ISO/IEC JTC1/SC2/WG2 N2748
2004-4-29

Universal Multiple Octet Coded Character Set
International Organization for Standardization
Organisation internationale de normalisation
Международная организация по стандартизации

Doc Type: **Working Group Document**

Title: Proposal on encoding New Tai Lue

Source: **China**
Status: **Expert contribution**
Action: **For consideration by JTC1/SC2/WG2**

The New Tai Lue is improved and simplified from traditional Dai scripts, mainly used in Xishuanbanna Prefecture, China. From 1950s', a great number of books, magazines and newspapers in Dai language have been published. It is also used in school education. New Tai Lue is a main tool to write and read for Dai people.

The simplified New Tai Lue has 42 consonants, which are divided into 2 groups: high tone and low tone. Compound consonants those already not been used in oral language are canceled, while 4 labialised consonants are kept. There are 9 monophthongs dividing into long and short. The short vowels are presented by combining the long vowels with the symbol " ". There are 13 diphthongs. New Tai Lue has 7 consonantal codas, represented by a small circle in the end of related consonants. There are 6 checked syllable tones and 3 unchecked syllable tones in New Tai Lue. There are two symbols represent the other tones besides the high tone and low tone of consonants.

Encoding ways

- **Cosonants**

Most consonants of New Tai Lue scripts are single symbols besides 10 consonants are combined with two symbols. The low tone consonants are combined by high tone consonant with a hat "^" over it. The consonants are represented by high tone consonants with the

symbol “ᨠ” under it. The 7 low tone consonants mentioned above are corresponding to high tone consonants. They belong to same phoneme. The high tone consonant ᨡ also is combined by consonant with symbol “ᨠ” under it, but they are not corresponding. The mood particle ᨣ is combined by another mood particle ᨤ with the symbol “ ᨠ ” under it. There are 2 combined consonants ᨡ、ᨡ that the shapes are changed. It’s better to keep the combined characters rather than design an underneath symbol for each character. Based on the situation described above, we encode New Tai Lue in separate way, i.e., use the consonants already existing combine with overhead symbol and underneath symbol to represent the combined characters. We don’t encode the combined characters as single character. This principle is widely accepted by Dai people. In Dai custom, these characters are not independent, but combined characters.

- **Vowels**

Monophthongs also are encoded in combining way. Long vowels combine with symbol “ —ᨠ ”represent short vowels. Vowels combined with two characters are encoding in combining way. For example, the vowel “ᨣᨣ— ”is encoded by combining two single vowel “ ᨣ— ”, vowel “ᨣ—ᨦ ” is encoded by combining single vowel “ ᨣ— ” and “ —ᨦ ”. The diphthong “ ᨣ—ᨧ ” is encoded by combining “ ᨣ— ” and “ —ᨧ ”. This is the best way to encode the vowels in New Tai Lue. In some proposals on encoding New Tai Lue, the encoding way is not consistent. For example, in proposal N2634, vowels “ ᨣ—ᨦ ” and “ ᨣ—ᨧ ” are encoded in combining way, while the vowel “ ᨣᨣ— ” is not.

- **Symbols**

There are two symbols in the proposal, one is overhead mark “^”, an other is underneath mark “ ᨠ ”. New characters are presented by combining the symbols with existing consonants.

Ordering of characters

- **Consonants**

The consonants are ordered according to several properties in Dai tradition.

- Point of articulation: glottal, velar, apical, labial
- Manner of articulation: plosive (unaspirated, aspirated), affricate (unaspirated, aspirated), nasal, fricative, semi-vowel
- Tone class: high, low

The traditional way of ordering the consonants is divided the 42

consonants into 8 groups according to the point of articulation. The consonants with same point of articulation (usually 3 consonants) are taken as one sub-group. Among the sub-group the consonants are ordered according to the manner of articulation. At last, the consonants are ordered according to the tone class. The group of consonants with high tone is listed before the group of consonants with low tone. In New Tai Lue, 6 out of the 8 original consonants are canceled, only the high glottal and low glottal are kept. This ordering has linguistic continuity. In language study practice in hundreds of years, this ordering is considered the best one for mastering the point of articulation, the manner of articulation and the inner principle of the language.

- **Vowels**

The vowels are ordered in the turn of monophthong, diphthong and consonantal coda. The vowels can be put before or after a consonant in a syllable, but the consonantal coda should be after any consonant and vowel. The consonantal coda is the changing of the shape of corresponding consonant, i.e., to add a circle in the end. It belongs to either consonant character or vowel character. It is a special symbol. There for the consonantal coda should be ordered after the characters, follow the compound vowels.

