

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 additional characters for Vedic in the UCS
Source: UC Berkeley Script Encoding Initiative (Universal Scripts Project)
Authors: Michael Everson and Peter Scharf
Status: Liaison Contribution
Date: 2008-08-21

This proposal requests the encoding of a number of characters for Vedic use. It also requests to move three characters currently on the ballot as PDAM 6.2 into different code positions. In order to make clear the changes we propose for the ballot, we will discuss the changes block by block.

U+0900..U+097F Devanagari

U+094E ◌ DEVANAGARI VOWEL SIGN PRISHTHAMATRA E. We reiterate our request from N3383R (2008-03-06) and N3366 (2007-10-18) to encode this character. We remain convinced that this single character used in combination with other vowel signs is all that is necessary to correctly represent *e*, *ai*, *o*, and *au* in prsthmatra orthography. Our view is that using U+0947 for this with a radical glyph difference for prsthmatra orthography would simply be shifting the spoofing concern to a far more dangerous place. Protocols could restrict the use of U+094E DEVANAGARI VOWEL SIGN PRISHTHAMATRA E; they can't do the same for U+0947 DEVANAGARI VOWEL SIGN E. 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: ॢ is the *pr̥ṣṭhamātrā* rendering of के *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: ॣ 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: । 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: ॥ is the *pr̥ṣṭhamātrā* rendering of कौ *kau*.

U+0973 ◌ DEVANAGARI SIGN PUSHPIKA. We recommend that this character be moved to U+A8F8.

U+0974 ◌ DEVANAGARI CARET. We recommend that this character be moved to U+A8FA.

U+1CD0..U+1CFF Vedic Extensions

U+1CD4 ◌ VEDIC TONE YAJURVEDIC KASHMIRI SVARITA. This character is described at length in N3493 (L2/08-219). In northern manuscripts of the Mādhyandina recension of the *Vājasaneyi-samhitā* to indicate a dependent svarita, which is a falling tone from high to low. It marks the

tone on a vowel that occurs after a high-toned (udātta) vowel. In *Maitrāyaṇīsamhitā* its occurrence is further limited in being either final or followed by monotone (ekāśruti). In manuscripts it is marked in red by a horizontal line at mid-height at the right side of orthographic syllables.

U+1CE9 ॐ VEDIC SIGN ANUSVARA ANTARGOMUKHA (= *vaidika anusvaara antarmukha*) was described in N3366 and should be added to this position.

U+1CEA ॐ VEDIC SIGN ANUSVARA BAHIRGOMUKHA (= *vaidika anusvaara naagaphanna*) was described in N3366 and should be added to this position.

U+1CEB ॐ VEDIC SIGN ANUSVARA VAMAGOMUKHA (= *vaidika anusvaara vaamamukha*) was described in N3366 and should be added to this position.

U+1CEC ॐ VEDIC SIGN ANUSVARA VAMAGOMUKHA WITH TAIL (= *vaidika anusvaara vaamamukha savakra*) was described in N3366 and should be added to this position.

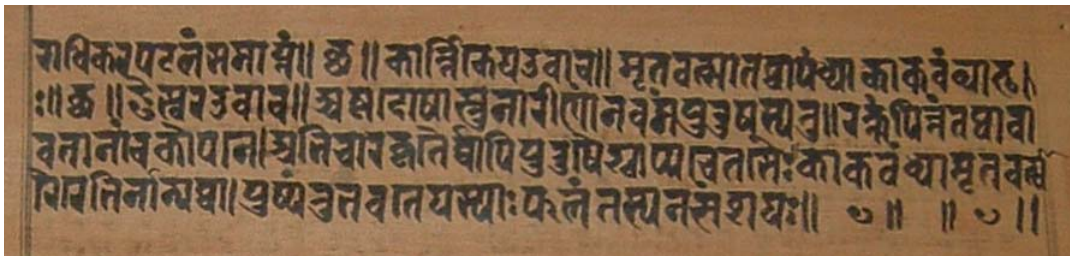
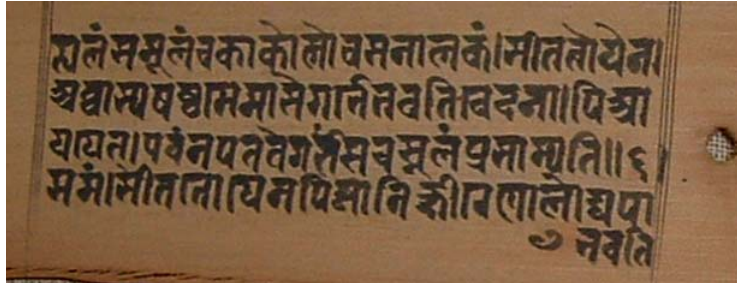
U+1CF1 ॐ VEDIC SIGN ARDHAVISARGA (= *vaidika jihvaamuliya upadhamaaniya*) should be moved down one position to U+1CF2.

U+1CF1 § VEDIC SIGN ANUSVARA UBHAYATO MUKHA (= *vaidika anusvaara ubhayato mukhaa*) shown in N3383R at U+A8F8 as DEVANAGARI SIGN ANUSVARA UBHAYATO MUKHA should be added to this position.

