

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

Doc Type: Working Group Document
Title: Final revised proposal to encode the Saurashtra script in the UCS
Source: Michael Everson and Jeyakumar Chinnakkonda Krishnamoorthy
Status: Individual Contribution
Date: 2005-08-09

Introduction

Saurashtra is an Indo-European language, related to Gujarati and spoken by about 310,000 people (1997 India Missions Association) in southern India, mostly in the area around Madurai, Salem, and Thanjavur cities. The SIL *Ethnologue* describes it as “an Indo-European island surrounded by Dravidian languages”. Since the end of the 19th century the Telugu, Tamil, Devanagari, and Saurashtra scripts have been used to publish books in Saurashtra. At present, the Tamil script is most commonly used to write Saurashtra, augmented with the use of superscript digits and a colon to indicate sounds not available in the Tamil script.

The Saurashtra script is of the Brahmic type; its modern version was developed in the late 1800s. Like Tamil, Saurashtra does not make use of complex consonant clusters, preferring instead to mark the killed vowel with a visible subscript virama. An exception to this is the conjunct ્kṣa, which is formed of KA + VIRAMA + SSA in this encoding, and which is sorted as a unique letter in older dictionaries. Early Saurashtra text makes use of additional conjuncts, which can be handled with the usual Brahmic shaping rules. So although, apart from ્kṣa, the virama is always visible in modern Saurashtra, it is best to encode modern Saurashtra with ZERO WIDTH NON-JOINER making the virama visible, rather than to consider it a general visible mark. This supports both older and newer Saurashtra text. Input methods may make it easy to type both the VIRAMA and the ZWNJ with a single keystroke.

Collating order

The usual Sanskrit ordering applies to Saurashtra: A AA I II U UU RU RUU LU LUU E EE AI O OO AU AM AH KA KHA GA GHA NGA CA CHA JA JHA NYA TTA TTHA DDA DDHA NNA TA THA DA DHA NA PA PHA BA BHA MA YA RA LA VA SHA SSA SA HA LLA KSSA

Glyph placement

The vowel signs in Saurashtra follow the consonant to which they are applied. The vowels *-i* and *-ī*, however, are typographically joined to the top right corner of their consonant. Examples: ્ka + ્i = ્ki (not ્kī), ્ka + ્ī = ્kī (not ્kī), and ્ṭa + ્i = ્ṭi (not ્ṭī), ્ṭa + ્ī = ્ṭī (not ્ṭī).

Punctuation and digits

In the samples viewed at present, European COMMA, FULL STOP, and QUESTION MARK are used, most commonly. Both SAURASHTRA DANDA and SAURASHTRA DOUBLE DANDA are used in poetry; SAURASHTRA DANDA is also used in prose. Digits have distinctive forms.

Saurashtra UPAKSHARA

A letter **𑌨** called UPAKSHARA is unique to Saurashtra, and does not have an equivalent in the Devanagari, Tamil, or Telugu scripts. It functions like the Tamil *āytam*, in that it is used to modify other letters to represent sounds not found in the basic Brahmic alphabet. In a chart from *Bashabhimāni*, four Saurashtra letters are shown in combination with this unidentified letter, which we transliterate *H* here: **𑌨𑌢** *rH*, **𑌨𑌤** *lH*, **𑌨𑌦** *mH*, **𑌨𑌨** *nH*. In this chart, the visible VIRAMA is used after the liquids and nasals and before the UPAKSHARA (and the Tamil letters are shown with the TAMIL SIGN VIRAMA) as reproduced here; but in running text the visible virama is not shown. The related language Marathi also has aspirated liquids and nasal such as these, which are shown as conjuncts with *ha*: **𑌨𑌢** *rha*, **𑌨𑌤** *lha*, **𑌨𑌦** *mha*, **𑌨𑌨** *nha*. Saurashtra UPAKSHARA functions as an independent letter, killing the inherent vowel of the letter to which it is applied; it also takes the dependent vowel when used. In the publication *Padārthun*, the words **𑌨𑌢𑌢𑌢** *nHāna* and **𑌨𑌢𑌢𑌢** *nHīna* are transliterated into Tamil-based orthography as **𑌢𑌢𑌢𑌢** *nā:na* and **𑌢𑌢𑌢𑌢** *nī:na*. **𑌨𑌢𑌢𑌢** *mHor* and **𑌨𑌢𑌢𑌢** *nHannō* are transliterated as **𑌢𑌢𑌢𑌢** *mo:r* and **𑌢𑌢𑌢𑌢** *na:nnō*. The colon-like mark is used also in Tamil transliterations of this letter in *Bhāṣābhimāni*, where **𑌨𑌢𑌢𑌢** *nHittak* is written **𑌢𑌢𑌢𑌢** *nī:ttak*. This colon-like mark is not encoded in the Tamil block, and the exact meaning of either it or the Saurashtra letter is uncertain. Another example of a word written with the UPAKSHARA is **𑌨𑌢𑌢𑌢** *nHīnāv* (note again there is no visible VIRAMA drawn under the first letter). The Tamil transliteration given is **𑌢𑌢𑌢𑌢** *nHīnāv*, apparently with the wrong vowel in the first syllable, and with the (as yet unencoded for Tamil) colon. The Tamil translation given is **𑌢𑌢𑌢𑌢** *vazakkai* ‘bald’. The Hindi word for this is given (in Tamil script!) as **𑌢𑌢𑌢𑌢** *ka³ñjā*, i.e., *gañjā* ‘bald’. In L2/03-256 Daniel Kai has suggested that the aspirates be encoded as digraphs, but this is unnecessary. Often considerable space can be seen between the liquid or nasal and the UPAKSHARA.

