

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

**Doc Type: Working Group Document****Title: Documentation on legacy encodings of the Old Lisu script****Author: China****Status: Member Contribution****Action: For consideration by JTC1/SC2/WG2 and UTC****Date: 2007-12-29**

**1. Introduction.** This document is written as part of the discussion surrounding L2/07-344. It describes three legacy encodings (fonts) of the Old Lisu script. All three were created in Thailand, one by Andy G. Thomson and two by David L. Morse. There are reportedly several other fonts created in China in use by two Lisu newspapers (*Tuanjie Bao* and *Nujiang Bao*) in Yunnan, but no encoding details are available.

**2. Tag Ledza Lisu.** Andy Thomson created this font between 1995 and 1997. There may be four or five type styles that he designed, and this is one of them. Table 1 shows the font encoding. It can be observed that:

- (1) Old Lisu characters are encoded in alphabetical order leaving no ASCII semantics in the range 0040..007E.
- (2) Tone letters (0068..006D) are separately encoded with tones *mya cya* and *mya bo* assigned their own code points (006A and 006B) and the rest of the set designed to look different from ASCII punctuation 002C, 002E, 003A and 003B.

There are many publications using this font. An organisation in Chiang Mai, Thailand has a small hymn book, some Bible-study books and a book translated from English. A literature centre near Mandalay, Myanmar has a New Testament ready for printing as well as a few booklets.<sup>1</sup> David Bradley of La Trobe University, Australia has about 700 pages of materials published using a Macintosh version of this font. These include a song book (Bradley 2000) and a dictionary of about 400 pages (Bradley 2005). In the original PC version, Thomson published a number of primers, readers, glossaries of Bible terms and other literature covering about 500 pages, and there was a New Testament done some years ago. Various Lisu people that Bradley knows still use this font and have published a variety of commentaries and other Christian literature in Thailand and Myanmar covering at least another 500 pages. Altogether there are about 2,000 pages of materials recorded and published in this encoding.

**3. Lisu-AC1.** This font was created between 1993 and 1996 and is one in an earlier series of fonts that David Morse designed. The series employs an encoding that was in use from the days of Apple II to early PC days (up to Microsoft Windows 98). Table 2 shows the font encoding. It can be observed that:

- (1) Old Lisu characters are encoded in sorting order without retaining any ASCII semantics in the ranges 0041..005A and 0061..007A.
- (2) Tone letters (006F..0074) are separately encoded from ASCII punctuation 002C, 002E, 003A and 003B with tones *mya cya* and *mya bo* assigned their own code points (0071 and 0072).

---

<sup>1</sup> The centre switched to using David Morse's *F* font series (Section 4) in 2003 for ease of typing.

Morse used to have data for about 120 books recorded in this encoding, but he later converted them into a new encoding. According to Morse, these earlier fonts were not very widely used, since they required a keyboarding program (KRF, licensed from Canada) and users found it cumbersome to set up. All users have now reportedly switched to Morse's *F* font series with a new encoding as described below.

**4. LisuFCA1.** This is one of the newer (*F*-series) fonts that David Morse designed. Based on real-life publication usage, the idea is to have a font where no additional keyboarding program is necessary and for which the input method is intuitive and easy to learn. Additionally, the goal is that the font can be used not only in Microsoft Word but also in Microsoft Excel and Adobe PageMaker with no data corruption or character shifts (especially with PageMaker) when files are exchanged between programs. Table 3 shows the font encoding. It can be observed that:

- (1) Old Lisu characters are encoded in a keyboard-friendly way without retaining any ASCII semantics in the ranges 0041..005A and 0061..007A.
- (2) Characters at positions 002C, 002E, 003A and 003B are used to represent tone letters in addition to their usual ASCII semantics.
- (3) Tones *mya cya* and *mya bo* have their own code points at 0048 and 0059, respectively.

