



Title: Analysis of the Indian proposal to encode Lepcha script in the UCS  
 Source: Michael Everson  
 Action: For consideration by UTC  
 Date: 2005-02-04

My analysis of the proposal for Lepcha in L2/04-397 suggests that it is technically inappropriate for encoding. My grounds for saying so are as follows:

1. The organization of the code table is based on the Devanagari/Indic scripts encoded elsewhere in the standard. It is unnecessary to waste space by providing empty holes for characters which do not exist in Lepcha, particularly as space in the BMP is so rare.
2. The proposal contains a number of characters which can be represented by other characters in the proposal. Encoding the independent vowels would simply create normalization requirements which are otherwise not required for Lepcha. It is worth noting that the Shipmoo font used to prepare L2/04-397 does not contain precomposed independent vowels.
3. The Indian proposal presupposes a fully Brahmic-model for Lepcha, though the script is only Brahmic in having a set of dependent vowel matras. Lepcha does not have a virama, and its use of conjoined –YA and –RA, its use of final consonant letters, and its use of a set of complex consonant letters in a non-decomposable –LA are all departures from the original Brahmic model. Lepcha follows the Tibetan encoding model more closely than it does the Indic model; that in itself should inform us of how we need to represent the vowels in Lepcha.
4. The proposal omits LEPCHA LETTER WA and a number of punctuation characters.
5. The sorting order for Lepcha is syllabic; this can be approximated in the code table in the usual way with regard to the relative order of the main letters. The order is (approximately) *ka kak kang kat kan kap kam kar kal ki kik king kit kin kip kim kir kil ku kuk kung kut kun kup kum kur kul ke kek keng ket ken kep kem ker kel ko kok kong kot kon kop kom kor kol* followed by the same series in *kya*, followed by the same series in *kra*, followed by the same series in *krya*, followed by the same series in *kla*, followed by the same series in *klya*, then proceeding to the next letter *kha*.

The table below gives a list of the characters in the order they appear in L2/04-397, followed by the corresponding code positions in my current proposal to encode Lepcha. The code table for that proposal is attached.

			1CA2		LEPCHA LETTER WA
			1CBB	l	LEPCHA CHONG
			1CBC	ll	LEPCHA NYAT CHA CHONG
			1CBD	x	LEPCHA CER WA
			1CBE	⊗	LEPCHA CER WA TSHOK
			1CBF		LEPCHA TSHOK

1A01	◉ᄁ	LEPCHA SIGN NYINDO	1CB5	◉ᄁ	LEPCHA CONSONANT SIGN NYINDO
1A02	◉	LEPCHA SIGN KANG	1CB4	◉	LEPCHA CONSONANT SIGN KANG
1A04	ᄁ	Lepcha Letter SHORT A	1CA3	ᄁ	LEPCHA LETTER A
1A05	ᄁᄂ	Lepcha Letter A	1CA3 1CB6		
1A06	ᄁᄃ	Lepcha Letter AA	1CA3 1CA6		
1A07	ᄁᄄ	Lepcha Letter I	1CA3 1CA7		
1A08	ᄁᄅ	Lepcha Letter II	1CA3 1CA7 1CB6		
1A09	ᄁᄆ	Lepcha Letter U	1CA3 1CAA		
1A0A	ᄁᄇ	Lepcha Letter UU	1CA3 1CAB		
1A0F	ᄁᄈ	Lepcha Letter E	1CA3 1CAC		
1A13	ᄁᄉ	Lepcha Letter O	1CA3 1CA8		
1A14	ᄁᄊ	Lepcha Letter AU	1CA3 1CA9		
1A15	ᄁᄋ	Lepcha Letter KA	1C80	ᄁᄋ	LEPCHA LETTER KA
1A16	ᄁᄌ	Lepcha Letter KHA	1C82	ᄁᄌ	LEPCHA LETTER KHA
1A17	ᄁᄍ	Lepcha Letter GA	1C83	ᄁᄍ	LEPCHA LETTER GA
1A19	ᄁᄎ	Lepcha Letter NGA	1C85	ᄁᄎ	LEPCHA LETTER NGA
1A1A	ᄁᄏ	Lepcha Letter CA	1C86	ᄁᄏ	LEPCHA LETTER CA
1A1B	ᄁᄐ	Lepcha Letter CHA	1C87	ᄁᄐ	LEPCHA LETTER CHA
1A1C	ᄁᄑ	Lepcha Letter JA	1C88	ᄁᄑ	LEPCHA LETTER JA
1A1E	ᄁᄒ	Lepcha Letter NYA	1C89	ᄁᄒ	LEPCHA LETTER NYA
1A1F	ᄁᄓ	Lepcha Letter TTA	1C80 1CB7 1CA5		
1A20	ᄁᄔ	Lepcha Letter TTHA	1C9D 1CB7 1CA5		
1A21	ᄁᄕ	Lepcha Letter DDA	1C83 1CB7 1CA5		
1A24	ᄁᄖ	Lepcha Letter TA	1C8A	ᄁᄖ	LEPCHA LETTER TA