In 2003, Peri Bhaskararao suggested that the UPAKSHARA could be analyzed as a combining form of HA. No additional information in support of that position has been provided to overturn the analysis we have given above. Compare again Saurashtra **𑌨𑌢** *rH*, with its Tamil transliteration **𑌢𑌢** *rh* shown above. For Saurashtra were exactly equivalent to that, we would see **𑌨𑌢** *rh*. The equivalent Devanagari sequence would be **𑌨𑌢** *rh* with two viramas; all three of these would be encoded RA + VIRAMA + ZWNJ + HA + VIRAMA + ZWNJ. For **𑌨𑌢** to be equivalent to Devanagari **𑌨𑌢** or **𑌨𑌢**, it would have to be encoded RA + VIRAMA + ZWNJ + VIRAMA + HA + VIRAMA + ZWNJ. Accordingly, now, in 2005, we reiterate our proposal for UPAKSHARA to be encoded as a unique Saurashtran character.

Glyph variants

The shape of the vowel signs for *-l̄* and *-l̄* is noteworthy. In one teaching book it appears that full-size glyphs are used, with a tilde beneath marking the long vowel: **𑌢𑌢** *kl̄*, **𑌢𑌢** *k̄l̄* (compare **𑌢𑌢** *kr̄* and **𑌢𑌢** *k̄r̄*). In earlier sources, however, the glyph for *-l̄* is reduced, and length is marked with *-ā*; the *Sangudhari* font supports this practice: **𑌢𑌢** *kl̄*, **𑌢𑌢** *k̄l̄* (compare *Sangudhari* **𑌢𑌢** *kr̄* and **𑌢𑌢** *k̄r̄*).

Implementations

There are a number TrueType implementations of the modern Saurashtra script in circulation on the internet. To date, no computer implementation of the older version of the script has been created. Previously, a letterpress was used to print *Bhashabimani* magazine. As printing technology advanced, the need for a Saurashtra script font was felt, and was created by local publishers. The *Suresh* font, designed by Suresh Thimma Ramdas, is used to print *Bhāṣābhimāni*; a later corrected version of that font is *Kuber*, which is used here. (The name *Kuber* is derived from “Kubendiran”, the name of the editor of this magazine.)

References

Norihiko Uchida, *Language of the Saurashtrans in Tirupati*. (In Latin script.)

Norihiko Uchida. *Saurashtra-English dictionary*. ISBN 3447030550. (In Latin script.) (Uchida is a Japanese scholar who went to Madurai to research Saurashtra. He did his research at Madurai Kamarajar University. The local language in Madurai is Tamil; in the course of his research, he learned Saurashtra well and is said to speak Saurashtra as well as a native Saurashtrian.)

சுப்ரமணியன், ஸ்ரீ தாடா. 2000. *ஸவ்ராஷ்ட்ர பாத்² மாலொ*. மதுரை: ப்ரியா பதப்பகம் ப்ரியா டெக்னிகல் இன்ஸ்டிடியூட். [Śrī Tāṭā. Cupramaniyaṇ. 2000. *Savrāṣṭra pātha mālo = Savrāṣṭra pāt²a mālo*. Madurai: Priyā Patappakam Priyā Teknikal Inṣṭīyut]

ભહાશભિમાની *Bhāṣābhīmāni*.

Unicode Character Properties

Rick McGowan provided the following text in L2/03-225.

