

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: 2005-09-21

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, Saurashtra is most often written in the Tamil script, 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; early Saurashtra text makes use of conjuncts, which can be handled with the usual Brahmic shaping rules. The modern version, developed in the 1880s, has undergone some simplification. Like Tamil, modern Saurashtra does not use complex consonant clusters, marking instead the killed vowel with a visible subscript virama; **𑌕** *ka* is formed of KA + VIRAMA. An exception to this is the conjunct **𑌕𑌖** *kṣa*, which is formed of KA + VIRAMA + ZWJ + SSA in this encoding, and which is sorted as a unique letter in older dictionaries. Because, apart from *kṣa*, the virama is always visible in modern Saurashtra, it is best to encode modern Saurashtra with the virama visible by default, and to use ZERO WIDTH JOINER to force conjunct behaviour. Thus this encoding supports both older and newer Saurashtra text. A keyboard specification may make it easy to type **𑌕𑌖** *kṣa* with a single keystroke comprising all four characters KA + VIRAMA + ZWJ + SSA.

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: **𑌕** *ka* + **𑌖** *-i* = **𑌕𑌖** *ki* (not **𑌕**), **𑌕** *ka* + **𑌗** *-ī* = **𑌕𑌗** *kī* (not **𑌕**), and **𑌘** *ṭa* + **𑌖** *-i* = **𑌘𑌖** *ṭi* (not **𑌘**), **𑌘** *ṭa* + **𑌗** *-ī* = **𑌘𑌗** *ṭī* (not **𑌘**). Vowel matras are also applied to the UPAKSHARA (see below).

Digits and punctuation

Digits have distinctive forms. 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.

Glyph variants

The shape of the vowel signs for *-ḷ* and *-ḻ* is noteworthy. In one teaching book it appears that full-size glyphs are used, with a tilde-like mark beneath indicating the long vowel: **𑌕𑌘** *kḷ*, **𑌕𑌙** *kḻ* (compare **𑌕𑌚** *kṛ* and **𑌕𑌛** *kṝ*). In earlier sources, however, the glyph for *-ḷ* is reduced, and length is

marked with \bar{a} ; the *Sangudhari* font supports this practice: **𑌕𑌃** *kl*, **𑌕𑌃̄** *kḷ* (compare *Sangudhari* **𑌕𑌃** *kr* and **𑌕𑌃̄** *kṛ*).

Saurashtra UPAKSHARA

A character **𑌕** called UPAKSHARA, which we transliterate *H* here, is unique to Saurashtra, and does not have an equivalent in the Devanagari, Tamil, or Telugu scripts. It functions in some regards like the Tamil *āytam*, in that it is used to modify other letters to represent sounds not found in the basic Brahmic alphabet. It differs from *āytam* because it is a dependent consonant—a kind of consonant matra. In Figure 7, a chart from *Bhāṣābhīmāni*, four Saurashtra letters are shown in combination with the UPAKSHARA: **𑌕𑌃** *rH*, **𑌕𑌃** *lH*, **𑌕𑌃** *mH*, **𑌕𑌃** *nH*. In this chart, the visible VIRAMA is drawn below the liquids and nasals and before the UPAKSHARA, though in the encoding, it is, for example, RA + UPAKSHARA + VIRAMA. (In that same chart, the Tamil letters are shown with the Tamil *pulli* on two letters: **𑌕𑌃̄** *rh*, **𑌕𑌃̄** *lh*, **𑌕𑌃̄** *mh*, **𑌕𑌃̄** *nh*., a different transliteration and encoding model; here we have RA + VIRAMA + HA + VIRAMA.) Since in running text the visible virama is not shown, it is clear that the UPAKSHARA modifies the consonant as a whole. The related language Marathi also has aspirated liquids and nasal such as these, which are shown as conjuncts with *ha*: **𑌕𑌃** *rha*, **𑌕𑌃** *lha*, **𑌕𑌃** *mha*, **𑌕𑌃** *nha*—a different model again. Saurashtra UPAKSHARA functions as an dependent consonant sign, also taking the dependent vowel when used. Here is a the vowel repertoire in Saurashtra, Latin, and Tamil:

𑌕𑌃	𑌕𑌃	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄
<i>mH</i>	<i>mHa</i>	<i>mHā</i>	<i>mHi</i>	<i>mHī</i>	<i>mHu</i>	<i>mHū</i>	<i>mHḥ</i>	<i>mHe</i>	<i>mHai</i>	<i>mHō</i>	<i>mHo</i>	<i>mHau</i>
𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄
𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄	𑌕𑌃̄

In the publication *Padārthun*, the words **𑌕𑌃̄** *nHāna* and **𑌕𑌃̄** *nHīna* are transliterated into Tamil-based orthography as **𑌕𑌃̄** *nā:na* and **𑌕𑌃̄** *nī:na*. **𑌕𑌃̄** *mHor* and **𑌕𑌃̄** *nHannō* are transliterated as **𑌕𑌃̄** *mo:r* and **𑌕𑌃̄** *na:nnō*. The colon-like mark is used also in Tamil transliterations of this letter in *Bhāṣābhīmāni*, where **𑌕𑌃̄** *nHittak* is written **𑌕𑌃̄** *ttak*. This colon-like mark is not encoded in the Tamil block; one could consider it a new sort of VISARGA, perhaps. Another example of a word written with the UPAKSHARA is **𑌕𑌃̄** *nHīnāv*. The Tamil transliteration given is **𑌕𑌃̄** *nHīnāv*, apparently with the wrong vowel in the first syllable, and with the colon-like mark. The Tamil translation given is **𑌕𑌃̄** *vazakkai* ‘bald’. The Hindi word for this is given (in Tamil script!) as **𑌕𑌃̄** *ka³ñjā*, i.e., *gañjā* ‘bald’. In L2/03-256 Ka‘ōnohi Kai has suggested that the aspirates be encoded as digraphs, but this seems unnecessary. In some implementations, considerable space can be seen between the liquid or nasal and the UPAKSHARA, but this is probably an artifact of font technology.

In a document in 2003 (L2/03-277), 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. Though the chart in *Bhāṣābhīmāni* (shown in Figure 7) gives **𑌕𑌃̄** *rh* as the killed Tamil equivalent to Saurashtra **𑌕𑌃** *rH*, transliterations elsewhere would suggest **𑌕𑌃̄** *rh*. For Saurashtra to be equivalent to **𑌕𑌃̄** *rh*, we might expect to see **𑌕𑌃̄** *rh*, that is, killed *ra* + killed *ha*. The equivalent Devanagari sequence would be **𑌕𑌃̄** *rh* with two viramas. We believe that the evidence suggests that **𑌕𑌃** *rH* (with killed vowel) and **𑌕𑌃̄** *rha* (without killed vowel, equivalent to **𑌕𑌃̄** *rha*) would be distinct in the Saurashtra script. Consider the following transliteration opportunity. A Sanskrit phrase **𑌕𑌃̄** *Om simha Kaṇapataye namaha* appears in Tamil script on a web page at <http://sripambattisiddhar.com/tamil/poojai/vinayaga.htm>; note the consonant cluster **𑌕𑌃̄** *mha*. In Saurashtra this could be

written ॐ सिंहा गणपतये नमः *Om simha Gaṇapataye namaḥ*, assuming that Sanskrit *mha* is not the aspirate *mHa*. Contrast this with the pair ॠःॠि *mu:ri* and ॠॠॠ *mHuri*—both examples from *Bhāṣābhīmāni*. (We are aware that this argument would be stronger if we had an actual *pūja* text in Saurashtran script to show the transliteration practice for Sanskrit.) In any case, now, in 2005, we reiterate our proposal for UPAKSHARA to be encoded as a unique Saurashtran character.