This is the encoding into which Morse has converted all his older files (some 120 books) that used the old encoding (Section 3). There are currently 15 type styles, of which about 10 are still in trial mode. This Lisu *F* font series has been very well accepted and widely distributed to Lisu and other users in Thailand and Myanmar, and even to some Lisu in India as well as to the Lisu in China. A Lisu bi-monthly in Myanmar has been using this encoding since late 1997, producing cumulatively about 4,800 pages in all. A wide range of other commentaries, hymn books, primers and scripture reprints or revisions also use it. The total number of pages published in this encoding is certainly well over 6,000 pages. There is also continual progress going on. Morse has recently helped format a 980-page Bible that others translated and typed. His company has three new books ready for printing, one of them over 300 pages. They are also halfway finished with revising a 1,300-page Lisu Bible. Besides, there are two Bible-study books ready for printing in Myanmar, and a general health book in Lisu is being revised (about 280 pages). Altogether there are at least another 3,000 pages in various stages of the publication process.

**5. Conclusion.** All legacy encodings from Thailand encode the Old Lisu alphabet as a distinct set separate from ASCII Latin. Two out of three encodings are alphabetical and have separate code points for Old Lisu tone letters from 002C, 002E, 003A and 003B. One of these two is no longer in use. Several legacy fonts are being used in China, but no encoding information is available.

## 6. Bibliography

Bradley, David. 2005. *Southern Lisu dictionary*. ms. Published in James A. Matisoff, ed., *STEDT monograph series*, vol. 4 (Berkeley, CA: University of California, 2006).

\_\_\_\_\_. 2006. Personal interview by Adrian Cheuk. 10 April 2006.

Bradley, David, ed. 2000. *Lisu bride price song*. Transcribed by David Fish. Melbourne, Australia: La Trobe University.

## 7. Acknowledgements

This research was made possible by the help of SIL International, with Adrian Cheuk of SIL East Asia Group conducting the main research and drafting the document and Martin Hosken of SIL Non-Roman Script Initiative and Payap University, Thailand providing technical guidance. Special thanks go to David Morse, expert and mother-tongue Lisu speaker, and David Bradley, renowned scholar in Lisu and related languages, for information about and access to legacy fonts of the Old Lisu script.

# Tables

FONTLAB FONT TABLE															
Font: Tag-LedzaLisu															11/20/07 14:50:17
															Page 1/2
0020	0021	0022	0023	0024	0025	0026	0027	0028	0029	002A	002B	002C	002D	002E	002F
	!	"	#	\$	%	&	'	(	)	*	+	,	-	.	/
0030	0031	0032	0033	0034	0035	0036	0037	0038	0039	003A	003B	003C	003D	003E	003F
0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
0040	0041	0042	0043	0044	0045	0046	0047	0048	0049	004A	004B	004C	004D	004E	004F
B	P	D	D	T	L	G	K	K	J	C	C	Z	F	F	M
0050	0051	0052	0053	0054	0055	0056	0057	0058	0059	005A	005B	005C	005D	005E	005F
N	L	S	R	R	V	V	H	G	J	W	X	Y	A	V	E
0060	0061	0062	0063	0064	0065	0066	0067	0068	0069	006A	006B	006C	006D	006E	006F
E	I	O	U	U	L	D	B	.	,	..	.,	:	;	-	=
0070	0071	0072	0073	0074	0075	0076	0077	0078	0079	007A	007B	007C	007D	007E	007F
-.	-	'	,	'	"	;	-	-	.	ˆ	:	ƒ			
FFFF0000	FFFF0001	FFFF0002	00A2	00A3	00A4	00A5	00A6	00A7	00A8	00AC	00B0	00B6	00B7	0152	0153
□			"	„	ø				}	└	□	¶	.	)	∩
0160	0161	0192	02C6	02DC	2013	2014	2018	2019	201A	201C	201D	201E	2020	2021	2022
-3-	0	_	—	A	-	—	'	'	┌	“	”	—	┌	—	•

Table 1: Encoding of Tag Ledza Lisu by Andy G. Thomson. Note the range 0040..007E in which no ASCII semantics are retained.

