L2/26-037

Proposal to encode the Leke script into the Unicode Standard

Frank van de Kasteelen

January 18, 2026

1 Introduction

The Leke (*leke*) script, also known as the Chicken Scratch script (*laicangwee*) is used to write the Pwo Eastern Karen (*hploung*) language [kjp] spoken by just over a million people in Myanmar and some 50 000 people in Thailand. It was invented in the early 19th century and seems to have been influenced by the Ancient Mon and Myanmar writing systems.

This document is an expansion of the earlier proposal by Fickle and Hosken (2013) L2/13-116.

2 Usage

The Leke script is mainly used by members of the Ariya millenarian buddhist sect, but is also used by other Eastern Pwo peoples. There are currently three other competing writing systems in use for Pwo Eastern Karen: two based on the Myanmar writing system and one using the Thai script. The buddhist monastic Myanmar script is most commonly used for Eastern Pwo. The Christian missionary Myanmar script has some limited usage, but is mostly associated with Western Pwo and the Thai script is only used by refugees in Thailand.

The script has been in continuous use since it's invention and is currently being taught throughout the region.

3 Writing System

The Leke script is an abugida, like many other scripts in the region. It creates all the vowel sounds and tones necessary for the language by combining

different vowel signs, final consonant signs and tone marks to any given consonant. Leke also possesses four dedicated medial consonant signs for glides between the initial consonant and the vowel.

3.1 Consonants

Each consonant sound of the Eastern Pwo language is represented by a unique consonant sign. Some consonants can occur at the end of a syllable, in which case they lose their consonantal quality and are used as vowel signs. These will be treated below.

3.1.1 Initial Consonants

The following consonants are used for the onset:

Name	**Glyph**	**IPA**	**Name**	**Glyph**	**IPA**
LETTER KA	ɔ	k	LETTER YA	≴	j
LETTER HKA	/	kʰ	LETTER RA	ঌ	r
LETTER NGA	ꝛ	ŋ	LETTER LA	ɛ̃	l
LETTER CA	ㅅ	tɕ / s	LETTER WA	ꝙ	w
LETTER HCA	୨	tɕʰ / sʰ	LETTER THA	ꝏ	θ
LETTER NYA	ɕ	ɲ	LETTER GHA	ɞ	ɣ
LETTER TA	ɞɩ	t	LETTER HA	⊙	h
LETTER HTA	∩	tʰ	LETTER KHA	ɛ̃	x
LETTER NA	ɕ	n	LETTER HWA	⋈	ɰ
LETTER SA	ϱ	ɕ	LETTER A	×	ʔ
LETTER PA	ɜ	p	LETTER BA	⊙	ɓ
LETTER HPA	//	pʰ	LETTER DA	ɷ	ɗ
LETTER MA	≷	m			

Table 1: Leke consonants

One letter, U+11B95 LEKE LETTER HWA, has a variant with dots inside the character ⋈ vs. ⋈, but this is merely a stylistic choice and does not require seperate encoding.

3.1.2 Medial Consonants

Leke has a set of four medial consonants used for glides. These signs should be encoded anatomically as they have been in the Myanmar script and other similar writing systems. It would not make sense to add a virama-like character.

Glyph	Leke Name	Unicode Name	IPA
		PA YA	-j-
		PA LA	-l-
		PA WA	-w-
		RAI YA	-r-

Table 2: Leke medial consonants

A base consonant can, in some cases, carry up to two medial consonant signs. They are then placed next to each other, as in:

= U+11B80 + U+11BA9 + U+11BAB

= U+11B80 + U+11BAC + U+11BAB

3.2 Vowels

The Leke script uses a set of nine vowel signs and four tone marks, which are combined to create all the vowel and tone phonemes of the language. The vowel signs are:

Glyph	Leke Name	Unicode Name
		DWAI THA
		TANG DWAI
		TANG KANG LI
		SEUNG NU CUI
		THAEU WEE KLEU
		DWAI THAEU WEE
		CEE CANG
		DWAI THA HPA HTU
		TANG PI HPA HTU

Table 3: Leke vowel signs

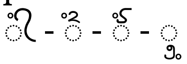
3.2.1 Vowels with Zero Initial

Leke does not have a set of independent vowels. Instead, initial vowels are written using × U+11B96 LEKE LETTER A with the required vowel and tone sign(s).

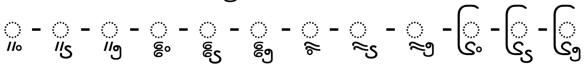
3.2.2 Vowel Placement

The Leke vowel system is very similar to the one for Myanmar, using a combination of vowel signs, final consonant signs and tone marks to create all the necessary vowel-tone combinations. Vowel signs can be placed before, above, after or below a consonant and tone marks can likewise be placed above, below and after a consonant.

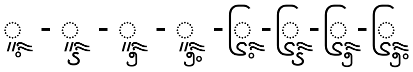
When a vowel and tone co-occur above a consonant, the tone is placed to the left of the vowel, but when they co-occur below a consonant, the tone is placed to the right:



In combination with medial consonants, vowels and tones usually come below to the right:



In case of double medial consonants, they are placed in the middle below:



Medial RAI YA can be adapted to accommodate for vowel signs above the consonant:
- - - - - , but - -

3.2.3 Finals

The Leke script has a set of two killers, that suppress the inherent vowel of the letter they attach to. These letters are then read as the coda of the syllable they follow. They are most readily compared to MYANMAR SIGN ASAT, which functions in the same way. They are always visibly displayed and do not form conjuncts.

Glyph	Leke Name	Unicode Name
		LANG HPU
		LANG DU

Table 4: Leke killers

The sign LANG HPU is used in combination with KA, HKA, NGA, CA, NYA, NA, PA, MA, YA, GHA, HA, KHA, HWA and A, while the sign LANG DU is used exclusively with TA and WA. The other consonants never occur in the syllable coda and therefore don't combine with either of these killers.

Since the distribution of the two killers is purely context dependent, they could be unified into a single killer with different appearance depending on the preceding consonant. Since the two signs are perceived as separate by the user community and since there is a history of them being separated in current ASCII and PUA based fonts, they have been proposed as two distinct signs.

LANG DU is a relatively new addition to the script. In older texts, it may be preferable to write LANG HPU under 11B86 and 11B90 as well, which also speaks for keeping them separate characters.

Phonetically, final stops and fricatives were once used to indicate a final glottal stop. These glottal stops are no longer pronounced in modern Pwo Eastern Karen, so often these final consonants now go completely unpronounced. In some cases they do change the vowel sound of their syllable. For example, KA + CA + LANG HPU is not pronounced /kac/, as might be expected, but rather /kɪ/.

The final nasals can all be used to indicate final /ɴ/ and depending on the letter used, can also modify the preceding vowel. For example KA + NA +

LANG HPU is pronounced /kaɴ/, but KA + NGA + LANG HPU is pronounced /kaiɴ/.

The final glides <y> and <w> are used to create front and back vowels, respectively. Examples can be foundd in table 7.

