

UTC Members

Submitted by Peter Scharf, 2007-10-18

The following lists the significant differences between L2/07-230 (N3290), presented to the UTC scripts committee in Redmond Washington on 7-8 August 2007, and L2/07-343 (N3366). Minor corrections and editorial revisions conditioned by these changes are not included. Hence, while the elimination of the *ṛṣṭhamātrā* character from characters proposed in the Devanagari block is mentioned, the change of "56" to "55" is not mentioned where the number of characters proposed is mentioned in the introduction in section 1 of p. 1 nor is the change of "25" to "24" where the number of characters in the proposed Vedic Extensions block is mentioned later in the same paragraph.

p. 2, sec. 1.1 As penultimate sentence before "While the names", added: "Tone marks are encoded as combining marks following the vowel they modify in canonical order and preceding visarga and nasal characters."

p. 3, sec. 2, end (after the characters) Added: "In order to accommodate the use of these characters with scripts other than those to which their names suggest they belong, we propose to change their script properties to "script=common".

This proposal is mentioned also on p. 10, sec. 10 for the characters:

0951
0952
0CF1
0CF2

p. 4, sec. 4.1, replaced "The proposed Sāmavedic characters and digits do not." by "The proposed Sāmavedic characters and digits constitute a minor closed subset of the Devanagari character set and do not carry the same significance as the independent characters they resemble."

p. 4 after "The following eighteen...block.", added: "These characters and digits follow the syllable they modify; in the canonical character sequence they will come subsequent to a vowel, or virāma but prior to a visarga or nasal character. In Sāmaveda recitation, final consonants, visarga, and certain nasals may be syllabified and hence carry their own tone markers just as full-fledged syllables with vowel characters do."

p. 4 paragraph 2, after "employed in superscript" added "or subscript", after "superscripted" added "or subscripted". At the end of the sentence added "(See Figure 4.1S.)" In the next sentence also after "superscripted" added "or subscripted".

p. 17 after Figure 4.1R, added "Figure 4.1S. Figures 4.1Sa and b are samples showing interlinear annotation of *Jaiminīya-Sāmagāna* using full character sequences below the line of text. The usage of character sequences to mark phrasal melody contrasts sharply with the use of superscript character diacritics per syllable to mark the tone dynamics of the syllable with which the diacritic is associated in the annotation of *Sāmagāna* in the Kāuthuma and Rāṇāyanīya traditions."

Figure 4.1Sa from file SVJaiShrautaP99v1.57.11.tif
Figure 4.1Sb from file SVJaiShrautaP99v1.57.11b.tif

p. 5, sec. 4.1., COMBINING DEVANAGARI DIGIT EIGHT Added: "Although the superscript digit ऌ is not used either in *Sāmavedasamhitā* or *Sāmagāna*, every other digit is used and hence is being proposed for encoding in the present document. If the digit ऌ is excluded, it will present an inelegant gap in the series of superscript digits. We propose to encode the character in order to complete the series."

p. 5, sec. 4.2, VEDIC TONE KARSHANA Added: "The karshana sign has a distinctive shape with a thick rising diagonal and thin descending diagonal that distinguish it from the generic circumflex [U0302]."

p. 5, sec. 4.2, VEDIC TONE SHARA Added: "The shara 'arrow' is a single combining sign, not a combination of a vertical line and caret, and not a spacing sign. It also has a distinctive shape with a thick rising diagonal and thin descending diagonal that distinguish it from generic arrows."

p. 6, sec. 4.2, VEDIC TONE PRENKHA Added: "The prenkha has a distinctive shape with angular ends that distinguish it from the generic macron [U0304]."

p. 6, sec. 5.1, VEDIC TONE DOT BELOW Added: "The dot is shaped like a diamond in Devanāgarī unlike the generic dot below [U0323]. Unlike the nukta, which appears left of center and higher, it is centered below the orthographic syllable that has a combining vowel character above, below, or to the left. If the syllable contains a combining vowel character *ā ī o au* to the right of the rightmost consonant sign in an orthographic syllable, it is centered below the orthographic syllable, the consonant sign, or the vowel character *ā ī o au* it modifies. The dot below cannot be merged with nukta because each canonical sequence consonant + nukta is equivalent to an Arabic character."

