

Request to Change Pulli Representation in the Proposed Vatteluttu Encoding

Author: Cibu C Johny (cibucj@gmail.com)

Date: 2021 Feb

Action: For consideration by Script Ad Hoc and UTC

Abstract

The document [L2/16-068](#) proposes to encode Vatteluttu characters as per the Indic model for modern scripts like Tamil. Even though we can appreciate the potential developer conveniences in adopting such an encoding model, modern Indic is not the most intuitive model for certain characters like *pulli*. In this document, I request that the *pulli* may follow an encoding model where it could be applied to both consonants, vowels and vowel signs, acting as a vowel reducer.

*Pull*i in Vatteluttu

The *pulli* diacritic appears as a combining dot at the top or top-right corner of a base character. The base can be certain independent vowels, consonants or <consonant, vowel sign> clusters.

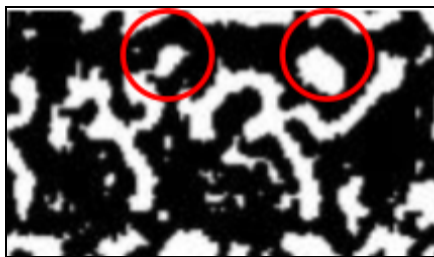


Fig 1: *Pull*i on top of KA and K-O in Velvikudi plates (769CE)¹

*Pull*i diacritic, when used, explicitly indicates vowel reduction and this function is analogous to that of *breve* (\sim) in Latin. When *pulli* is placed on a vowel sign or on a <consonant, vowel sign> cluster, it explicitly marks the corresponding vowel as a short one. When placed on a consonant, its default vowel is shortened or removed. All are vowel reduction mechanisms. In Fig 1, *pulli* represents both vowel removal and reduction in the K-K-O conjunct. The first *pulli* removes the default vowel /a/ from KA. The second *pulli* indicates that the Dependent Vowel Sign O on KA should be reduced. In all these three cases, *pulli* is the same character

¹ [Sastri, H.krishna, Epigraphia Indica. Vol 17](#)

indicating the same phonemic function of vowel reduction, interacting in the same way with the base character. Please see [L2/03-249](#) (Brahmi Encoding Proposal) for the details of evolution of *pulli* from Tamil Brahmi.

The number of inscriptions with *pulli* is quite limited and are from an older set of inscriptions that are found only in Tamil Nadu. Even though Vatteluttu inscriptions existed for an many more centuries in the Kerala region, usage of *pulli* is yet to be found in that time frame. The vast majority of inscriptions found in Kerala and from Tamil Nadu in this later time period did not use *pulli* and do not explicitly mark vowel reduction and leave it to the reader to disambiguate. That is, since *pulli* was not in use, VOWEL E/O could be read as short or long E/O, depending on the context. Similarly, consonants can be read with the default vowel /a/ or as a dead consonant, depending on the text context. This ambiguity resolution is further helped by the meta information like the geographical location of the inscription and its epigraphical time period and other details.

Please note that *pulli* is not just a virama which deletes the default vowel from Brahmic consonants. *Pulli* is a vowel reducer which includes the function of traditional virama in some contexts.

Pulli in the Current Proposal

Pulli is treated in three different ways in the proposed encoding in [L2/16-068](#), depending on whether the base is an independent vowel, a dependent vowel, or a consonant. In case of a consonant, *pulli* becomes Indic Virama character, in its own right. However, for the independent and dependent vowel forms, both vowel character and *pulli* character is atomically encoded as a composed form, as listed below:

𑀧𑀺	VATTELUTTU VOWEL LETTER E
𑀧𑀻	VATTELUTTU VOWEL LETTER EE
𑀧𑀼	VATTELUTTU VOWEL LETTER AI
𑀧𑀽	VATTELUTTU VOWEL LETTER O
𑀧𑀾	VATTELUTTU VOWEL LETTER OO

- ✓◌̇ VATTELUTTU VOWEL SIGN E
- ✓◌̈ VATTELUTTU VOWEL SIGN EE
- ◌̆ VATTELUTTU VOWEL SIGN AI
- ✓◌̊ VATTELUTTU VOWEL SIGN O
- ✓◌̋ VATTELUTTU VOWEL SIGN OO

Issue with the Current *Puḷḷi* Representation

The treatment of *puḷḷi* in the current proposal is analogous to an encoding model where ‘ä’ (breve over ‘a’) is encoded atomically and ‘č’ (breve over ‘c’) is encoded as a decomposed form. This treatment is not consistent with the *puḷḷi* behavior as explained above.

