# Request for editorial updates to Indic scripts

Srinidhi A and Sridatta A

Tumakuru, India

srinidhi.pinkpetals24@gmail.com, sridatta.jamadagni@gmail.com

April 17, 2017

## 1 Devanagari

**1.1 A8F8 DEVANAGARI SIGN PUSHPIKA** represents the flower and this symbol is commonly used at the end of texts and also as filler. The Code chart annotates it as vaidika pushpika. This character is not limited to Vedic texts and Sanskrit. This sign also occurs in classical Sanskrit and other languages and scripts.
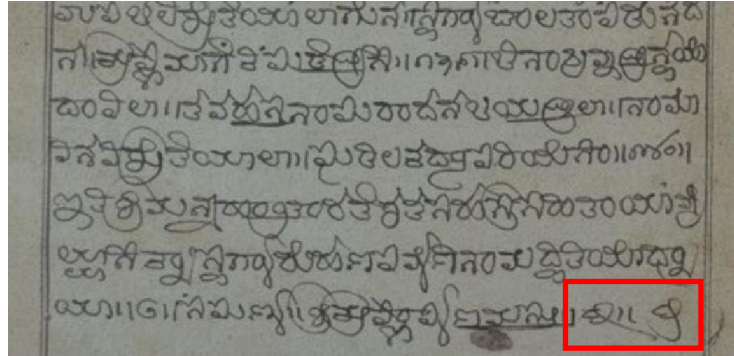
### Action Requested

- Remove the term vaidika pushpikaa from Code chart.
- In Script Extensions extend the use of DEVANAGARI SIGN PUSHPIKA to Kannada and Newa. Currently Tigalari and Nandinagari are not encoded. Add these to Script Extensions after their encoding.
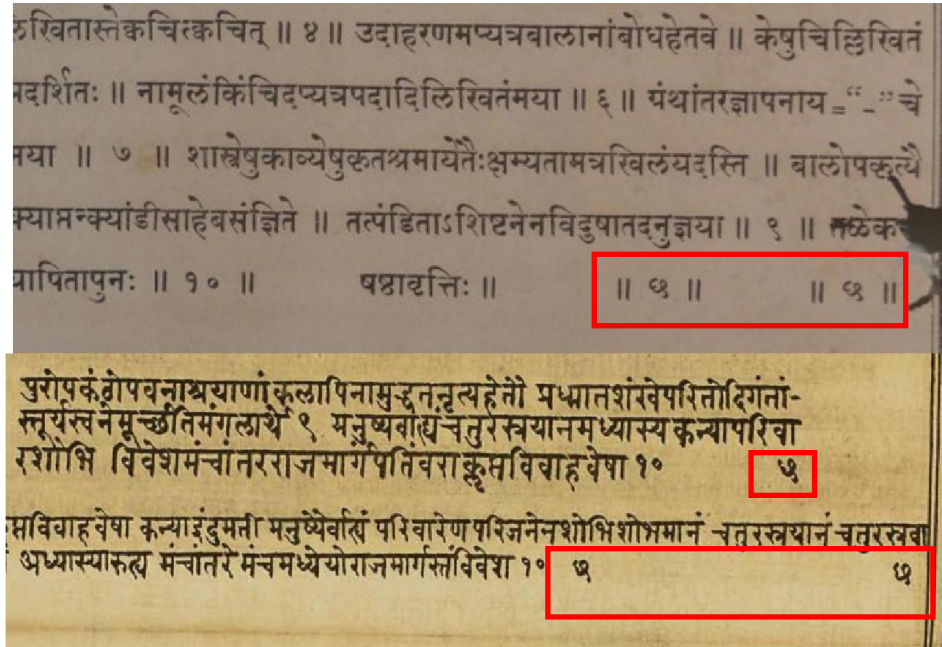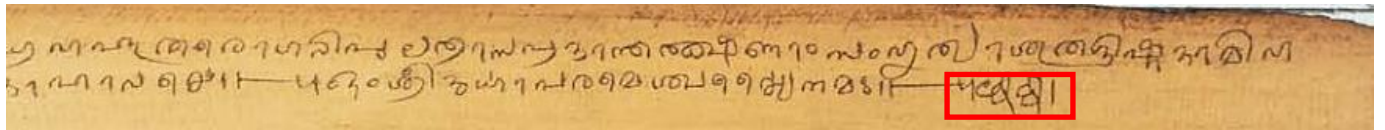


