Changes to Core Specification for Indic scripts for Unicode 11.0

Srinidhi A
srinidhi.pinkpetals24@gmail.com

Sridatta A
sridatta.jamadagni@gmail.com

December 8, 2017

This document provides changes to Core Specification for Indic scripts. It was partially discussed in L2/17-098. Based on the recommendations of Script Ad Hoc, a separate document is presented here.
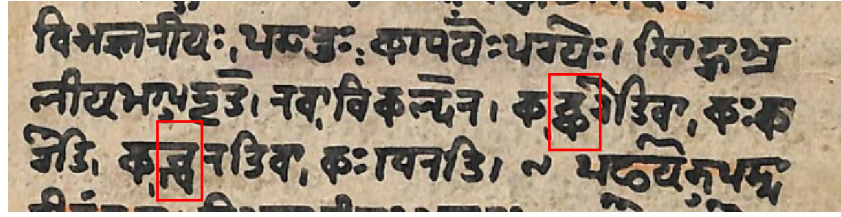
1 Sharada

In Sharada section of Core Specification, the following representation for Jihvamuliya is presented.

$$U+111C2 \; \boxed{\text{ᰤ}} \; jihvamuliya + U+11191 \; \text{क} \; ka \rightarrow \text{क̆}$$

$$U+111C2 \; \boxed{\text{ᰤ}} \; jihvamuliya + U+11192 \; \text{ख} \; kha \rightarrow \text{ख̆}$$

However Jihvamuliya is not written above the consonant, instead it makes ligature with the consonant

as . A sample of their representations is shown in below manuscript.

Jihvamuliya is various manuscripts. (from IndoSkript).

L2/17-255 asked *"Does the change of shape for KA take place because of the jihvamuliya or because KA is subjoined below another consonant?"*

In consonant clusters these letters behave similar to other consonants. The ligature occurs because KA is subjoined below another consonant.

Action Requested Correct the shape of combination of Jihvamuliya and Ka/Kha in core specification.

2 Kannada

In Kannada section of Core Specification, the following representation for Jihvamuliya is presented.

U+0CF1 𑿀 *jihvamuliya* + U+0C95 ಕ *ka* → 𑿀ಕ

U+0CF2 ೲ *upadhmaniya* + U+0CAB ಫ *pha* → ೲಫ

Dependent vowels signs can also be added to the stack:

U+0CF1 𑿀 *jihvamuliya* + U+0C95 ಕ *ka* + U+0CBF ಿ *vowel sign i* → 𑿀ಕಿ

Action Requested: ẖka should not have Talekattu (or headstroke). Correct the shape of ẖka from 𑿀ಕ

to 𑿀ಕ (see L2/13-242 for attestations).

3 Sora Sompeng

Sora Sompeng is currently categorized as Abugida by Unicode. The following details are provided in page 614 of Core Specification.

*"**Encoding Structure**. The Sora Sompeng script is structured as an abugida. The consonant letters contain an inherent vowel. There are no conjunct characters for consonant clusters, and there is no visible vowel killer to show the deletion of the inherent vowel. The reader must determine the presence or absence of the inherent schwa based on recognition of each word. The character repertoire does not match the phonemic repertoire of Sora very well."*

The proposal L2/09-189R says *"Sora Sompeng is of the Brahmic type: the consonant letters contain an inherent vowel."*, but doesn't provide attestations where the consonant letters contain inherent vowel as well as the actual usage of the script. The only attestation provided is an alphabet chart where the consonants are transcribed as s, t, b etc. without inherent vowel.

However after examining primary sources like books published by native speakers and name plates, it was concluded that Sora Sompeng is structurally an alphabet and cannot be categorized as abugida. It does not contain inherent vowel, virama, consonant conjuncts and vowel signs. Letters are represented using combination of consonants and vowels. For example Ka is represented as 𑃜𑃠 (KAH+AH) and Hargl (meaning ornament in Sora) is written as 𑃦𑃠𑃝𑃜𑃗 (HAH+AH+RAH+GAH+LAH).

