

# Finalizing the Grantha virama model

Shriramana Sharma, jamadagni-at-gmail-dot-com, India

2010-Oct-10

At this stage of encoding Grantha, the only major issue left unresolved is the Grantha virama model. Normally the use and implementation of the virama is straightforward in Indic scripts. However, in Grantha this issue has been complicated by the existence of various forms of writing vowelless consonants and the claim by some parties that such variation in writing implies an underlying semantic difference. This document intends to clarify the ground reality and ensure the primary goal of the Grantha Unicode encoding, viz proper representation of the orthographic nature and semantics of the script as used to write its native language. For those who suggest that the Unicode encoding of this script should be able to cater to other non-native languages, this document suggests viable alternatives which would not adversely affect the primary goal of encoding this script.

## §1. Finalizing the language represented by Grantha

To finalize the Grantha virama model, one must decide clearly as to what language is represented by Grantha. Grantha is used to represent Sanskrit. Sanskrit is script-agnostic, and as such the only function of the virama in any script used to write Sanskrit is to render consonants vowelless, possibly creating ligatures and/or conjoining forms.

Some opine that Grantha is also used to represent Malayalam. In Malayalam written in the modern Malayalam script there are alternative methods of writing vowelless consonants with potential semantic distinction involved. Therefore, these people argue that such distinction is (to be) preserved in Malayalam written in Grantha as well.

If only Sanskrit is represented by Grantha, no underlying encoding distinction between various potential forms of writing vowelless consonants is warranted. If Malayalam is also represented, then since encoding distinction between various vowelless consonant forms is made for Malayalam in Unicode, it may also be appropriate for Grantha.

The finalizing arguments in this matter are provided below.

### *§1.1. Modern Grantha for Sanskrit and Sanskrit alone*

The Grantha script as has been recognized for encoding in Unicode is used for the writing of Sanskrit and Sanskrit alone. The vast corpus of manuscripts and printed material in the

script relates to Sanskrit. The only real-world, current and widespread usage of the script is for the representation of Sanskrit – whether that is by manuscriptologists working on preserving, editing and re-publishing works present in manuscripts, or by the Vedic and Sanskrit scholars of Tamil Nadu, Sri Lanka and elsewhere who use this script in their daily life for religious and some non-religious purposes.

Even by those who claim the use of this script for languages other than Sanskrit, specifically for Malayalam, no attestations have been provided in the long more-than-a-year period in which the discussions of the Grantha Unicode encoding have been taking place. If today any “new” handwritten samples of the same are provided, the question as to why no such samples were provided hitherto despite the elapse a more-than-sufficient period of opportunity for doing so will remain. The question will also stand that if the representation of the Malayalam language in the Grantha script is really accepted by the native user community why not even one publication of the same is given as attestation.

Thus Sanskrit is the only language that the Grantha Unicode encoding should primarily be able to represent properly as it is the only language that Modern Grantha, which the present encoding is based on, is attested to represent. *Without conflict with this*, the encoding may provide for the transcription of non-Sanskrit sounds as well.

### *§1.2. Archaic Malayalam script, called Malabar-Grantha, used for Malayalam*

The fact that there was an archaic writing system called “Malabar Grantha” or “Malayalam Grantha” leads to some parties to claim that “Grantha is used for writing Malayalam” (my italics to emphasize the implication of current usage by the present tense). The absence of attestation for the same was however noted above.

How this “Malayalam Grantha” script was born and became the modern Malayalam script has been lucidly and concisely described in Wikipedia with proper academic references, and therefore I am content to quote from the relevant passages: (from ref 1):

*... one form of the Grantha script, originally used in the Chola kingdom, was imported into the southwest coast of India in the 8<sup>th</sup> or 9<sup>th</sup> century, which was then modified in course of time in this secluded area, where communication with the east coast was very limited. This script was ... originally only applied to write Sanskrit. ... In Malabar, this writing system was termed Arya-eluttu ... meaning “Arya writing” (Sanskrit is Aryan while Malayalam is Dravidian) ... Vatteluttu was in general use, but was not suitable for literature where many Sanskrit words were used. Like Tamil-Brahmi, it was originally used to*

write Tamil, and as such, did not have letters for ... consonants used in Sanskrit but not used in Tamil. For this reason, Vatteluttu and the Grantha script were sometimes mixed, as in Manipravalam. ... Thunchath Ezhuthachan, a poet from around the 17<sup>th</sup> century, used Arya-eluttu to write his Malayalam poems based on Classical Sanskrit literature. For a few letters missing in Arya-eluttu (la, la, ra), he used Vatteluttu. His works ... popularized Arya-eluttu as a script to write Malayalam.

Thus the “Malayalam Grantha” script was an ancestral form of the modern Malayalam script dating back to the 17<sup>th</sup>/18<sup>th</sup> century. The 1772 CE publication *Alphabetum Grandonico-Malabaricum Sive Samscrudonicum* (ref 2) documents this script under this name “Malayalam Grantha” (“Grandonico-Malabaricum” in Latin) for the writing of Malayalam.

However, by the 19<sup>th</sup> century this script started to be recognized as merely the “Malayalam” script, as seen in the 1837 CE publication *Ancient and Modern Alphabets of the Popular Hindu Alphabets of the Southern Peninsula of India* (ref 3). This same reference labels the version of the Grantha script of the time as merely “Grantha” without any qualifying adjectives. It is clear that this author who has done painstaking work in documenting various written forms of the South Indian scripts of his time opines that “Grantha” (without adjectives) used for Sanskrit is distinct from “Malayalam” used for Malayalam:

Grantha	ക	കൃ	ഉ	ഊ	ഉ	ഉ	ഈ	ഈ	ഈ	ഈ	ഈ			ര	ര	ര	ര
Malayalam	ന	ന	ന	ന	ന	ന	ന	ന	ന	ന	ന	ന	ന	ന	ന	ന	ന
Grantha	ട	ട	ട	ട	ട	ട	ട	ട	ട	ട	ട	ട	ട	ട	ട	ട	ട
Malayalam	ട	ട	ട	ട	ട	ട	ട	ട	ട	ട	ട	ട	ട	ട	ട	ട	ട

That even at the time of this compilation (early 19<sup>th</sup> century) Grantha was marked as used only for Sanskrit (and not for Malayalam) is evident from the absence of the short vowels E/O (seen above) and that of the Dravidian consonants (later on in the book) under “Grantha”. The title of the 18<sup>th</sup> century book describing Malayalam Grantha also identifies (“sive”) “Grandonico-Malabaricum” with “Samscrudonicum” which also tallies with the Wikipedia description above of “Arya eluttu” or “Sanskrit writing” being *borrowed* to write Malayalam. This only further strengthens the association of Grantha with Sanskrit and the fact that such a Sanskrit writing system *separated off* into a Malayalam writing system.

Now this Malayalam script of the 19<sup>th</sup> century underwent further changes such as:

- 1) changing its representation of the short and long vowels E/EE and O/OO,
- 2) gaining distinct written forms for TTTA and NNNA in limited scholarly use,

- 3) script reform regarding ligatures as those with the vowel signs U/UU,
- 4) losing its dot-reph,
- 5) and finally and most importantly (for now) degeneration of the writing of the samvruthokaram using vowel sign U with chandrakala to using only the chandrakala, hence resulting in the birth of semantic distinction between the two different ways of representing some vowelless consonants in the script

... to become the modern Malayalam script. It is clear that this script has come a long way since the days it was labeled “Malabar Grantha”. This modern Malayalam script, far different from any archaic or modern version of the Grantha script used for writing Sanskrit, is the only script that is used for the popular writing of Malayalam today. Thus the Grantha script – as in the 9<sup>th</sup> century when it was imported into Kerala, as recognized in the early 19<sup>th</sup> century by ref 3 as noted above, and as recognized by all learned Vedic/Sanskrit scholars today in the 21<sup>st</sup> century – which is in use for the writing of Sanskrit, is *not* used for writing Malayalam.

