

## Changes to ScriptExtensions.txt for Indic characters for Unicode 11.0

Srinidhi A  
srinidhi.pinkpetals24@gmail.com

Sridatta A  
sridatta.jamadagni@gmail.com

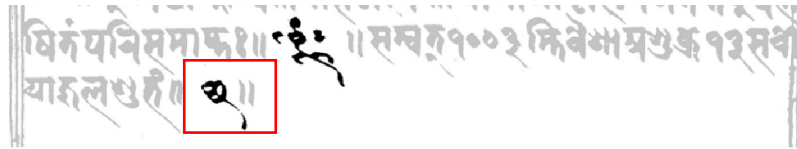
December 8, 2017

This document provides changes to ScriptExtensions.txt for Indic characters. It was discussed in L2/17-098. Based on the recommendations of Script Ad Hoc, a separate document is presented here.

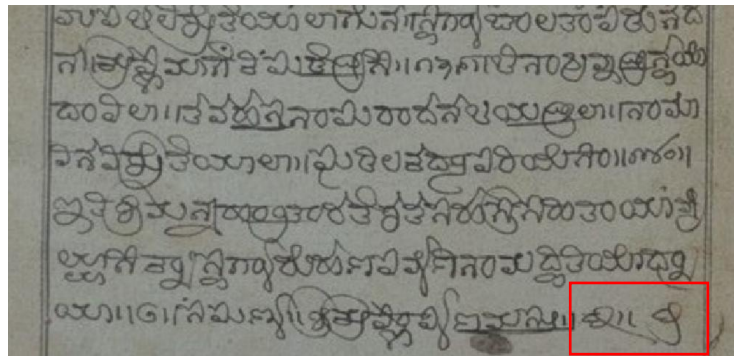
### 1 Devanagari

#### A8F8 ॐ DEVANAGARI SIGN PUSHPIKA

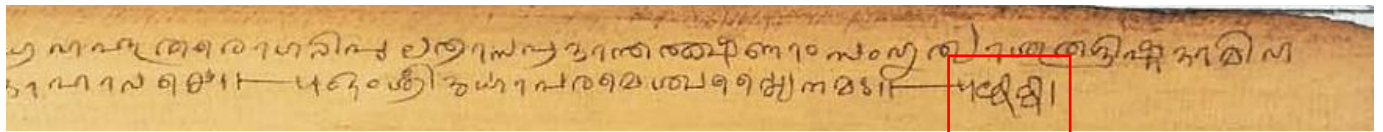
It represents a flower and this symbol is commonly used at the end of texts.



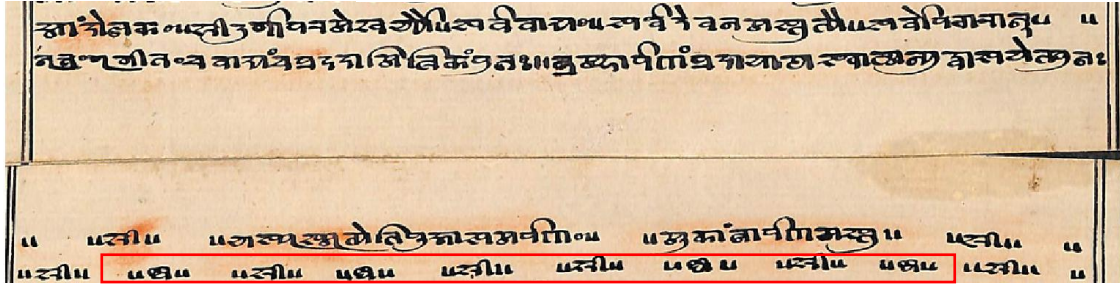
Pushpika in Newa script (from Figure 41 of L2/12-003).



Pushpika in Kannada script.



Pushpika in Tigalari.



Pushpika in Nandinagari.

