Proposed Property Changes for Ardhavisargas

Sriniidhi A  
srinidhi.pinkpetals24@gmail.com

Sridatta A  
sridatta.jamadagni@gmail.com

March 6, 2018

The document L2/18-059 recommended to explore the interaction of following characters.

- U+1CF2 \text{VEDIC SIGN ARDHAVISARGA}
- U+1CF3 \text{VEDIC SIGN ROTATED ARDHAVISARGA}

Ardhavisarga marks either \textit{jihvamuliya}, a velar fricative occurring only before the unvoiced velar stops \textit{ka} and \textit{kha}, or the \textit{upadhmaniya}, a bilabial fricative occurring only before the unvoiced labial stops \textit{pa} and \textit{pha}.

L2/18-059 asks “\textit{Is the character in the Bengali samples U+1CF6 or a variant of U+1CF5 VEDIC SIGN JIHVAMULIYA?}” The character used in Bengali is not a variant of U+1CF5 VEDIC SIGN JIHVAMULIYA as it used for both \textit{jihvamuliya} and \textit{upadhmaniya}. It is an ardhavisarga character.

Interaction with the characters

In addition to Consonants, \textit{visarga}, \textit{anusvara}, \textit{candrabindu}, vowel signs etc can also occur preceding the ardhavisarga as in कुवोः \textit{ᳲ} कू पौ क्पृच्छ, नूँपाहि, नूँपाहि, दुःख अन्नकरण. After the Ardhavisarga spaces, \textit{ka/kha} and \textit{pa/pha} can occur.

When ardhavisarga is used independently it can be enclosed by spaces, parentheses, square brackets and similar punctuation marks.

Currently the characters are analyzed as spacing combining mark (gc=Mc). The two characters are vowelless consonants. As indicated in L2/09-343, there is no compelling reason to give these characters GC=Mc instead of Lo. Line breaking occurs both before and after the Ardhavisarga. The current Indic_Syllabic_Category is Visarga, however Ardhavisarga’s behavior is different from Visarga. Consonant_Dead is the appropriate category as it indicates a vowelless consonant.

The following property changes are recommended for above two ardhavisargas.

<table>
<thead>
<tr>
<th>Property values</th>
<th>Current</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>General category</td>
<td>Mc</td>
<td>Lo</td>
</tr>
<tr>
<td>Line Break</td>
<td>CM</td>
<td>AL</td>
</tr>
<tr>
<td>Indic_Syllabic_Category</td>
<td>Visarga</td>
<td>Consonant_Dead</td>
</tr>
</tbody>
</table>
Suggestions on the extended character properties like ID Start, XID Start, Grapheme Cluster Break, Word Break, Sentence Break etc for the above two characters are requested from UTC.

As documented in L2/17-424 Bengali, Kannada, Telugu and Tirhuta be added to the scripts in ScriptExtensions.txt for U+1CF2 VEDIC SIGN ARDHAVISARGA.

Also, the dotted circle ◯ should not be displayed in Code chart.
Figure 1. कख and पफ from a Sanskrit Grammar book.

Figure 2. A Sanskrit manuscript showing ardhavisarga in beginning of the line in the word कपौ.
Figure 3. Occurrences of Ardhavisarga in a Sanskrit text in Telugu script.
Figure 4. Occurrences of ardhavisarga in the beginning of the word/akshara.
Figure 5. Use of ardhavisarga enclosed by spaces and parentheses in the Sanskrit book संधिं: written by G Mahabaleshwara Bhat published by Samskrita Bharati, Bengaluru.
Figure 6. Use of Ardhisarga at the end the words and line in the book तैरीय-मन्त्रकोश:
published by Sri Ramakrishna Math, Chennai.
Figure 7. Use of ardhavisarga and rotated ardhavisarga enclosed by spaces and square brackets in the Sanskrit book शिष्याप्रथमकृतः written by Shriramana Sharma (2010).

Figure 8. Use of Ardhavisarga in the phrases कुप्पोः, कः कः, नृंपाहि and नृंपाहि in the Sanskrit Grammatical work Madhyasiddhantakaumudi.