(https://upload.wikimedia.org/wikipedia/commons/c/cd/Mangei_Gomango.JPG)

Nameplate of Mangei Gomango who devised Sora Sompeng script at Odisha Tribal Development Society Bhubaneswar. Here Guru Mangei Gomango is written as ꓱꓴꓦꓴ �10ꓸꓣꓞ0ꓮ1 ꓞ2ꓸ0ꓣ2.
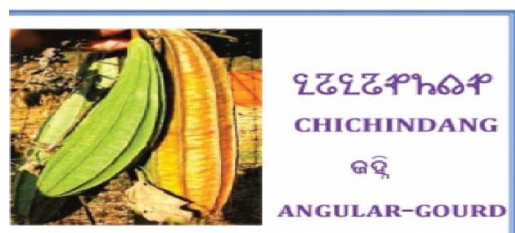

A book called Savara Laglngn Abirnaba depicted above shows Combination of Consonants with vowels. Words in Sora Sompeng from the same book.[1]

	SIING
	SANGSANG
	SEJAB
	SIING
	SURGA
	SONETI

	TANGNIY
	TANGNIY
	TREDA
	TIRIYUNG
	TUDUM
	TOGI

BUBUTENG
ଉମର
BLACK BEE

CHICHINDANG
ଜହି
ANGULAR-GOURD

Words in Sora sompeng from a learning book written by Krushna Sabar.[2]

---

[1] http://savara.bharatavani.in/book/savara-laglngn-abirnaba/

Action Requested: Change the kind of writing system from abugida to alphabetic in Chapters 6.1, 15.14 and elsewhere. Update the section Encoding Structure of chapter 15.14 appropriately.

4 VEDIC SIGN NIHSHVASA

Section 4.5 of L2/09-372 and L2/09-298 informs U+1CD3 separates sections between which a pause is disallowed. The annotation in the code chart has been corrected as suggested in L2/09-372, but explanation in section 12.1 pg. 465 of the core specification "U+1CD3 VEDIC SIGN NIHSHVASA indicates where a breath may be taken." is misleading. This should be corrected to "Separates sections of Sama Vedic singing between which a pause is disallowed" as in the code chart.

5 Ligature Forms for Ra + Vocalic Liquids

The Core specification gives following information of Ligature Forms for Ra + Vocalic Liquids. However these sequences are not restricted Devanagari, in fact it is attested in most Indic scripts which have repha. It is customary to write them as the independent_vocalic_vowel with a repha than ra+Vowel signs as noted in L2/12-106.

For the benefit of font developers these formations may be mentioned in individual chapters in the Core Specification or in this Devanagari section as 'also attested in other Indic scripts'.

Ligature Forms for Ra + Vocalic Liquids. The phonological sequence /r vocalic_r/, expressed with the character sequence <U+0930 ra, U+0943 vocalic_r>, can graphically appear as either of two forms, as shown in the first row of *Table 12-4*. It may appear as the full independent vowel form of the vocalic_r, with a superscript *repha* form of the ra (V + RA$_{sup}$): ऋ̊. Alternatively, it may appear as the full letter form of the ra with the subscript, dependent form of the vocalic_r (RA$_n$ + V$_{vs}$): र̗. Similarly, the phonological sequences with the other vocalic sounds (*rr, l, ll*) have two written forms, as shown in *Table 12-4*.

Table 12-4. RA + Vocalic Letter Ligature Forms

[2] http://savara.bharatavani.in/book/savara-language-learning-book/

ಇಂದ್ರೋ ವಹ್ನಿಃ ಪಿತೃಪತಿರ್ನೈ ಋರ್ತೋ ವರುಣೋ ಮರುತ್ ।
ಕುಬೇರ ಈಶಃ ಪತಯಃ ಪೂರ್ವಾದೀನಾಂ ದಿಶಾಂ ಕ್ರಮಾತ್ ॥   ೯೨

ಐರಾವತಃ ಪುಂಡರೀಕೋ ವಾಮನಃ ಕುಮುದೋಂಜನಃ ।
ಪುಷ್ಪದಂತಸ್ಸಾರ್ವಭೌಮಸ್ಸುಪ್ರತೀಕಷ್ಟ ದಿಗ್ಗಜಾಃ ॥   ೯೩

ಕರಿಣ್ಯೋಭ್ರಮುಕಪಿಲಾಪಿಂಗಲಾನುಪಮಾಃ ಕ್ರಮಾತ್ ।
ತಾಮ್ರಕರ್ಣೀ ಶುಭ್ರದಂತೀ ಚಾಂಗನಾ ಚಾಂಜನಾವತೀ ॥   ೯೪

ಕ್ಲೀಬಾವ್ಯಯಂ ತ್ವಪದಿಶಃ ದಿತೋರ್ಮಧ್ಯೇ ವಿದಿಕ್ ಸ್ತ್ರಿಯಾಮ್ ।
ಅಭ್ಯಂತರಂ ತ್ವಂತರಾಲಂ ಚಕ್ರವಾಲಂ ತು ಮಂಡಲಮ್ ॥   ೯೫