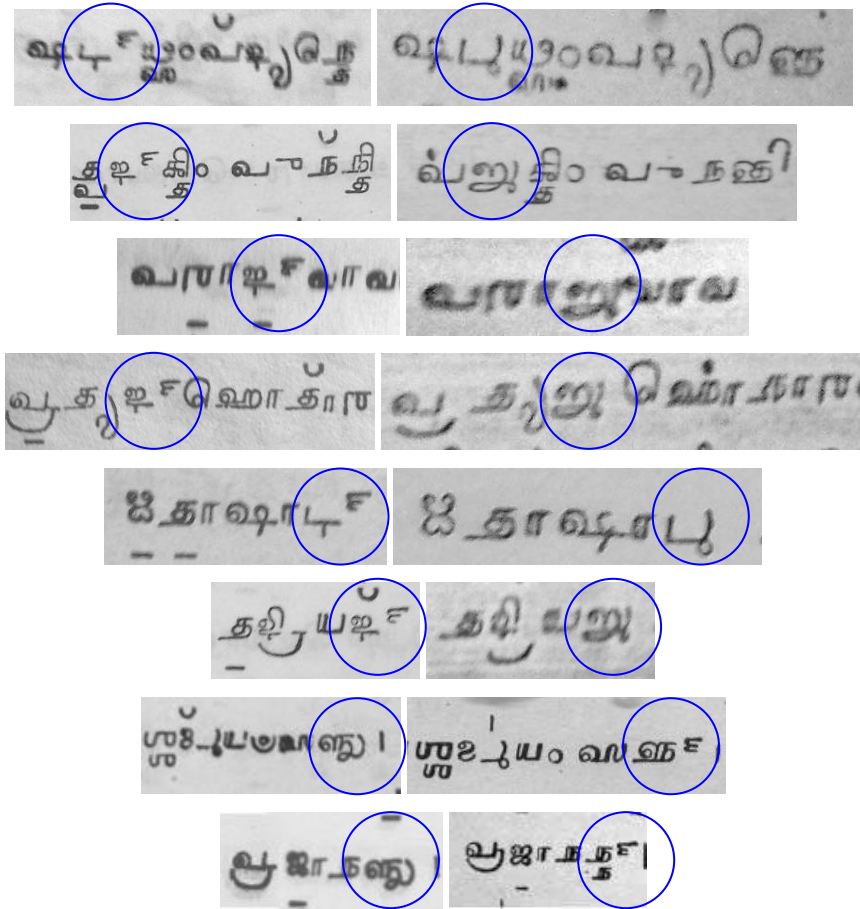
The “Malabar Grantha” script of the 17<sup>th</sup>/18<sup>th</sup> centuries should be analysed as an archaic variant of the modern Malayalam script, since in both cases the language represented is Malayalam, rather than as a variant of the modern Grantha script, which is used for writing Sanskrit and Sanskrit only as documented above and elsewhere. Even if one insists on analysing Malabar Grantha as a form of Grantha, it is evident that it (Malabar Grantha) did not have any semantic distinction between various vowelless consonant forms since the degeneration of the samvruthokaram is a very recent phenomenon of the past 50 years or less. Thus at *no* point in time when *any* form of Grantha was allegedly used to write Malayalam did any semantic distinction between various vowelless forms exist.

## §2. Finalizing the semantics of Grantha virama forms

### §2.1. *Mutual semantic equivalence*

As should be clear from the foregoing discussion, the Grantha script as recognized today for encoding is used to represent Sanskrit only. As Sanskrit is a script-agnostic language as has been variously documented, only the phonetic content of any written sequence of Sanskrit will provide the meaning and not the visible forms. Thus any and all different forms of writing vowelless consonants in any script being used to write Sanskrit – for the present discussion, in Grantha – will be totally semantically identical with each other and with vowelless consonant forms of any other script used to write Sanskrit.

The above fact has been repeatedly stated in my various documents including my Grantha proposal L2/09-372, on pp 22 and 25 of which I have clearly illustrated the mutual semantic equivalence of the various Grantha virama forms with proper attestations:



The references for these samples and the respective meanings that are preserved despite variation in written forms were provided in my proposal. These samples clearly illustrate that the various virama forms are used in free alteration as per typographic style.

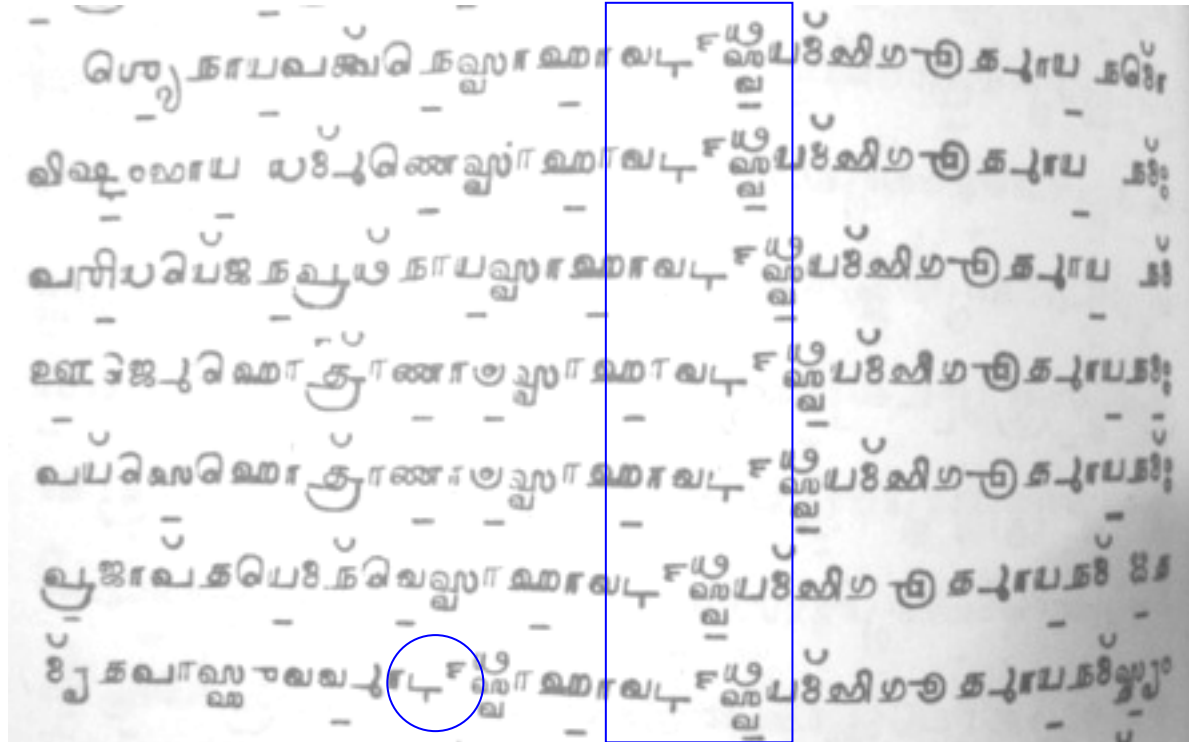
To further emphasize and clarify this equivalence, I now provide comparisons of (Vedic) Sanskrit texts printed in three different scripts:

- 1) Grantha, the script under discussion,
- 2) Devanagari, the script most widely used for Sanskrit as a base reference,
- 3) Kannada, as another South Indian script.

These samples will illustrate that the virama forms of Grantha are all full equivalents of and hence semantically identical to the (single) “overt” virama forms of the other Indic scripts. (The virama forms belonging to the different scripts in the samples have been marked for easy identification and comparison.) This will hence transitively prove that all the various virama forms of Grantha are all mutually semantically equal, since their equivalents in other scripts are unified in both written form and semantics. (A = B & B = C implies A = C.)

§2.2. Spacing virama forms in Grantha and equivalents in Devanagari and Kannada

Grantha



Ref 7 p 306

Devanagari

श्येनाय॒ पत्वंने॒ स्वाहा॒ वट्थस्वयम॑भिगूर्ताय॒ नमो॑ विष्टम्भाय॒ धर्मणे॒ स्वाहा॒ वट्थस्वयम॑भि-  
गूर्ताय॒ नमः॑ परि॒धये॑ जन॒प्रथ॑नाय॒ स्वाहा॒ वट्थस्वयम॑भिगूर्ताय॒ नम॑ ऊ॒र्जे हो॒त्राणा॑ꣳ स्वाहा॒ वट्थस्व-  
यम॑भिगूर्ताय॒ नमः॑ पर्य॒से हो॒त्राणा॑ꣳ स्वाहा॒ वट्थस्वयम॑भिगूर्ताय॒ नमः॑ प्र॒जाप॑तये॒ मन॑वे॒ स्वाहा॒ वट्-  
थस्वयम॑भिगूर्ताय॒ नम॑ ऋ॒तमृ॑तपास्सुवर्वाट्थस्वाहा॒ वट्थस्वयम॑भिगूर्ताय॒ नम॑स्तृ॒पन्ता॑ꣳ हो॒त्रा म॑धो-

