Date: 2014-08-07

Proposal to encode Kannada Sign Spacing Candrabindu

Vinodh Rajan vrs3@st-andrews.ac.uk

Introduction

The Badaga language is a minority language in the Indian state of Tamil Nadu. It is spoken by approximately 400 000 people around the Nilgiris district in Tamil Nadu and is closely related to the Kannada language. However, Badaga does not have any script of its own. At present, it is commonly written using the Tamil script. Recently, there have been attempts to create a separate Badaga script [1].

In the late 19th century, the Christian missionaries attempted to use the Kannada script to write Badaga (perhaps due the language's similarity to Kannada). Several New Testament gospel books were translated into Badaga and published using the Kannada script [2] [3].

Badaga Orthography in Kannada Script

Badaga when written with Kannada orthography seems to use only un-aspirated consonants. It also uses subjoined Kannada LLLA, which probably represents the retroflex vowels unique to the Badaga language.

But the most interesting is the use of a spacing *Candrabindu* character, which probably represents nasalization.

(Sample Text from the Gospel of Luke in Badaga Language)

ಮಾಡು, ಯೇಕಾಂದಲೆ ಯೆಂಗೂ ಯೆಂಗ ತಪ್ಪುಗಾರರುಗ ಬುಟ್ಟನೆಯೋ꙰; ಇನ್ನು ಎಂಗ್ಟುವ ಸೋದನೆಗ ಸೇತದೆ ಕೇಡೊಂದ ತಪ್ಪಿಸು, ಎಂದು ಹ್ಡೇಗಿವಿ ಎನ್ನ꙰.

māḍu, yēkāṁdale yeṁgū yeṁga tappugāraruga buṭṭaneyō**m̐**; innu eṁg<u>l</u>uva sōdanega sētade kēḍoṁda tappisu, eṁdu h<u>l</u>ēgivi enna**m̐**.

It is proposed that this new character be named as "KANNADA SIGN SPACING CANDRABINDU" and encoded in the Kannada range of the UCS.

Character Properties

0C80; KANNADA SIGN SPACING CANDRABINDU;Lo;0;L;;;;;N;;;;;

Annotation

This character is to be annotated as "*used in the Badaga language*"

Encoding Kannada Letter Spacing Candrabindu

The Kannada block in the UCS already has a combining *Candrabindu* encoded at U+0C81. It might be argued that the proposed character be unified with the already encoded *Candrabindu* character. However, such unification is unwanted. Apart from the difference in the character property, U+0C81 is attested for the Vedic orthography in the Kannada script, whereas the proposed character is attested for the Badaga orthography. Simultaneous plain text representation of both orthographies would become impossible with such unification.