U+A8E0..U+A8FF Devanagari Extended

U+A8F8 ॐ DEVANAGARI SIGN PUSHPIKA should be added here, moved from U+0973.

U+A8F9 ॐ DEVANAGARI GAP FILLER should be added here. This character has not been previously proposed, It signifies any intentional gap in the space ordinarily occupied by text. Unlike the DEVANAGARI HEADSTROKE described below and used for missing or illegible akṣaras, the LINE FILLER signifies the opposite, namely, that the gap (or as shown below, the end of the text) is intentional and should not be taken as a lacuna in the text.



U+A8FA ॐ DEVANAGARI CARET should be added here, moved from U+0974.

U+A8FB ॐ DEVANAGARI HEADSTROKE (= *shirorekhaa*) should be added here. This character has not previously proposed, It is used in manuscripts copied from deteriorated originals to indicate small lacunae, that is, illegible gaps in the original due to obliteration of the text or destruction of the material on which it was written. For example, if the य *ya* in उभयतो *ubhayato* cannot be read, this character—which has a letter property—is used thus: उभ॒तो

ubha_to. The writer uses as many headstrokes as he estimates there to be orthographic syllables to have occupied the illegible space. Many Devanāgarī fonts include such a character. While manuscripts typically make the headstroke for each orthographic syllable (akṣara) as a separate unit, modern printed and digital Devanāgarī runs the headstroke together in a continuous line. Since it is essential to the character to be discrete and countable so that the reader knows how many missing characters are represented by counting the number of discrete headstrokes, the headstroke character should be made short, that is, non-joining when placed in sequence with itself or with any other character that has a headstroke included.

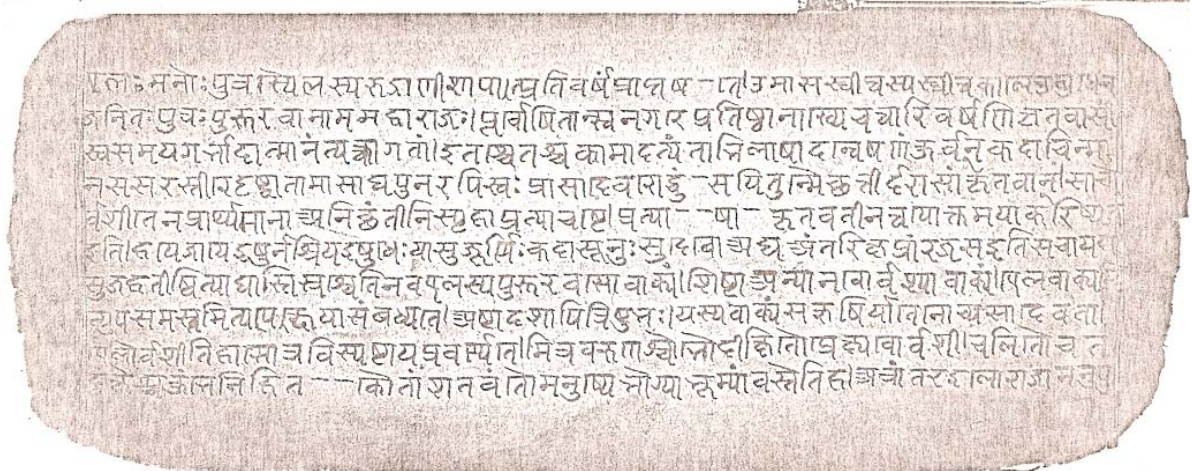


Figure 2 shows four examples, in lines 1, 4, 5, and 10 of the verso of folio 157 [Scharf photocopy p. 161 top] of the manuscript labelled by Macdonell (1886) as P1.

The code charts as affected by this proposal are attached.

