

Universal Multiple-Octet Coded Character Set  
 International Organization for Standardization  
 Organisation internationale de normalisation  
 Международная организация по стандартизации

**Doc Type: Working Group Document****Title: Proposal to encode 55 characters for Vedic Sanskrit in the BMP of the UCS****Source: Michael Everson and Peter Scharf (editors), Michel Angot, R. Chandrashekar, Malcolm Hyman, Susan Rosenfield, B. V. Venkatakrishna Sastry, Michael Witzel****Status: Individual Contribution****Action: For consideration by JTC1/SC2/WG2 and UTC****Replaces: N3290****Date: 2007-10-18**

**1. Introduction.** This document requests the addition to the UCS of 55 characters used chiefly in Vedic Sanskrit. Some of the characters are script-specific, but many are generic and are intended to be used with any script which conforms to the classic Brahmic script model. The document proposes the creation of a block for Vedic Extensions, a block for Devanāgarī Extensions, and the addition of several characters to the existing Devanāgarī block. The Vedic Extensions block will encode 24 Vedic characters generic to scripts that conform to the classic Brahmic script model. The Devanāgarī Extensions block will encode 25 Vedic characters specific to Devanāgarī. Six additions to the Devanāgarī block include two Vedic characters especially suited to inclusion in the Devanāgarī block, two Devanāgarī characters used for transcription of Avestan characters, and two signs common in Indic texts.

The document discusses the proposed additions in sections that group characters of related function. Sections 1.1 and 1.2 provide an overview of Vedic accentuation. Section 2 discusses the disposition of several Vedic characters already encoded in the Unicode standard. Sections 3-8 propose Vedic characters not presently encoded. Of these, sections 3-6 propose additional characters used for accentuation in each of the four major Vedic traditions, section 7 proposes characters related to visarga, and section 8 proposes nasal characters. Section 9 proposes the additions to the existing Devanāgarī block. Sections concerning character properties, bibliography, and figures follow, and tables showing the placement of the proposed additions are located at the end of the document.

**1.1. Tone in Vedic.** Indian linguists describe tone either as a feature of vowels, in which case it is shared by consonants in the same syllable, or directly as a feature of syllables. Vowels are marked for tone in Vedic as are certain non-vocalic characters that are syllabified in Vedic recitation (*visarga* and *anusvāra*). Vowels are categorized according to tone as *udātta* (high-toned or ‘acute’), *anudātta* (low-toned or ‘non-acute’), *svarita* (circumflexed or ‘modulated’), or *ekasruti* (monotone). A circumflexed vowel is generally described as dropping from high to low, and a series of syllables is monotone if devoid of relative distinction in tone.

Indian linguists describe a number of different types of svarita. A *dependent svarita* is one that results from the contextual raising of an anudātta and hence always follows an udātta. An *independent svarita*, which results from the lexical or post-lexical combination of an udātta vowel with a following anudātta vowel, is context-independent. An *aggravated independent svarita* is an independent svarita that is followed by an udātta or another independent svarita; its decline is steeper resulting in a lower tone at the end.

Due to tonal shift in the history of the language, various Vedic traditions differ concerning the surface tone that is recited for the underlying tone. In the common recension of *Ṛgveda*, for example, the last anudātta before an udātta is recited with low surface tone and the svarita has the highest surface tone.

Some of the same graphic symbols used for marking tone indicate different tones in different traditions. *Visarga* may be marked for all three tones, and *anusvāra* may be marked for high or low surface tone. Tone marks are encoded as combining marks following the vowel they modify in canonical order and preceding visarga and nasal characters. While the names given to the marks (both existing in the Unicode standard and hereunder proposed for addition to it) capture the usage in certain traditions, we describe basic parameters for the use of each character below and will detail further specifics in a technical note.

**1.2. Tone in the Sāmavedic tradition.** The Sāmavedic tradition is divided into three branches (Kauthuma, Rāṇāyanīya, Jaiminīya), each of them having its own way of naming, writing, and singing the texts. The signs vary also according to the manuscript traditions, the habits of the writers, and fonts available to printers. The Sāmaveda may be either recited or sung, with different systems of annotating each.

- a) **Recited.** The collection of the texts (*Sāmaveda-Saṁhitā*) is recited, like most of the Vedic Saṁhitā texts, with three tones (*svara*). The tones are marked with a digit, or letter, or digit with following letter, superscripted above the syllable being marked. *Udātta* (U), *svarita*, (S), and *anudātta* (A) are marked with <१>, <२> and <३> respectively. The letters <उ>, <र> and <क> are used for specific tonal sequences: <१>-< >-<२र> for the sequence U-U-S, <२उ>-< >-<३> for the sequence U-U-A and <३क>-<२र> for the sequence A-S (in which case S is an independant svarita). As in the other Vedic traditions, the tones that are not marked are inferred.
- b) **Sung.** When a *sāman* is sung in *Sāmagāna*, seven tones (*svara*) are used; they constitute a Sāmavedic scale. Six of them are indicated in the written and printed texts by digits from <१> to <६> in order from high to low; the seventh, and highest tone, is indicated in one of two ways, either by the numeral <१> or by the numeral <१ १>. If the seventh and highest tone is marked with the numeral <१> as is the first tone, the marking is ambiguous. The difference between them is usually inferable from the marking of a skip in descent on the subsequent syllable; in the few remaining cases, it is known by oral tradition.

The original text of the *Sāmaveda-Saṁhitā*, when it is sung, is also modified in different ways: shaking of the voice, prolongation of a vowel, modulation from one *svara* to another (with different cases of omission of one or several *svaras* of the scale), etc. All these modifications are marked with different characters: digits, avagraha, letters, other signs like the arrow, and so on. When a digit is used for different purposes in a particular tradition, one is superscript, the other not. In the annotational tradition of the Rāṇāyanīya school, syllables are added to the original text in line with the text and in the annotational tradition of the Jaiminīya school, syllables are added as superscripts in red above syllables in the original text. These signs are directly linked to the *mūdras* ‘hand-positions’ which, before the oral tradition was committed to writing, were the only means used to visually express musical motives on the text. The signs required to encode the superscript signs used in the Kauthuma tradition of annotation, and the one superscript sign used in the Rāṇāyanīya tradition are detailed below. (To encode characters used in the Jaiminīya tradition of annotation, nearly all of the characters of Grantha script, with proper combinatory mechanics for conjuncts, would have to be available in superscript.)

In combination, the combining digits and letters are displayed side by side, for example: क१ or क११. Ordinary digits may also bear diacritical marks, such as ३̣; we mention this because some current implementations may not permit such sequences, and they should.

**2. Characters already encoded.** Four characters already encoded in the UCS are intended to be used generically with any script which conforms to the classic Brahmic script model, despite the fact that they are encoded with script-specific names. The characters are:

U+0951 DEVANAGARI STRESS SIGN UDATTA would, if it were being encoded today, been named \*VEDIC TONE SVARITA, since that is its primary use. (Figure 2A)

- ◉ U+0952 DEVANAGARI STRESS SIGN ANUDATTA would, if it were being encoded today, been named \*VEDIC TONE ANUDATTA, since that is its primary use. (Figure 2B)
- ⌘ U+0CF1 KANNADA SIGN JIHVAMULIYA is used to mark *jihvāmūlīya* (a velar fricative [x] occurring only before unvoiced velar stops KA and KHA). (Figure 2C)
- U+0CF2 KANNADA SIGN UPADHMANIYA is used to mark *upadhmānīya* (which is a bilabial fricative [ɸ] occurring only before unvoiced labial stops PA or PHA). (Figure 2D)

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” (see §10).

**2.1** 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 U+0947 DEVANAGARI VOWEL SIGN E, U+0948 DEVANAGARI VOWEL SIGN AI, U+094B DEVANAGARI VOWEL SIGN O, and U+094C DEVANAGARI VOWEL SIGN AU:

- The character DEVANAGARI VOWEL SIGN E renders with *pr̥ṣṭhamātrā* to the left of the consonant cluster the character follows:  $\overline{\text{क}}$  is the *pr̥ṣṭhamātrā* rendering of कै *ke*.
- The character DEVANAGARI VOWEL SIGN AI renders with *pr̥ṣṭhamātrā* to the left and glyph for U+0947 DEVANAGARI VOWEL SIGN E above the consonant cluster the character follows:  $\overline{\text{के}}$  is the *pr̥ṣṭhamātrā* rendering of कै *kai*.
- The character DEVANAGARI VOWEL SIGN O renders with *pr̥ṣṭhamātrā* to the left and glyph for U+093E DEVANAGARI VOWEL SIGN AA to the right of the consonant cluster the character follows:  $\overline{\text{का}}$  is the *pr̥ṣṭhamātrā* rendering of को *ko*.
- The character DEVANAGARI VOWEL SIGN AU renders with *pr̥ṣṭhamātrā* to the left and glyph for U+093E DEVANAGARI VOWEL SIGN O to the right of the consonant cluster the character follows:  $\overline{\text{कौ}}$  is the *pr̥ṣṭhamātrā* rendering of कौ *kau*.

**3. Combining diacritic for the R̥gvedic tradition.** The following character is proposed for encoding in the Vedic Extensions block.

- ◉ VEDIC TONE RIGVEDIC KASHMIRI INDEPENDENT SVARITA is used to mark an independent svarita in the *R̥gveda Vāṣkala-Samhitā*. (Figure 3)

**4. Combining characters for the Sāmavedic tradition.** Howard (1986: 228-229) summarizes the significance of the digits <१>, <२> and <३>, and the characters <२>, <३> and <क> in the *Sāmaveda-Samhitā* as follows:

Numbers 1 and 3 always represent udātta and anudātta, respectively. Number 2 indicates svarita, but it denotes also an udātta syllable followed by anudātta. When two or more udātta syllables appear in succession, only the first is marked with 1, but the sign 2r is placed above the following svarita. If, however, an anudātta follows, 2u is placed above the first udātta syllable and the rest are left undesignated. In a series of anudātta syllables at the beginning of the line, only the first is marked with 3. An independent svarita has the sign 2r, and the preceding anudātta is marked 3k. Pracaya syllables have no markings.

He (1977: 120) summarizes the principle of the annotation in Sāmagāna as follows:

Each chant consists of a certain number of standard phrases, part of a repertoire of melodic fragments constituting all of the musical material belonging to a certain style of singing. These phrases recur over and over again, in various patterns, to form the several thousand sāmans. This recurrence of

melodic formulae is without doubt the *raison d'être* of the division into parvans, each of which corresponds to a specific musical phrase or motive. A melody-type is symbolized in the gānas by a particular syllable (in the case of the Rāṇāyanīyas), a certain sequence of numerals (in the case of the Kauthumas), or a specific sequence of syllables (in the case of the Jaiminīyas). In the latter two cases it is not the individual numeral or syllable which symbolizes always a specific melody-type; rather it is the arrangement of the numerals or syllables within a parvan which determines its musical content. This technique of patchwork composition (centonization) is characteristic also of the ancient Hebrew chant and some of the oldest Gregorian chants, the Tracts.

The following is based in part on Howard's (1977: 79-81) presentation of the details of the significance of specific marks in his tables 5–6.

**4.1. Combining digits and letters for the Sāmavedic tradition.** This section proposes 18 characters for encoding in the Devanāgarī Extensions block. The combining digits and alphabetic characters proposed in this section are smaller versions of the corresponding ordinary Devanāgarī digits and characters placed directly above individual characters with which they are associated as diacritics. They are analogous to the medieval superscript letter diacritics encoded at U+0363 and in the Combining Diacritical Marks Supplement block at U+1DC0..U+1DFF.

We have examined the possibility of representing these digits and characters using Ruby annotation (Encoding Sāmaveda with Ruby, Sanskrit Library Technical Note 1, <http://sanskritlibrary.org/VedicUnicode/SLTN1.pdf>). Although it would be possible to achieve the desired appearance using Ruby annotation, this would not support the characters as they are actually used. The characters in question are utilized in *Sāmaveda* as diacritics at the character level, not at a higher-level. They are associated with individual base characters as diacritics in precisely the same manner that other diacritic marks are associated with individual base characters. In the same way that a COMBINING MACRON is encoded separately even though it resembles an EN DASH, or that a COMBINING DOT ABOVE is encoded separately even though it resembles a FULL STOP, the proposed combining digits and characters ought to be encoded separately, even though they resemble ordinary digits and characters. The appearance of the COMBINING MACRON and COMBINING DOT ABOVE could equally well be achieved using Ruby just as much as the appearance of the Sāmavedic digits and characters could be. Ruby is an interlineation procedure that associates parallel independent character encodings each of which continues to carry its independent significance unchanged when stacked. 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. The fact that shapes similar to ordinary characters were chosen by the originators of the notation system as diacritics is irrelevant. Used as diacritics and placed above the characters they modify, they do not carry the same significance as the ordinary characters they resemble. Due to their unique placement and their function as diacritic marks at the character level, they are unique characters, and it is our strong opinion that it is best to encode them as such.