The tone diacritic THA LANG is placed below LANG HPU, but inside LANG DU: and .

3.2.4 Tones

The Easter Pwo language has four tones: high, mid, low and falling. These are marked with three different tone marks in Leke, the mid tone being left unmarked. The high tone can be indicated with two different signs, with one being used in combination with vowel signs above the consonant and the other on signs after the consonant. These are not seen as two variants of the same tone mark, however, and should be encoded separately.

Glyph	Leke Name	Unicode Name
		TANG PI
		THA LANG
		THA PI
		THA SOO

Table 5: Leke tones

3.2.5 List of Vowels and Tones

Below is a list of all the vowel combinations found in the Leke script. U+11B80 LEKE LETTER KA is used in place of the initial consonant. The first table gives only vowel signs, the second table gives vowel-final consonant combinations.

U+11BA6 DWAI THA HPA HTU and U+11BA7 TANG PI HPA HTU can optionally form a ligature , where TANG PI HPA HTU replaces the upper circle of DWAI THA HPA HTU.

Please be aware that the exact correspondences between sound and spelling in Leke are still being researched. The IPA values in the following tables are based on current understanding and may be subject to change.

transl.	IPA			◌ံ	◌့	◌း
A	a		ၥ			ၥး
A	a	◌ါ	ၥါ	ၥါံ	ၥါ့	ၥါး
II	i	◌ႋ	ၥႋ			
II	i	◌ီ	ၥီ	ၥီံ	ၥီ့	ၥီး
UI	ɨ	◌	ၥ			
EU	ɯ	◌ၳၳ	ၥၳၳ	ၥၳံ	ၥၳ့	ၥၳး
EE	e	ေ◌ၳ	ေၥၳ	ေၥၳံ	ေၥၳ့	ေၥၳး
E	ɛ	◌ꩻ	ၥꩻ		ၥꩻ့	
U	ʊ	ေ◌ါ	ေၥါ		ေၥါ့	
U	ʊ	ေ◌ၞ	ေၥၞ			
U	ʊ	ေ◌ၳၞ	ေၥၳၞ			
U	ʊ	ေ◌ၲၞ	ေၥၲၞ			
ANG	aɴ	◌ၳ	ၥၳ	ၥၳံ	ၥၳ့	ၥၳး

Table 6: Leke vowel combinations

transl.	IPA			[illegible]	[illegible]	[illegible]
AE	ə	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
AING	aiɴ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
I	ɪ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
ING	ɪɴ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
ANG	aɴ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
AEU	əʊ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
UNG	ʊɴ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
EING	eiɴ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
OUNG	oʊɴ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
OONG	oɴ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
E	ɛ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
O	ɔ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
AENG	əɴ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
AEU	əʊ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
EU	ɯ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
OO	o	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
UI	ɨ	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

transl.	IPA		
A	a	[illegible]	[illegible]
A	a	[illegible]	[illegible]
AI	ai	[illegible]	[illegible]
AU	aʊ	[illegible]	[illegible]
AU	aʊ	[illegible]	[illegible]
O	ɔ	[illegible]	[illegible]
AU	aʊ	[illegible]	[illegible]
AE	ə	[illegible]	[illegible]
AI	ai	[illegible]	[illegible]

Table 7: Leke vowel-final consonant combinations

3.2.6 Character sequencing

Syllables in the Leke script are to be entered in a logical order starting with the base consonant, then an optional medial consonant, optionally followed by the vowel signs going left, above, right, then below that consonant, or by the killer sign, and then optionally a tone mark. This does require some visual reordering of combining marks. The table below gives several examples in the prefered alphabetical order of vowel signs.

It is common practice in Leke teaching to give a syllable with all possible tonemarks to represent all different possibilities, so [illegible] stands for [illegible] or [illegible] or [illegible] or [illegible].

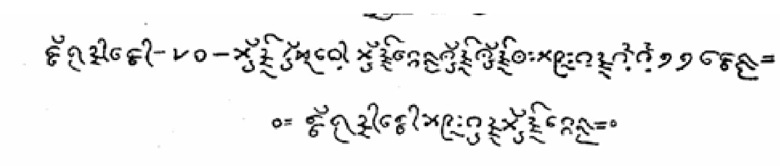
ၥ	ၥ	KA
[illegible]	ၥ + [illegible] (+ [illegible]/[illegible]/◌း)	KA + DWAI THA (+ TANG PI / THA LANG / THA PI)
[illegible]	ၥ + [illegible]	KA + TANG DWAI
[illegible]	ၥ + [illegible] (+ [illegible]/[illegible]/◌း)	KA + TANG KANG LI (+ TANG PI / THA LANG / THA PI)
[illegible]	ၥ + [illegible]	KA + SEUNG NU CUI
[illegible]	ၥ + [illegible] + [illegible] (+ [illegible]/[illegible]/◌း)	KA + TANG PI HPA HTU + THAEU WEE KLEU (+ TANG PI / THA LANG / THA PI)
[illegible]	ၥ + ေ◌ + [illegible] (+ [illegible]/[illegible]/◌း)	KA + DWAI THAEU WEE + TANG PI HPA HTU (+ TANG PI / THA LANG / THA PI)
[illegible]	ၥ + [illegible] (+ [illegible])	KA + CEE CANG (+ THA LANG)
[illegible]	ၥ + ေ◌ + [illegible] (+ [illegible])	KA + DWAI THAEU WEE + DWAI THA (+ THA LANG)
[illegible]	ၥ + ေ◌ + [illegible] (+ [illegible] / [illegible])	KA + DWAI THAEU WEE + DWAI THA HPA HTU (+ THA SOO / TANG PI HPA HTU)
[illegible]	ၥ + [illegible] (+ [illegible]/[illegible]/◌း)	KA + TANG PI HPA HTU (+ TANG PI / THA LANG / THA PI)
ၥး	ၥ + ◌း	KA + THA PI

Table 8: Visual reordering of characters in a syllable

3.3 Punctuation

The Leke script doesn't originally use punctuation. Spaces are used between phrases in a way similar to Latin commas and periods. A traditional 'end of section' mark [illegible] U+11BBA LEKE END OF SECTION is used. Examples can be found in figures 9 and 10 on pages 25 and 26.

In some educational material, a line, a double line and a small circle are used. For these U+002D HYPHEN-MINUS, U+2E30 RING DOT ∘ and U+2E40 DOUBLE HYPHEN ⹀ can be used. An annotation could be added to these characters. The figure below shows these characters in use.



Three consecutive rings ᪥᪥᪥ can be used used as a paragraph separator.

In modern texts, Latin punctuation marks, such as question marks, brackets, quotation marks and dashes are used. These will usually follow the Myanmar style with open circles.

3.4 Numerals

Leke has its own set of decimal digits: ᪀ ᪁ ᪂ ᪃ ᪄ ᪅ ᪆ ᪇ ᪈ ᪉.