The Saurashtra characters have properties similar to Devanagari and the other Indic scripts. Since Saurashtra uses no truly non-spacing marks, all of the combining classes are zero, with the exception of the virama. All of the letters are in the Letter other category, and have strong left-to-right directionality. The vowel signs all follow the consonants with which they are associated, and therefore are combining class zero as well, analogous to the Devanagari model. The virama has the same combining class (9) as other virama characters. The digits likewise are analogous to the Devanagari and other digits, with the same properties, including decimal digit values. The table below shows all of the properties defined in the UnicodeData.txt file.

```
AB80;SAURASHTRA SIGN ANUSVARA;Mc;0;L;;;;N;;;;;
AB81;SAURASHTRA SIGN VISARGA;Mc;0;L;;;;N;;;;;
AB82;SAURASHTRA LETTER A;Lo;0;L;;;;N;;;
AB83;SAURASHTRA LETTER AA;Lo;0;L;;;;N;;;
AB84;SAURASHTRA LETTER I;Lo;0;L;;;;N;;;
AB85;SAURASHTRA LETTER II;Lo;0;L;;;;N;;;
AB86;SAURASHTRA LETTER U;Lo;0;L;;;;N;;;
AB87;SAURASHTRA LETTER UU;Lo;0;L;;;;N;;;
AB88;SAURASHTRA LETTER VOCALIC R;Lo;0;L;;;;N;;;
AB89;SAURASHTRA LETTER VOCALIC RR;Lo;0;L;;;;N;;;
AB8A;SAURASHTRA LETTER VOCALIC L;Lo;0;L;;;;N;;;
AB8B;SAURASHTRA LETTER VOCALIC LL;Lo;0;L;;;;N;;;
AB8C;SAURASHTRA LETTER E;Lo;0;L;;;;N;;;
AB8D;SAURASHTRA LETTER EE;Lo;0;L;;;;N;;;
AB8E;SAURASHTRA LETTER AI;Lo;0;L;;;;N;;;
AB8F;SAURASHTRA LETTER O;Lo;0;L;;;;N;;;
AB90;SAURASHTRA LETTER OO;Lo;0;L;;;;N;;;
AB91;SAURASHTRA LETTER AU;Lo;0;L;;;;N;;;
AB92;SAURASHTRA LETTER KA;Lo;0;L;;;;N;;;
AB93;SAURASHTRA LETTER KHA;Lo;0;L;;;;N;;;
AB94;SAURASHTRA LETTER GA;Lo;0;L;;;;N;;;
AB95;SAURASHTRA LETTER GHA;Lo;0;L;;;;N;;;
AB96;SAURASHTRA LETTER NGA;Lo;0;L;;;;N;;;
AB97;SAURASHTRA LETTER CA;Lo;0;L;;;;N;;;
AB98;SAURASHTRA LETTER CHA;Lo;0;L;;;;N;;;
AB99;SAURASHTRA LETTER JA;Lo;0;L;;;;N;;;
AB9A;SAURASHTRA LETTER JHA;Lo;0;L;;;;N;;;
AB9B;SAURASHTRA LETTER NYA;Lo;0;L;;;;N;;;
AB9C;SAURASHTRA LETTER TA;Lo;0;L;;;;N;;;
AB9D;SAURASHTRA LETTER TTHA;Lo;0;L;;;;N;;;
AB9E;SAURASHTRA LETTER DDA;Lo;0;L;;;;N;;;
AB9F;SAURASHTRA LETTER DDHA;Lo;0;L;;;;N;;;
ABA0;SAURASHTRA LETTER NNA;Lo;0;L;;;;N;;;
ABA1;SAURASHTRA LETTER TA;Lo;0;L;;;;N;;;;
```

ABA2;SAURASHTRA LETTER THA;Lo;0;L;;;;N;;;;
 ABA3;SAURASHTRA LETTER DA;Lo;0;L;;;;N;;;;
 ABA4;SAURASHTRA LETTER DHA;Lo;0;L;;;;N;;;;
 ABA5;SAURASHTRA LETTER NA;Lo;0;L;;;;N;;;;
 ABA6;SAURASHTRA LETTER PA;Lo;0;L;;;;N;;;;
 ABA7;SAURASHTRA LETTER PHA;Lo;0;L;;;;N;;;;
 ABA8;SAURASHTRA LETTER BA;Lo;0;L;;;;N;;;;
 ABA9;SAURASHTRA LETTER BHA;Lo;0;L;;;;N;;;;
 ABAA;SAURASHTRA LETTER MA;Lo;0;L;;;;N;;;;
 ABAB;SAURASHTRA LETTER YA;Lo;0;L;;;;N;;;;
 ABAC;SAURASHTRA LETTER RA;Lo;0;L;;;;N;;;;
 ABAD;SAURASHTRA LETTER LA;Lo;0;L;;;;N;;;;
 ABAE;SAURASHTRA LETTER VA;Lo;0;L;;;;N;;;;
 ABAF;SAURASHTRA LETTER SHA;Lo;0;L;;;;N;;;;
 ABB0;SAURASHTRA LETTER SSA;Lo;0;L;;;;N;;;;
 ABB1;SAURASHTRA LETTER SA;Lo;0;L;;;;N;;;;
 ABB2;SAURASHTRA LETTER HA;Lo;0;L;;;;N;;;;
 ABB3;SAURASHTRA LETTER LLA;Lo;0;L;;;;N;;;;
 ABB4;SAURASHTRA LETTER UPAKSHARA;Lo;0;L;;;;N;;;;
 ABB5;SAURASHTRA VOWEL SIGN AA;Mc;0;L;;;;N;;;;
 ABB6;SAURASHTRA VOWEL SIGN I;Mc;0;L;;;;N;;;;
 ABB7;SAURASHTRA VOWEL SIGN II;Mc;0;L;;;;N;;;;
 ABB8;SAURASHTRA VOWEL SIGN U;Mc;0;L;;;;N;;;;
 ABB9;SAURASHTRA VOWEL SIGN UU;Mc;0;L;;;;N;;;;
 ABBA;SAURASHTRA VOWEL SIGN VOCALIC R;Mc;0;L;;;;N;;;;
 ABBB;SAURASHTRA VOWEL SIGN VOCALIC RR;Mc;0;L;;;;N;;;;
 ABBC;SAURASHTRA VOWEL SIGN VOCALIC L;Mc;0;L;;;;N;;;;
 ABBD;SAURASHTRA VOWEL SIGN VOCALIC LL;Mc;0;L;;;;N;;;;
 ABBE;SAURASHTRA VOWEL SIGN E;Mc;0;L;;;;N;;;;
 ABBF;SAURASHTRA VOWEL SIGN EE;Mc;0;L;;;;N;;;;
 ABC0;SAURASHTRA VOWEL SIGN AI;Mc;0;L;;;;N;;;;
 ABC1;SAURASHTRA VOWEL SIGN O;Mc;0;L;;;;N;;;;
 ABC2;SAURASHTRA VOWEL SIGN OO;Mc;0;L;;;;N;;;;
 ABC3;SAURASHTRA VOWEL SIGN AU;Mc;0;L;;;;N;;;;
 ABC4;SAURASHTRA SIGN VIRAMA;Mn;9;NSM;;;;N;;;;
 ABC5;SAURASHTRA DIGIT ZERO;Nd;0;L;;0;0;0;N;;;;
 ABC7;SAURASHTRA DIGIT ONE;Nd;0;L;;1;1;1;N;;;;
 ABC8;SAURASHTRA DIGIT TWO;Nd;0;L;;2;2;2;N;;;;
 ABC9;SAURASHTRA DIGIT THREE;Nd;0;L;;3;3;3;N;;;;
 ABCA;SAURASHTRA DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;
 ABCB;SAURASHTRA DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;
 ABCC;SAURASHTRA DIGIT SIX;Nd;0;L;;6;6;6;N;;;;
 ABCD;SAURASHTRA DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;
 ABCE;SAURASHTRA DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;
 ABCF;SAURASHTRA DIGIT NINE;Nd;0;L;;9;9;9;N;;;