Ref 3 p 108

Kannada

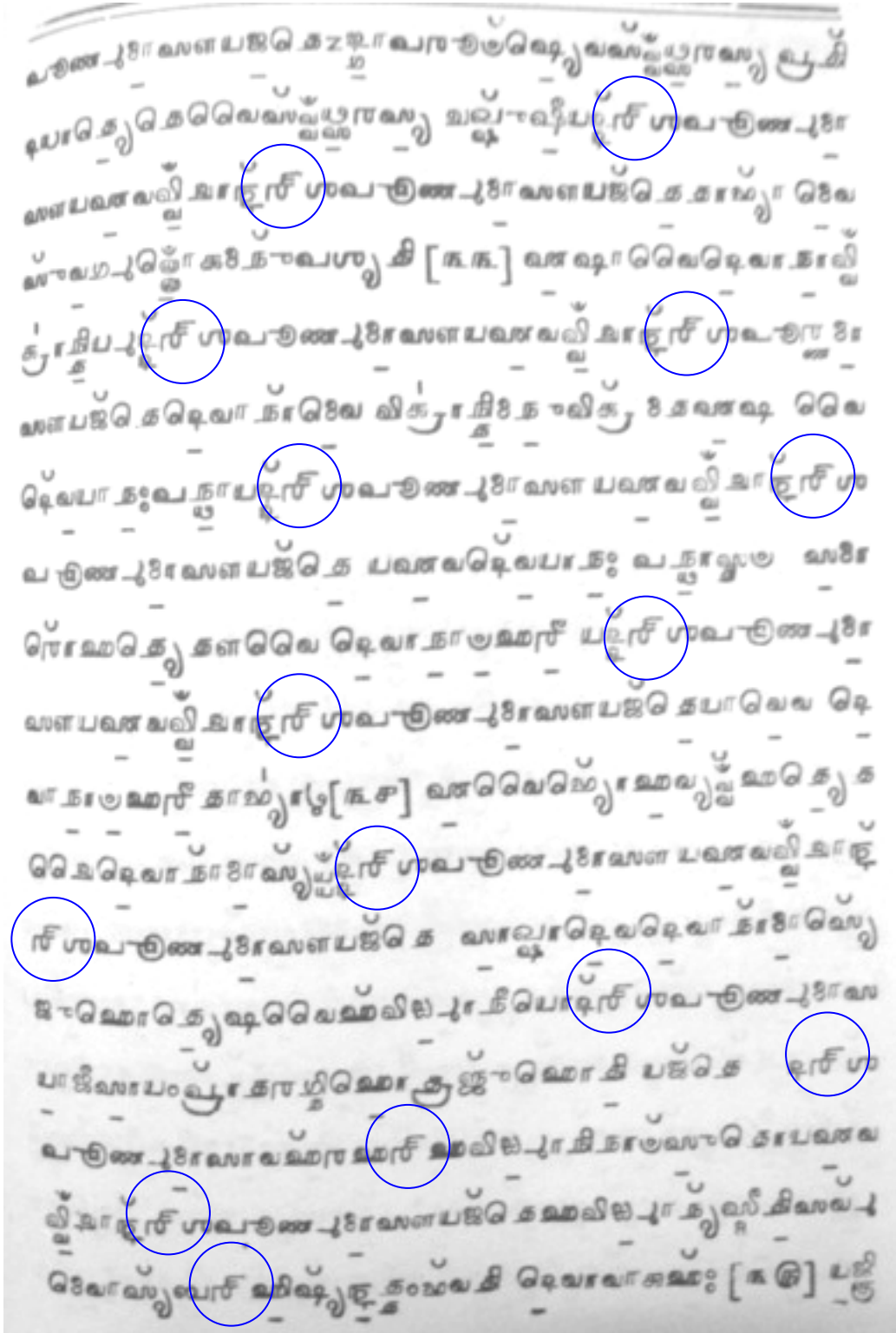
ಶ್ರೀನಾಯ ಪತ್ತನೇ ಸ್ವಾಹಾ ವಟ್‌ಧ್ವಯಮಭಿಗೂರ್ತಾಯ  
ನಮೋ ವಿಷ್ಣುಭಾಯ ಧರ್ಮಣೇ ಸ್ವಾಹಾ ವಟ್‌ಧ್ವಯಮಭಿ-  
ಗೂರ್ತಾಯ ನಮಃ ಪರಿಧಯೇ ಜನಪ್ರಥನಾಯ ಸ್ವಾಹಾ ವಟ್-  
ಧ್ವಯಮಭಿಗೂರ್ತಾಯ ನಮ ಊರ್ಜೇ ಹೋತ್ರಾಣಾಗ್  
ಸ್ವಾಹಾ ವಟ್‌ಧ್ವಯಮಭಿಗೂರ್ತಾಯ ನಮಃ ಪಯಸೇ  
ಹೋತ್ರಾಣಾಗ್ ಸ್ವಾಹಾ ವಟ್‌ಧ್ವಯಮಭಿಗೂರ್ತಾಯ ನಮಃ  
ಪ್ರಜಾಪತಯೇ ಮನವೇ ಸ್ವಾಹಾ ವಟ್‌ಧ್ವಯಮಭಿಗೂರ್ತಾಯ  
ನಮ ಋತಮೃತಪಾಸ್ಸುವರ್ವಾಟ್‌ಧ್ವಯಮಭಿ-

Ref 5 p 377



§2.3. Touching virama forms in Grantha and equivalents in Devanagari and Kannada

Grantha



र्णमासौ यजतेऽङ्गापरुष्येव संवत्सरस्य प्रति दधात्येते वै संवत्सरस्य चखुषी यद्दर्शपूर्णमासौ  
 य एवँविद्वान्दर्शपूर्णमासौ यजते ताभ्यामेव सुवर्गल्लोकमनु पश्य- ॥ ३३ ॥ तेषा वै देवाना-  
 वँविक्रान्तिर्यद्दर्शपूर्णमासौ य एवँविद्वान्दर्शपूर्णमासौ यजते देवानामेव विक्रान्तिमनु वि क्रमत  
 एष वै देवयानः पन्था यद्दर्शपूर्णमासौ य एवँविद्वान्दर्शपूर्णमासौ यजते य एव देवयानः पन्था-  
 स्तः समारोहत्येतौ वै देवानां हरी यद्दर्शपूर्णमासौ य एवँविद्वान्दर्शपूर्णमासौ यजते यावेव  
 देवानां हरी ताभ्या- ॥ ३४ ॥ मेवैभ्यो हव्यँवहत्येतद्वै देवानामास्यँय्यद्दर्शपूर्णमासौ य एवँवि-  
 द्वान्दर्शपूर्णमासौ यजते साख्यादेव देवानामास्यँ जुहोत्येष वै हविर्द्वानी यो दर्शपूर्णमासयाजी  
 सायंप्रातरग्निहोत्रञ्जुहोति यजते दर्शपूर्णमासावहर्हविर्द्वानिनां सुतो य एवँविद्वान्दर्शपूर्ण-  
 मासौ यजते हविर्द्वान्यस्मीति सर्वमेवाऽस्य बर्हिष्यन्दत्तं भवति देवा वा अहं- ॥ ३५ ॥ यज्ञियन्ना-

ಏವಾಪಸಾನೇ ವರೇ ದೇವಾನಾಮವ ಸ್ಯತ್ಕೇತಾನಿ ವಾ ಅಂಗಾ-  
 ಪರೂಗ್ಂಷಿ ಸಂವಧ್ಧರಸ್ಯ ಯದ್ಧರ್ಶಪೂರ್ಣಮಾಸೌ ಯ ಏವಂ  
 ವಿದ್ವಾನ್ ದರ್ಶಪೂರ್ಣಮಾಸೌ ಯಜತೇಽಂಗಾಪರೂಗ್ಂಷ್ಯೇವ  
 ಸಂವಧ್ಧರಸ್ಯ ಪ್ರತಿ ದಧಾತ್ಯೇತೇ ವೈ ಸಂವಧ್ಧರಸ್ಯ ಚಮ್ಪಿಷೀ  
 ಯದ್ಧರ್ಶಪೂರ್ಣಮಾಸೌ ಯ ಏವಂ ವಿದ್ವಾನ್ ದರ್ಶಪೂರ್ಣ-  
 ಮಾಸೌ ಯಜತೇ ತಾಭ್ಯಾಮೇವ ಸುವರ್ಗಂ ಲೋಕಮನು ಪಶ್ಯತಿ  
 (೨.೨) ಏಷಾ ವೈ ದೇವಾನಾಂ ವಿಕ್ರಾಂತಿಯದ್ಧರ್ಶಪೂರ್ಣಮಾಸೌ  
 ಯ ಏವಂ ವಿದ್ವಾನ್ ದರ್ಶಪೂರ್ಣಮಾಸೌ ಯಜತೇ ದೇವಾನಾ-  
 ಮೇವ ವಿಕ್ರಾಂತಿಮನು ವಿ ಕ್ರಮತ ಏಷ ವೈ ದೇವಯಾನಃ ಪಂಥಾ  
 ಯದ್ಧರ್ಶಪೂರ್ಣಮಾಸೌ ಯ ಏವಂ ವಿದ್ವಾನ್ ದರ್ಶಪೂರ್ಣ-  
 ಮಾಸೌ ಯಜತೇ ಯ ಏವ ದೇವಯಾನಃ ಪಂಥಾಸ್ತಗ್ಂ ಸಮಾ-  
 ರೋಹತ್ಯೇತೌ ವೈ ದೇವಾನಾಗ್ಂ ಹರೀ ಯದ್ಧರ್ಶಪೂರ್ಣಮಾಸೌ  
 ಯ ಏವಂ ವಿದ್ವಾನ್ ದರ್ಶಪೂರ್ಣಮಾಸೌ ಯಜತೇ ಯಾವೇವ  
 ದೇವಾನಾಗ್ಂ ಹರೀ ತಾಭ್ಯಾಂ (೨.೪) ಏವೈಭ್ಯೋ ಹವ್ಯಂ  
 ವಹತ್ಯೇತದ್ವೈ ದೇವಾನಾಮಾಸ್ಯಂ ಯದ್ಧರ್ಶಪೂರ್ಣಮಾಸೌ ಯ  
 ಏವಂ ವಿದ್ವಾನ್ ದರ್ಶಪೂರ್ಣಮಾಸೌ ಯಜತೇ ಸಾಖ್ನಾದೇವ