ಅಬ್ರಂ ಮೇಘೋ ವಾರಿವಾಹಃ ಸ್ತನಯಿತ್ನುರ್ಬಲಾಹಕಃ ।
ಧಾರಾಧರೋ ಜಲದರಸ ಟಿತ್ವಾನ್ಯಾರಿದೋಂಬುಭೃತ್ ॥   ೯೬

೨.   ಆಗ್ನೀಃ ಪೂರ್ವೇೞಿಭಿರ್ಋಷಿಭಿ—ರೀಡ್ಯೋ ನೂತನೈರಿಪಿತ।
     ಸ ದೇವಾँ ಏಹ ವಕ್ಷತಿ   ॥ ೨ ॥

     ಪುರಾಣ ಋಷಿಗಳು ಯಾವಾತನ ಸ್ತುತಿಯನ್ನು ಮಾಡಿದ್ದರೋ, ಈಗಿನ
ಋಷಿಗಳು ಯಾವಾತನ್ನು ಸ್ತುತಿಸುವರೋ, ಅಂಥ ಅಗ್ನಿಯು ದೇವತೆಗಳನ್ನು ಈ
ಯಜ್ಞದಲ್ಲಿ ಆಹ್ವಾನಿಸಲಿ.   ॥ ೨ ॥

ಖ್ರ್ rr̥ in Kannada.
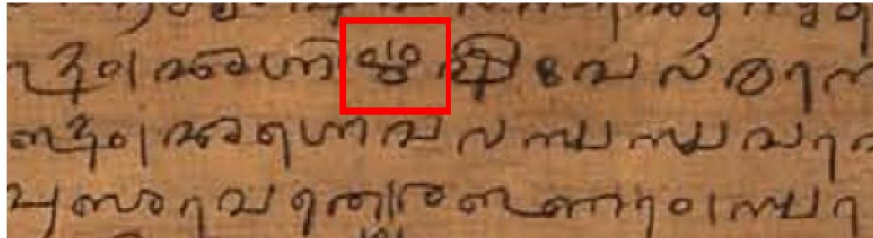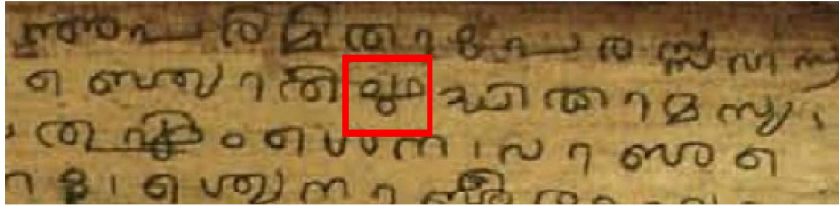
rr̥ in Tigalari.
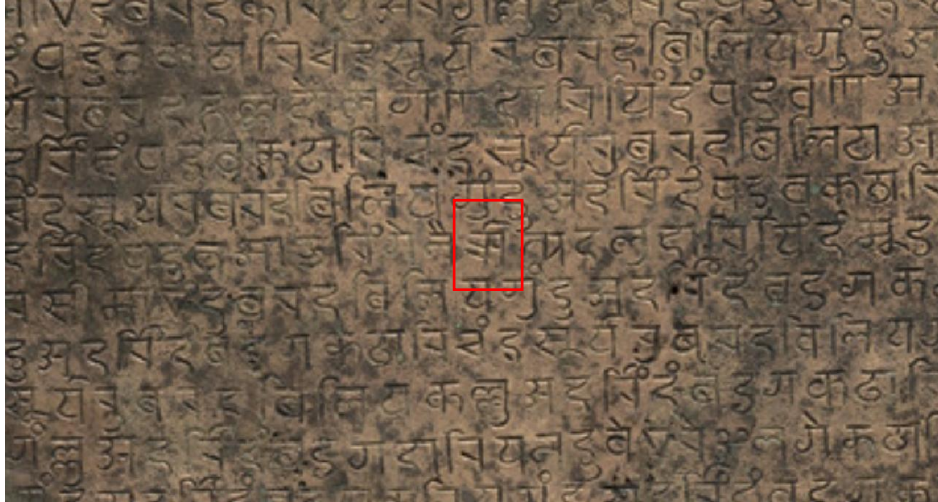
rṛ in Nandinagari.

