais.'
Kan pa a ra. 'Our father is coming.'

If *kan* does not have meaning by itself, it should be combined.

eg. *Kanmah ta a si*. 'It is ours.'
Kannih cu kan kal cang lai. 'We

will go.']

No parallel statement is made for the second person subject or possessor agreement markers *na* and *nan*, nor is there any reference to the object agreement markers *ka*, *kan*, *in* or *hna*. There is mention of the imperative plural subject marker *u*.

- 17.1 *u* timi biafang cu plural a si i amah tein a dir
khomi a si caah komh lo siseh.
eg. *Kal u law, a thra lai.*
- 17.2 *u* timi biafang cu amah lawngte a dir i sullam
a ngeih lo ahcun komh siseh.
eg. *Amah nakin ka upa deuh.*
Ka unau pawl an tin cang.
(pp. 27-28)

[If *u* is plural and can stand alone, it should not be combined.

- eg. *Kal u law, a thra lai.*
'You had better go.'

If *u* has no meaning by itself, it should be combined.

- eg. *Amah nakin ka upa deuh.*
'I am older than he is.'
Ka unau pawl an tin cang.
'My brothers shared.']

The pattern here resembles that above for *ka*: when *u* is an agreement marker (but not taken as a pronoun) it is a separate word, but the alternative given represent a distinct morpheme *u* meaning 'elder brother'. *Laica Trialning* thus consistently regards Lai agreement markers as separate words, and there is little opposition to this among Lais.

A second general case of some interest involves the Lai postpositions *ah* 'at' or 'to' and *in* 'from'. These have a variety of extended uses.

- 1.3 *ah* timi biafang cu Noun le Verb he aa chan
ahcun komh lo siseh.
eg. *Haka khua ah ka tlung lai.* (N)
Kan kal ah a thra deuh lai. (V)

1.4 *ah* timi biafang cu Preposition paoh he komh siseh.

eg. *Inn chungah a lut.*
Cabuai chungah va chia. (p. 1)

[If *ah* accompanies a noun or a verb, it should not be combined.

eg. *Haka khua ah ka tlung lai.*
 'I will arrive at Hakha.'
Kan kal ah a thra deuh lai.
 'It will be better if we go.'

Ah should be combined with any preposition.

eg. *Inn chungah a lut.*
 'She entered the house.'
Cabuai chungah va chia.
 'Put it in the desk.']

Here *ah* as an independent postposition is taken as a separate word, but combines with *chung* 'inside'. *Chung* is taken to be a 'preposition', but is probably better analyzed as a relational noun something like English *inside of*. In the same way, *in* is taken to be a separate word when a postposition by itself, but to combine with *cung* 'topside'; this latter is parallel to *chungah* above.

7.2 *in* timi biafang cu Noun le Verb he aa chan ahcun komh lo siseh.

eg. *Haka in ka tlung.*
A leen in kan leeng hlah.

7.3 *in* timi biafang cu Noun le Verb biafang hna chim loin Adverb, Preposition he a dang biafang hna he komh siseh.

eg. *Lunglawm tein kan i tong.* (Adv)
Cabuai cingin a tla. (Pre)
Fangvoi nakin facang a thaw
deuh. (Conjunction)
 (p.10)

[If *in* accompanies a noun or a verb, it should not be combined.

eg. *Haka in ka tlung.*
 'I arrived from Hakha.'
A leen in kan leeng hlah.

'Don't visit us.'

If *in* accompanies an adverb, preposition, or anything other than a noun or verb, it should be combined.

eg. *Lunglawm tein kan i tong.*
 'We met with pleasure.'
Cabuai cugin a tla.
 'It fell off the desk.'
Fangvoi nakin facang a thaw
deuh.
 'Rice is tastier than millet.']

In also combines with the comparative *nak* (*ah* does not; *nak* is considered a conjunction), and with *te* (an adverb?) in the productive adverbial form. Some people write *lunglawmte in*. See also the example given under principle 2 above; explicit indication of contraction is frowned upon in written Lai.

A third case involves the relative marker *mi*.

