

# Sikkim Bhutia Lepcha Apex Committee (SIBLAC)

Upper Syari, P. O. Deorali - 737102, East Sikkim

E-mail: [siblac.sikkim@gmail.com](mailto:siblac.sikkim@gmail.com) Website: [www.siblac.org](http://www.siblac.org) Ph. +91 9434144224

Dr. Ken Whistler  
Technical Director  
Unicode Consortium

Gangtok, 7th December 2015

## Lepcha Unicode - Published Standard

Dear Dr. Whistler,

best thanks for your communication of November 26, 2015, in response to our Error Report submitted to the Unicode Consortium. Please, accept our appreciation for your dedicated and elaborate explanations which give us the impression of being taken seriously. We hope not to overstress your patience with the following considerations with respect to what we regard as errors:

### 1. General Remarks

We are aware of the fact that document L2/05-158 by Everson et al. was a proposal submitted to the Unicode Technical Committee (UTC). Since the minutes published at

<http://www.unicode.org/L2/L2005/05180.htm#104-C24>

do not mention any amendments, we expected the proposal to be adopted in its entirety - apart from the controversy regarding U+1C35/KANG for which approval was being kept pending. Of course, we can not rule out other modifications agreed upon by the UTC, however, they have, to our knowledge, not been published and are intransparent if in existence. Considering the flawed paragraphs on the Lepcha script in v. 5 of the Unicode Standard, we imagine that the inaccuracies discussed below might as well be due to editorial faults rather than deliberate decisions taken by the UTC.

With due respect, we disagree in regard to corrigenda you consider ill-advised. There is absolutely no risk to run into complications for implementations and data corruption as, to date, no fully functional Unicode compliant Lepcha font is available. *Mainwaring Róng*, a font under development for SIBLAC and scheduled for release in April 2016, thus attempts to be the first working Unicode font for the script. Since the *Lepcha Language Kit* by XenoType is no longer available, there remain only two nominally Unicode compliant fonts with imperfect rendering of the complex script. That's why native writers still stick to either handwriting or using the *JG Lepcha* custom encoded font. For shortcomings in *SIL Mingzat* and *Noto Sans Lepcha*, please, see Annex B.

In February 2015, Microsoft introduced the initial release of a Universal Shaping Engine (USE) which may or may not gain wider acceptance but, in any case, represents a new approach promising genuine rendering of Lepcha script. We consider it advisable to remove errors and misconceptions in an early stage of development of that new technology.

As a matter of principle, the UTC will have to decide whether digital typography is committed to reflect the traditional native Lepcha script or whether it's the script which will have to fit in the requirements of the technology and comply with a smooth implementation.

## 2. Supposed Combination U+1C26U+1C36

It's true that Mainwaring (1876), discussing the vowel modifier RAN, mentions that

*„This sign is supposed to be confined to the sole use of the two vowels & a and & i, but I have seen it written in books (though this is far from being any criterion of correctness) over the & á, and when necessary, it ought to be so applied;“*

However, it's clear that he does not trust in the authenticity of this ,third variant of a' encountered in documents written during an era characterized by an almost complete lack of orthographical standards and, thus, does not list it in his chart of vowels.

Subsequent linguistic and typographic experts such as

- Karl Faulmann, 1880
- G.B. Mainwaring and Albert Grünwedel, 1898
- G.A. Grierson, 1909
- Erik Haarh, 1959
- David Diringer, 1968
- Ashit Chakraborty, 1978
- Florian Coulmas, 1996
- Peter T. Daniels and William Bright, 1996
- Daniel Kai, 2003
- Heleen Plaisier, 2007
- Piotr Klafkowski, 2012

fully coincide with Mainwaring insofar as their listings of vowels do not include AA-RAN. Likewise, the ALA-LC Romanization Table (2012), to mention just the most reputable system of romanization for the Lepcha script, does not provide a transliteration for the supposed vocalic combination U+1C26U+1C36.

While one may object that some of these - mainly Western - authors just borrowed from the findings of their forerunners - which, especially in the case of the encyclopaedic writers, appears likely to be the case - a look at native sources may shed decisive light on the matter.

Among the Lepchas, script and orthography is traditionally taught by means of syllabaries called Lazong (𑌎𑌔) representing highly systematic charts of all the existing syllables in Lepcha.