- **Other symbols**

The tone marks are ordered after the vowels. Then come the ten digits listed as 1, 2, 3, 4, 5, 6, 7, 8, 9. In New Tai Lue, the shape and size of the digit "0" is simillar with the character " ᧐ ". So they are encoded as one character. The last three are the character of mood particle, the overhead symbol and the underneath symbol.

Spelling rule

The spelling of New Tai Lue takes the consonants as the centre. Vowels can be put before or after consonants. When a consonant is spelling with a consonantal coda, and with tone mark in the end, the spelling order should be as following:

- if the vowel is before the consonant,

vowel + consonant + consonantal coda + tone mark

- if the vowel is after the consonant,

consonant + vowel + consonantal + tone mark

About the font

New Tai Lue has many kinds of fonts. The font in this proposal is widely used in China at the present time. It is developed by Beidafanzheng Group and used by Yunnan Nationality Publishing House.

Row XX : NEW TAI LUE

Dec	hex	Name
001	81	TAI LUE LETTER HIGH OA
002	82	TAI LUE LETTER HIGH KA
003	83	TAI LUE LETTER HIGH XA
004	84	TAI LUE LETTER HIGH NGA
005	85	TAI LUE LETTER LOW KA
006	86	TAI LUE LETTER LOW XA
007	87	TAI LUE LETTER LOW NGA
008	88	TAI LUE LETTER HIGH TSA
009	89	TAI LUE LETTER HIGH SA
010	8A	TAI LUE LETTER HIGH YA
011	8B	TAI LUE LETTER LOW TSA
012	8C	TAI LUE LETTER LOW SA
013	8D	TAI LUE LETTER LOW YA
014	8E	TAI LUE LETTER HIGH TA
015	8F	TAI LUE LETTER HIGH THA
016	90	TAI LUE LETTER HIGH NA
017	91	TAI LUE LETTER LOW TA
018	92	TAI LUE LETTER LOW THA
019	93	TAI LUE LETTER LOW NA
020	94	TAI LUE LETTER HIGH PA
021	95	TAI LUE LETTER HIGH PHA
022	96	TAI LUE LETTER HIGH MA
023	97	TAI LUE LETTER LOW PA
024	98	TAI LUE LETTER LOW PHA
025	99	TAI LUE LETTER LOW MA
026	9A	TAI LUE LETTER HIGH FA
027	9B	TAI LUE LETTER HIGH LA
028	9C	TAI LUE LETTER LOW FA
029	9D	TAI LUE LETTER LOW VA
030	9E	TAI LUE LETTER LOW LA
031	9F	TAI LUE LETTER HIGH HA
032	A0	TAI LUE LETTER HIGH DA
033	A1	TAI LUE LETTER HIGH BA
034	A2	TAI LUE LETTER LOW HA
035	A3	TAILUE VOWEL SIGN VOWEL SHORTENER
036	A4	TAI LUE VOWEL SIGN AA
037	A5	TAI LUE VOWEL SIGN II
038	A6	TAI LUE VOWEL SIGN U
039	A7	TAI LUE VOWEL SIGN UU

Dec	hex	Name
040	A8	TAI LUE VOWEL SIGN E
041	A9	TAI LUE VOWEL SIGN OO
042	AA	TAI LUE VOWEL SIGN O
043	AB	TAI LUE VOWEL SIGN UE
044	AC	TAI LUE VOWEL SIGN AY
045	AD	TAI LUE VOWEL SIGN AAY
046	AE	TAI LUE VOWEL SIGN UY
047	AF	TAI LUE VOWEL SIGN OOY
048	B0	TAI LUE VOWEL SIGN OY
049	B1	TAI LUE VOWEL SIGN UEY
050	B2	TAI LUE VOWEL SIGN IY
051	B3	TAI LUE LETTER FINAL V
052	B4	TAI LUE LETTER FINAL NG
053	B5	TAI LUE LETTER FINAL N
054	B6	TAI LUE LETTER FINAL M
055	B7	TAI LUE LETTER FINAL K
056	B8	TAI LUE LETTER FINAL D
057	B9	TAI LUE LETTER FINAL B
058	BA	TAI LUE TONE MARK-1
059	BB	TAI LUE TONE MARK-2
060	BC	TAI LUE DIGIT ONE
061	BD	TAI LUE DIGIT TWO
062	BE	TAI LUE DIGIT THREE
063	BF	TAI LUE DIGIT FOUR
064	C0	TAI LUE DIGIT FIVE
065	C1	TAI LUE DIGIT SIX
066	C2	TAI LUE DIGIT SEVEN
067	C3	TAI LUE DIGIT EIGHT
068	C4	TAI LUE DIGIT NINE
069	C5	TAI LUE SIGN AND
070	C6	TAI LUE COMPONENT PART OF LOW VOICE LETTERS
071	C7	TAI LUE COMPONENT PART