Figures

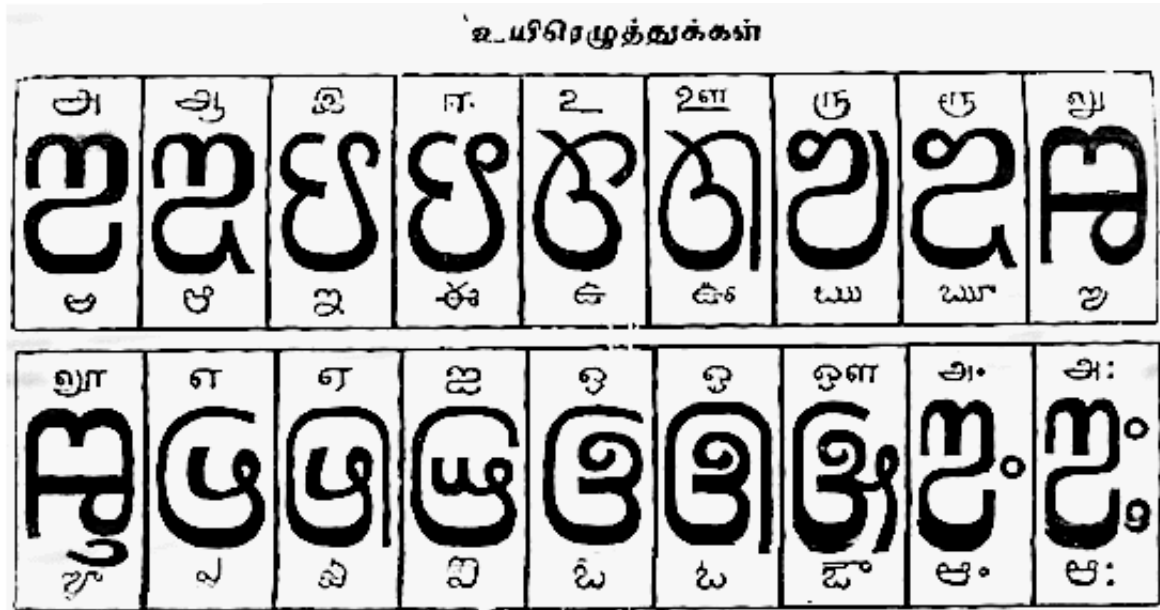


Figure 1. Saurashtra vowels, with equivalents in Tamil and Telugu scripts.