ದೇವಾನಾಮಾಸ್ಯೇ ಒುಹೋತ್ಯೇಷ ವೈ ಹವಿರ್ಧಾನೀ ಯೋ  
 ದರ್ಶಪೂರ್ಣಮಾಸಯಾಜೇ ಸಾಯಂಪ್ರಾತರಗ್ನಿಹೋತ್ರಂ  
 ಒುಹೋತಿ ಯಜತೇ ದರ್ಶಪೂರ್ಣಮಾಸಾವಹುರಹವಿ-  
 ಧಾನಿನಾಗ್ಂ ಸುತೋ ಯ ಏವಂ ವಿದ್ವಾನ್ ದರ್ಶಪೂರ್ಣ-  
 ಮಾಸೌ ಯಜತೇ ಹವಿರ್ಧಾನೈಸ್ತೀತಿ ಸರ್ವಮೇವಾಸ್ಯ ಒರಹಿಷ್ಯಂ  
 ದತ್ತಂ ಭವತಿ ದೇವಾ ವಾ ಅಹಃ (೨.೫) ಯಜ್ಞಿಯಂ ನಾವಿಂದನ್



தெதிரீயயஜுஸ்யூகூகூணெ (க-பு க)

க்ஹத்யு | நெராத்ரலஸஸ்தகைநெஜீஹத்யு | சிநஸஸ்தகை  
 நெஜீஹத்யு | வாகலஸஸ்தகைநெஜீஹத்யு | சூய்ஸூ  
 ய்நெய்யத்யு | சூய்யுஜூய்யத்யு | சூய்யுஜூவ்த  
 யெய்யத்யு | பூணவ்யு | பூணெய்யத்யு | பூணுஜூய்  
 யத்யு | [க] பூணுஜூவ்தயெய்யத்யு | உவ்யுஸஸவ்யு  
 நெய்யத்யு | உவ்யுஜூய்யத்யு | உவ்யுஜூவ்தயெய்யத்யு  
 யத்யு | நெராத்ரஸூநெராத்ரெய்யத்யு | நெராத்ரூஜூய்  
 யத்யு | நெராத்ரூஜூவ்தயெய்யத்யு | தளநெவளஸூகூ  
 சிநிநள | கலய்தநெநீவீயுஸ் | கலய்த ஜூநூஷீ |  
 [ச] உஷிநெஜூஸூஸூய்யத்யு | பூணாஸூஸூய்யு |  
 பூஜூஸூய்ய யஜூநெய | திரஸூஸூணம் | திரஸூஸூகூ |  
 கவநூதளஸூஸூகூள ஸஹாஸூநா | ஸூகூஸூ ஸூஸூ  
 ஸி | சிநிஸூஸூஸூஸூ | ஸூஸூஸூஸூஸூ விஸூஸூஸூ | ஸூ  
 ஸூஸூஸூஸூஸூ ஸூஸூஸூஸூஸூ | ஸூஸூஸூஸூஸூஸூஸூ  
 திரஸூஸூஸூஸூ | தஸூ ஸூஸூஸூஸூஸூஸூஸூஸூஸூஸூ || [ஊ]



ब्रह्म सन्धत्तन्तन्मै जिन्वतम् । स्यत्रऽ सन्धत्तन्तन्मै जिन्वतम् । इषऽ सन्धत्तन्ताम्मै  
 जिन्वतम् । ऊर्जऽ सन्धत्तन्ताम्मै जिन्वतम् । रुषिऽ सन्धत्तन्ताम्मै जिन्वतम् । पुष्टिऽ सन्धत्तन्ता-  
 म्मै जिन्वतम् । प्रजाऽ सन्धत्तन्ताम्मै जिन्वतम् । पशून्धत्तन्तान्मै जिन्वतम् ॥ स्तुतोऽसि  
 जनधाः । देवास्त्वा शुक्रपाः प्रणयन्तु ॥ [१] सुनीलः प्रजाः प्रजनयन्परीहि । शुक्रश्शुक्रशोचिषा ॥  
 स्तुतोऽसि जनधाः । देवास्त्वा मन्धिषाः प्रणयन्तु ॥ सुप्रजाः प्रजाः प्रजनयन्परीहि । मन्धी मन्धि-  
 शोचिषा ॥ सञ्जग्मानौ दिव आ पृथिव्याऽऽयुः । सन्धत्तन्तन्मै जिन्वतम् । प्राणऽ सन्धत्तन्तम्मै  
 जिन्वतम् । अपानऽ सन्धत्तन्तम्मै जिन्वतम् ॥ [२] व्यानऽ सन्धत्तन्तम्मै जिन्वतम् । चक्षुस्स-  
 न्धत्तन्तन्मै जिन्वतम् । श्रोत्रऽ सन्धत्तन्तन्मै जिन्वतम् । मनस्सन्धत्तन्तन्मै जिन्वतम् । वाचऽ  
 सन्धत्तन्ताम्मै जिन्वतम् ॥ आयुस्स्थ आयुर्मै धत्तम् । आयुर्यज्ञाय धत्तम् । आयुर्यज्ञपतये धत्तम् ।  
 प्राणस्थः प्राणम्मै धत्तम् । प्राण्यैज्ञाय धत्तम् ॥ [३] प्राण्यैज्ञपतये धत्तम् । चक्षुस्स्थश्चक्षुर्मै  
 धत्तम् । चक्षुर्यज्ञाय धत्तम् । चक्षुर्यज्ञपतये धत्तम् । श्रोत्रं स्थश्श्रोत्रम्मै धत्तम् । श्रोत्र्यैज्ञाय  
 धत्तम् । श्रोत्र्यैज्ञपतये धत्तम् । तौ देवौ शुक्रामन्धिनौ । कल्पयन्तन्दैवीर्विशः । कल्पयन्तम्मा-  
 नुषीः ॥ [४] इषमूर्जमस्मासु धत्तम् । प्राणान्पशुर्षु । प्रजाम्मरिं च यजमाने च ॥ निरस्तश्शण्डः ॥  
 निरस्तो मर्कः ॥ अपनुत्तौ शण्डामर्कौ सहामुना ॥ शुक्रस्य समिदसि ॥ मन्धिनस्समिदसि ॥ स  
 प्रथमस्सद्भृतिर्विश्वकर्मा । स प्रथमो मित्रो बरुणो अग्निः । स प्रथमो बृहस्पतिश्चिवित्वान् । तस्मा  
 इन्द्राय सुतमाजुहोभि ॥ [५] ॥ नयन्त्वपानऽ सन्धत्तन्तम्मै जिन्वतम्प्राण्यैज्ञाय धत्तम्मानुषीरग्निद्वे च ॥ १॥