4 Rendering

4.1 Stacking and Placement

In traditional, handwritten documents, the consonant letters NA and SA both extend below the base line. For SA this means that vowels and tones are placed a bit to the right, next to the descender. For NA this means that vowels and tones can often be placed inside the character. When the space inside the character is taken up by a medial consonant, the vowel and tone will be written underneath.

NYA is the only consonant letter that traditionally extends above x-height. In this case, vowels and tones are also often placed inside the character. In cases where this would create overlap, the top line of NYA is shortened to accommodate for the vowel.

Modern fonts have, due to technical constraints, simplified these cases by making SA, NA and NYA fit x-height. The table below highlights some of the differences.

The table below shows some of the differences between the traditional and modern font styles.

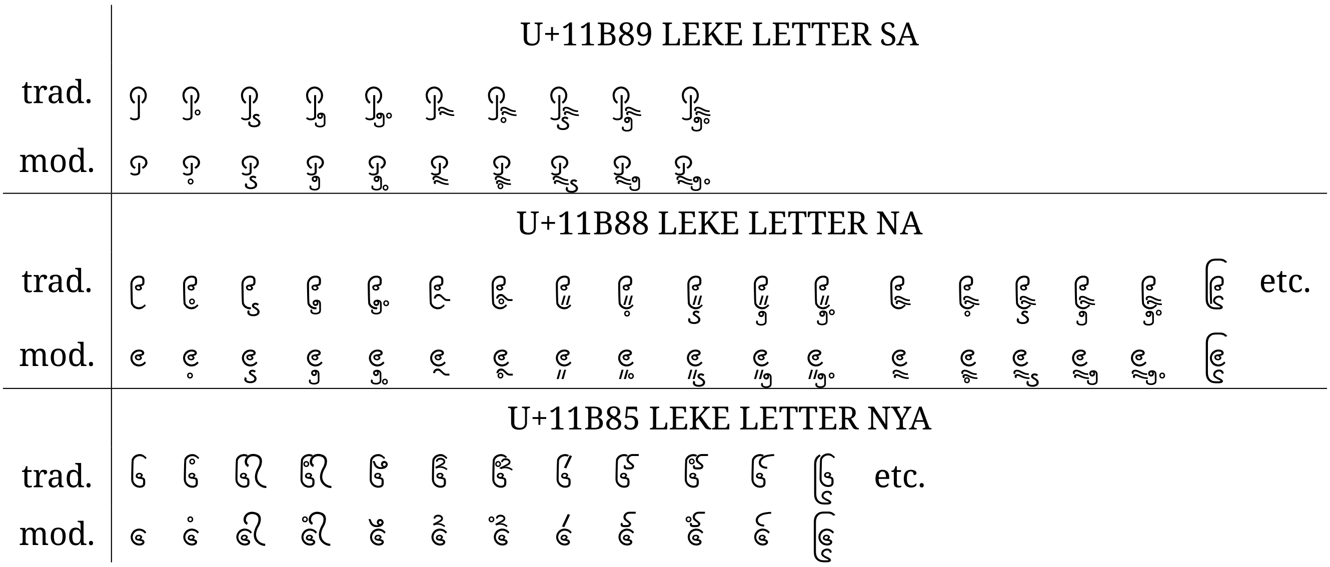


Table 9: Differences between traditional and modern font styles

It is important to note that the modern style is not the preferred style. It is used in modern documents due to the technical constraints of ASCII-based fonts, but the more traditional style is still used in handwriting.

4.2 Line Breaking and Space

Similar to Myanmar, spacing is not used between words, but rather between phrases. Line breaking can occur after a phonetic syllable, meaning that consonants appended with [illegible] LANG DU or [illegible] LANG HPU can never start a new line, e.g. [illegible] = [illegible] | [illegible] | or [illegible] = [illegible] | [illegible] | [illegible] |, where each | represents a line breaking opportunity.

Closing punctuation marks, such as],), ’, ” and [illegible] cannot start a new line.

4.3 Encoding model

Leke will be encoded as an indic script according to the Brahmi model. It does not contain any obligatory ligatures, but does have two pre-base dependent signs, as well as several combining marks that interact when combined. Especially in the traditional representation, several letters require some shaping behavior. The ccc values of all combining marks must be 0, since their placement depends on the preceding letter or sign.

5 Unicode Character Data

5.1 Collation

ↄ KA < ⵔ HKA < ⵒ NGA < ⵑ CA < ⵐ HCA < ⵏ NYA < ⵍ TA < ⵎ HTA < ⵋ NA < ⵖ SA < ⵗ PA < ⵘ HPA < ⵙ MA < ⵚ YA < ⵛ RA < ⵜ LA < ⵝ WA < ⵞ THA < ⵟ GHA < ⵠ HA < ⵡ KHA < ⵢ HWA < ⵣ A < ⵤ BA < ⵥ DA < ◌ⵯ SIGN DWAI THA < ◌ⵦ SIGN TANG PI < ◌ⵧ SIGN THA LANG < ◌⵨ SIGN THA PI < ◌⵩ SIGN TANG DWAI < ◌⵪ SIGN TANG KAN̊G LI < ◌ⵓ SIGN SEUNG NU CUI < ◌ⵖ SIGN THAEU WEE KLEU < ⵏ◌ SIGN DWAI THAEU WEE < ◌ⵕ SIGN CEE CANG < ◌ⵘ SIGN DWAI THA HPA HTU < ◌ⵙ SIGN TANG PI HPA HTU < ◌ⵚ LEKE SIGN THA SOO < ◌ⵛ SIGN PA YA < ◌ⵜ SIGN PA LA < ◌ⵝ SIGN PA WA < ◌ⵞ SIGN RAI YA < ◌ⵟ SIGN LANG HPU < ◌ⵠ SIGN LANG DU

