

Subject : Comments on the “Proposal to Add Tamil Nukta Character-L2/15-256”

Authors: Naga.Elangovan (nelango5@gmail.com), Dr.Krishnan Ramasamy (iraamaki@bsnl.in), N.D.Logasundaram (selvindls61@gmail.com)

Date : 22-Jan-2016

To : UTC

We have the following comments on the proposal L2/15-256.

- 1) Proposal assumes that Chetti (ctt) is the dominant language amongst the other languages in reference. This appears to be incorrect as the Wayanad Chetti language community has only 5,000 native speakers according to http://www-01.sil.org/iso639-3/cr_files/2006-030_ctt.pdf . Whereas, the Irula (iru) and Betta Kurumba (xub) have 200,000 and 32,000 native speakers respectively.(https://en.wikipedia.org/wiki/Irula_language, https://en.wikipedia.org/wiki/Betta_Kurumba_language). Further, referring to the Proposal Summary Form (section C, item 2), it can be observed that the proposer has made communication with only the Chetti language community and not with others.

Therefore, the basis of the evaluation does not seem to be correct.

- 2) Referring to the code chart of ‘Combining Diacritical Marks’ <http://www.unicode.org/charts/PDF/U0300.pdf>, it can be observed that a diacritic, usually called under-dot, (herein as Nukta) viz., U+0323 already exists, albeit in a different place from the Tamil block. Nukta just means a diacritical mark in North Indian language like Hindi. The proposed U+0BBC is just a replica of the one already existing at U+0323. Should we be duplicating these diacritics in every language block? Then where would be the end? Further, the double-dot Nukta and a single-ring Nukta also exist in U+0324 and U+0325 respectively.

From [U0300.pdf](#):

- 0323 ◌ COMBINING DOT BELOW
= nang
- IPA: closer variety of vowel
 - Americanist and Indo-Europeanist: retraction or retroflexion
 - Semiticist: velarization or pharyngealization
 - Vietnamese tone mark
- 02D4 ◌ modifier letter up tack
- 0324 ◌◌ COMBINING DIAERESIS BELOW
- IPA: breathy-voice or murmur
- 02B1 ħ modifier letter small h with hook
- 0325 ◌ COMBINING RING BELOW
- IPA: voiceless
 - vocalic (in Latin transliteration of Indic sonorants)
 - Madurese

Therefore, we do not see any valid reason to include Nukta into the Tamil block as suggested in the proposal.

- 3) If anyone wants to develop a font inputting software to enable the writing of these requirements, then he/she can inherit U+0323 to U+0325 and make the Nuktas work for them as required.

In fact using MS Word, the same Chetti (ctt) text given in the proposal is produced as below very easily and simply.

பீயிக்கு

It should be noted that the dot under the character பீ is just the “U+0323 COMBINING DOT BELOW”. Therefore, when MS Word itself is enabling the writing of these texts very easily and simply, there is no reason to encode a Nukta character into the Tamil block. Having already the common and standard set of Combining Diacritical Marks, combining the diacritic mark with a Tamil character is just a software work for the requirement in the proposal, and it cannot become a replication of encoding.

Therefore, based on the discussions above, adding Nukta into Tamil Code Chart is nothing but unnecessary replication, and hence we request UTC to not encode Nukta into Tamil Code Chart.

.....end.....