Kannada

ಹಂಃ ಓಮ್ || ಬ್ರಹ್ಮ ಸಂಧತ್ತಂ ತನ್ನೆ ಜಿನ್ನತಮ್ | ಬ್ರಿತ್ತಗ್ಂ  
 ಸಂಧತ್ತಂ ತನ್ನೆ ಜಿನ್ನತಮ್ | ಇಷಗ್ಂ ಸಂಧತ್ತಂ ತಾಂ ಮೇ  
 ಜಿನ್ನತಮ್ | ಉರ್ಜಗ್ಂ ಸಂಧತ್ತಂ ತಾಂ ಮೇ ಜಿನ್ನತಮ್ |  
 ರಯಗ್ಂ ಸಂಧತ್ತಂ ತಾಂ ಮೇ ಜಿನ್ನತಮ್ | ವೃಷ್ಟಿಗ್ಂ ಸಂಧತ್ತಂ  
 ತಾಂ ಮೇ ಜಿನ್ನತಮ್ | ಪ್ರಜಾಗ್ಂ ಸಂಧತ್ತಂ ತಾಂ ಮೇ ಜಿನ್ನತಮ್ |  
 ಪಶೂನ್ಯಂಧತ್ತಂ ತನ್ನೆ ಜಿನ್ನತಮ್ | ಸ್ತುತೋಸಿ ಜನಧಾ |  
 ದೇವಾಸ್ತಾ ಶುಕ್ರವಾಃ ಪ್ರಣಯಂತು || ೧ || ಸುವೀರಾಃ ಪ್ರಜಾಃ  
 ಪ್ರಜನಯನ್ವರೀಹಿ | ಶುಕ್ರಶುಕ್ರಶೋಚಿವಾ | ಸ್ತುತೋಸಿ ಜನಧಾ |  
 ದೇವಾಸ್ತಾ ಮನ್ರಿವಾಃ ಪ್ರಣಯಂತು | ಸುಪ್ರಜಾಃ ಪ್ರಜಾಃ  
 ಪ್ರಜನಯನ್ವರೀಹಿ | ಮನ್ರಿ ಮನ್ರಿಶೋಚಿವಾ | ಸಂಜಗ್ನಾಸೌ ದಿವ ಆ  
 ಪೃಥಿವ್ಯಾಯುಃ | ಸಂಧತ್ತಂ ತನ್ನೆ ಜಿನ್ನತಮ್ | ಪ್ರಾಣಗ್ಂ ಸಂಧತ್ತಂ  
 ತಂ ಮೇ ಜಿನ್ನತಮ್ | ಆವಾನ್ಗ್ಂ ಸಂಧತ್ತಂ ತಂ ಮೇ ಜಿನ್ನತಮ್ |  
 || ೨ || ವ್ಯಾನ್ಗ್ಂ ಸಂಧತ್ತಂ ತಂ ಮೇ ಜಿನ್ನತಮ್ | ಚಮ್ಪಿಸ್ಸಂಧತ್ತಂ

ತನ್ನೆ ಜಿನ್ನತಮ್ | ಶೋತ್ರಗ್ಂ ಸಂಧತ್ತಂ ತನ್ನೆ ಜಿನ್ನತಮ್ |  
 ಮನಸ್ಸಂಧತ್ತಂ ತನ್ನೆ ಜಿನ್ನತಮ್ | ವಾಚಗ್ಂ ಸಂಧತ್ತಂ ತಾಂ ಮೇ  
 ಜಿನ್ನತಮ್ | ಅಯುಸ್ತ ಅಯುರ್ಮೇ ಧತ್ತಮ್ | ಅಯುರ್ಯಜ್ಞಾಯ  
 ಧತ್ತಮ್ | ಅಯುರ್ಯಜ್ಞಪತಯೇ ಧತ್ತಮ್ | ಪ್ರಾಣಸ್ತು ಪ್ರಾಣಂ  
 ಮೇ ಧತ್ತಮ್ | ಪ್ರಾಣಂ ಯಜ್ಞಾಯ ಧತ್ತಮ್ | ೩ || ಪ್ರಾಣಂ  
 ಯಜ್ಞಪತಯೇ ಧತ್ತಮ್ | ಚಮ್ಪಿಸ್ತು ಚಮ್ಪಿರ್ಮೇ ಧತ್ತಮ್ | ಚಮ್ಪಿ-  
 ಯಜ್ಞಾಯ ಧತ್ತಮ್ | ಚಮ್ಪಿಯಜ್ಞಪತಯೇ ಧತ್ತಮ್ |  
 ಶೋತ್ರಗ್ಂ ಸ್ತು ಶೋತ್ರಂ ಮೇ ಧತ್ತಮ್ | ಶೋತ್ರಂ ಯಜ್ಞಾಯ  
 ಧತ್ತಮ್ | ಶೋತ್ರಂ ಯಜ್ಞಪತಯೇ ಧತ್ತಮ್ | ತೌ ದೇವೌ ಶುಕ್ರಾ-  
 ಮನ್ರಿನೌ | ಕಲ್ಪಯಂತಂ ದೈವೀರ್ವಿಶಃ | ಕಲ್ಪಯಂತಂ ಮಾನುಷೀಃ || ೪ ||  
 ಇಪಮೂರ್ಜಮಸ್ಯಾಸು ಧತ್ತಮ್ | ಪ್ರಾಣಾನ್ರಶುಷು | ಪ್ರಜಾಂ ಮಯಿ  
 ಚ ಯಜಮಾನೇ ಚ | ನಿರಸ್ತಶ್ಚಂಡಃ | ನಿರಸ್ತೋ ಮರ್ಕಃ | ಅಪನುಶ್ಚಿ  
 ಶಂಡಾಮರ್ಕೌ ಸಹಾಮುನಾ | ಶುಕ್ರಸ್ಯ ಸಮಿದಸಿ | ಮನ್ರಿನಸ್ಯಮಿದಸಿ |  
 ಸ ಪ್ರಥಮಸ್ಸಂಕೃತಿರ್ವಿಶ್ವಕರ್ಮಾ | ಸ ಪ್ರಥಮೋ ಮಿತ್ರೋ ವರುಣೋ  
 ಅಗ್ನಿಃ | ಸ ಪ್ರಥಮೋ ಬೃಹಸ್ಪತಿರ್ವಿಶ್ವಾನ್ | ತಸ್ಮಾ ಇಂದ್ರಾಯ  
 ಸುತಮಾ ಜುಹೋಮಿ || ೫ ||

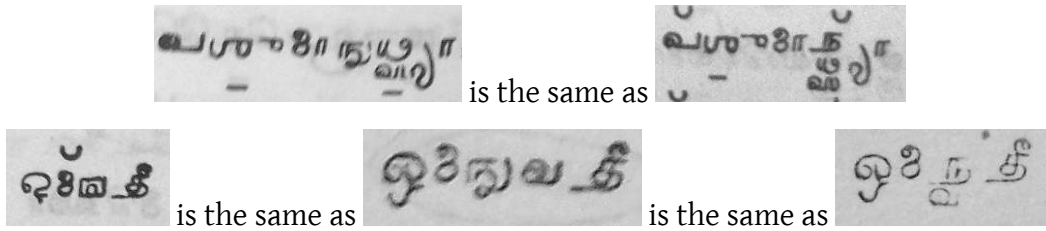
Ref 6 pp 1, 2

## §2.5. Discussion

From the foregoing samples, it will be clear that the three virama forms in Grantha – spacing, touching and ligated – are used in exactly the same way as the overt virama forms of Devanagari or Kannada. That is, wherever an overt virama form would be used in Devanagari or Kannada, any one of three virama forms can occur in Grantha.

In Devanagari or Kannada an overt virama form can occur: 1) at the end of a word, 2) as fallback for absence of sufficient conjoining forms or 3) by explicit choice of the writer to break a consonant cluster into separate orthographic syllables. The use of the virama forms of Grantha are seen in all three cases as seen in the samples of the ligated, spacing and touching virama forms respectively. However, it is not to be construed that ligating virama are only used at the end of sentences, spacing forms are only used as fallback etc, since I have previously shown the interchangeability of these virama forms.

Here are some more samples (provided with proper references on pp 22 and 23 of my proposal) that show the use of ligated virama forms in the middle of words and in free alteration with ligatures or conjoining forms:



Now the samples on the preceding pages are of the same Vedic texts (Taittirīya Saṃhitā and Brāhmaṇam) printed in different scripts; thus there is no scope for thinking that there is any difference in meaning between the Devanagari/Kannada and Grantha versions.

I must also remark in passing that though the ligated virama forms of Malayalam called chillaksharams have been analysed as conjoining forms, in Grantha none of the three virama forms is to be analysed as a conjoining form as in fact they are used as *fallback* for the *absence of conjoining forms* as shown above.

Thus it should be clear that each virama form of a consonant in Grantha is in its own right a full equivalent and cognate of the overt virama forms of Devanagari or Kannada or any other Indic script used to write Sanskrit. Any suggestion that there is a difference in meaning between the various virama forms of Grantha would imply that since Devanagari or Kannada or other Indic scripts do not have more than one virama form, they are incapable of representing Sanskrit properly, which is absurd.





As the virama form chosen for each consonant is merely a matter of typographic style, even if another virama form per consonant were chosen in another typographic style there would be no difference in the implied meaning and such an alternative presentation form would be equally valid. This variation is just as in the Bengali vowel sign ligature situation.