The superscript characters and digits proposed concern the Kauthuma and Rāṇāyanīya traditions of Sāmaveda and Sāmagāna. The situation in the Jaiminīya tradition is quite different and the Jaiminīya tradition *is* suitable for representation using Ruby. Unlike the other two traditions where superscripted characters are used as diacritics proper to the character over which they are written, in the Jaiminīya tradition the full range of characters that appears in the script is employed in superscript or subscript and the superscripted or subscripted character sequence names a figure that applies to a textual phrase rather than an individual character. (See Figure 4.1S.) Since the Jaiminīya tradition does not employ the superscripted or subscript characters as character diacritics, we acknowledge that Ruby *is* the appropriate means to represent the Jaiminīya system of Sāmaveda annotation.

We propose that only Devanāgarī combining digits 0-9 and alphabetic characters used in the Kauthuma and Rāṇāyanīya traditions of *Sāmaveda* be added to the BMP, although there may be parallels in other scripts. It is true that all of combining digits 0-9 and alphabetic characters proposed below in this section

would have the same significance for Sāmaveda in other scripts as they do in Devanāgarī. Moreover, editions of *Sāmaveda* are known to have been published in several other major Indic scripts in the past. However, searches for publications of Sāmaveda in scripts other than Devanāgarī at major Indic book dealers in Delhi, Varanasi, and Bangalore proved unsuccessful. The representation of *Sāmaveda* in other major Indic scripts may be desired by teachers, students, and practitioners of Vedic recitation, and they may request Vedic extensions of this kind. But due to the rarity of Sāmaveda publications in scripts other than Devanāgarī and the narrower circle of usage of such publications, it may be appropriate that future proposals for parallel extension blocks for Vedic characters in Roman or Indic scripts other than Devanāgarī be added to the SMP rather than the BMP. The representation of Sāmaveda in Devanāgarī, on the other hand, has considerably wider usage. Devanāgarī has become standard for the publication of Sanskrit texts both within and outside of India and is regularly taught to non-Indians learning Sanskrit. It is therefore deemed appropriate to add a Devanāgarī extension block to the BMP.

The following eighteen characters are proposed for encoding in the Devanagari Extensions block. 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.

- ० ○ COMBINING DEVANAGARI DIGIT ZERO is used to mark a long vowel that is not augmented (*vr̥ddha*) in the Rāṇāyanīya tradition of Sāmagāna. (Figure 4.1A)
- १ ○ COMBINING DEVANAGARI DIGIT ONE is used to mark an udātta in *Sāmaveda-Saṁhitā*, and, in Sāmagāna, to mark the first tone (*prathama*) or seventh tone (*kr̥ṣṭa*), or, written as a superscript over a numeral in line in the text (which indicates a secondary tone), to indicate that the tone signified by the numeral is held for one mora. (Figure 4.1B)
- २ ○ COMBINING DEVANAGARI DIGIT TWO is used to mark an independent svarita, or, it occurs followed by an <३> over the first of two udātta vowels followed by an anudātta in *Sāmaveda-Saṁhitā*. In Sāmagāna it is used to mark the second tone (*dviṭīya*). (Figure 4.1C)
- ३ ○ COMBINING DEVANAGARI DIGIT THREE is used to mark an anudātta in *Sāmaveda-Saṁhitā*, and the third tone (*tr̥ṭīya*) in Sāmagāna. It may be followed by superscript <४>. (Figure 4.1D)
- ४ ○ COMBINING DEVANAGARI DIGIT FOUR is used to mark the fourth tone (*catūrtha*) in Sāmagāna. (Figure 4.1E)
- ५ ○ COMBINING DEVANAGARI DIGIT FIVE is used to mark the fifth tone (*mandra* or *pañcama*) in Sāmagāna. (Figure 4.1F)
- ६ ○ COMBINING DEVANAGARI DIGIT SIX is used to mark an atisvarya tone in Sāmagāna. (Figure 4.1G)
- ७ ○ COMBINING DEVANAGARI DIGIT SEVEN is used to mark brief recitation (*abhiḡīta*) in Sāmagāna. (Figure 4.1H)
- ८ ○ COMBINING DEVANAGARI DIGIT EIGHT has not been found in Sāmagāna, but we propose it here for the sake of completing the logical set. Although the superscript digit ८ is not used either in *Sāmaveda-saṁhitā* or *Sāmagāna*, every other digit is used and hence is being proposed for encoding in the present document. If the digit ८ 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.
- ९ ○ COMBINING DEVANAGARI DIGIT NINE is used to indicate bending or sinking (*namana*) in Sāmagāna. (Figure 4.1J)
- अ ○ COMBINING DEVANAGARI LETTER A is used to mark brief recitation (*abhiḡīta*) in Sāmagāna. (Figure 4.1K)
- उ ○ COMBINING DEVANAGARI LETTER U is used to mark an udātta in Böhtlingk and Roth's St. Petersburg Sanskrit-English lexicon, and, following a superscript <२>, to indicate the first of two successive udāttas followed by an anudātta in *Sāmaveda-Saṁhitā*. (Figure 4.1L)
- क ○ COMBINING DEVANAGARI LETTER KA is used, after a superscript <३>, to mark an anudātta preceding an independent svarita in *Sāmaveda-Saṁhitā*. (Figure 4.1M)

- ◉ ऩ COMBINING DEVANAGARI LETTER NA is used in South Indian manuscripts to mark bending or sinking (*namana*) in Sāmagāna. (Figure 4.1N)
- ◉ ड COMBINING DEVANAGARI LETTER PA is used in Sāmagāna instead of <ॡ> as a superscript over a numeral (a numeral in-line in the text indicates a secondary tone) to mark vibrato (*preṅkha*). (Figure 4.1O)
- ◉ ढ COMBINING DEVANAGARI LETTER RA is used, following a superscript 2, to mark an independent svarita in *Sāmaveda-Saṃhitā*, and in Sāmagāna, alone or following a superscript numeral 1-5, to mark a long (*dīrgha*) vowel that is not augmented (*vṛddha*). (Figure 4.1P)
- ◉ ढि COMBINING DEVANAGARI LETTER VI is used to mark a musical motive called *vinata* in Sāmagāna. Although the character resembles U+0935 DEVANAGARI LETTER VA combined with U+093F DEVANAGARI VOWEL SIGN I, it is used as a unitary diacritic in Sāmaveda. (Figure 4.1Q)
- ◉ ढऱ COMBINING DEVANAGARI SIGN AVAGRAHA is used in Sāmagāna to mark the omission or skipping of a tone in a descending scale (*atikrama*), or the musical motive called *vinata*, or bending or sinking (*namana*). (Figure 4.1R)

**4.2. Combining diacritics for the Sāmavedic tradition.** The following four characters are proposed for encoding in the Vedic Extensions block.

- ◉ ^ VEDIC TONE KARSHANA is used in Sāmagāna, as a superscript over a numeral in line in the text (which indicates a secondary tone), to indicate continuous progression or slide (*karṣaṇa*) of the tone signified by the numeral, or over a syllable to indicate bending or sinking (*namana*), or occasionally the musical motive involving descent from a primary second tone to a secondary third tone (*praṇata*). The karshana sign has a distinctive shape with a thick rising diagonal and thin descending diagonal that distinguish it from the generic circumflex U+0302. It is related in shape to the next character discussed here, VEDIC TONE SHARA. (Figure 4.2A)
- ◉ ↑ VEDIC TONE SHARA is used in Sāmagāna to mark skipping (*atikrama*), usually (in 52/56 instances appearing above an in-line 2 after a superscript 1) from *kruṣṭa* to *dviṭīya*. 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. (Figure 4.2B)
- ◉ \_ VEDIC TONE PRENKHA is a horizontal line used in Sāmagāna as a superscript over a character or numeral (a numeral in-line in the text indicates a secondary tone) to mark vibrato (*preṅkha*): ̄. The PRENKHA has a distinctive shape with angular ends that distinguish it from the generic macron U+0304. (Figure 4.2C)
- ◉ // VEDIC SIGN NIHSHVASA is a spacing character used to indicate to the performer where a breath can be conveniently taken. (Figure 4.2D)

## 5. Combining diacritics for the Yajurvedic tradition.

**5.1. General.** The following nine characters are proposed for encoding in the Vedic Extensions block.

- ◉ ˆ VEDIC TONE YAJURVEDIC INDEPENDENT SVARITA is used to mark an independent svarita (not aggravated) following an anudatta in the *Śuklayajurveda Mādhyandina-Saṃhitā*, and in the *Atharvaveda Paippalāda-Saṃhitā*. (Figure 5.1A)
- ◉ ˆ VEDIC TONE YAJURVEDIC KATHAKA INDEPENDENT SVARITA is used to mark an independent svarita (not aggravated) in the *Kṛṣṇayajurveda Kāṭhaka-Saṃhitā*. (Figure 5.1B)
- ◉ ˆ VEDIC TONE YAJURVEDIC AGGRAVATED INDEPENDENT SVARITA is used to mark an aggravated independent svarita the *Śuklayajurveda Mādhyandina-Saṃhitā* and in the *Kṛṣṇayajurveda Kāṭhaka-Saṃhitā*. (Figure 5.1C)

- ◉ VEDIC TONE CANDRA BELOW is used to mark an independent svarita (not aggravated) in the *Kṛṣṇayajurveda Kāṭhaka-Saṃhitā*, and an independent svarita (not aggravated) followed by an anudātta or ekaśruti in the *Kṛṣṇayajurveda Maitrāyaṇī-Saṃhitā*. It is also used instead of DEVANAGARI STRESS SIGN ANUDATTA to indicate low surface tone in *Śatapathabrāhmaṇa*. (Figure 5.1D)
- || VEDIC TONE DOUBLE SVARITA is used to mark a long (*dīrgha*) svarita. (Figure 5.1E)
- ≡ VEDIC TONE TRIPLE SVARITA is used to mark a dependent svarita followed by an anudātta in *Kṛṣṇayajurveda Maitrāyaṇī-Saṃhitā*. (Figure 5.1F)
- ◉ VEDIC TONE DOT BELOW is used to mark a dependent svarita in *Yajurveda Kāṭhaka-Saṃhitā* and *Atharvaveda Paippalāda-Saṃhitā*, and also to mark the first ekaśruti after an independent svarita in the latter. The dot is shaped like a diamond in Devanāgarī unlike the generic dot below U+0323. 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. (Figure 5.1G)
- ◉ VEDIC TONE KATHAKA ANUDATTA BELOW is used to mark an anudātta in *Yajurveda Kāṭhaka-Saṃhitā* and *Atharvaveda Paippalāda-Saṃhitā*. (Figure 5.1H)
- ◉ VEDIC TONE YAJURVEDIC AGGRAVATED INDEPENDENT SVARITA SCHROEDER is used to mark aggravated independent svarita in Schröder's edition of the *Kṛṣṇayajurveda Kāṭhaka-Saṃhitā*. The sign has a distinctive shape with a thick rising diagonal and thin descending diagonal that distinguish it from the generic circumflex below U+032D. (Figure 5.1I)

**5.2 Śatapathabrāhmaṇa.** The following two characters are proposed for encoding in the Vedic Extensions block.

- ◉ VEDIC TONE THREE DOTS BELOW is used to mark a surface low pitch corresponding to an underlying pre-pause udātta, i.e. one that occurs immediately before a pause or mediated by a single syllable before a pause, or followed by an udātta after the pause in Weber's edition of the *Śatapathabrāhmaṇa*. Doubled stacked, it is followed by a svarita after the pause. (Figure 5.2A)
- ◉ VEDIC TONE TWO DOTS BELOW is used to mark a surface low pitch corresponding to an underlying pre-pause udātta, i.e. one that occurs immediately before a pause or mediated by a single syllable before a pause, followed by an udātta or independent svarita after the pause; an (immediately) pre-pause anudātta, followed by an independent svarita after the pause in the *Śatapathabrāhmaṇa*. (Figure 5.2B)

**6. Combining diacritic for the Atharvavedic tradition.** The following character is proposed for encoding in the Vedic Extensions block.

- ◉ S VEDIC TONE ATHARVAVEDIC INDEPENDENT SVARITA is used to mark an independent svarita (not aggravated) in the *Atharvaveda Śaunakīya-Saṃhitā*. It is distinct from the avagraha U+093D 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 SVARATA rises above the headbar at the upper right and descends below the base line at the lower left. (Figure 6)

**7. Ardhavisarga and combining diacritics for visarga.** The following four characters are proposed for encoding in the Vedic Extensions block. The first three are tone markers that appear in red in Vedic manuscripts, just as other tone markers do. They combine only with the VISARGA, following it in the text stream. Encoding these three accent signs as combining characters distinct from the visarga U+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 ॐ 0903 + ॐ 1CE1 encodes a *svarita visarga*, ॐ 0903 + ॐ 1CE2 encodes an *udātta visarga*, and ॐ 0903 + ॐ 1CE3 encodes an *anudātta visarga*. VEDIC TONE VISARGA UDATTA and VEDIC TONE ANUDATTA VISARGA sometimes appear together combined on a VISARGA in final position, thus: ॐ 0903 + ॐ 1CE2 + ॐ 1CE3 (Figure 8Kb).

