

## Naming Tamil Symbols in SMP – *Vallinam* Characters

Dr. Naga Ganesan ([naa.ganesan@gmail.com](mailto:naa.ganesan@gmail.com))

This document requests the **proper spelling of Tamil words** for archaic symbols to be encoded in SMP of the Unicode Standard.

- 1.0 As per our previous request, all the archaic symbols are in SMP Tamil block, and the Pipeline of future characters has been updated recently. Proposed characters for Tamil fractions and symbols (all in SMP) in the current ISO/IEC 10646 ballot can be found here. <http://www.unicode.org/L2/L2014/14271-n4637.pdf>
- 2.0 In the last UTC meeting, the decision was taken by UTC not to use any digits within Tamil words to be used in Character names in SMP. We have seen the letter written by Dr. Lisa Moore to Tamil Nadu Government. Also, note that in any Indian language word, this practice has been followed, For example, words like *danda*, *karshana*, *viraama* etc., do not have any digits within them. It is good that same practice will be continued for Tamil symbols in SMP block as well.
- 3.0 Linguistically correct representation of *Vallinam* (ஃ, ழ, ல், ழ், ழ்) letters of Tamil words in SMP Character names for Tamil symbols is needed [1]. This is because Vallinam letters, *k*, *c*, *t*, *th*, *p* have not only voiced *g*, *j*, *d*, *dh*, *b* sounds but they also have fricative sounds. That is why linguists such as Prof. K. Nachimuthu, Dr. Vasu Renganathan have written as comments from INFITT in 2014. Here is INFITT Chair's document from 2014 and Tamil linguist at University of Pennsylvania giving the summary of words that need a letter change in each. <http://www.unicode.org/L2/L2014/14212-tamil-names.pdf>

Please implement these needed changes (L2/14-212) in the repertoire in Tamil character names and annotations. The changes requested for Vallinam characters are very minimal and follows Tamil linguistic principles.

Thanks,  
Naga Ganesan

Reference [1]: In the next page, I give quotes from the book by Prof. R. E. Asher, Edinburgh. As two examples, the quotes describes the various sounds of Tamil vallinam letters ஃ & ழ, including the fricatives, of Vallinam letters. The book is: *Croom Helm Descriptive Grammars, TAMIL*, published in 1985

### 3.1.2.1.1. Plosives and affricates

#### (1) k

[k] voiceless dorso-velar plosive in utterance-initial position (word-initial  $\text{க}$ - which is not utterance-initial is sometimes realised as a more laxly articulated consonant): e.g. kuṭu 'give(imp)' [kəṭə]. In this position there may be slight, nondistinctive, aspiration;

[k:] long voiceless dorso-velar plosive as the realisation of the intervocalic sequence -kk-: e.g. pakkam 'side' [pak:ṣ]. In rapid speech the closure may be shorter than the duration suggested by the length mark, but the consonant will always be a stop and voiceless;

[x] voiceless dorso-velar fricative in intervocalic position: makan 'son' [maxɛ̃]. In other dialects a glottal fricative, [h] or [ɦ], occurs in this position;

[g] voiced dorso-velar plosive occurs medially in the sequence -ṅk-: vaṅku 'buy' [va:ŋgɪ];

#### (2) c

[tʃ] voiceless lamino-palatoalveolar affricate in utterance-initial position in a few words: e.g. cantiran 'moon' [tʃaṅṅɪrɛ̃], cantanam 'sandalwood paste' [tʃaṅṅənɔ̃], catṇi 'chutney' [tʃatṇɪ], cil 'exclamation expressive of disgust or contempt' [tʃɪ:] (though note the alternative pronunciations [sɪ:tʃɪ:] and [tʃɪ:tʃɪ:] in the reduplicated form of this, cilccil);

212

[tʃ̌] voiceless lamino-palatoalveolar affricate with the stop element potentially slightly prolonged, as the realisation of the intervocalic sequence -cc-: pacce 'green' [patʃ̌ɛ̃];

[s] voiceless lamino-alveolar fricative in (i) intervocalic position: paci 'hunger' [pasɪ], kocu 'mosquito' [kosə], and (ii) word-initial position: cante 'market' [saṅgɛ̃], colu 'say' [sol:ə]. In words such as those referred to in (ii), initial  $\text{c}$ - is for some speakers realised as the affricate in certain grammatical environments, most important of which (in terms of frequency) are preceding accusative case, preceding dative case, and preceding adverb in -aa. Lack of pause at such word junctures is a necessary condition of the articulation of [tʃ̌]. An alternative to [s] in some dialects is voiceless lamino-postalveolar fricative [ɸ];

[dʒ] voiced lamino-palatoalveolar affricate occurs medially in the sequence -ṅc-: neṅcu 'heart' [neṅdʒɪ], otaṅcutu 'it broke' [otaṅdʒɪṭɪ];

[ʃ] voiceless apico-palatal fricative when preceded by a retroflex stop: kaṭci '(political) party' [kaṭʃɪ], veṭci 'name of a shrub' [veṭʃɪ]. For such words there is an alternative, probably more common, pronunciation with a voiceless lamino-palatoalveolar affricate: [katʃ̌ɪ], [veṭʃ̌ɪ], for which the appropriate broad transcription would be kaccɪ and