

# Malayalam: Comparison of Dot-Reph Proposals(Re: L2/08-411)

Cibu Johny cibu@google.com

2008-08-08

## Update 1: South Asia Subcommittee Report (L2/08-411)

### E.2.1:

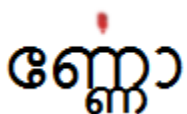
In addition to the alternatives presented in the document, we considered a spacing character that occurs before the consonant, but is rendered as a dot above the consonant, rather than a combining mark occurring after the consonant. It's a variation on D, let's call it D'.

Ideally, the usual mechanisms, RA Virama {ZWJ/ZWNJ} Cons, would be preferred. However, those mechanisms are already used in legacy documents for the representation of chillus. For that reason, and considering keyboard interactions and collations, D' seems the most promising path.

We recommend to solicit input from the Government of India, and make a resolution at the February 2009 UTC meeting.

### Response from the author:

The proposal by the committee is excellent, if it works with various rendering systems. Please note that, the MALAYALAM REPH being encoded needs to work as a preceding combining mark to a consonant. For example, two part vowel signs should go around MALAYALAM REPH and its consonant/cluster. As show below:



/rnnO/ <REPH, NNA, VIRAMA, NNA, OO SIGN>

## Introduction

Dot style Reph is used to represent Malayalam letter <RA> as a dot over the consonant following it. That is, in the sequence <Malayalam Akshara X, r, Akshara Y>, The consonant 'r' is represented as a dot over Y. Please see scans below. As of Unicode 5.1, there is no stable mechanism to represent dot-reph. Following are the various proposals had been put forward.

# Proposals

A	RA, VIRAMA, ZWJ, Consonant
B	RRA, VIRAMA, ZWJ, Consonant
C	Chillu-RR, VIRAMA, Consonant
D	Consonant, Dot-Reph combining mark (need to be encoded for this proposal)
E	RA, VIRAMA, Consonant
F	RRA, VIRAMA, Consonant

## Problems

Above proposals share some common problems. They are listed below.

### 1. C1-conjoining and C2-conjoining forms need to be differentiated (applicable to E, F)

Both C1-conjoining and C2-conjoining forms of the cluster needs to be supported and differentiated in plain text. These proposals would make C2-conjoining form unable to be represented. Example: C2-conjoining ligature of <RA, Virama, YA> is a very common cluster found in Malayalam words for sun, wife, subject etc.

### 2. Deviation from standard definition of Virama (applicable to C)

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'*. However, this is a phonetic property. In Malayalam, Virama is used with different phonetic properties as well. For example, Samvruthokaram sequence is <U-sign, Virama> 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. Deviation from common-man's understanding about Dot-reph (applicable to A, B, E, F)

Common man's understanding of dot-reph is that it is a short form of Chillu-RR.

### 5. Deviation from etymology of Dot-reph (applicable to B, C, F)

Etymologically dot-reph is a short form for <RA, VIRAMA>.

### 6. Uses ignorable characters (applicable to A, B)

Using ignorable characters would be problematic when they get ignored in many usages or implementations. That would make the text illegible.

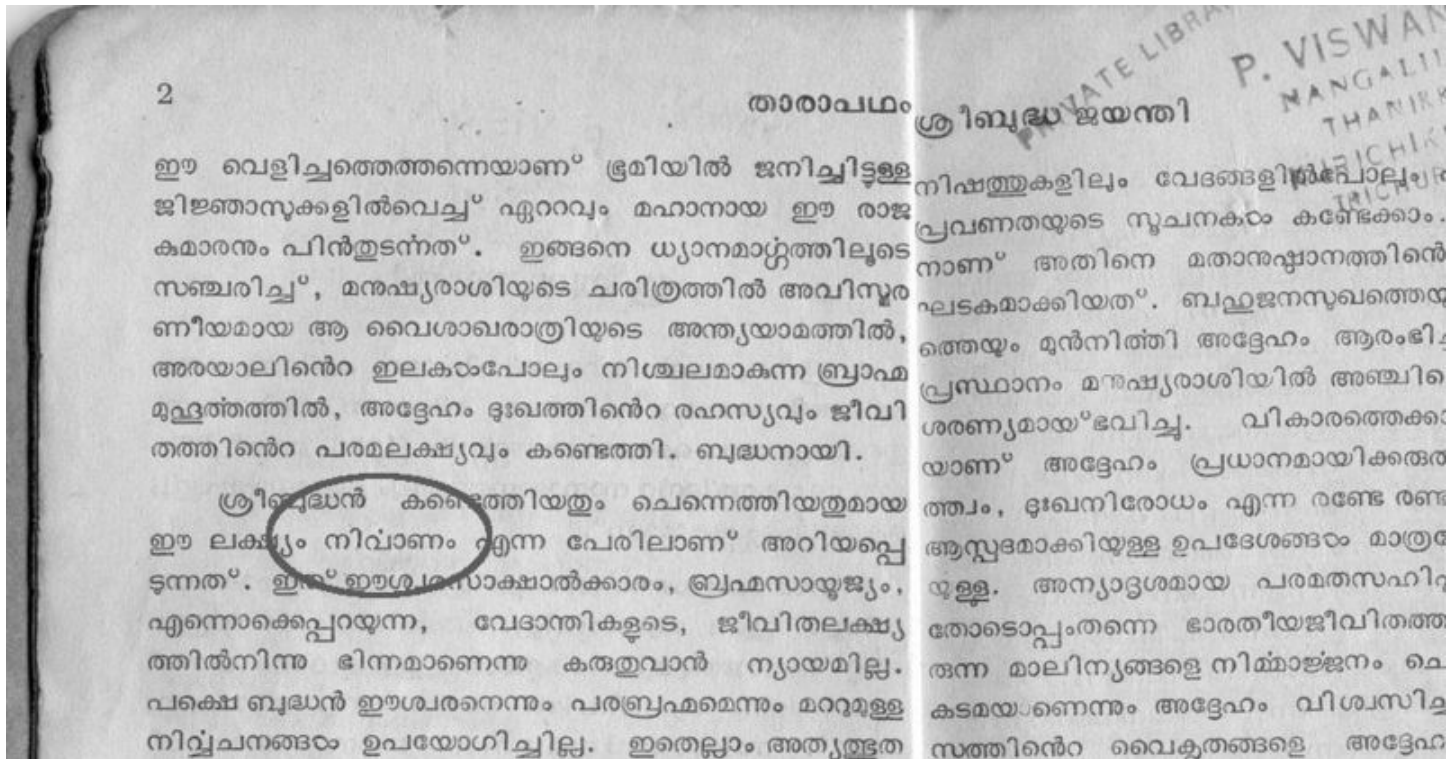
## 7. Collation issues (applicable to D)

In the cluster involving dot-reph, users expect dot-reph to be the first component. However, Unicode can only have a combining mark as the second component of the cluster. This would require, all clusters involving dot-reph to be listed in the collation tables explicitly as contraction-expansion sequences. Approximate count of such clusters present in existing texts are around 40.

## 8. Conflicts with existing data (applicable to A)

<RA, Virama, ZWJ> would be forming Chillu-RR to support existing data.

### 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 <sup>A Malayalam English Dictionary By Herрман Gundert, 1872</sup> 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). സൂലൂക്കൻ 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, SSA - year 1872

പരവശപ്പെടുന്നു.  
പരവുന്നു.  
പരിചാരകൻ.  
പരിഭാഷ.  
പരിഹാരം.  
പരിമാസം.  
പരീക്ഷിക്കുന്നു.  
പരക്കൻ.  
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