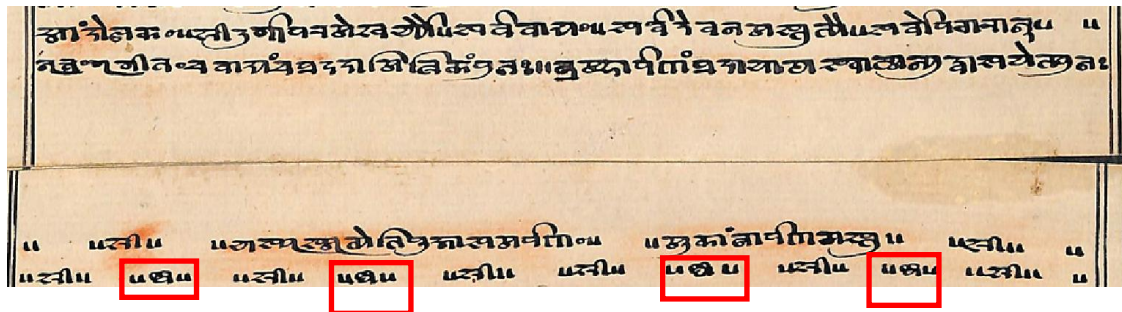Use of Pushpika in Classical Sanskrit texts of Devanagari.
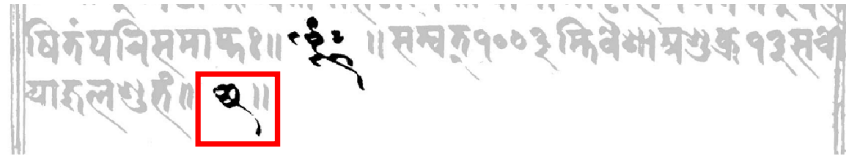
Pushpika in Newa script (from Figure 41 of L2/12-003).



Pushpika in Kannada script.



Pushpika in Tigalari.



Pushpika in Nandinagari.

## 1.2 A8FA DEVANAGARI CARET

This editorial mark occurs in majority of Indian scripts and languages. This isn't a Vedic specific sign.

**Action Requested:** Remove the term vaidika from Code chart.

**1.3 Representation of *rya* in North Indian languages.**

In Devanagari orthography of North Indian languages like Braj Bhasha, Rajasthani, Malvi etc have special rendering of Rya. Post-base form of ya is preferred over repha. This is exemplified in Braj Bhasha reader of Rupert Snell.

(c) Conjuncts of *r* with *y*, common in perf. ptc. from verb roots in -*r*, are written not with a superscript *repha* but with a modified य, appearing in the printed text as e.g. करचौ.

Examples of such forms are given below
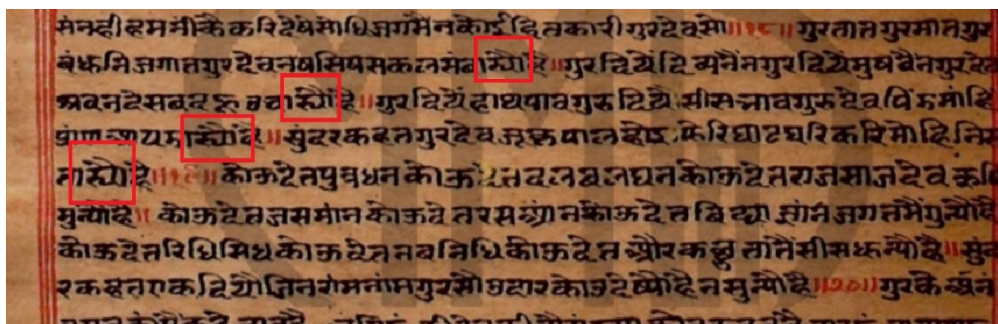


सब तजि भजिऐं नंदकुमार ।
और भजे तैं काम सरै नहिं, मिटै न भव जंजार ॥
जिहिं जिहिं जोनि जन्म धरचौ बहु जोरचौ अघ कौ भार ।
तिहिं काटन कौं समरथ हरि कौ तीछन नाम कुठार ॥
बेद परान भागवत गीता. सब कौ यन्न मत सार ।

BRAJ BHĀKHĀ.                    ,                    (DISTRICT, BULANDSHAHR.)

एक आदमी-के दो लड़के है । छोटे-ने कह्यौ बापू हमारा हिस्सा हमें दे-दे । उस-ने अपना हिस्सा वा-को बाँट-देस्यो । छोटो थोरे-ही दिन-में अपनो माल जमा परदेस-को ले-के चलो गयो । वहाँ सब लुँगाड़पने-में बरबाद कर्यो । जब सब बरबाद कर चुक्यो वा देस-में जबरा अकाल पर्यो । वा भूखो कंगाल हो-गयो । वा एक कोई-की नौकर हो-गयो । वा-ने सुअरन चुगाने-पे नौकर कर-दियो । जब वा-को कोई कुछ दे-नाई तो वो जो सूअर चरे-हे खोकटा वा-सि पेट भरे-हे ॥

In order to represent both forms in plain text U + 200D zero width joiner is to be placed immediately after the ra to obtain the ya-phalaa. The repha form is rendered when no ZWJ is present as shown in the following example. Similar methods are adopted in Bengali and Kannada which have repha.