p. 19, Figure 5.1G Replaced the comma after *Samhitā* with a period and added, "Figure 5.1Ga shows". Changed the Figure number 5.1G to 5.1Ga. Added after "verse 16.104.6." the following: "In figure 5.1Gb of page 23 of Raghu Vira's edition, it is clear that the distinctive shape of the dot below is a diamond in contrast to the squarish dot representing the period between digits of verse numbers and the round dot used for ellipsis in the margin." Figure 5.1G is in file PSRaghuViraP23dotsdiamonds.tif.

p. 6, sec. 5.1, VEDIC TONE YAJURVEDIC AGGRAVATED INDEPENDENT SVARITA SCHROEDER changed "independent svarita (not aggravated) to "aggravated independent svarita".

Added: "The sign has a distinctive shape with a thick rising diagonal and thin descending diagonal that distinguish it from the generic circumflex below [U032D]."

p. 7, sec. 6 Added: "It is distinct from the avagraha [U093D] both in function and appearance. Unlike the avagraha, which is a spacing character, the sign is a combining mark because, like other accents, it is a modifier of the preceding vowel after which it occurs in canonical order. Whereas the upper right end of the avagraha meets the headbar, the ATHARVAVEDIC INDEPENDENT SVARITA rises above the headbar at the upper right and descends below the base line at the lower left."

p. 7, sec. 7 Changed "They combine with" to "They combine only with". After "in the text stream", added "Encoding these three accent signs as combining characters distinct from the visarga [0903] permits color differentiation in rich text to capture the traditional practice of writing accents in a colored ink different from the black ink employed for the base text. The sequence 0903 + 1CE1 encodes a *svarita visarga*, 0903 + 1CE2 encodes an *udātta visarga*, and 0903 + 1CE3 encodes an *anudātta visarga*."

p. 7, sec. 8, last sentence of introductory paragraph After "proposed", insert ", one for encoding in the Devanagari block and ten". After the end of the sentence, added: "Although the four characters DEVANAGARI SIGN DOUBLE CANDRABINDU VIRAMA, DEVANAGARI SIGN CANDRABINDU TWO, DEVANAGARI SIGN CANDRABINDU THREE, and DEVANAGARI SIGN CANDRABINDU AVAGRAHA combine

the candrabindu with a candrabindu virama, the digits २ or ३, and avagraha, we propose the encoding of precomposed characters because it avoids the difficulties that would arise from encoding these as sequences. Neither combining candrabindu nor spacing chandrabinu plus

candrabinu virama, २, ३, or avagraha provides the correct appearance if the font does not contain a precomposed glyph. In actual usage, a spacing candrabindu never occurs followed horizontally by these four signs. In the case of the double candrabindu virama, two candrabindus never occur side by side in ordinary Devanāgarī usage. Candrabindu does occur above a preceding *akṣara*

followed by २ or ३ or avagraha (figure 8Eb). To produce such placement would require utilizing the sequence: character + combining candrabindu + the digit or avagraha. To produce the

candrabinu above the २ or ३ or avagraha, however, would require reversing the sequence: २ or ३ or avagraha + combining candrabindu. But this sequence would leave the digits or avagraha at full size and the ends of the candra above the headbar. In practice, the ends of the candra appear at the

headbar when it occurs over an avagraha (figure 8G) and the digits २ and ३ appear reduced when the candra appears over them (figure 8Fa, 8Fc)."

p. 22, Figure 8E. Added the first image in the file ChandrabinduCombos2007July27.rtf as 8Eb and renumber 8E as 8Ea.

p. 22, Figure 8F, replaced Figure 8Fa with the similar image (2nd image in the file ChandrabinduCombos2007July27.rtf). Added the 3rd image in the file as 8Fc.

p. 7, sec. 8, DEVANAGARI SIGN INVERTED CANDRABINDU Replaced "Although proposed for the Devanagari block, it may be appropriate to name this vedic sign inverted candrabindu." by "Although used primarily in Devanāgarī, the sign is used to represent Vedic texts in other scripts as well; therefore, its script property = "common".

p. 29, removed INVERTED CANDRABINDU A8F2 and moved to Devanagari block 0900 (per Ken Whistler's suggestion).

p. 8, sec. 9, DEVANAGARI VOWEL SIGN PRISHTHAMATRA E. Moved from here to sec. 2, removed from the code table on p. 26, and from the names list on p. 27. Also in the reference to "6" or "six" additions to the devanagari block, changed "6" or "six" to "5" or "five". Similarly, in references to the total number of characters proposed, changed "56" or "fifty-six" to "55" or "fifty-five".

