

Title:	Comments on N2634, Proposal to encode the Tai Lue script in the BMP of the UCS
Doc. Type:	Expert contribution
Source:	Peter Constable, Microsoft
Date:	October 16, 2003
Action:	For consideration by JTC1/SC2/WG2, UTC
References:	WG2/N2634 (=L2/03-321), N2371 (=L2/01-348), N2242, L2/99-243
Distribution:	WG2 members, UTC members

Document N2634 provides a complete proposal for encoding the Tai Lue (Xishuangbanna Dai) script. This document provides comments on that proposal.

Character inventory and encoding model

The character repertoire and encoding model proposed in N2634 is appropriate and complete for this script.

Ordering of characters

The proposal does not discuss collation issues explicitly. While the encoding order of characters should not be strictly tied to collation order (due to collation, in general, involving greater complexity than can be reflected in code point order of characters alone, as well as variation between cultures), the code point order will influence implementers, and so is an important consideration.

There is some variation in Tai Lue documents and descriptions of Tai Lue with respect to ordering, though one ordering of consonant and vowels seems to prevail. The ordering of consonants in this proposal is corroborated in various sources.

The ordering of vowel symbols can be seen in the following sample:¹

¹ Note that the columns in this chart come in sets of three: the left column has a phonetic transcription, the middle column has the Tai Lue vowel sign, and the right-hand column contains the older writing convention for the Tai Lue language, using a variation of the Lanna script.

国际音标	新	老	国际音标	新	老
a	-: 低组字母	-: 低组字母用	ai	ꠘ-	ꠘ-
			a:i	-ꠘ	-ꠘ
	单用高组字母	单用高组字母	ui	-ꠘ	ꠘ、ꠘ
			oi	-ꠘ	ꠘ
a:	-ꠘ	-ꠘ、-ꠘ	ꠘi	-ꠘ	ꠘ
i	-ꠘ	ꠘ	ui	-ꠘ	ꠘꠘ
i:	-ꠘ	ꠘ	oi	ꠘ-ꠘ	ꠘꠘꠘ
u	-ꠘ	ꠘ	au	-ꠘ	ꠘꠘ
u:	-ꠘ	ꠘ	a:u	-ꠘꠘ	-ꠘ
e	ꠘ-:	ꠘ-:	iu	-ꠘꠘ	ꠘ、ꠘ
e:	ꠘ-	ꠘ-	eu	ꠘ-ꠘ	-ꠘ、ꠘ
ɛ	ꠘꠘ-:	ꠘꠘ-:	ɛu	ꠘꠘ-ꠘ	ꠘꠘ
ɛ:	ꠘꠘ-	ꠘꠘ-	-m	-ꠘ	-ꠘ[am]、ꠘ、-ꠘ[am]
o	ɛ-:	ɛ-:	-n	-ꠘ	ꠘ[an]、ꠘ、-ꠘ[an]
o:	ɛ-	ɛ-、ꠘ、ꠘ-ꠘ	-ŋ	-ꠘ	ꠘ[aŋ]、ꠘ、-ꠘ[aŋ]
ɔ	-ꠘ:	ꠘꠘ、ꠘ	-p	-ꠘ	ꠘ[ap]、ꠘ、-ꠘ[ap]
ɔ:	-ꠘ	ꠘ、ꠘ	-t	-ꠘ	ꠘ[at]、ꠘ、-ꠘ[at]
ɯ	-ꠘ:	ꠘ:	-k	-ꠘ	ꠘ[ak]、ꠘ、-ꠘ[ak]
ɯ:	-ꠘ	ꠘ			
ə	ꠘ-ꠘ:	ꠘꠘ:			
ə:	ꠘ-ꠘ	ꠘꠘ、ꠘꠘꠘ			

Figure 1. Tai Lue vowel chart and comparison with Lanna script (SCEM, p. 73)²

The following sample shows a vowel chart that is organized differently. The bottom row contains the diphthongs, however, and if these are placed in order following all the basic vowel symbols, then going through the remainder of the chart by columns from left to right yields exactly the same order as seen in Figure 1.

-:	a	-ꠘ:	-ꠘ	ꠘ-:	ꠘꠘ-:	ɛ-:	-ꠘ:	-ꠘ:	ꠘ-ꠘ:
	a	i	u	e	ɛ	o	ɔ	ɯ	ə
ꠘ-	-ꠘ	-ꠘ	-ꠘ	ꠘ-	ꠘ-	ꠘ-	-ꠘ	-ꠘ	ꠘ-ꠘ
ai	ai	ui				oi	oi	ui	yi

Figure 2. Tai Lue vowel chart (PSCC, p. 165)³

² The chart in Figure 1 gives a different order for final consonants than is indicated in the proposal. This source is an exception in this regard, however; the ordering of final consonants in the proposal is corroborated in PSCC and LA1.