Conjunct characters

As noted above, the virama is by default visible in Saurashtra, and ZWJ is added to force a conjunct; ZWNJ will also force a visible virama:

ॠ KA	+	◌̣	+	ॡ SSA	=	ॠॡ <i>kṣa</i>		
ॠ KA	+	◌̣	+	ZWJ	+	ॡ SSA	=	ॠॡ <i>kṣa</i>
ॠ KA	+	◌̣	+	ZWNJ	+	ॡ SSA	=	ॠॡ <i>kṣa</i>

Older Saurashtra texts have not been available to us, but the following conjunct forms are known as attested in Figure 2 below:

ॠ SHA	+	◌̣	+	ॠ RA	=	ॠॠ <i>śra</i>								
ॠ SHA	+	◌̣	+	ZWJ	+	ॠ RA	=	ॠॠ <i>śra</i>						
ॡ SSA	+	◌̣	+	ॠ TTHA	=	ॡॠ <i>ṣṭha</i>								
ॡ SSA	+	◌̣	+	ZWJ	+	ॠ TTHA	=	ॡॠ <i>ṣṭha</i>						
ॠ TA	+	◌̣	+	ॠ YA	=	ॠॠ <i>tya</i>								
ॠ TA	+	◌̣	+	ZWJ	+	ॠ YA	=	ॠॠ <i>tya</i>						
ॡ SSA	+	◌̣	+	ॠ TTA	+	◌̣	+	ॠ RA	=	ॡॠॠ <i>ṣṭra</i>				
ॡ SSA	+	◌̣	+	ZWJ	+	ॠ TTA	+	◌̣	+	ZWJ	+	ॠ RA	=	ॡॠॠ <i>ṣṭra</i>

An implementation of older Saurashtra will entail further study as to the specific consonant clusters used.

Implementations

Ka‘ōnohi Kai implemented Saurashtra for Mac OS X 10.2 in 2003, based on a font by Jason Glavy. This commercial product was based on my exploratory Unicode Saurashtra proposal (published at www.evertype.com/standards/iso10646/pdf/saurashtra.pdf). The encoding model and repertoire used by XenoType is a subset of the model used in *this* proposal; it is clear that the model works well for Saurashtra. There are also two 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 *Bhāṣābhīmāni* 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 *Bhāṣābhīmāni*; 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.) Traditional implementation in lead type was known from the end of the 19th century. According to Norihiko Učida:

Saurashtrians had their own beautiful and systematic script, the origin of which is not known. At the end of the last century T. M. Rama Ray reformed the script and cast types for it. He published many books in

this script. Later he introduced some changes and standardized it. which is now used by some people as Saurashtra script. Since the end of the nineteenth century books have been printed in Telugu, Tamil, Devanagari and Saurashtra scripts. When Gandhiji preached for one common script for India, one group of the community decided to abandon the Saurashtra script in favour of Devanagari and destroyed the matrix designed by Rama Ray.

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 UPAKSHARA 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