10. Unicode Character Properties. Character properties are proposed here. These should be rigorously checked.

- 0900;DEVANAGARI SIGN INVERTED CANDRABINDU;Mn;0;NSM;;;;;N;;;;;
- 094E;DEVANAGARI VOWEL SIGN PRISHTHAMATRA E;Mc;0;L;;;;;N;;;;;
- 0955;DEVANAGARI VOWEL SIGN CANDRA LONG E;Mn;0;NSM;;;;;N;;;;;
- 0979;DEVANAGARI LETTER ZHA;Lo;0;L;;;;;N;;;;;
- 097A;DEVANAGARI LETTER HEAVY YA;Lo;0;L;;;;;N;;;;;
- 1CD0;VEDIC TONE KARSHANA;Mn;230;NSM;;;;;N;;;;;
- 1CD1;VEDIC TONE SHARA;Mn;230;NSM;;;;;N;;;;;
- 1CD2;VEDIC TONE PRENKHA;Mn;230;NSM;;;;;N;;;;;
- 1CD3;VEDIC TONE NIHSHVASA;Po;0;L;;;;;N;;;;;
- 1CD4;VEDIC TONE YAJURVEDIC KASHMIRI SVARITA;Mn;1;NSM;;;;;N;;;;;
- 1CD5;VEDIC TONE YAJURVEDIC AGGRAVATED INDEPENDENT SVARITA;Mn;220;NSM;;;;;N;;;;;
- 1CD6;VEDIC TONE YAJURVEDIC INDEPENDENT SVARITA;Mn;220;NSM;;;;;N;;;;;
- 1CD7;VEDIC TONE YAJURVEDIC KATHAKA INDEPENDENT SVARITA;Mn;220;NSM;;;;;N;;;;;
- 1CD8;VEDIC TONE CANDRA BELOW;Mn;220;NSM;;;;;N;;;;;
- 1CD9;VEDIC TONE YAJURVEDIC KATHAKA INDEPENDENT SVARITA SCHROEDER;Mn;220;NSM;;;;;N;;;;;
- 1CDA;VEDIC TONE DOUBLE SVARITA;Mn;230;NSM;;;;;N;;;;;
- 1CDB;VEDIC TONE TRIPLE SVARITA;Mn;230;NSM;;;;;N;;;;;
- 1CDC;VEDIC TONE KATHAKA ANUDATTA;Mn;220;NSM;;;;;N;;;;;
- 1CDD;VEDIC TONE DOT BELOW;Mn;220;NSM;;;;;N;;;;;
- 1CDE;VEDIC TONE TWO DOTS BELOW;Mn;220;NSM;;;;;N;;;;;
- 1CDF;VEDIC TONE THREE DOTS BELOW;Mn;220;NSM;;;;;N;;;;;
- 1CF0;VEDIC TONE RIGVEDIC KASHMIRI INDEPENDENT SVARITA;Mn;230;NSM;;;;;N;;;;;
- 1CF1;VEDIC TONE ATHARVAVEDIC INDEPENDENT SVARITA;Mc;0;L;;;;;N;;;;;
- 1CE2;VEDIC TONE VISARGA SVARITA;Mn;1;NSM;;;;;N;;;;;
- 1CE3;VEDIC TONE VISARGA UDATTA;Mn;1;NSM;;;;;N;;;;;
- 1CE4;VEDIC TONE REVERSED VISARGA UDATTA;Mn;1;NSM;;;;;N;;;;;
- 1CE5;VEDIC TONE VISARGA ANUDATTA;Mn;1;NSM;;;;;N;;;;;
- 1CE6;VEDIC TONE REVERSED VISARGA ANUDATTA;Mn;1;NSM;;;;;N;;;;;
- 1CE7;VEDIC TONE VISARGA UDATTA WITH TAIL;Mn;1;NSM;;;;;N;;;;;
- 1CE8;VEDIC TONE VISARGA ANUDATTA WITH TAIL;Mn;1;NSM;;;;;N;;;;;
- 1CE9;VEDIC SIGN ANUSVARA ANTARGOMUKHA;Lo;0;L;;;;;N;;;;;
- 1CEA;VEDIC SIGN ANUSVARA BAHIRGOMUKHA;Lo;0;L;;;;;N;;;;;
- 1CEB;VEDIC SIGN ANUSVARA VAMAGOMUKHA;Lo;0;L;;;;;N;;;;;
- 1CEC;VEDIC SIGN ANUSVARA VAMAGOMUKHA WITH TAIL;Lo;0;L;;;;;N;;;;;

1CED;VEDIC SIGN TIRYAK;Mn;220;NSM;;;;;N;;;;;
 1CEE;VEDIC SIGN HEXIFORM LONG ANUSVARA;Lo;0;L;;;;;N;;;;;
 1CEF;VEDIC SIGN LONG ANUSVARA;Lo;0;L;;;;;N;;;;;
 1CF0;VEDIC SIGN RTHANG LONG ANUSVARA;Lo;0;L;;;;;N;;;;;
 1CF1;VEDIC SIGN ANUSVARA UBHAYATO MUKHA;Lo;0;L;;;;;N;;;;;
 1CF2;VEDIC SIGN ARDHAVISARGA;Mc;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 SPACING CANDRABINDU;Lo;0;L;;;;;N;;;;;
 A8F3;DEVANAGARI SIGN CANDRABINDU VIRAMA;Lo;0;L;;;;;N;;;;;
 A8F4;DEVANAGARI SIGN DOUBLE CANDRABINDU VIRAMA;Lo;0;L;;;;;N;;;;;
 A8F5;DEVANAGARI SIGN CANDRABINDU TWO;Lo;0;L;;;;;N;;;;;
 A8F6;DEVANAGARI SIGN CANDRABINDU THREE;Lo;0;L;;;;;N;;;;;
 A8F7;DEVANAGARI SIGN CANDRABINDU AVAGRAHA;Lo;0;L;;;;;N;;;;;
 A8F8;DEVANAGARI SIGN PUSHPIKA;Po;0;L;;;;;N;;;;;
 A8F9;DEVANAGARI GAP FILLER;Po;0;L;;;;;N;;;;;
 A8FA;DEVANAGARI CARET;Mn;0;NSM;;;;;N;;;;;
 A8FB;DEVANAGARI HEADSTROKE;Lo;0;L;;;;;N;;;;;

| | 090 | 091 | 092 | 093 | 094 | 095 | 096 | 097 |
|---|-----------|-----------|-----------|------------|------------|------------|-----------|------------|
| 0 | ै 0900 | ऐ 0910 | ठ 0920 | र 0930 | ी 0940 | ॐ 0950 | ऋ 0960 | ० 0970 |
| 1 | ँ 0901 | ऑ 0911 | ड 0921 | र 0931 | ु 0941 | ं 0951 | ॠ 0961 | ॱ 0971 |
| 2 | ं 0902 | ओ 0912 | ढ 0922 | ल 0932 | ्र 0942 | ्र 0952 | ॡ 0962 | ँ 0972 |
| 3 | ः 0903 | ओ 0913 | ण 0923 | ळ 0933 | ृ 0943 | े 0953 | ॢ 0963 | |
| 4 | ऐ 0904 | औ 0914 | त 0924 | ळ 0934 | ृ 0944 | े 0954 | । 0964 | |
| 5 | अ 0905 | क 0915 | थ 0925 | व 0935 | ँ 0945 | ँ 0955 | ॥ 0965 | |
| 6 | आ 0906 | ख 0916 | द 0926 | श 0936 | े 0946 | | ० 0966 | |
| 7 | इ 0907 | ग 0917 | ध 0927 | ष 0937 | े 0947 | | १ 0967 | |
| 8 | ई 0908 | घ 0918 | न 0928 | स 0938 | ै 0948 | ऋ 0958 | २ 0968 | |
| 9 | उ 0909 | ङ 0919 | न 0929 | ह 0939 | ॉ 0949 | ख 0959 | ३ 0969 | ज़ 0979 |
| A | ऊ 090A | च 091A | प 092A | | ो 094A | ग 095A | ४ 096A | ष 097A |
| B | ऋ 090B | छ 091B | फ 092B | | ो 094B | ज 095B | ५ 096B | ग 097B |
| C | ॠ 090C | ज 091C | ब 092C | ्र 093C | ौ 094C | ड 095C | ६ 096C | ज 097C |
| D | ँ 090D | झ 091D | भ 092D | ऽ 093D | ्र 094D | ढ 095D | ७ 096D | २ 097D |
| E | ऐ 090E | ञ 091E | म 092E | ा 093E | ि 094E | फ़ 095E | ८ 096E | ड 097E |
| F | ए 090F | ट 091F | य 092F | ि 093F | | य 095F | ९ 096F | ब 097F |

Based on ISCII 1988**Various signs**

| | | |
|------|---|--|
| 0900 | ॐ | DEVANAGARI SIGN INVERTED CANDRABINDU = vaidika adhomukha candrabindu |
| 0901 | ँ | DEVANAGARI SIGN CANDRABINDU = anunasika → 0310 ॐ combining candrabindu |
| 0902 | ं | DEVANAGARI SIGN ANUSVARA = bindu |
| 0903 | ः | DEVANAGARI SIGN VISARGA |

Independent vowels

| | | |
|------|---|---|
| 0904 | ॐ | DEVANAGARI LETTER SHORT A |
| 0905 | अ | DEVANAGARI LETTER A |
| 0906 | आ | DEVANAGARI LETTER AA |
| 0907 | इ | DEVANAGARI LETTER I |
| 0908 | ई | DEVANAGARI LETTER II |
| 0909 | उ | DEVANAGARI LETTER U |
| 090A | ऊ | DEVANAGARI LETTER UU |
| 090B | ऋ | DEVANAGARI LETTER VOCALIC R |
| 090C | ॠ | DEVANAGARI LETTER VOCALIC L |
| 090D | ॡ | DEVANAGARI LETTER CANDRA E |
| 090E | ॢ | DEVANAGARI LETTER SHORT E for transcribing Dravidian short e |
| 090F | ए | DEVANAGARI LETTER E |
| 0910 | ऐ | DEVANAGARI LETTER AI |
| 0911 | ऑ | DEVANAGARI LETTER CANDRA O |
| 0912 | औ | DEVANAGARI LETTER SHORT O for transcribing Dravidian short o |
| 0913 | ओ | DEVANAGARI LETTER O |
| 0914 | औ | DEVANAGARI LETTER AU |

Consonants

| | | |
|------|----|--|
| 0915 | क | DEVANAGARI LETTER KA |
| 0916 | ख | DEVANAGARI LETTER KHA |
| 0917 | ग | DEVANAGARI LETTER GA |
| 0918 | घ | DEVANAGARI LETTER GHHA |
| 0919 | ङ | DEVANAGARI LETTER NGA |
| 091A | च | DEVANAGARI LETTER CA |
| 091B | छ | DEVANAGARI LETTER CHA |
| 091C | ज | DEVANAGARI LETTER JA |
| 091D | झ | DEVANAGARI LETTER JHA |
| 091E | ञ | DEVANAGARI LETTER NYA |
| 091F | ट | DEVANAGARI LETTER TTA |
| 0920 | ठ | DEVANAGARI LETTER TTHA |
| 0921 | ड | DEVANAGARI LETTER DDA |
| 0922 | ढ | DEVANAGARI LETTER DDHA |
| 0923 | ण | DEVANAGARI LETTER NNA |
| 0924 | त | DEVANAGARI LETTER TA |
| 0925 | थ | DEVANAGARI LETTER THA |
| 0926 | द | DEVANAGARI LETTER DA |
| 0927 | ध | DEVANAGARI LETTER DHA |
| 0928 | न | DEVANAGARI LETTER NA |
| 0929 | न् | DEVANAGARI LETTER NNNA for transcribing Dravidian alveolar n ≡ 0928 न 093C ॱ |
| 092A | प | DEVANAGARI LETTER PA |
| 092B | फ | DEVANAGARI LETTER PHA |
| 092C | ब | DEVANAGARI LETTER BA |

| | | |
|------|---|---|
| 092D | भ | DEVANAGARI LETTER BHA |
| 092E | म | DEVANAGARI LETTER MA |
| 092F | य | DEVANAGARI LETTER YA |
| 0930 | र | DEVANAGARI LETTER RA |
| 0931 | ॠ | DEVANAGARI LETTER RRA for transcribing Dravidian alveolar r half form is represented as “Eyelash RA” ≡ 0930 र 093C ॱ |
| 0932 | ल | DEVANAGARI LETTER LA |
| 0933 | ळ | DEVANAGARI LETTER LLA |
| 0934 | ॡ | DEVANAGARI LETTER LLLA for transcribing Dravidian l ≡ 0933 ळ093C ॱ |
| 0935 | व | DEVANAGARI LETTER VA |
| 0936 | श | DEVANAGARI LETTER SHA |
| 0937 | ष | DEVANAGARI LETTER SSA |
| 0938 | स | DEVANAGARI LETTER SA |
| 0939 | ह | DEVANAGARI LETTER HA |

Various signs

| | | |
|------|---|--|
| 093C | ॱ | DEVANAGARI SIGN NUKTA for extending the alphabet to new letters |
| 093D | ऽ | DEVANAGARI SIGN AVAGRAHA |

Dependent vowel signs

| | | |
|------|---|--|
| 093E | ा | DEVANAGARI VOWEL SIGN AA |
| 093F | ि | DEVANAGARI VOWEL SIGN I stands to the left of the consonant |
| 0940 | ी | DEVANAGARI VOWEL SIGN II |
| 0941 | ु | DEVANAGARI VOWEL SIGN U |
| 0942 | ू | DEVANAGARI VOWEL SIGN UU |
| 0943 | ॠ | DEVANAGARI VOWEL SIGN VOCALIC R |
| 0944 | ॡ | DEVANAGARI VOWEL SIGN VOCALIC RR |
| 0945 | ॢ | DEVANAGARI VOWEL SIGN CANDRA E = candra |
| 0946 | ॣ | DEVANAGARI VOWEL SIGN SHORT E for transcribing Dravidian vowels |
| 0947 | । | DEVANAGARI VOWEL SIGN E |
| 0948 | ॥ | DEVANAGARI VOWEL SIGN AI |
| 0949 | ० | DEVANAGARI VOWEL SIGN CANDRA O |
| 094A | ॡ | DEVANAGARI VOWEL SIGN SHORT O for transcribing Dravidian vowels |
| 094B | ॢ | DEVANAGARI VOWEL SIGN O |
| 094C | ॣ | DEVANAGARI VOWEL SIGN AU |

Sign

| | | |
|------|---|--|
| 094D | ् | DEVANAGARI SIGN VIRAMA = halant (the preferred Hindi name) suppresses inherent vowel |
|------|---|--|

Archaic dependent vowel sign

| | | |
|------|---|--|
| 094E | ० | DEVANAGARI VOWEL SIGN PRISHTHAMATRA E character has historic use only |
|------|---|--|

Various signs

| | | |
|------|---|--|
| 0950 | ॐ | DEVANAGARI OM |
| 0951 | ॠ | DEVANAGARI STRESS SIGN UDATTA mostly used for Rigvedic svarita, with rare use for Yajurvedic udatta |
| 0952 | ॡ | DEVANAGARI STRESS SIGN ANUDATTA |

| | | |
|------|----|--|
| 0953 | ◌̄ | DEVANAGARI GRAVE ACCENT |
| 0954 | ◌̆ | DEVANAGARI ACUTE ACCENT |
| 0955 | ◌̇ | DEVANAGARI VOWEL SIGN CANDRA LONG E used in transliteration of Avestan |

Additional consonants

| | | |
|------|----|---|
| 0958 | क़ | DEVANAGARI LETTER QA ≡ 0915 क़093C ◌̣ |
| 0959 | ख़ | DEVANAGARI LETTER KHHA ≡ 0916 ख़093C ◌̣ |
| 095A | ग़ | DEVANAGARI LETTER GHHA ≡ 0917 ग़093C ◌̣ |
| 095B | ज़ | DEVANAGARI LETTER ZA ≡ 091C ज़093C ◌̣ |
| 095C | ड़ | DEVANAGARI LETTER DDDHA ≡ 0921 ढ़093C ◌̣ |
| 095D | ढ़ | DEVANAGARI LETTER RHA ≡ 0922 ढ़093C ◌̣ |
| 095E | फ़ | DEVANAGARI LETTER FA ≡ 092B फ़093C ◌̣ |
| 095F | य़ | DEVANAGARI LETTER YYA ≡ 092F य़093C ◌̣ |