1A25	ཐ	Lepcha Letter THA	1C8B	ཐ	LEPCHA LETTER THA
1A26	ཎ	Lepcha Letter DA	1C8C	ཎ	LEPCHA LETTER DA
1A28	ཡ	Lepcha Letter NA	1C8D	ཡ	LEPCHA LETTER NA
1A2A	འ	Lepcha Letter PA	1C8E	འ	LEPCHA LETTER PA
1A2B	པ	Lepcha Letter PHA	1C90	པ	LEPCHA LETTER PHA
1A2C	ཅ	Lepcha Letter BA	1C93	ཅ	LEPCHA LETTER BA
1A2E	མ	Lepcha Letter MA	1C95	མ	LEPCHA LETTER MA
1A2F	ཙ	Lepcha Letter YA	1C9A	ཙ	LEPCHA LETTER YA
1A30	འ	Lepcha Letter RA	1C9B	འ	LEPCHA LETTER RA
1A32	ཟ	Lepcha Letter LA	1C9C	ཟ	LEPCHA LETTER LA
1A35	ཐ	Lepcha Letter VA	1C9F	ཐ	LEPCHA LETTER VA
1A36	ཅ	Lepcha Letter SHA	1CA1	ཅ	LEPCHA LETTER SHA
1A38	ཟ	Lepcha Letter SA	1CA0	ཟ	LEPCHA LETTER SA
1A39	ཏ	Lepcha Letter HA	1C9D	ཏ	LEPCHA LETTER HA
1A3C	འ	Lepcha Sign NUKTA	1CB7	འ	LEPCHA SIGN NUKTA
1A3D	འ	Lepcha Sign RAN	1CB6	འ	LEPCHA SIGN RAN
1A3E	འ	Lepcha Vowel Sign AA	1CA6	འ	LEPCHA VOWEL SIGN AA
1A3F	འ	Lepcha Vowel Sign I	1CA7	འ	LEPCHA VOWEL SIGN I
1A40	འ	Lepcha Vowel Sign II	1CA7 1CB6		
1A41	འ	Lepcha Vowel Sign U	1CAA	འ	LEPCHA VOWEL SIGN U
1A42	འ	Lepcha Vowel Sign UU	1CAB	འ	LEPCHA VOWEL SIGN UU
1A47	འ	Lepcha Vowel Sign E	1CAC	འ	LEPCHA VOWEL SIGN E
1A4B	འ	Lepcha Vowel Sign O	1CA8	འ	LEPCHA VOWEL SIGN O
1A4C	འ	Lepcha Vowel Sign AU	1CA9	འ	LEPCHA VOWEL SIGN OO
1A55	ཐ	Lepcha Letter TSA	1C97	ཐ	LEPCHA LETTER TSA
1A56	ཟ	Lepcha Letter TSHA	1C98	ཟ	LEPCHA LETTER TSHA