5.2 Character properties

```
11B80;LEKE LETTER KA;Lo;0;L;;;;;N;;;;
11B81;LEKE LETTER HKA;Lo;0;L;;;;;N;;;;
11B82;LEKE LETTER NGA;Lo;0;L;;;;;N;;;;
11B83;LEKE LETTER CA;Lo;0;L;;;;;N;;;;
11B84;LEKE LETTER HCA;Lo;0;L;;;;;N;;;;
11B85;LEKE LETTER NYA;Lo;0;L;;;;;N;;;;
11B86;LEKE LETTER TA;Lo;0;L;;;;;N;;;;
11B87;LEKE LETTER HTA;Lo;0;L;;;;;N;;;;
11B88;LEKE LETTER NA;Lo;0;L;;;;;N;;;;
11B89;LEKE LETTER SA;Lo;0;L;;;;;N;;;;
11B8A;LEKE LETTER PA;Lo;0;L;;;;;N;;;;
11B8B;LEKE LETTER HPA;Lo;0;L;;;;;N;;;;
11B8C;LEKE LETTER MA;Lo;0;L;;;;;N;;;;
11B8D;LEKE LETTER YA;Lo;0;L;;;;;N;;;;
11B8E;LEKE LETTER RA;Lo;0;L;;;;;N;;;;
11B8F;LEKE LETTER LA;Lo;0;L;;;;;N;;;;
11B90;LEKE LETTER WA;Lo;0;L;;;;;N;;;;
11B91;LEKE LETTER THA;Lo;0;L;;;;;N;;;;
11B92;LEKE LETTER GHA;Lo;0;L;;;;;N;;;;
11B93;LEKE LETTER HA;Lo;0;L;;;;;N;;;;
11B94;LEKE LETTER KHA;Lo;0;L;;;;;N;;;;
11B95;LEKE LETTER HWA;Lo;0;L;;;;;N;;;;
11B96;LEKE LETTER A;Lo;0;L;;;;;N;;;;
11B97;LEKE LETTER BA;Lo;0;L;;;;;N;;;;
11B98;LEKE LETTER DA;Lo;0;L;;;;;N;;;;
11B9C;LEKE SIGN DWAI THA;Mc;0;L;;;;;N;;;;
11B9D;LEKE SIGN TANG PI;Mn;0;NSM;;;;;N;;;;
```

```
11B9E;LEKE SIGN THA LANG;Mn;0;NSM;;;;;N;;;;
11B9F;LEKE SIGN THA PI;Mc;0;L;;;;;N;;;;
11BA0;LEKE SIGN TANG DWAI;Mn;0;NSM;;;;;N;;;;
11BA1;LEKE SIGN TANG KANG LI;Mn;0;NSM;;;;;N;;;;
11BA2;LEKE SIGN SEUNG NU CUI;Mn;0;NSM;;;;;N;;;;
11BA3;LEKE SIGN THAEU WEE KLEU;Mn;0;NSM;;;;;N;;;;
11BA4;LEKE SIGN DWAI THAEU WEE;Mc;0;L;;;;;N;;;;
11BA5;LEKE SIGN CEE CANG;Mn;0;NSM;;;;;N;;;;
11BA6;LEKE SIGN DWAI THA HPA HTU;Mc;0;L;;;;;N;;;;
11BA7;LEKE SIGN TANG PI HPA HTU;Mn;0;NSM;;;;;N;;;;
11BA8;LEKE SIGN THA SOO;Mn;0;NSM;;;;;N;;;;
11BA9;LEKE SIGN PA YA;Mn;0;NSM;;;;;N;;;;
11BAA;LEKE SIGN PA LA;Mn;0;NSM;;;;;N;;;;
11BAB;LEKE SIGN PA WA;Mn;0;NSM;;;;;N;;;;
11BAC;LEKE SIGN RAI YA;Mc;0;L;;;;;N;;;;
11BAD;LEKE SIGN LANG HPU;Mn;0;NSM;;;;;N;;;;
11BAE;LEKE SIGN LANG DU;Mn;0;NSM;;;;;N;;;;
11BB0;LEKE DIGIT ZERO;Nd;0;L;;0;0;0;N;;;;;
11BB1;LEKE DIGIT ONE;Nd;0;L;;1;1;1;N;;;;;
11BB2;LEKE DIGIT TWO;Nd;0;L;;2;2;2;N;;;;;
11BB3;LEKE DIGIT THREE;Nd;0;L;;3;3;3;N;;;;;
11BB4;LEKE DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;;
11BB5;LEKE DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;;
11BB6;LEKE DIGIT SIX;Nd;0;L;;6;6;6;N;;;;;
11BB7;LEKE DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;;
11BB8;LEKE DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;;
11BB9;LEKE DIGIT NINE;Nd;0;L;;9;9;9;N;;;;;
11BBA;LEKE END OF SECTION;Po;0;L;;;;;N;;;;;
```

5.3 Indic categories

```
 # Indic_Syllabic_Category=Vowel_Dependent
11B9C         ;  Vowel_Dependent   # Mc        LEKE SIGN DWAI THA
11BA0..11BA3  ;  Vowel_Dependent   # Mn  [4]   LEKE SIGN TONG DWAI
                                               ..LEKE SIGN THAEU
                                               WEE KLEU
11BA4         ;  Vowel_Dependent   # Mc        LEKE SIGN DWAI
                                               THAEU WEE
11BA5         ;  Vowel_Dependent   # Mn        LEKE SIGN CEE CANG
11BA6         ;  Vowel_Dependent   # Mc        LEKE SIGN DWAI THA
                                               HPA HTU
11BA7         ;  Vowel_Dependent   # Mn        LEKE SIGN TONG PI
                                               HPA HTU
```

```
 # Indic_Syllabic_Category=Consonant
11B80..11B98  ;  Consonant          # Lo  [25]  LEKE LETTER KA..
                                               LEKE LETTER DA

 # Indic_Syllabic_Category=Consonant_Medial
11BA9..11BAB  ;  Consonant_Medial   # Mn  [3]   LEKE SIGN PA YA..
                                               LEKE SIGN PA WA
11BAC         ;  Consonant_Medial   # Mc        LEKE SIGN RAI YA

 # Indic_Syllabic_Category=Tone_Mark
11B9D..11B9E  ;  Tone_Mark          # Mn  [2]   LEKE SIGN TANG PI..
                                               LEKE SIGN THA LANG
11B9F         ;  Tone_Mark          # Mc        LEKE SIGN THA PI
11BA8         ;  Tone_Mark          # Mn        LEKE SIGN THA SOO

 # Indic_Syllabic_Category=Pure_Killer
11BAD..11BAE  ;  Pure_Killer        # Mn  [2]   LEKE SIGN LANG HPU
                                               ..LEKE SIGN LANG DU

 # Indic_Syllabic_Category=Number
11BB0..11BB9  ;  Number             # Nd  [10]  LEKE DIGIT ZERO..
                                               LEKE DIGIT NINE
```

5.4 Positional categories

Notes:

The following character is assigned the positional category Top, but may have different positions in some cases:
* U+11B9D LEKE SIGN TANG PI has contextually variable placement in Leke.

The following characters are all assigned the positional category Bottom, but may have different positions in some cases:
* U+11B9E LEKE SIGN THA LANG, 11BA2 LEKE SIGN SEUNG NU CUI and 11BA3 LEKE SIGN THAEU WEE KLEU have contextually variable placement in Leke.

```
 # Indic_Positional_Category=Top
11B9D         ;  Top     # Mn       LEKE SIGN TANG PI
11BA0..11BA1  ;  Top     # Mn  [2]  LEKE SIGN TANG DWAI..LEKE SIGN
                                    TANG KANG LI
11BA5         ;  Top     # Mn       LEKE SIGN CEE CANG
11BA7..11BA8  ;  Top     # Mn  [2]  LEKE LEKE SIGN TANG PI HPA
                                    HTU..LEKE SIGN THA SOO
```

```
 # Indic_Positional_Category=Bottom
11B9E          ;  Bottom  # Mn       LEKE SIGN THA LANG
11BA2..11BA3   ;  Bottom  # Mn  [2]  LEKE SIGN SEUNG NU CUI..LEKE SIGN
                                     THAEU WEE KLEU
11BA9..11BAB   ;  Bottom  # Mn  [3]  LEKE SIGN PA YA..LEKE SIGN PA WA
11BAD..11BAE   ;  Bottom  # Mn  [2]  LEKE SIGN LANG HPU..LEKE SIGN LANG
                                     DU

 # Indic_Positional_Category=Right
11B9C          ;  Right   # Mc       LEKE SIGN DWAI THA
11B9F          ;  Right   # Mc       LEKE SIGN THA PI
11BA6          ;  Right   # Mc       LEKE SIGN DWAI THA HPA HTU

 # Indic_Positional_Category=Left
11BA4          ;  Left    # Mc       LEKE SIGN DWAI THAEU WEE
11BAC          ;  Left    # Mc       LEKE SIGN RAI YA
```