- ॐ VEDIC TONE VISARGA SVARITA is used to show that a visarga has a svarita tone. (Figure 7A)
- ॐ VEDIC TONE VISARGA UDATTA is used to show that a visarga has an udātta tone. (Figure 7B)
- ॐ VEDIC TONE VISARGA ANUDATTA is used to show that a visarga has an anudātta or pracaya tone. (Figure 7C)
- ॐ VEDIC TONE ARDHAVISARGA is used to mark either *jihvāmūlīya* (which is a velar fricative [x] occurring only before unvoiced velar stops KA or KHA) or *upadhmānīya* (which is a bilabial fricative [ɸ] occurring only before unvoiced labial stops PA or PHA). (Figure 7D)

**8. Nasals.** Indian phonetic treatises describe a number of phonetic distinctions in the articulation of nasals. First they distinguish between nasalized vowels, nasalized semivowels, nasal stops, and anusvāra. Ancient Vedic treatises (Prātiśākhya) describe the nasalization of vowels; nasalized semivowels *y*, *v*, and *l*; and two lengths of anusvāra: short (*hrasva*) and long (*dīrgha*). Long anusvāra occurs after short vowels, and short anusvāra occurs after long vowels. In addition to short and long anusvāra, medieval phonetic texts (Śikṣā) describe a heavy (*guru*) anusvāra, and a two-mora (*dvimātra*) anusvāra, and one treatise describes a prolonged (*pluta*) anusvāra. The heavy anusvāra occurs before a conjunct consonant, and the guru anusvāra occurs before a consonant followed by vocalic *r*. The *Pratijñāsūtra* prescribes that *gm* occurs in place of anusvāra before *r* or a spirant and has a three-fold distinction: short (after a long vowel), long (after a short vowel), and heavy (before a conjunct). Most Śikṣās give the name *raṅga* to a two-mora vowel with modulation of tone (*kampa*) in the middle and nasalization at the end. The *Mallaśarmakṛtā Śikṣā* describes several distinctions in the length of nasalized vowels, ranging from one to six mora. Those of four, five and six mora are called *raṅga*, *mahāraṅga*, and *atiraṅga*, respectively, and are followed by a pause in recitation. Different traditions mark varieties of nasals differently using the symbols below and others. The following eleven characters are proposed, one for encoding in the Devanagari block and ten for encoding in the Vedic Extensions block.

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 candrabindu plus candrabindu 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 candrabindu 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).

- ◌̣ DEVANAGARI SIGN INVERTED CANDRABINDU is used to mark *anusvāra* before spirants in Schröder’s edition of the *Kṛṣṇayajurveda Kāṭhaka-Saṃhitā*. (Figure 8A). Although used primarily in Devanāgarī, the sign is used to represent Vedic texts in other scripts as well; therefore, its script property = “common”.
- ◌̣ DEVANAGARI SIGN SPACING CANDRABINDU is a spacing mark used to mark *anusvāra*. It is lower than U+0910 DEVANAGARI SIGN CANDRABINDU and occurs in-line at the level of the Devanagari headbar. (Figure 8B)
- ◌̣ DEVANAGARI SIGN CANDRABINDU VIRAMA is used to mark *anusvāra*. (Figure 8C)
- ◌̣ DEVANAGARI SIGN DOUBLE CANDRABINDU VIRAMA is used to mark *anusvāra* before a spirant initial in a consonant cluster. (Figure 8D)
- ◌̣ DEVANAGARI SIGN CANDRABINDU TWO is used to mark a vowel prolonged to two mora with nasalization. (Figure 8E)
- ◌̣ DEVANAGARI SIGN CANDRABINDU THREE is used to mark a vowel prolonged to three mora with nasalization. (Figure 8F)
- ◌̣ DEVANAGARI SIGN CANDRABINDU AVAGRAHA is used to mark *anusvāra*. (Figure 8G)
- ◌̣ VEDIC SIGN ANTARGOMUKHA is used, with a bindu added on top, to mark short *anusvāra* after a long vowel. (Figure 8H)
- ◌̣ VEDIC SIGN BAHIRGOMUKHA is used, with a bindu or candrabindu added on top, to mark *anusvāra* or nasalization. (Figure 8I)
- ◌̣ VEDIC SIGN SAJIHVA BAHIRGOMUKHA is used, with a bindu or candrabindu added on top, to mark *anusvāra* or nasalization. (Figure 8J)
- ◌̣ VEDIC SIGN LONG ANUSVARA is used to mark a long *anusvāra* after a short vowel. (Figure 8K)

NOTE: Several of the characters above could be considered to be sequences of other characters. The glyphs for the CANDRABINDU characters are all aligned equally, at the height of the headline (and not above it) as shown next to the *ka* as follows: क̣ २̣ ३̣ २̣̣ ३̣̣ ३̣̣̣. There is no question that the first two of these here need to be encoded as spacing characters, but one might argue that the last four could be composed with a base character and COMBINING CANDRABINDU. An argument (which seems to us to be particularly strong) against this would be the typical glyph representation of such sequences (all shown here at 14 points): क̣ क̣̣ क̣̣̣ क̣̣̣̣ क̣̣̣̣̣ क̣̣̣̣̣̣ क̣̣̣̣̣̣̣ क̣̣̣̣̣̣̣̣ क̣̣̣̣̣̣̣̣̣ क̣̣̣̣̣̣̣̣̣̣. Note how the true AVAGRAHA connects with the KA while the CANDRABINDU AVAGRAHA does not. We prefer the unique encodings.

**9. Additions for Devanāgarī.** The following five characters are proposed as additions to the existing Devanāgarī block.

- ◌̣ DEVANAGARI VOWEL SIGN CANDRA LONG E is used in Devanagari transcriptions of Avestan to mark the long schwa ē. (DEVANAGARI VOWEL SIGN CANDRA E is used to mark the regular schwa ə.) (Figure 9B)
- ◌̣ DEVANAGARI SIGN PUSHPIKA is used as a placeholder or “filler”, often flanked by double dandas (Figure 9C)
- ◌̣ DEVANAGARI CARET is used to mark the insertion point of omitted text and to mark word division. 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. It is a zero-width spacing character centered on the point: ̣ which is used between orthographic syllables: क̣ो̣को̣ *koko*. (Figure 9D)
- ◌̣ DEVANAGARI LETTER ZHA is used in Devanagari transcriptions of Avestan to mark the voiced palatal fricative [ʒ]. (Figure 9E)

य DEVANAGARI LETTER HEAVY YA is used to mark an affricated glide [dʒ], as in कुर्यात् *kuryāt* ‘one would do’, written in other dialects कुर्यात् *kuryāt*. The distinction is similar to that made in Bengali orthography between য DEVANAGARI LETTER YA [dʒ] and ঞ BENGALI LETTER YYA [j]. (Figure 9F)

## 10. Unicode Character Properties. Character properties are proposed here.

094E;DEVANAGARI VOWEL SIGN PRISHTHAMATRA E;Mc;0;L;;;;;N;;;;;  
0955;DEVANAGARI VOWEL SIGN CANDRA LONG E;Mn;0;NSM;;;;;N;;;;;  
0973;DEVANAGARI SIGN PUSHPIKA;Po;0;L;;;;;N;;;;;  
0974;DEVANAGARI SIGN DIVIDER;Mn;0;NSM;;;;;N;;;;;  
0979;DEVANAGARI LETTER ZHA;Lo;0;L;;;;;N;;;;;  
097A;DEVANAGARI LETTER HEAVY YA;Lo;0;L;;;;;N;;;;;

1CD0;VEDIC TONE NIHSHVASA;Mn;230;NSM;;;;;N;;;;;  
1CD1;VEDIC TONE KARSHANA;Mn;230;NSM;;;;;N;;;;;  
1CD2;VEDIC TONE SHARA;Mn;230;NSM;;;;;N;;;;;  
1CD3;VEDIC TONE PRENKHA;Mn;230;NSM;;;;;N;;;;;  
1CD4;VEDIC TONE DOUBLE SVARITA;Mn;230;NSM;;;;;N;;;;;  
1CD5;VEDIC TONE TRIPLE SVARITA;Mn;230;NSM;;;;;N;;;;;  
1CD6;VEDIC TONE YAJURVEDIC INDEPENDENT SVARITA;Mn;220;NSM;;;;;N;;;;;  
1CD7;VEDIC TONE YAJURVEDIC AGGRAVATED INDEPENDENT SVARITA;Mn;220;NSM;;;;;N;;;;;  
1CD8;VEDIC TONE YAJURVEDIC KATHAKA INDEPENDENT SVARITA;Mn;220;NSM;;;;;N;;;;;  
1CD9;VEDIC TONE YAJURVEDIC KATHAKA INDEPENDENT SVARITA SCHROEDER;Mn;220;NSM;;;;;N;;;;;  
1CDA;VEDIC TONE CANDRA BELOW;Mn;220;NSM;;;;;N;;;;;  
1CDB;VEDIC TONE ATHARVAVEDIC INDEPENDENT SVARITA;Mc;0;L;;;;;N;;;;;  
1CDC;VEDIC TONE RIGVEDIC KASHMIRI INDEPENDENT SVARITA;Mn;230;NSM;;;;;N;;;;;  
1CDD;VEDIC TONE THREE DOTS BELOW;Mn;220;NSM;;;;;N;;;;;  
1CDE;VEDIC TONE TWO DOTS BELOW;Mn;220;NSM;;;;;N;;;;;  
1CDF;VEDIC TONE DOT BELOW;Mn;220;NSM;;;;;N;;;;;  
1CE0;VEDIC TONE KATHAKA ANUDATTA;Mn;220;NSM;;;;;N;;;;;  
1CE1;VEDIC TONE SVARITA VISARGA;Mn;1;NSM;;;;;N;;;;;  
1CE2;VEDIC TONE UDATTA VISARGA;Mn;1;NSM;;;;;N;;;;;  
1CE3;VEDIC TONE ANUDATTA VISARGA;Mn;1;NSM;;;;;N;;;;;  
1CE4;VEDIC SIGN ARDHAVISARGA;Lo;0;L;;;;;N;;;;;  
1CE5;VEDIC SIGN ANTARGOMUKHA;Lo;0;L;;;;;N;;;;;  
1CE6;VEDIC SIGN BAHIRGOMUKHA;Lo;0;L;;;;;N;;;;;  
1CE7;VEDIC SIGN SAJIHVA BAHIRGOMUKHA;Lo;0;L;;;;;N;;;;;  
1CE8;VEDIC SIGN LONG ANUSVARA;Lo;0;L;;;;;N;;;;;

A8E0;COMBINING DEVANAGARI DIGIT ZERO;Mn;230;NSM;;;;;N;;;;;  
A8E1;COMBINING DEVANAGARI DIGIT ONE;Mn;230;NSM;;;;;N;;;;;  
A8E2;COMBINING DEVANAGARI DIGIT TWO;Mn;230;NSM;;;;;N;;;;;  
A8E3;COMBINING DEVANAGARI DIGIT THREE;Mn;230;NSM;;;;;N;;;;;  
A8E4;COMBINING DEVANAGARI DIGIT FOUR;Mn;230;NSM;;;;;N;;;;;  
A8E5;COMBINING DEVANAGARI DIGIT FIVE;Mn;230;NSM;;;;;N;;;;;  
A8E6;COMBINING DEVANAGARI DIGIT SIX;Mn;230;NSM;;;;;N;;;;;  
A8E7;COMBINING DEVANAGARI DIGIT SEVEN;Mn;230;NSM;;;;;N;;;;;  
A8E8;COMBINING DEVANAGARI DIGIT EIGHT;Mn;230;NSM;;;;;N;;;;;  
A8E9;COMBINING DEVANAGARI DIGIT NINE;Mn;230;NSM;;;;;N;;;;;  
A8EA;COMBINING DEVANAGARI LETTER A;Mn;230;NSM;;;;;N;;;;;  
A8EB;COMBINING DEVANAGARI LETTER U;Mn;230;NSM;;;;;N;;;;;  
A8EC;COMBINING DEVANAGARI LETTER KA;Mn;230;NSM;;;;;N;;;;;  
A8ED;COMBINING DEVANAGARI LETTER NA;Mn;230;NSM;;;;;N;;;;;  
A8EE;COMBINING DEVANAGARI LETTER PA;Mn;230;NSM;;;;;N;;;;;  
A8EF;COMBINING DEVANAGARI LETTER RA;Mn;230;NSM;;;;;N;;;;;  
A8F0;COMBINING DEVANAGARI LETTER VI;Mn;230;NSM;;;;;N;;;;;  
A8F1;COMBINING DEVANAGARI SIGN AVAGRAHA;Mn;230;NSM;;;;;N;;;;;  
A8F2;DEVANAGARI SIGN INVERTED CANDRABINDU;Mn;0;NSM;;;;;N;;;;;  
A8F3;DEVANAGARI SIGN SPACING CANDRABINDU;Lo;0;L;;;;;N;;;;;  
A8F4;DEVANAGARI SIGN CANDRABINDU VIRAMA;Lo;0;L;;;;;N;;;;;  
A8F5;DEVANAGARI SIGN DOUBLE CANDRABINDU VIRAMA;Lo;0;L;;;;;N;;;;;  
A8F6;DEVANAGARI SIGN CANDRABINDU TWO;Lo;0;L;;;;;N;;;;;  
A8F7;DEVANAGARI SIGN CANDRABINDU THREE;Lo;0;L;;;;;N;;;;;  
A8F8;DEVANAGARI SIGN CANDRABINDU AVAGRAHA;Lo;0;L;;;;;N;;;;;