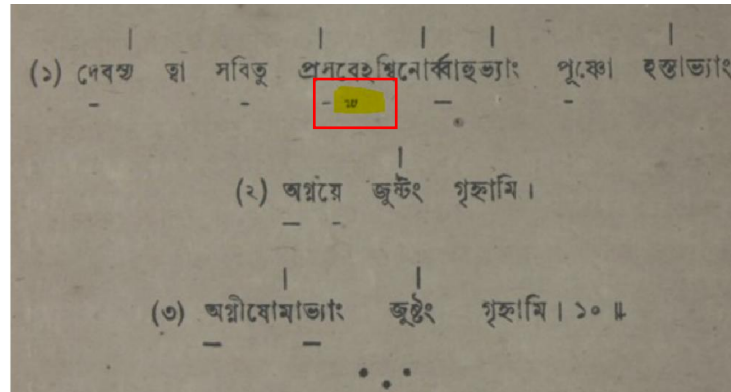
**Action Requested** In Script Extensions extend the use of DEVANAGARI SIGN PUSHPIKA to Kannada and Newa. Currently Tigalari and Nandinagari are not encoded. Add these to Script Extensions after their encoding.

## 2 Vedic Characters

### 2.1 Veda in Bengali

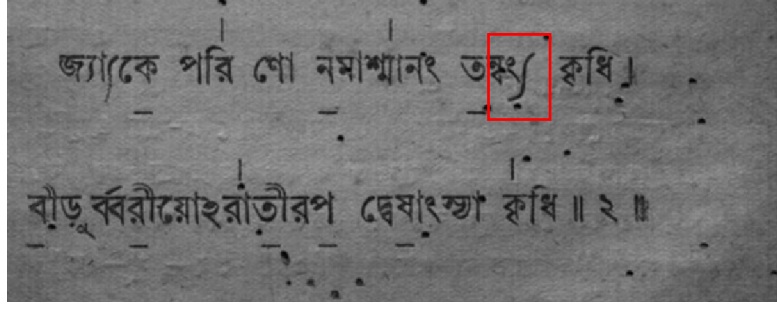
Various Vedic signs are employed in Bengali script. The details are given below. The attestations are from Endangered Archives Programme.

#### 2.1.1 Shukla-yajurvedic characters



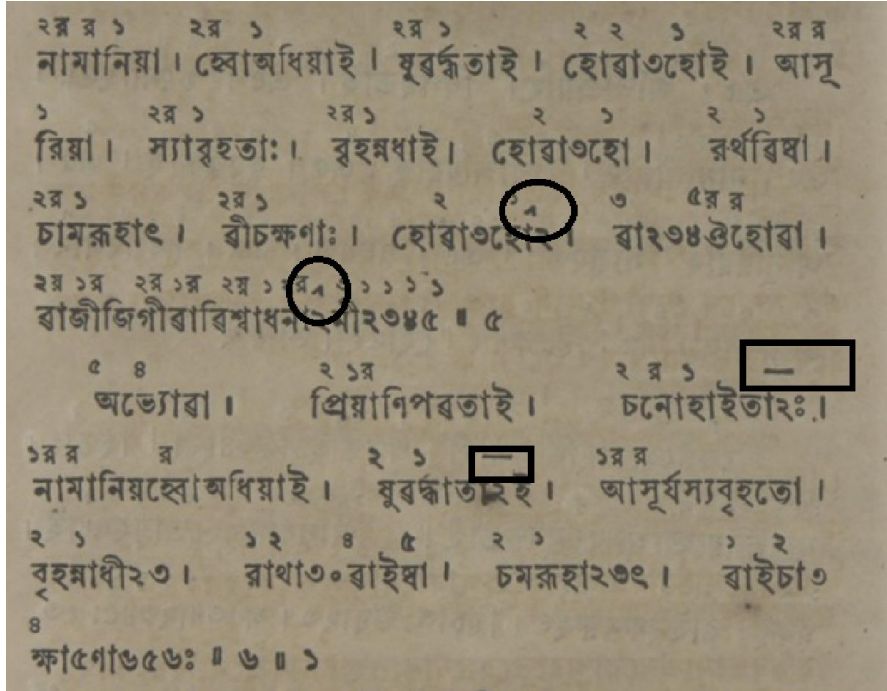
Attestation for 1CD5 VEDIC TONE YAJURVEDIC AGGRAVATED INDEPENDENT SVARITA.



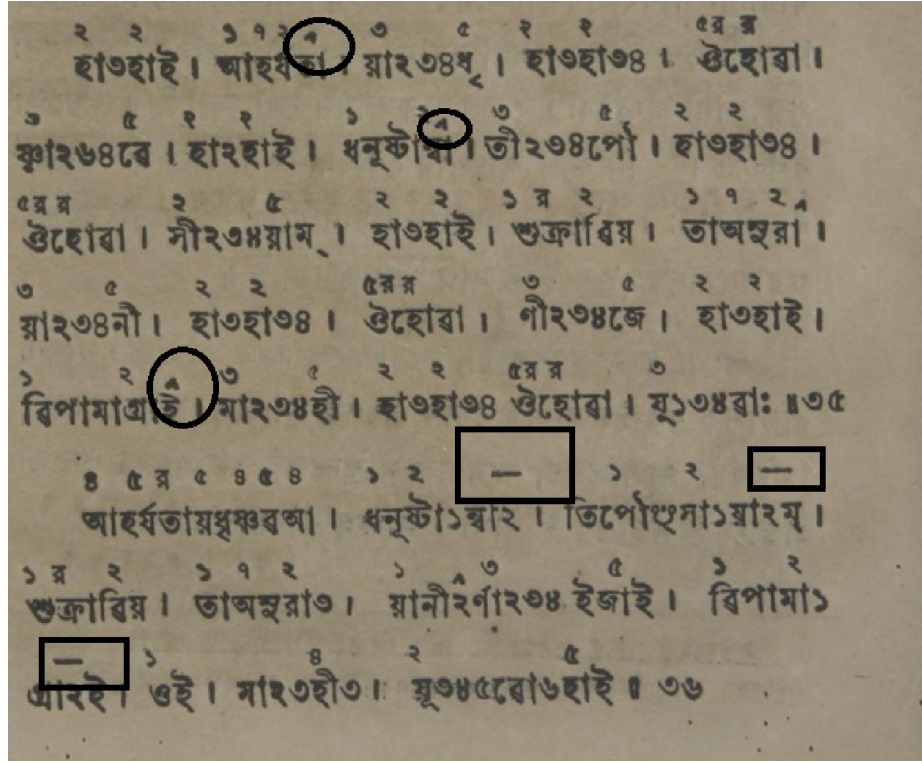


Attestation for 1CE1 VEDIC TONE ATHARVAVEDIC INDEPENDENT SVARITA.

### 2.1.3 Samavedic characters in Bengali



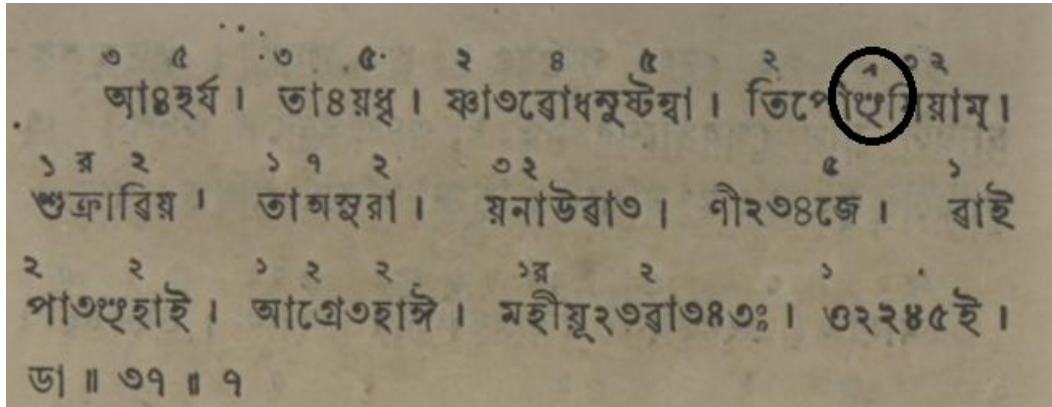


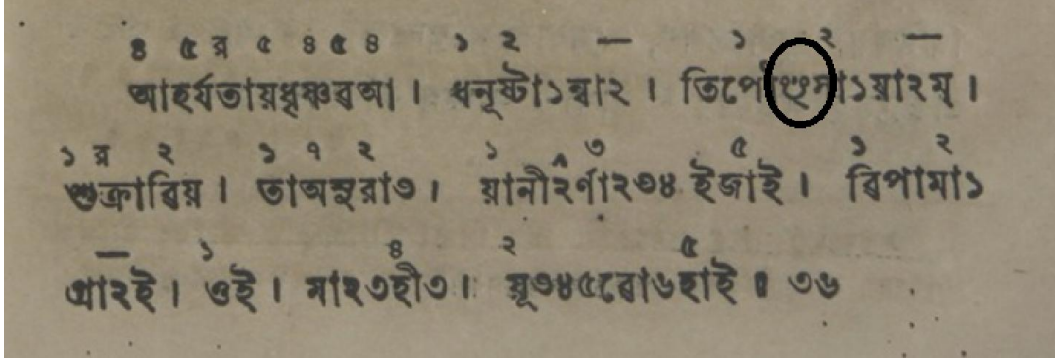


#### Attestations for 1CD0 and 1CD2.

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](#))

#### 2.1.4 Vedic Anusvara





Attestations for 1CEA and 1CED.

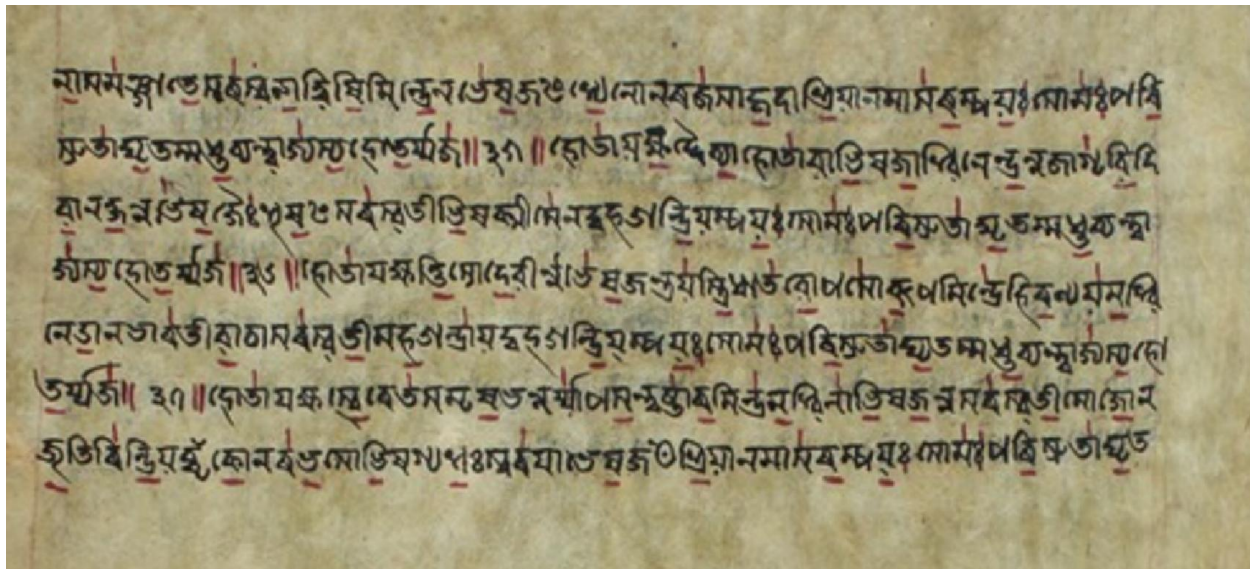
L2/17-255 asked “For U+1CEA VEDIC SIGN ANUSVARA BAHIRGOMUKHA, the examples provided in 4.1.4 on page 12 show a dot between the two loops. Is it a combining mark? Part of the character or a random dot?”

It is not a random dot, it is anusvara. Since Vedic texts were also published in Devanagari from Kolkata, It is probable that publishers have reused same glyph as in Devanagari. Though anusvara in Bengali is distinct from Devanagari.

**Nasalization Marks.** A set of spacing marks and one combining mark, U+1CED VEDIC SIGN TIRYAK, are encoded in the range U+1CE9..U+1CF1. They describe phonetic distinctions in the articulation of nasals. The *gomukha* characters from U+1CE9..U+1CEC may be combined with U+0902 DEVANAGARI SIGN ANUSVARA or U+0901 DEVANAGARI SIGN CANDRABINDU. U+1CF1 VEDIC SIGN ANUSVARA UBHAYATO MUKHA may indicate a *visarga*

**Action requested:** 1CD5, 1CD6, 1CD8, 1CE1, 1CD0, 1CD2, 1CEA and 1CED should be extended for Bengali script.

## 2.2 Veda in Tirhuta



**Action requested:** The use of 0951 DEVANAGARI STRESS SIGN UDATTA and 0952 DEVANAGARI STRESS SIGN ANUDATTA should be extended to Tirhuta in Script extensions.

### 2.3 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.

L2/17-255 asked “The examples for U+1CD0, U+ 1CD2, and U+1CD4 appear to have solid evidence, but we request more evidence be provided for U+1CD3, which is a spacing character and seems to be represented in L2/15-164 by the following:”

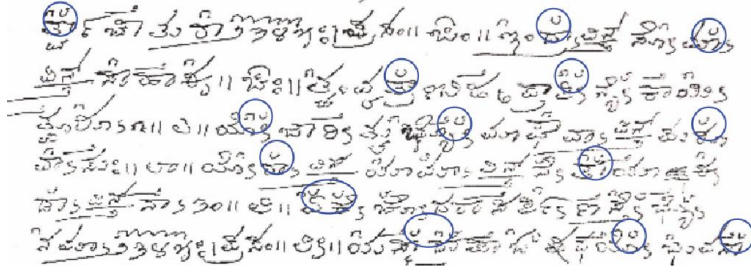
॥ २ ॥  
 ଶ୍ରୀମଦ୍ଭଗବତ୍ ॥

The spacing character used is Double Danda. The combining character written above digit 5 is U+1CD3.

೧೬೯.೩. ಕಾಢಶಿಯಾ | ನಶ್ವಾಢಿಇತ್ಯಾಢಿಆಭುವಾತ್ || ಉ  
ಸದಾವೃಧಸ್ಸ | ಖಾ | ಔಢಿಹೋಹಾಇ | ಕಯಾಢ.  
ಪಾಇ || ಪೃಯಾಹೋಢಿ | ಹಿಮಾಆಢಿ || ವಾ  
|| ಡಿ  
ತೋಢಿಇಹಾಇ ||

Use of  $U+1CD0 \hat{\circ}$ ,  $U+1CD2 \bar{\circ}$  and  $U+1CD3 \circ''$ .

ಅಭಿತ್ಯಾಪ್ಯಪ್ರಭಾಸುತಾಃ || ಸೂತೌ ಪೃಥಾ | ಮಿಪಾಣತಾ  
 ಂಗಿಯಾಂಕಿ || ತೃಮ್ವಾವಾಂಗಿಯಾಂಕಿ || ಪ್ನಹಾಂಗಿ  
 ಉವಾಯೇಂಗಿ || ಮಾಂಗಿಳಿದಾಮ್ ||



Attestations for U+1CF4 ◌ VEDIC TONE CANDRA ABOVE.

**Action requested:** Use of 1CD0, 1CD2, 1CD3 and 1CF4 should be extended to Kannada script in Script extensions.

## 2.4 Veda in Odia

1CDA VEDIC TONE DOUBLE SVARITA is employed in Odia Vedic texts. Here signs similar to ‘ ‘ ‘ ‘ is used due typographic limitation.

ଓଁ ତତ୍ତ୍ଵଂ ଯୋରାବୁଂ ଶମହେ । ଗାତୁଂ ଯୁଜାୟ । ଗାତୁଂ  
 ଯୁଜାୟତୟେ । ଦୈବଂ ସ୍ଵସ୍ଵସ୍ଵରଂ ସୁ ନଃ । ସ୍ଵସ୍ଵମାନ୍ନୁଷେଭ୍ୟଃ ।  
 ଭୃଷ୍ଂ ଜିଂଗାତୁ ଭେଷଜମ୍ । ଶନ୍ନୋଂ ଅସୁ ଦ୍ଵିପଦେଂ । ଶମ୍  
 ଚତୁଂଷ୍ଠଦେ । ଓଁ ଶାନ୍ତିଶାନ୍ତିଶାନ୍ତିଶାନ୍ତିଃ ॥

Purushasukta in Odia script.<sup>1</sup>

**Action requested:** Extend the use of 1CDA to Odia in Script extensions.

<sup>1</sup> [http://www.odia.org/sanskrit/vishnu\\_purushasukta\\_odi.pdf](http://www.odia.org/sanskrit/vishnu_purushasukta_odi.pdf)



## 2.5 U+1CF2 VEDIC SIGN ARDHAVISARGA

ఇచ్చాతిగ్గక్ష్మప్రవచనీయవంజ్ఞస్సాస్మ్యత్ - అతిదేవాన్ కృష్ణః ఖిలపింగదాన్తసంభావనా  
 న్వవస్సగహ్వాసముచ్చయేషు॥ ౧॥౮॥౯౬॥ ఏషుద్యోత్యేష్వపిరు క్తనంజ్ఞస్సాస్మ్యత్ - స  
 ర్విషోఽపిస్యాత్ అనుచస్సత్వాన్నవః-సంభావనాయాంలిజ్ - తస్యాపవవిషయభూ  
 తెభవనైకర్మదౌర్లభ్యప్రయు క్తన్దర్లభ్యంద్యో తయన్న పిశబ్దస్సాస్మ్యదిత్య నేనసంబ  
 ధ్యతే - సర్విచ ఇతివక్ష్మీతుఅపిశబ్దబలెనగమ్యమానస్యచిన్దోరవయవావయవిభావసంబ  
 ధై-ఇయమేవహ్యపిశబ్దస్యచహర్లద్యో తకతానామ - ద్వితీయాతు సేహప్ర వర్తతే-స  
 ర్విషావిన్దనాయోగోనత్వపి నేత్యుక్తత్వాత్ - అపిస్తుయాద్విష్టుమ్- సంభావనంశక్తు  
 త్కర్షమావిషక్తుమత్యుక్తిః-అపిస్తుహి- అన్వవసర్గకామచారానుజ్ఞా- ధిగ్గేవదత్త  
 మపిస్తుయాద్విష్టువలమ్-గహ్వా-అపిసింఞ్చఅపిస్తుహిసముచ్చయేషు కాలాధ్వనోరత్యన్త  
 సంయోగే॥౧॥౩॥౫॥ఇహద్వితీయాస్యాత్ - మాసజ్కల్యాణీ మాసమధీతే-మాసం-గుడ  
 ధానాః- క్రోశంకుటిలానదీ- క్రోశమధీతే- క్రోశంగిరిః-అత్యన్తసంయోగేకేమ్-మా  
 సస్యద్విరధీతే- క్రోశస్యైకదేశేచర్వతఃఖిలస్వతన్తగ్గకర్తా ॥ ౧॥౮॥౫౮॥ క్రియా  
 యాంపాతం త్ర్యేణవివక్షీతోఽర్థగ్గకర్తాస్యాత్ ఖిలసాధకతమజ్కరణమ్॥౧॥౮॥౫౯॥

Use of Ardhavisarga in Telugu from the book Siddhantakaumudi. Ardhavisarga used in phrase like svatantrahkartā.

In addition to Jihvamuliya and Upadhmaniya characters, modern texts of Kannada and Bengali scripts also use U+1CF2 VEDIC SIGN ARDHAVISARGA.

ವಯಶ್ಚತೇ ಪತತಿ, ಗೇ  
 ಉಷ್ಣ ಪಾರಸ್ವತ  
 ವಯಃ - ಚತ - ತೇ -  
 ಚತು - ಪಾತ - ಅಮ  
 ವಯಶ್ಚತೇ ಪತತಿ, ಗೇ  
 ವಯಶ್ಚತೇ ಪತತಿ, ಗೇ  
 ವಯಃ - ಚತ - ತೇ -  
 ಚತು - ಪಾತ - ಅಮ

Use of Ardhavisarga in Kannada (From p 2 of L2/15-158).

অনুবাদ - কবর্গ কিংবা পবর্গ (পরে) থাকিলে বিসর্গের স্থানে (যথাক্রমে) ঙ্ক এবং ঙপ আদেশ হয়, চকার হেতু বিসর্গ ও হইবে। (যথা-) নুং ঙপাহি, নুং পাহি, নুং ঙ পাহি, নুং পাহি। নুন্ পাহি।।

আলোচনা - 'নুর্ + পাহি' এই অবস্থায় 'অত্রানুসিকঃ পূর্বসা তু বা' এই সূত্রের দ্বারা বিকল্পে 'ক' এর পূর্বে অনুসাসিক আদেশ করিয়া 'নুর্ + পাহি' এই অবস্থায় 'তশ্মিন্মিতি নির্দিষ্টে পূর্বসা' - এই সূত্রের সহায়তায় 'খরবসানয়োর্বিসর্জনীয়ঃ' এই সূত্রের দ্বারা পদান্ত রেফের বিসর্গ করিয়া 'নুং + পাহি' - এই অবস্থায় 'তশ্মিন্মিতি নির্দিষ্টে পূর্বসা' 'যথাসংখ্যামনুদেশঃ সমানাম্' এই দুই সূত্রের সহায়তায় 'কৃণোঃ ঙ্ক ঙপৌ চ' এই সূত্রের দ্বারা বিসর্গের উপস্থানীয় আদেশ করিয়া 'নুং ঙ পাহি' - এই একটি রূপ সিদ্ধ হয়। উপস্থানীয় আদেশ না হইলে অর্থাৎ সূত্রে চকার থাকায় বিসর্গ থাকিলে 'নুং পাহি' - এই দ্বিতীয় রূপ সিদ্ধ হয়। আবার অনুসাসিকের অভাব পক্ষে অর্থাৎ 'নুর্ + পাহি' এই অবস্থায় রেফের পূর্বের বর্ণের পরে অনুসারের আগম করিয়া 'নুং + পাহি' - এই অবস্থায় 'তশ্মিন্মিতি -' সূত্রের সহায়তায় 'খরবসানয়োর্বিসর্জনীয়ঃ' - এই সূত্রের দ্বারা রেফের বিসর্গ আদেশ করিয়া নুং + পাহি এই অবস্থায় - 'তশ্মিন্মিতি নির্দিষ্টে পূর্বসা' 'যথাসংখ্যামনুদেশঃ সমানাম্' - এই দুই সূত্রের সহায়তায় 'কৃণোঃ ঙ্ক ঙপৌ চ' - সূত্রের দ্বারা বিসর্গের উপস্থানীয় আদেশ করিয়া 'নুং ঙ পাহি' এই তৃতীয় রূপ সিদ্ধ হয়, সূত্রে চকার থাকায় উপস্থানীয় অবস্থাপক্ষে বিসর্গ থাকিয়া 'নুংপাহি' - এই চতুর্থ রূপ সিদ্ধ

এর ছক্-

ক	ডা	ঞ	ক	ঙ	ক	ক
ত	হ	ম	জ	ঙ	ক	ক
খ	ফ	ঙ	জ	ঙ	ক	ক
দ	ব	ণ				
ধ	ঢা	ম				
ন	ম	ং				
ল	ঙ প					
স			ফ			

Use of Ardhasvarga in Bengali in a printed book of Siddhantakaumudi.



Use of Ardhasvarga in a Tirhuta manuscript.

**Action requested:** Update the Script Extensions to extend U+1CF2 VEDIC SIGN ARDHAVISARGA used in Bengali, Kannada, Telugu and Tirhuta scripts.

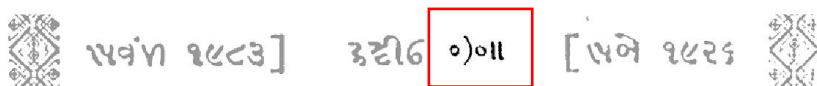
### 3 Khojki

Page 10 of Khojki proposal L2/11-021 identified the use of fractions, place holding marks, currency marks etc. These may be represented using the fraction and currency signs already encoded in the Common Indic Number Forms block.

#### 4.15 Number Forms and Unit Marks

Number forms and unit marks are found in Khojki documents (see Figure 3). The most common characters are signs for writing fractions: eg. 𑂔 U+A833 NORTH INDIC FRACTION ONE SIXTEENTH, 𑂕 U+A831 NORTH INDIC FRACTION ONE HALF, etc.; placeholder marks: eg. 𑂖 U+A836 NORTH INDIC QUARTER MARK; and currency marks: eg. 𑂗 U+A838 NORTH INDIC RUPEE MARK.

In the specimen below, the parenthesis-like character represents ) U+A838 NORTH INDIC RUPEE MARK. The number form represents the currency value of '8 ānā'.



These characters may be represented using those already encoded in the ‘Common Indic Number Forms’ block (U+A830..U+A83F).<sup>10</sup>

**Action Requested:** In Script Extensions extend the use of A830..A839 to Khojki.

## 4 Grantha

U+0BF3 **௫** TAMIL DAY SIGN is also commonly used as a sign of auspiciousness. It is hence termed in Tamil “*pillaiyār sulī*”, literally “the whorl of Ganesha”, referring to the spiral in its shape. It corresponds to siddham characters of other Indic scripts. In addition to Tamil **௫** also used in Grantha.

**No. 116.**

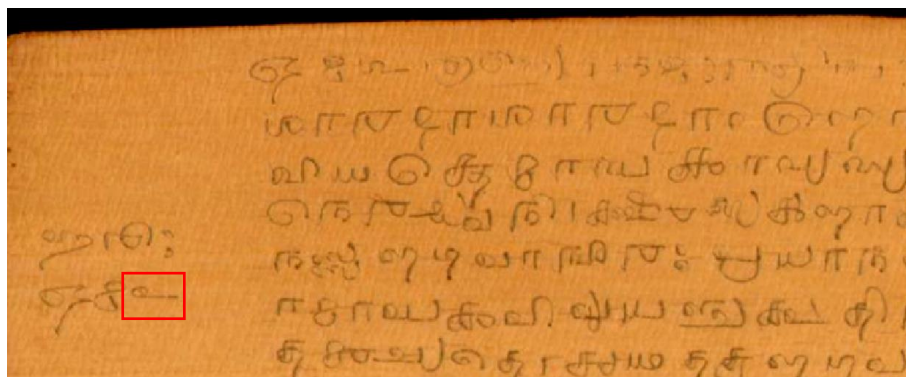
(No. 56—T. T.)

[On the south wall of the first prākāra of Tirumala Temple.]

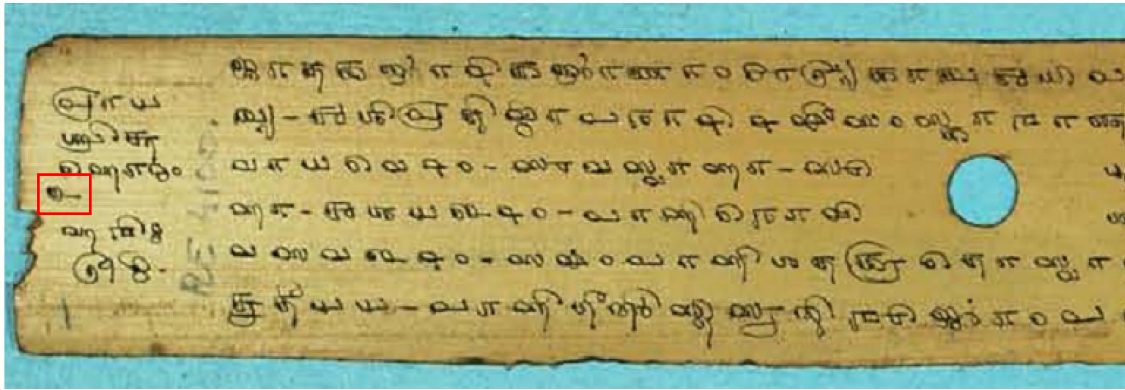
**Text**

1. **உ** ஸாஸ்திரம் || \* || வெகுண்டதோஷத் வெங்கடநிரொரெவெதி  
மொளரொயழ வாணிதீழமுயதெ தடின்தித [உ]தி ஸாவாஃதொவா-  
லிவ 1
2. [ஹதூ]மொஸௌயாடிவாவழிசுகி தவெஸ் ஃதநஸி முனெய்யுதெநொ  
மொஸௌலிவதிஜீயநதநஸனொஷணழ ஹதிஷணழி || க || \*

(From Tirumala Tirupati Devasthanams Inscriptions Vol 1: Early Inscriptions).







ॐ in Grantha manuscripts.

**Action Requested:** In Script extensions extend the use of 0BF3 to Grantha.