Also, unlike the Vedic *Candrabindu,* the proposed character combines with Anusvara. It is for this reason the character is proposed as "*Lo*" rather than "*Mc*" (similar to U+A8F2 which is also a spacing *Candrabindu*). Hypothetically, we might also discover a similar cognate Vedic usage of the proposed character in the future.
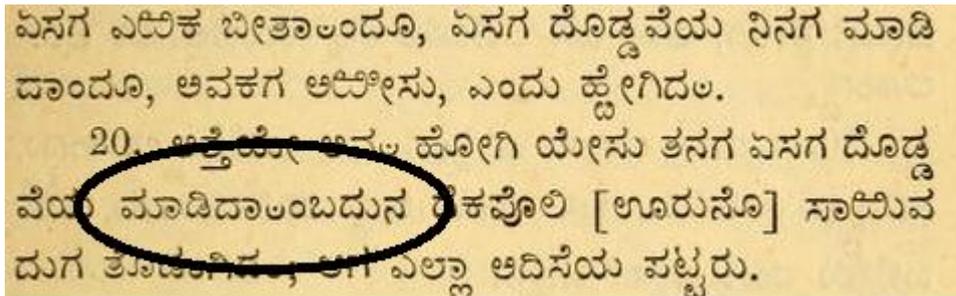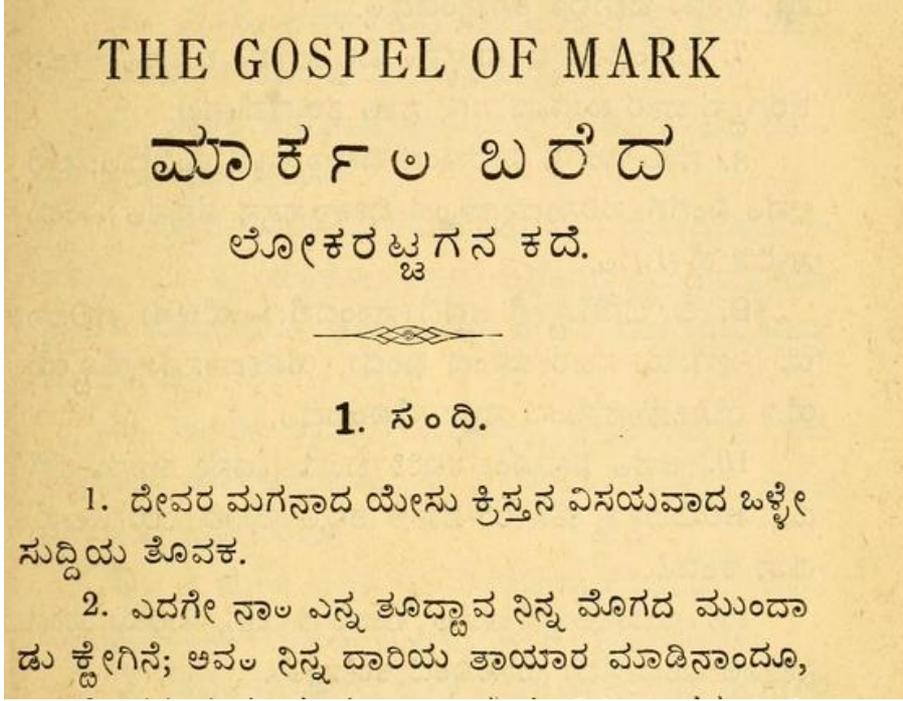

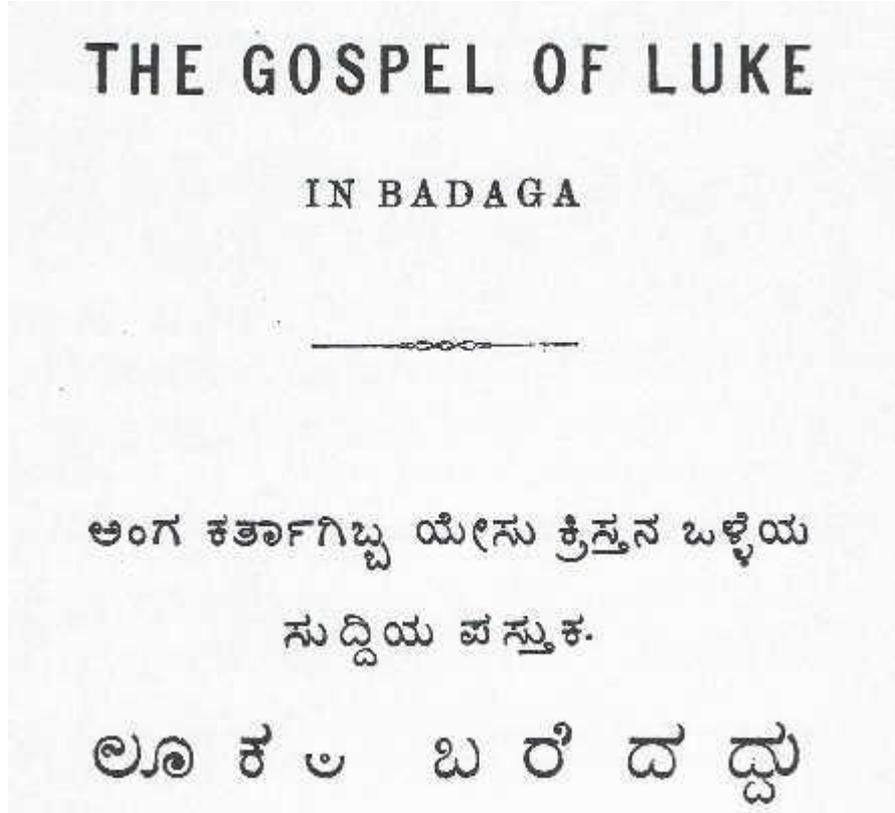The encircled word reads: *māḍidā m̐ mbaduna*

It is to be noted that several scripts in Unicode have both spaced and combining forms of the same character. In fact, in the main Indic block, Devanagari has a combining *Candrabindu* U+0901 and also spacing *Candrabindu* U+A8F2. The recently encoded Grantha script also has a combining *Anusvara* U+11300 (proposed) and spacing *Anusvara* U+11301.

