

## Comments on Devanagari short vowels

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

2010-Nov-27

This document is with reference to Anshuman Pandey's documents L2/09-320, 377 and 403. In L2/09-320 Anshuman had proposed seven characters for the Devanagari block, but this proposal "fell through the cracks" when it came into conflict with the Govt of India's Kashmiri vowels proposal (represented in its final form by L2/09-389 and which had its fruit in Unicode 6.0). While Anshuman has not yet re-proposed his characters after this, I urge the UTC to take recognition of at least some aspects of his proposal which are not encoding-specific. I also place on record my comments on some of his characters.

### §1. Annotations and editorial remarks in the code chart

One important point that Anshuman brought to light in L2/09-320 is that those characters currently named DEVANAGARI LETTER/VOWEL SIGN SHORT E/O are not really "for Dravidian transliteration" only, but were originally introduced by Hoernle for the Bihari languages Bhojpuri, Magadhi and Maithili, after which "the best Benares pandits" adopted them. It is possible that these characters are annotated "for Dravidian" as a figment left over from ISCII wherein it was noted (probably with consideration for only the major Indian languages) that these are "for transliteration from Dravidian".

As it is, Anshuman's evidence is sufficient, in my opinion, to immediately remove the annotation "for transliteration from Dravidian", or to at least *primarily* annotate these as "for Bhojpuri, Magadhi and Maithili" and *secondarily* on transliteration.

Similarly, in L2/09-377 Anshuman points out that while the UE/UUE characters are limited to Kashmiri, the OE/OOE and AW characters are *not*, and that again, they are used for Bhojpuri, Magadhi and Maithili as well.

As such, the Unicode 6.0 code chart unfortunately ignores Anshuman's feedback and labels *all* these characters – including the OE/OOE/AW characters – as "for Kashmiri" – either as an annotation or as a sub-heading in the code chart description. This is despite Anshuman requesting in L2/09-403 "the characters should not be described in the names list as being specific to Kashmiri".

If it is considered important to indicate the languages for which these characters are used (say in order to avoid kneejerk questions as to why these characters are present

when they are not used for the major languages like Hindi), they should be annotated appropriately and not with a reference to Kashmiri only.

There are two options here. *Either* the subheadings “Dependent vowel signs for Kashmiri” etc should be modified to *add* the names of the other languages. *Or*, the subheadings must be made general and per-character annotations should be added as in Phags-Pa (to give a geographically distant but conceptually close example). I believe the latter would be more appropriate to accommodate future additions. Rather than make subheadings of unwieldy length, one had better make annotations.

Further, in L2/09-403 Anshuman had stressed that the AW characters should instead be named SHORT AU characters. Given that the characters are now encoded as AW, they should get an annotation reflecting Anshuman’s intention – that these should be annotated as representing the short diphthong /au/ when used for the Bihari languages.

## §2. Characters proposed by L2/09-320

Now among the characters proposed by Anshuman in L2/09-320, two (LETTER/VOWEL SIGN AW aka SHORT AU) are already encoded courtesy of the Kashmiri proposal. I therefore restrict my comments to the other characters.

### §2.1. Short AI

ॐ DEVANAGARI LETTER SHORT AI  
ॐ DEVANAGARI VOWEL SIGN SHORT AI

Of the characters proposed, I feel the short AI characters are the best candidates for encoding, with their shapes coming into line with the existing Devanagari characters. They will also complete the set of short vowels for E/O/AI/AU.

However, if the short AU characters are named AW, should the short AI characters be named AY?! The nomenclature rules as applicable here are not clear to me. If these are then to be named AY, then they should receive an annotation to the effect that they are used for short /ai/. It would be much more straightforward to use the word SHORT AI in the name as proposed by Anshuman, but then that would then not correspond with AW.

*However*, if AW was named that way only in an effort to imitate the Kashmiri pronunciation, then it would be appropriate to retain the name SHORT AI for this character since it has no such issues.

I recommend that this character pair is encoded, and with the name SHORT AI.

## §2.2. Reversed E

### 𑀓 DEVANAGARI LETTER REVERSED E

Of all the proposed characters, I feel this is *not* a valid candidate for encoding. It does *not* fit in with the glyph pattern for Devanagari, it does *not* have a distinct corresponding vowel sign, and to top it all off, it is an easy confusable with the regular E. Already I see many people on blogs and website using the short E vowel signs instead of the regular E ones for Sanskrit verses (where there is no short E sound). The reason is silly: “it looks stylish”.

Similarly, if this “reversed E” is encoded, it would give way for people to start using that for similar silly reasons thereby creating unnecessary problems with text search – you don’t get E when you expect it.

Given that Hoernle’s system is what has been accepted already for short E/O/AU for real-world usage, it would be best to keep to that for AI as well and avoid wavering and shifting. Thus the “Grierson’s scheme for short vowels” argument will not stand, unless there is a compelling academic need to distinctly encode texts that use Grierson’s scheme.

I therefore recommend that this character is *not* encoded. If at all any text needs to be presented that uses this scheme, fonts should be sufficient for the variation.

## §2.3. Neutral and sonant vowel signs

### ◌̣ DEVANAGARI SIGN NEUTRAL VOWEL

### ◌̤ DEVANAGARI SIGN SONANT VOWEL

These two characters are in the grey area. They are said to represent the articulation in Bihari of a vowel that would not be articulated in Hindi, given the same written form. However, is this character needed by the natives of that language to learn/pronounce their language properly, or is it merely a device for the reader familiar with Hindi to remember the proper pronunciation for Bihari? The evidence presented seems in favour of the latter. Thus this character is merely like an IPA diacritic – to clarify pronunciation.

What is however *not* clear to me is why two distinct nomenclatures and characters are needed. There seems to be no phonemic difference asserted between the sounds that are enunciated in the penultimate and ultimate positions where they would *not* be enunciated in Hindi – they are merely the inherent vowel /a/ albeit perhaps different allophones of it. Given this, is it really needed to distinguish these two? Again, these seem merely phonetic markers. Personally, I feel these should be encoded only if these need to be used by either natives (which is unlikely) or are needed for scholarly studies in this field.

As I have insufficient data to judge on this, I do not make a specific recommendation regarding these two characters. I only remark that if at all these characters are encoded, some thought should be given to whether they should receive GC=Lm (like 0971 DEVANAGARI SIGN HIGH SPACING DOT which is also a sound modifier) or the suggested GC=Mc.

### §3. Conclusion

An important aspect not to be forgotten is that if at all these characters are to be encoded, they must perforce go to the Devanagari Extended block as there is only one lone empty space in the original Devanagari block.

In the interests of completing the set of short vowels, a proposal to encode at least the short AI characters after Hoernle's system would be welcome.

Apart from this, given that the editorial changes that are needed in the Devanagari code chart do not require any follow-up proposal, I request that they be done expediently to ensure the proper representation of the characters that are already encoded.

-0-