ROW ××: NEW DAILUE

	××8	××9	××A	××B	××C	××D
0	000	016	032	048	064	
1	001	017	033	049	065	
2	002	018	034	050	066	
3	003	019	035	051	067	
4	004	020	036	052	068	
5	005	021	037	053	069	
6	006	022	038	054	070	
7	007	023	039	055	071	
8	008	024	040	056		
9	009	025	041	057		
A	010	026	042	058		
B	011	027	043	059		
C	012	028	044	060		
D	013	029	045	061		
E	014	030	046	062		
F	015	031	047	063		

****

****

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646

A. Administrative

1. Title:
Proposing on Encoding New Tai Lue
2. Requester's name:
China
3. Requester type (Member body/Liaison/Individual contribution):
Member body
4. Submission date:
2004/4/13
5. Requester's reference (if applicable):
N967, N2044, N2242R, N2634
6. Choose one of the following:
This is a complete proposal:
Yes

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:
New Tai Lew
b. The proposal is for addition of character(s) to an existing block:
No

2. Number of characters in proposal:
71

3. Proposed category (select one from below - see section 2.2 of P&P document):
A-Contemporary

4. Proposed Level of Implementation (1, 2 or 3)
(see Annex K in P&P document):
Level 2
Is a rationale provided for the choice?
Yes
If Yes, reference:
New Tai Lue requrees Level 2 implementation as other Brahmic scripts do.

5. 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

6. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

PUC FOUNDER (MSC) BERHAD （Beijing Founder Electronic Co., Ltd）

If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

yjh@founder.com.cn , No.9, 5 Street, Shangdi, Information Industry Base, Haidian District, Beijing 100085, China

7. 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

8. 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)?

No

9. 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 http://www.unicode.org/Public/UNIDATA/UCD.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? Yes

If YES explain

This is a revised proposal

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?
Yunnan Nationality Publishing Company, Tai language group of Yunnan people's broadcasting station
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:
New Tai Lew is used by Dai people in Yunnan province. See N 2242R for more information

4. The context of use for the proposed characters
(type of use; common or rare)
It is widely used to write Tai Lue language by Dai people.
Reference:
_____

5. Are the proposed characters in current use by the user community?
Yes
If YES, where? Reference:
Yunnan Province, China

6. After giving due considerations to the principles in the P&P document
must the proposed characters be entirely in the BMP?
Yes
If YES, is a rationale provided?
Yes
If YES, reference:
Contemporary use and accordance with the roadmap.

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 an existing character?
No

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:
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
No

12. Does the proposal contain characters with any special properties such as control function or similar semantics?
No

13. Does the proposal contain any Ideographic compatibility character(s)?
No

(၅) ကြာရှေ့ ဂလမ္ဘေ ဘန်း ယာဒှ ငါ့ခွေ မှာ ဣာဒျ ကာဒှ ဈဉ်မှေ ဈကှ သပူ့ ကဒှ

ဥဒကှ စာကှ ဂသ ငြ ဂမှ ဂပူ့ ဈကှ , ဗျာေ ဍာေ လကှ ဈဉ်မှေ ဈကှ သပူ့ ကဒှ ;

ဥပှ ဍဖြ ဂဒရှ ဘဉာဒှ ယာ , ပရှေ ခြာ့ ငတ ;

သာရှေ ဍပူ့ ဂပူ့ ဈကှ ဗျာေ သာရှေ ဣာဒျ ဏပူ့ , ဒဖဒှေ ဍပူ့ ဈဉ်မှေ ဈကှ ဗျာေ ဣာဒျ ဂဉာဒှ , ဗျာေ ဇဒေ မှဖြ ကာဒှ ဘိုဒှ ပရှေ ဘဉရှ ကဒှ ဂဖြေ ဂဒေ ဈဉ်မှေ ဈကှ သပူ့ ကဒှ ဥဒှေ ;