On the same lines, it is proposed that a new character, KANNADA SIGN SPACING CANDRABINDU be encoded to support the minority Badaga orthography.

Additional Attestations

THE GOSPEL OF MARK

ಮಾರ್ಕಂ ಬರೆದ

ಲೋಕರಟ್ಟಗನ ಕದೆ.

◇◇◇

1. ಸಂದಿ.

1. ದೇವರ ಮಗನಾದ ಯೇಸು ಕ್ರಿಸ್ತನ ವಿಸಯವಾದ ಒಳ್ಳೇ ಸುದ್ದಿಯ ತೊವಕ.

2. ಎದಗೇ ನಾಲ ಎನ್ನ ತೂಡ್ಡಾವ ನಿನ್ನ ಮೊಗದ ಮುಂದಾ ಡು ಕ್ಟೇಗಿನೆ; ಅವಂ ನಿನ್ನ ದಾರಿಯ ತಾಯಾರ ಮಾಡಿನಾಂದೂ,

First chapter from the Gospel of Mark

THE GOSPEL OF LUKE

IN BADAGA

ಆಂಗ ಕರ್ತಾಗಿಬ್ಚ ಯೇಸು ಕ್ರಿಸ್ತನ ಒಳ್ಳೆಯ

ಸುದ್ದಿಯ ಪಸ್ತುಕ.

ಲೂಕಂ ಬರೆದದ್ಧು

First page from the Gospel of Luke

References

1. Badaga Script: http://www.omniglot.com/writing/badaga.htm
2. The Gospel of Mark in Badaga:
   https://archive.org/details/gospelofmarkinba00sike
3. The Gospel of Luke in Badaga:
   http://www.christusrex.org/www1/pater/wolf3-pdf/badaga-luke-nt.pdf

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM

A. Administrative

1. Title:
Proposal to encode KANNADA SIGN SPACING CANDRABINDU
2. Requester's name:
Vinodh Rajan
3. Requester type (Member body/Liaison/Individual contribution):
Individual Contribution
4. Submission date:
18th July 2014
5. Requester's reference (if applicable):
6. (Choose one of the following:)
This is a complete Proposal

B. Technical - General

1. (Choose one of the following:)
a. This proposal is for a new script (set of characters):
No
b. The proposal is for addition of character(s) to an existing block:
Yes. Kannada block
2. Number of characters in proposal:
One
3. Proposed category
Category B1 (Specialized Small)
4. Is a repertoire including character names provided?:
Yes
 a. If YES, are the names in accordance with the 'character naming guidelines' in Annex K of ISO/IEC 10646-1?
 Yes
 b. Are the character shapes attached in a reviewable form?
 Yes
6. Who will provide the appropriate computerized font (ordered preference: True Type, PostScript or 96x96 bit-mapped format) for publishing the standard?
 Vinodh Rajan
 If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:
Vinodh Rajan, Fontforge
7. References:
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?
Yes
b. Are published examples (such as samples from newspapers, magazines, or other sources) of use 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.

C. Technical - Justification

1. 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?
Wg Cdr Bellie Jayaprakash, native Badaga speaker and webmaster of http://badaga.co/.
If YES, available relevant documents?
The communication was through e-mail.
2. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?
Digitization of Badaga texts in Kannada script
3. The context of use for the proposed characters (type of use; common or rare)
Rare
4. Are the proposed characters in current use by the user community?
No
If YES, where?
5. After giving due considerations to the principles in N 1352 must the proposed characters be entirely in the BMP?
Yes
If YES, is a rationale provided?
It is an addition to the Kannada block encoded in the BMP
If YES, reference:
6. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?
N/A. A single character is proposed
7. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?
No
8. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?
Yes.
If YES, reference:
It might be unifiable with U+0C81 KANNADA SIGN CANDRABINDU. **The unification issue is discussed in the proposal**
9. Does the proposal include use of combining characters and/or use of composite sequences (see clause 4.11 and 4.13 in ISO/IEC 10646-1)?
No
10. Does the proposal contain characters with any special properties such as control function or similar semantics?
No