The following four characters are proposed to have their script property changed to “Common”:

0951;DEVANAGARI STRESS SIGN UDATTA;Mn;230;NSM;;;;;N;;;;;  
0952;DEVANAGARI STRESS SIGN ANUDATTA;Mn;220;NSM;;;;;N;;;;;  
0CF1;KANNADA SIGN JIHMULIYA;So;0;ON;;;;;N;;;;;  
0CF2;KANNADA SIGN UPADHMUNIYA;So;0;ON;;;;;N;;;;;

## 11. Bibliography.

- Bhaṭṭācārya, Satyavrata Sāmaśramin. (1871-78). *Sāmavedasamhitā*. 5 volumes. Calcutta: Asiatic Society.  
Reprinted: 1983, Delhi: Munshiram Manoharlal.
- Böhtlingk, Otto and Rudolph Roth. (1852-1855). *Sanskrit-Wörterbuch*. 7 vols. St. Petersburg: Kaiserlichen Akademie der Wissenschaften. Reprint: Motilal Banarsidass Publishers Private Ltd. Delhi (1990).

- Deshpande, Madhav M. (1997). *Śaunakīya Caturādhyāyikā: A Prātiśākhya of the Śaunakīya Atharvaveda: With the commentaries Caturādhyāyībhāṣya, Bhārgava-Bhāskara-Vṛtti and Pañcasandhi: Critically edited, translated & annotated*. HOS 52. Cambridge, Mass.: Department of Sanskrit and Indian Studies, Harvard University. Distributed by Harvard University Press.
- Dreyfuss, Henry. 1984. *Symbol sourcebook: an authoritative guide to international graphic symbols*. New York: John Wiley & Sons. ISBN 0-471-28872-1
- Gauḍa, Daulatarāma Śāstrī. (1965). *Śukla Yajur Veda: Vājasaneyisamhitā*. Vidyābhavana Saṁskṛta Granthamālā 129. Varanasi: Chowkhambā Vidyābhavana.
- Gauḍa, Daulatarāma Śāstrī. (n.d.). *Śuklayajurveda Samhitā*. Banaras: Babu Thakur Prasad Gupta.
- Ghosal, S. N. (1964). *Vajasaneyi Pratisakhya. Part I, Text with Translation & Critical Notes: With the English Translation of A. Weber's Introduction to the Text*. Indian Studies Past & Present. Calcutta: Quality Printers & Binders. [Chapters 1-2.]
- Gosvāmī, Bijanabihārī. (1977). *Yajurveda Samhitā [Śukla & Kṛṣṇa]*. Calcutta.
- Goswami, Sitanath & Himansu Narayan Chakravarti. (1974). *Ṛk-samhitā*. Sanskrit Pustak Bhandar, Bidhan Sarani, Calcutta - 6.
- Howard, Wayne. (1986). *Veda recitation in Varanasi*. Motilal Banarsidass, Delhi.
- Howard, Wayne. (1988). *Mātrālakṣaṇam: text, translation, extracts from the commentary, and notes, including references to two oral traditions of south India*. Motilal Banarsidass, Delhi.
- Howard, Wayne. (1988). *The decipherment of Samavedic notation of the Jaiminiyas*. Finnish Oriental Society, Helsinki, Finland.
- Howard, Wayne. (1977). *Sāmavedic Chant*. Yale University Press, New Haven.
- Jois, Gomatham Ramanuja (ed.) (n.d.). *Yajurāraṇyaka*. [in Telugu script.] Mysore.
- Jois, Gomatham Ramanuja, ed. (1902). *Yajurveda Samhitā*. [in Telugu script.] Mysore.
- Kashikar, K. G, ed. (1994). *Śrautakoṣa: Based on the Samhitās, the Brāhmaṇas, the Āraṇyakas and the Śrautasūtras*. Vol.II, Sanskrit Section. Part II, *The seven Soma-sacrifices subsequent to the Agniṣṭoma*. Pune: Vaidika Saṁśodhana Maṇḍala.
- McPherson, Hugh. (1924). "The Oriya Alphabet". *Journal of Bihar and Orissa Research Society* (Mar-Jun).
- Nemāni, Venkaṭa Narasimhaśāstrī & Vaḍlamūḍi, Gopālakṛṣṇayya. (1982). *Samūla Śrīmadāndhra Ṛgveda Samhitā*. Vol. 3. Tirumala Tirupati Devasthānamulu, Tirupati.
- Nemāni, Venkaṭa Narasimhaśāstrī & Vaḍlamūḍi, Gopālakṛṣṇayya. (1985). *Samūla Śrīmadāndhra Ṛgveda Samhitā*. Vol. 4. Tirumala Tirupati Devasthānamulu, Tirupati.
- Pai, B.S., ed. (1956). *Kannaḍa Ṛgveda Samhiteyu*. Bhāga 1. [in Kannada script.] Indological Series: Kannada Editions I, Laxman Babani Pai Memorial Publications, Jaichamraj Nagar, Hubli.
- Raghu Vira, ed. (1936-1941). *Atharvavedīyā Paippalāda-Samhitā*. 3 vols. Mehar Chand Lachhman Das Sanskrit and Prakrit series; 4. Lavapuram: Sarasvati Vihara.
- Rāmeśvarāvadhānī, P.S., ed. (1994). *Taittirīya Samhitā - Prathamakāṇḍa: 1-2 Prapāṭhaka*. [in Kannada script.] Kṛṣṇa Yajurveda Kannaḍa Prakāśana Saṁpuṭa 2. Jyoti Sāṁskṛtika Pratiṣṭhāna, Banagalore.
- Rao, H.P. Venkata, ed. (1954). *Ṛgveda Samhitā*. Bhāga - 22. [in Kannada script.] Śrī Jayacāmarājendra Vedaratnamālā, Mysore.
- Rastogi, Shrimati Indu, ed. trans. (1967). *The Śuklayajuḥ-prātiśākhya of Kātyāyana*. Forward by Mangal Deva Shastri [her father]. Kashi Sanskrit Series 179. Varanasi: Chowkhamba Sanskrit Series Office.
- Roy, Satyacharan, (n.d.). *Sāmaveda-samhitā (Āraṇya Parva)* (in Bengali script). Brajeshwar Roy Publication, Calcutta.
- Sāmaveda Samhitā Part I*. [in Grantha script.] (1985). Śrī Govinda Dikṣitar Puṇya Smaraṇa Samiti (Reg.), Kumbakonam - 612081.
- Śarmmā, Ananta Tripāṭhī. (1976). *Atharvaveda Samhitā: Śaunakīya*. Śiromaṇi Press, Brahmapura.
- Śatapathabrāhmaṇa. *According to the Mādhyandina Recension with the Vedārthaprakāśa bhāṣya of Sāyaṇācārya and supplemented by the commentary of Harisvāmin*. 5 volumes. Volume 5, *Bṛhadāraṇyakopaniṣad with an additional commentary of Vāsudeva Brahma Bhagavat*. (1987). Delhi: Gian Publishing House.
- Śāstrī, Gaṅgādhara A.N., ed. (n.d.). *Taittirīya Yajurvedīya Yajurāraṇyaka Āraṇyaka, Upaniṣat, Ekāgnikāṇḍa Mantra Sahita*. [in Kannada script.] Bangalore.
- Sastri, Mahadeva A. and K. Rangacharya. (1984). *The Taittirīyasamhitā of the Black Yajurveda with the commentary of Bhaṭṭa Bhāskara Miśra*. 10 volumes. Mysore: Government Oriental Library.
- Sastri, Mahadeva A. & K. Rangacharya. (1900-02). *The Taittirīya Āraṇyaka: with the commentary of Bhaṭṭa Bhāskara Miśra*. 3 Vols (Bound in one). Mysore State Government Oriental Library Series. [Reprint: Motilal Banarsidass, Delhi (1985).]

- Śastrī, Nārāyaṇa T.M. (1926). *Taittirīyāraṇyakam Kāṭhakahāgasahitam, Drāviḍapāṭhakramayātam ca*. [in Grantha script.] Śaradāvilāsamudrākṣarasālā, Kumbakonam.
- Sātvalekar, Śrīpād Dāmodara, ed. (1983). *Kṛṣṇayajurvedīya Kāṭhaka-Saṃhitā*. Pārḍi, Gugarāt: Svādhyāya Maṇḍalam, 4th ed.
- Sātvalekar, Śrīpād Dāmodara, ed. (1985). *Ṛk-Saṃhitā*. Pārḍi, Gugarāt: Svādhyāya Maṇḍalam.
- Sātvalekar, Śrīpād Dāmodara, ed. (1990). *Kṛṣṇayajurvedīya Taittirīya-Saṃhitā*. Pārḍi, Gugarāt: Svādhyāya Maṇḍalam, 5th ed.
- Sātvalekar, Śrīpād Dāmodara, ed. (n.d.). *Kṛṣṇayajurvedīya Maitrāyaṇī-Saṃhitā*. Pārḍi, Gugarāt: Svādhyāya Maṇḍalam, 4th ed.
- Sātvalekar, Śrīpād Dāmodara, ed. (n.d.). *Śuklayajurvedīya Kāṇva-Saṃhitā*. Pārḍi, Gugarāt: Svādhyāya Maṇḍalam, 4th ed.
- Scheftelowitz, Isidor. 1966. *Die Apokryphen des Ṛgveda*. Hildesheim: Georg Olms Verlagsbuchhandlung.
- Schroeder, Leopold von, ed. (1881-1886). *Maitrāyaṇī-Saṃhitā: die Saṃhitā der Maitrāyaṇīya-śākhā*. Leipzig: Verlag der Deutschen Morgenländischen Gesellschaft. [Reprint: Wiesbaden: F. Steiner (1970-1972)].
- Schroeder, Leopold von, ed. (1900-1912). *Kāṭhaka: die Saṃhitā der Kāṭha-śākhā*. 4 vols. 4th vol: *Index Verborum* by Richard Simon. Leipzig: Verlag der Deutschen Morgenländischen Gesellschaft. [Reprint: Wiesbaden: F. Steiner (1970-1972)].
- Sharma, B. R., ed. (2000). *Sāmaveda Saṃhitā of the Kauthuma School: with padapāṭha and the commentaries of Madhava, Bharatasvāmin and Sayana*. 3 vols. Cambridge, Mass.: Harvard University Press.
- Sharma, Candradhara, ed. (1937). *Mādhyandinaśākhīyam Śatapathabrāhmaṇam*. Banaras: Acyutagrānthamālākāryālaya.
- Shastri, A. Mahadeva, ed. (1911). *Taittirīya Brāhmaṇa with the commentary of Bhaṭṭa Bhāskara Miśra*. 4 vols. Mysore: Mysore State Government Oriental Library Series, Bibliotheca Sanskrita No. 36. [Reprint: Delhi: Motilal Banarsidass, 1985].
- Shastri, Mangal Deva. (1926). *A Comparison of the contents of the Ṛgveda, Vājasaneyi, Taittirīya and Atharva-Prātiśākhya*. Princess of Wales Sarasvati bhavana Studies, vol. 5. Banaras.
- Shastri, Mangal Deva, ed. trans. *The Ṛgveda-prātiśākhya with the commentary of Uvaṭa: edited from original manuscripts, with introduction, critical and additional notes, English translation of the text and several appendices and indices*. 3 vols. Vol. I, *Introduction, original text of the Ṛgveda-prātiśākhya in stanza-form, supplementary notes and several appendices*. Varanasi: Vaidika Svādhyāya Mandira, 1959. Vol. II, *Text in sūtra-form and commentary with critical apparatus*. Allahabad: The Indian Press, 1931. Vol. III, *English translation of the text, additional notes, several appendices and indices*. Lahore: Moti Lal Banarsi Das [Motilal Banarsidass], 1937.
- Shrouthi, Ramamurthy, ed. 1998. *Samaveda (prakṛti bhāgaḥ)*. Sringeri: Sri Sharada Peetham.
- Sontakke, N. S. and Kashikar, C. G., eds. (1972). *Ṛgveda-Saṃhitā, with the commentary of Sāyanācārya*. 5 vols. 2d. Pune: Tilak Maharashtra Vidyapith, Vaidika Saṃśodhana Maṇḍala.
- Sontakke, N. S. and T. N. Dharmadhikari. *Śrautakośa*. Vol. 2, Sanskrit section part 1, *Agniṣṭoma with Pravargya*. Pune: Vaidika Saṃśodhana Maṇḍala, 1970.
- Surya Kanta, ed. trans. (1939). *Atharvapraṭiśākhya with critical introduction and notes*. Lahore: Mehar Chand Lachhman Das. [Reprint: Delhi: Mehar Chand Lachhman Das, 1968].
- Surya Kanta, ed. (1943). *Kāṭhaka Saṃkalana: saṃskṛtagranthebhyah saṃgrhitāni kāṭhaka brāhmaṇa, kāṭhaka śrautasūtra, kāṭhaka gṛhyasūtrāṇām uddharaṇāni*. Lahore: Mehar Chand Lachhman Das. [Reprint: Delhi: Mehar Chand Lachhman Das, 1981].
- Sukthankar, S. Vishnu and S. K. Belvalkar eds. (1925-1959). *The Mahābhārata*. 19 vols. Poona: Bhandarkar Oriental Research Institute.
- Śuklayajurveda Kāṇvasaṃhitā*. [in Kannada script.] (1985). Śrī Śuklayajuśśākhā Trust (Reg.), Śrī Yājñavalkyāśrama, Chamarajpet, 3rd main road, Bangalore - 560050.
- Svāmi, Cidānanda, ed. (1989). *Sasvara Kṛṣṇayajurvedīya Taittirīyasamhitā (Caturtham Pañcamam Kāṇḍam)*. Saṃpuṭa 2. [in Kannada script.] Śrī Rāmakṛṣṇāśrama, Bangalore.
- Taittirīyasamhitā*. [in Grantha script.] Kumbakonam Publication
- Thakur, Paritosh, ed. (n.d.). *Ṛgveda-saṃhitā*. Bhāga 1. [in Bengali script.] Veda Prakashan, Calcutta - 26.
- Yājuṣa Mantra Ratnākaram*. [in Grantha script.] (2002). Heritage India Educational Trust, 6, Sanskrit College Street, Chennai - 600004.
- Weber, Albrecht. (1972). *Śuklayajurvede Vājasaneyisaṃhitā*. Chaukhamba Sanskrit Series 103. Varanasi: Chaukhamba Sanskrit Series Office.
- Weber, Albrecht. (1855). *The Ṣatapathabrāhmaṇa in the Mādhyandina-ṣākhā with extracts from the commentaries of Sāyaṇa, Harisvāmin and Dvivedagaṅga*. Berlin. Reprinted 1964, Varabasu: Chowkhamba.