र ra + ்‌ virama + य ya → र्य

र ra + ்‌ virama + ZWJ + य → र्‍य

र ra + ZWJ + ்‌ virama + य → **र्य**

**Action Requested:** Mention the representation of rya with and without repha and their approach to render correctly in Devanagari chapter of Core Specification.

### 1.4 U+0904 ऄ DEVANAGARI LETTER SHORT A

As noted in L2/09-321 there isn't any information regarding the phonetic value and usage of the character. The character is used to denote short e similar to ऒ ओ औ where the signs are graphically composed of अ and the respective vowel sign. The character is used in many sources such as Hindi translations along with Devanagari transliteration of Literature of Kannada, Telugu, Tamil, Malayalam and Kashmiri languages published by Bhuvan Vani Trust, Lucknow.



ऄ in Kamba Ramayana of Tamil.



ऄ in Telugu śrī pōtana mahābhāgavatamu.

It is also used in recent books containing Awadhi language to indicate short e.



ऄहि विधि बीते बरष षट सहस बारि आहार।
संबत सस सहस्र पुनि रहे समीर अधार॥
बरष सहस दश त्यागेउ सोऊ। ठाढ़े रहे एक पद दोऊ॥

(मा. १.१४४, १४५.१)

4

ऍ in  Śrīrāmacaritamānasa-kathā mānasamēṃ tīsa kathāĕ̃ (containing Awadhi) by Svāmī Rāmabhadrācārya (2017).

**Action Requested:** Annotate in Code chart denoting the usage for short e.

### 1.5 Tibetan in Devanagari

In Himalayan regions like Nepal, Tibetan language is infrequently written Devanagari. The following characters are employed. The manuscript is from EAP676/3/6.

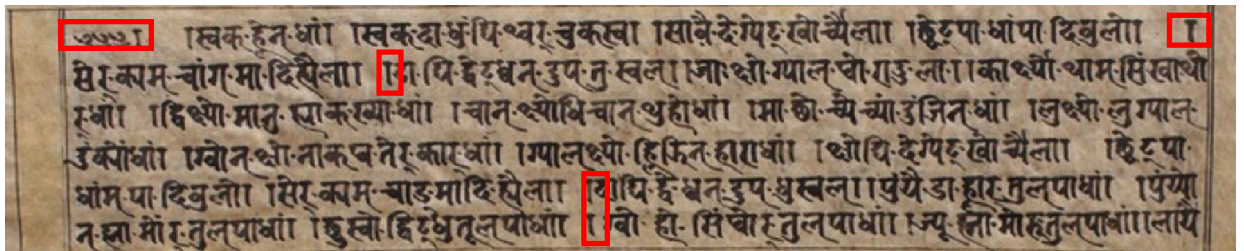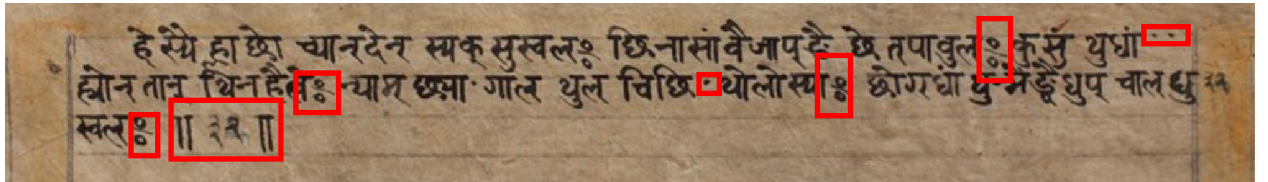0F04 ༄ TIBETAN MARK INITIAL YIG MGO MDUN MA

0F05 ༅ TIBETAN MARK CLOSING YIG MGO SGAB MA

0F0B ་ TIBETAN MARK INTERSYLLABIC TSHEG

0F0D ། TIBETAN MARK SHAD

0F0E ༎ TIBETAN MARK NYIS SHAD

0F14 ༔ TIBETAN MARK GTER TSHEG





**Action Requested:** Extend the use of above mentioned Tibetan characters to Devanagari in script extensions.
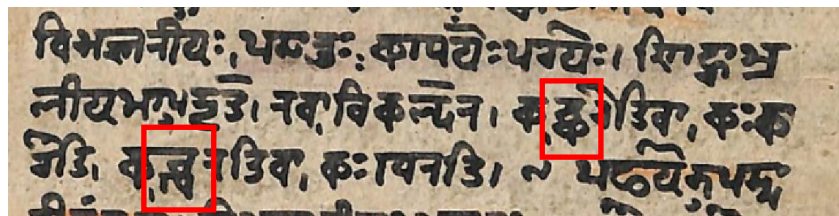
## 2 Sharada and Kannada

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

U+111C2 ⬚ *jihvamuliya* + U+11191 क *ka* → कॅ

U+111C2 ⬚ *jihvamuliya* + U+11192 प *kha* → पॅ

However Jihvamuliya is not prefixed above the consonant, instead it makes ligature with the consonant as 𑅂. A sample of their representations is shown in below manuscript.