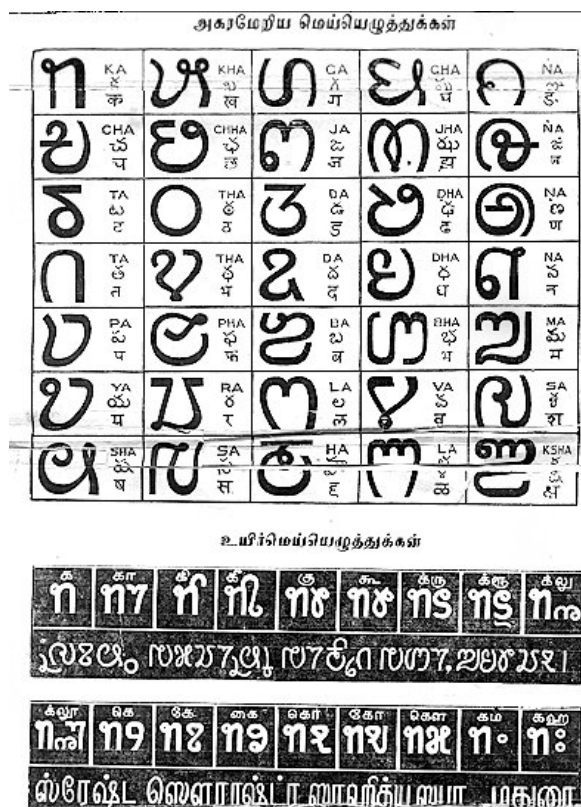


Figure 2. The text in the bottom half of this figure is given in Saurashtra and Tamil script; the Saurashtra appears to be in an early form of the script as it shows conjunct consonants. It is rendered here with modern orthography and transliteration:

ஸ்ரேஷ்ட ஸௌராஷ்டிர ஸாஹித்ய ஸபா, மதுரை
 śrēṣṭha saurāṣṭra sāhitya sabhā, machurai
 ஸ்ரேஷ்ட ஸௌராஷ்டிர ஸாஹித்ய ஸபா, மதுரை
 śrēṣṭha saurāṣṭra sāhitya sapā, maturai

அஜாது

[illegible]

ஒபந்

௩ ^{க¹} க	௪ ^{க²} க க	௫ ^{க³} க க	௬ ^{க⁴} க க	௭ ^{க⁵} க க
௮ ^{ச¹} ச	௯ ^{ச²} ச ச	௧௦ ^{ச³} ச ச	௧௧ ^{ச⁴} ச ச	௧௨ ^{ச⁵} ச ச
௧௩ ^{ட¹} ட	௧௪ ^{ட²} ட ட	௧௫ ^{ட³} ட ட	௧௬ ^{ட⁴} ட ட	௧௭ ^{ட⁵} ட ட
௧௮ ^{த¹} த	௧௯ ^{த²} த த	௨௦ ^{த³} த த	௨௧ ^{த⁴} த த	௨௨ ^{த⁵} த த
௨௩ ^{ப¹} ப	௨௪ ^{ப²} ப ப	௨௫ ^{ப³} ப ப	௨௬ ^{ப⁴} ப ப	௨௭ ^{ப⁵} ப ப
௨௮ ^{ய¹} ய	௨௯ ^{ய²} ய ய	௩௦ ^{ய³} ய ய	௩௧ ^{ய⁴} ய ய	௩௨ ^{ய⁵} ய ய
௩௩ ^{ஷ¹} ஷ	௩௪ ^{ஸ¹} ஸ	௩௫ ^{ஹ¹} ஹ	௩௬ ^{ள¹} ள	௩௭ ^{ஶ¹} ஶ

Figure 3. Saurashtra letters, with equivalents in Tamil and Devanagari scripts.

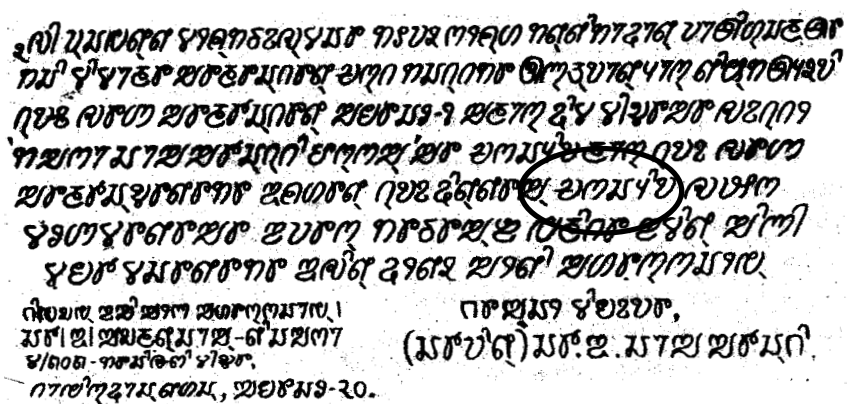


Figure 4. Handwritten pamphlet showing the UPAKSHARA character. The word is **ચળરહિયા** *calarhiya*. The SAURASHTRA DANDA is used three times in the small text to the lower left.

