

Malayalam: Comparison of Dot-Reph Proposals in 5.1

Cibu C. Johny (cibu@google.com)



Umesh Nair (umesh@google.com)

Even though a proposal for Malayalam dot-reph and an alternative proposal are given in <http://www.unicode.org/versions/Unicode5.1.0>, their relative merits and problems are not discussed in that document. We are trying to list them below:

1. Problems with both of the proposals

Deviates from the model common to rest of the Indic. That is, $\langle RA, \textit{Virama}, \textit{consonant} \rangle$ does not form the dot-reph when consonant is YA or VA.

2. Problems specific to proposed encoding

- Does not have a representation for dot-reph forms of  (Dot-Reph on YA - Table 4a Row 4) and  (Dot-Reph on VA - Table 4b Row 4). Dot-Reph on VA is found in prints before 1970. Even though no evidence is presented for Dot-Reph on YA, it is definitely possible in the writing system.
- When consonant is YA or VA, $\langle RA, \textit{Virama}, \textit{consonant} \rangle$ has to be rendered differently as a special case.



3. Problems specific to alternate encoding

It is violating the common understanding that *'The virama sign nominally serves to suppress the inherent vowel of the consonant to which it is applied'*. First of all, this is a phonetic property. In Malayalam, Virama is used with different phonetic properties as well. For example, Samvruthokaram sequence is $\langle U\text{-sign}, \textit{Virama} \rangle$ in which there is no inherent vowel to be removed. Also, Chillus are new entities not found in rest of the Indic. The interaction of Virama with Chillus is yet to be defined.

4. Advantage of alternate encoding

It goes well with common user understanding that dot-reph is the short notation of Chillu-RR.

5. Following sequences cannot be used for Dot-Reph

- $\langle RA, \textit{Virama}, \textit{ZWJ}, \textit{Consonant} \rangle$ forming dot-reph, since $\langle RA, \textit{Virama}, \textit{ZWJ} \rangle$ should be forming Chillu RR to support existing data.
- $\langle RA, \textit{Virama}, \textit{YA} \rangle$ forming dot-reph. Currently, $\langle RA, \textit{Virama}, \textit{YA} \rangle$ is used for  which is a very common cluster found in Malayalam words for sun, wife, subject etc. We may not force the usage of ZWJ in such cases. That is, we must avoid proposing $\langle RA, \textit{ZWJ}, \textit{Virama}, \textit{YA} \rangle$ as the only sequence forming  .

6. Other potential uses of alternate proposal

There is evidence for subscribing Chillus. The sequence <Chillu, Virama, Consonant> could be used for rendering those visuals.

7. More proposals and their problems

7.1 <RA, VIRAMA, ZWJ, consonant>

- <RA, Virama, ZWJ> is already used up by Chillu-RR for supporting existing data.
- Fallback to Chillu-RR is not always valid. For example, സൂര്യൻ is 'Sun' in traditional orthography. If dot-reph is falling back to Chillu-RR, then we would get സൂര്യൻ which is an incorrect spelling. The correct spelling in modern orthography is സൂര്യൻ with full RA and C2-conjoining YA.

7.2. <RA, VIRAMA, ZWJ, Consonant> when Consonant is YA or VA and <RA, VIRAMA, Consonant> elsewhere

- <RA, Virama, ZWJ> is already used up by Chillu-RR for supporting existing data. However, the impact is only in the context of YA or VA.
- Like in 7.1, fallback to Chillu-RR is not always valid.
- Overhead of special case processing.

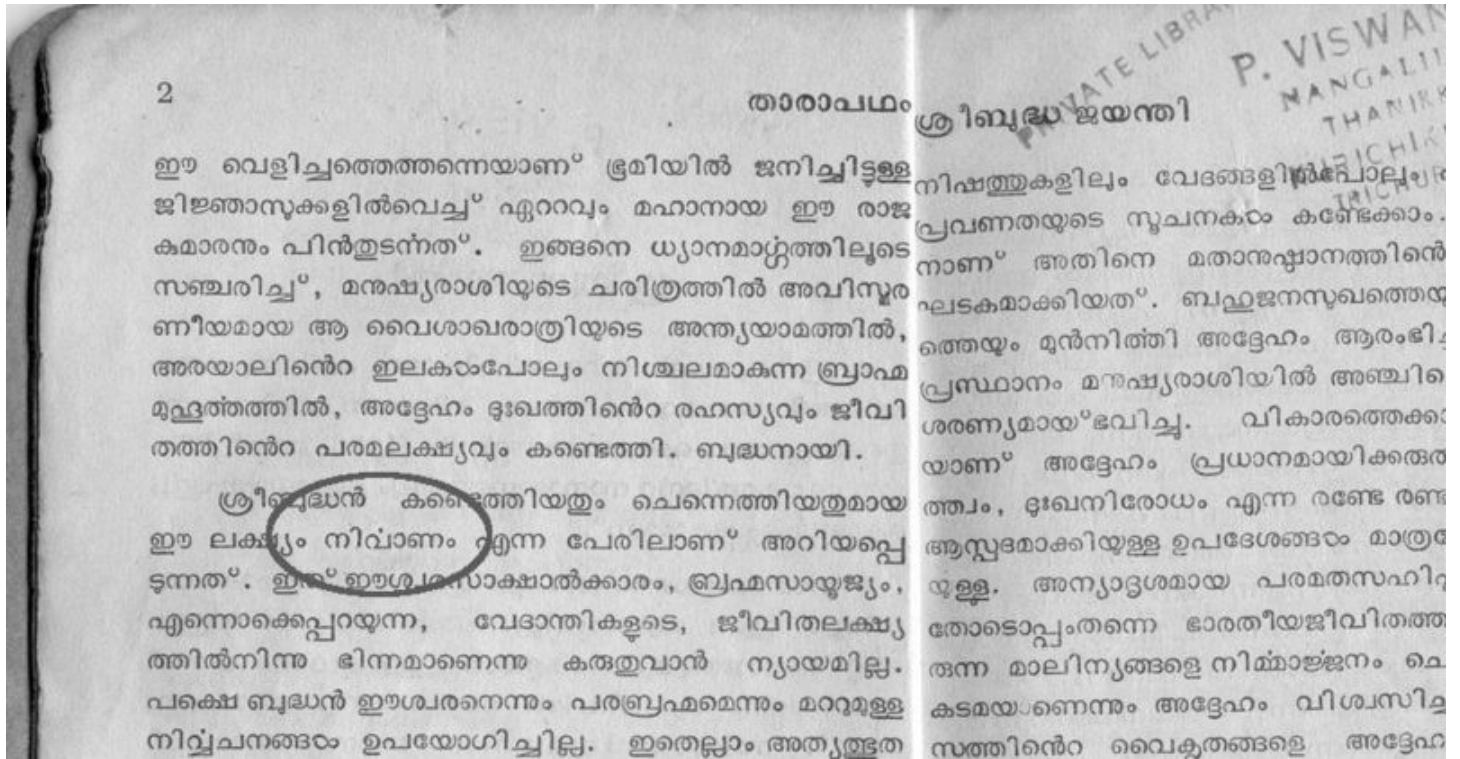
7.3 <RA, VIRAMA, ZWJ, ZWJ, consonant>

- Logic being <RA, Virama, ZWJ> would be Chillu-RR and it is joined with consonant by yet another ZWJ.
- One Problem would be, ignorable nature of zwj.

7.4 <Chillu-RR, ZWJ, Consonant>

- Dot Reph is always part of the grapheme cluster of Consonant. So all rendering engines and other implementations would need to change the existing C (V C)*(V|Vs)* pattern for finding an Indic grapheme cluster. This is a major change. (C = Consonants+Chillus, V = Virama, Vs = Vowel sign)
- Like in 7.1, fallback to Chillu-RR is not always valid.

Scans



Dot-Reph over single VA - from Malayalam textbook for Grade 10 in the year 1966

പാർവ്വണം pārvanam S. Connected with പാർവ്വം, as പാ'ണേരു the full-moon.

പാർവ്വതി pārvati S.(പാർവ്വതം)=മലമാതൃ; N.pr. f.

പാർവ്വത്യം pārvatyam (Tdbh. of പ്രാവൃത്യം? as if fr. പ്രവൃത്തി q. v.; but T. Te. V1. have pārapatyam, C. pā ^{A Malayalam English Dictionary By Herрман Gundert, 1872} nagement, as if fr. പാക്ക). 1. Local government, administration മോർ വില്ലുന്നതായേ ഉൾരിലേ പാർവത്യം എന്തിനു prov., also പാറോത്യം; നാട്ടിൽ കാരോരുത്തർ പാറോവത്തി ചെയ്യൂ, പാറവത്യം ചെയ്യുന്നവൻ TR. 2. Trav. a subordinate revenue situation.

പാർവ്വതുകാരൻ a subordinate Collector & Magistrate (Trav. = അധികാരി). സൂലൂക്കൻറ പാർവ്വതുകാരന്മാർ, പാർവ്വതുകാരൻ TR.(In Mal. a district about as large as 4 amśams (അംശം 2) together, was formerly under 1 പാർവ്വതുകാരൻ: f.i. പാർവ്വതിക്കാർ പണത്തിനു വരാറായി No. said only by very old people).

പാർശ്വാപ് P. Pādshāh, The Mogul emperor

Dot-Reph over single VA - year 1872

പരവശപ്പെടുന്നു.
 പരവുന്നു.
 പരിചാരകൻ.
 പരിഭാഷ.
 പരിഹാരം.
 പരിഹാസം.
 പരീക്ഷിക്കുന്നു.
 പരീക്കൻ.
 പർവ്വതം.
 പലം.

v. n. To be distressed.
 v. n. To extend.
 s. m. A man servant, a courtier.
 s. n. A translation.
 s. n. A remedy, a means of averting.
 s. n. Ridicule.
 v. a. To try, to test.
 adj. Coarse.
 s. n. A mountain.
 } s. n. A weight of gold or silver, the
 100 part of a Toolam.

Dot-Reph over single VA from year 1851

97

ആൾമാരി—ആഴക

- l. ആൾമാരി ālmāri Port. = അൽമാരി Almeida.
- y ആൾമാർ ālvār & ആഴമാർ T. M. C. Te.
- 2 ("the rulers" or "the humble") Title of the
- b 12 Vishnu apostles. see ആഴമാർ.
- n ആഴകൾ ālakku (T. C. ആഴക്കൾ) A handful =
- a ചോങ്ങ (prh. from അര ഉഴക്ക) the 1/4 of a നാ
- :- ഴി =
- r. a. പത്തൊന്നു a half ālakku. A. Nāi.

A Malayalam and English Dictionary
 By H. Gundert, 1872 Mangalore, by C. Stolz - Basel Mission
 and Tract Depsitory, Trubner & Co. - London.

Subscripting LLA-chillu