Additional vowels for Sanskrit

| | | |
|------|----|----------------------------------|
| 0960 | ठ | DEVANAGARI LETTER VOCALIC RR |
| 0961 | ड | DEVANAGARI LETTER VOCALIC LL |
| 0962 | ◌̈ | DEVANAGARI VOWEL SIGN VOCALIC L |
| 0963 | ◌̉ | DEVANAGARI VOWEL SIGN VOCALIC LL |

Generic punctuation for scripts of India

These punctuation marks are for common use for the scripts of India despite being named "DEVANAGARI".

| | | |
|------|--|---|
| 0964 | | DEVANAGARI DANDA = purna viram phrase separator |
| 0965 | | DEVANAGARI DOUBLE DANDA = deergh viram |

Digits

| | | |
|------|---|------------------------|
| 0966 | ० | DEVANAGARI DIGIT ZERO |
| 0967 | १ | DEVANAGARI DIGIT ONE |
| 0968 | २ | DEVANAGARI DIGIT TWO |
| 0969 | ३ | DEVANAGARI DIGIT THREE |
| 096A | ४ | DEVANAGARI DIGIT FOUR |
| 096B | ५ | DEVANAGARI DIGIT FIVE |
| 096C | ६ | DEVANAGARI DIGIT SIX |
| 096D | ७ | DEVANAGARI DIGIT SEVEN |
| 096E | ८ | DEVANAGARI DIGIT EIGHT |
| 096F | ९ | DEVANAGARI DIGIT NINE |

Devanagari-specific additions

| | | |
|------|----|----------------------------------|
| 0970 | ◌̣ | DEVANAGARI ABBREVIATION SIGN |
| 0971 | ◌̤ | DEVANAGARI SIGN HIGH SPACING DOT |

Additional vowel for Marathi

| | | |
|------|----|---------------------------------------|
| 0972 | अँ | DEVANAGARI LETTER CANDRA A Marathi |
|------|----|---------------------------------------|

Additional consonants

| | | |
|------|----|---|
| 0979 | ज़ | DEVANAGARI LETTER ZHA used in transliteration of Avestan |
| 097A | य़ | DEVANAGARI LETTER HEAVY YA used for an affricated glide → 09AF य़ bengali letter ya |

Sindhi implosives

These are added from Amendment 3 to 10646:2003.

| | | |
|------|----|-----------------------|
| 097B | ग़ | DEVANAGARI LETTER GGA |
| 097C | ज़ | DEVANAGARI LETTER JJA |

















































Glottal stop

| | | |
|------|----|--|
| 097D | ◌̥ | DEVANAGARI LETTER GLOTTAL STOP used for writing Limbu in Devanagari a glyph variant has the connecting top bar |
|------|----|--|

Sindhi implosives

These are added from Amendment 3 to 10646:2003.

| | | |
|------|----|------------------------|
| 097E | ड़ | DEVANAGARI LETTER DDDA |
| 097F | ब़ | DEVANAGARI LETTER BBA |

| | 1CD | 1CE | 1CF |
|---|---|---|---|
| 0 |  1CD0 |  1CE0 |  1CF0 |
| 1 |  1CD1 |  1CE1 |  1CF1 |
| 2 |  1CD2 |  1CE2 |  1CF2 |
| 3 |  1CD3 |  1CE3 |  |
| 4 |  1CD4 |  1CE4 |  |
| 5 |  1CD5 |  1CE5 |  |
| 6 |  1CD6 |  1CE6 |  |
| 7 |  1CD7 |  1CE7 |  |
| 8 |  1CD8 |  1CE8 |  |
| 9 |  1CD9 |  1CE9 |  |
| A |  1CDA |  1CEA |  |
| B |  1CDB |  1CEB |  |
| C |  1CDC |  1CEC |  |
| D |  1CDD |  1CED |  |
| E |  1CDE |  1CEE |  |
| F |  1CDF |  1CEF |  |

Tone marks for the Samaveda

- 1CD0 ◌̇ VEDIC TONE KARSHANA
= vaidika saamasvara karshanna
- 1CD1 ◌̈ VEDIC TONE SHARA
= vaidika svarita uurdhva shara
- 1CD2 ◌̄ VEDIC TONE PRENKHA
indicates vibrato
= vaidika saamasvara prenkha
- 1CD3 ~ VEDIC SIGN NIHSHVASA
= vaidika saamagaana yogakaala