5.5 Linebreaking properties

```
11B80..11B98   ;  AL  # Lo  [25]  LEKE LETTER KA..LEKE LETTER DA
11B9C          ;  CM  # Mc        LEKE SIGN DWAI THA
11B9D..11B9E   ;  CM  # Mn  [2]   LEKE SIGN TANG PI..LEKE SIGN THA
                                  LANG
11B9F          ;  CM  # Mc        LEKE SIGN THA PI
11BA0..11BA3   ;  CM  # Mn  [4]   LEKE SIGN TANG DWAI..SIGN THAEU
                                  WEE KLEU
11BA4          ;  CM  # Mc        LEKE SIGN DWAI THAEU WEE
11BA5          ;  CM  # Mn        LEKE SIGN CEE CANG
11BA6          ;  CM  # Mc        LEKE SIGN DWAI THA HPA HTU
11BA7..11BAB   ;  CM  # Mn  [5]   LEKE SIGN TANG PI HPA HTU..LEKE
                                  SIGN PA WA
11BAC          ;  CM  # Mc        LEKE SIGN RAI YA
11BAD..11BAE   ;  SA  # Mn  [2]   LEKE SIGN LONG HPU..LEKE SIGN LONG
                                  DU
11BB0..11BB9   ;  NU  # Nd  [10]  LEKE DIGIT ZERO..LEKE DIGIT NINE
11BBA          ;  BA  # Po        LEKE END OF SECTION
```

5.6 Character table

	11B8	11B9	11BA	11BB
0	11B80	11B90	11BA0	11BB0
1	11B81	11B91	11BA1	11BB1
2	11B82	11B92	11BA2	11BB2
3	11B83	11B93	11BA3	11BB3
4	11B84	11B94	11BA4	11BB4
5	11B85	11B95	11BA5	11BB5
6	11B86	11B96	11BA6	11BB6
7	11B87	11B97	11BA7	11BB7
8	11B88	11B98	11BA8	11BB8
9	11B89		11BA9	11BB9
A	11B8A		11BAA	11BBA
B	11B8B		11BAB	
C	11B8C	11B9C	11BAC	
D	11B8D	11B9D	11BAD	
E	11B8E	11B9E	11BAE	
F	11B8F	11B9F		

5.7 Character list

Letters

11B80		LEKE LETTER KA
11B81		LEKE LETTER HKA
11B82		LEKE LETTER NGA
11B83		LEKE LETTER CA
11B84		LEKE LETTER HCA
11B85		LEKE LETTER NYA
11B86		LEKE LETTER TA
11B87		LEKE LETTER HTA
11B88		LEKE LETTER NA
11B89		LEKE LETTER SA
11B8A		LEKE LETTER PA
11B8B		LEKE LETTER HPA
11B8C		LEKE LETTER MA
11B8D		LEKE LETTER YA
11B8E		LEKE LETTER RA
11B8F		LEKE LETTER LA
11B90		LEKE LETTER WA
11B91		LEKE LETTER THA
11B92		LEKE LETTER GHA
11B93		LEKE LETTER HA
11B94		LEKE LETTER KHA
11B95		LEKE LETTER HWA
11B96		LEKE LETTER QA
11B97		LEKE LETTER BA
11B98		LEKE LETTER DA
11B99		<reserved>
11B9A		<reserved>
11B9B		<reserved>

Vowels and Tones

11B9C		LEKE SIGN DWAI THA
11B9D		LEKE SIGN TANG PI
11B9E		LEKE SIGN THA LANG
11B9F		LEKE SIGN THA PI
11BA0		LEKE SIGN TANG DWAI
11BA1		LEKE SIGN TANG KANG LI
11BA2		LEKE SIGN SEUNG NU CUI
11BA3		LEKE SIGN THAEU WEE KLEU
11BA4		LEKE SIGN DWAI THAEU WEE
11BA5		LEKE SIGN CEE CANG
11BA6		LEKE SIGN DWAI THA HPA HTU
11BA7		LEKE SIGN TANG PI HPA HTU
11BA8		LEKE SIGN THA SOO
11BA9		LEKE SIGN PA YA
11BAA		LEKE SIGN PA LA
11BAB		LEKE SIGN PA WA
11BAC		LEKE SIGN RAI YA
11BAD		LEKE SIGN LONG HPU
11BAE		LEKE SIGN LONG DU
11BAF		<reserved>

Digits

11BB0		LEKE DIGIT ZERO
11BB1		LEKE DIGIT ONE
11BB2		LEKE DIGIT TWO
11BB3		LEKE DIGIT THREE
11BB4		LEKE DIGIT FOUR
11BB5		LEKE DIGIT FIVE
11BB6		LEKE DIGIT SIX
11BB7		LEKE DIGIT SEVEN
11BB8		LEKE DIGIT EIGHT
11BB9		LEKE DIGIT NINE

Punctuation

11BBA		LEKE END OF SECTION
11BBB		<reserved>
11BBC		<reserved>
11BBD		<reserved>
11BBE		<reserved>
11BBF		<reserved>

6 Bibliography

References

Fickle, Erich and Martin Hosken (2013). *Revised Proposal to add the Leke Script in the SMP of the UCS*.

Kato, Atsuhiko (2009). "A basic vocabulary of Htoklibang Pwo Karen with Hpa-an, Kyonbyaw, and Proto-Pwo Karen forms". In: *Asian and African Languages and Linguistics* 4, pp. 169–218.

— (2019). "The middle marker in Pwo Karen". In: *Reports of the Keio Institute of Cultural and Linguistic Studies* 50, pp. 21–62.

— (2020). "A note on the quantifier float in Pwo Karen". In: *Reports of the Keio Institute of Cultural and Linguistic Studies* 51, pp. 173–188.

— (2021). "Pwo Karen writing systems". In: *Reports of the Keio Institute of Cultural and Linguistic Studies* 52, pp. 23–55.

— (2022). "Eastern Pwo Karen verb particles indicating 'up' and 'down'". In: *Journal of Research Institute* 63, pp. 143–176.

Kelly, Piers (Oct. 2018). "The art of not being legible: Invented writing systems as technologies of resistance in mainland Southeast Asia". In: *Terrain* 70.

Stern, Theodore (1968). "Three Pwo Karen Scripts: A study of alphabet formation". In: *Anthropological Linguistics* 10.1.