10.8 *mi* timi biafang cu 'minung' tinak sullam a
 ngeih ahcun komh lo siseh.
 eg. *Mi zakhat an i pum.*

10.9 *mi* timi biafang cu Verb, Adjective le Adverb
 he aa chan ahcun komh siseh.
 eg. *A rami pahra an ti.* (V.)
Aa dawhmi pahnih an i tel. (Adj.)
A thaw lomi rawl cu ei hlah.
 (Adv.) (p. 17)

[If *mi* means 'person', it should not be combined.

eg. *Mi zakhat an i pum.*
 'A hundred people gathered.'

If *mi* goes with a verb, adjective or adverb, it should be combined.

eg. *A rami pahra an ti.*
 'Ten people came.'
Aa dawhmi pahnih an i tel.
 'The two pretty ones participated.'
A thaw lomi rawl cu ei hlah.
 'Don't eat food which doesn't taste good.']

Mi is an independent noun meaning 'human being' in 10.8; it occurs in many compounds, such as *mipa* 'man' or *mithra* 'good person'. 10.9 refers to the relative marker, which may be derived from this noun. A noun phrase like *aa dawhmi pahnih* 'the two pretty ones' is understood to have human reference, but there is no overt head noun (i. e. nothing like English *one*). In *a thaw lomi rawl* 'food which doesn't taste good' the head noun *rawl* 'food' is non-human. Here it is combined with the negative *lo*, which also appears in *Laica Trialning*.

9.14 *lo* timi biafang cu a tanglei bantukin hman
siseh.

eg. *A nu he an i lo.*
Amah cu a ra bal lo.

9.15 *lo* timi biafang cu aa changmi biafang he
komh tikah Noun a chuah ahcun komh
siseh.

eg. *Mithralo cu hawi an ngei kho lo.*
Zolo kan tu; Lailo le tulo zong kan
vat. (p. 15)

[*Lo* should be used as below.

eg. *A nu he an i lo.*
'He looks like his mother.'
Amah cu a ra bal lo.
'He hasn't arrived yet.'

When *lo* combines with another word to make a
noun, it should be combined.

eg. *Mithralo cu hawi an ngei kho lo.*
'Bad people have no friends.'
Zolo kan tu; Lailo le tulo zong kan
vat.
'We sow Zo fields; we clear Lai
fields and also forest
fields.']

The negative *lo* appears in the second example sentence in 9.14. The first contains a verb *lo* 'resemble', presumably unrelated. The negative *lo* does appear in compound nouns like *mithralo* 'bad person' in 9.15 (the other *lo* here is also negative). But the *lo* in the second example sentence is a third (unrelated) noun meaning 'field'.

Some people dislike 'word combinations' like *lomi* in 10.9, on the grounds that they are simply not semantically coherent words. This view is expressed forcefully by Antony Ngun Uk (1996).

CACC nih a phunphun in kan holh, kan catrialning, kan phunglam le kan hla te hna a thranchonak aa zuam i a rauh hlan ah kan catrialning hi sawi awk um lo in a thra cang lai ti in ka zumh. Asinain 'in' le 'mi' cu biakomh (affix) ah an hman i bia dang he an komh thluahmah tikah a rel zong a har i sullam zong biafang pakhat ca in a ngei rua lo ti in ka ruah. Abiana ah 'lo in' ti ko ding cu 'loin' an ti i, 'lo mi' ti awk cu 'lomi' ti an trial. Biafang pakhat cu amahte in a dir kho lai; sullam a ngei lai kan ti lioah, biafang pahnih a si ko mi 'loin' hi zeidang bia um lo in amah lawng bak in hmu u law, asiloah mi nih a sullam in hal hna sehlaw zeitindah nan lehpiak hna hnga? Cu bantuk thriamthriam in 'lomi' zong? Miphun dang nih an ka hal sual ahcun, "Hitihin cun ka hngal ve lo. Asinain catrialtu nih 'lo in' ti in a trial ahcun a sullam cu 'without' a si. Cun 'from the field' ti zong a si kho'", ka ti hna lai. (p. 19)