Unicode Character Properties

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.

```
A880;SAURASHTRA SIGN ANUSVARA;Mc;0;L;;;;;N;;;;;
A881;SAURASHTRA SIGN VISARGA;Mc;0;L;;;;;N;;;;;
A882;SAURASHTRA LETTER A;Lo;0;L;;;;;N;;;;;
A883;SAURASHTRA LETTER AA;Lo;0;L;;;;;N;;;;;
A884;SAURASHTRA LETTER I;Lo;0;L;;;;;N;;;;;
A885;SAURASHTRA LETTER II;Lo;0;L;;;;;N;;;;;
A886;SAURASHTRA LETTER U;Lo;0;L;;;;;N;;;;;
A887;SAURASHTRA LETTER UU;Lo;0;L;;;;;N;;;;;
A888;SAURASHTRA LETTER VOCALIC R;Lo;0;L;;;;;N;;;;;
A889;SAURASHTRA LETTER VOCALIC RR;Lo;0;L;;;;;N;;;;;
A88A;SAURASHTRA LETTER VOCALIC L;Lo;0;L;;;;;N;;;;;
A88B;SAURASHTRA LETTER VOCALIC LL;Lo;0;L;;;;;N;;;;;
A88C;SAURASHTRA LETTER E;Lo;0;L;;;;;N;;;;;
A88D;SAURASHTRA LETTER EE;Lo;0;L;;;;;N;;;;;
A88E;SAURASHTRA LETTER AI;Lo;0;L;;;;;N;;;;;
A88F;SAURASHTRA LETTER O;Lo;0;L;;;;;N;;;;;
A890;SAURASHTRA LETTER OO;Lo;0;L;;;;;N;;;;;
A891;SAURASHTRA LETTER AU;Lo;0;L;;;;;N;;;;;
A892;SAURASHTRA LETTER KA;Lo;0;L;;;;;N;;;;;
A893;SAURASHTRA LETTER KHA;Lo;0;L;;;;;N;;;;;
A894;SAURASHTRA LETTER GA;Lo;0;L;;;;;N;;;;;
A895;SAURASHTRA LETTER GHA;Lo;0;L;;;;;N;;;;;
A896;SAURASHTRA LETTER NGA;Lo;0;L;;;;;N;;;;;
A897;SAURASHTRA LETTER CA;Lo;0;L;;;;;N;;;;;
A898;SAURASHTRA LETTER CHA;Lo;0;L;;;;;N;;;;;
A899;SAURASHTRA LETTER JA;Lo;0;L;;;;;N;;;;;
A89A;SAURASHTRA LETTER JHA;Lo;0;L;;;;;N;;;;;
A89B;SAURASHTRA LETTER NYA;Lo;0;L;;;;;N;;;;;
A89C;SAURASHTRA LETTER TTA;Lo;0;L;;;;;N;;;;;
A89D;SAURASHTRA LETTER TTHA;Lo;0;L;;;;;N;;;;;
A89E;SAURASHTRA LETTER DDA;Lo;0;L;;;;;N;;;;;
A89F;SAURASHTRA LETTER DDHA;Lo;0;L;;;;;N;;;;;
A8A0;SAURASHTRA LETTER NNA;Lo;0;L;;;;;N;;;;;
A8A1;SAURASHTRA LETTER TA;Lo;0;L;;;;;N;;;;;
A8A2;SAURASHTRA LETTER THA;Lo;0;L;;;;;N;;;;;
A8A3;SAURASHTRA LETTER DA;Lo;0;L;;;;;N;;;;;
A8A4;SAURASHTRA LETTER DHA;Lo;0;L;;;;;N;;;;;
A8A5;SAURASHTRA LETTER NA;Lo;0;L;;;;;N;;;;;
A8A6;SAURASHTRA LETTER PA;Lo;0;L;;;;;N;;;;;
A8A7;SAURASHTRA LETTER PHA;Lo;0;L;;;;;N;;;;;
A8A8;SAURASHTRA LETTER BA;Lo;0;L;;;;;N;;;;;
```