தருது	
க ன க	க்லா னௌ -
கா ன௪ கா	கெ ன௬ -
கி னீ கி	கே ன௪ கே
கீ னி கி	கை ன௬ கை
கு ன௪ கு	கொ ன௪ -
கூ ன௪ கூ	கோ ன௪ கோ
க்ரு ன௪ க்ரு	கௌ ன௪ கௌ
க்ரு ன௪ -	கம் ன௪ க்
க்லு ன௪ -	கஹ ன௪ க:

Figure 5. Saurashtra vowels, with equivalents in Tamil and Devanagari scripts.

**ஸவ்ராஷ்ட்ர மாதா
ஸ௪ஸ௪ஸ௪ ஸ௪ஸ௪**

1. ஸ௪ ஸ௪ ஸ௪ஸ௪ஸ௪ ஸ௪ஸ௪ |
ஸ௪ ஸ௪ ஸ௪ஸ௪ஸ௪ ஸ௪ஸ௪ ||
2. ஸ௪ஸ௪ ஸ௪ஸ௪ஸ௪ ஸ௪ ஸ௪ஸ௪ ஸ௪ஸ௪
ஸ௪ஸ௪ ஸ௪ஸ௪ ஸ௪ஸ௪ ஸ௪ஸ௪ ஸ௪ஸ௪ |
ஸ௪ஸ௪ ஸ௪ஸ௪ ஸ௪ ஸ௪ஸ௪ ஸ௪ஸ௪ ஸ௪ஸ௪
ஸ௪ஸ௪ஸ௪ ஸ௪ஸ௪ ஸ௪ஸ௪ ஸ௪ஸ௪ ஸ௪ஸ௪ (ஸ௪ஸ௪) ||
3. ஸ௪ஸ௪ஸ௪ ஸ௪ஸ௪ஸ௪ ஸ௪ஸ௪ஸ௪ ஸ௪ஸ௪ஸ௪
ஸ௪ஸ௪ ஸ௪ ஸ௪ஸ௪ ஸ௪ஸ௪ஸ௪ ஸ௪ஸ௪ |
ஸ௪ஸ௪ ஸ௪ஸ௪ ஸ௪ ஸ௪ஸ௪ஸ௪ ஸ௪ஸ௪ ஸ௪ஸ௪
ஸ௪ஸ௪ஸ௪ ஸ௪ஸ௪ ஸ௪ஸ௪ ஸ௪ஸ௪ ஸ௪ஸ௪ (ஸ௪ஸ௪) ||

- 1 ஜய ஜய ஸவ்ராஷ்ட்ர மாதா,
ஜய ஜய ஸவ்ராஷ்ட்ர மாதா.
- 2 ஸௌவு தே³ஷுனுக் தூ ராணி மாதா
ஸு¹ரிது குலம் தி³யெ ஸவ்ராஷ்ட்ர மாதா.
ஸௌவு ஜெ²லும் க²டி³ தொகொ நமு மாதா
ஸௌந்தம்மு ப²ய்லொ தூஸ் ஸவ்ராஷ்ட்ர மாதா (ஜய ஜய)
- 3 காவ்யமு பூ⁴ஷுனுன் ஸஸர் ஸஸர் கெ¹ரிகின்
க¹லுஸு மீ தொகொ ஸவ்ராஷ்ட்ர மாதா.
ஜீ¹வ்யமு லெக்கு யே, தெ⁴ய்ரியம் ஜு²க்கு ஸே
ஜெ³துரு ப²ய்லொ தூஸ் ஸவ்ராஷ்ட்ர மாதா (ஜய ஜய)

Figure 6. Saurashtra poem with both SAURASHTRA DANDA and SAURASHTRA DANDA. A Tamil transliteration of the poem is shown below.

ரோமன் எழுத்து - ஸௌராஷ்டிர மொழி

- ஜாபாலி

ரோமன் (ஆங்கில) எழுத்துகளைக் கொண்டு ஸௌராஷ்டிர மொழியை எழுதப் பயன்படுத்தும் முறை. வெளிநாடுகளில் வாழும் ஸௌராஷ்டிரர்கள் இம் முறையினைப் பயன்படுத்தி தம் உறவினர்களுக்கும் நண்பர்களுக்கும் கடிதம் மற்றும் E-mail-களை அனுப்பலாம். ஏன், இங்குள்ளவர்களும் வயது முதிர்ந்த - புது வகை வரி வடிவத்தினைக் கற்றுக் கொள்ள முடியாது என்று நம்புவர்கள் - இதனைப் பயன்படுத்தலாமே.