7 Acknowledgements

I owe many thanks to Sher Nay Moo (Paung Mla Mla) for sharing his insights into the script of his community and for closely cooperating with me on this project, to Atsuhiko Kato for sharing his private resources regarding the script with me, to Martin Hosken, who allowed me to take over his project and provided feedback and to Lorna Evans for her insights in the proposal process.

8 Materials

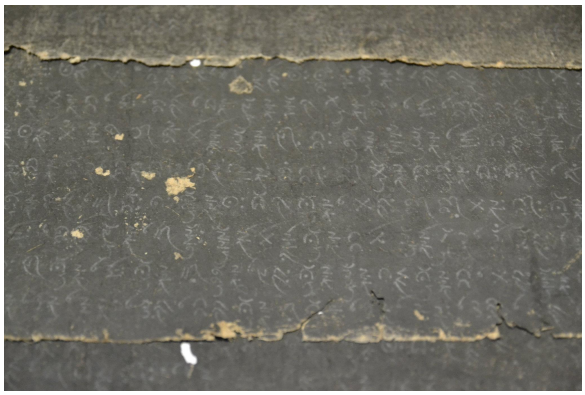


Figure 1: Old Leke black bark paper text. One of the earliest examples of Leke writing.

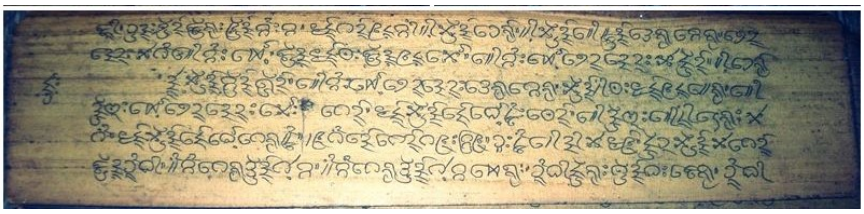


Figure 2: Excerpt from an old Leke leaf paper book.

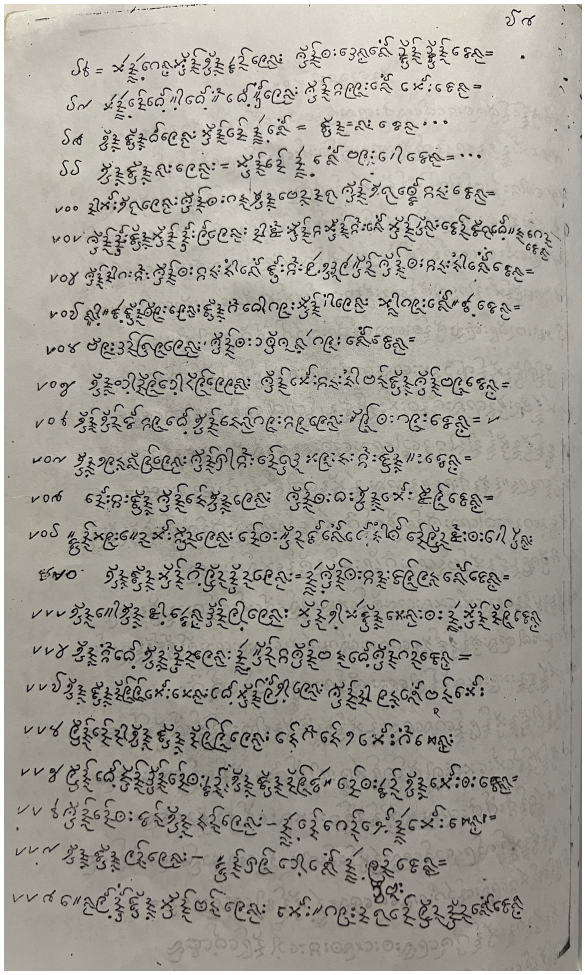


Figure 3: Handwritten Leke document using ◦◦◦ and ▫. *Eastern Pwo Karen Literature and Culture group*

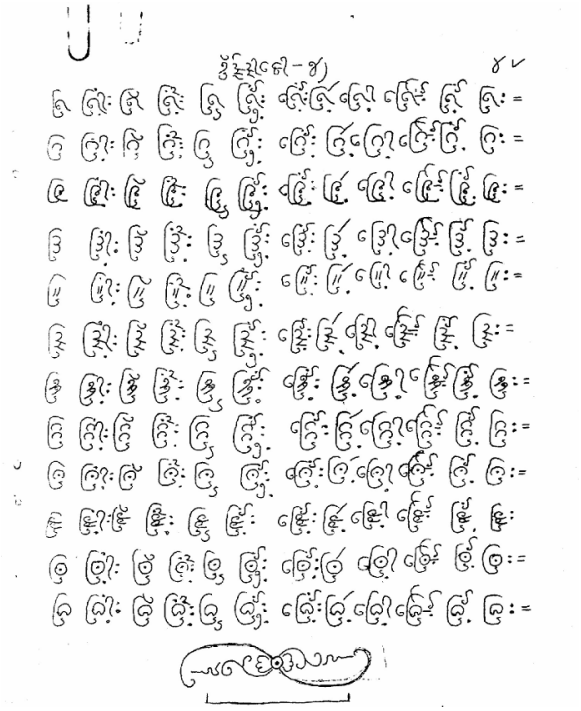


Figure 4: Page from a Leke copybook showing several consonant-vowel combinations with medial RAI YA. *Eastern Pwo Karen Literature and Culture group*

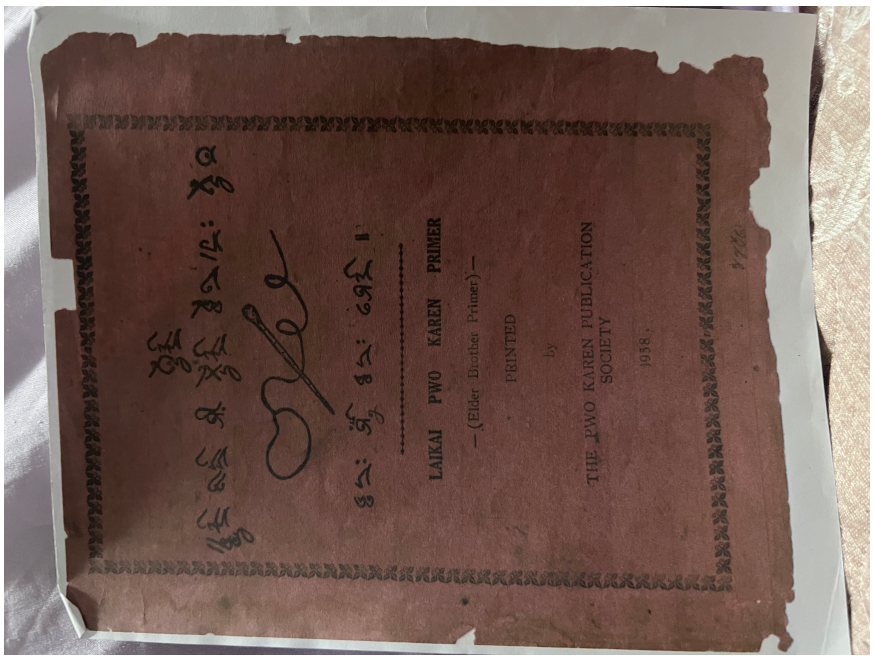


Figure 5: Cover of a Leke primer from 1958.

