

Review of a Proposal placed before UTC bearing No. L2/03-273

Reviewed by
Peri Bhaskararao

Title of the Proposal: Proposal to add Tamil grantha character SHA

Dated: July 29, 2003

1. Purpose of the Proposal:

The proposal under review is for inclusion of Tamil ‘Grantha’ SHA. It should be pointed out that Devanagari script as used for Sanskrit language possesses ‘akṣaras’ for four fricatives. In conventional romanization (for linguistic and philological purposes), the four fricatives are transcribed as: {ṣ} {ś} {s} {h}. A transcriptional variant of {ṣ} is {ṣ̣}. In a very old romanization method it was also transcribed as {ç}.

All these four letters occur in the native vocabulary of all the Indo-Aryan languages (which are derivatives of Sanskrit). However, there has been some merger of the first three in different combinations in some of the modern Indo-Aryan languages. But they are maintained in the writing systems of these languages.

The situation of Dravidian languages is somewhat different. {ṣ} and {ś} occur only in words originally borrowed from Sanskrit and later from other Indo-Aryan languages as well as from English. Malayalam, Telugu and Kannada have internalized the letters for all these four. ISCII included letters for {ś} {s} {h} but left out the letter for {ṣ} in Tamil. This resulted in a major deficiency in the code – for instance, there is no way of representing the backing string of a very important ‘akṣara’ in the language viz., {ṣrī}. Probably this is taken care of by some glyph somewhere outside the U+0B80 to U+BFF range. This position needs to be corrected.

The proposal under review correctly points out that the letter {ṣ} is an essential letter in writing and reading Tamil documents. Though it occurs in words of Indo-Aryan origin, its correct usage is important for the integrity of the document. It cannot be substituted by any nearby letter such as {s} for maintaining the correct pronunciation.

Although its ‘Proposed Name’ is given as ‘TAMIL LETTER SHA’, the title of the Proposal calls it ‘Tamil grantha character SHA’. Usage of the term ‘Grantha character’ has wider ramifications. I strongly recommend that we reserve the term ‘Grantha character’ for a character that will appear in the full set of ‘Grantha Characters’ when they are encoded separately as a complete set. The characters U+0B9C and U+0BB7 to U+0BB9 though occur only in borrowed vocabulary in Tamil language are simply called ‘Tamil letters’. They are thoroughly internalized into Tamil script. The glyph shapes of the above characters resemble those of the glyphs of the corresponding characters in ‘Grantha script’. The other characters in ‘Grantha script’ considerably differ in their glyph shapes from the present-day Tamil characters. ‘Grantha characters’ when encoded will be nearer to ‘Tulu characters’ which also need to be taken up sometime.

Hence, the proposed character should be simply called ‘Tamil Letter SHA’ and the Proposal be treated as ‘Proposal to add Tamil character SHA’.

Combinations of {ṣ}:

The proposal discusses ‘Data Processing and ‘Ligature’ issues of this character. In general this character behaves like {ṣ} in its combination with vowel-mātrās as well clustering with other consonants. As mentioned in §1 above, the ‘akṣara’ {ṣrī} requires a separate glyph. There is also a separate glyph for {ṣri} as in the word {āṣṛita} (e.g., in the phrase {āṣṛita-rakṣaka-ṣrī-rāma}). Both {ṣrī} and {ṣri} will be treated as Grapheme-compounds and the compound-glyphs will be derived by rule.

Recommendation: The proposed character for Tamil {ṣ} be given the status of a character and be placed at U+0BB6 as requested by the proposers.

Prof. *Peri Bhaskararao*, Ph.D.
ILCAA, Tokyo University of Foreign Studies
Asahi-cho 3-11-1, Fuchu-shi
TOKYO 183-8534, Japan
Phones: +81-42-330-5607 and +81-42-330-5609
Fax: +81-42-330-5607
Email: bhaskar@aa.tufs.ac.jp
<http://www3.aa.tufs.ac.jp/~bhaskar/index.html>
[http: www.gicas.jp](http://www.gicas.jp)