'Word Combination' in Lai'

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:

\begin{array}{cccc}
\text{p} & \text{t} & \text{c} & \text{tl} \\
\text{ph} & \text{th} & \text{ch} & \text{thl} \\
\text{b} & \text{d} & \text{t} & \text{k} \\
\text{th} & \text{kh} \\
\end{array}

Syllable Initial Fricatives:

\begin{array}{ccc}
\text{f} & \text{s} & \text{h} \\
\text{v} & \text{z} \\
\end{array}

Syllable Initial Sonorants:

\begin{array}{cccc}
\text{m} & \text{n} & \text{l} & \text{r} \\
\text{hm} & \text{hn} & \text{hl} & \text{hr} \\
\text{ng} & \text{hng} \\
\end{array}

Vowels:

\begin{array}{cccccc}
\text{a} & \text{e} & \text{i} & \text{o} & \text{u} & \text{aw} \\
\end{array}

Diphthongs:

\begin{array}{cccc}
\text{ai} & \text{ao} & \text{au} & \text{awi} \\
\text{ei} & \text{eo} & \text{eu} \\
\text{ia} & \text{io} & \text{iu} \\
\text{oi} & \text{uo} \\
\text{ua} & \text{ui} & \text{uo} \\
\end{array}

Triphthongs:

\begin{array}{cccc}
\text{iai} & \text{ia} & \text{iau} \\
\text{iei} & \text{iao} & \text{ia} \\
\text{uai} & \text{ua} & \text{uau} \\
\text{uei} & \text{ue} \\
\end{array}

Syllable Final Stops:

\begin{array}{ccc}
\text{p} & \text{t} & \text{k} \\
\text{h} \\
\end{array}
Syllable Final Sonorants:

\[
\begin{array}{cccc}
m & n & l & r & ng \\
mh & nh & lh & rh & ngh \\
\end{array}
\]

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 syllable\(^4\). 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 Betlehem 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)

'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 kas 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 caa cu a tuak i Bethlehem