

Proposal to encode A8FE DEVANAGARI SIGN SPACING ANUSVARA

Shriramana Sharma, jamadagni-at-gmail-dot-com, India

2015-Jun-06

This is a proposal to encode one character in the Devanagari Extended block for Samavedic:



A8FE DEVANAGARI SIGN SPACING ANUSVARA

This is in contrast to the regular anusvara for this script 0902 ः DEVANAGARI SIGN ANUSVARA as also to the various Vedic anusvara-s seen in Samavedic. On the other hand, this is parallel to the spacing anusvara-s in other Indic scripts which are attested for Vedic (0982 Bengali, 0B02 Oriya, 0C02 Telugu, 0C82 Kannada, 0D02 Malayalam, 11302 Grantha) and glyphically identical to all of them except Bengali. However it should be positioned vertically centered with the Devanagari digits identical to 0966 ० DEVANAGARI DIGIT ZERO.

§1. Discussion

The regular Devanagari anusvara 0902 ः is non-spacing. This poses a problem when composing texts of the Sama Veda since these use digits on the mainline to denote *svara*-s: (Below, we refer to the written representation of the linguistic pattern [C*]V as “syllable”).

1) In the Ṛc-s (verses) a *kampa* or “aggravated” *svarita svara* is marked by a 2 above the syllable (or inferred as continued from a previous syllable), a KA or avagraha above the syllable, and a digit 3 following the syllable (see L2/09-372 pp 13 and 14 and L2/15-162 p 4). The digit 3 here denotes the *anudātta svara* in the latter part of the *kampa*.

2) In the *Sāman*-s (melodies), secondary *svara*-s in which a syllable’s vowel should be continued to be sung are marked by digits following the syllable.

In both these cases, any anusvara following the vowel has to be marked after the *anudātta* or secondary *svara*-s since *svara*-s are intrinsic to the vowel whereas the anusvara is not. Marking the anusvara before the digits would imply that it is the anusvara which takes the extra *svara*-s, which is false. Thus while signs denoting the anusvara, visarga etc normally directly follow the syllable, in such Vedic texts they have to come after any digits following the syllable. (Text shaping engines should be aware of this requirement and properly process such sequences. See also my Grantha proposal L2/09-372 p 40.)

Now, this is not much of a problem in scripts where the regular anusvara is a spacing mark, since it may then simply be input in encoded text after the digits. However, when the regular anusvara is non-spacing, it cannot be dislocated from its base. In this case, a spacing version needs to be used in these particular contexts.

Now Devanagari and Gujarati are the two major Indic scripts having the regular anusvara as non-spacing. Of these, Devanagari is attested to use a spacing circular anusvara in these contexts in contrast with the regular non-spacing dot-like anusvara. Based on these attestations, we propose that DEVANAGARI SIGN SPACING ANUSVARA be encoded.

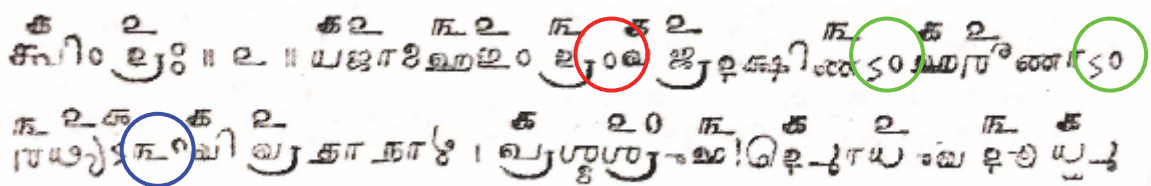
Note that the attestations come from three corners of India (Kolkata/West Bengal, Kanchipuram/Tamil Nadu and Jamnagar/Gujarat) and disparate time frames (1870s, 1976 and 2012). As such, this is not a one-time experiment by a particular publisher.

Parallel attestations for comparison are provided from Grantha where the regular anusvara is spacing and is hence used both directly after the syllable and after digits.

§2. Attestations

Note: Red, blue and green circles respectively mark the regular non-spacing anusvara, the proposed spacing anusvara and Vedic anusvara-s in Devanagari. In Samavedic texts, the first two in fact denote the homorganic nasal of the following consonant whereas it is the Vedic anusvara-s which denote the actual linguistic anusvara. The corresponding parts in Grantha are marked with the same colours.

