

Naming the Tamil fractions and symbols based on the Tamil Lexicon

Shriramana Sharma, jamadagni-at-gmail-dot-com, India

2013-Nov-03

This is w.r.t. the following documents:

L2/13-047	Revised proposal to encode Tamil fractions and symbols	Sharma
L2/13-077	Comments on above	Ganesan
L2/13-161	Document echoing sec 1 of above	TVA
L2/13-175	Response to the Ganesan/TVA suggestions	Sharma
L2/13-193	Comments on above	Ganesan
L2/13-210	Recommendations on Indic documents	Anderson et al

(secs 7 and 8 of the above relate to this issue)

In his latest document L2/13-193, Ganesan says the TN Govt has constituted a committee of experts in this regard and that they have endorsed the request for changes made in the TVA's L2/13-161 which in fact echoes his previous suggestions from L2/13-077. The official communication by the TVA (TN Govt's IT Dept) does not mention any such committee. Further, to my knowledge Ganesan is not the authorized spokesperson for the TN Govt nor has he given reference to any official GOTN announcement regarding such a committee. Nevertheless, I will disregard his unverifiable claims and now examine the issue alone.

While neither Ganesan's earlier document nor the TVA's document mentions any academic sources as basis for the "transcription standard" they recommend, Ganesan's recent document seems to imply that the Tamil Lexicon is this basis. Anderson et al hence recommend that a spelling based on such dictionary usage be followed.

I have no objection to naming the words based on the Tamil Lexicon's transliteration. My point is only that any naming policy should be consistently followed. I have explained in detail in my previous document L2/13-175 N4477 as to how such consistency is absent in the recommendations of Ganesan/TVA w.r.t. the c/s, k/g and t/d spelling issues. I will not repeat those detailed explanations here.

In L2/13-210 §7, Anderson et al say “we recommend ... the current spelling described above”. It is unclear whether they refer to a consistent policy based on the Tamil Lexicon or the spellings recommended by Ganesan/TVA, since the two are not the same.

It would however be elementary and acceptable to adopt a consistent spelling based on the Tamil Lexicon’s transliteration. I present such a spelling policy below.

A consistent naming policy based on the Tamil Lexicon

The transliteration followed by the Tamil Lexicon (TL), at least going by its online version at <http://dsal.uchicago.edu/dictionaries/tamil-lex/>, seems to be effectively the same as the international standard ISO 15919 as applied to the Tamil script. One could simply remove the diacritics from such a transliteration of the word to obtain a simple recognizable name. Based on this policy, the following spelling would result:

	TL / ISO 15919	Current	TVA	TL/ISO minus diacs
11FD0	கீ <i>kīl</i>	KIIZH	-	KIL
11FD1	ஐ <i>nel</i>	NEL	-	NEL
11FD2	ஞ <i>cuvaṭu</i>	SUVADU	CUVADU	CUVATU
11FD3	ற <i>ālākku</i>	AAZHAAKKU	-	ALAKKU
11FD4	றி <i>uri</i>	URI	-	URI
11FD5	று <i>mūvuḷakku</i>	MUUVUZHAKKU	-	MUVULAKKU
11FD6	று <i>patakku</i>	PADAKKU	-	PATAKKU
11FD7	று <i>mukkurūṇi</i>	MUKKURUNI	-	MUKKURUNI
11FD8	று <i>paicā</i>	PAISAA	-	PAICA
11FD9	று <i>aṇā</i>	ANAA	-	ANA
11FDA	று <i>kācu</i>	KAASU	-	KACU
11FDB	று <i>paṇam</i>	PANAM	-	PANAM
11FDC	று <i>poṇ</i>	PON	-	PON
11FDD	று <i>varākaṇ</i>	VARAAGAN	VARAAKAN	VARAKAN
11FDE	று <i>pāram</i>	BAARAM	PAARAM	PARAM
11FDF	று <i>kajam</i>	GEJAM	KESAM	KAJAM (see below)
11FE0	று <i>kuli</i>	KUZHI	-	KULI

11FE1	வெ	<i>vēli</i>	VELI	-	VELI
11FE8	மு	<i>mutaliya</i>	MUDALIYA	-	MUTALIYA
11FE9	வக	<i>vakaiyarā</i>	VAGAIYARAA	VAKAIYARAA	VAKAIYARA
11FEA	சீ	<i>cirañcīvi</i>	CIRANJIVI	-	CIRANCIVI
11FEB	பீ	<i>piḷḷai</i>	PILLAI	-	PILLAI
11FEC	ரீ	<i>rāja</i>	RAAJA	-	RAJA
11FED	கீ	<i>kku</i>	KKU	-	KKU
11FEE	யூ	<i>yum</i>	YUM	-	YUM
11FEF	வூ	<i>vum</i>	VUM	-	VUM

This model also avoids the unsightly repeated consonants which would result if the existing UCS conventions were applied to these words. (See L2/13-175 p 6.)

In this connection of repeated consonants, Anderson et al mention only liquids and nasals and seem to have forgotten the retroflex plosive TTA. The correct categorization would be in terms of retroflex and alveolar consonants. Anderson et al invite suggestions from the TVA as to how these may be transliterated. This seems to indicate that their previous words “we recommend ... the current spelling described above” indeed refer to the Ganesan/TVA suggestions.

However, as noted, the Ganesan/TVA suggestions are not consistent with any model, even the Tamil Lexicon. OTOH the model described above follows the Tamil Lexicon and the ISO 15919 international standard, is consistent, and does not require special consideration to be accorded to any characters (i.e. retroflexes etc) – the policy being simply to use the Tamil Lexicon / ISO 15919 spelling minus the diacritics.

The final point is regarding the name of 11FDF currently named TAMIL SIGN GEJAM. Ganesan/TVA request that it be renamed to KESAM. However, this is inconsistent with the TL and ISO 15919, which do not use /s/ for ஃ in intervocalic position. Further, as explained in L2/13-176, the *primary* form of the word as per the TL is not *kecam* or *kejam* but it is *kajam* (see <http://dsalrv02.uchicago.edu/cgi-bin/philologic/getobject.pl?c.2:1:6406.tamillex.2509528>). This spelling is hence to be preferred. Since it has no diacritics, it may be directly used as per the policy described above.

-0-0-0-