

Comment on L2/12-231: Tamil symbols and fractions

Dr. Naga Ganesan (naa.ganesan@gmail.com)

(1) There is already a proposal (in 2010 documents) to encode the Tamil fractions in SMP (L2/ 10-408) from INFITT, and happy to note that the archaic Tamil fractions and symbols are proposed to be in SMP in this later proposal also. There will be few more archaic symbols to be added in the SMP block in the future. All the symbols in L2/12-231 are very archaic and not in use in any contemporary Tamil books or newspapers. Hence, it is recommended that these archaic symbols are grouped together in one SMP block, and not a few in the BMP Tamil block prejudicing some symbols over others. Separating just a few archaic signs out of the total number of archaic signs is not to be recommended for encoding into TUS.

(2) Tamil credit sign – Is duplicate encoding needed?

S. Sharma is proposing duplicate encoding for 2 letters in Brahmi (L2/12-226), one letter (II) in Malayalam (L2/12-225). Per comment in L2/12-233 , these duplicate encoding code points in Brahmi and Malayalam blocks need to be avoided. Otherwise, the Unicode standard will have many code points for every glyph variant of characters in Indic scripts. To avoid duplicate encoding of letters like these, the original Brahmi proposal itself has statements on these glyph variants.

When I wrote the Tamil OM sign proposal, the many glyph variants of OM was mentioned, and just one OM sign code point handles many variants. Also, in Grantha original proposal, the capability to write archaic Grantha works need to have Indo-Aryan and Dravidian letters (not just the ability to write Indo-Aryan as in pipeline now), and Grantha script to write Dravidian languages such as old Tamil and old Malayalam, there has to be two Grantha fonts in Unicode encoding – just like we have Antiqua and Fraktur style fonts in European languages.

Similarly, for Tamil encoding, there is already a TAMIL CREDIT SIGN, and there is no need for an additional Tamil credit sign now as proposed in L2/12-231. From printed books of the last 200 years, Tamil credit sign is well attested and are already documented in

L2/12-150. Tamil credit sign contains letter, VA (ᱠ) just like Tamil debit sign contains letter, PA (ᱡ) due to their names in Tamil language. These letter shapes, VA and PA, are

embedded in current TUS charts and used widely, and if there is any glyph shape change in the future for Tamil credit sign in TUS charts, the evidence in Tamil printed books documented in L2/12-150 should be used. Other hand written shapes of Tamil credit & vowel signs, not attested in printed books or in the major Tamil fonts in Unicode can be handled in user fonts.

It is recommended that the duplicate code point for Tamil credit sign in Sharma's document, L2/12-231 be removed from encoding. Also, there is no need for 2 code points for the number sign, the two glyph variants for the Tamil number sign are very minor shape variations, and should be handled in fonts if the archaic shape is needed.

It is also requested that all of these archaic symbols be moved to SMP block, and not a small subset of these archaic signs into Tamil block in BMP.

(3) I have not had time to go thru' the annotations, but careful thought is needed to annotate these archaic signs (all of them should be in SMP). For example, I read annotations like "aaga" and "poga" which may suit Telugu or Kannada well, but not Tamil language. These annotations, if getting into TUS, should be corrected as "aaka" and "poka" respectively, according to modern Tamil linguistic scholarship. (The intervocalical –k- is a fricative, & not g). So, more work needs to be done regarding annotation texts and their spellings. And, also in character names written in English, this Tamil linguistically correct principle in retaining intervocalical –k- should be followed: vakaiyaraa and varaakan (NOT vagaiyaraa and varaagan etc.,) as in Sharma's proposal. This is standard practice to write –k- instead of –g- in all University level academic publications both the West (Europe, USA) or in Japan and also in India.

Summary:

- (1) all these archaic symbols are recommended to be in SMP & not some split into BMP.
- (2) Tamil in TUS should not have duplicate code points for credit and number signs.
- (3) Names and annotations in English should retain –k- (& not –g-) for intervocalical க as the modern academic practice of writing Tamil names.