The only proper way of handling this situation would be to allow CONSONANT + VIRAMA to be displayed as the default virama form of the consonant using font tables. To quote the South Asian Committee's document L2/10-167 p 4:

*If the representations are truly equivalent and there is truly no semantic difference ..., then the rendering ... is a freely variable choice that can be worked out in font software. No underlying difference in encoding is warranted. The encoding model is therefore “pure” in terms of using the virama, which fits with the “script agnostic” nature of the Sanskrit text as documented.*

Thus,

CONSONANT + VIRAMA → DEFAULT VIRAMA FORM

where the default virama form may be ligated, touching or spacing as per the style. This is just as in Bengali old-style orthography the sequence CONSONANT + VOWEL SIGN U will render as a ligature for some consonants and non-ligated form for others. This is also as in Tamil orthography the same sequence CONSONANT + VOWEL SIGN U/UU will render as a ligature for the native Tamil consonants but with a distinct glyph for the vowel sign for the “Grantha consonants” JA, SHA, SSA, SA and HA (and the ligature K·SSA). Thus such a variation in the rendering of equivalent sequences is nothing unique to Grantha.

### *§3.2. Achieving the display of specific virama forms*

The parallel of the Grantha virama form situation and Bengali vowel sign ligature situation has been noted. Now just as in the case of Bengali vowel sign ligatures a mechanism is provided to unambiguously request a specific presentation form, it might also be useful to have such a mechanism for the various Grantha virama forms.

One should however remember that it is neither possible nor the practice in Unicode to provide an unambiguous mechanism of requesting every presentation form. Such mechanisms are provided only as far as it is both meaningful and possible to do so.

For example, and particularly relevant to the present discussion, there is no mechanism to specifically request consonant cluster ligatures in Indic, since the character intended for joining – ZWJ – is used in Indic to request conjoining forms and not ligatures.

The formation of ligatures is up to the fonts. “Individual fonts may provide fewer or more ligature forms”, to quote TUS 5.2 p 273. There are only mechanisms to break ligatures and request “lesser” presentation forms such as conjoining forms or overt virama forms.

### §3.2.1. Achieving ligated virama forms

Similarly, in the present situation of the virama forms of Grantha, when the default vowelless representation of a particular consonant in an orthographic style is a ligated virama form, a font for that orthographic style should by default provide that ligated form. As in the case of Devanagari consonant cluster ligatures, a particular font may provide such forms to a greater or lesser degree as per the style. If a particular font does not provide the amount of ligatures required by a particular style, it is neither the fault of the encoding model nor can the encoding model be expected to compensate for it. Thus it is the duty of a Grantha font to provide the required ligated virama forms by glyph substitution.

It should here be noted that not all consonants in Grantha are attested to have ligated virama forms, while touching and spacing forms are attested for all consonants. (For more details see my Grantha proposal L2/09-372 §5.3.3 pp 23-24.) Even among the ligated virama forms that are attested, most are highly archaic and to be found in manuscripts only, and most modern readers would not even recognize them. This is because only about seven (those of TA, NA, MA and N·NA and more rarely TTA, NGA and KA) are to be seen in contemporary orthography as defined by what is seen in printings of the past century (as noted in my proposal L2/09-372 p 21).

Thus whichever ligated virama forms are practically desired to be shown to the user will be provided by default by the font, just as in the case of the consonant cluster ligatures of Devanagari. This is also alike the situation of old-style Bengali orthography wherein the desired consonant-vowel sign ligatures are shown by default.

### §3.2.2. Achieving touching virama forms

Those consonants which do not take ligated virama forms in a particular orthographic style may either be presented in touching or spacing virama form. Among these two kinds of forms, the touching virama forms will also be shown as default – for the particular consonants which must take them – by appropriate glyph substitution.

### §3.2.3. Summary of rendering

Thus, in a font that caters to the “modern” typographic style of Grantha as attested to by the samples on p 17 the rendering would go as follows:

$$1) \text{ ক} + \text{◌} = \text{ক্}, \text{ খ} + \text{◌} = \text{খ্}, \text{ গ} + \text{◌} = \text{গ্}, \text{ ঘ} + \text{◌} = \text{ঘ্}$$

$$2) \text{ ঙ} + \text{◌} = \text{ঙ}, \text{ চ} + \text{◌} = \text{চ্}$$

$$3) \text{ ছ} + \text{◌} = \text{চ্ছ}, \text{ জ} + \text{◌} = \text{জ্}, \text{ ঝ} + \text{◌} = \text{ঝ্} \dots$$

By thus allowing the sequence CONSONANT + VIRAMA to be rendered as the default virama form per consonant for all consonants, one automatically ensures the proper representation of the Sanskrit language in the script and hence direct compatibility and equivalence with other Indic scripts like Devanagari and Kannada as seen and illustrated before. The interconversion of Sanskrit texts between Grantha and other Indic scripts will be a matter of elementary one-to-one mapping. Texts converted to Grantha from other scripts will automatically achieve their desired appearance by the use of appropriate fonts.

#### §3.2.4. Unambiguous mechanism for requesting spacing virama forms

Now just as in the case of old-style Bengali orthography, ZWNJ can be used to prevent the consonant and virama from glyphically combining. At the font level, the presentation of CONSONANT + VIRAMA as either a ligated or a touching virama form is handled in the same way i.e. by glyph substitution. Such a glyph substitution will be prevented by inserting a ZWNJ between the consonant and virama as CONSONANT + ZWNJ + VIRAMA (just as it is inserted between the consonant and vowel sign in Bengali):

$$1) \text{ ক} + \text{ZWNJ} + \text{◌} = \text{ক্}, \text{ খ} + \text{ZWNJ} + \text{◌} = \text{খ্},$$

$$\text{ গ} + \text{ZWNJ} + \text{◌} = \text{গ্}, \text{ ঘ} + \text{ZWNJ} + \text{◌} = \text{ঘ্}$$

$$2) \text{ ঙ} + \text{ZWNJ} + \text{◌} = \text{ঙ}, \text{ চ} + \text{ZWNJ} + \text{◌} = \text{চ্}$$

#### §3.2.5. Unambiguous mechanism for requesting touching or ligated virama forms

In this model there is no way to particularly request the formation of a ligated or touching virama form. It is not possible to suggest that, just as ZWNJ was used as per the old-style Bengali orthography model, ZWJ can be used as per the new-style Bengali orthography model to request a ligated or touching virama form, since the sequence CONSONANT + ZWJ + VIRAMA already has a defined function – that of requesting C2-conjoining forms of the

following consonant, especially in a South Indian script like Grantha which heavily uses those forms. The rendering of the sequence CONSONANT + ZWJ + VIRAMA when isolated would be different from that of the same sequence when part of a consonant cluster CONSONANT + ZWJ + VIRAMA + C2. If this is not considered a problem, then ZWJ may be used this way to request a ligated or touching form, so:

CONSONANT + ZWJ + VIRAMA → LIGATED OR TOUCHING VIRAMA FORM

This mechanism however does not make it possible to distinguish between ligated and touching forms. There is however no urgent need for that since there is no usage context where ligated and touching forms need to be discriminated between in presentation.

Of course, if it is not possible to give the sequence ZWJ + VIRAMA the above new definition, the matter may be simply dropped since the ligated and touching forms are to be provided by default anyway just as consonant cluster ligatures are. There is no need for a mechanism for requesting touching or ligated virama forms just as no urgent need has ever been experienced or voiced by anyone in the case of consonant cluster ligatures.

Even the sequence of CONSONANT + ZWNJ + VIRAMA to prevent ligated or touching forms and hence effectively request spacing virama forms is only provided for hypothetical use. By the use of an appropriately designed font one would be able to achieve the desired virama form for each consonant without any problems and without needing to use any invisible or abnormal character to select a particular virama form among them.

### §3.2.6. Conformance of this model to Unicode standards

Note that this model does not prescribe the use of ZWJ/ZWNJ to achieve the default appearance of the text. Thus there is no room for anyone to accuse this model of complicating the existing Indic joiner situation. The model prescribed here is perfectly in accordance with the directives on ZWJ/ZWNJ in TUS 5.2 p 504:

*The ZWJ and ZWNJ are designed for marking the unusual cases where ligatures or cursive connections are required or prohibited. These characters are not to be used in all cases where ligatures or cursive connections are desired; ... they are meant only for overriding the normal behavior of the text.*

### §3.3. Virama forms within consonant clusters

I should note here that even though ZWJ or ZWNJ will not be needed or used to select a particular virama form, if ZWNJ is used for breaking up consonant clusters, one of the three virama forms will certainly appear.

For example,

1) வஷ்டு + ெ + கூரஃ = வஷ்டுகூரஃ

வஷ்டு + ெ + ZWNJ + கூரஃ = வஷ்டுெ + ZWNJ + கூரஃ = வஷ்டுெகூரஃ

2) ஁ர + ெ + ஸ = ஁ஸு

஁ர + ெ + ZWNJ + ஸ = ஁ரெ + ZWNJ + ஸ = ஁ரெஸ

3) ஁ந + ெ + வதீ = ஁நஙதீ

஁ந + ெ + ZWNJ + வதீ = ஁நு + ZWNJ + வதீ = ஁நுவதீ

(Note: Compare the darśa and omanvatī renderings with the samples on pp 10 and 16.)