**Action Requested**

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

- Change the Indic_Syllabic_Category of 111C2 SHARADA SIGN JIHVAMULIYA from Consonant_Prefixed to Consonant_With_Stacker.

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

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

U+0CF2 ∞ *upadhmaniya* + U+0CAB ಫ *pha* → ಫ಼

Dependent vowels signs can also be added to the stack:

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

**Action Requested:** ẖka should not have Talekattu (or headstroke). Correct the shape of ẖka from ☒ಕ to ☒ಕ (see L2/13-242 for attestations).
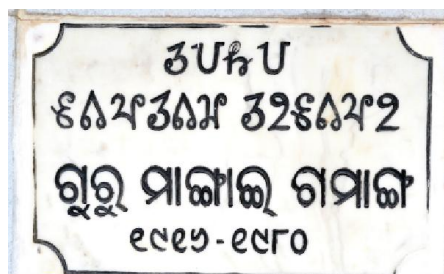
**3 Sora Sompeng**

Sora Sompeng is currently categorized as Abugida by Unicode. The following details are provided in page 604 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 proposal L2/09-189R claims *"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 ꪖꪰ (KAH + AH) and Hargl (meaning ornament in Sora) is written as ꪝꪰꪮꪖꪜ (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 ꪖ�꪿U ꪝꪰꪮꪖꪝ ꪖꪒꪝꪰꪮꪲ.

"𝑓" (SA)

𝑓 + ∅ = 𝑓∅ , 𝑓 + 𝛿 = 𝑓𝛿, 𝑓 + =𝑓𝑧

S + A = SA , S + E = SE , S + I = SI

𝑓 + U = 𝑓U , 𝑓 + 2 = 𝑓2

S + U = SU , S + O = SO
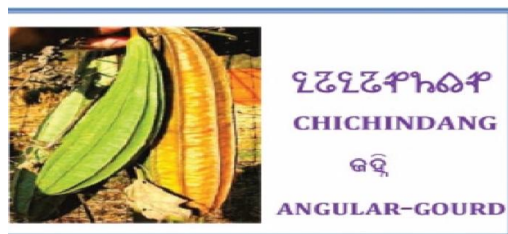
𝟛 ( TA )

𝟛 + ∅ = 𝟛∅ , 𝟛 + 𝛿 = 𝟛𝛿, 𝟛 + 𝑧 = 𝟛𝑧

T + A = TA , T + E = TE, T + I = TI

𝟛 + U = 𝟛U , 𝟛 + 2 = 𝟛2 .

T + U = TU , T + O = TO .

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



| 𝑓𝑧𝑧𝑓 | SIING | 𝟛∅𝑓𝟑𝑧𝑦 | TANGNIY |
|---|---|---|---|
| 𝑓∅𝑓𝑓∅𝑓 | SANGSANG | 𝟛∅𝑓𝟑𝑧𝑦 | TANGNIY |
| 𝑓𝛿𝑧∅�ʒ | SEJAB | 𝟛ℎ𝛿ℎ∅ | TREDA |
| 𝑓𝑧𝑧𝑓 | SIING | 𝟛𝑧ℎ𝑧𝑦U𝑓 | TIRIYUNG |
| 𝑓U ℎ𝛿𝛿 | SURGA | 𝟛UℂℎU𝟪 | TUDUM |
| 𝑓2𝟑𝛿𝟛 | SONETI | 𝟛2𝛿𝟪 | TOGI |