A sample of a Lazong from 1945 is available online at

[http://eap.bl.uk/database/overview\\_item.a4d?catId=49309;r=28145](http://eap.bl.uk/database/overview_item.a4d?catId=49309;r=28145)

It does not include any syllables containing U+1C26U+1C36.

Lepcha scholar Azuk Tamsangmoo reports on a Lazong discovered recently at

<http://aachulay.blogspot.com/2010/09/secret-and-concealed-lazaong-book-of.html>

illustrating the structure of the syllabary with a reproduction of the first few blocks. None contains syllables with AA-RAN.

Finally, it may be mentioned that the elaborate *Lepcha-English Encyclopaedic Dictionary* by Lepcha scholar K.P. Tamsang (1980) does not list a single entry with U+1C26U+1C36 in a collection extending over more than one thousand pages.

The cited wording by Everson et al. (L2/05-158)

*„... the diacritical mark RAN (^) - only after the inherent vowel or -a or -i, never with any of the other vowels - ...“*

may be unfortunate and ambiguous, however, we'd expect the authors to have used the transcript ,aa', had they meant U+1C26. Hence, we read the passage „or -a" as a clarification of the meaning of „inherent vowel".

It escaped our notice, so far, that the same document contains, indeed, syllables in Lepcha script containing AA-RAN (in the example you mention as well as later on in the Lazong-like enumeration of syllables). We regard this to be an obvious error for the reasons given above.

### 3. Lepcha Syllabic Structure

The LEPCHA SIGN RAN (U+1C36) is a modifier having effect on the pronunciation (length and accent) of the vowels. The sign is, thus, clearly associated with the vowel signs. Handwritten Lepcha documents reveal that the LEPCHA VOWEL SIGN OO is nothing short of a LEPCHA VOWEL SIGN O + RAN merged into one character when the first movable types were cast (Calcutta, around 1850). Likewise, the LEPCHA VOWEL SIGN UU - conjunct in handwriting as well - may be seen as a LEPCHA VOWEL SIGN U topped by a rudimentary RAN. Besides, it's noteworthy that R.K. Sprigg, the paramount expert on Lepcha linguistics, does not hesitate to speak of „*the superscript rân component of the í symbol*" (1983, p. 314, para 2). This may suffice to stress the tight connection between vowel signs and U+1C36.

Therefore, input of RAN immediately after the vowel signs as advocated by Kai (2003) and proposed by Everson et al. (2005) makes sense. We are not fully confident to interpret your remarks in appropriate way, but we have the impression that the UTC decided to reverse the final two elements of the syllabic structure to ease implementation in font production. Modern technology, however, provides ways and means to implement reordering in a fail-safe manner. Moreover, Lepcha script requires reordering in many cases anyway. So, why should we shy away from reordering in connection with RAN?

Based on the definition provided in the Unicode Standard v. 1, p. 16 (we could not find a more recent definition) stating that

*„In a Unicode encoding, text is stored in sequential order in the backing store. Logical or backing store order corresponds to the order in which text is typed on the keyboard ..."*

we confess some difficulty in following your differentiation between typing order and backing store. We suggest that phonology, typing order and backing store sequence should, ideally, be one and the same.

Lastly, we are surprised at the interpretation of the term FINAL CONSONANT - a term coined and adopted in the process of approval by the UTC - which in consequence of the reversed order obtains the meaning of a ‚SOMETIMES FINAL' consonant.

Since - for the reasons mentioned earlier - there are hardly any Unicode encoded Lepcha documents ‚out in the open', we believe it's still time to get things put straight. On a longer term, we consider Unicode, i.e. seamless interchangeability, to be a crucial contribution towards the preservation and further development of the Lepcha heritage for which the experts at the Unicode Consortium deserve our everlasting gratitude.

We hope that our remarks are expressed in a generally understandable way and recommend the points raised for your kind consideration.