- Whitney, William Dwight, ed. (1871). *Taittirīya Prāṭisākhya: with its commentary, the Tribhāshyaratna: text, translation, and notes*. JAOS 9: 1-469. [Reprint: Delhi, 1973]
- Whitney, William Dwight, ed. trans. (1905). *Atharva-Vedasamhitā: Text with English translation, Mantra Index and Names of Ṛṣis and Devatās*. Cambridge, Mass: Harvard University. [Reprint: Delhi: Nag Publishers, 1987].
- Witzel, Michael. (1565 CE). *Maitrāyaṇī Samhitā*, Kāṇḍa I (mss).
- Witzel, Michael. (n.d.). *Śatapatha Brāhmaṇa*, Kāṇḍa I (mss).
- Witzel, Michael. 1974. “On some unknown systems of marking the vedic accents”, in *Vishvabandhu Commemoration Volume. Vishveshvaranand Indological Journal* (VIJ) 12: 472-502.

## 12. Acknowledgements

This project was made possible by grants from the U.S. National Science Foundation, which funded the International Digital Sanskrit Library Integration Project (at Brown University, grant no. 0535207), and from the U.S. National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at UC Berkeley). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the National Endowment for the Humanities.

## 13. Figures.

### 2. Characters already encoded.

**Figure 2A.** U+0951 DEVANAGARI STRESS SIGN UDATTA primarily used as svarita but also as udātta in some Vedic schools. In figure 2Aa, the vertical stroke represents a svarita in Satvalekar’s edition of the *Ṛgveda* 1.1.1, as it does in figure 2Ab *Ṛgveda-Samhitā*, Poleman manuscript 4 / Houghton Indic Ms 636, folio 5 verso. In figure 2Ac, on the other hand, the same character represents an udātta in Raghu Vira’s edition of the *Atharvaveda Paippalāda-Samhitā* 1.30.6.

होतारं

Figure 2Aa

॥पुनानोवर्धनो गिरः॥

Figure 2Ab

क इदं

Figure 2Ac

**Figure 2B.** U+0952 DEVANAGARI STRESS SIGN ANUDATTA. Figure 2Ba shows *Ṛgveda* 1.82.1 in Satvalekar’s edition. Figure 2Bb is taken from folio 5 verso of Poleman manuscript 4 / Houghton Indic Ms 636 *Ṛgveda-Samhitā*.

यदा

Figure 2Ba

॥पुनानोवर्धनो गिरः॥

Figure 2Bb

Figure 2C. U+0CF1 KANNADA SIGN JIHVAMULIYA.

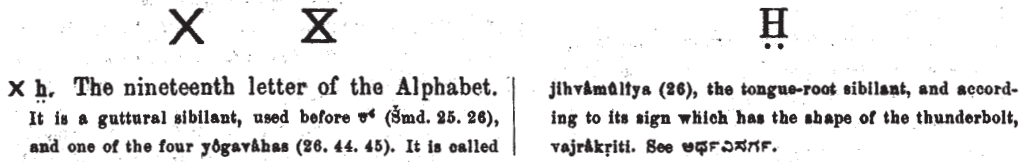


Figure 2C

Figure 2D. U+0CF2 KANNADA SIGN UPADHMANIYA.

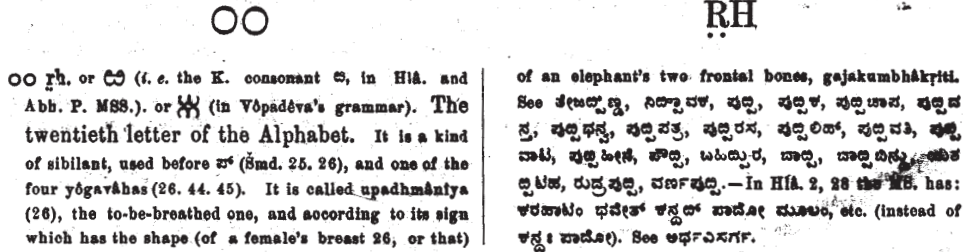


Figure 2D

### 3. Combining diacritic for the Ṛgvedic tradition.

Figure 3. VEDIC TONE RIGVEDIC KASHMIRI INDEPENDENT SVARITA. Figures 3a and 3b show the character in Sontakke's edition of the *Ṛgveda Khilāni*, verse RVKh 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 *Ṛgveda Khilani*.



Figure 3a

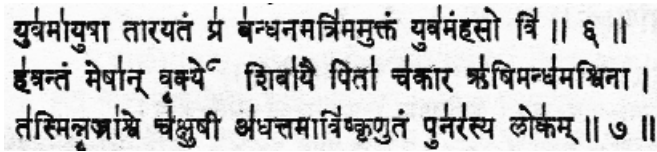


Figure 3b

3a: Der primäre Svarita wird durch ein Häkchen (̣) über dem svaritierten Laut angedeutet. So ruht der eigentliche Haken auf dem svaritierten *y* bzw. *v*, während die Verlängerung der einen Seite des Häkchens sich über dem nächsten Vokal befindet. „Im Kāthaka finden wir eine kräftige Kurve über der Linie als Bezeichnung des primären Svarita“, dagegen steht in der Maitr.S. die Kurve unter der Linie (Schroeder, Maitr. XXX).

Figure 3c

Beispiele für den Svarita ad 3a: 5, 45, 1 *sv̄ar*; 7, 59, 12 *tryāmbakam*; 7, 34, 17 *budh̄nyo r̄ṣé*; 7, 34, 13 *vȳetu*; 7, 34, 4 *dh̄r̄ṣv̄ asmai*; 7, 29, 2 *jūṣān̄orv̄ācinó*; 7, 25, 4 *h̄ndra*; 7, 24, 5 *div̄iva*; 7, 1, 2 *nȳ r̄n̄van*; 8, 26, 10 *sv̄ r̄ṣe*; ad 3b: 5, 66, 2 *amurya3m ās̄āte*; 6, 29, 3 *sva3r̄n̄á*; 7, 30, 3 *nya3gn̄ih̄*; 7, 38, 3 *namasya3śc̄áno*; 7, 79, 1 *vyu3śá*.

Figure 3d

#### 4. Combining characters for the Sāmavedic tradition.

##### 4.1. Combining digits and letters for the Sāmavedic tradition.

**Figure 4.1A.** COMBINING DEVANAGARI DIGIT ZERO in *Sāmaveda*, *Kouthama śākhā*, *Uha Uhya Gana*, Vol. I. <http://www.vedamu.org/>.

ॐ-योञ्जिरिचुभ्योमधवा ।  
योरा विन् यामा विन् नैतायज्ञ  
१२४ २ ३ ५२ १२ २४ २  
लि-येकेचोदरसपिणस्त्रोभ्योनमा.

Figure 4.1A

**Figure 4.1B.** COMBINING DEVANAGARI DIGIT ONE in Dandekar's edition of the *Śrautakośa* Sanskrit Section, Vol. II, Part II. Figure 4.1Ba is from p. 206; figure 4.1Bb is from p. 12. Figure 4.1Bc shows the two digits combined as the number 11. The canonical character order to generate ॐ with ११ above will be the following: ॐ 090A + ० A8E1 + ० A8E1.

१  
माइमा

Figure 4.1Ba

१ १ १ १  
२ ३ ४ ५

Figure 4.1Bb

अतिक्रमः चतुर्विधः— ११ ३ ५  
१. कुट्टः प्रथममतिक्रामति— ऊऽ२ । न आगाऽ२३४ही ।

Figure 4.1Bc

**Figure 4.1C.** COMBINING DEVANAGARI DIGIT TWO, taken from Dandekar's edition of the *Śrautakośa*, Sanskrit Section, Vol. II, Part II, p. 206. In figure 4.1Cb, the canonical character order to generate र with २२ above will be the following: र 0930 + ा 093E + ० A8E2 + ० A8EF. In figure 4.1Cc, the canonical character order to generate आ with २५ above will be the following: आ 0906 + ० A8E2 + ० A8F1.

२  
स्यचा

Figure 4.1Ca

२२ १  
राधो

Figure 4.1Cb

२९  
आ

Figure 4.1Cc

**Figure 4.1D.** COMBINING DEVANAGARI DIGIT THREE in samples taken from Dandekar's edition of the *Śrautakośa*, Sanskrit Section, Vol. II, Part II, p. 237. In figure 4.1Db, the canonical character order to generate नू with ३र above will be the following: न 0928 + ू 0942 + ३ A8E3 + ३ A8EF.

३ २ ३ २  
अदा ३४ ब्धस्सू

Figure 4.1Da

३र २  
नू ना

Figure 4.1Db

**Figure 4.1E.** COMBINING DEVANAGARI DIGIT FOUR in Dandekar's edition of the *Śrautakośa*, Sanskrit Section, Vol. II, Part II, p. 206.

४  
गा

Figure 4.1E

**Figure 4.1F.** COMBINING DEVANAGARI DIGIT FIVE in Dandekar's edition of the *Śrautakośa*, Sanskrit Section, Vol. II, Part II, p. 206. In figure 4.1Fb, the canonical character order to generate वा with ५र above will be the following: व 0935 + ा 093E + ५ A8E5 + ५ A8EF.

५  
दहाउवा

Figure 4.1Fa

५र ५र  
ताओवा

Figure 4.1Fb

**Figure 4.1G.** COMBINING DEVANAGARI DIGIT SIX in Dandekar's edition of the *Śrautakośa*, Sanskrit Section, Vol. II, Part II, p. 237.

६ २  
ओ३हा

Figure 4.1G



**Figure 4.1H.** COMBINING DEVANAGARI DIGIT SEVEN in Dandekar’s edition of the *Śrautakośa*, Sanskrit Section, Vol. II, Part II, p. 12.