Now what has happened here is that the ZWNJ has broken the consonant cluster into two parts to be rendered separately. The preceding and following parts are rendered as if the other did not exist, and hence the vowelless consonant preceding the ZWNJ is rendered as its default virama form. This is in accordance with TUS 5.2 p 506:

*... A ZWNJ after a sequence of consonant plus virama requests that conjunct formation be interrupted, usually resulting in an explicit virama on that consonant. ... ZWNJ will normally have the desired effect naturally for most fonts without any change, as it simply obstructs the normal ligature/cursive connection behavior.*

The Indic rendering chapter at TUS 5.2 p 266 also states (underlining mine):

*... the Unicode Standard adopts the convention of placing the character U+200C ZERO WIDTH NON-JOINER immediately after the encoded dead consonant that is to be excluded from conjunct formation. In this case, the virama sign is always depicted as appropriate for the consonant to which it is attached.*

Note the underlined part. The *appropriate* way of depicting vowelless consonants in Grantha may be any one of three virama forms as per the typographic style. We have also demonstrated that all these three forms are full equivalents of the single overt virama forms of other Indic scripts like Devanagari and that they are not conjoining forms. Thus the sequence CONSONANT + VIRAMA + ZWNJ can and should be rendered as the default virama

form of the particular consonant in Grantha. Only then will the attested full equivalence of the Grantha virama forms to those of other Indic scripts be preserved.

### *§3.4. Summary of preferred virama model*

The preferred virama model described above may be termed the “single virama model” in view of the need to distinguish between other virama models that have been (and will be presently) considered for Grantha.

In this model, the basic sequence CONSONANT + VIRAMA will be rendered as the default virama form of the consonant in the desired typographic style:

$$\text{CONSONANT} + \text{VIRAMA} \rightarrow \text{DEFAULT VIRAMA FORM}$$

Ligated or touching virama forms are to be provided by the font by default, so the way of achieving those forms is to use an appropriate font. ZWNJ can however break such consonant-virama combinations to effectively request a spacing virama form:

$$\text{CONSONANT} + \text{ZWNJ} + \text{VIRAMA} \rightarrow \text{SPACING VIRAMA FORM}$$

When CONSONANT + VIRAMA occurs as part of a consonant cluster, the normal Indic behaviour of consonant ligatures or conjoining forms (stacks in Grantha) will occur:

$$\text{C1} + \text{VIRAMA} + \text{C2} \rightarrow \text{LIGATURES/STACKS}$$

When such clusters are broken up by ZWNJ, the default virama form of the consonant will once more appear:

$$\text{C1} + \text{VIRAMA} + \text{ZWNJ} + \text{C2} \rightarrow \text{DEFAULT VIRAMA FORM OF C1} + \text{C2}$$

## §4. Other possible models

For the sake of completeness and to emphasize why the single virama model described in §3 above is the “preferred” one, we here briefly consider other virama models as well.

The “preferred” model as described above has the one apparent shortcoming that it does not provide a specific sequence to unambiguously request ligated or touching virama forms. It relies on the font to provide those virama forms. Other models may be able to provide specific sequences for specific forms. We will consider those models one by one.

### *§4.1. Non-decomposing two-virama model*

This model proposes two virama characters, a regular virama denoted by just VIRAMA and a second virama termed LIGATING VIRAMA. This model is as follows:

$$\text{CONSONANT} + \text{VIRAMA} \rightarrow \text{SPACING VIRAMA FORM}$$
$$\text{CONSONANT} + \text{LIGATING VIRAMA} \rightarrow \text{LIGATED VIRAMA FORM}$$

This model was proposed by the author of L2/09-345. It rejects the existence of touching virama forms (see L2/10-062 and 10-154). It also does not preserve the semantic equivalence between the spacing and ligated virama forms claiming that semantic non-equivalence is the fact just as in Malayalam.

**Pros:**

There are straightforward sequences for spacing and ligated virama forms.

**Cons:**

This model does not support touching virama forms in any way, despite sufficient attestation for the same. (See p 9 of the present document and L2/09-372 pp 21 and 35.)

Further and more importantly, this model goes against the semantic equivalence of the virama forms which has been proven above. Even if it is suggested that said equivalence need not be supported at the encoding level and that higher protocols may be used for this, the encoding dichotomy prevents the basic sequence CONSONANT + VIRAMA from being rendered as the default virama form per consonant. Thereby the interconvertibility of Sanskrit texts between Grantha and other Indic scripts is effectively crippled.

While composing Sanskrit texts in other Indic scripts one never has to use any *abnormal* characters other than the default vowels, consonants and virama. In Grantha, one would have to use the *abnormal* character that is the LIGATING VIRAMA which is abnormal because it does not fit into the pan-Indic model and hence affects interconvertibility between Grantha and other Indic scripts as noted above. Even though other Brahmic scripts like Khmer have two virama characters, the interoperability of Grantha with other *Indic* scripts is a major issue which must be provided for to ensure the proper representation of the script as used for denoting its native language viz Sanskrit.

If CONSONANT + VIRAMA cannot be rendered as the default virama form per consonant, and if the LIGATING VIRAMA is to be exclusively used for achieving ligated virama forms, it would mean that the typographic style is being implemented at the encoding level and that the encoding is moving from a character-based system to a glyph-based one which goes against the basic principles of Unicode.

For all the above reasons, this model is *not* acceptable for Grantha Unicode.

**Note:**

This model initially gave the second virama the Malayalam-based name of CHILLU MARKER. I had objected to the name in my various documents since the word 'chillu' is



unknown among the native Grantha user community. The author of this model then (in L2/10-303) accepted the name change to LIGATING VIRAMA.

Whatever be the name, the model is the same, since the second virama is distinct from the regular virama at the encoding level, being devoid of a canonical decomposition to the regular virama. This results in the restriction of CONSONANT + VIRAMA to the spacing virama forms with undesired effects as described above. This is why I call this the “non-decomposing two-virama model”.

#### *§4.2. Decomposing two-virama model*

The non-decomposing two-virama model described above does not allow the sequence CONSONANT + VIRAMA to be presented as the default virama form per consonant. It implies a semantic distinction between the various virama forms at the encoding level. It also does not support the attested touching virama forms of Grantha. These faults may be rectified if the LIGATING VIRAMA were provided a canonical decomposition to the regular VIRAMA resulting in the “decomposing two-virama model” as described below:

CONSONANT + VIRAMA → DEFAULT VIRAMA FORM

CONSONANT + ZWNJ + VIRAMA → SPACING VIRAMA FORM

CONSONANT + LIGATING VIRAMA → TOUCHING VIRAMA FORM

CONSONANT + ZWJ + LIGATING VIRAMA → LIGATED VIRAMA FORM

This model was proposed by the present author in L2/09-372 (p 35). It intends to provide unambiguous mechanisms for requesting each attested virama form in Grantha and at the same time maintain the two important constraints of allowing CONSONANT + VIRAMA to be presented as the default virama form and (thereby) preserving the semantic equivalence of the various virama forms at the encoding level. The explanatory document L2/09-375 provides (on p 5) justification for each sequence given above.

#### **Pros:**

This model provides unambiguous mechanisms to request all three virama forms.

This model allows CONSONANT + VIRAMA to be presented as the default virama form which is important for pan-Indic compatibility as previously demonstrated.

This model preserves the semantic equivalence between the various virama forms, as even though the LIGATING VIRAMA is encoded separately, it will become equivalent to the regular VIRAMA for purposes like text search and collation thanks to the decomposition.

It is the decomposition that also enables the CONSONANT + VIRAMA to be presented as the default virama form. Without the decomposition, allowing CONSONANT + VIRAMA to be

presented as any other virama form than spacing, i.e. as touching or ligated, would cause security problems in IDN etc as the use of CONSONANT + [ZWJ +] LIGATING VIRAMA would also produce the same display. The decomposition removes this security problem.

#### Cons:

The variegated (albeit logical) combinations of ZWJ/ZWNJ and VIRAMA/LIGATING VIRAMA may not be straightforward for everyone to understand. People have therefore accused this model of confusing the existing joiner situation in Indic (L2/10-154 p 6).

The use of CONSONANT + LIGATING VIRAMA (without the intervening ZWJ) to produce any virama form other than what has been called ‘ligated’ is also unacceptable to some, who relate the word LIGATING in the name LIGATING VIRAMA to the ligated virama forms.

The requested decomposition of the LIGATING VIRAMA may not be entirely in line with the existing Unicode standards for providing such decomposition. Till now, canonical decomposition is only provided in Unicode when there are encoding-wise different but glyphically and semantically equivalent character sequences for composing the same text. While the LIGATING VIRAMA would be semantically equivalent to the regular virama, it has no glyphic form, being merely a technical device like 17D2 KHMER SIGN COENG. Therefore the decomposition might be not entirely self-evident or in line with pan-Unicode standards.