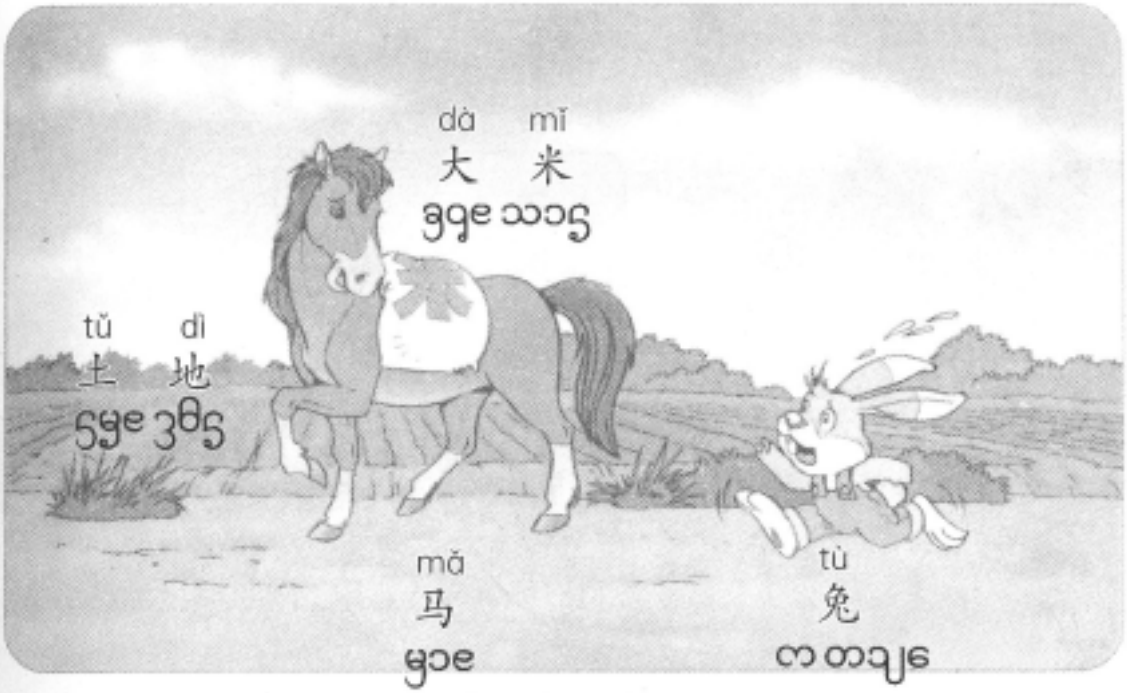
ငြ ဂမှ စကှ ဇဒှေ ဈကှ ဂဂဣာရှ ကဒှ ဗဉာရှ ဗမှ ကဒှ ။ ဂဒှ ဝာေ ဣာဒျ ဗာကှ ၈ စဒေ ဂမှေ ဘန်း ယာဒှ ငလကှ ဂဘဒှေ လာမှ စပှ , ငြ ဂမှ ဂမှဖြေ ဒြ ဈဉ်မှေ ဈကှ သပူ့ ကဒှ ယာမှ ဥဒှေ စကှ ဇဒှေ ဇဒှေ ရဉာရှ ဗာရှ ဘိုဒ ဥဒ ဇမှ ဍပူ့ ဂဖြေ ;

ပရှေ ဈဉ်မှေ ဂဂြရှေ သပူ့ ကဒှ , ဂဂြရှေ သပူ့ စကှ ဇဒှေ ဥန်း ဥရျေ ငဣာဒှ ဇသ ;

ကကှ ဖန်း ကာဒှ ဈဉ်မှေ ဈကှ သပူ့ ကဒှ ဘိုဒှ ပရှေ ငဣာဒှ ဇသ , စကှ ဇဒှေ ဇဒှေ ရဉာရှ ဗာရှ ဘိုဒ ဥဒ ဇမှ ဍပူ့ ဂဖြေ , ကြာရှေ ကာဒှ ဂဣာဖြေ ဈဉ်မှေ ဈကှ သပူ့ ကဒှ ဎ် ဣာဒ ဇဒှေ ဘန်း ယာဒှ ငလကှ ဂဘဒှေ ဎ် ဘန်း ယာဒှ ဂဂမှရှ ။

ဘဒ္ဒလင်ဖရုင်မှ

	ဘဒ္ဒ			ဘဒ်ဒ်	
က	ခ	ဟ္ခ	ဂ	ဂ်	၅
ဝ	သ	ဟ်	ဒ	ဒ်	ယ
တ	၅	ဟ္ဒ	ဖ	ဓ	၄
ပ	ဗ	ဟ္ဗ	ဟ္ဗ	ဘ	မ
စ်	ဟ္ဗ	ဟ္ဗ	ကြ	ဝ	လ
ဟ	ဒ	ပ	၈	ဒ်ဒ်	ပ်ပ်
ဟ္ဗ	ခ္ဓ		ဂ္ဓ	ဂ္ဓ်	



qīng qīng de

轻 轻 地

xiǎo tù xiǎo tù qīng qīng tiào
小 兔 小 兔 轻 轻 跳，
xiǎo gǒu xiǎo gǒu màn màn pǎo
小 狗 小 狗 慢 慢 跑，
yào shì cǎi téng le xiǎo cǎo
要 是 踩 疼 了 小 草，
wǒ jiù bù gēn nǐ men hǎo
我 就 不 跟 你 们 好。

大 米 土 地 马

儿歌作者郑春华。