Malayalam/Dot-Reph

From Unicode discussion

< Malayalam

```
Cibu C. Johny
cibu@google.com
2007-01-30
```

1 Introduction

2 Current Status

Unicode 5.0 says the Dot-Reph is <RA, VIRAMA, C>. This encoding produces following unacceptable results:

1. The word കാൎയ്യം is rendered in a modern font as കാര്യ്യം.
2. The word പാൎവ്വതി is rendered in a modern font as പാര്വ്വതി.

Instead of the current encoding, author proposes to encode Dot-Reph atomically to meet following objectives:

- A correct text written using traditional orthography font should get displayed in a modern orthography font (and vice verse) without spelling mistakes or unreadable rendering.

- Traditional writer should have the flexibility to mix Dot-Reph and Chillu-RA/RRA in the text without switching the font.

3 Arguments

3.1 Dot-Reph is not same as Chillu-RA/RRA or <RA, VIRAMA>

Dot-reph is not used in modern script. This does not mean that it has been replaced by Chillu-RA/RRA or <RA, VIRAMA>. Instead, modern script devised a different method (transcribe) to write words which had Dot-Reph in them.

Both of the following examples provide the evidence that, equivalences or fall-backs between Dot-Reph and Chillu-RA/RRA or <RA, VIRAMA> are not reliable.

3.1.1 Example 1

A traditional writer using old orthography font, writes /kaaryam/ (meaning: matter, subject) as: കാൎയ്യം. All following potential fallbacks for a modern orthography font are incorrect:

- കാരൎയ്യം
- കാർയ്യം
- കാര്യ്യം

That is, the word should be rendered as കാൎയ്യം in any type of font.

In this case, the transcribing of <RA, YA, YA> from old orthography to new is made as <RA, YA> to get കാര്യം.

3.1.2 Example 2

Suppose Dot-Reph is made equivalent to Chillu-RA/RRA and a young writer using modern font writes /paarvathi/ as പാർവതി. A person using traditional (old orthography) font would see the incorrect rendering as പാവ്തി while he expects പാവ്വതി.

In this case, the transcribing of <RA, VA, VA> from old orthography to new is made as <Chillu-RA/RRA, VA>.

4 Proposed Solution

4.1 Encode Dot-Reph atomically

As of now, there is no reliable mapping between Dot-Reph and Chillu-RA/RRA or RA_dead. Transcribing is the only option between traditional and modern texts. So the words with Dot-Reph should be displayed as it is, irrespective of the nature of font. That is guaranteed only by encoding Dot-Reph.

This will give complete control for the writer to choose the style he want. This paradigm is similar to the decision to independently encode Chillu-RA/RRA.

4.1.1 Problems

Need to separate Dot-Reph from its base character RA and Chillu-RA/RRA.

5 Alternate Solution

5.1 Mandate Dot-Reph implementation in a font

Unicode can mandate a Dot-Reph implementation in a Malayalam font. Thus following produce Dot-Reph in any Malayalam font.

- <ര, VIRAMA, യ, VIRAMA, യ>
- <ര, VIRAMA, വ, VIRAMA, വ>
- <ര, VIRAMA, ഭ, VIRAMA, ഭ>
- ...

This would not conflict with Chillu-RA/RRA since it is being atomically encoded.

Retrieved from "http://unicode.wikia.com/wiki/Malayalam/Dot-Reph"