७  
अवा

**Figure 4.1H**

**Figure 4.1I.** COMBINING DEVANAGARI DIGIT EIGHT ABOVE.

No occurrences of this character have yet been found.

**Figure 4.1J.** COMBINING DEVANAGARI DIGIT NINE in Dandekar’s edition of the *Śrautakośa*, Sanskrit Section, Vol. II, Part II, p. 12.

९  
यो

**Figure 4.1J**

**Figure 4.1K.** COMBINING DEVANAGARI LETTER A in Dandekar’s edition of the *Śrautakośa*, Sanskrit Section, Vol. II, Part II, p. 239.

अ  
कृता

**Figure 4.1K**

**Figure 4.1L.** COMBINING DEVANAGARI LETTER U ABOVE. Figure 4.1La shows the character in conjunction with the preceding digit 2 in B. R. Sharma’s edition of the *Sāmaveda*, pū 3.6.5. Figure 4.1Lb shows the character unconjoined in Bötlingk and Roth’s *Sanskrit Wörterbuch*, p. 831/832. In figure 4.1La, the canonical character order to generate च with ३ above followed by ना with २उ above will be the following: च 091A + ३ A83E + न 0928 + ा 093E + ३ A8E2 + २ A8EB.

३२उ ३ २  
चनास्मद्गातिः

**Figure 4.1La**

३ष्टका — ३ष्टि

**Figure 4.1Lb**

**Figure 4.1M.** COMBINING DEVANAGARI LETTER KA in B. R. Sharma’s edition of the *Sāmaveda*, pū 1.5.8. In figure 4.1M, the canonical character order to generate त with ३क above will be the following: त 0924 + ३ A8E3 + ३ A8EC, and to generate न्वा with २र above will be: न 0928 + ३ 094D + व 0935 + ा 093E + ३ A8E2 + ३ A8EF.

३क २र  
त न्वा

**Figure 4.1M**

**Figure 4.1N.** COMBINING DEVANAGARI LETTER NA in Dandekar’s edition of the *Śrautakośa*, Sanskrit Section, Vol. II, Part II. p. 287. 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.

१ न  
सियौ

Figure 4.1N

**Figure 4.1O.** COMBINING DEVANAGARI LETTER PA in Dandekar’s edition of the *Śrautakośa*, Sanskrit Section, Vol. II, Part II. p. 250.

१ प  
वा २ तौ

Figure 4.1O

**Figure 4.1P.** Samples showing COMBINING DEVANAGARI LETTER RA in Dandekar’s edition of the *Śrautakośa*, Sanskrit Section, Vol. II, Part II, p. 206. Figure 4.1Pb shows the character in conjunction with a preceding digit 2, and figure 4.1Pc shows it in conjunction with a preceding digit 5. In figure 4.1Pb, the canonical character order to generate रा with २र above will be the following: र 0930 + ा 093E + ० A8E2 + ० A8EF. In figure 4.1Pc, the canonical character order to generate वा with ५र above will be the following: व 0935 + ा 093E + ० A8E5 + ० A8F1.

२ र ५  
धाओवा

Figure 4.1Pa

२र १  
रा धौ

Figure 4.1Pb

२ र ५र  
ताओवा

Figure 4.1Pc

**Figure 4.1Q.** COMBINING DEVANAGARI LETTER VI. Figures 4.1Qa and b show the character in Shrouthi 1998: 182 and 184, and figures 4.1Qc and 4.1Qd show the full pages to demonstrate the context of the examples in *Sāmagāna*.

१२ १वि १ २ र  
आतिह्वाराऽ२२सिरऽहिया ।

Figure 4.1Qa

१ र २१वि १  
नस्तोममपुत्राऽ२२रजस्तुराऽ२३म् ॥

Figure 4.1Qb

१२२२१ २३२१ २२ ३  
 गोभिष्टेव । णमभिवाऽ२३ ॥ सयामसिहोऽ२३४इ ॥ डा ॥ \* (दी३-प१०-मा५-)३५(णु)।११५८।  
 ५७६-१. ॥ यशासि त्रीणि त्रयाणां सोम उष्णिक सोमः ॥  
 ५ र २१२ २ २ ३ ५ १२१वि १ २ र  
 पवतेहा ॥ र्यतोहरिः । औऽ३होऽ३१इ । औऽ२होऽ२३४वा । आतिह्वराऽ२सिरहिया ।

Figure 4.1Qc

५८०-१. ॥ कार्णश्रवसानि त्रीणि त्रयाणां कर्णश्रवा ककुबन्धः ॥  
 ५२२२ १ ३ ५२२ ३ ५ १२१२ १२वि १ २  
 आसोतापा ॥ राऽ२इषाऽ२३४औहोवा । चाऽ२३४ता । अम्बन्नस्तोमममूराऽ२रजस्तुरम् ॥ ओयेऽ३ ।

Figure 4.1Qd

**Figure 4.1R.** Samples showing COMBINING DEVANAGARI SIGN AVAGRAHA ABOVE. Figure 4.1Ra shows avagraha conjoined with the preceding digit 2 in Dandekar's edition of the *Śrautakośa*, Sanskrit Section, Vol. II, Part II, p. 206. Figure 4.1Rb-d show two examples of superscript avagraha unconjoined on p. 622 of Samaśrami's edition of *Sāmagāna*. In figure 4.1Ra, the canonical character order to generate आ with २५ above will be the following: आ 0906 + ० A8E2 + ० 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: ह 0939 + ी 0940 + ० A8F1.

२५ र र  
 आइमाउवा

Figure 4.1Ra

५ १  
 हीइया

Figure 4.1Rb

५ १  
 दिवऐरहीइया ।

Figure 4.1Rc

२ १ र र ५ २ १  
 या । सस्त्रिर्वाजन्दिवेदिवऐरहीइया । गोषाता । यि  
 २ २ १ ५ ४ ५ ३ ४ ५ र  
 रा१२३ । श्वसो२३४वा । आ५सो६चायि(२) ॥ सम्मिहो  
 २ १ ३ ५ १२ १ ५ २  
 अरूपे । हीऐही२३४या । षोभुवऐरहीऐरहीइया ।  
 १२ र २ १ ५ १ १  
 छपस्थाभिर्नधेनुभिरैरहीऐरहीइया । सायिदाञ्छया-

Figure 4.1Rd

**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 Kauthuma and Rāṇāyanīya traditions.

साजैसं [ १.५७.११ ] ≡ साकौसं [ १.६.२.२.११ ]  
 अञ्जतायिव्यञ्जतायिसमञ्जातायि । क्रतुर्हिहान्तीमध्वाभ्यञ्जतायि ।  
 ककी कि त्य क टथ श क्यका कित्या कि टथ श  
 सिन्धोरुच्छ्वासेष त यान्तामुक्षाणाम् । हि रण्यपावाःपशुमाप्सुगृभ्णताबु । वा ॥  
 कि की टथा क्यका किका कीकि चिचाश तख  
 [ प्राम ३.९.३१ ]

Figure 4.1Sa

क्रतुर्हिहान्तीमध्वाभ्यञ्जतायि  
 क्यका कित्या कि टथ श

Figure 4.1Sb

#### 4.2. Combining diacritics for the Sāmavedic tradition.

**Figure 4.2A.** Samples showing VEDIC TONE KARSHANA in Dandekar's edition of the *Śrautakośa*, Sanskrit Section, Vol. II, Part II. Figure 4.2Aa shows the character over a digit on p. 12. Figure 4.2Ab shows the character conjoined with a preceding digit 2 over an alphabetic character sequence on p. 239. In figure 4.2Ab, the canonical character order to generate ते with २ and the *karṣaṇa* above will be the following: त 0924 + ० 0947 + ० A8E2 + ० 1CD1. Note that the ५ above the धट precedes the visarga in canonical order.

३ ^ ३  
 मा २ हा

Figure 4.2Aa

२ ^ ३ ४ ५  
 ते सुमेधाः

Figure 4.2Ab

**Figure 4.2B.** VEDIC TONE SHARA in Dandekar's edition of the *Śrautakośa*, Sanskrit Section, Vol. II, Part II, p. 252. 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 2 in canonical order.

१ २ ^ ३  
 स्वर्नाः शू २ :

Figure 4.2B

**Figure 4.2C.** VEDIC TONE PRENKHA ABOVE. Figure 4.2Ca shows the character in Dandekar's edition of the *Śrautakośa*, Sanskrit Section, Vol. II, Part II, p. 12. Figure 4.2Cb shows the character connecting over several in-line characters in Rāmamūrtiśrauti's edition of Sāmaveda, *Kouthama śākhā, Uha Uhya Gana*, Vol. I. <http://www.vedamu.org/>. The prenkha character follows each syllable or other character over which it appears, in the following order: न 0928 + ० 0943 + ० A8E2 + भ 092D + ा 093E + ० 1CD3 + ० 093D + ० 1CD3 + २ 0968 + ० 1CD3.

भरा॑ रे॒न्ता॑

Figure 4.2Ca

वाँऽर॑३४सी॑ ॥ नृमाँऽर॑ इः ॥

Figure 4.2Cb

**Figure 4.2D.** VEDIC SIGN NIHSHVASA in Dandekar's edition of the *Śrautakośa*, Sanskrit Section, Vol. II, Part II, p. 239.

ना॑३ना॑५इ॑

Figure 4.2D

## 5. Combining diacritics for the Yajurvedic tradition.

### 5.1. General

**Figure 5.1A.** Samples showing VEDIC TONE YAJURVEDIC INDEPENDENT SVARITA. Figure 5.1Aa shows the character in *Śuklayajurveda Mādhyandina-Saṁhitā* 18.64, edited by Daulata Rāma Gauḍa and published by Caukhamba. Figure 5.1Ab shows it in Raghu Vira's edition of the *Atharvaveda Paippalāda-Saṁhitā*, verse 14.2.8.

नय॑स्व॒हे॒वेषु॑

Figure 5.1Aa

तन्वा॑ श॒त॒मया॑

Figure 5.1Ab

**Figure 5.1B.** Samples showing VEDIC TONE YAJURVEDIC KATHAKA INDEPENDENT SVARITA Figure 5.1Ba is in Poleman manuscript 93 / Houghton MS Indic 371, *Rudrajāpya*, folio 4 recto. Figure 5.1Bb is from Schröder's edition of the *Kṛṣṇayajurveda Kāṭhaka-Saṁhitā*, verse 24.4.

॥ ब्राह्मणस्य॑ मुखमासी॑षा

Figure 5.1Ba

वीर्या॑णि

Figure 5.1Bb

**Figure 5.1C.** Samples showing VEDIC TONE YAJURVEDIC AGGRAVATED INDEPENDENT SVARITA. Figure 5.1Ca shows the character in Daulata Rāma Gauḍa's edition of the *Śuklayajurveda Mādhyandina-Saṁhitā*, verse 38.17, published by Caukhamba. Figure 5.1Cb shows it in Satvalekar's edition of the *Kṛṣṇayajurveda Kāṭhaka-Saṁhitā*, verse 1.4.

अ॒भी॒म॒म्मा॒हि॒मा

Figure 5.1Ca

उर्बुन्तरिक्षं

Figure 5.1Cb

**Figure 5.1D.** Samples showing VEDIC TONE CANDRA BELOW. Figure 5.1Da shows the character in Satvalekar's edition of the *Kṛṣṇayajurveda Kāthaka-Saṁhitā*, verse 24.4; figure 5.1Cb, in Satvalekar's edition of the *Kṛṣṇayajurveda Maitrāyaṇī-Saṁhitā*, verse 1.2.9; and figure 5.1Cc, in the *Mādhyandina Śatapathabrāhmaṇa*, verse 1.1.1.16, published by Gian Publishing.

व्यृद्धा वां

Figure 5.1Da

वीर्याणि प्रवोच ५

Figure 5.1Db

देवान् ह वै षड्नेन

Figure 5.1Dc

**Figure 5.1E.** VEDIC TONE DOUBLE SVARITA in *Nakshatra Sutra*, TS 3.5.1.2.  
<http://www.sanskritdocuments.org/>.

रोहिणीवेतु पत्नी'। विश्वरूपा

Figure 5.1E

**Figure 5.1F.** VEDIC TONE TRIPLE SVARITA in *Kṛṣṇayajurveda Maitrāyaṇī-Saṁhitā*, Witzel manuscript 1571ce, folio 61 verso.

त्रिष्टुवन्ती मा जे मती

Figure 5.1F

**Figure 5.1G.** VEDIC TONE DOT BELOW in Raghu Vira's edition of the *Atharvaveda Paippalāda-Saṁhitā*. Figure 5.1Ga shows verse 16.104.6. 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.