On p. 2 or 3, after sec. 2 Characters already encoded, added the following:

"2.1 **Pr̥ṣṭhamātrā orthography.** In order to produce the dependent vowel signs for /e/, /o/, /ai/ and /au/ in *pr̥ṣṭhamātrā* orthography, font designers are requested to implement the following variant renderings for the characters 0947 devanagari vowel sign e, 0948 devanagari vowel sign ai, 094B devanagari vowel sign o, and 094C devanagari vowel sign au.

The character DEVANAGARI VOWEL SIGNE renders with *pr̥ṣṭhamātrā* to the left of the consonant cluster the character follows.

The character DEVANAGARI VOWEL SIGN AI renders with *pr̥ṣṭhamātrā* to the left and glyph for 0947 devanagari vowel sign e above the consonant cluster the character follows.

The character DEVANAGARI VOWEL SIGN O renders with *pr̥ṣṭhamātrā* to the left and glyph for 093E devanagari vowel sign aa to the right of the consonant cluster the character follows.

The character DEVANAGARI VOWEL SIGN AU renders with *pr̥ṣṭhamātrā* to the left and glyph for 094B devanagari vowel sign o to the right of the consonant cluster the character follows.

[See Unicode 5, p. 651 Malayalam vowels 0D4A-0D4C. These characters have glyphs flanking the hollow dotted circle.]"

p. 24, fig. 9A. Replaced the paragraph with the following:

"Samples showing the *pr̥ṣṭhamātrā* system of writing the vowels *e*, *o*, *ai*, and *au* in Witzel manuscript 1250CE of the *Vājasaneyī-Saṁhitā*. Figure 9Aa illustrates the vowels *o* and *e* respectively in the sequence *ṣoḍaśine* equivalent to Devanāgarī षोडशिने. Figures 9Ab and 9Ac illustrate the vowel *ai* in the sequences *upaimi*, equivalent to Devanāgarī उपैमि, and *tasmāi*, equivalent to Devanāgarī तस्मै. Figure 9Ad illustrates the vowel *au* in the sequence *mahīdyauh*, equivalent to Devanāgarī महीद्यौः."

Changed "Figure 9Ab" to "Figure 9Ad". Added the image in file WitzelMssVS1.6Pai.tif as Figure 9Ab, and the image in file WitzelMssVS1.6Smai.tif as figure 9Ac.

p. 8, sec. 9, DEVANAGARI VOWEL SIGN CANDRA LONG E

p. 8, sec. 9, DEVANAGARI CARET Added: "The divider sign has a distinctive shape with a thin descending diagonal and thick rising diagonal that distinguish it from the generic caret U+2038. Ken Whistler suggested making this a spacing character, calling it "devanagari caret", that we explain that it is a "zero-width spacing character centered on the point". Note that it also differs from the generic caron U+02C7 which is used as a superscript centered above a character. Technical implementation: It should be placed so as not to crach into characters that occur above the headbar like raised *r*, *o*, *anusvāra*, etc.

p. 13, Figure 3. Called Figure 3 Figure 3a and added as Figure 3b the image in

RVKhSontakke-v1.12.7b.tif and circle or point to the umbrella handle on वृक्ष्ये in line 2. Added as Figure 3c the image in RVKhp48para3A.tif and point to or circle the umbrella handle following H√schken in line 1. Added as Figure 3d the image in RVKhp49examples.tif and point to or circle the umbrella handle over svar line 1 and over vyetu in line 2. Edit the text above the figures after "Figure 3." After small caps "svarita" put a period and insert before "in", the following: "Figures 3a and 3b show the character" and added after "1.11.4", "and 1.12.7. Figures 3c and 3d show Scheftelowitz's (1966: 48-49) description and examples of the character in her Roman edition of the Rgveda Khilani."

