## Beyond Roman fonts: Extra dimensions in Malayalam fonts

Hussain KH, Rajeesh KV, Aravind Rajendran

### Abstract

The Malayalam alphabet consists of 51 basic characters which combine to form more than 900 conjuncts (ligatures) in traditional script. Unlike Roman scripts, Malayalam has two kinds of ligatures, namely Horizontal and Vertical conjuncts. Vertical conjuncts in Malayalam are unusual among other Indic scripts, and almost never seen in Roman scripts. Anatomically there are two parts in vertical conjuncts — Above Character and Below Character.

Vertical conjuncts demand careful attention in space management in Malayalam types. Accommodating the below character in the below-base(line) space when adhering to Roman metrics poses serious inconveniences. Compressing all levels into a single level harms the shapes of around 700 vertical conjuncts, as Malayalam poses extreme cases of divergence while dealing with vertical conjuncts. Designing vertical conjuncts results in many deviations from the accepted norms of Roman typography. Deviating from Roman metrics poses problems of point sizes when typesetting documents using Malayalam and Latin text together.

Thus, creating original Malayalam script fonts while satisfying dimensions of Roman fonts exerts formidable pressure on the designers. Malayalam typography looks at geometrical consistency in a different way than Roman typography.

## 1 Malayalam script and Rachana

Malayalam, the mother tongue of 45 million people in Kerala, the southernmost state in India, is one of the 22 official languages of India. It is a 1600-year-old Dravidian language, and its script is classified as abugida, or alphasyllabary. Unicode support for Malayalam script was in GNU/Linux systems around 2002 and MS Windows XP added support in 2004. Since then Malayalam language technology has seen significant advances, thanks to Swathanthra Malayalam Computing (`http://www.smc.org.in`) and its Rachana font based on the traditional script.

**Figure 1**: Ligatures
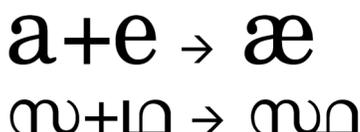


**Figure 2**: Horizontal and vertical conjuncts.
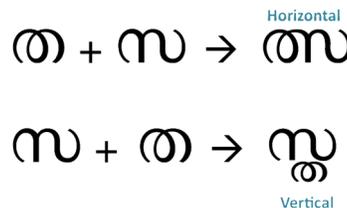


Unicode encodes the basic characters in Malayalam [1] while all the complex conjunct characters are supported by OpenType shaping rules (see Appendix A for information about the glyphs in the Rachana font). The Malayalam script also requires 'complex text shaping' support from shaping engines (notably HarfBuzz, used by X$_{\text{E}}$T$_{\text{E}}$X and (as of 2020) LuaLAT$_{\text{E}}$X) to properly shape its conjuncts and vowel symbol combinations.

### 1.1 Old vs. New script

'New lipi' or Reformed script of Malayalam, which tried to simplify the original script by breaking many conjuncts and vowel combinations, appeared in the 1970s as an official government effort to use Malayalam with English typewriters and Linotype typesetting. ASCII fonts with 140 characters based on the New script were later popularized by desktop publishing (DTP). New script is still in vogue in typesetting, though it is only a subset of the original script with 900 conjuncts standardized for printing by Benjamin Bailey in 1824. New lipi with its limited conjuncts can easily be accommodated in fonts with Roman font metrics/dimensions.

### 1.2 Rachana

The language campaign named 'Rachana' in 1999 began advocating for use of the traditional script in Malayalam computing. After the advent of Unicode Malayalam in 2002, the traditional script (Original Script), popularly known as 'Old Script', has had a steadily growing presence in the Web and in typesetting and printing. The traditional orthography demands more space beneath the baseline for vertical conjuncts. Designing fonts for original script