A8A9;SAURASHTRA LETTER BHA;Lo;0;L;;;;;N;;;;;
A8AA;SAURASHTRA LETTER MA;Lo;0;L;;;;;N;;;;;
A8AB;SAURASHTRA LETTER YA;Lo;0;L;;;;;N;;;;;
A8AC;SAURASHTRA LETTER RA;Lo;0;L;;;;;N;;;;;
A8AD;SAURASHTRA LETTER LA;Lo;0;L;;;;;N;;;;;
A8AE;SAURASHTRA LETTER VA;Lo;0;L;;;;;N;;;;;
A8AF;SAURASHTRA LETTER SHA;Lo;0;L;;;;;N;;;;;
A8B0;SAURASHTRA LETTER SSA;Lo;0;L;;;;;N;;;;;
A8B1;SAURASHTRA LETTER SA;Lo;0;L;;;;;N;;;;;
A8B2;SAURASHTRA LETTER HA;Lo;0;L;;;;;N;;;;;
A8B3;SAURASHTRA LETTER LLA;Lo;0;L;;;;;N;;;;;
A8B4;SAURASHTRA CONSONANT SIGN UPAKSHARA;Mc;0;L;;;;;N;;;;;
A8B5;SAURASHTRA VOWEL SIGN AA;Mc;0;L;;;;;N;;;;;
A8B6;SAURASHTRA VOWEL SIGN I;Mc;0;L;;;;;N;;;;;
A8B7;SAURASHTRA VOWEL SIGN II;Mc;0;L;;;;;N;;;;;
A8B8;SAURASHTRA VOWEL SIGN U;Mc;0;L;;;;;N;;;;;
A8B9;SAURASHTRA VOWEL SIGN UU;Mc;0;L;;;;;N;;;;;
A8BA;SAURASHTRA VOWEL SIGN VOCALIC R;Mc;0;L;;;;;N;;;;;
A8BB;SAURASHTRA VOWEL SIGN VOCALIC RR;Mc;0;L;;;;;N;;;;;
A8BC;SAURASHTRA VOWEL SIGN VOCALIC L;Mc;0;L;;;;;N;;;;;
A8BD;SAURASHTRA VOWEL SIGN VOCALIC LL;Mc;0;L;;;;;N;;;;;
A8BE;SAURASHTRA VOWEL SIGN E;Mc;0;L;;;;;N;;;;;
A8BF;SAURASHTRA VOWEL SIGN EE;Mc;0;L;;;;;N;;;;;
A8C0;SAURASHTRA VOWEL SIGN AI;Mc;0;L;;;;;N;;;;;
A8C1;SAURASHTRA VOWEL SIGN O;Mc;0;L;;;;;N;;;;;
A8C2;SAURASHTRA VOWEL SIGN OO;Mc;0;L;;;;;N;;;;;
A8C3;SAURASHTRA VOWEL SIGN AU;Mc;0;L;;;;;N;;;;;
A8C4;SAURASHTRA SIGN VIRAMA;Mn;9;NSM;;;;;N;;;;;
A8CE;SAURASHTRA DANDA;Po;0;L;;;;;N;;;;;
A8CF;SAURASHTRA DOUBLE DANDA;Po;0;L;;;;;N;;;;;
A8D0;SAURASHTRA DIGIT ZERO;Nd;0;L;;0;0;0;N;;;;;
A8D1;SAURASHTRA DIGIT ONE;Nd;0;L;;1;1;1;N;;;;;
A8D2;SAURASHTRA DIGIT TWO;Nd;0;L;;2;2;2;N;;;;;
A8D3;SAURASHTRA DIGIT THREE;Nd;0;L;;3;3;3;N;;;;;
A8D4;SAURASHTRA DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;;
A8D5;SAURASHTRA DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;;
A8D6;SAURASHTRA DIGIT SIX;Nd;0;L;;6;6;6;N;;;;;
A8D7;SAURASHTRA DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;;
A8D8;SAURASHTRA DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;;
A8D9;SAURASHTRA DIGIT NINE;Nd;0;L;;9;9;9;N;;;;;

References

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

Norihiko Učida. *Saurashtra-English dictionary*. ISBN 3447030550. (In Latin script.) (Učida 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.)

பெரியாழ்வார் *Padārthun*. 1998. A magazine published in Madurai.

பெரியாழ்வார் *Bhāṣābhīmāni*. 2001. A magazine published in Madurai.

சுப்ரமணியன், ஸ்ரீ தாடா. 2000. *சுப்ரமணியன் பாத் மாலை* = ஸவ்ராஷ்ட்ர பாட² மாலொ. மதுரை: ப்ரியா பதப்பகம் ப்ரியா டெக்னிகல் இன்ஸ்டிடியூட். [Śrī Tāṭā. Cupramaṇiyaṅ. 2000. *Savrāṣṭra pātha mālo* = *Savrāṣṭra pāṭ²a mālo*. Madurai: Priyā Patappakam Priyā Ṭekṇikal Inṣṭīyut]

Acknowledgements

This project was made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at UC Berkeley).