0020	0021	0022	0023	0024	0025	0026	0027	0028	0029	002A	002B	002C	002D	002E	002F
	!	"	#	..	.,	-	'	(	)	*	+	,	-	.	/
0030	0031	0032	0033	0034	0035	0036	0037	0038	0039	003A	003B	003C	003D	003E	003F
0	1	2	3	4	5	6	7	8	9	:	;	..;	=	.,;	?
0040	0041	0042	0043	0044	0045	0046	0047	0048	0049	004A	004B	004C	004D	004E	004F
-.	B	P	d	D	T	l	G	K	K	J	C	C	Z	F	F
0050	0051	0052	0053	0054	0055	0056	0057	0058	0059	005A	005B	005C	005D	005E	005F
M	N	L	S	R	R	v	V	H	G	r	[	;	]	^	_
0060	0061	0062	0063	0064	0065	0066	0067	0068	0069	006A	006B	006C	006D	006E	006F
,	X	W	Y	A	V	E	I	E	O	U	U	L	D	B	.
0070	0071	0072	0073	0074	0075	0076	0077	0078	0079	007A	007B	007C	007D	007E	007F
,	..	.,	:	;	-	-.	=	-	'	Q	“	;	”	~	
FFFF0000	FFFF0001	FFFF0002	0080	0081	008D	008E	008F	0090	009D	009E	00A0	00A1	00A2	00A3	00A4
□			<sup>128</sup> *	<sup>129</sup> *	<sup>141</sup> *	<sup>142</sup> *	<sup>143</sup> *	<sup>3.</sup> Y	<sup>157</sup> *	<sup>158</sup> *		†	◌̇	<sup>3.</sup>	⌋
00A5	00A6	00A7	00A8	00A9	00AA	00AB	00AC	00AD	00AE	00AF	00B0	00B1	00B2	00B3	00B4
4̄	5̄	<sup>3.</sup> U	-	-	-	-	⌋	-	A	E	O	U	Y	1̄	2̄

Table 2: Encoding of Lisu-AC1 by David L. Morse. Note the loss of ASCII semantics in the ranges 0041..005A and 0061..007A.

0020	0021	0022	0023	0024	0025	0026	0027	0028	0029	002A	002B	002C	002D	002E	002F
	!	"	#	\$	%	&	'	(	)	*	+	,	-	.	/
0030	0031	0032	0033	0034	0035	0036	0037	0038	0039	003A	003B	003C	003D	003E	003F
0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
0040	0041	0042	0043	0044	0045	0046	0047	0048	0049	004A	004B	004C	004D	004E	004F
@	A	B	C	D	E	F	G	..	-.	J	K	L	■	●	-
0050	0051	0052	0053	0054	0055	0056	0057	0058	0059	005A	005B	005C	005D	005E	005F
P	★	R	;	T	U	V	;	"	.,	“	[	\	]	^	-
0060	0061	0062	0063	0064	0065	0066	0067	0068	0069	006A	006B	006C	006D	006E	006F
,	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
0070	0071	0072	0073	0074	0075	0076	0077	0078	0079	007A	007B	007C	007D	007E	007F
P	Q	R	S	T	U	V	W	X	Y	Z	{		}	~	
FFFF0000	FFFF0001	FFFF0002	FFFF0004	FFFF0005	00A4	00A7	00A9	00AB	00AC	00AE	00B6	00B7	00C8	00C9	00CA
□					¤	§	©	«	¬	®	¶	·		∞	∞
00D2	00D3	00D4	00D5	00D6	00D7	00DC	00DD	00DE	00DF	00E0	00E6	00E7	00E8	00E9	00EA
■	■	■	■	■	■	●	●	●	●	●	<u>0</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>

Table 3: Encoding of LisuFCA1 by David L. Morse. Note the same two ranges in which no ASCII sematics are retained.