𝟑U𝟑U 𝟛𝛿𝑓
BUBUTENG
ଭ୍ରମର
BLACK BEE



𝜀𝑧𝜀𝑧𝑓ℎ∅𝑓
CHICHINDANG
ଜହ୍ନି
ANGULAR-GOURD

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

---

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

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

8

**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.
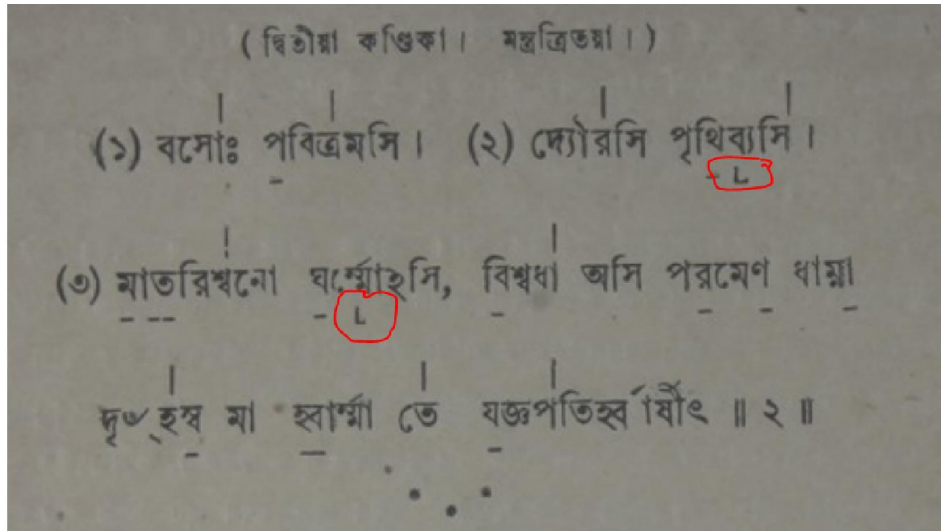
## 4 Vedic Extensions

### 4.1 Veda in Bengali

The following Vedic characters- 1CD5, 1CD6, 1CD8, 1CE1, 1CD0, 1CD2, 1CEA and 1CED should be extended for Bengali script.
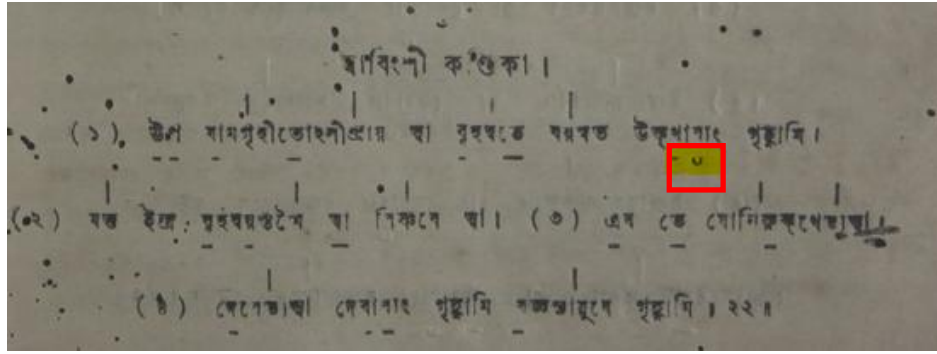
### 4.1.1 Shukla-yajurvedic characters
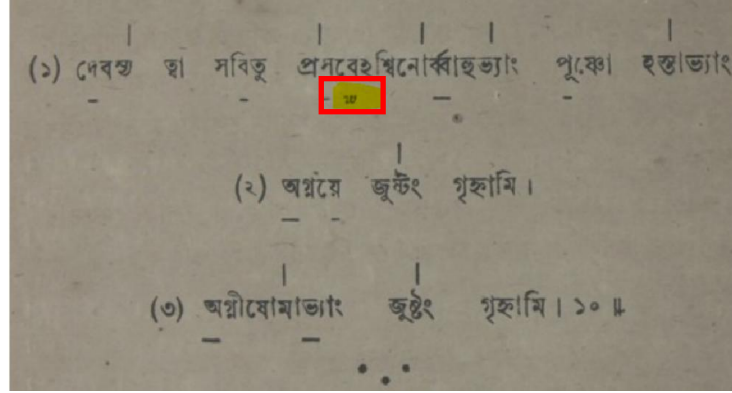


Attestation for 1CD5 VEDIC TONE YAJURVEDIC AGGRAVATED INDEPENDENT SVARITA



Attestation for 1CD6 VEDIC TONE YAJURVEDIC INDEPENDENT SVARITA

Attestation for 1CD8 VEDIC TONE CANDRA BELOW

Here latin w,L are used due to lack of proper glyphs.