**Figure 3**: Metrics of below base character.
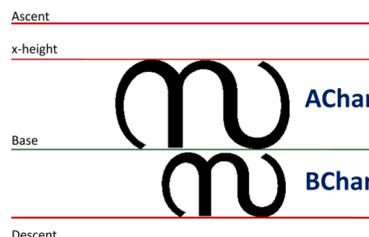
**Figure 4**: Below-base parts and space in Roman and Malayalam (basic characters) type metrics.
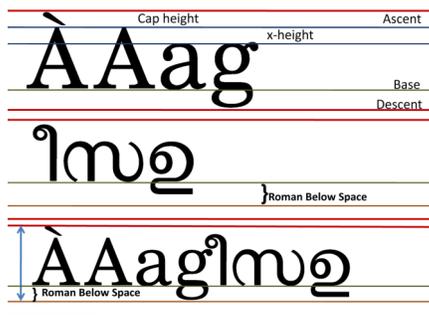
satisfying dimensions of Roman fonts exerts a lot of pressure on the designers.

In this discussion typography of Original Script (Old Lipi) is considered, since it is a superset of all variations of New Lipi that exist in various fonts used for DTP. Typesetting is now shifting from New Lipi to Old Lipi, thanks to growing usage of Unicode.

As mentioned, the Malayalam alphabet consists of 51 basic characters which combine to form more than 900 conjuncts (ligatures) in traditional script. The Rachana font, first designed in 1999 (so-called ASCII[1]) for the campaign for traditional script, underwent major modifications in 2000 and especially in 2006 when the font became Unicode-compliant and distributed under the GNU GPL. Now, after a long period of 15 years, all 1000+ glyphs in Rachana have been totally redesigned, taking more graphical liberties with fewer constraints exerted by Roman typography. This paper explores new possibilities in Malayalam typography and how far beyond Roman typography it can go without compromising either aesthetics or functionality.

---

[1] Before the advent of Unicode, Indic script fonts followed an encoding similar to that of Latin scripts by laying out characters in an 8-bit table, known as ISCII. To accommodate the 900+ characters needed, Rachana was originally designed as a set of six fonts. The typesetter then manually switches fonts in a DTP program to pick specific characters.

**Figure 5**: Beyond Roman type metrics.

Hussain KH, Rajeesh KV, Aravind Rajendran
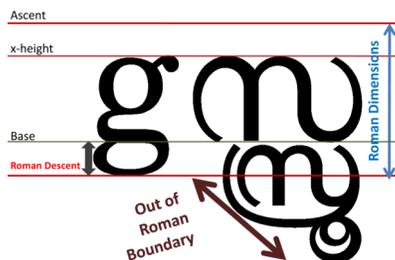
**Figure 6**: Below-base levels 1–5.

(a) Below-base levels 1–3.

(b) Below-base levels 4 and 5.
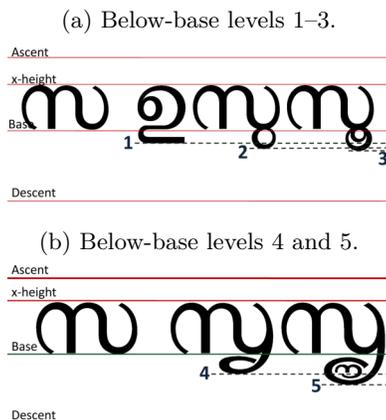
## 2 Ligatures/Conjuncts

Similar to Roman types, Malayalam also has ligatures, known as conjuncts. They are formed by combining basic characters (see an example in Fig. 1). The considerations of moving away from Roman typography mainly relate to conjuncts, which amount to nearly 20 times the number of basic characters.

### 2.1 Conjuncts: horizontal and vertical

Differing from Roman fonts, conjuncts in Malayalam are formed in two ways — horizontal and vertical. For example, Basic characters ഥ (tha) and സ (sa) combine horizontally to form ഥസ (thsa) i.e., ഥ + സ → ഥസ, whereas സ and ഥ combine vertically, സ + ഥ → സ്ഥ (stha). See Fig. 2 for a depiction of this.