The ordering of basic vowel signs in N2634 follows the order shown in Figure 1 and Figure 2. The ordering of diphthong vowel signs differs, however: in N2634, TAI LUE VOWEL SIGN IY 𑜉𑜂𑜆 is placed third among the diphthongs, whereas it should occur last. The position of this character after all other vowels is seen in Figure 1 and Figure 2 above, as well as in the following sample:

	-ꨀ	-ꨁ			ꨀ-ꨀ	ꨀꨀ-ꨀ	ꨀ-ꨀ	-ꨀꨀ	-ꨁ	ꨀ-ꨁ	ꨀ-ꨁ
	-ꨂ	-ꨃ	-ꨄ	-ꨅ	ꨀ-	ꨀꨀ-	ꨀ-	-ꨂ	-ꨃ	ꨀ-ꨂ	ꨀ-ꨃ
	ꨂ-	-ꨃ		-ꨄ			-ꨅ	-ꨆ	-ꨇ	ꨀ-ꨇ	ꨀ-ꨇ
-ꨈ	ꨈ	ꨉ			ꨀ-ꨈ	ꨀꨀ-ꨈ				ꨀ-ꨉ	ꨀ-ꨉ
-ꨊ	ꨊ	ꨋ	ꨌ		ꨀ-ꨊ	ꨀꨀ-ꨊ	ꨀ-ꨊ	ꨀ	ꨋ	ꨀ-ꨌ	ꨀ-ꨌ
-ꨍ	ꨍ	ꨎ	ꨏ		ꨀ-ꨍ	ꨀꨀ-ꨍ	ꨀ-ꨍ	ꨀ	ꨎ	ꨀ-ꨏ	ꨀ-ꨏ
-ꨐ	ꨐ	ꨑ	ꨒ		ꨀ-ꨐ	ꨀꨀ-ꨐ	ꨀ-ꨐ	ꨀ	ꨑ	ꨀ-ꨒ	ꨀ-ꨒ
-ꨓ	ꨓ	ꨔ	ꨕ	ꨖ	ꨀ-ꨓ	ꨀꨀ-ꨓ	ꨀ-ꨓ	ꨀ	ꨔ	ꨀ-ꨕ	ꨀ-ꨕ
-ꨗ	ꨗ	ꨘ	ꨙ		ꨀ-ꨗ	ꨀꨀ-ꨗ	ꨀ-ꨗ	ꨀ	ꨘ	ꨀ-ꨙ	ꨀ-ꨙ
-ꨛ	ꨛ	ꨜ	ꨝ		ꨀ-ꨛ	ꨀꨀ-ꨛ	ꨀ-ꨛ	ꨀ	ꨜ	ꨀ-ꨝ	ꨀ-ꨝ
-ꨟ	ꨟ	ꨠ	ꨡ		ꨀ-ꨟ	ꨀꨀ-ꨟ	ꨀ-ꨟ	ꨀ	ꨠ	ꨀ-ꨡ	ꨀ-ꨡ

Figure 3. Tai Lue vowel chart: TAI LUE VOWEL SIGN IY as last diphthong (LA1, p. 36)

Thus, I recommend re-assignment of diphthong vowel characters as follows:

Proposed character	Revised code point
19C8 TAI LUE VOWEL SIGN IY 𑜉𑜂𑜆	19CC
19C9 TAI LUE VOWEL SIGN UY 𑜉𑜃𑜆	19C8
19CA TAI LUE VOWEL SIGN OY 𑜉𑜄𑜆	19C9
19CB TAI LUE VOWEL SIGN OY 𑜉𑜄𑜆	19CA
19CC TAI LUE VOWEL SIGN UEY 𑜉𑜅𑜆	19CB

Table 1. Recommended reassignment of code points for diphthongs

Names of characters

There are four characters for which the names proposed in N2634 should be revised:

19BC TAI LUE VOWEL SIGN VOWEL SHORTENER 𑜉𑜂𑜆. This character, like the cognate character in Thai and Lao scripts, has two functions: it can be used to represent the short vowel /a/ in open syllables, and it can be used in combination with other vowel signs to denote shortened vowels. Use in an open syllable to represent the short vowel /a/ is seen in the following examples:

³ The chart in Figure 2 has a minor typographic error in that, for the diphthong uui, they have used the character 𑜉𑜂𑜆, whereas it should be 𑜉𑜃𑜆.