VOWELS - உயிரெழுத்துகள்

ஐ	ஐ	ஐ	ஐ	ஐ	ஐ	ஐ	ஐ	ஐ	ஐ
அ	ஆ	இ	ஈ	உ	ஊ	எ	ஏ	ஒ	ஓ
A	Ā	I	Ī	U	Ū	E	Ē	O	Ō

CONSONANTS - மெய்யெழுத்துகள்

க	க்	க்	க்	ங்	ச்	ச்	ஜ்	ஜ்	ஞ்
K	Kh	G	Gh	Ng	C	Ch	J	Jh	Ņg
த்	த்	த்	த்	ண்	த்	த்	த்	த்	ந்
T	Th	D	Dh	N	T	Th	D	Dh	N
ப்	ப்	ப்	ப்	ம்	ய்	ர்	ல்	வ்	ஸ்
P	Ph	B	Bh	M	Y	R	L	V	S
ஷ்	ஸ்	ஹ்	ள்	க்ஷ்	ர்ஹ்	ல்ஹ்	ம்ஹ்	ந்ஹ்	
Sh	S	H	L	Ksh	Rh	Lh	Mh	Nh	

CONSONANT ADDED VOWELS - உயிர் மெய்யெழுத்துகள்

க	கா	கி	கீ	கு	கூ	கெ	கே	கொ	கோ
KA	KĀ	KI	KĪ	KU	KŪ	KE	KĒ	KO	KŌ

EXAMPLES - உதாரணங்கள்

பாஷாபிமானி	இயெ அநுசரி அம்தெநு லேகுந் லிக்குவொ
BhĀShĀBhīmāṇi	IYE ANUCARI AMTENU LĒKhUN LIKKUVO
ஸவ்ராஷ்டிர மணி	ஒண்டெதிம் ஸிக்குல்டுவாயி
SAVRĀShĪTRA MAṇi	OṇTĒDIM SIKKULĀUVĀYi

தீபாவளி மசை 2001

Figure 7. Saurashtra vowels, with equivalents in Tamil. The UPAKSHARA character is shown in combination with killed consonants, indicated with the arrow.

Row A8: SAURASHTRA

	A88	A89	A8A	A8B	A8C	A8D
0	◌̣	ଐ	ଐ	ଐ	ଐ	୦
1	◌̤	ଐ	ଠ	ଠ	ଠ	୧
2	ଐ	ଐ	ଐ	ଐ	ଐ	୨
3	ଐ	ଐ	ଐ	ଐ	ଐ	୩
4	ଐ	ଐ	ଐ	୦	୍	୪
5	ଐ	ଐ	ଐ	ଐ		୫
6	ଐ	ଐ	ଐ	ଐ		୬
7	ଐ	ଐ	ଐ	ଐ		୭
8	ଐ	ଐ	ଐ	ଐ		୮
9	ଐ	ଐ	ଐ	ଐ		୯
A	ଐ	ଐ	ଐ	ଐ		
B	ଐ	ଐ	ଐ	ଐ		
C	ଐ	ଐ	ଐ	ଐ		
D	ଐ	ଐ	ଐ	ଐ		
E	ଐ	ଐ	ଐ	ଐ	।	
F	ଐ	ଐ	ଐ	ଐ	॥	

