

Proposal to encode MALAYALAM LETTER CHILLU M

Cibu Johny, cibu@google.com
2014-Jan-08

Introduction

In the Malayalam script, a Chillu or Chillaksharam is a special vowel-less form of a consonant. In Unicode, these special vowel-less forms of Malayalam are given their own codepoints: 0D7A to 0D7F as Chillu NN, N, RR, L, LL and K (in that order):

ൺ റ്റ ൽ ള്ല ഴ്ല ക്ക

MALAYALAM LETTER CHILLU LLL is already proposed and in the pipeline with codepoint 0D4F and is at stage 5 as of writing this document.

ഴ

An additional Chillu character is proposed in a separate document L2/14-XXX: MALAYALAM LETTER CHILLU Y:

യ്

Now a yet another Chillu character is proposed: MALAYALAM LETTER CHILLU M:

മ്

While this Chillu M is not in use today, it has been attested in Malayalam documents from 1730(approximate) to 1851, by multiple authors, in manuscript and print as well. So far it is seen to be used for Sanskrit text only. It must be encoded for recording those documents.

Attestations

Attestations from various sources are listed below as sources organized as subsections.

Grammatica Grandonica¹, Manuscript by Johann Ernst Hanxleden S.J. (1681-1732)

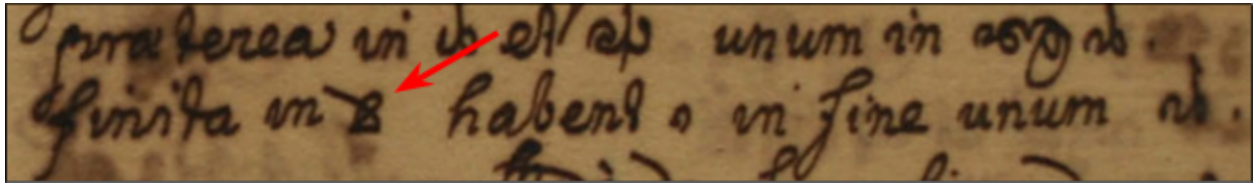


Image 1: Indicates word final *m* appears as chillu M.

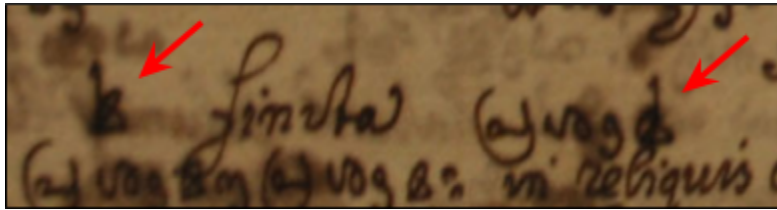


Image 2: Indicates word final *m* appears as chillu M and shows an example word $\langle \text{പ്രശം}$$\langle \text{chillu M}>$.

Malayā|abhāṣā Vyākaraṇaṁ (1851), Lithograph by Rev. Dr. Hermann Gundert

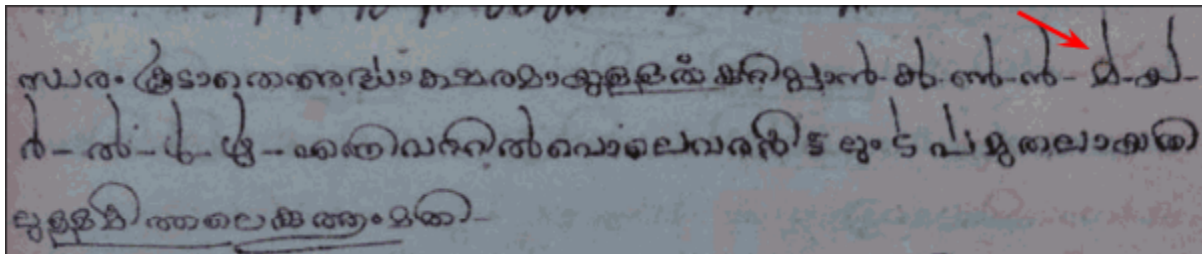


Image 3. Page 3 of the lithograph lists various Chillus of Malayalam script.

¹ http://opus.kobv.de/ubp/volltexte/2013/6321/pdf/hanxleden_grammatica.pdf pages 60 and 70.

South Indian Scripts in Sanskrit Manuscripts and Prints² (2001), by Reinhold Grünendahl

<i>Prepausal Consonants</i>			
k	ക & ക & ക	r	ര & ര
ṅ	ങ & ങ	l	ല & ല
t	ത & ത & ത	ḷ	ഴ & ഴ
n	ന & ന	ḻ	ഴ
m	മ		

Image 4: Lists various chillus in practice

Aspects of Manuscript Culture in South India³ (2012) by Saraju Rath

Author Dr. Anna Esposito describes the difference between Anusvara and Chillu M.

bit arbitrary. The end of a sentence is at times indicated by the dissolution of the *sandhi*, by the preservation of the *visarga* or by a final *m* with *virāma* instead of the *anusvāra*.¹⁷ After the end of longer text

¹⁷ The *anusvāra* is indicated by a small centred circle (o), the *m* with *virāma* looks like a small copy of the character *ma* (മ) with the diagonal line protruding at the top (ḻ).

Nasals in combination with consonants: A nasal, in combination with a palatal, cerebral, dental or *r*, is written with ligature. In combination with a guttural, however, a nasal is usually replaced by the *anusvāra* (o) in order to avoid difficult ligatures. For the same reason, either the *anusvāra* or the special sign for *m* with *virāma* (ḻ) appears before labials. Additionally, the *m* with *virāma* can sometimes be found before *v*.

Image 5

² <http://books.google.com/books?id=ApAn2YZlz6wC&lpg=PP1&pg=PA92> page 92

³ <http://books.google.com/books?id=FaZs3T8hRT0C&lpg=PP1&pg=PA85> pages 85, 86, 89

Dr. Anna is referring to 400 year old manuscripts like the one shown below⁴:

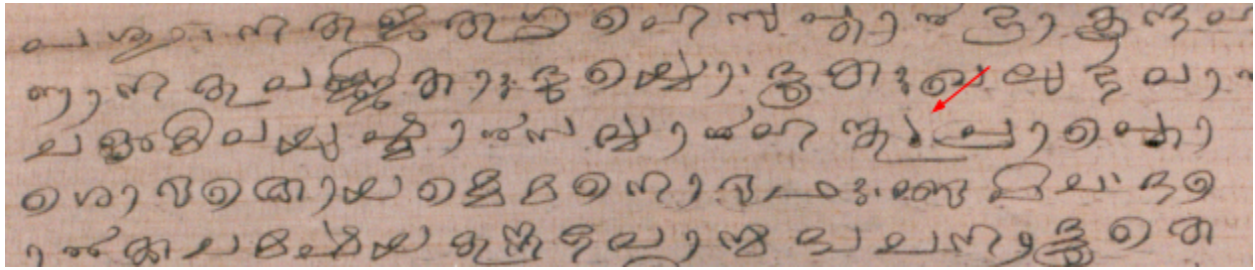


Image 6: The word using Chillu M being: ഹന്തം<chillu M> (*hantum*)

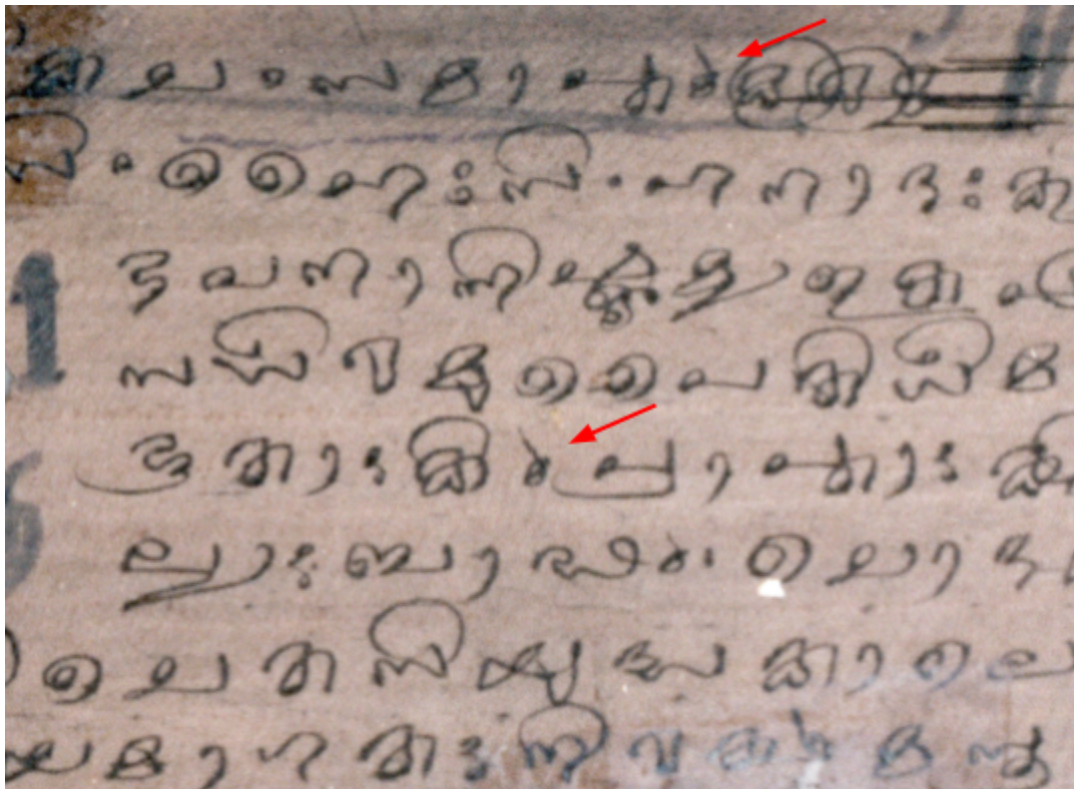


Image 7: The words using Chillu M: സമപ്തം<chillu M> (*samaptam*), കി<chillu M> (*kim*)

Based on the above attestations we can see that this character was in use in the later half of the 18th century and the early half of the 19th century. Since then it has been replaced by the Anusvara: ൦൦ (0D02)

However, given that the chillu representation of vowelless consonants has a distinct encoded form than their virama representation⁵, for faithfully recording as digital text the original content of such old prints/manuscripts as cited above, it is justified and necessary to encode a MALAYALAM LETTER CHILLU M.

⁴ <http://www.bhasa.indologie.uni-wuerzburg.de/mss/TRV3168/103.html> which is part of the Bhasa project at <http://www.bhasa.indologie.uni-wuerzburg.de/>

⁵ Proposal to encode Malayalam Chillaksharams, L2/06-207.

Choice of codepoint

It is suggested that this character may be placed together with the Chillu LLL which is in stage 5. For that, Chillu LLL may be moved to OD56 and Chillu Y be placed above that in the same column at OD55. Chillu M could be placed above that in the alphabetic order. The proposed codepoint is **OD54**.

Malayalam Code Chart

	OD0	OD1	OD2	OD3	OD4	OD5	OD6	OD7
0	◌̇	ഐ	ഠ	ര	ീ		ഋ	ധ
1	◌̈		ഡ	റ	ു		ൠ	ൡ
2	◌̉	ഓ	ഌ	ല	ൂ		ൣ	൤
3	ഃ	ഔ	ണ	ള	ൃ		൦	ൠ
4		ഔ	ത	ഴ	ൂ	ൡ		ൢ
5	അ	ക	ഥ	വ		യ	ൣ	൤
6	ആ	ഖ	ഭ	ശ	ഌ	൦		ൡ
7	ഇ	ഗ	ധ	ഷ	ഌ	ൠ	ൡ	ൢ
8	ഈ	ഘ	ന	സ	ഌ	ൣ	൤	൥
9	ഉ	ങ	ണ	ഹ		൦	ൠ	ൡ
A	ഊ	ച	പ	ട	ഌ	ൡ	ൢ	ൣ
B	ഋ	ൠ	ൡ	ൢ	ൣ	൤	൥	൦
C	ൠ	ൡ	ൢ	ൣ	൤	൥	൦	ൠ
D		ൡ	ൢ	ൣ	൤	൥	൦	ൠ
E	ൡ	ൢ	ൣ	൤	൥	൦	ൠ	ൡ
F	ൡ	ൢ	ൣ	൤	൥	൦	ൠ	ൡ

Legend

White: Published as of Unicode 6.3.

Green: Accepted and at advanced stage not permitting changes.

Yellow: Fractions, accepted.

Orange: Chillu LLL accepted but preferably to be moved to group with Chillu Y.

Blue: New location proposed for Chillu LLL.

Purple: Locations of ongoing proposals of Anusvara Above, Chillus M and Y, Vertical Bar Virama and Symbol Para.

Red: Location proposed for Circular virama.

Proposed character with properties



OD54;MALAYALAM LETTER CHILLU M;Lo;0;L;;;;N;;;;

Thanks

Sri Nidhi for providing the scans of various manuscripts which are key attestations for this proposal.

Official Proposal Summary Form

(Based on N3902-F)

A. Administrative

1. Title

Proposal to encode MALAYALAM LETTER CHILLU M

2. Requester's name

Cibu Johny

3. Requester type (Member body/Liaison/Individual contribution)

Individual contribution

4. Submission date

2014-Jan-08

5. Requester's reference (if applicable)

6. Choose one of the following: This is a complete proposal (or) More information will be provided later

This is a complete proposal.

B. Technical – General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters), Proposed name of script

No

1b. The proposal is for addition of character(s) to an existing block, Name of the existing block

Yes, Malayalam

2. Number of characters in proposal

1 (one)

3. Proposed category

Category B1, specialized small (for this character)

4. Is a repertoire including character names provided?

Yes

4a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?

Yes

4b. Are the character shapes attached in a legible form suitable for review?

Yes

5. Fonts related:

a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?

Michael Everson

b. Identify the party granting a license for use of the font by the editors (include address, e-mail etc.)

Michael Everson, everson@evertype.com

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

Yes

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes

7. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

Yes

8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script.

See detailed proposal.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

No

2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

Yes

2b. If YES, with whom?

The author is himself one of the user community.

2c. If YES, available relevant documents

None specifically. The matter was discussed via email.

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

Those (scholars and enthusiasts) who desire to store old Malayalam documents using this character.

4a. The context of use for the proposed characters (type of use; common or rare)

Rare

4b. Reference

See detailed proposal.

5a. Are the proposed characters in current use by the user community?

No. This character was in historical use.

5b. If YES, where?

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

Yes

6b. If YES, is a rationale provided?

It belongs in the Malayalam block which is in the BMP.

6c. If YES, reference

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

Only one character is proposed.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

No

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences?

No

11b. If YES, is a rationale for such use provided?

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?

13c. If YES, reference: