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 ☲ 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.
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 1CD6 VEDIC TONE YAJURVEDIC INDEPENDENT SVARITA.

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

2.1.2 Atharvavedic character
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
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
Vajasneeya Samhita in Tirhuta (from EAP790/21/101).

**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:”

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

Use of U+1CD0, U+1CD2 and U+1CD3. 
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

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

![Purushasukta in Odia script](http://www.odia.org/sanskrit/vishnu_purushasukta_odi.pdf)

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

---

1 http://www.odia.org/sanskrit/vishnu_purushasukta_odi.pdf
2.5 U+1CF2 VEDIC SIGN ARDHAVISARGA

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

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

Use of Ardhavisarga in Kannada (From p 2 of L2/15-158).
Use of Ardhavisarga in Bengali in a printed book of Siddhantakaumudi.

Use of Ardhavisarga 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. \text{U+\text{A833 NORTH INDIC FRACTION ONE SIXTEENTH}}, \text{U+\text{A831 NORTH INDIC FRACTION ONE HALF}}, etc.; placeholder marks: eg. \text{U+\text{A836 NORTH INDIC QUARTER MARK}}; and currency marks: eg. \text{U+\text{A838 NORTH INDIC RUPEE MARK}}.

In the specimen below, the parenthesis-like character represents \text{U+\text{A838 NORTH INDIC RUPEE MARK}}. The number form represents the currency value of ‘8 நா’. These characters may be represented using those already encoded in the ‘Common Indic Number Forms’ block (U+\text{A830-A83F}).

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

4 Grantha

\text{U+\text{0BF3 ட்}}: TAMIL DAY SIGN is also commonly used as a sign of auspiciousness. It is hence termed in Tamil “pillaiyar sulji”, 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 ட் also used in Grantha.

\text{No. 116.}
\text{(No. 56—T. T.)}
\text{[On the south wall of the first prākāra of Tirumala Temple.]}  
\text{Text}

1 \text{ட் [இய]} \text{தமிழ் நாள் விந்தை பெருக்குகள் எச்சூர் புனித தமிழ் நாள் விந்தை பெருக்குகள் எச்சூர்}

2 \text{ட் இய தமிழ் நாள் விந்தை பெருக்குகள் எச்சூர் புனித தமிழ் நாள் விந்தை பெருக்குகள் எச்சூர்}

(From Tirumala Tirupati Devasthanams Inscriptions Vol 1: Early Inscriptions).
Action Requested: In Script extensions extend the use of 0BF3 to Grantha.