Rc 334



१ २ ३ २ ३ १ २ ४ १ २ ४ २ ५ १ २
यजामह इन्द्रं वज्रं दक्षिणं चरिणं प्रथ्या ३० विव्रतानाम्
१२ २२ ३ १ २ ३ १ २ ४ १ २ २२ ३ १ २ ४ १ २ २२
प्रश्नश्रुभिर्दोषु वदूँ धाभु वद्विसेनाभिर्भयमानो विराधसा ३१०२

க. சு.மி.நி.தா.பு. || வு.ணீ.ஹா.உ.ஹா.தா.ரா.த.உ.ந.வீ.
 ஸ்வே.த.ஸா.பு. || சு.ஸ்யா.த.உ.ந.ஹா.சு.ஸு.த.உ.ஹா.வா.ஸ்யா.
 ஸ.க.த.பு.உ.யா.த.உ.ந.ஹா.த.உ.சு.ந.ஸு.த.உ.சு.ந.ஹ.||யா.||தீ.சு.

அ.நி.ந்.தா.ம். | வு.ணீ.ம.ஹா.இ. | ஹோ.தா.ரா.உ.ந.வீ. |

ஸ்வே.த.ஸா.ம். | அ.ஸ்ய. யா.உ.ந.ஜா. | அ. | ஹோ.தா.வா. |

ஸ்யா.ஸு.க்ரு.த.ம். | இ.தா.உ.ந.மா.உ.ந.ஜா. | ஹோ.தா.உ.ந.பு.இ. | தா. ||யா. || உ.

அ.நி.ந்.தா.ம். || வு.ணீ.ம.ஹா.இ. | ஹோ.தா.உ.ந.வீ. | ஸ்வே.த.ஸா.ம். ||
 அ.ஸ்ய.யா.உ.ந.ஜா. | அ. | ஹோ.தா.வா. || ஸ்யா.ஸு.க்ரு.த.ம். | இ.தா.உ.
 ந.மா.உ.ந.ஜா. | ஹோ.தா.உ.ந.பு.இ. || தா. ||

அ.நி.ந்.தா.ம். || வு.ணீ.ம.ஹா.இ. | ஹோ.தா.உ.ந.வீ. | ஸ்வே.த.ஸா.ம். || அ.ஸ்ய.யா.உ.ந.ஜா. |
 அ. | ஹோ.தா.வா. || ஸ்யா.ஸு.க்ரு.த.ம். | இ.தா.உ.ந.மா.உ.ந.ஜா. | ஹோ.தா.உ.ந.பு.இ. || தா. ||

(Note: The usage of the international digits as svāra markers in the second Devanagari samples shown in these pages is not to be taken as grounds for encoding these separately! They are merely a substitute used by the printer possibly due to absence of appropriately sized Devanagari digits which are in fact the appropriate characters to be used here.)

दूताऽ३०वोऽ३विश्वावेदसाम् ॥ हव्यवाहाम् । आमाऽ२र्त्ताऽ२३४याम् ॥ याजिष्ठम् ।
ऋ । जासेऽ३हाइ ॥ गिरा । औऽ३होवा । होऽ५इ ॥ डा ॥

4 5 0 4 4 3 2 2 3 0 5 0 2 0 1 0 0 अ
प्रभूर्जयंताम् ॥ महाऽ३४३० विपोधाम् । मूरैर्मूरं पुरान्दमीऽ२३-
2 3 2 2 3 0 5 3 2 2 3 5 4 5
णाम् ॥ नयाऽ३४३न्तगीर्भिः । वनाऽ३४३धियंथाः ॥ हरि-
4 5 5 2 1 1 1 1
श्मश्रुन्नवर्मणाऽ६ । हाउवा ॥ धनऽ३र्चीऽ२३४५म् ॥

प्राभूर्जायंताम् ॥ माहाऽ३४३०विपोधाम् । मूरैरमूरं पुरान्दार्माऽ२३णाम् ॥
नायाऽ३४३न्तगीर्मिः । वानाऽ३४३धियंधः ॥ हारिश्मश्रुन्नवर्मणाऽ६॥ हाउवा ॥
धनाऽ३र्चीऽ२३४५म् ॥ (दी.८। प.८। मा.७)११(डे।१२१)

प्रत्यस्मैपिपाहाउ ॥ आइषाऽऽताइ । वाइश्वानिवाइ । दूषेऽऽहाऽऽइ । भाऽऽरा ।
आराऽऽगामा ॥ याजाऽऽहाऽऽ । ग्माऽऽयाइ ॥ अपाऽऽर । श्वाऽऽदाऽऽर४-
औहोवा ॥ ध्वानेऽऽनाराऽऽर४५ः ॥ (दी.७। प.११। मा.९)११(चो । ६०४)

§3. Sources for above attestations

Grantha: Kauthuma Sama Veda, AVMSSGDPS Samiti, Kumbakonam, reprinted 1985.