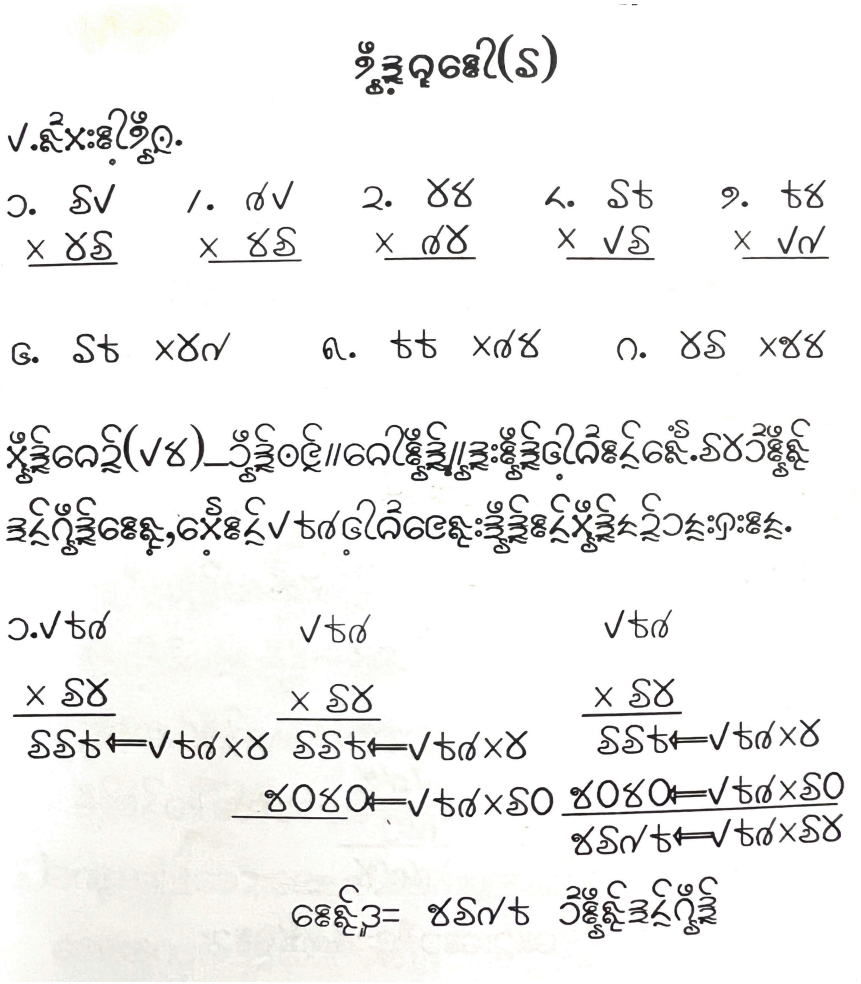


Figure 6: Page from a Leke maths textbook. *Done Taw Nee Education Department (2001)*

Figure 7: Page from a Leke book of poems. *Bangkok Karen Lietchawwait Association (2019)*

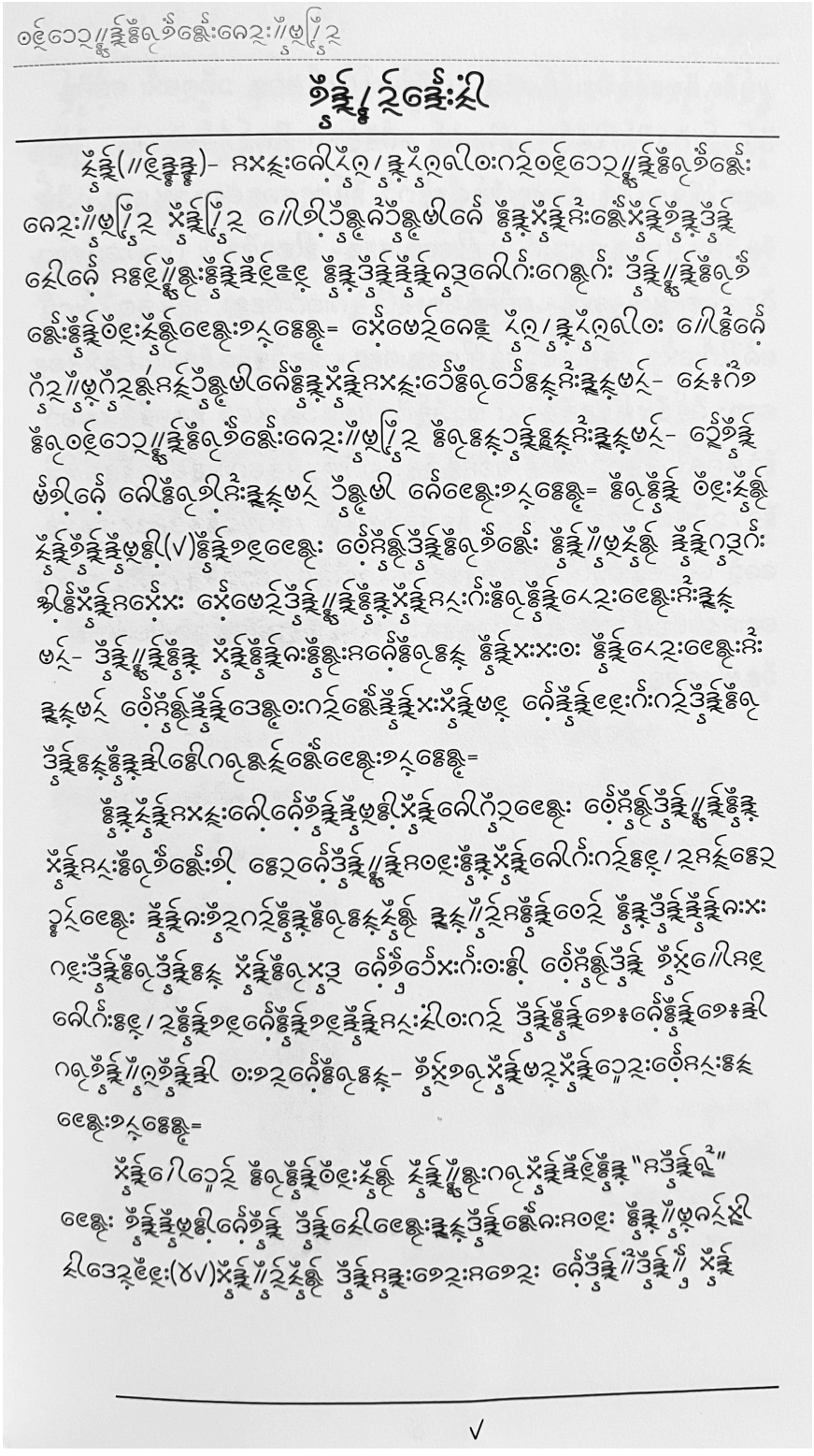


Figure 8: Page from a Leke book introduction. *Bangkok Karen Lietchawwait Association (2019)*