Figures

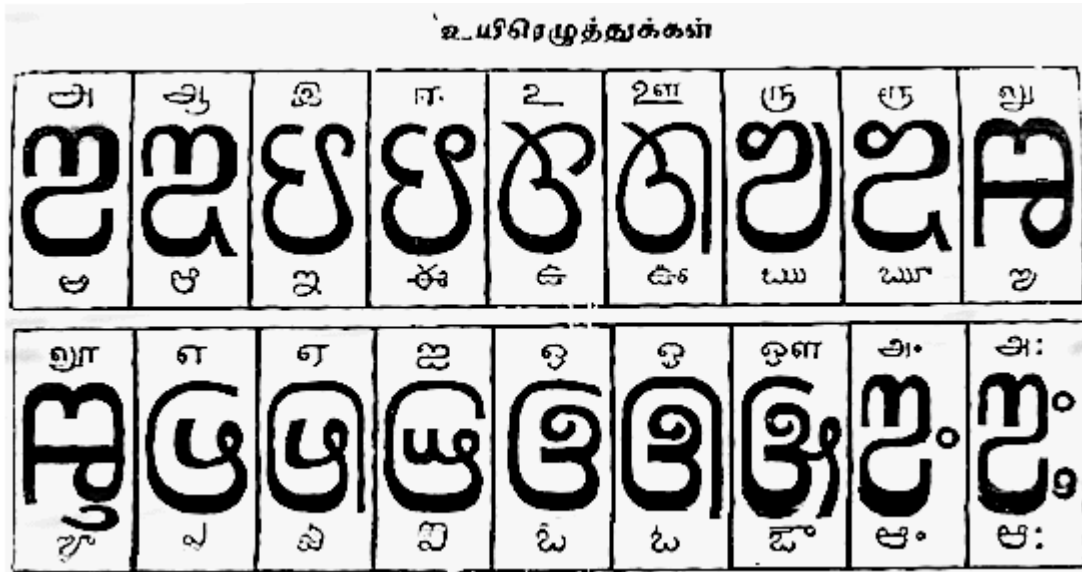


Figure 1. Saurashtra vowels, with equivalents in Tamil and Telugu scripts.

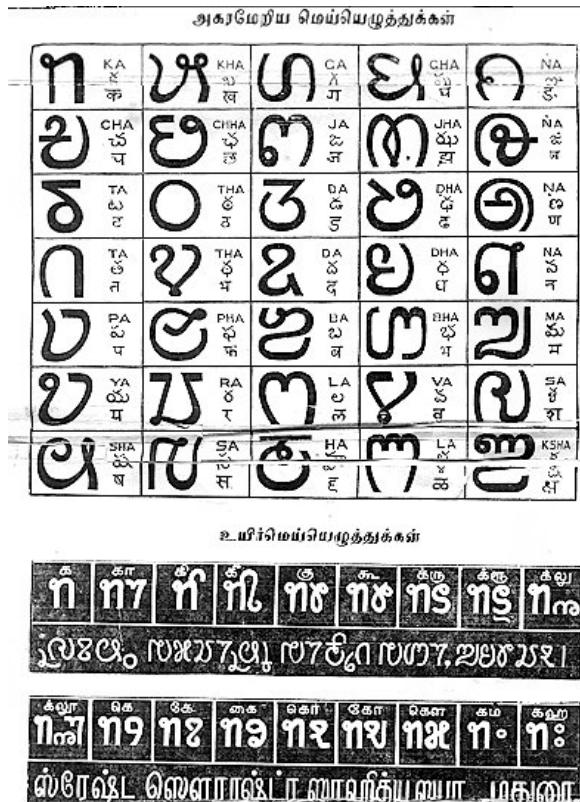


Figure 2. The text in the bottom half of this figure is given in Saurashtra and Tamil script. Note the unusual forms of -ḷ and -ḻ. The Saurashtra in the second line is in an early form of the script as it shows conjunct consonants. It is rendered here with modern orthography and transliteration:

ஸ்ரேஷ்ட, ஸாஹித்ய, ஶபா, மதுரை |
 ஶ்ரேஷ்ட ஶாஹித்ய ஶபா, மதுரை |
 śrēṣṭha saurāṣṭra sāhitya sabhā, machuro |
 ஶ்ரேஷ்ட ஶாஹித்ய ஶபா, மதுரை |
 srēṣṭha saurāṣṭra sāhītya sapā, maturai