1A57	ཨ	Lepcha Letter DZA	1C99	ཨ	LEPCHA LETTER DZA
1A58	ཀ	Lepcha Letter KLA	1C81	ཀ	LEPCHA LETTER KLA
1A59	ཁ	Lepcha Letter GLA	1C84	ཁ	LEPCHA LETTER GLA
1A5A	ཎ	Lepcha Letter PLA	1C8F	ཎ	LEPCHA LETTER PLA
1A5B	ཌ	Lepcha Letter FLA	1C92	ཌ	LEPCHA LETTER FLA
1A5C	ཟ	Lepcha Letter BLA	1C94	ཟ	LEPCHA LETTER BLA
1A5D	ཡ	Lepcha Letter MLA	1C96	ཡ	LEPCHA LETTER MLA
1A5E	མ	Lepcha Letter HLA	1C9E	མ	LEPCHA LETTER HLA
1A5F	ཎ	Lepcha Letter FA	1C91	ཎ	LEPCHA LETTER FA
1A66	༠	LEPCHA DIGIT ZERO-TEE	1CC0	༠	LEPCHA DIGIT ZERO
1A67	༡	LEPCHA DIGIT ONE-KAAT	1CC1	༡	LEPCHA DIGIT ONE
1A68	༢	LEPCHA DIGIT TWO-NYET	1CC2	༢	LEPCHA DIGIT TWO
1A69	༣	LEPCHA DIGIT THREE-SAAM	1CC3	༣	LEPCHA DIGIT THREE
1A6A	༤	LEPCHA DIGIT FOUR-FALEE	1CC4	༤	LEPCHA DIGIT FOUR
1A6B	༥	LEPCHA DIGIT FIVE-FANGOO	1CC5	༥	LEPCHA DIGIT FIVE
1A6C	༦	LEPCHA DIGIT SIX-TARAOK	1CC6	༦	LEPCHA DIGIT SIX
1A6D	༧	LEPCHA DIGIT SEVEN-KAKYAOK	1CC7	༧	LEPCHA DIGIT SEVEN
1A6E	༨	LEPCHA DIGIT EIGHT-KAKU	1CC8	༨	LEPCHA DIGIT EIGHT
1A6F	༩	LEPCHA DIGIT NINE-KAKYAOT	1CC9	༩	LEPCHA DIGIT NINE
1A72	ཀྱ	Lepcha Consonant Sign K	1CAD	ཀྱ	LEPCHA CONSONANT SIGN K
1A73	ཀྲ	Lepcha Consonant Sign NG ( <i>duplicate of 1A02</i> )	1CB4	ཀྲ	( <i>Duplicate</i> )
1A74	ཀླ	Lepcha Consonant Sign T	1CB3	ཀླ	LEPCHA CONSONANT SIGN T
1A75	ཀྴ	Lepcha Consonant Sign N	1CB0	ཀྴ	LEPCHA CONSONANT SIGN N
1A76	ཀྵ	Lepcha Consonant Sign P	1CB1	ཀྵ	LEPCHA CONSONANT SIGN P
1A77	ཀྶ	Lepcha Consonant Sign M	1CAE	ཀྶ	LEPCHA CONSONANT SIGN M
1A78	ཀྷ	Lepcha Consonant Sign R	1CB2	ཀྷ	LEPCHA CONSONANT SIGN R

1A79	◌ <sup>◌</sup>	Lepcha Consonant Sign L	1CAF	◌ <sup>◌</sup>	LEPCHA CONSONANT SIGN L
1A7A	◌ <sup>◌</sup>	Lepcha Consonant Sign Y	1CA4	◌ <sup>◌</sup>	LEPCHA LETTER CONJOINED YA
1A7B	◌ <sup>◌</sup>	Lepcha Consonant Sign R ( <i>duplicate name</i> )	1CA5	◌ <sup>◌</sup>	LEPCHA LETTER CONJOINED RA