[I believe that CACC has been trying to improve our language, writing system, culture and songs in various ways and to perfect our writing system in the near future. However, I think when they use 'in' and 'mi' as affixes and combine them indiscriminately with other words, not only are they hard to read, but they don't have the meaning of a single word. For example, instead of 'lo in' they write 'loin' and instead of 'lo mi', they write 'lomi'. Though we say that one word can stand by itself and have meaning, if you see 'loin' which is itself two words without any other words as context, or someone asks you its meaning, how would you interpret it for them? And what about 'lomi'? If a foreigner should happen to ask me about it, I would say to them, "I don't know like this. But if the writer writes 'lo in' then it means 'without'. It could also mean 'from the field'".]

A fourth case involves interrogative constructions with *dah*.

4.1 *dah* timi biafang cu amah lawng a dir i sullam a ngeih lo ahcun komh siseh.

eg. *ahodah*, *zeidah*, *zeicaahdah*,
zeitindah

4.2 *dah* timi biafang cu, in, he, ah timi biafang he aa chan ahcun, komh lo siseh.

eg. *Khuazei in dah na rat?*
Aho he dah nan kal lai?
Khuazei ah dah nan inn a si? (p.7)

[If *dah* stands by itself and has no meaning, it should be combined.

eg. *ahodah* 'who?', *zeidah* 'what?',
zeicaahdah 'why?', *zeitindah*
'how?'

If *dah* accompanies *in* 'from', *he* 'with' or *ah* 'to, at', it should not be combined.

eg. *Khuazei in dah na rat?*
'Where did you come from?'
Aho he dah nan kal lai?
'Who are you going with?'
Khuazei ah dah nan inn a si?
'Where is your house?']

The syntax and semantics of *dah* has been discussed by Lehman (1995). When the quantifying element is directly followed by *dah*, or in small number of other cases (perhaps determined by where English has a single interrogative word) the components are combined as in 4.1; otherwise they are not. The list of elsewhere cases in 4.2 is far from complete; in particular head nouns may appear, as in the following cases.

aho nih dah 'by whom?'
zei cauk dah 'what book?'
zei bantuk thingkung dah 'what kind of tree?'

It appears that Lai interrogatives are noun phrases with full syntactic capability; but when this capability is not exercised, a morphological condensation takes place.

Interrogatives are not the only noun phrases like this in Lai. Demonstratives also can mark both the beginning and end of a noun phrase, as in 10.3.

10.2 *mah* timi biafang cu aa changmi biafang he
komh tikah sullam pakhat lawng a ngeih
ahcun komh siseh.
eg. *keimah*, *namgah*, *amah*

10.3 *mah* timi biafang cu amah lawng a dir i
sullam a ngeih ahcun komh lo siseh.
eg. *mah nu hi* (p.16)

[If *mah* has only a single meaning when combined with the word it accompanies, it should be combined.

eg. *keimah* 'I', *namgah* 'you', *amah*
'he/ she'

If *mah* has meaning by itself, it should not be combined.

eg. *mah nu hi* 'this woman']

Here *mah* is the NP initial demonstrative, and *hi* the NP final one. Though it is not explicitly mentioned in *Laica Trialning*, without a head noun, these two would combine.

mahhi 'this'

See the brief discussion of *mah* as a pronoun forming element in Lehman with Van Bik (1997). *Laica Trialning* also does not consider the other demonstratives *cu* 'that', *kha* 'that', *khi* 'that' and *hi* 'this', all four of which may occupy either position.

cu mipa cu 'that man'

Here too, if no head noun appears, a demonstrative pronoun-like combination is normal.

cucu 'that'

There are also a few other elements which are taken to enter into such morphological combinations.

cuticun 'in that way'
cuka ah 'there'

Here *ti* may be related to the verb *ti* 'say', and *ka* is the locative morpheme referred to under 8.2 above.

A final case, which is not discussed directly in *Laica Trialning* at all (presumably because there is no morpheme which characterizes it uniquely) is compound verbs like *lunglawm* 'be happy'. Lai is rich in such verbs, as illustrated in Van Bik (1998). In finite forms, the two components will be separated by at least an agreement marker and thus not combined.