It needs to be considered that both the scenarios—vowels with *puḷḷi* and consonants with *puḷḷi*—occur with similar frequency in texts, if not, consonants with *puḷḷi* are found more. So if vowels with *puḷḷi* are encoded atomically, then consonants with *puḷḷi* should also be treated in the same fashion.

To put things in context, the current proposal ([L2/16-068](#)) seems to be biased towards Velvikudi plates. Velvikudi plates represent only one tradition among many that existed across geographies and time periods. Specifically, the number of inscriptions with *puḷḷi* as in Velvikudi plates is limited to those found only in Tamil Nadu in the 8th century. Please note that Vatteluttu existed till the 11th century all across Tamil Nadu and Kerala, and continued for more centuries in Kerala. So *puḷḷi* with vowels cannot be considered as the main mechanism for representing vowels E and O.

Moreover, current proposal would require to add following canonical equivalences or Do-Not-Use table in the standard, to guide users or implementations to the prescribed encoding as multiple sequences that could be visually identical are present:

Use	Canonical Equivalence / Do Not Use
INDEPENDENT VOWEL E	<INDEPENDENT VOWEL EE, VIRAMA>

INDEPENDENT VOWEL O	<INDEPENDENT VOWEL OO, VIRAMA>
DEPENDENT VOWEL SIGN E	<DEPENDENT VOWEL SIGN EE, VIRAMA>
DEPENDENT VOWEL SIGN O	<DEPENDENT VOWEL SIGN OO, VIRAMA>

Requested Change on the *Puḷḷi* Representation

To address the above mentioned issues in the current *puḷḷi* representation, we have two choices:

1. *Puḷḷi* encoded as a separate combining mark; without atomically encoding any of the composed forms with *puḷḷi*.
2. Every attested <vowel, *puḷḷi*>, <vowel sign, *puḷḷi*>, and <consonant, *puḷḷi*> composed forms are atomically encoded.

To be closer to the Indic encoding model, it would be prudent to go with the first approach which is also the treatment of *virama* in Brahmi in its original proposal [L2/03-249](#). That is, separately encoded *puḷḷi* combining mark can be attached to an independent vowel, dependent vowel sign, or a consonant. Virama getting attached to vowels or dead consonants has a precedence in Malayalam. This would mean following changes to the proposal in [L2/16-068](#):

Current Proposal (L2/16-068)	Requested Change
VIRAMA	PULLI
INDEPENDENT VOWEL E	Not to be encoded. Represented as <INDEPENDENT VOWEL E, PULLI>
INDEPENDENT VOWEL O	Not to be encoded. Represented as <INDEPENDENT VOWEL O, PULLI>
DEPENDENT VOWEL SIGN E	Not to be encoded. Represented as <DEPENDENT VOWEL SIGN E, PULLI>
DEPENDENT VOWEL SIGN O	Not to be encoded. Represented as <DEPENDENT VOWEL SIGN O, PULLI>
INDEPENDENT VOWEL EE	INDEPENDENT VOWEL E
INDEPENDENT VOWEL OO	INDEPENDENT VOWEL O
DEPENDENT VOWEL SIGN EE	DEPENDENT VOWEL SIGN E

DEPENDENT VOWEL SIGN OO	DEPENDENT VOWEL SIGN O
-------------------------	------------------------

Since the script is not encoded yet, we can match the encoding model as much as possible with the behavior of the script. We will not get this choice once it is encoded.

Acknowledgements

I sincerely thank Jan Kucera for directing me to Velvikudi plates and pointing out recent changes in Brahmi. I also thank Sunil V.S. for his very helpful review of this document and great suggestions.

References

1. [L2/03-249](#): Proposal for the Encoding of Brāhmī in Plane 1 of ISO/IEC 10646
2. [L2/16-068](#): Preliminary proposal to encode Vatteluttu in Unicode
3. [L2/19-402](#): Proposal to Encode 6 Characters in the Brahmi Block

* END *