ரோமன் எழுத்து - லெவராஷ்டிர மொழி

- ஜாபாலி

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

VOWELS - உயிரெழுத்துகள்

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

CONSONANTS - மெய்யெழுத்துகள்

ॠ	ॡ	ॢ	ॣ	।	॥	०	ॠ	ॡ	ॢ
க்	க் ²	க் ³	க் ⁴	ங்	ச்	ச் ²	ஜ்	ஜ் ²	ஞ்
K	Kh	G	Gh	Ng	C	Ch	J	Jh	Ņg
ॠ	ॡ	ॢ	ॣ	।	॥	०	ॠ	ॡ	ॢ
ட்	ட் ²	ட் ³	ட் ⁴	ண்	த்	த் ²	த் ³	த் ⁴	ந்
T	Th	D	Dh	Ñ	T	Th	D	Dh	N
ॠ	ॡ	ॢ	ॣ	।	॥	०	ॠ	ॡ	ॢ
ப்	ப் ²	ப் ³	ப் ⁴	ம்	ய்	ர்	ல்	வ்	ஸ்
P	Ph	B	Bh	M	Y	R	L	V	Ś
ॠ	ॡ	ॢ	ॣ	।	॥	०	ॠ	ॡ	ॢ
ஷ்	ஸ்	ஹ்	ள்	க்ஷ்	ர்ஹ்	ல்ஹ்	ம்ஹ்	ந்ஹ்	
Sh	S	H	L	Ksh	Rh	Lh	Mh	Nh	

CONSONANT ADDED VOWELS - உயிர் மெய்யெழுத்துகள்

ॠ	ॡ	ॢ	ॣ	।	॥	०	ॠ	ॡ	ॢ
க	கா	கி	கீ	கு	கூ	கெ	கே	கொ	கோ
KA	KĀ	KI	KĪ	KU	KŪ	KE	KĒ	KO	KŌ

EXAMPLES - உதாரணங்கள்

ॠॡॢॣ।॥०ॠॡॢ	ॠॡॢॣ।॥०ॠॡॢ
பாஷாபிமானி	இயே அநுசரி அம்தெநு லேகுந் லிக்குவொ
Bhāshābhīmāni	IYE ANUCARI AMTENU LĒKHUN LIKKUVO
ॠॡॢॣ।॥०ॠॡॢ	ॠॡॢॣ।॥०ॠॡॢ
ஸவ்ராஷ்டிர மணி	ஒண்டெதிம் சிக் குல்டுவாயி
SAVRĀSHĪTRĀ MAṆĪ	OṆṬEDIM SIKKULĪUVĀYĪ

தீபாவளி மலர் 2001

Figure 7. Sample from the magazine *Bhāṣābhīmāni*, showing Saurashtra letters with equivalents in Tamil and Latin. The consonants are shown with their vowels killed; the UPAKSHARA character is shown in combination with killed consonants, indicated with the arrow. Below that the vowel matras are given in combination with consonant ॠ ka.

Row A8: SAURASHTRA

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

G = 00
P = 00

Row A8: SAURASHTRA

hex	Name
A880	SAURASHTRA SIGN ANUSVARA
A881	SAURASHTRA SIGN VISARGA
A882	SAURASHTRA LETTER A
A883	SAURASHTRA LETTER AA
A884	SAURASHTRA LETTER I
A885	SAURASHTRA LETTER II
A886	SAURASHTRA LETTER U
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 CONSONANT SIGN 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

hex	Name
A8D9	SAURASHTRA DIGIT NINE
A8DA	(This position shall not be used)
A8DB	(This position shall not be used)
A8DC	(This position shall not be used)
A8DD	(This position shall not be used)
A8DE	(This position shall not be used)
A8DF	(This position shall not be used)

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

2005-09-21

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)

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 (Evertyping). 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 (Evertyping). 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?

Jeyakumar C. K. is a user himself, and has been in touch with Mr Kubendiran, editor of *ભજાંબહિમંની* *Bhāṣābhīmānī*.

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)

Saurashtra script is used fairly rarely, but it is enjoying a renaissance, and certainly has some commercial value (there are even advertisements in the monthly magazine *Bhāṣābhīmāni* for “Desk Top Publishing in English, Tamil, Hindi, Malayalam, Sourashtra”). Both the Saurashtran and Tamil scripts are used to write Saurashtra. Use of Saurashtra script is primarily cultural and religious, though some private schools are now teaching the script to children.

4b. Reference

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

Yes.

5b. If YES, where?

In the Indian states Tamil Nadu and Andhra Pradesh. The monthly magazine *Bhāṣābhīmāni* is published in Madurai. Much of this magazine is written in Tamil and in Saurashtra in Tamil script, but Saurashtra script is also used in it.

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. Superficial

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