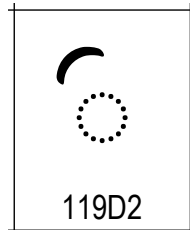


Reconsidering **gc** and **InPC** values of **U+119D2 NANDINAGARI VOWEL SIGN I**

To: Unicode Technical Committee
 From: Liang Hai / 梁海 <lianghai@gmail.com>
 Date: 9 January 2019



1. Proposal

Change **gc** (**General_Category**) and **InPC** (**Indic_Positional_Category**) values of U+119D2 NANDINAGARI VOWEL SIGN I (newly encoded in Unicode 12.0) from **Mn** and **Top** (as of the Unicode 12.0 beta review) to **Mc** and **Left**.

2. Background

In the original proposal L2/17-162, [Final proposal to encode Nandinagari in Unicode](#) (Anshuman Pandey, 2017), the Nandinagari vowel sign *i* was proposed to have **gc** = **Mn** in §6.1 *Character Properties* (page 16) and **InPC** = **Top_And_Left** in §6.4 *Positional Categories* (page 17). Also, an introductory paragraph was included in §4.2 *Vowel signs* (page 4):

The [...] VOWEL SIGN I is positioned at the top-left edge of letters that have headstrokes. For other letters it hangs above the top-left portion of the body. The style of writing the sign varies considerably, and in rapid writing the stroke may be positioned to the left of the base and may appear as an elongated curve.

As it was pointed out by Pandey, the attested shapes of the encoded written form fall in a wide spectrum, varying from a typical non-spacing stroke to a stroke that stands on the baseline and significantly participates in the inline relationship between letters with its own advancing.

- Pandey explained to the author in a phone call that he chose the above-base variant for the representative glyph because it is the more characteristic variant for the Nandinagari script.
- An examination of the original proposal's appended figures shows that Figures 1–3, 7, 10, 12–13, 16, and 18 appear to have pretty typical pre-base variants of vowel sign *i*, while the variants in Figures 4–6, 8, 11, 14–15, 19–20, and 31 do not exhibit significant advancing. (Other figures either do not have this structure or list both variants without recognizing the default style.) Therefore both above- and

pre-base variants appear to be common. See this document's *Appendix* for some extracted figures that have been annotated.

Instead of the originally proposed `Mn/Top_And_Left`, the Unicode 12.0 beta has `Mn/Top`. This change appears to be a result of the data file owners' understanding.

- Note that `gc` and `InPC` are coupled under the current practice. For any `InPC` values involving `Left` and/or `Right`, the `gc` value shall be `Mc (Spacing_Mark)`; while if purely `Top` and/or `Bottom`, `gc` shall be `Mn (Nonspacing_Mark)`.
- Therefore the original `Mn/Top_And_Left` is not actually a valid proposal.

3. Analysis

The choice between `InPC` values is assumably based on at least the following two factors:

1. *Visual placement* in relation to the base.
 - In particular, whether the written form (or a part of it) occupies its own inline space between those typical inline letters.
2. *Intended shaping* of whether to be reordered or not.
 - Although the Unicode Standard does not specify how the exact reordering process should be done in “higher-level protocols”, whether an encoded character is meant to be reordered is indeed a fundamental part of the Unicode's encoding model(s) for Indic scripts.
 - An appropriately defined to-be-reordered character can take advantage of the architectures that have enabled existing characters, then implementations can be simplified and less error-prone.
 - To-be-reordered combining marks for Indic scripts have been effectively defined by `InPC` values that involve `Left`.

Taking the whole spectrum of variants into consideration, the Nandinagari vowel sign *i*'s variants suggest a reordered character. Even the non-spacing variants can benefit from being reordered, as they are apparently written at the leftmost end of a syllable. U+115B8 SIDDHAM VOWEL SIGN E is an existing `InPC = Left` character that is nearly non-spacing.