Devanagari: (in order of appearance of samples after Grantha in each page)

Kauthuma Sama Veda:

- 1) ed. Satyavrata Samashrami, pub. Asiatic Society, Kolkata, 1871-75.
- 2) ed. R Shankaranarayana Shrauti, pub. Shri Chandrashekharendra Sarasvati Shashtyabdapurti Trust, Kanchipuram, 1976.

Ranayaniya Sama Veda:

- 3) ed. Vijaya Kumar Sharma, pub. Shri Sharada Peeth Vidya Sabha, Jamnagar, Gujarat, 2012.

§4. Unicode Character Properties

A8FE;DEVANAGARI SIGN SPACING ANUSVARA;Mc;0;L;;;;;N;;;;;

The code point A8FE is chosen since it is the first available in the Devanagari Extended block after A8FC and A8FD in the current pipeline. Note that only one more codepoint A8FF is available in this block for future use.

The rest of the given properties are based on those of the cognate character 11302 GRANTHA SIGN ANUSVARA. Other properties such as Indic categories, line-breaking etc should also be taken as that of 11302. For collation, the character is to be treated as equivalent to 0902 DEVANAGARI SIGN ANUSVARA in whose place it is used when a spacing form is needed.

Annotate as: “used after digits indicating secondary svāra-s in Samavedic”.

§5. Alternatives

I considered proposing a generic character for the Vedic Extensions block since the shape is generic. However, it would then have script=Inherited permitting use with (the majority of the) scripts which already have a spacing anusvara. This may lead to multiple encoded representations of the same text. I thus decided to avoid doing that.

Another alternative is to merely annotate 0966 DEVANAGARI DIGIT ZERO as “used for an anusvara after digits indicating secondary svāra-s in Samavedic”, since it is the exact shape that is required and the other digits are already used for the svāra-s. Devanagari deviates in any case by requiring a separate character for these contexts, so it may well be 0966 to avoid encoding a new and confusable character. I leave it to the UTC to decide this.

§6. Official Proposal Summary Form

(Based on N4502-F)

A. Administrative

1. Title

Proposal to encode A8FE DEVANAGARI SIGN SPACING ANUSVARA

2. Requester's name

Shriramana Sharma

3. Requester type (Member body/Liaison/Individual contribution)

Individual contribution

4. Submission date

2015-Jun-06

5. Requester's reference (if applicable)

6. Choose one of the following: This is a complete proposal (or) More information will be provided later

This is a complete proposal.

B. Technical – General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters), Proposed name of script

No

1b. The proposal is for addition of character(s) to an existing block, Name of the existing block

Devanagari Extended

2. Number of characters in proposal

1 (one)

3. Proposed category

Category A, contemporary

4. Is a repertoire including character names provided?

Yes

4a. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&P document?

Yes

4b. Are the character shapes attached in a legible form suitable for review?

Yes

5. Fonts related:

a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?

Shriramana Sharma (though this should not be necessary as any South Indic anusvara can be adapted)

b. Identify the party granting a license for use of the font by the editors (include address, e-mail etc.)

Shriramana Sharma

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

No

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes

7. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

Yes

8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script.

See detailed proposal.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

No

2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

Yes

2b. If YES, with whom?

Various scholars of the Kauthuma Sama Veda in Tamil Nadu (contact details available if needed);
Vidvan Parameshvara Bhat, scholar of the Ranayaniya Sama Veda, Uttara Kannada District, Karnataka
2c. If YES, available relevant documents

None specifically. The matter was discussed in person and via phone.

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

Those who wish to digitize Devanagari texts of the Kauthuma Sama Veda.

4a. The context of use for the proposed characters (type of use; common or rare)

Common in the context of Devanagari texts of the Kauthuma Sama Veda.

4b. Reference

See detailed proposal.

5a. Are the proposed characters in current use by the user community?

Yes

5b. If YES, where?

Where there are scholars of the Kauthuma Sama Veda - Tamil Nadu, Gujarat, Varanasi etc

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

Yes

6b. If YES, is a rationale provided?

Yes

6c. If YES, reference

The character belongs to the Devanagari Extended block where sufficient space is available.

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

Only one character is proposed.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character?

The character is similar in appearance to 0966 DEVANAGARI DIGIT ZERO. It is similar in appearance and function to other Indic (esp. South Indic) anusvara-s.

10b. If YES, is a rationale for its inclusion provided?

Yes

10c. If YES, reference

The practice is to encode anusvara characters per-script.

11a. Does the proposal include use of combining characters and/or use of composite sequences?

Yes

11b. If YES, is a rationale for such use provided?

Yes

11c. If YES, reference

The character is by nature a combining mark.

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?

13c. If YES, reference:

-O-O-O-