p. 13, Figure 4.1B. After "Figure 4.1Bc shows...number 11.", added: "The canonical character order to generate ऊ with ११ above will be the following: 090A + A8E1 + A8E1."

p. 13, Figure 4.1C. Added: "In figure 4.1Cb, the canonical character order to generate रा with र above will be the following: 0930 + 093E + A8E2 + A8EF. In figure 4.1Cc, the canonical character order to generate आ with र above will be the following: 0906 + A8E2 + A8F1."

p. 14, Figure 4.1D. Added: "In figure 4.1Db, the canonical character order to generate नू with र above will be the following: 0928 + 0942 + A8E3 + A8EF."

p. 14, Figure 4.1F. Added: "In figure 4.1Fb, the canonical character order to generate वा with र above will be the following: 0935 + 093E + A8E5 + A8EF."

p. 15, Figure 4.1L. Added: "In figure 4.1La, the canonical character order to generate च with र above followed by ना with र above will be the following: 091A + A83E + 0928 + 093E + A8E2 + A8EB."

p. 16, Figure 4.1M. Added: "In figure 4.1M, the canonical character order to generate त with र above will be the following: 0924 + A8E3 + A8EC, and to generate न्वा with र above will be: 0928 + 094D + 0935 + 093E + A8E2 + A8EF." Added a thin pencil line in 4.1M showing the alignment of र over the त.

p. 16, Figure 4.1N. Added: "Note: The fact that the typography has raised the न so that it does not crash into the यौ confirms that Sāmaveda superscripts are syllable specific annotations."

p. 16, Figure 4.1P. Added: "In figure 4.1Pb, the canonical character order to generate रा with र above will be the following: 0930 + 093E + A8E2 + A8EF. In figure 4.1Pc, the canonical character order to generate वा with र above will be the following: 0935 + 093E + A8E5 + A8F1."

p. 16, Figure 4.1Q. After small caps "vi", added a period and replaced the remaining text with "Figures 4.1Qa and b show the character in Shrouthi 1998: 182 and 184, and figures 4.1Qc and d show the full pages to demonstrate the context of the examples in *Sāmāgāna*." Replaced Figure 4.1Q and the heading with the following:

Figure 4.1Qa from the file SVprakftiP182ShrouthiVinataLinesP.tif
Figure 4.1Qb from the file SVprakftip184ShrouthiVinataLinesP.tif
Figure 4.1Qc from the file SVprakftiP182Shrouthi.tif
Figure 4.1Qd from the file SVprakftiP184Shrouthi.tif

p. 17, Deleted figure 4.1Rb and added figures 4.1Rb, c, and d from the following files in the folder: SVSamaSramiAvagrahas

4.1a is SamaSramiSVP622.tif

4.1b is SamaSramiSVP622Avagraha.tif

4.1c is SamaSramiSVP622AvagrahaWith2.tif

p. 17, Figure 4.1R. Replaced "Figure 4.1Rb shows it unconjoined in Samasrami's edition of the Samaveda, p. 620" with the following: Figure 4.1Rb-d show two examples of superscript avagraha unconjoined on p. 622 of Samaśrami's edition of *Sāmagāna*.

Added: "In figure 4.1Ra, the canonical character order to generate आ with २५ above will be the following: 0906 + A8E2 + A8F1. Figures 4.1Rb and 4.1Rc both show the avagraha over the syllable *hī* on lines 2 and 5 respectively of the page in Figure 4.1Rd. The canonical character order to generate the syllable will be the following: 0939 + 0940 + A8F1."

p. 17, Figure 4.2A. Added: "In figure 4.2Ab, the canonical character order to generate ते with २ and the *karṣaṇa* above will be the following: 0924 + 0947 + A8E2 + 1CD1. Note that the ५ above the धा precedes the visarga in canonical order."

p. 17, Figure 4.2B. Added: "The *śara* over the शू is a single arrow, not a sequence of vertical line and caret with which it is shown here due to inadequate typography. The *śara* precedes the २ in canonical order."

p. 17, Figure 4.2C. Added: "The prenkha character follows each syllable or other character over which it appears, in the following order: 0928 + 0943 + A8E2 + 092D + 093E + 1CD3 + 093D + 1CD3 + 0968 + 1CD3."