

## Spelling changes for Tamil fractions and symbols

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

2013-Dec-20

### Supersedes:

N4477 L2/13-175 – Response to L2/13-161 on naming Tamil fractions and symbols

N4478 L2/13-176 – Request to change one character name from N4430 L2/13-047

## Introduction

My proposal N4430 L2/13-047 to encode 55 new Tamil characters has been approved by the UTC and WG2 committees at the meetings in 2013 May and June respectively. These characters are currently in the pipeline in Stage 3. This is a request to change the spelling of sixteen character names from the list that has been approved.

## Background

In my original proposal L2/12-231 p 61 I had clarified the following: I proposed the Tamil name for characters when it cannot be translated into English. In doing so, since there is no de jure standard mapping Tamil letters to plain ASCII, I used the de facto conventions employed by natives when writing Tamil words in “plain English”.

Subsequent to this, the UTC received the feedback document L2/13-161 asking for some of the spellings to be changed. However, as I have discussed in N4477 L2/13-175, the changes requested do not adhere to any consistent naming scheme. In N4478 L2/13-176, I had separately requested one character name to be changed for other reasons.

Since then, the UTC preliminary review committee commenting on the requested changes has noted in L2/13-210 §7:

*The ... recommendation seems to be based on dictionary forms, such as those in the University of Madras Tamil Lexicon... By relying on the entries in a standardized reference (such as a dictionary), the spelling in the Unicode Standard will be as consistent as possible.*

In my feedback document L2/13-217, I had explained how it is indeed possible to arrive at a consistent naming policy based on the transliteration employed in the Tamil Lexicon. The present document is a formal request for the spelling changes based on that policy.

## A consistent naming policy based on the Tamil Lexicon

The transliteration scheme followed in the Tamil Lexicon is the same as ISO 15919 applied to Tamil. Thus, by the simple expedient of removing the diacritic marks from the ISO 15919 form, one can arrive at consistent (and unique) names for all the characters in question.

Based on this policy, the spelling of the following characters will change as follows. For brevity and clarity, I have not repeated the adjectives prefixed to each name.

		<b>Current</b>	<b>ISO 15919</b>	<b>Newly proposed</b>
<b>TAMIL FRACTION DOWNSCALING FACTOR</b>				
11FD0	கீ	KIIZH	<i>kīl</i>	KIL
		<b>Current</b>	<b>ISO 15919</b>	<b>Newly proposed</b>
<b>TAMIL SIGN</b>				
11FD2	சூ	SUVADU	<i>cuvaṭu</i>	CUVATU
11FD3	ஶ	AAZHAAKKU	<i>ālākku</i>	ALAKKU
11FD5	ஶ	MUUVUZHAKKU	<i>mūvulakku</i>	MUVULAKKU
11FD6	ஶ	PADAKKU	<i>patakku</i>	PATAKKU
11FD8	பச	PAISAA	<i>paicā</i>	PAICA
11FD9	அணு	ANAA	<i>aṇā</i>	ANA
11FDA	கூ	KAASU	<i>kācu</i>	KACU
11FDD	வர	VARAAGAN	<i>varākaṇ</i>	VARAKAN
11FDE	பு	BAARAM	<i>pāram</i>	PARAM
11FDF	கி	GEJAM	<i>kajam</i>	KAJAM
11FE0	கூ	KUZHI	<i>kūli</i>	KULI
11FE8	மு	MUDALIYA	<i>mutaliya</i>	MUTALIYA
11FE9	வக	VAGAIYARAA	<i>vakaiyarā</i>	VAKAIYARA
11FEA	சீ	CIRANJIIVI	<i>cirañcīvi</i>	CIRANCIVI
11FEC	ரீ	RAAJA	<i>rāja</i>	RAJA

It would however be advisable to retain the originally proposed spellings as informative aliases for convenience of recognition by native users.

The spelling of the following character names *does not change* as a result of the policy:

		<b>Current</b>	<b>ISO 15919</b>	<b>Newly proposed</b>
<b>TAMIL SIGN</b>				
11FD1	இ	NEL	<i>nel</i>	NEL
11FD4	ஊரி	URI	<i>uri</i>	URI
11FD7	முகுருணி	MUKKURUNI	<i>mukkuruni</i>	MUKKURUNI
11FDB	பாணம்	PANAM	<i>paṇam</i>	PANAM
11FDC	பொன்	PON	<i>poṇ</i>	PON
11FE1	வெளி	VELI	<i>vēli</i>	VELI
11FEB	பிள்ளை	PILLAI	<i>piḷḷai</i>	PILLAI

		<b>Current</b>	<b>ISO 15919</b>	<b>Newly proposed</b>
<b>TAMIL ABBREVIATION</b>				
11FED	குகு	KKU	<i>kku</i>	KKU
11FEE	யும்	YUM	<i>yum</i>	YUM
11FEF	வும்	VUM	<i>vum</i>	VUM

For the convenience of the editors, I have attached the updated UnicodeData.txt and NamesList.txt entries which need to be added to the present published standard. These include the informative aliases requested above.

Note that the above changes affect only characters proposed for the newly requested Tamil Supplement block in the SMP and do not affect characters proposed for the existing Tamil block in the BMP.

-0-0-0-