### 2.2 Below-base characters

#### 2.2.1 AChar and BChar

Vertical conjuncts demand careful attention in space management of Malayalam types. Anatomically, there are two parts in vertical conjuncts — Above

**Figure 7**: Below-base levels 6–9.

(a) Below-base levels 6–8.
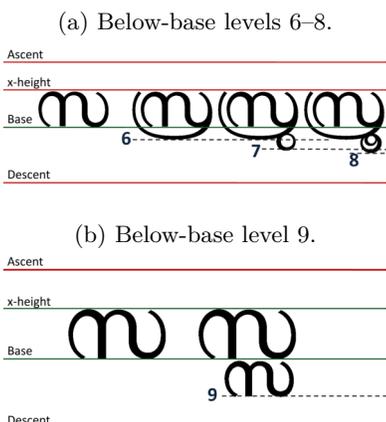
(b) Below-base level 9.

**Figure 8**: Below-base levels 10–18.

(a) Below-base levels 10–12.



(b) Below-base levels 13–15.



(c) Below-base levels 16–18.
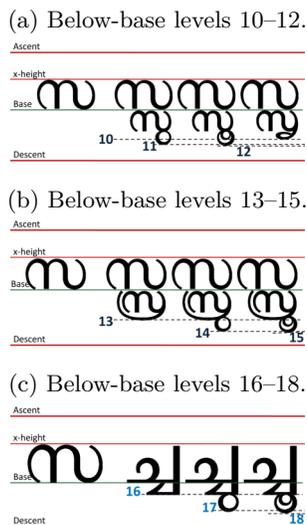


Character (AChar) and Below Character (BChar). While designing glyphs, AChars are placed above the baseline, filling the x-height. BChars are placed below the baseline and are always smaller in size than AChars, following the pattern in handwriting and calligraphy (see Fig. 3).

### 2.2.2 Below-base space

Parts of glyphs going below the baseline occurs in Roman glyphs, e.g., the letters 'g' and 'j'. This happens in Malayalam as well and its basic characters (adopted in the Malayalam Unicode chart) are fit well in the above and below spaces allocated as in normal Roman types (see Fig. 4).

### 3 Objective of the paper

When it comes to vertical conjuncts in Malayalam, accommodating BChars in the below-base space allocated in Roman metrics poses significant problems. The peculiar behaviour of 'space grabbing' of vertical conjuncts, as we'll see, is in perpetual collision with the Roman below-base space and BChars often go below the Roman descent boundary (an ex-

**Figure 9**: Extra descent in Rachana.
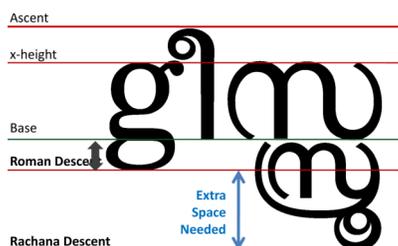


**Figure 10**: Below base space: varying vs. constant.
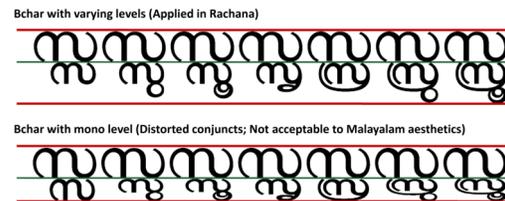


**Figure 11**: Equalizing levels 1, 4 and 6.



ample is in Fig. 5). The main objective of the paper is to show how this is circumvented in the typography of Rachana, going beyond Roman metrics.

### 3.1 Levels in below-base

The large vertical space demanded by BChars below the baseline is a challenge in Malayalam typography which in principal should be specially treated for appropriate leading (interline space). Owing to consonant-vowel pairing, BChars have 18 different heights/levels, explored in Figs. 6, 7 and 8.

### 3.2 Treating BChar levels

From Figs. 6a to 8c, it is evident that the vertical space between baseline and descent allotted in normal Roman fonts is too little to accommodate BChars. Malayalam glyphs require nearly the same space above and below the baseline (shown in Fig. 9). Below-base space cannot be compensated by taking from the Above-base space since some of the Malayalam vowel signs use the full cap-height; i.e., space above the base is needed exactly as in Roman fonts. Below-base space allocations in Roman type is woefully inadequate for Malayalam, leading to some unfortunate treatments in designing vertical conjuncts.