#### Note:

To satisfy those who relate the word LIGATING in the name LIGATING VIRAMA to these forms as mentioned above, a variant of this model can be suggested in which CONSONANT + LIGATING VIRAMA produces the ligated virama forms, leaving the touching forms without an unambiguous representation. Or, CONSONANT + LIGATING VIRAMA may be allowed to present either ligating or touching forms with the display of either being up to the font. This leaves *both* ligated and touching forms without a mutually exclusive unambiguous representation.

### §4.3. Discussion

Of the two alternative virama models for Grantha presented above, both of which employ two virama characters, the non-decomposing model is entirely unacceptable as it does not preserve the representation in the Grantha script of its native language Sanskrit. The South Indic committee document L2/10-299 which advocated the name change of the second virama from CHILLU MARKER to LIGATING VIRAMA states:

*GRANTHA SIGN LIGATING VIRAMA ... provides a way to represent the consonant-virama ligations for Sanskrit. If it also provides a way to handle Malayalam transliteration, that is another point in its favor, but **the proposed solution***

*focuses on addressing the Sanskrit situation, a primary concern raised in the documents above (and in the feedback documents in #5 Grantha User-feedback, below).*

This heartens me because it puts the transliteration of Malayalam in the secondary position that it should take, and attaches primary importance to the writing of Sanskrit. However, this document does not speak clearly about whether the decomposition of the second virama to the regular one which is needed for Sanskrit is provided or not. While it was I who advocated the LIGATING VIRAMA for Sanskrit (as against a CHILLU MARKER for Malayalam) I have been from the very beginning very careful as to insist on the decomposition as without it proper representation of Sanskrit would be impossible:

From my original proposal L2/09-372 p 35:

*... we suggest the encoding of at least one other virama, which should be non-spacing (as against the spacing nature of the default virama) and **most importantly be canonically decomposable to the default virama in order to maintain semantic equivalence.***

From my “Grantha virama ligatures clarification” document L2/09-375 p 5:

*encode a separate ligating virama **which should nevertheless be canonically decomposable to the regular virama in order to maintain semantic equivalence.***

From L2/10-267 p 12:

*If the same chillu marker were to be renamed as a ligating virama **and provided a decomposition to the regular virama**, then I would have no objection to there being such a character that CONSONANT + THAT CHARACTER produces a virama ligature.*

Therefore while the South Indic committee in L2/10-299 recommended the name change from CHILLU MARKER to LIGATING VIRAMA based on my comments in L2/10-267 quoted above, I must remind the South Indic committee and the UTC that a mere name change cannot “address the Sanskrit situation” which revolves around the matter of semantic equivalence for which canonical decomposition is needed if a second virama is encoded.

I also quote the original South Indic committee document L2/10-167 which discussed the chillu marker/ligating virama matter and pronounced (on p 4):

*... we recommend the use of the GRANTHA SIGN VIRAMA and, for chillu forms, fonts (and ligature tables) be employed. That is, **we recommend against***

*encoding any chillus, chillu marker, or any special means of indicating the presence or absence of chillu forms. If the representations are truly equivalent and there is truly no semantic difference between chillu and non-chillu representation of texts, then the rendering as chillu or conjunct is a freely variable choice that can be worked out in font software. No underlying difference in encoding is warranted. The encoding model is therefore “pure” in terms of using the virama, which fits with the “script agnostic” nature of the Sanskrit text as documented.*

If a second virama character is to be encoded and yet there should be – at least *effectively* – no underlying difference in encoding, the canonical decomposition is absolutely necessary. If however the canonical decomposition is not possible for the reasons I myself stated in the “cons” section of §4.2 above or for any other reasons, there should be no second virama character as it will certainly cripple the representation of Sanskrit in the Grantha script.

## §5. Representing Malayalam

For those who want to represent the archaic form of the Malayalam script which was called Malabar Grantha, there is always the existing Malayalam encoding in the BMP. By the use of an appropriate font, this Malayalam encoding will be best suited to represent any form of the Malayalam script, being as it is equipped with the scholarly-use TTTA and NNNA, dot-reph, and most importantly for the various parties concerned here the atomically encoded and semantically different chillu characters.

## §6. Conclusion

As there is no proof that the Malayalam script is at all written in Grantha today, and since it is universally accepted by all accredited manuscriptologists and academically qualified native scholars using Grantha (see L2/10-233, L2/10-283 and L2/10-285) that the Grantha script is used for Sanskrit only, the Grantha encoding must not be tailored for representing Malayalam but it must be tailored to fit in with the pan-Indic model of scripts used to represent Sanskrit – especially Devanagari which is the de facto universal standard for Sanskrit – and it is obvious that Devanagari has only one virama character.

That any purpose of transliteration of non-native languages must take a secondary position to the representation of the native language of a script is obvious. If a CHILLU MARKER or in its alternate name the non-decomposing LIGATING VIRAMA is encoded ostensibly

for the purposes of Malayalam transliteration, the representation of the native language Sanskrit in the script would be crippled. This would be analogous to getting your king checkmated in trying to save a bishop in chess and hence should not be done. Furthermore, I have pointed out as far back as L2/09-316 p 8 predating my own Grantha proposal that seeing as there are no attested ligated virama forms in Grantha for RRA, LA or LLA, it is impossible to transliterate one-to-one the corresponding chillus of Malayalam. Thus even the goal of transliterating Malayalam one-to-one is impossible. What meaning would then there be in encoding either a CHILLU MARKER or non-decomposing LIGATING VIRAMA by the name of Malayalam transliteration purposes?

The single virama model, which has been presented as the preferred model for handling the virama forms of Grantha is devoid of the problems associated with either of the two-virama models. The alleged shortcoming of the single virama model, viz not providing unambiguous sequences for the ligated and touching virama forms is not really a shortcoming, since:

- 1) There is no urgent need for such unambiguous sequences as the desired default appearance can be achieved by the use of appropriate fonts.
- 2) Even if there is said to be such a need, the two-virama models which attempt to provide such sequences still depend on fonts to support them. So the fact that the single virama model depends on appropriate fonts is not its unique fault.
- 3) Even if one attempts to design such appropriate fonts, not all Grantha consonants are attested to have ligated virama forms and hence any sequence that requests ligated virama forms in the two-virama models would be meaningless for those consonants. In the single-virama model, the desired virama form for each consonant is automatically achieved by the single consistent and straightforward sequence CONSONANT + VIRAMA and so there are no meaningless encoded sequences.

Thus the single virama model as justified and described in detail in this document, and succinctly summarised in §3.4 on p 23, is recommended as the preferred virama model for Grantha, being devoid of faults and compliant with the restrictions that are implied by the semantics of the Grantha script and its virama forms as described in §2.

Those who desire the representation of Malabar Grantha of the 18<sup>th</sup> century, which is the only script form carrying the name Grantha attested to be used for Malayalam, can best achieve all their desired intentions and purposes using the Malayalam encoding and an appropriate font. Hence the semantics of Grantha should not be sacrificed for those ends.

## §7. References

- 1) Wikipedia article on Malayalam script;  
[http://en.wikipedia.org/w/index.php?title=Malayalam\\_script&oldid=386389303#Grantha](http://en.wikipedia.org/w/index.php?title=Malayalam_script&oldid=386389303#Grantha)
- 2) Alphabetum Grandonico-Malabaricum Sive Samscrudonicum, Rome, 1772;  
<http://books.google.com/books?id=qmETAAAAQAAJ&dq=Alphabetum%20Grandonico-Malabaricum%20Sive%20Samscrudonicum>
- 3) Taittirīya Saṃhitā, Devanagari script, Ed: Dr R Krishnamurti Shastri and R Ganeshvara Dravid, Shri Nrisimha Priya Trust, Mylapore, Chennai, 2003.
- 4) Taittirīya Brāhmaṇam, Devanagari script, Ed: Dr R Krishnamurti Shastri and R Ganeshvara Dravid, Shri Nrisimha Priya Trust, Mylapore, Chennai, 2003.
- 5) Taittirīya Saṃhitā part 1, Kannada script, Ed: Swami Chidananda, Ramakrishna Ashram, Bangalore, 2000.
- 6) Taittirīya Brāhmaṇam part 1, Kannada script, Ed: Swami Chidananda, Ramakrishna Ashram, Bangalore, 2000.
- 7) Taittirīya Saṃhitā, Grantha script, Heritage India Educational Trust, Mylapore, Chennai, 2006.
- 8) Taittirīya Brāhmaṇam, Grantha script, Heritage India Educational Trust, Mylapore, Chennai, 2006.

-o-o-o-