Note that `Top_And_Left` is not a good choice for Nandinagari vowel sign *i* because the value seems to be typically applicable to split vowel signs, which are multi-sided and both parts of them tend to correspond to other single-sided signs:

U+0B48 ORIYA VOWEL SIGN AI (ꣳ)

U+0DDA SINHALA VOWEL SIGN DIGA KOMBUVA (◌්)

U+17BE KHMER VOWEL SIGN OE (្រ)

U+114BB TIRHUTA VOWEL SIGN AI

While the rest two are of less understood scripts and thus probably are not good examples to follow (the Siddham one is arguably also a split vowel sign):

U+1C29 LEPCHA VOWEL SIGN OO
U+115B9 SIDDHAM VOWEL SIGN AI

Therefore **InPC** = **Left** is the appropriate choice, and consequently there shall be **gc** = **Mc**.

The author has also consulted Srinidhi A and Sridatta A, and both of them agree with the author's proposal on **gc** and **InPC**. They have further suggested the representative glyph should take the pre-base variant and plan to submit a proposal later.

4. Acknowledgements

The author would like to thank Srinidhi A, Sridatta A, Anshuman Pandey, and the Script Ad Hoc group for their feedback. Srinidhi and Sridatta also kindly provided additional samples of Nandinagari vowel sign *i*'s variants, however the author eventually decided to limit figures to the ones available in the original proposal, for the sake of a simpler context.

Appendix

All figures below are extracted from the original proposal with original figure numbers.

Selective occurrences of vowel sign *i* are highlighted with blue circles.

Figure 12, pre-base:

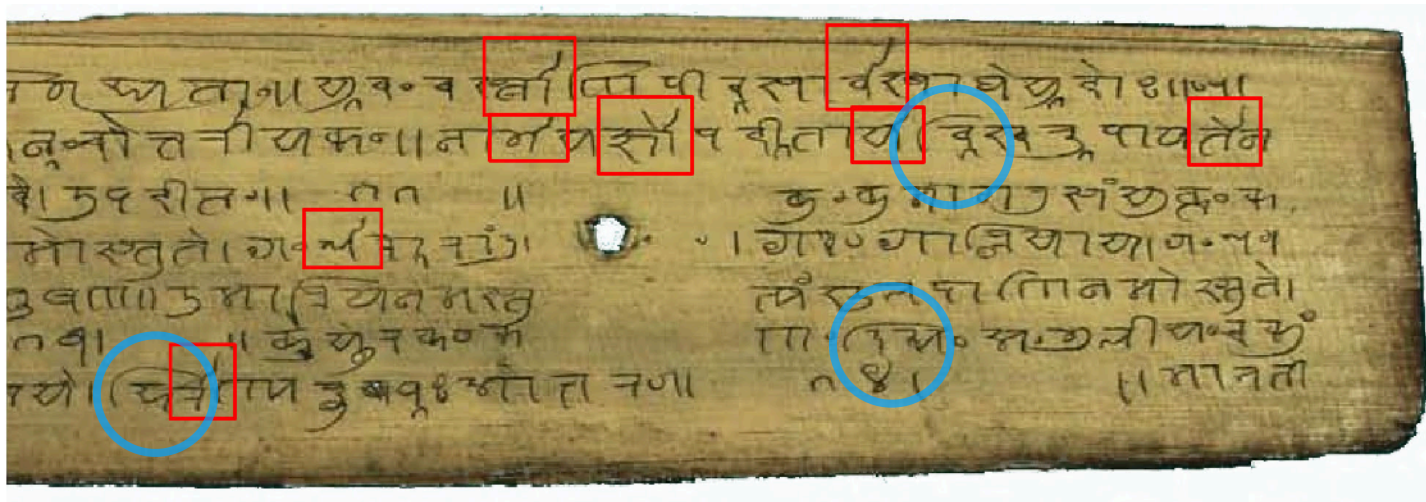


Figure 16, pre-base:

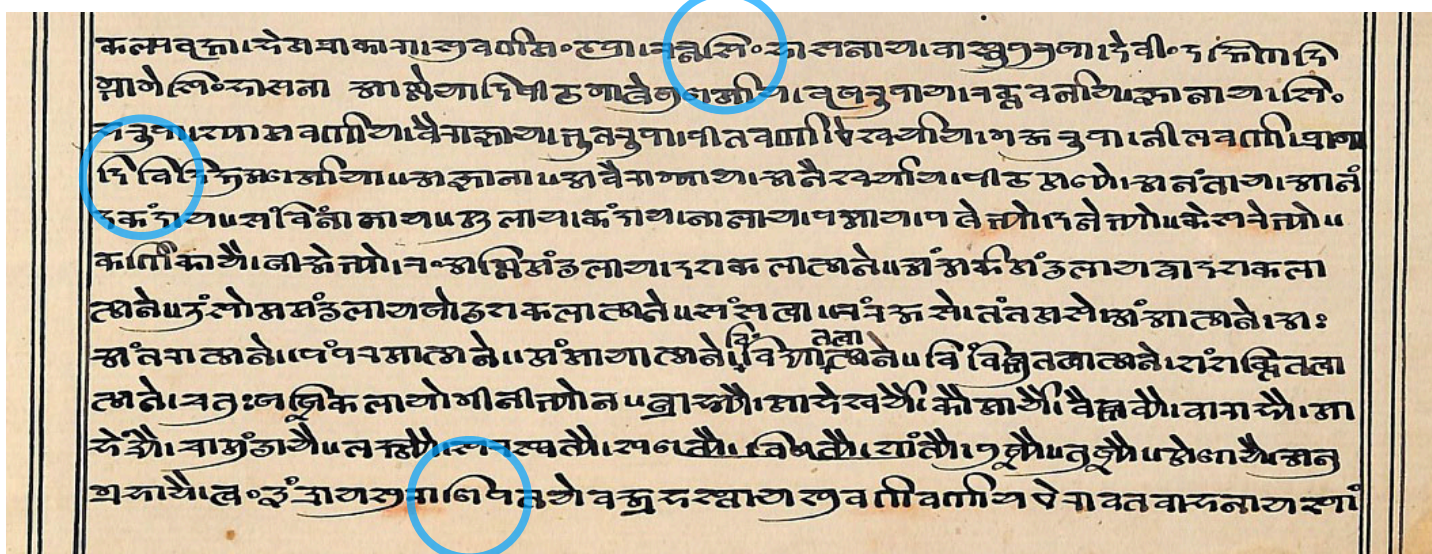


Figure 3, pre-base:

नमः शानिकला कोटिके लुमाना कपा
 लकल्या विजाय राने वा खत्रिसमस्र नुवना सदा सपा
 संमद्रा पाजा वी जत वा मेश्वर वामन दृक् कल्याणा ववा
 धीश्वर गजा गोकुल दत्र विपुलाय कां त्रद त्र वै वि वैधवा य
 लुका कुल क मल मा त्रंड क द न व रंड क लिंगा को द ड ग ड मा
 त्रंड प को वि वि रा पं ग धी र सु ख व त्रि ग य दि श्या व ट्टा ज न
 त्रि ग य सां का र क न र व त्रि ग य म श क र ल व द वि
 सां मं त शि ग व मा र कां का वा त्रु दि स न टां क र नि गृ क र
 प्या गी न व त्रु व मी व रा ह लीं न वृ त्त र श र त्रा जा व
 नि वृ त्त त र श मा मी लं क र शि शो म वं शो द व शि व र ना नां व व
 क व त्रि क ल्या गा व रा स व स क द्या वि नो दे न मा कं क री त्रि द क्षिा
 दि सा वां दि वि क जे य टा त्रं वि ज वा क र शि दे ज प गी शी त्र क ट

Figure 6, above-base:

मन्त्रा विस्तारं चण्ड मन्त्रेण विष्णवे चैव चण्ड मन्त्रेण विष्णवे चैव चण्ड मन्त्रेण
 उहउं अ के वा न म् उं तो विष्णवे चैव चण्ड मन्त्रेण विष्णवे चैव चण्ड मन्त्रेण
 पा वै न द श्या सै हं उं तो विष्णवे चैव चण्ड मन्त्रेण विष्णवे चैव चण्ड मन्त्रेण
 कु विष्णवे चैव चण्ड मन्त्रेण विष्णवे चैव चण्ड मन्त्रेण विष्णवे चैव चण्ड मन्त्रेण
 शि त्रु उ वै श टो वा विष्णवे चैव चण्ड मन्त्रेण विष्णवे चैव चण्ड मन्त्रेण
 को शि अ वि सु ना रा म ० प क ल्ले वि का वै ल्ले शि उ अ प री च
 ६ च वै न सी शो क टो र श हि न उ म् द शि शी भा ॥ ६ र पो च ६ र म ३ उ ६ उ