One of the solutions for accommodating BChars in Roman-descent is to squeeze all levels into a single level. This is graphically possible but produces distorted heterogeneous characters totally unacceptable to Malayalam aesthetics. In fact this kind of deformative practice is unacceptable to the typography of any script in the world. For instance, see the single-level BChars in Fig. 10.

**Figure 12**: Equalizing levels 7–9.

**Figure 13**: Equalizing below-base levels 10 and 13.
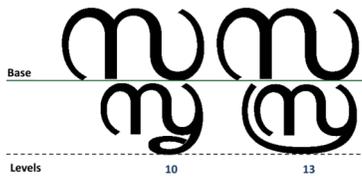
**Figure 14**: Equalizing below-base levels 11, 12, 14 and 15.
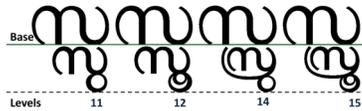
### 3.2.1 Equalizing near heights

Compressing all levels into one obviously harms the shapes of most vertical conjuncts in Malayalam. Reducing the number of levels can be achieved by another method. In earlier versions of Rachana, attempts were made to reduce vertical levels by equalizing near-heights, which reduced the number of vertical levels from 18 to 6. By this it was hoped to attain almost evenly sized BChars in vertical conjuncts. It could be a potential way to contain the unusual leading in Malayalam typesetting caused by lacunae in vertical conjuncts (Figs. 11, 12, 13 and 14).

### 3.3 Typographic deviations of BChar

Even though vertical conjuncts have two parts, it must be remembered that both are integral parts of a single character. Naturally one may expect both parts to follow the same typographic characteristics, but in reality they differ (rather, are forced to differ). Unfortunately, it is found to be impossible to keep the same types in above and below parts. The practice adopted in designing Rachana results in violations of basic rules of Roman typography (see Fig. 15).

Let us consider creating the vertical conjunct ന്ന which consists of the same character ന in above and below parts.

**Figure 15**: Typographic deviations of single conjunct ന്ന.

**Figure 16**: Typographic deviation of same character ന in the conjunct ന്ന.

(a) Superimposed AChar (red) and BChar (gray). Notice the curves marked with 1, 2 and 3.
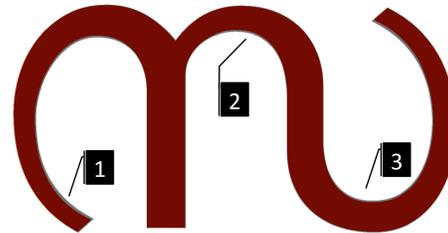
(b) Enlarged parts of Fig. 16a.
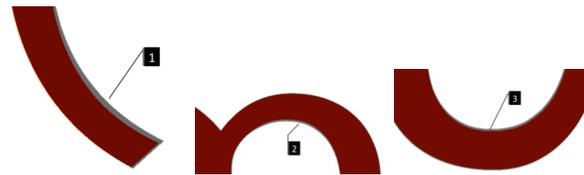
As seen in Fig. 15(1), when a same-sized AChar is placed in the below-base area, the conjunct produced is strikingly disproportional. The same shape with the same size produces an optical illusion of an oversized BChar. A BChar should invariably be smaller than the AChar in Malayalam orthography.

In Fig. 15(2), a 60% uniformly scaled BChar strictly adheres to the type design of AChar, but produces an unbalanced shape. The small BChar appears to suffer from pressure under the big AChar.

Fig. 15(3) shows a more balanced shape, achieved with non-uniform scaling. This BChar is designed with 70% horizontal and 60% vertical scaling. It produces a more pleasing effect compared to uniform scaling, at the same time not increasing the vertical size. More or less this proportion is followed in all versions of Rachana. Here a wider BChar 'supports' its counterpart in AChar and helps to achieve legibility at lower point sizes (10pt or 11pt) while typesetting.