Signs for Yajurvedic

- 1CD4 ♂ VEDIC SIGN YAJURVEDIC KASHMIRI SVARITA
also used in the Vajasaneyi Madhyandina Samhita
- 1CD5 ♀ VEDIC TONE YAJURVEDIC AGGRAVATED INDEPENDENT SVARITA
= vaidika svarita adho nyubja
- 1CD6 ♀ VEDIC TONE YAJURVEDIC INDEPENDENT SVARITA
= vaidika svarita adhah konna
- 1CD7 ♀ VEDIC TONE YAJURVEDIC KATHAKA INDEPENDENT SVARITA
= vaidika svarita adho vakra rekhaa
- 1CD8 ♀ VEDIC TONE CANDRA BELOW
= vaidika svarita adho'rdha vakra
- 1CD9 ♀ VEDIC TONE YAJURVEDIC KATHAKA INDEPENDENT SVARITA SCHROEDER
= vaidika svarita adhah samyukta rekhaa
- 1CDA ◌̇ VEDIC TONE DOUBLE SVARITA
= vaidika svarita uurdhva dvi rekhaa
- 1CDB ◌̈ VEDIC TONE TRIPLE SVARITA
= vaidika svarita uurdhva tri rekhaa
- 1CDC ◌̄ VEDIC TONE KATHAKA ANUDATTA
= vaidika svarita adho rekhaa
- 1CDD ◌̇ VEDIC TONE DOT BELOW
= vaidika svarita adho bindu

Tone marks for the Satapathabrahmana

- 1CDE ◌̇ VEDIC TONE TWO DOTS BELOW
= vaidika svarita adho dvi bindu
- 1CDF ◌̇ VEDIC TONE THREE DOTS BELOW
= vaidika svarita adhas tri bindu

Tone mark for the Rigveda

- 1CE0 ♂ VEDIC TONE RIGVEDIC KASHMIRI INDEPENDENT SVARITA
= vaidika uurdhva vakra rekhaa

Tone mark for the Atharvaveda

- 1CE1 ♂ VEDIC TONE ATHARVAVEDIC INDEPENDENT SVARITA
= vaidika svarita dvi vakra khaanda

Diacritics for visarga

- 1CE2 ♂ VEDIC SIGN VISARGA SVARITA
= vaidika madhyarekhaa
- 1CE3 ♂ VEDIC SIGN VISARGA UDATTA
= vaidika visarga dakshinnatah uurdhvaga
- 1CE4 ♂ VEDIC SIGN REVERSED VISARGA UDATTA
= vaidika visarga vaamatah uurdhvaga

- 1CE5 ♂ VEDIC SIGN VISARGA ANUDATTA
= vaidika visarga vaamatah adhoga

- 1CE6 ♂ VEDIC SIGN REVERSED VISARGA ANUDATTA
= vaidika visarga dakshinnatah adhoga

- 1CE7 ♂ VEDIC SIGN VISARGA UDATTA WITH TAIL
= vaidika visarga dakshinnatah uurdhva vakra

- 1CE8 ♂ VEDIC SIGN VISARGA ANUDATTA WITH TAIL
= vaidika visarga vaamatah adho vakra

Marks of nasalization

- 1CE9 ♂ VEDIC SIGN ANUSVARA ANTARGOMUKHA
= vaidika anusvaara antarmukha
- 1CEA ♂ VEDIC SIGN ANUSVARA BAHIRGOMUKHA
= vaidika anusvaara naagaphanna
- 1CEB ♂ VEDIC SIGN ANUSVARA VAMAGOMUKHA
= vaidika anusvaara vaamamukha
- 1CEC ♂ VEDIC SIGN ANUSVARA VAMAGOMUKHA WITH TAIL
= vaidika anusvaara vaamamukha sa-vakra
- 1CED ◌̇ VEDIC SIGN TIRYAK
= vaidika tiryak
- 1CEE ♂ VEDIC SIGN HEXIFORM LONG ANUSVARA
= vaidika anusvaara anugamii
- 1CEF ♂ VEDIC SIGN LONG ANUSVARA
= vaidika anusvaara dakshinnamukha
- 1CF0 ♂ VEDIC SIGN RTHANG LONG ANUSVARA
= vaidika anusvaara ttha-sadrisha
- 1CF1 ♂ VEDIC SIGN ANUSVARA UBHAYATO MUKHA
= vaidika anusvaara ubhayato mukha

Ardhavisarga

- 1CF2 ♂ VEDIC SIGN ARDHAVISARGA
= vaidika jihvaamuuliiya upadhmaaniya

| | A8E | A8F |
|---|-----------|-----------|
| 0 | ० A8E0 | ॠ A8F0 |
| 1 | ॡ A8E1 | ॡ A8F1 |
| 2 | ॢ A8E2 | ॢ A8F2 |
| 3 | ॣ A8E3 | ॣ A8F3 |
| 4 | । A8E4 | । A8F4 |
| 5 | ॥ A8E5 | ॥ A8F5 |
| 6 | ० A8E6 | ० A8F6 |
| 7 | १ A8E7 | १ A8F7 |
| 8 | ॡ A8E8 | ॡ A8F8 |
| 9 | ॢ A8E9 | ॢ A8F9 |
| A | अ A8EA | अ A8FA |
| B | उ A8EB | उ A8FB |
| C | क A8EC | |
| D | न A8ED | |
| E | प A8EE | |
| F | र A8EF | |

Cantillation marks for the Samaveda

| | | |
|------|----|--|
| A8E0 | ◌̐ | COMBINING DEVANAGARI DIGIT ZERO = vaidika saamasvara anka shuunya |
| A8E1 | ◌̑ | COMBINING DEVANAGARI DIGIT ONE = vaidika saamasvara anka eka udaatta |
| A8E2 | ◌̒ | COMBINING DEVANAGARI DIGIT TWO = vaidika saamasvara anka dvi svarita |
| A8E3 | ◌̓ | COMBINING DEVANAGARI DIGIT THREE = vaidika saamasvara anka tri anudaatta |
| A8E4 | ◌̔ | COMBINING DEVANAGARI DIGIT FOUR = vaidika saamasvara anka chatur |
| A8E5 | ◌̕ | COMBINING DEVANAGARI DIGIT FIVE = vaidika saamasvara anka pancha |
| A8E6 | ◌̖ | COMBINING DEVANAGARI DIGIT SIX = vaidika saamasvara anka shatt |
| A8E7 | ◌̗ | COMBINING DEVANAGARI DIGIT SEVEN = vaidika saamasvara anka sapta |
| A8E8 | ◌̘ | COMBINING DEVANAGARI DIGIT EIGHT = vaidika saamasvara anka ashta |
| A8E9 | ◌̙ | COMBINING DEVANAGARI DIGIT NINE = vaidika saamasvara anka nava |
| A8EA | ◌̚ | COMBINING DEVANAGARI LETTER A = vaidika saamasvara abhinihita |
| A8EB | ◌̛ | COMBINING DEVANAGARI LETTER U = vaidika saamasvara u |
| A8EC | ◌̜ | COMBINING DEVANAGARI LETTER KA = vaidika saamasvara ka |
| A8ED | ◌̝ | COMBINING DEVANAGARI LETTER NA = vaidika saamasvara namana |
| A8EE | ◌̞ | COMBINING DEVANAGARI LETTER PA = vaidika saamasvara prannatam |
| A8EF | ◌̟ | COMBINING DEVANAGARI LETTER RA = vaidika saamasvara ra |
| A8F0 | ◌̠ | COMBINING DEVANAGARI LETTER VI = vaidika saamasvara vinata |
| A8F1 | ◌̡ | COMBINING DEVANAGARI SIGN AVAGRAHA = vaidika saamasvara diirghiibhaava |

| | | |
|------|----|--|
| A8F9 | ◌̢ | DEVANAGARI GAP FILLER used to indicate that a gap in a manuscript or text is not a lacuna |
| A8FA | ◌̣ | DEVANAGARI CARET = vaidika trutikaa zero-advance character centred on the point between two orthographic syllables |
| A8FB | ◌̤ | DEVANAGARI HEADSTROKE = shirorekhaa used to indicate uncertain manuscript readings |

Marks of nasalization

| | | |
|------|----|--|
| A8F2 | ◌̥ | DEVANAGARI SIGN SPACING CANDRABINDU = vaidika candrabindu |
| A8F3 | ◌̦ | DEVANAGARI SIGN CANDRABINDU VIRAMA = vaidika anusvaara candrabindu tiryak |
| A8F4 | ◌̧ | DEVANAGARI SIGN DOUBLE CANDRABINDU VIRAMA = vaidika anusvaara dvi candrabindu tiryak |
| A8F5 | ◌̨ | DEVANAGARI SIGN CANDRABINDU TWO = vaidika anusvaara candrabindu sa-dvi |
| A8F6 | ◌̩ | DEVANAGARI SIGN CANDRABINDU THREE = vaidika anusvaara candrabindu sa-tri |
| A8F7 | ◌̪ | DEVANAGARI SIGN CANDRABINDU AVAGRAHA = vaidika anusvaara candrabindu sa-avagraha |

Editorial marks

| | | |
|------|----|--|
| A8F8 | ◌̫ | DEVANAGARI SIGN PUSHPIKA = vaidika pushpikaa used as a placeholder or “filler” often flanked by double dandas |
|------|----|--|

A. Administrative

1. Title

Proposal to encode additional characters for Vedic in the UCS

2. Requester's name

Michael Everson and Peter Scharf

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

Individual contribution.

4. Submission date

2008-08-21

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)

No.

1b. Proposed name of script

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

Yes

1d. Name of the existing block

Devanagari, Vedic Extensions, Devanagari Extended

2. Number of characters in proposal

8 new characters.

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 A.

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)?

Yes.

8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. 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; cf. N3366 and 3383R.

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 (editor), Michel Angot, R. Chandrashekar, Malcolm Hyman, Susan Rosenfield, B. V. Venkatakrisna Sastry, Michael Witzel

2c. If YES, available relevant documents

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

Sanskritists and other scholars.

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

Used historically and in modern editions.

4b. Reference

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

Yes.

5b. If YES, where?

Scholarly 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 similar 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?

No.

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

10c. If YES, reference

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?