Source:
http://eap.bl.uk/database/overview_item.a4d?catId=187010;r=31647
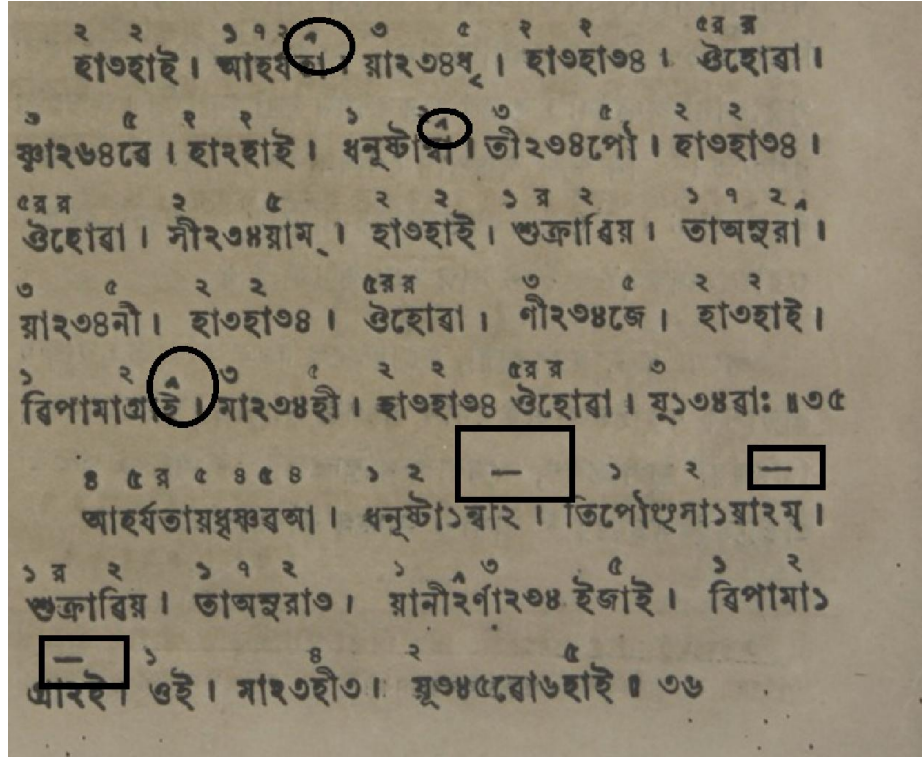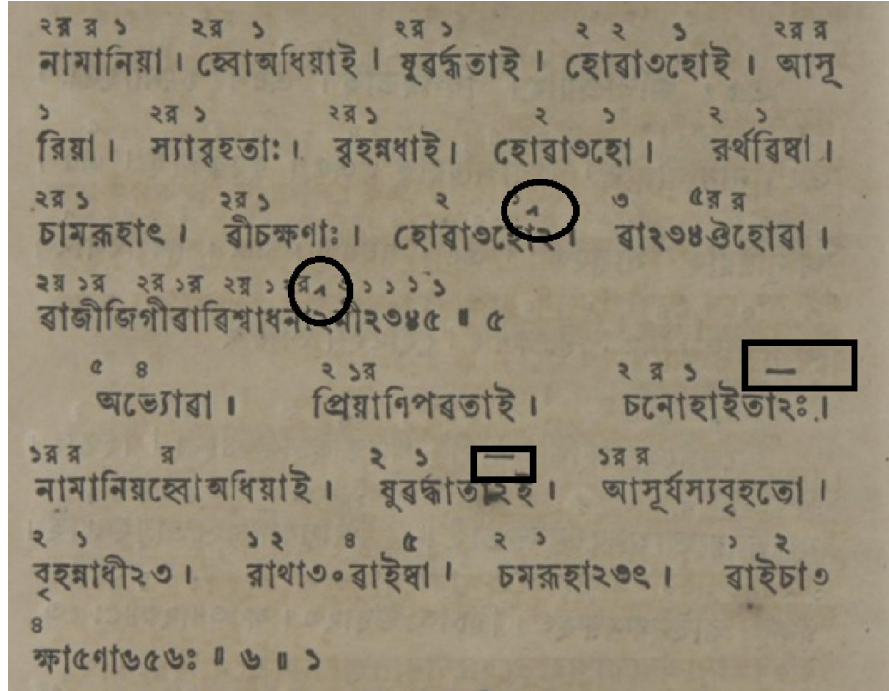http://eap.bl.uk/database/overview_item.a4d?catId=187012;r=13093