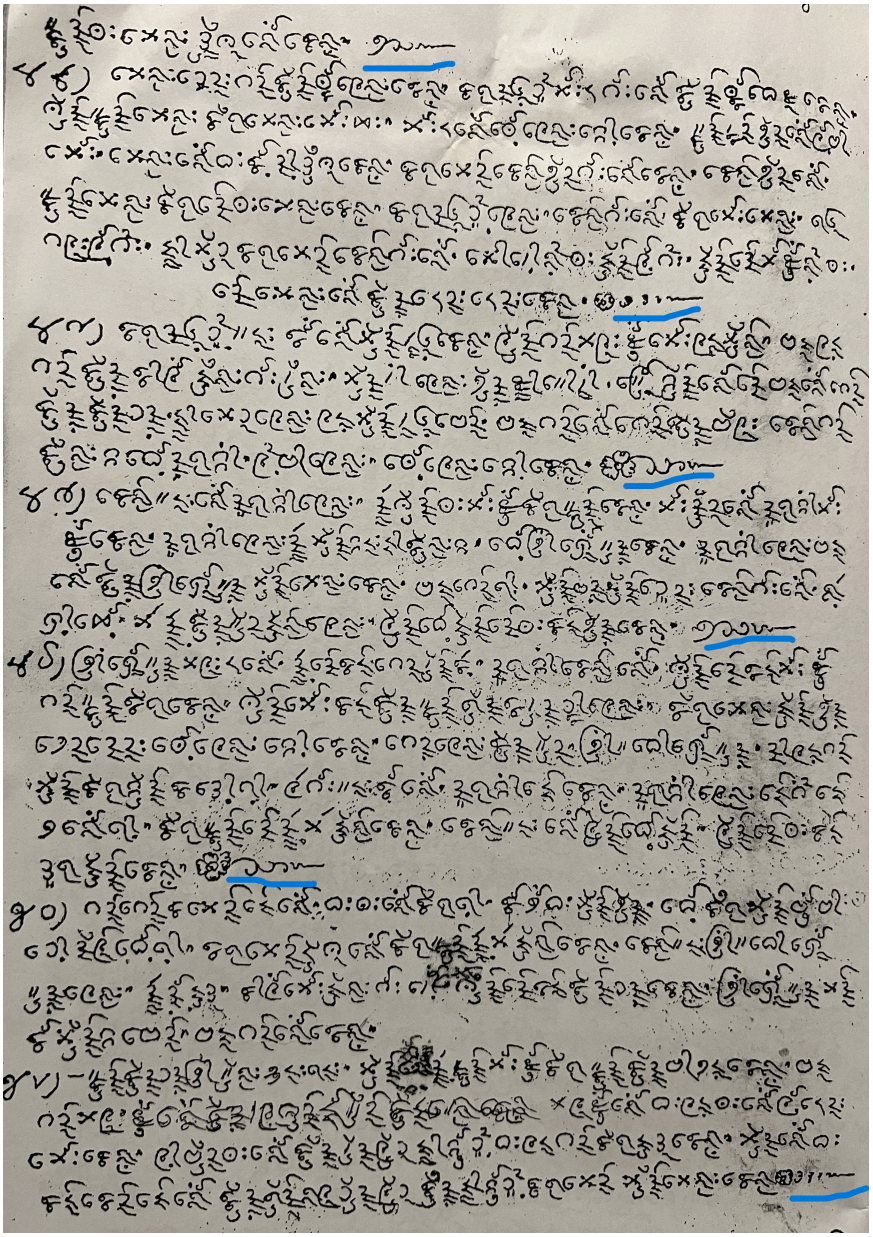


Figure 9: Example of the section mark in an ornate style.

Figure 10: Example of the section mark in a simple style

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]
Please fill all the sections A, B and C below.
Please read Principles and Procedures Document (P & P) from http://std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html **for guidelines and details before filling this form.**
Please ensure you are using the latest Form from http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html**.**
See also http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

A. Administrative

1. **Title:** *Proposal to encode Leke*
2. Requester's name: *F. van de Kasteelen, Sher Nay Moo*
3. Requester type (Member body/Liaison/Individual contribution): *Individual contribution*
4. Submission date: *18-12-2024*
5. Requester's reference (if applicable):
6. Choose one of the following:
   This is a complete proposal: *X*
   (or) More information will be provided later:

B. Technical – General

1. Choose one of the following:
   a. This proposal is for a new script (set of characters): *Yes*
      Proposed name of script: *Leke*
   b. The proposal is for addition of character(s) to an existing block:
      Name of the existing block:
2. Number of characters in proposal: *55*
3. Proposed category (select one from below - see section 2.2 of P&P document):
   A-Contemporary A | B.1-Specialized (small collection) | B.2-Specialized (large collection)
   C-Major extinct | D-Attested extinct | E-Minor extinct
   F-Archaic Hieroglyphic or Ideographic | G-Obscure or questionable usage symbols
4. Is a repertoire including character names provided? *Yes*
   a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document? *Yes*
   b. Are the character shapes attached in a legible form suitable for review? *Yes*
5. Fonts related:
   a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?
   t.b.d.
   b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
6. References:
   a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? *Yes*
   b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? *Yes*
7. Special encoding issues:
   Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)? *Yes*
8. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see Unicode Character Database (http://www.unicode.org/reports/tr44/) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

[1] Form number: N4502-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? *No*
   If YES explain
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)? *Yes*
   If YES, with whom? *Sher Nay Moo*
   If YES, available relevant documents:
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? *Yes*
   Reference: *1. / 2.*
4. The context of use for the proposed characters (type of use; common or rare) *Common*
   Reference: *2.*
5. Are the proposed characters in current use by the user community? *Yes*
   If YES, where? Reference: *Myanmar, 1. / 2.*
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? *No*
   If YES, is a rationale provided?
   If YES, reference:
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? *Yes*
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? *No*
   If YES, is a rationale for its inclusion provided?
   If YES, reference:
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? *No*
   If YES, is a rationale for its inclusion provided?
   If YES, reference:
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character? *No*
    If YES, is a rationale for its inclusion provided?
    If YES, reference:
11. Does the proposal include use of combining characters and/or use of composite sequences? *Yes*
    If YES, is a rationale for such use provided? *Yes*
    If YES, reference: *3.3*
    Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
    If YES, reference:
12. Does the proposal contain characters with any special properties such as control function or similar semantics? *No*
    If YES, describe in detail (include attachment if necessary)
13. Does the proposal contain any Ideographic compatibility characters? *No*
    If YES, are the equivalent corresponding unified ideographic characters identified?
    If YES, reference: