

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: Working Group Document
Title: Response to Chinese contribution N3238, “Proposing on Encoding Old Tai Lue”
Source: Irish and UK National Bodies
Status: National Body contribution
Action: For consideration by JTC1/SC2/WG2
Date: 2007-04-11

The Irish and UK National Bodies have reviewed the Chinese National Body’s contribution N2328 and would like to make the following comments.

N3238 proposes that the encoding for Old Tai Lue, and therefore for Northern Thai and Khün, be unified with the New Tai Lue encoding already in Unicode (U+1980–U+19DF). As was already stated in N3169 (“Lanna ad-hoc report” by Chen Zhuang, Michael Everson, Martin Hosken, Wei Lin-Mei), the encoding model for New Tai Lue would have to be very different to cater for the needs of Old Tai Lue. New Tai Lue simplifies the script immensely following a basically linear structure having no diacritics and only reordered vowels. Old Tai Lue, on the other hand has a complex subjoining model, whether flattened as N3238 proposes or made explicit as in N3207 (“Revised proposal for encoding the Lanna script in the BMP of the UCS”, by Michael Everson, Martin Hosken, and Peter Constable).

In examining N3238 we find that nearly everything listed can already be encoded using N3207 and there is much that is in N3207 that is missing from N3238. N3238 lists a number of glyph variants, giving them separate codes. The question remains as to whether certain sets of codes are variants or are truly contrastive. We hope in the ensuing discussion to be able to resolve whether the following sets are merely font or context variants of each other or need to be encoded separately in some way.

- 1A94 *low la* and 1A95 *low la* (what is 1A95?)
- 1AA5 *vowel sign ooy* and 1AA6 *vowel sign ooy*, and with them 1AA8 *part y* and 1AA9 *part y*.
- 1AAB *part p*, 1AAC *part th* and 1AAD *part p*
- 1AB6 *final n* and 1AB7 *final n* (what is 1AB7?)
- 1ABD *final m* and 1ABE *final m*
- 1AC7 *final t* and 1AC8 *final t*
- 1AD0 *final t* and 1AD3 *final t*
- 1AD4 *final p* and 1AD9 *final p*
- 1AE5 *digit one* and 1AE4 *digit one* or 1AE3 *digit one*.

It would greatly help discussion if examples of the existence of two characters in a set existing in the same text could be found, Otherwise we must presume that the characters in a set are variants of the same character. We give below our analysis of the mappings between N3238 and N3207:

1A80	high i	1A50	LETTER I
1A81	low ii	1A51	LETTER II
1A82	low uu	1A53	LETTER UU
1A83	low ee	1A54	LETTER EE
1A84	high xa	1A22	LETTER HIGH KXA
1A85	low xa	1A25	LETTER LOW KHA
1A86	high sa	1A29	LETTER KHUN HIGH CHA
1A87	high sa	1A5F	LETTER GREAT SA
1A88	high sa	1A49	LETTER HIGH SHA
1A89	high sa	1A4A	LETTER HIGH SSA
1A8A	high sa	1A2E	LETTER LOW CHA
1A8B	high ya	1A4C	1A60 1A43 la + sakot + high ya
1A8C	low ya	1A2F	LETTER LOW NYA
1A8D	high ta	1A30	LETTER RATA
1A8E	high tha	1A31	LETTER HIGH RATHA
1A8F	low tha	1A33	LETTER LOW RATHA
1A90	low na	1A34	LETTER RANA
1A91	low na	1A39	1A63 LETTER NA (var. NA? or presentation form NAA?)
1A92	low pa	1A3E	LETTER LOW PA
1A93	low ha	1A56	LETTER LOW HA
1A94	low la	1A4D	LETTER LLA
1A95	low la	1A4D,	style variant of letter lla?
1A96	vsign aa	1A64	VOWEL SIGN TALL AA
1A97	vsign i	1A65	VOWEL SIGN I
1A98	vsign ii	1A66	VOWEL SIGN II
1A99	vsign ue	1A67	VOWEL SIGN UE
1A9A	vsign ue	1A68	VOWEL SIGN UUE
1A9B	vsign ue	1A45	LETTER RUE
1A9C	vsign u	1A69	VOWEL SIGN U
1A9D	vsign u	1A69	VOWEL SIGN U
1A9E	vsign u	1A69	VOWEL SIGN U
1A9F	vsign uu	1A6A	VOWEL SIGN UU
1AA0	vsign uu	1A6A	VOWEL SIGN UU
1AA1	vsign uu	1A6A	VOWEL SIGN UU
1AA2	vsign oa	1A73	VOWEL SIGN OA ABOVE
1AA3	vsign oa	1A6C	VOWEL SIGN OA BELOW
1AA4	vsign oa	1A6C	VOWEL SIGN OA BELOW
1AA5	vsign ooy	1A60	1A48 1A60 1A43 sakot + wa + sakot + high ya
1AA6	vsign ooy	1A60	1A48 1A60 1A43, var. sakot + wa + sakot + high ya?
1AA7	vsign ooy	1A60	1A48 1A60 1A43 sakot + wa + sakot + high ya
1AA8	part y	1A60	1A43 sakot + ya
1AA9	part y	1A60	1A43, variant of sakot + ya?
1AAA	part y	1A6D	VOWEL SIGN OY
1AAB	part p	1A5E	HIGH RATHA OR LOW PA
1AAC	part th	1A5E,	variant or high ratha or low pa?
1AAD	part p	1A5E,	variant of high ratha or low pa?
1AAE	part v	1A60	1A48 sakot + wa
1AAF	final ng	1A74	SIGN MAI KANG
1AB0	final ng	1A5A	SIGN MAI KANG LAI
1AB1	final ng	1A65	1A74 vowel sign i + sign mai kang
1AB2	final ng	1A6B	VOWEL SIGN O
1AB3	final n	1A60	1A39 sakot + na
1AB4	final n	1A59	CONSONANT SIGN MEDIAL LA
1AB5	final n	1A59	CONSONANT SIGN MEDIAL LA
1AB6	final n	1A60	1A46 sakot + la
1AB7	final n	1A60	1A4b, variant of sakot + la?
1AB8	final n	1A60	1A44 sakot + ra
1AB9	final n	1A60	1A2F sakot + low nya
1ABA	final n	1A34	1A60 1A2F rana + sakot + low nya
1ABB	final n	1A60	1A34 sakot + rana
1ABC	final m	1A63	1A74 sakot + vowel sign aa + sign mai kang
1ABD	final m	1A60	1A41 sakot + ma

1ABE	final	m	1A60	1A41	variant of sakot + ma?
1ABF	final	k	1A60	1A20	sakot + high ka
1AC0	final	k	1A60	1A23	sakot + low ka
1AC1	final	k	1A60	1A21	sakot + high kha
1AC2	final	k	1A60	1A25	sakot + low kha
1AC3	final	t	1A60	1A32	sakot + da
1AC4	final	t	1A60	1A27	sakot + high ca
1AC5	final	t	1A60	1A29	sakot + khun high cha
1AC6	final	t	1A60	1A2A	sakot + low ca
1AC7	final	t	1A60	1A2E	sakot + low cha
1AC8	final	t	1A60	1A2E	variant of sakot + low cha
1AC9	final	t	1A60	1A30	sakot + rata
1ACA	final	t	1A60	1A31	sakot + high ratha
1ACB	final	t	1A60	1A33	sakot + low ratha
1ACC	final	t	1A60	1A35	sakot + high ta
1ACD	final	t	1A60	1A36	sakot + high tha
1ACE	final	t	1A60	1A37	sakot + low ta
1ACF	final	t	1A60	1A38	sakot + low tha
1AD0	final	t	1A60	1A4B	variant of sakot + high sa ?cf 1AD3
1AD1	final	t	1A60	1A49	sakot + high sha
1AD2	final	t	1A60	1A4A	sakot + high ssa
1AD3	final	t	1A60	1A4B	sakot + high sa
1AD4	final	p	1A60	1A3C	sakot + high pha
1AD5	final	p	1A60	1A3B	sakot + high pa
1AD6	final	p	1A60	1A3B	sakot + high pa
1AD7	final	p	1A60	1A3E	sakot + low pa
1AD8	final	p	1A60	1A40	sakot + low pha
1AD9	final	p	1A60	1A3C	variant of sakot + high pha ?cf 1AD4
1ADA	tone-1		1A75	SIGN	TONE-1
1ADB	tone-2		1A76	SIGN	TONE-2
1ADC	tone-2		1A62	1A75	vowel sign mai sat + sign tone-1
1ADD	tone-2		1A62	1A76	vowel sign mai sat + sign tone-2
1ADE	part hong		1A58	CONSONANT	SIGN MEDIAL RA
1ADF	wordrepeat		1A7B	SIGN	MAI SAM
1AE0	punct		1A7A	SIGH	RA HAAM
1AE1	punct		1AA7	SIGN	MAI YAMOK
1AE2	punct		1AAA	SIGN	SATKAAN
1AE3	digit	1	1A81	DIGIT	ONE
1AE4	digit	1	1A91	THAM	DIGIT ONE
1AE5	digit	1	1A49	LETTER	HIGH SHA
1AE6	digit	2	1A92	THAM	DIGIT TWO
1AE7	digit	3	1A93	THAM	DIGIT THREE
1AE8	digit	4	1A94	THAM	DIGIT FOUR
1AE9	digit	5	1A96	THAM	DIGIT SIX
1AEA	digit	6	1A95	THAM	DIGIT FIVE
1AEB	digit	7	1A97	THAM	DIGIT SEVEN
1AEC	digit	8	1A98	THAM	DIGIT EIGHT
1AED	digit	9	1A89	DIGIT	NINE
1AEE	digit	9	1A99	THAM	DIGIT NINE