

To: Thiru T.K.Ramahandran, Principal Secretary to the Government of Tamil Nadu
From: Unicode Technical Committee (UTC)
Author: Lisa Moore
Date: August 17, 2017
Subject: U+1133B COMBINING BINDU BELOW

Principal Secretary Ramachadran,

In response to your letter sent via email on 7/28/2017, the UTC included your comments in our discussions on possible ballot comments on PDAM 1.3 of ISO/IEC 10646.

The UTC is aware of the earlier proposal from the government of Tamil Nadu to support the need for a combining dot below (nukta) for Badaga text by using U+0323 COMBINING DOT BELOW. However, for best rendering behavior, an element of text needs to have both an appropriate image and appropriate properties. If a character does not have the most appropriate properties there will be implementation issues, including:

- *Maintenance issues.* Rendering engines will not function well without extra implementation effort, leading to challenging maintenance issues experienced by both the user community and the software providers.
- *Unexpected combing character class.* General Indic rendering systems require that a diacritic used as a nukta must be a combining mark with a combining class of 7. The character U+0323 COMBINING DOT BELOW has a combining class of 220, so this character is not a good candidate.
- *Breaking script runs.* The diacritic that is used must not unnecessarily break script runs in Tamil. The best results will be achieved if the character has a script value that is often used with the Tamil script, such as "Tamil" or "Grantha".
- *Lack of keyboard support.* The diacritic must be easily available on Badaga language keyboards and display correctly with the set of fonts that support Badaga text display.
- *Confusion of one and two dot forms.* If single and double dot forms were unified, it would lead to confusion in use and data representation. Though the function of both forms is the same, users might expect to be able to use one form or the other.

Based on all of these considerations, we supported the original proposal to encode a new character U+0BBC TAMIL SIGN NUKTA. It could have the correct combining class, the correct appearance, and correct behavior for script runs (assuming it was also assigned Script=Tamil). It would be straight forward to add it to keyboards and fonts.

However, we heard feedback from the government of Tamil Nadu that you preferred to not have the new character encoded in the Tamil block. As an accommodation, the UTC decided to instead to encode the new character, U+1133B COMBINING BINDU BELOW in the Grantha block, with the properties that would most easily meet the needs of the Badaga language community. Both the Tamil block and the Grantha block are good candidates for such a character.

Therefore, to support the Badaga language community with a diacritic that their language requires, we have analyzed input from your organization and the many technical issues and decided that to best accommodate your input and the implementation requirements, we will add the U+1133B COMBINING BINDU BELOW as a compromise location that will still work well in Indic rendering engines.

If you would like to discuss further, please let me know.

Lisa Moore
Chair, Unicode Technical Committee