जिह्वाया आस्याय च

Figure 5.1Ga

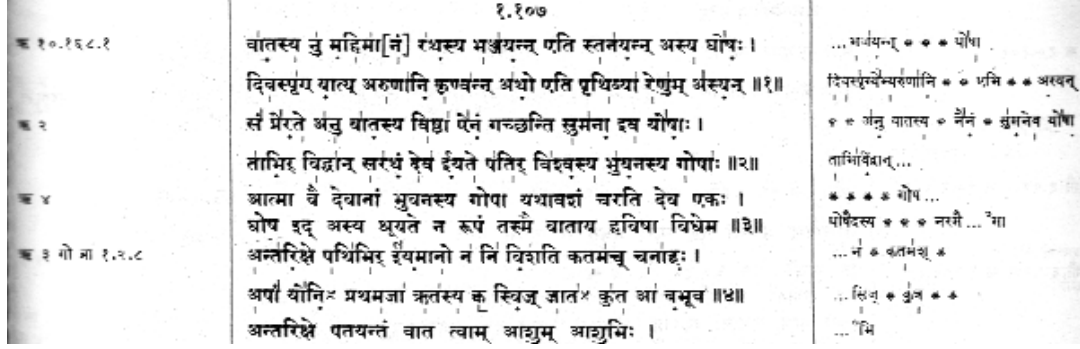


Figure 5.1Gb

**Figure 5.1H.** VEDIC TONE KATHAKA ANUDATTA BELOW in Raghu Vira's edition of the *Atharvaveda Paippalāda-Saṁhitā*, verse 2.18.1.

या देवी

Figure 5.1H

**Figure 5.1I.** VEDIC TONE YAJURVEDIC KATHAKA INDEPENDENT SVARITA SCHROEDER in Raghu Vira's edition of the *Atharvaveda Paippalāda-Saṁhitā*, verse 2.18.1.

उर्वन्तरिक्षं वीहीन्द्रस्य

Figure 5.1I

## 5.2 Śatapathabrāhmaṇa.

**Figure 5.2A.** Samples showing VEDIC TONE THREE DOTS BELOW in Weber's edition of the *Śatapathabrāhmaṇa*. Figure 5.2Aa is taken from ŚBr 9.2.3.26. Figure 5.2Ab shows the three dots doubled and stacked in ŚBr 4.2.1.13.

सम्पदेव । षट्

Figure 5.2Aa

यदुक्थ्यः । सोऽस्यैष

Figure 5.2Ab

**Figure 5.2B.** VEDIC TONE TWO DOTS BELOW in the Vedic Yantrālaya edition of the *Śatapathabrāhmaṇa* as shown by Yudhiṣṭhira Mīmāṁsaka 1964.

मानुऽपुषैर्न देवत्रा ॥७॥

Figure 5.2B

## 6. Combining diacritic for the Atharvavedic tradition.

**Figure 6.** Samples showing VEDIC TONE ATHARVAVEDIC INDEPENDENT SVARITA. Figure 6Aa is taken from Whitney's edition of the *Atharvaveda Śaunakīya-Saṁhitā*, verse 1.1.1; figures 6Ab and 6Ac, from other editions.

तन्वोऽ अद्य

Figure 6a

यातुधान्यः

Figure 6b

सोर्वर्धत स महानभवत् स महादेवोभवत् ॥४॥

Figure 6c

### 7. Combining diacritics for visarga.

**Figure 7A.** Samples showing VEDIC TONE SVARITA VISARGA. Figures 7Aa and 7Ab show the character in verses 4.25 and 1.31 in Daulata Rāma Gauḍa's edition of *Śuklayajurveda Mādhyandina-Saṁhitā*, as published by Gupta. Figure 7Ac shows the character in red, in accordance with the custom of marking all accents in red, in Poleman manuscript 93 / Houghton MS Indic 371, *Rudrajāpya*, folio 3 verso.

सवितुर्ब्रह्मसुव

Figure 7Aa

सवितारमोण्योःकृविक्रंतुम

Figure 7Ab

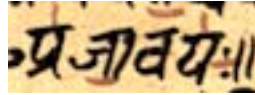


Figure 7Ac

**Figure 7B.** Samples showing VEDIC TONE UDATTA VISARGA. Figure 7Ba shows the character in verse 1.21 in Daulata Rāma Gauḍa's edition of *Śuklayajurveda Mādhyandina-Saṁhitā*, as published by Gupta. Figure 7Bb shows the character in red, in accordance with the custom of marking all accents in red, in Poleman manuscript 93 / Houghton MS Indic 371, *Rudrajāpya*, folio 4 verso.

सवितुःप्रसुवे

Figure 7Ba

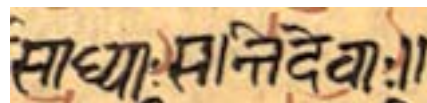


Figure 7Bb

**Figure 7C.** Samples showing VEDIC TONE ANUDATTA VISARGA. Figure 7Ca shows the character in verse 1.21 in Daulata Rāma Gauḍa's edition of *Śuklayajurveda Mādhyandina-Saṁhitā*, as published by Gupta.



Figure 7Cb shows the character in red, in accordance with the custom of marking all accents in red, in Poleman manuscript 93 / Houghton MS Indic 371, *Rudrajāpya*, folio 4 verso.

ओपधीभिःसमोपधयो

Figure 7Ca

विष्णुकर्मणाःसमवर्तता

Figure 7Cb

**Figure 7D.** TELUGU SIGN ARDHAVISARGA in Gomatham's Telugu edition of the *Taittirīya-Samhitā*, verse TS 1.1.24.

నగపితుంకృణ

Figure 7D

### 8. Nasals.

**Figure 8A.** DEVANAGARI SIGN INVERTED CANDRABINDU in *Kṛṣṇayajurveda Kāthaka-Samhitā* 9.7 in Schröder's edition.

भेषजं सुभेषजं

Figure 8A

**Figure 8B.** DEVANAGARI SIGN SPACING CANDRABINDU in Rāmamūrtiśrauti's edition of *Sāmaveda, Kouthama śākhā*, Uha Uha Gana, Vol. I. <http://www.vedamu.org/>.

५ ३ ४६ ५६  
फ-प्रस-७६[५३]इच्छात्कमायता ॥

Figure 8B

**Figure 8C.** DEVANAGARI SIGN CANDRABINDU VIRAMA in *Taittirīya-Samhitā* 5.6.1.2 in Shastri's edition.

सर्वाः अग्नीः

Figure 8C

**Figure 8D.** Samples showing DEVANAGARI SIGN DOUBLE CANDRABINDU VIRAMA. Figure 8Da and 8Db show it in *Taittirīya-Samhitā* 5.6.6.26 and 5.7.11.42 in Shastri's edition. 8Dc shows it in *Taittirīya Brāhmaṇa* 1.1.3.20 in Shastri's edition.

सु॒स्कृत्यात्मना

Figure 8Da

भूयो॒स्यान्न॑ सु॒ स्यादित्य

Figure 8Db

कथमिदं स्यादिति

Figure 8Dc

Figure 8E. DEVANAGARI SIGN CANDRABINDU TWO in Houghton Ms. Indic 62, folio 4 recto.

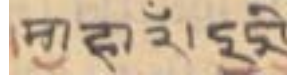


Figure 8Ea

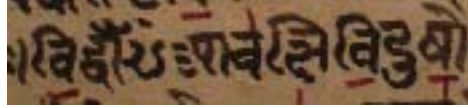


Figure 8Eb

Figure 8F. Samples showing DEVANAGARI SIGN CANDRABINDU THREE. Figure 8Fa shows the character in *R̥gveda-Samhitā* 10.146.1 in Satvalekar's edition. Figure 8Fb shows it in Poleman manuscript 163 / UP 2021, *Aitareya Āraṇyaka*, *Pañcāranyaka*.

विन्दतीँ ॥ १ ॥

Figure 8Fa

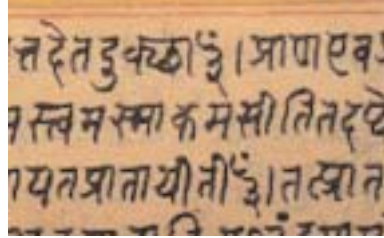


Figure 8Fb

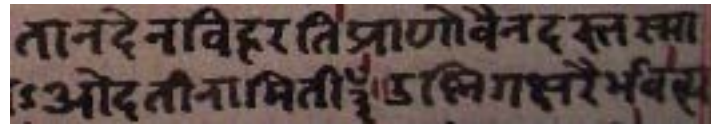


Figure 8Fc

Figure 8G. DEVANAGARI SIGN CANDRABINDU AVAGRAHA in Poleman manuscript 100 / Houghton MS Indic 133, *Śatarudriya*, folio 9 verso.

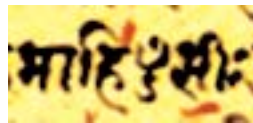


Figure 8G

Figure 8H. Samples showing VEDIC SIGN ANTARGOMUKHA. Figure 8Ha shows it in *Śuklayajurveda Mādhyandina-Samhitā* 1.21 in Daulata Rāma Gauḍa's edition published by Gupta. Figure 8Hb shows it in the same passage of the same text published by Caukhamba.

जर्गतीभिःपृच्छयन्तां॑सम्मधुमतीर्मा

Figure 8Ha

जर्गतीभिःपृच्छयन्तां॑सम्मधुमतीर्मा

Figure 8Hb

**Figure 8I.** Figures 8Ia and 8Ib show VEDIC SIGN BAHIRGOMUKHA in Poleman manuscript 3474 / UP 2032, *Rudraprārambha*, folio 2 verso.

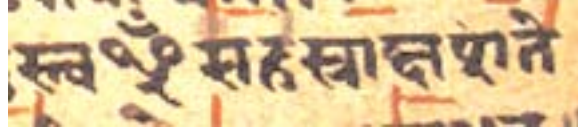


Figure 8Ia

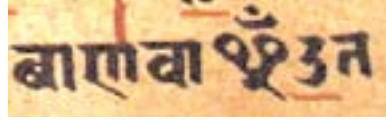


Figure 8Ib

**Figure 8J.** Samples showing VEDIC SIGN SAJIHVA BAHIRGOMUKHA. Figure 8Ja shows the character without bindu in the Acyutagrathamālā edition of the *Śuklayajurveda Mādhyandina Śatapathabrāhmaṇa* 1.1.1.3. Figure 8Jb shows it with bindu in *Śatapathabrāhmaṇa* 1.1.2.4 in the same edition. Figure 8Jc shows it with candrabindu in Gian Publishing edition of the *Śuklayajurveda Mādhyandina Śatapathabrāhmaṇa*, page 83.

स॒थ॒स्थि॒ते

Figure 8Ja

र॒क्षा॑थं॒स्य॒तो

Figure 8Jb

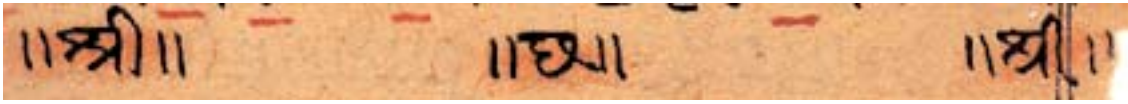
कृ॒ ॐ

Figure 8Jc

**Figure 8K.** Sample showing VEDIC SIGN LONG ANUSVARA. Figure 8Ka shows the character in *Śuklayajurveda Mādhyandina-Samhitā* 5.43 in Daulata Rāma Gauḍa's edition as published by Gupta, and figure 8Kb shows it in *Śuklayajurveda Mādhyandina-Samhitā* 4.1 in Daulata Rāma Gauḍa's edition as published by Caukhamba. The latter also shows VEDIC TONE VISARGA UDATTA and VEDIC TONE VISARGA ANUDATTA combined on a visarga in final position. Figure 8Kc shows a typeface imitation of the character using the Devanāgarī digit <६> with a bindu in *Śuklayajurveda Mādhyandina Śatapathabrāhmaṇa* 1.1.3.11 in Gian Publishing's edition. Figures 8Kd and 8Ke show the same in *Śuklayajurveda Mādhyandina Śatapathabrāhmaṇa* 1.2.1.18 and 1.4.1.39, in the Acyutagrathamālā edition.

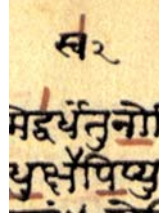


**Figure 9C.** DEVANAGARI SIGN PUSHPIKA in Poleman manuscript 4554 / Houghton MSIndic 133, *Devīrahasya*, folio 7 recto.



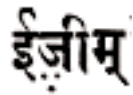
**Figure 9C**

**Figure 9D.** Samples showing DEVANAGARI SIGN DIVIDER. in Poleman manuscript 4 / Houghton MSIndic 636, *Ṛgveda-Saṁhitā*, folio 5 verso indicates that the characters in the top margin are to be inserted at the insertion point.