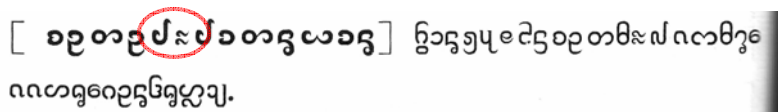


Figure 4. Short /a/ in open syllable: ꠘ (Jinxhiang 1994, p.50)

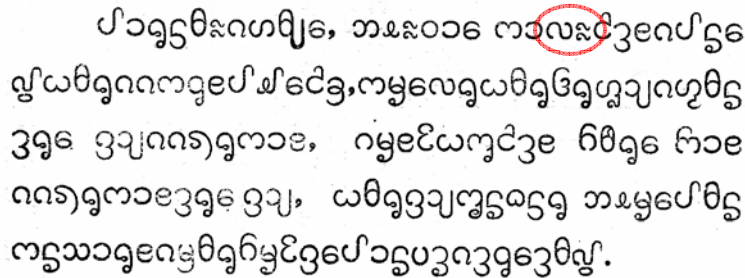


Figure 5. Short /a/ in open syllable: ꠘ (LA1, p. 35)

The primary identity of this Tai Lue character (as for the corresponding Thai and Lao characters) is in its function of representing short /a/. For instance, note that, in the vowel charts in Figure 1, Figure 2 and Figure 3, it is presented as short /a/, and is paired with the sign for long /a:/, ꠘ. Thus, I recommend that this character be renamed as TAI LUE VOWEL SIGN A.

19C2 TAI LUE VOWEL SIGN E ꠘ. This character represents a distinct vowel quality then 19C1 TAI LUE VOWEL SIGN EE ꠘ, not a shortened counterpart to it as the name suggests. I recommend that this character be renamed as TAI LUE VOWEL SIGN AE.

19DE TAI LUE SIGN AND ꠘ and **19DF TAI LUE SIGN RESULTATIVE** ꠘ. These characters are abbreviations for the frequently-used morphemes /ɛ/ ‘and’, and /ɛ:w/. The latter is a particle that indicates perfective aspect (completion), not a resulting state.⁴ Thus, if the name of the character were to be based on its meaning, it should be called “perfective” rather than “resultative”.

Rather than basing names of these characters on English glosses of the morphemes they denote, however, I recommend that they have names indicative of their pronunciation. This would be consistent with the names used in earlier proposals submitted by the China national body (N2242, N2371), and in earlier contributions from Michael Eversion (N2044) and the current author (L2/99-243). Thus, 19DE TAI LUE SIGN LAE, and 19DF TAI LUE SIGN LAEW.

Conclusion

With the exception of a few details that have been noted, this proposal is sound. I recommend that it be approved by WG2 and UTC with the recommended revisions discussed herein.

⁴ The term *resultative* is used in linguistics to refer to an expression that describes an action along with a resulting state caused by that action; e.g., “He boiled the kettle dry.”

References

- [SCEM] 中国社会科学院民族研究所, 国家民族事务委员会文化宣传司 (Zhongguo shehui kexueyuan minzu yanjiusuo, guojia minzu shiwu weiyuanhui wenhua xuanchuansi = Chinese Academy of Social Sciences, Institute for Nationality Studies; Department of Cultural Propagation, State Ethnic Affairs Commission). 1991. 中国少数民族文字 (*Zhongguo shaoshu minzu wenzi = Scripts of Chinese Ethnic Minorities*). 北京 (Beijing): 中国藏学出版社 (Zhongguo zangxue chubanshe = China Tibetan Studies Publishing House).
- [PSCC] 戴庆厦 许寿椿 高喜奎 主编 (Dai Qingsha, Xu Shouchuan, Gao Xikui, editors-in-chief). 1991. 中国各民族文字与电脑信息处理 (*Zhongguo ge minzu wenzi yu diannao xinxi chuli = Processing the scripts of China on the computer*). 北京 (Beijing): 中央民族学院出版社 (Zhongyang minzu xueyuan chubanshe = Central Nationalities Institute Publishing House).
- ၆၇၅၆၅၆၅၆ (搜集整理者 = Dao Jinxiang). 1994. ဣၼၼၼၼၼၼၼၼၼၼၼၼ (古傣语词语译翻 = *Dictionary of ancient Dai words and expressions*). 昆明 (Kunming): ယၼၼၼၼၼၼၼၼၼၼၼၼ (云南民族出版社出版 = Yunnan Nationalities Press).
- [LA1] 西双版纳州教育局编 (Xishuangbanna Prefectural Education Bureau). 1980. ဣၼၼၼၼၼၼၼၼ, ဣၼၼၼၼၼၼ (语文, 第一册 = Language arts, book 1). (西双版纳傣文小学课本 = Xishuangbanna Dai primary school textbook.) 昆明 (Kunming): 云南民族出版社出版 (Yunnan Nationalities Press).
- L2/99-243 Constable, Peter. 1999. "Proposal for encoding New Tai Lue script in Unicode/ISO-IEC 10646."
- N2242 China. 2000. "Proposal for encoding Xishuang Banna Dai script on BMP of ISO/IEC 10646."
- N2371 China. 2001. "Proposal summary form for Dai scripts."
- N2634 Eversion, Michael. 2003. "Proposal to encode the Tai Lue script in the BMP of the UCS."