Please observe, non-uniform scaling of BChar produces a *different type*! This can be verified by superimposing a same-sized AChar and BChar. As seen in Figs. 16a and 16b, the curve of BChar (light gray in colour) often varies from AChar (red online, dark gray in print) due to non-uniform scaling. This

**Figure 17**: An extreme case of below base conjunct.

Hussain KH, Rajeesh KV, Aravind Rajendran

**Figure 18**: Scaling adjustments for glyph in Fig. 17.

(a) BChar with 70%–60% scaling.
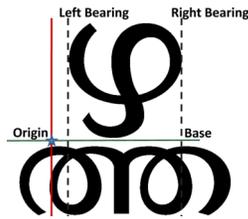


(b) BChar with 45%–50% scaling improves interchniaracter spacing.



**Figure 19**: Design adjustments for an extreme case of BChar.

(a) Collision.



(b) Larger interchniaracter space.



(c) Unconventional reduction.
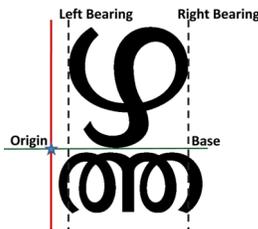


is a clear instance of typographic deviation between AChar and BChar.

This kind of typography with two kinds of types in the same glyph is almost unheard of in Roman types. Rachana takes the liberty to deviate from these accepted norms.

### 3.3.1  Extreme cases

Some vertical conjuncts are shaped deviating more from the usual 70%–60% proportion, depicted in Fig. 17.

If 70%–60% scaling is applied, the BChar extends far beyond the left bearing and right bearing, resulting in collisions with neighbour characters (Fig. 18a). If kerning is adjusted to avoid this, white space to the left and right of AChar produces a 'space effect' (Fig. 19b).

The only solution for these extreme cases is to apply a different proportion to BChar. In Rachana, 45%–50% scaling instead of the usual 70%–60% is applied in designing these types of vertical conjuncts to preserve normal character spacing (see Fig. 18b). This kind of elasticity applied in Malayalam breaks all established rules of typography. It perhaps does not occur even in typography of other Indic scripts.

All these considerations show that Malayalam fonts cannot be designed according to the metric calculations of Roman typography. This is more or less the case with all Indic scripts, due to the abundance of conjuncts. Malayalam poses extreme cases of divergence while dealing with vertical conjuncts. Different proportions applied to different conjunct formations in the same font completely contravenes

principles formulated for Roman text types; however, these complexities are perhaps comparable to Roman typography employed in typesetting mathematics (see Fig. 20). There too, different types in a single set are not tolerated.

### 3.3.2  A page typeset in Rachana

The Rachana font is reimagined and redesigned using many levels of descent, and yet this doesn't cause serious issues with leading. The overall aesthetics and readability are in fact improved. A sample document typeset using X∃TEX is shown in Fig. 21.

## 4  Conclusions

Vertical conjuncts in Malayalam are unique compared to Roman scripts and other Indic scripts. Designing vertical conjuncts results in many deviations from accepted norms of Roman typography. Even

**Figure 20**: Typesetting mathematics using the Monotype 4-line system. *Source: Daniel Rhatigan.*

**Figure 21**: Malayalam text typeset with Rachana.

മലയാളികൾ എഴുത്തുംപോലെ ഉച്ചരിക്കുന്നവരാണെന്നും എല്ലാ ഉച്ചാരണത്തിനും നമുക്ക് എഴുത്തുണ്ട് എന്നും തീരെ തെറ്റായ ഒരു ധാരണ പലരും പറഞ്ഞു പരത്തുന്നുണ്ട്. ഇത്തരം ഒരേർപ്പാട് ഒരു ഭാഷയ്ക്കും ഇല്ല, ആവശ്യവും ഇല്ല. എഴുതിക്കണ്ടാൽ ഉച്ചാരണം തിരിച്ചറിയണം ഉച്ചരിക്കുന്നതു കേട്ടാൽ എഴുതാനും ആകണം—അത്രമതി. അതിനെന്തു വേണം? പ്രസക്തമായ അർഥവ്യാവർത്തനങ്ങൾ എഴുതിക്കാൻ ഒത്ത ലിപി വേണം. പ്രസക്തമായ അർഥവ്യാവർത്തനങ്ങൾ എഴുതിക്കാട്ടാൻ ലിപി ഇല്ലാത്തതാണ് ഇംഗ്ലീഷിലെ എഴുത്തുരീതിയുടെ ഒരു കുഴപ്പം. ഒറ്റസ്വരങ്ങളും ഇരട്ടസ്വരങ്ങളും ചേർന്ന് ഇംഗ്ലീഷിൽ 20 എണ്ണമുള്ളതായി ഓക്സ്ഫഡ് ലേണേഴ്സ് ഡിക്ഷനറി പറയുന്നു. ലിപികളോ a, e, i, o, u എന്ന് അഞ്ചെണ്ണം മാത്രം. ആവശ്യത്തിന് ലിപിയില്ലാത്തതുപോലെ വേണ്ടതിലേറെ ലിപികൾ ഉണ്ടുതാനും. Q എപ്പോഴും u ചേർന്നേ വരൂ. Qu എന്നതിന് ഉച്ചാരണം എപ്പോഴും kw എന്ന മാത്രം. X എന്നതിന് ks എന്നതിന്റെ ഉച്ചാരണമാണ്. മലയാളത്തിലും ഇത്തരം കുഴപ്പങ്ങൾ കുറഞ്ഞ അളവിൽ ഇല്ലായ്മയില്ല. കന്നിയിലെ ന്ന അല്ല കന്നിയിലേതിന്. ഒരേ ലിപി, രണ്ടു വ്യത്യസ്തമായ ഉച്ചാരണവും. ലിപിക്കാരവിന്റെ കാര്യം ഇങ്ങനെ. ലിപിക്കൂട്ടലിന്റെ കാര്യമോ?

**Figure 22**: Sample glyphs of Rachana.



within a font, difficult situations arise when following a single type or rule.

'Leading' in Malayalam typesetting is a serious concern, especially when using fonts like Rachana based on traditional orthography. Leading should be kept as small as possible, and as a result it exerts pressure on Malayalam font designers due to abundant height variations in BChars. Holding a fixed proportion in BChar hurts leading. The 'negotiations and adjustments' in proportion are often applied in BChars but this results in different calculations for ascent–descent and point size estimates. That is why Malayalam font makers are often forced to explore different metrics. Other Indic scripts with moderate below characters/diacritical marks are luckier in this aspect. Deviating from Roman dimensions poses problems of point sizes when typesetting documents using Malayalam and Latin fonts together.

Rachana in earlier versions attempted to group and equalize many below-base levels but later promoted natural proportions in shapes rather than 'forced' proportions. For all intents and purposes Malayalam typography deviates from geometrical consistency set for Roman typography. Proper balance in glyph composition in Malayalam can only be achieved by embracing a different mindset, going beyond Roman typography.

## A Glyphs in the Rachana traditional orthography font

A sample of the glyph set designed for the Rachana font is shown in Fig. 22. The complete list, which contains the comprehensive character set for traditional orthography is available at `http://rachana.org.in/docs/Rachana-conjuncts.pdf`.

## References

[1] Unicode Consortium. Malayalam Unicode chart, Unicode standard 13.0 [online]. `https://www.unicode.org/charts/PDF/U0D00.pdf`

⬦ Hussain KH
  Rachana Institute of Typography
  `hussain (at) rachana dot org dot in`
  `http://rachana.org.in/`

⬦ Rajeesh KV
  Rachana Institute of Typography
  `rajeesh (at) rachana dot org dot in`
  `http://rachana.org.in/`

⬦ Aravind Rajendran
  STM Document Engineering
  `aravind (at) stmdocs dot in`
  `https://stmdocs.in/`