G = 00
P = 00

Row A8: SAURASHTRA

hex	Name	hex	Name
A880	SAURASHTRA SIGN ANUSVARA	A8D9	SAURASHTRA DIGIT NINE
A881	SAURASHTRA SIGN VISARGA	A8DA	(This position shall not be used)
A882	SAURASHTRA LETTER A	A8DB	(This position shall not be used)
A883	SAURASHTRA LETTER AA	A8DC	(This position shall not be used)
A884	SAURASHTRA LETTER I	A8DD	(This position shall not be used)
A885	SAURASHTRA LETTER II	A8DE	(This position shall not be used)
A886	SAURASHTRA LETTER U	A8DF	(This position shall not be used)
A887	SAURASHTRA LETTER UU		
A888	SAURASHTRA LETTER VOCALIC R		
A889	SAURASHTRA LETTER VOCALIC RR		
A88A	SAURASHTRA LETTER VOCALIC L		
A0BB	SAURASHTRA LETTER VOCALIC LL		
A88C	SAURASHTRA LETTER E		
A88D	SAURASHTRA LETTER EE		
A88E	SAURASHTRA LETTER AI		
A88F	SAURASHTRA LETTER O		
A890	SAURASHTRA LETTER OO		
A891	SAURASHTRA LETTER AU		
A892	SAURASHTRA LETTER KA		
A893	SAURASHTRA LETTER KHA		
A894	SAURASHTRA LETTER GA		
A895	SAURASHTRA LETTER GHA		
A896	SAURASHTRA LETTER NGA		
A897	SAURASHTRA LETTER CA		
A898	SAURASHTRA LETTER CHA		
A899	SAURASHTRA LETTER JA		
A89A	SAURASHTRA LETTER JHA		
A89B	SAURASHTRA LETTER NYA		
A89C	SAURASHTRA LETTER TTA		
A89D	SAURASHTRA LETTER TTHA		
A89E	SAURASHTRA LETTER DDA		
A89F	SAURASHTRA LETTER DDHA		
A8A0	SAURASHTRA LETTER NNA		
A8A1	SAURASHTRA LETTER TA		
A8A2	SAURASHTRA LETTER THA		
A8A3	SAURASHTRA LETTER DA		
A8A4	SAURASHTRA LETTER DHA		
A8A5	SAURASHTRA LETTER NA		
A8A6	SAURASHTRA LETTER PA		
A8A7	SAURASHTRA LETTER PHA		
A8A8	SAURASHTRA LETTER BA		
A8A9	SAURASHTRA LETTER BHA		
A8AA	SAURASHTRA LETTER MA		
A8AB	SAURASHTRA LETTER YA		
A8AC	SAURASHTRA LETTER RA		
A8AD	SAURASHTRA LETTER LA		
A8AE	SAURASHTRA LETTER VA		
A8AF	SAURASHTRA LETTER SHA		
A8B0	SAURASHTRA LETTER SSA		
A8B1	SAURASHTRA LETTER SA		
A8B2	SAURASHTRA LETTER HA		
A8B3	SAURASHTRA LETTER LLA		
A8B4	SAURASHTRA LETTER UPAKSHARA		
A8B5	SAURASHTRA VOWEL SIGN AA		
A8B6	SAURASHTRA VOWEL SIGN I		
A8B7	SAURASHTRA VOWEL SIGN II		
A8B8	SAURASHTRA VOWEL SIGN U		
A8B9	SAURASHTRA VOWEL SIGN UU		
A8BA	SAURASHTRA VOWEL SIGN VOCALIC R		
A8BB	SAURASHTRA VOWEL SIGN VOCALIC RR		
A8BC	SAURASHTRA VOWEL SIGN VOCALIC L		
A8BD	SAURASHTRA VOWEL SIGN VOCALIC LL		
A8BE	SAURASHTRA VOWEL SIGN E		
A8BF	SAURASHTRA VOWEL SIGN EE		
A8C0	SAURASHTRA VOWEL SIGN AI		
A8C1	SAURASHTRA VOWEL SIGN O		
A8C2	SAURASHTRA VOWEL SIGN OO		
A8C3	SAURASHTRA VOWEL SIGN AU		
A8C4	SAURASHTRA SIGN VIRAMA		
A8C5	(This position shall not be used)		
A8C6	(This position shall not be used)		
A8C7	(This position shall not be used)		
A8C8	(This position shall not be used)		
A8C9	(This position shall not be used)		
A8CA	(This position shall not be used)		
A8CB	(This position shall not be used)		
A8CC	(This position shall not be used)		
A8CD	(This position shall not be used)		
A8CE	SAURASHTRA DANDA		
A8CF	SAURASHTRA DOUBLE DANDA		
A8D0	SAURASHTRA DIGIT ZERO		
A8D1	SAURASHTRA DIGIT ONE		
A8D2	SAURASHTRA DIGIT TWO		
A8D3	SAURASHTRA DIGIT THREE		
A8D4	SAURASHTRA DIGIT FOUR		
A8D5	SAURASHTRA DIGIT FIVE		
A8D6	SAURASHTRA DIGIT SIX		
A8D7	SAURASHTRA DIGIT SEVEN		
A8D8	SAURASHTRA DIGIT EIGHT		

A. Administrative

1. Title

Final revised proposal to encode the Saurashtra script in the UCS.

2. Requester's name

Michael Everson and Jeyakumar Chinnakkonda Krishnamoorthy

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

Individual contribution.

4. Submission date

2005-08-08

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.

Proposed name of script

Saurashtra.

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

No.

1b. Name of the existing block

2. Number of characters in proposal

79

3. Proposed category (see section II, Character Categories)

Category A.

4a. Proposed Level of Implementation (1, 2 or 3) (see clause 14, ISO/IEC 10646-1: 2000)

Level 2.

4b. Is a rationale provided for the choice?

Yes.

4c. If YES, reference

Brahmic script with vowel signs.

5a. Is a repertoire including character names provided?

Yes.

5b. If YES, are the names in accordance with the naming guidelines in Annex L of ISO/IEC 10646-1: 2000?

Yes.

5c. Are the character shapes attached in a legible form suitable for review?

Yes.

6a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Jeyakumar C. K. (www.palkar.org) & Michael Everson (Everson Typography). TrueType.

6b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Jeyakumar C. K. (www.palkar.org) & Michael Everson (Everson Typography). Fontographer.

7a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

Yes.

7b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes.

8. 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.

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

See Unicode properties above.

C. Technical – Justification

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

Yes. See N2607.

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?

Mr Kubendiran, editor of ભજાશ્ત્રીય ભાષાશાસ્ત્ર Bhāṣāśāstrīyā.

2c. If YES, available relevant documents

N/A.

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

Yes.

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

This character is used to write the Saurashtran language.

4b. Reference

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

Yes.

5b. If YES, where?

In Tamil Nadu, Andhra Pradesh, Gujarat states of India. A monthly magazine called *Bhāṣābhīmāni* is published in Madurai, India.

6a. After giving due considerations to the principles in Principles and Procedures document (a WG 2 standing 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.

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

Yes.

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?

Yes.

11c. If YES, reference

Brahmic vowel signs.

12a. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

12b. If YES, reference

13a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

13b. If YES, describe in detail (include attachment if necessary)

14a. Does the proposal contain any Ideographic compatibility character(s)?

No.

14b. If YES, is the equivalent corresponding unified ideographic character(s) identified?

14c. If YES, reference