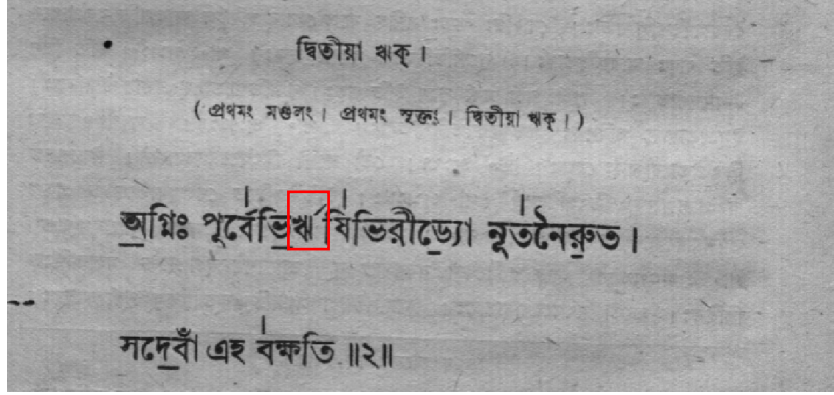When 𝖀 vowel sign vocalic r occurs with ra, the latter is written as *repha* and placed above the arrow-head of the vowel letter:

𝖀 *rṛ*   <𝖘 ra, 𝖀 vowel sign vocalic r>

rṛ in Bhaiksuki (from L2/14-091).

ଋ॔ rṛ in Odia.

6

ঋ rr̥ in Bengali.

rr̥ in Telugu. Note that modern orthography does not use repha. It is attested in older texts.

ઋૅ rr̥ in Gujarati

Similar forms may also exist in other scripts which have repha like Tirhuta, Newa, Siddham and so on.
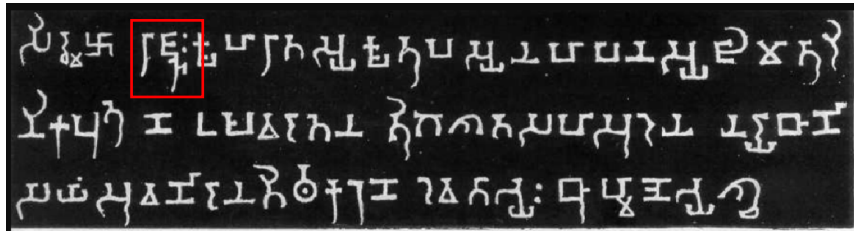
6 Brahmi

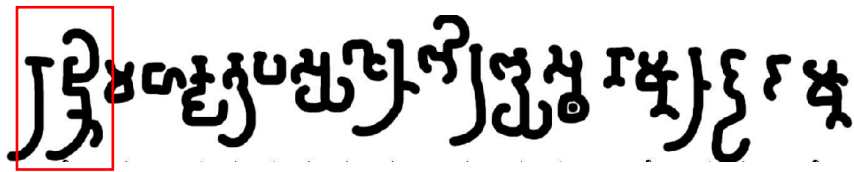In Brahmi chapter jña is written as ᚷ as seen below.

Figure 14-1. Consonant Ligatures in Brahmi

ᚷ + ◌̄ + ᚧ → ᚧ   sva
11032   11046   1102F

+ + ◌̄ + ᚧ → ᚧ   kṣa
11013   11046   11031

E + ◌̄ + ᚧ → ᚷ   jña
1101A   11046   1101C

However it appears to be written in stack of ja E and ña ᚧ as ᚷ. It is requested to update appropriately.

Nasik inscription, 1ˢᵗ century CE.

Junagadh rock inscription of Rudradaman, circa 130–150 CE.

Use of jña in other Brahmi texts.

In page 557

"U+11002 brahmi sign visarga is used to write syllable-final voiceless /h/; *that is, [x] and [f]. The velar and labial allophones of /h/, followed by voiceless velar and labial stops respectively, are sometimes written with separate signs U+11003 brahmi sign jihvamuliya and U+11004 brahmi sign upadhmaniya.*"

Should be corrected as "U+11002 brahmi sign visarga is used to write syllable-final voiceless /h/. The velar and labial allophones of /h; that is, [x] and [ɸ], followed by voiceless velar and labial stops respectively, are sometimes written with separate signs U+11003 brahmi sign jihvamuliya and U+11004 brahmi sign upadhmaniya."