### 4.1.2 Atharvavedic character
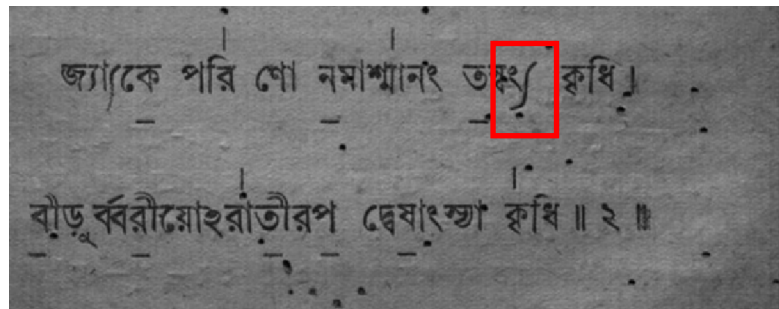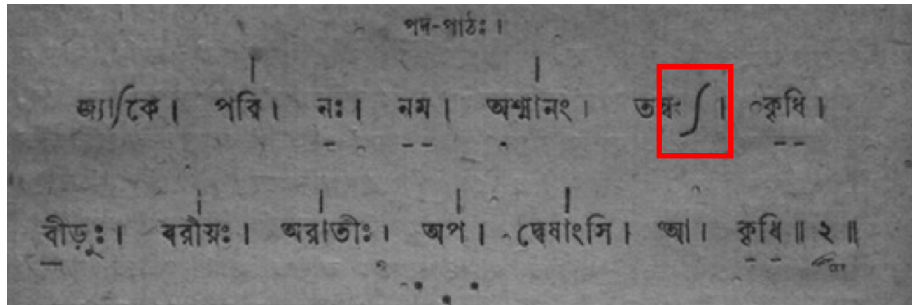




Attestation for 1CE1 VEDIC TONE ATHARVAVEDIC INDEPENDENT SVARITA

Source: http://eap.bl.uk/database/overview_item.a4d?catId=186923;r=17465
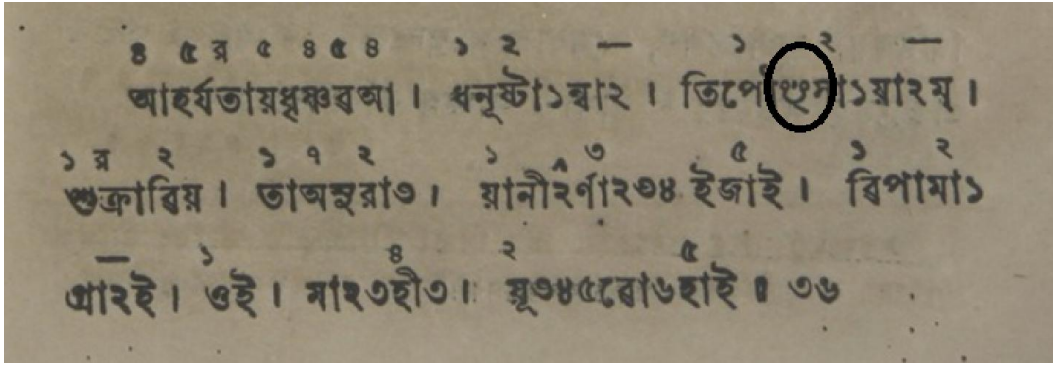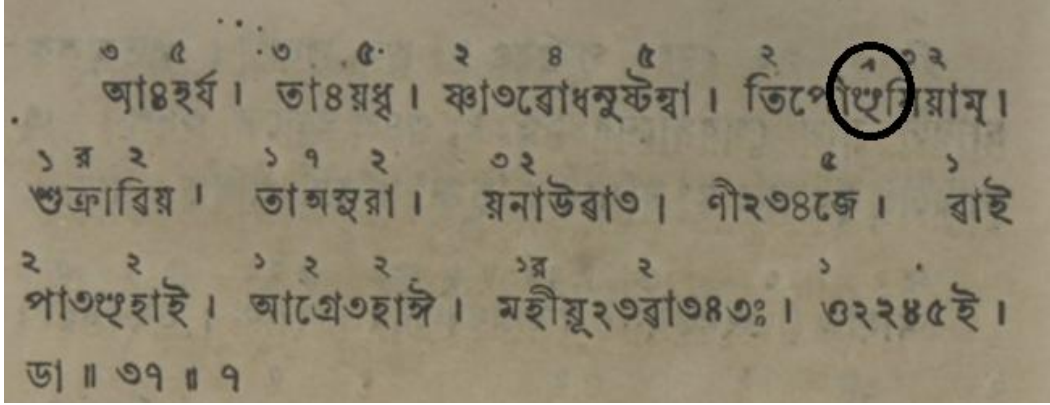
### 4.1.3 Samavedic characters in Bengali



Attestations for 1CD0 and 1CD2

No attestation is provided for 1CD1 VEDIC TONE SHARA which indicates *dīpta śruti*, since *dīpta śruti* is indicated by VEDIC TONE ATHARVAVEDIC INDEPENDENT SVARITA (see L2/15-164)