Figure 11, largely above-base:

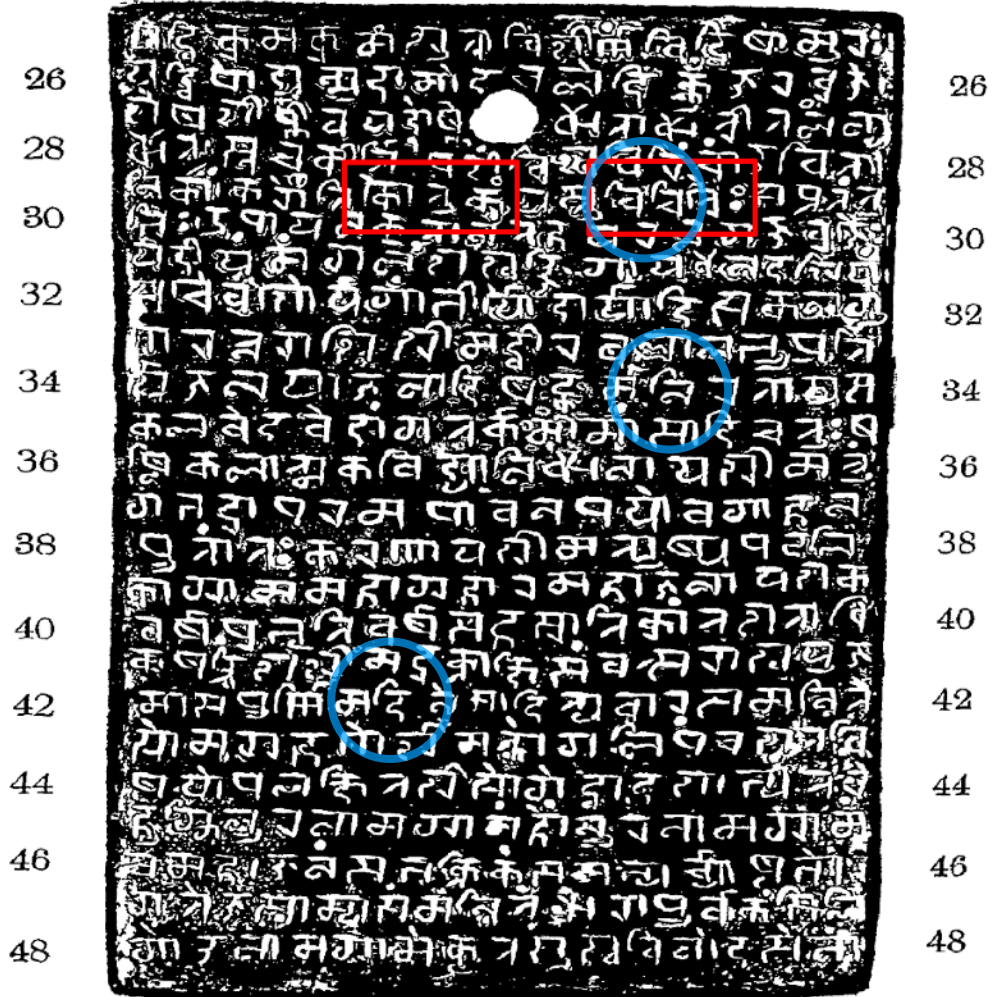
ल मी या क्क म या क्क वै ना ग्गा या रै र म् यो या क्क ल मी या क्क
 ना या इ ती क ० व प त्रु कां श ध या म् प त्रु कां श ल ० श रु नी धी
 श म् मी या तो न र ती का त वे व तै वि षु ० म शी ग म् रा वा
 ॥ - ति श्र ० त र्का ना क र्णा मी क न व म ० व र्णा उ क्क
 वृ षि श्रु ० श्रु ० वृ ० द शि वि नो य को वै व ता ग ० वी क्क ०
 गुं ० श क्क ० मी ० मी ल क्क ० म म ल को ली पा ल म् म की री
 ल ० म श य न ग म् रा वा पी त वै क्क ये वि नि यो ग श

Figure 15, largely above-base:

Plate IV-b.

No. 28 (Cont.)

Second plate, I side



R. S. Panchamulika.