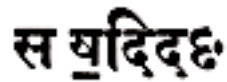
**Figure 9D**

**Figure 9E.** DEVANAGARI LETTER ZHA in Kanga's edition of the *Avesta*, yazna 41.3



**Figure 9E**

**Figure 9F.** DEVANAGARI LETTER HEAVY YA in the Acyutagrānṭhamālā edition of the *Śuklayajurveda Mādhyandina Śatapathabrāhmaṇa*, verse 1.1.3.4.



**Figure 9F**

Row 09: DEVANAGARI

	090	091	092	093	094	095	096	097
0	ँ	ऐ	ठ	र	ी	ॐ	ऋ	०
1	ँ	ऑ	ड	र	ु	ं	ॠ	॰
2	ं	ओ	ढ	ल	२	३	ॣ	अँ
3	ः	ओ	ण	ळ	५	६	१	छ
4	ऐ	औ	त	ळ	७	८	।	ँ
5	अ	क	थ	व	ँ	ँ	॥	
6	आ	ख	द	श	े		०	
7	इ	ग	ध	ष	े		१	
8	ई	घ	न	स	ै	२	३	
9	उ	ङ	न	ह	ॉ	ख	३	ज़
A	ऊ	च	प		ो	ग	४	ष
B	ऋ	छ	फ		ो	ज़	५	ग
C	ॠ	ज	ब	ि	ौ	ड	६	ज़
D	ँ	झ	भ	ऽ	्र	ढ	७	२
E	ऐ	ञ	म	ा		फ़	८	ड
F	ए	ट	य	ि		य	९	ब

G = 00  
P = 00

**Row 09: DEVANAGARI**

hex	Name
00	DEVANAGARI SIGN INVERTED CANDRABINDU
01	DEVANAGARI SIGN CANDRABINDU
02	DEVANAGARI SIGN ANUSVARA
03	DEVANAGARI SIGN VISARGA
04	DEVANAGARI LETTER SHORT A
05	DEVANAGARI LETTER A
06	DEVANAGARI LETTER AA
07	DEVANAGARI LETTER I
08	DEVANAGARI LETTER II
09	DEVANAGARI LETTER U
0A	DEVANAGARI LETTER UU
0B	DEVANAGARI LETTER VOCALIC R
0C	DEVANAGARI LETTER VOCALIC L
0D	DEVANAGARI LETTER CANDRA E
0E	DEVANAGARI LETTER SHORT E
0F	DEVANAGARI LETTER E
10	DEVANAGARI LETTER AI
11	DEVANAGARI LETTER CANDRA O
12	DEVANAGARI LETTER SHORT O
13	DEVANAGARI LETTER O
14	DEVANAGARI LETTER AU
15	DEVANAGARI LETTER KA
16	DEVANAGARI LETTER KHA
17	DEVANAGARI LETTER GA
18	DEVANAGARI LETTER GHA
19	DEVANAGARI LETTER NGA
1A	DEVANAGARI LETTER CA
1B	DEVANAGARI LETTER CHA
1C	DEVANAGARI LETTER JA
1D	DEVANAGARI LETTER JHA
1E	DEVANAGARI LETTER NYA
1F	DEVANAGARI LETTER TTA
20	DEVANAGARI LETTER TTHA
21	DEVANAGARI LETTER DDA
22	DEVANAGARI LETTER DDHA
23	DEVANAGARI LETTER NNA
24	DEVANAGARI LETTER TA
25	DEVANAGARI LETTER THA
26	DEVANAGARI LETTER DA
27	DEVANAGARI LETTER DHA
28	DEVANAGARI LETTER NA
29	DEVANAGARI LETTER NNNA
2A	DEVANAGARI LETTER PA
2B	DEVANAGARI LETTER PHA
2C	DEVANAGARI LETTER BA
2D	DEVANAGARI LETTER BHA
2E	DEVANAGARI LETTER MA
2F	DEVANAGARI LETTER YA
30	DEVANAGARI LETTER RA
31	DEVANAGARI LETTER RRA
32	DEVANAGARI LETTER LA
33	DEVANAGARI LETTER LLA
34	DEVANAGARI LETTER LLLA
35	DEVANAGARI LETTER VA
36	DEVANAGARI LETTER SHA
37	DEVANAGARI LETTER SSA
38	DEVANAGARI LETTER SA
39	DEVANAGARI LETTER HA
3A	(This position shall not be used)
3B	(This position shall not be used)
3C	DEVANAGARI SIGN NUKTA
3D	DEVANAGARI SIGN AVAGRAHA
3E	DEVANAGARI VOWEL SIGN AA
3F	DEVANAGARI VOWEL SIGN I
40	DEVANAGARI VOWEL SIGN II
41	DEVANAGARI VOWEL SIGN U
42	DEVANAGARI VOWEL SIGN UU
43	DEVANAGARI VOWEL SIGN VOCALIC R
44	DEVANAGARI VOWEL SIGN VOCALIC RR
45	DEVANAGARI VOWEL SIGN CANDRA E
46	DEVANAGARI VOWEL SIGN SHORT E
47	DEVANAGARI VOWEL SIGN E
48	DEVANAGARI VOWEL SIGN AI
49	DEVANAGARI VOWEL SIGN CANDRA O
4A	DEVANAGARI VOWEL SIGN SHORT O
4B	DEVANAGARI VOWEL SIGN O
4C	DEVANAGARI VOWEL SIGN AU
4D	DEVANAGARI SIGN VIRAMA
4E	(This position shall not be used)
4F	(This position shall not be used)
50	DEVANAGARI OM
51	DEVANAGARI STRESS SIGN UDATTA
52	DEVANAGARI STRESS SIGN ANUDATTA
53	DEVANAGARI GRAVE ACCENT
54	DEVANAGARI ACUTE ACCENT
55	DEVANAGARI VOWEL SIGN CANDRA LONG E
56	(This position shall not be used)
57	(This position shall not be used)
58	DEVANAGARI LETTER QA

Group 00

hex	Name
59	DEVANAGARI LETTER KHHA
5A	DEVANAGARI LETTER GHHA
5B	DEVANAGARI LETTER ZA
5C	DEVANAGARI LETTER DDDHA
5D	DEVANAGARI LETTER RHA
5E	DEVANAGARI LETTER FA
5F	DEVANAGARI LETTER YYA
60	DEVANAGARI LETTER VOCALIC RR
61	DEVANAGARI LETTER VOCALIC LL
62	DEVANAGARI VOWEL SIGN VOCALIC L
63	DEVANAGARI VOWEL SIGN VOCALIC LL
64	DEVANAGARI DANDA
65	DEVANAGARI DOUBLE DANDA
66	DEVANAGARI DIGIT ZERO
67	DEVANAGARI DIGIT ONE
68	DEVANAGARI DIGIT TWO
69	DEVANAGARI DIGIT THREE
6A	DEVANAGARI DIGIT FOUR
6B	DEVANAGARI DIGIT FIVE
6C	DEVANAGARI DIGIT SIX
6D	DEVANAGARI DIGIT SEVEN
6E	DEVANAGARI DIGIT EIGHT
6F	DEVANAGARI DIGIT NINE
70	DEVANAGARI ABBREVIATION SIGN
71	DEVANAGARI SIGN HIGH SPACING DOT
72	DEVANAGARI LETTER CANDRA A
73	DEVANAGARI SIGN PUSHPIKA
74	DEVANAGARI SIGN DIVIDER
75	(This position shall not be used)
76	(This position shall not be used)
77	(This position shall not be used)
78	(This position shall not be used)
79	DEVANAGARI LETTER ZHA
7A	DEVANAGARI LETTER HEAVY YA
7B	DEVANAGARI LETTER GGA
7C	DEVANAGARI LETTER JJA
7D	DEVANAGARI LETTER GLOTTAL STOP
7E	DEVANAGARI LETTER DDDA
7F	DEVANAGARI LETTER BBA

Plane 00

Row 09





**Row A8: DEVANAGARI EXTENDED**

	A8E	A8F
0	० ◌̐	वि ◌̐
1	१ ◌̐	ॡ ◌̐
2	२ ◌̐	ॢ ◌̐
3	३ ◌̐	ॣ ◌̐
4	४ ◌̐	। ◌̐
5	५ ◌̐	॥ ◌̐
6	६ ◌̐	० ◌̐
7	७ ◌̐	ॠ ◌̐
8	८ ◌̐	
9	९ ◌̐	
A	अ ◌̐	
B	उ ◌̐	
C	क ◌̐	
D	न ◌̐	
E	प ◌̐	
F	र ◌̐	

hex	Name
E0	COMBINING DEVANAGARI DIGIT ZERO
E1	COMBINING DEVANAGARI DIGIT ONE
E2	COMBINING DEVANAGARI DIGIT TWO
E3	COMBINING DEVANAGARI DIGIT THREE
E4	COMBINING DEVANAGARI DIGIT FOUR
E5	COMBINING DEVANAGARI DIGIT FIVE
E6	COMBINING DEVANAGARI DIGIT SIX
E7	COMBINING DEVANAGARI DIGIT SEVEN
E8	COMBINING DEVANAGARI DIGIT EIGHT
E9	COMBINING DEVANAGARI DIGIT NINE
EA	COMBINING DEVANAGARI LETTER A
EB	COMBINING DEVANAGARI LETTER U
EC	COMBINING DEVANAGARI LETTER KA
ED	COMBINING DEVANAGARI LETTER NA
EE	COMBINING DEVANAGARI LETTER PA
EF	COMBINING DEVANAGARI LETTER RA
F0	COMBINING DEVANAGARI LETTER VI
F1	COMBINING DEVANAGARI SIGN AVAGRAHA
F2	DEVANAGARI SIGN SPACING CANDRABINDU
F3	DEVANAGARI SIGN CANDRABINDU VIRAMA
F4	DEVANAGARI SIGN DOUBLE CANDRABINDU VIRAMA
F5	DEVANAGARI SIGN CANDRABINDU TWO
F6	DEVANAGARI SIGN CANDRABINDU THREE
F7	DEVANAGARI SIGN CANDRABINDU AVAGRAHA
F8	(This position shall not be used)
F9	(This position shall not be used)
FA	(This position shall not be used)
FB	(This position shall not be used)
FC	(This position shall not be used)
FD	(This position shall not be used)
FE	(This position shall not be used)
FF	(This position shall not be used)

## A. Administrative

1. Title

**Proposal to encode characters for Vedic Sanskrit in the BMP of the UCS**

2. Requester's name

**Michael Everson and Peter Scharf (editors), Michel Angot, R. Chandrashekar, Malcolm Hyman, Susan Rosenfield, B. V. Venkatakrishna Sastry, Michael Witzel**

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

**Individual contribution.**

4. Submission date

**2007-10-18**

5. Requester's reference (if applicable)

6. Choose one of the following:

6a. This is a complete proposal

**Yes.**

6b. More information will be provided later

**No.**

## B. Technical – General

1. Choose one of the following:

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

**Yes.**

1b. Proposed name of script

**Vedic Extensions, Devanagari Extended.**

1c. The proposal is for addition of character(s) to an existing block

**Yes**

1d. Name of the existing block

**Devanagari.**

2. Number of characters in proposal

**55 (5, 25, 25).**

3. Proposed category (A-Contemporary; B.1-Specialized (small collection); B.2-Specialized (large collection); C-Major extinct; D-Attested extinct; E-Minor extinct; F-Archaic Hieroglyphic or Ideographic; G-Obscure or questionable usage symbols)

**Category B.1.**

4a. Is a repertoire including character names provided?

**Yes.**

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

**Yes.**

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

**Yes.**

5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

**Michael Everson.**

5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

**Michael Everson, Fontographer.**

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

**Yes.**

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)?

**No.**

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. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database <http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

**See above.**

## C. Technical – Justification

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

**Yes, some of the characters have been proposed to the UTC by the Indian National Body.**

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?

**Peter Scharf (editors), Michel Angot, R. Chandrashekar, Malcolm Hyman, Susan Rosenfield, B. V. Venkatakrishna Sastry, Michael Witzel**

2c. If YES, available relevant documents

**Co-authors**

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

**Indologists, Indo-Europeanists, teachers, students, and practitioners of Vedic recitation, Hindus.**

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

**Used historically and liturgically.**

4b. Reference

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

**Yes.**

5b. If YES, where?

**Scholarly and religious publications.**

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

**Accordance with the Roadmap. Keep with other Indic characters.**

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

**No.**

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 an existing character?

**Yes.**

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

**Yes.**

10c. If YES, reference

**DEVANAGARI SIGN CANDRABINDU is a combining character and VEDIC SIGN CANDRABINDU is a non-combining character which is located below the Devanagari headbar.**

11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?

**Yes.**

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

**No.**

11c. If YES, reference

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

**No.**

11e. If YES, reference

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?