11

Source: http://eap.bl.uk/database/overview_item.a4d?catId=187038;r=21411

### 4.1.4 Vedic Anusvara





Attestations for 1CEA and 1CED

Source: http://eap.bl.uk/database/overview_item.a4d?catId=187038;r=21411

### 4.2 Veda in Tirhuta

The use of 0951 DEVANAGARI STRESS SIGN UDATTA and 0952 DEVANAGARI STRESS SIGN ANUDATTA should be extended to Tirhuta.
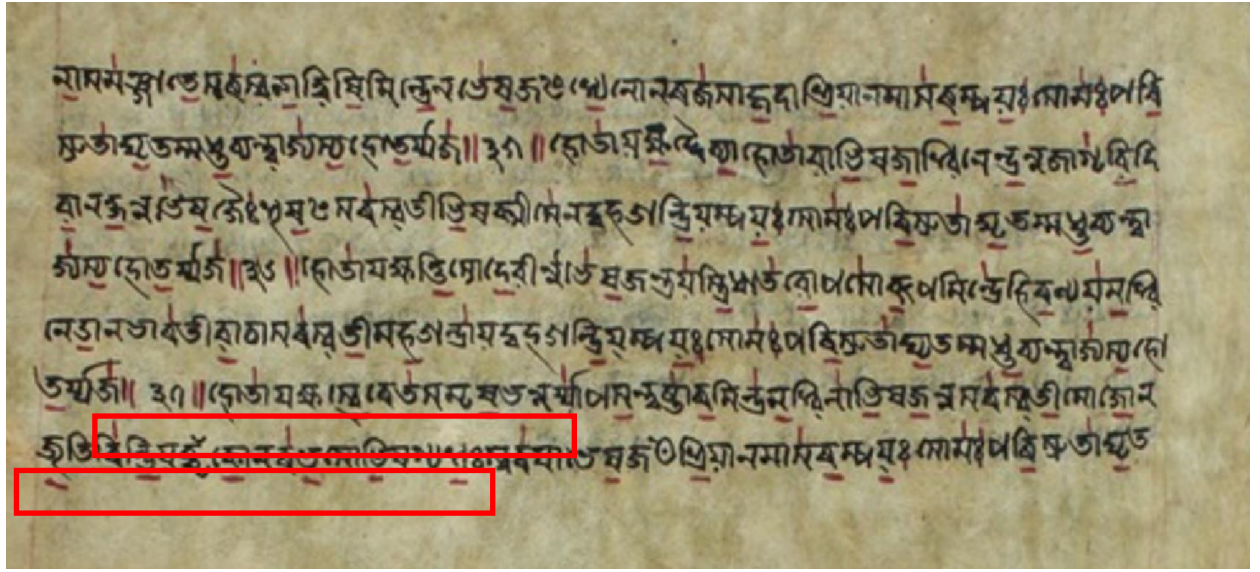
Vajasaneya Samhita in Tirhuta (from EAP790/21/101).

**4.3 Samavedic characters in Kannada.**

The pages 4 and 5 of L2/15-164 provided the use of 1CD0, 1CD2, 1CD3 and 1CF4 in Kannada script.

Extend the use of these characters to Kannada.

**4.4 Veda in Odia**

1CDA VEDIC TONE DOUBLE SVARITA is employed in Odia Vedic texts. Extend the use of 1CDA to Odia.



Purushasukta in Odia script.[3]

---

[3] http://www.odia.org/sanskrit/vishnu_purushasukta_odi.pdf

## 4.5  U+1CF2 VEDIC SIGN ARDHAVISARGA

Extend the use of Ardhavisarga to Telugu



Use of Ardhavisarga in Siddhanta Kaumudi.

Section 4.11 of L2/11-175R mentions the use of Ardhavisarga in Tirhuta. Extend the use of Ardhavisarga to Tirhuta.

**Action requested:** Update the Script Extensions to extend aforementioned characters used in Bengali, Tirhuta, Kannada, Odia and Telugu.

L2/09-342, L2/15-164 etc. demonstrate that the Ardhavisarga is not limited to Vedic and Vyakarana texts. The Ardhavisarga is commonly found in manuscripts, books, transcription of inscriptions etc.

**Action requested:** Remove vaidika jihvaamuuliiya upadhmaaniiya in the annotation.

## 4.6 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 Brahmi

This is about minor glyph changes in Brahmi digits 30 and 40. Like other letters headstroke in 30 appears in later stages of Brahmi. Earliest sources do not have the headstroke. The earliest sources of 40 have symmetrical form. Similar shape is employed in Segoe UI Historic font as ᚷ.



1105D  ᨔ  BRAHMI NUMBER THIRTY
1105E  ᨝  BRAHMI NUMBER FORTY

Current glyphs



1105D-  ᨔ      1105E- ᨝

Proposed glyphs



Earliest examples of 30 and 40 (from Indoskript project).



30 and 40 in Brahmi.
Current glyph of 30 with headstroke is found from 2-3 century CE
(from Indian Paleography of Bühler).

**Action Requested:** Change the glyphs of 1105D and 1105E to match their earliest sources.