Very sincerely

Stefan Daehler, Research and IT Associate, SIBLAC

Mail: [admin@siblac.org](mailto:admin@siblac.org)

Phone: 0041 79 478 2209

## **ANNEX A**

### **Sources cited:**

Ashit Chakraborty, *Read Lepcha*, New Delhi, 1978

Florian Coulmas, *The Blackwell Encyclopedia of Writing Systems*, Malden, 1996

Peter T. Daniels, William Bright, *The World's Writing Systems*, Oxford, 1996

David Diringer, *The Alphabet; a key to the history of mankind*, New York, 1968

Karl Faulmann, *Buch der Schrift*, Wien, 1880

G.A. Grierson, *Linguistic Survey of India*, Calcutta, 1909

Erik Haarh, *The Lepcha Script*, in: *Acta Orientalia*, Copenhagen, 1959

Daniel J. Kai, *Introduction to the Lepcha Script*, Seattle, 2003

Piotr Klafkowski, *Towards a Rong/Lepcha and Tibetan comparative project. From the manuscripts of Athing Joseph Rongong*, in: *Zwierciadło etnologiczne*, Szczecin, 2012

G.B. Mainwaring, *Grammar of the Róng (Lepcha) Language, as it exists in the Dorjeling and Sikim Hills*, Calcutta, 1876

G.B. Mainwaring, Albert Grünwedel, *Dictionary of the Lepcha-Language*, Berlin, 1898

Heleen Plaisier, *A Grammar of Lepcha*, Leiden, 2007

R. K. Sprigg, *Hooker's expenses in Sikkim: an early Lepcha text*, in: *Bulletin of the School of Oriental and African Studies*, Vol. 46, No. 02, Cambridge, 1983

K.P. (Khárpú) Tamsang, *The Lepcha-English Encyclopaedic Dictionary*, Kalimpong, 1980

## **ANNEX B**

### **Comparison of fonts:**

<sup>1</sup> OpenType font developed by Darjeeling Unlimited for SIBLAC, pre-release version.

<sup>2</sup> TrueType font timestamped September 25, 2013

<sup>3</sup> TrueType font timestamped November 12, 2015

Rendering of the two presently available fonts is faulty - regardless of the input order - as soon as RAN occurs in combination with final consonants. Hence, the restoration of the logical syllabic structure would cause little or rather no loss of existing impeccable documents.

**For the sake of honesty:** Table created with Nisus Writer in OS X 10.8 with identical display in OpenOffice, Pages and TextEdit. According to the limited tests we were able to undertake, the third column (SIL Mingzat with input according to UTC) renders correctly (presumably due to Graphite support introduced) in OpenOffice 4/Windows 7 and higher and with Unicode compliant text editors in OS X 10.10 and higher.

Mainwaring Róng <sup>1</sup> Expected Rendering	Romanization Translation	SIL Mingzat <sup>2</sup> Typing Order UTC	SIL Mingzat <sup>2</sup> Typing Order Everson	Noto Sans Lepcha <sup>3</sup> Typing Order UTC	Noto Sans Lepcha <sup>3</sup> Typing Order Everson
ᳵ᳚	krit hunger	~ ᳵ᳚	ᳵ᳚	ᳵ᳚	ᳵ᳚
ᳵ᳚	ngák look, observe	~ ᳵ᳚	ᳵ᳚	ᳵ᳚	ᳵ᳚
ᳵ᳚	nyín milk	~ ᳵ᳚	ᳵ᳚	ᳵ᳚	ᳵ᳚
ᳵ᳚	tatsât period of time	ᳵ᳚	ᳵ᳚	ᳵ᳚	ᳵ᳚
ᳵ᳚	trínchen kindness	~ ᳵ᳚	ᳵ᳚	ᳵ᳚	ᳵ᳚
ᳵ᳚	drámdyân break down	~ ᳵ᳚	ᳵ᳚	ᳵ᳚	ᳵ᳚
ᳵ᳚	pátmú Tibetan	~ ᳵ᳚	ᳵ᳚	ᳵ᳚	ᳵ᳚
ᳵ᳚	fík tear, pull apart	~ ᳵ᳚	ᳵ᳚	ᳵ᳚	ᳵ᳚
ᳵ᳚	bíklóng ox	~ ᳵ᳚	ᳵ᳚	ᳵ᳚	ᳵ᳚
ᳵ᳚	mân meat	~ ᳵ᳚	ᳵ᳚	ᳵ᳚	ᳵ᳚