Ka lung a lawm. 'I am happy.'

But when nominalized, for example, as in principle 1 above, they will be combined.

lunglawmnak 'happiness'

Similar examples are

Ka tha a ba. 'I am tired.'

which appears in 8.1 above, and

A thin a hung ngaingai. 'He was very angry.'

which appears in translations (a) and (b). In the former, hyphens are used in a way not sanctioned by *Laica Trialning*.

Notes

¹I am grateful to Samuel Ngun Ling for teaching me what I know about Lai, and to Albert Ceuhlun and Kenneth Van Bik for assistance with the preliminary version of this paper.

²For some discussion of the development of Lai orthography, see Bawi Hu (to appear).

³The Lai retroflex stops are standardly written with a dot below the letter 't', as in the Indian tradition. This orthographic device is inconvenient for all printing processes I have access to, including the word processor used to prepare this manuscript. In all subsequent Lai cited herein these stops are given as 'tr' and 'thr', but it should be assumed that the quoted originals either make no distinction between retroflexes and dentals, or use the standard dotted 't'.

⁴Lai does not have a developed linguistic terminology, and the term *biafang* is used to correspond to the English 'word', 'morpheme' or 'syllable'.

⁵The 1978 Lai bible was printed under difficult circumstances, and proof-reading was not always very thorough. Such examples as *ahngalh* 'realized' were presumably intended by the translator not to be combined, while others like *um mi* 'were', to be combined. There are frequent inconsistencies in 'word combination'; even in this short passage, both *mifimhna* and *mifim hna* 'wise men' appear.

⁶Every syllable has been separated with the exception of the Biblical names 'Herod' and 'Bethlehem'.

⁷The examples given are traditional Lai songs, and contain obsolete words whose precise meaning is not clear to many people. I will not attempt to translate them.

⁸The original has *tha* here and constitutes an example of the orthographic issue raised in n.3. The example could also be interpreted as *an tha* 'their energy'. I choose to interpret *an* here as subject agreement rather than possessor agreement. See also the example below in 8.1, which is not ambiguous.

References

- Antony Ngun Uk. 1996. Laica (Written Lai). *Muko Magazine*. Yangon: Lai Baptist Church. Pp. 18-21.
- Bedell, George. 1995. Agreement in Lai. *Papers from the Fifth Annual Meeting of the Southeast Asian Linguistics Society*, ed. by S. Chelliah and W. de Reuse. Pp. 21-32.
- Chin Association for Christian Communication (CACC) Adult Literacy Project. 1998. *Upa Lai Carel Cawnnak* (Lessons in Reading and Writing Lai for Adults). Hakha.
- Chin Christian Literature Society (CCLS). no date. *Laica Trialning* (Lai Orthography). Hakha. (CCLS was in existence 1988-93).
- Lehman, F. K. (Chit Hlaing). 1995. On the Use of *dah* in Lai Chin Questions and the Operator Syntax of Functors', *Papers from the Fifth Annual Meeting of the Southeast Asian Linguistics Society*, edited by S. Chelliah and W. de Reuse. Pp. 211-31.
- Lehman, F. K. (Chit Hlaing) and Kenneth Van Bik. 1997. Notes on Lai Chin Personal Pronouns and Overt Case Marking. *Studies in the Linguistic Sciences* 27.2:81-86.
- Pu Bawi Hu. [to appear]. Literacy and language in Chin State, Myanmar. *Papers from the Sixth Annual Meeting of the Southeast Asian Linguistics Society*.
- Strait, Chester U. (tr.). 1950. *Lai Baibal Ca* (Lai New Testament). Rangoon: British and Foreign Bible Society.
- Van Bik, David (tr.). 1978. *Lai Baibal Thiang* (The Holy Bible in Lai). Calcutta: United Bible Societies.
- Van Bik, Kenneth. 1998. Lai Psycho-collocation. *Linguistics of the Tibeto-Burman Area* 21.1:201-33.

