L2/17-182

Comments on encoding the Tigalari script

Srinidhi A and Sridatta A

Tumakuru, India

srinidhi.pinkpetals24@gmail.com, sridatta.jamadagni@gmail.com

May 26, 2017

This document provides comments on L2/16-241 Preliminary proposal to encode Tigalari script in Unicode and L2/16-342 Recommendations to UTC #149 November 2016 on Script Proposals.

Tigalari is a major Indic script developed from Brahmi predominantly used to write Sanskrit. We also find few examples of Kannada and Tulu languages. The proposal is well researched, but we request to incorporate the comments provided in this document in upcoming revision of the proposal. The comments provided here are based on meticulous study of original Tigalari manuscripts.

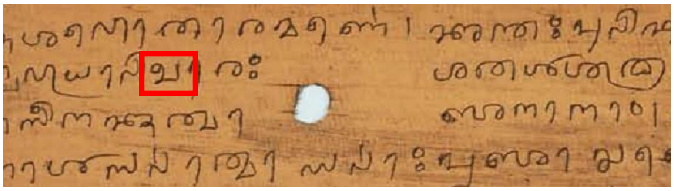
Several folios shown here are from the digitized collections of the French Institute of Pondicherry (IFP).

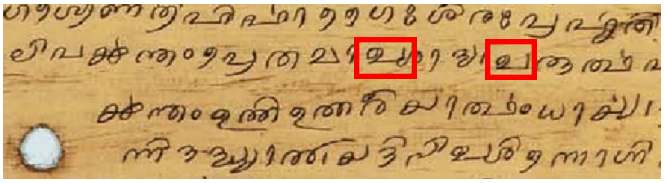
1 Glyph shapes

Since Tigalari was never employed in Print, we find minor variations in handwriting of manuscripts. However we see similarity in glyph of characters in their orthography. Some of the shapes of characters are erroneous whereas some require minor corrections. Most commonly used forms can be considered as Standard glyphs. Also see character inventory which is provided at the end of the document.

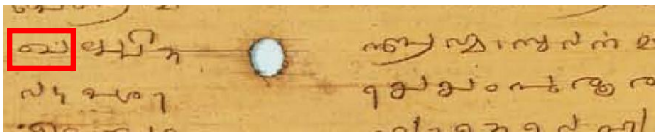
1.1 The characters which require major glyph change

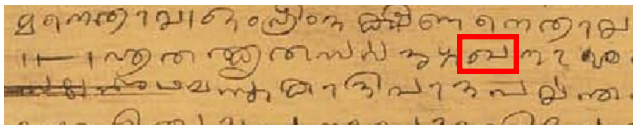
KHA The proposal provides 𑒖 as glyph shape. KHA is usually written as 𑒅. The current glyph is confusable with CA which is also written as 𑒖. 𑒅 is recommended as it the major glyph found in manuscripts.





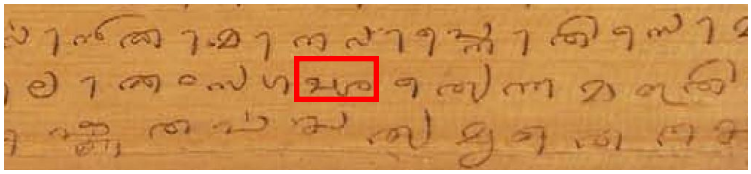
The above folios shows ഛ used for CA.

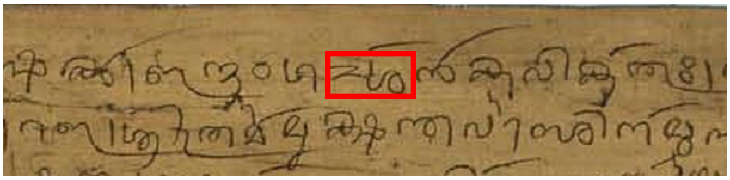




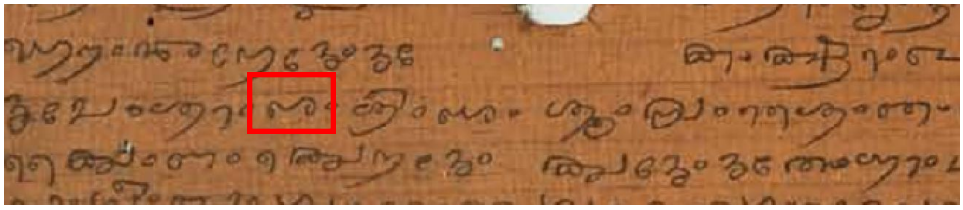
The above folios shows ഖ used for KHA.

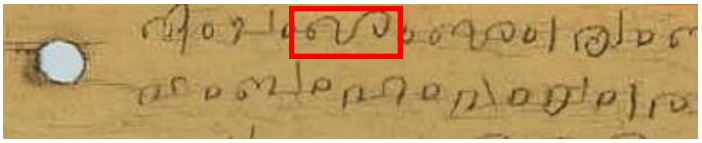
CHA The current glyph ച്ഛ does not represent CHA. It represents the conjunct *ccha* (CA+VIRAMA+CHA). CHA is written as ഛ .





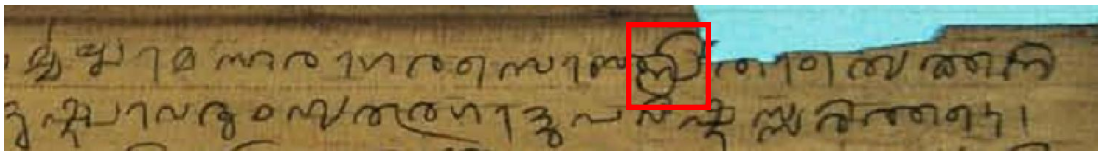
ച്ഛ used for *ccha.*

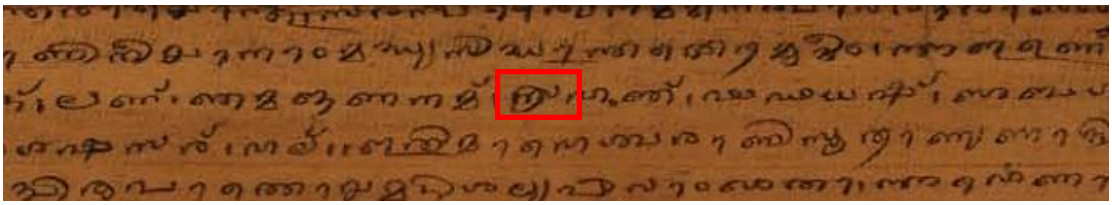


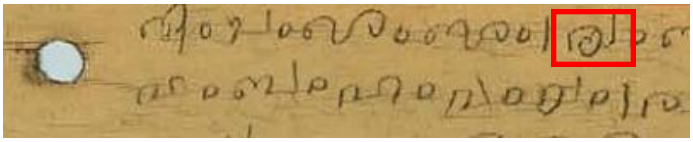


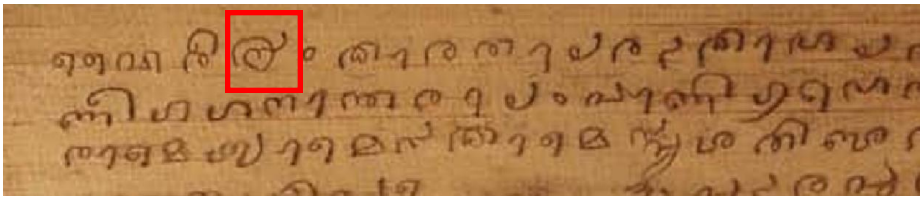
ഛ used for CHA.

JHA The current glyph ഝ is uncommon. ഩ is the major glyph attested in manuscripts. The same should be used in Code chart. The current glyph is likely a variation of the Standard form.

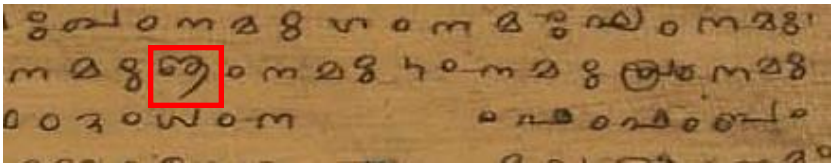


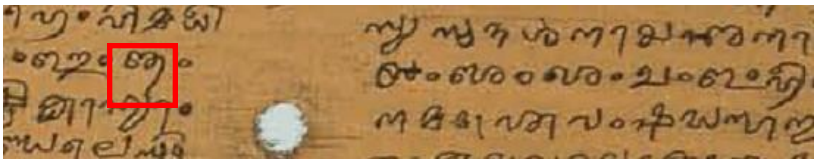




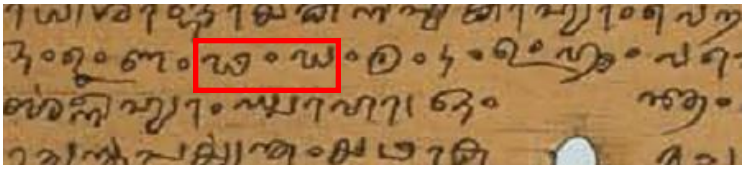


NYA ഞ is not attested in manuscripts as seen by the authors. ഋ is the glyph used in manuscripts.

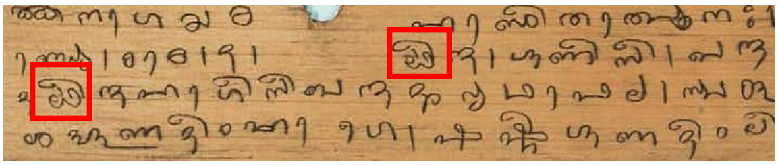


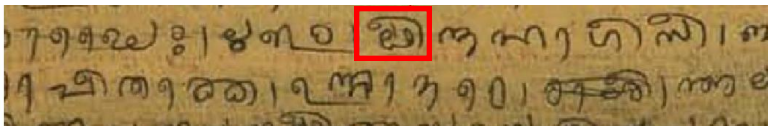


DDA and DDHA The proposal provides ഡ and ഢ as glyph shapes and identifies ഡ and ഢ as alternate forms. ഡ and ഢ are the major glyphs used in manuscripts. The current glyphs are not generally used.

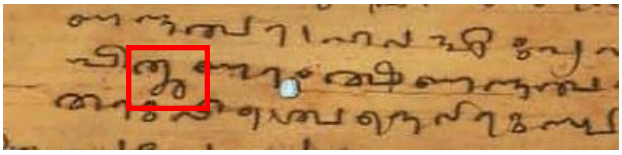


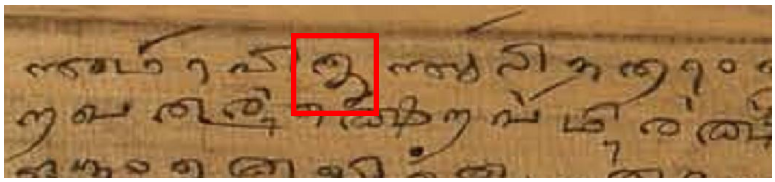
RRA ൟ is the major glyph. ൞ is cursive variation found in few manuscripts. ൟ should be employed in Code chart.

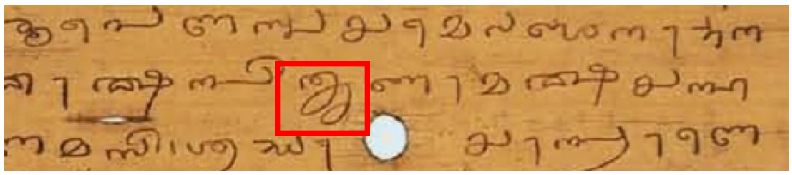




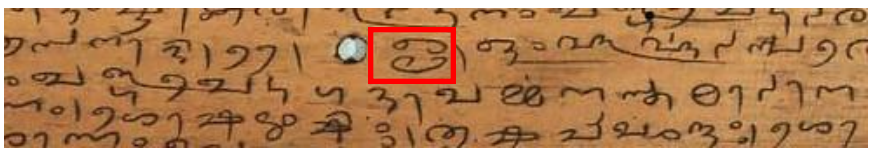
VOWEL SIGN VOCALIC RR The current glyph ൃ appears in Malayalam manuscripts. The form in Tigalari is different from that of Grantha and Malayalam. It is generally written as ഀ.

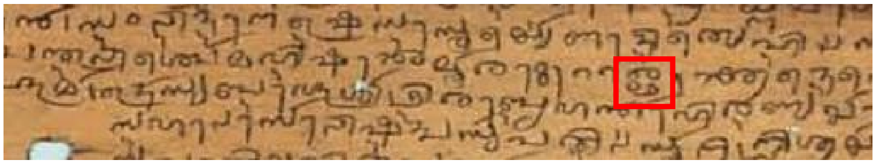


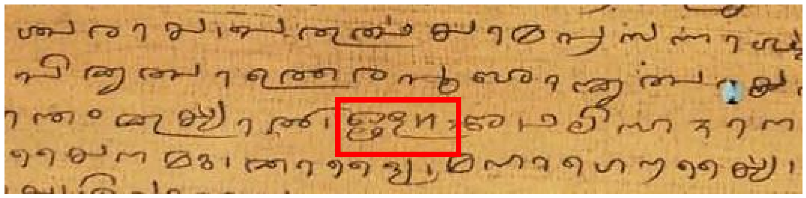




113D2 The current glyph 𑏒 appears to be a handwritten variant. The original glyph is 𑏓 which is composed of two semi-circular like elements facing each other.

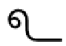






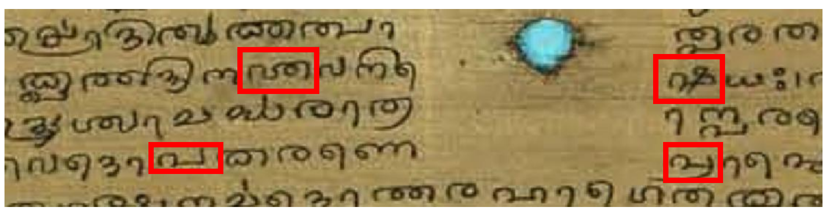
1.2 The characters for which minor glyph changes are suggested

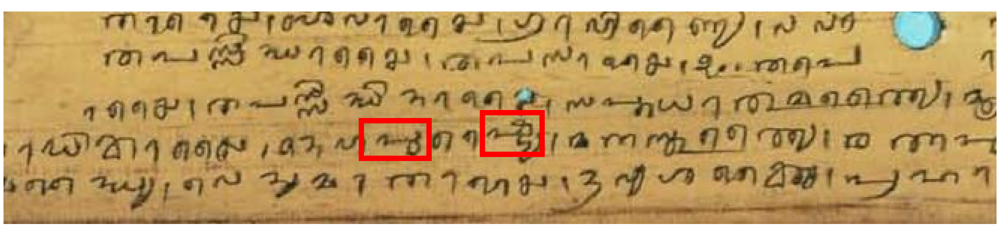
The shapes of characters listed below are correct, but some minor changes are suggested to keep uniformity in shapes with other characters.

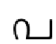
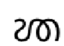
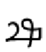
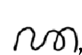
UU  is formed from the combination of  and  . The glyph  is recommended.

EE, AI and SSA

The Letters PA, SSA, EE and AI belong to same class with some components graphically similar.

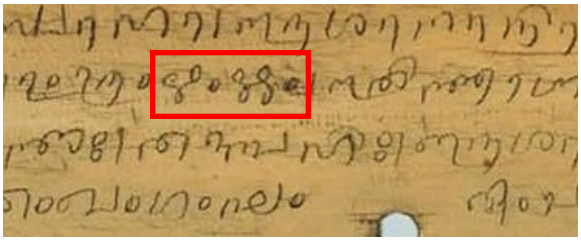




Here we can notice the shapes of PA, EE and SSA should be similar. However current  ,  and  are dissimilar in shapes. The glyphs of EE and SSA may be changed to  ,  . Correspondingly AI can be changed.

VOCALIC RR

The knot in  appears to be bigger. It can be written similar to  as shown in figure below.



Other characters

It is natural in handwriting, some characters lean towards left or right. It is suggested that Standardized glyphs may be linear in appearance.

Character name	Current glyphs	Suggested glyphs
OO	ൗ	ൗ
AU	ൗൗ	ൗൗ
NGA	ങ	ങ
JA	ഝ	ഝ
TTA	ഺ	ഺ
NNA	ണ	ണ
DA	ദ	ദ
BA	ബ	ബ
VA	വ	വ
SHA	ശ	ശ
SA	സ	സ

2 Vowel signs

2.1 Contextual vowel signs of U and UU

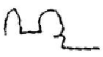
Page 8 provided information on TIGALARI VOWEL SIGN U & TIGALARI VOWEL SIGN UU which change their shape depending on the consonant it combines with. But complete list was not given. This section gives details of such combinations.

VOWEL SIGN U - ◌ു is the default sign used as in pu. Some consonants have two and sometimes three contextual forms. For example ku is written as either കു or കു. All forms are equivalent. Similar to Malayalam ന ണ, nu and ṇu take forms ന and ന.

Consonant	◌	◌ു	◌ൂ
ku	കു	കു	–
gu	ഗു	ഗു	–
chu	ചു	ചു	–
ju	ജു	ജു	–
ṭu	ടു	–	ടൂ
ṇu	ണു	ണു	ണൂ
tu	തു	തു	–
nu	നു	–	നൂ
bhu	ഭു	ഭു	–
ru	രു	രു	രൂ
śu	ശു	ശു	–
hu	ഹു	ഹു	–

VOWEL SIGN UU- ◌ൂ is default sign as used in pū. ◌ു is used for some consonants .

Consonant	◌ു	◌ൂ
kū	കൂ	–
gū	ഗൂ	–
chū	ചൂ	–
jū	ജൂ	–
ṭū	ടൂ	ടൂ

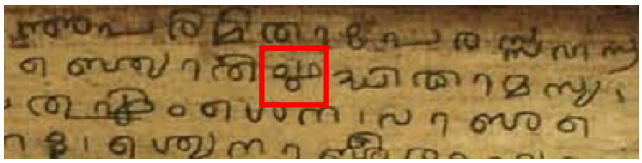
ṇū		
tū		–
nū		
bhū		–
rū		
śū		–
hū		–

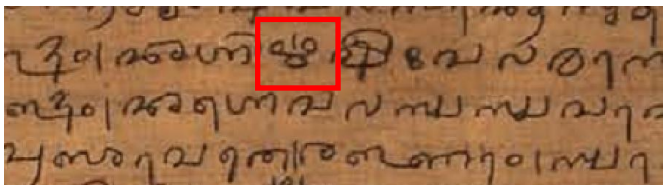
The forms  and  are rarely seen.

2.2 Contextual form vowel sign Vocalic R

Similar to Devanagari, Kannada, Bhaiksuki and so on, when VOWEL SIGN VOCALIC R occurs with RA, the sequence is written as repha placed above the vowel VOCALIC R.

Similarly, the phonological sequences can be extended to the other vocalic sounds (rr, l, ll)



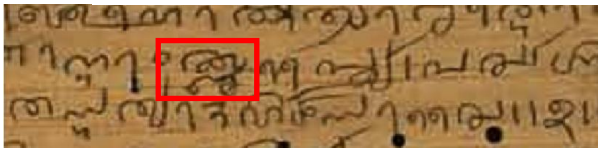


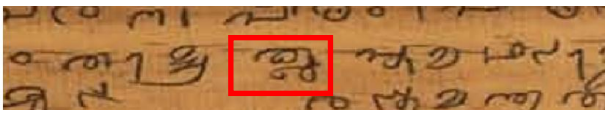
2.3 Vowel signs of Vocalic L & LL

The proposal identifies Vocalic L & LL in Tigalari has three alternate forms.

1. Placed directly below:

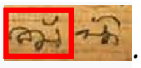
These forms are major glyphs and are commonly attested in manuscripts. The same forms should be used in Code chart. The forms of VOWEL SIGN VOCALIC L  as seen in manuscripts.





VOWEL SIGN VOCALIC LL as seen in manuscripts 

2. Placed below and to the right; ligating :

This cannot be considered as an alternate form as it a conjunct which is also seen in other south Indian scripts. kḷ is sometimes written by this sequence kl (KA+ Virama+ LA+ Virama) as seen here  .

The script ad hoc report L2/16-342 says

"To handle representation of vocalic L and vocalic LL, the situation in Bengali (§12.2 in TUS, p. 467ff.) may be used as a guide. In Bengali, a font implementation can choose whether the ligature of the C + vowel combination is the default. If the non-ligated form is the default, then ZWJ can be used as a hint to request the ligature form. If the ligated form is the default for a font implementation, then ZWNJ can be used to block the ligature (see figures 12-11 and 12-12 in TUS)."

The above representations involving ZWJ and ZWNJ are not required as it is not an alternate form.

3. Placed above the base and to the right:

Attestations were not provided in the proposal for these alternate forms. We have not seen these forms in the manuscripts. The script ad hoc report says *"The post-base form does not require a ZWNJ, but should be encoded as an independent vowel."* Even if these forms are attested, these should be considered as variants and managed at font level as in Grantha. Independent encoding is not required.

3 Virama

The proposal says "Virama behaves as a ZWJ character between two consonants creating a conjunction character by default similar to other Indic scripts."

This is confusing as Virama and ZWJ are two distinct terms. Instead it can be written as

The Virama has three functions. It is used as a halanta for marking the absence of the inherent vowel of a consonant letter and a control character that is used for producing conjuncts. It is also used to represent short un-rounded ‘u/uu' in Tulu similar to Samvrutokaram.

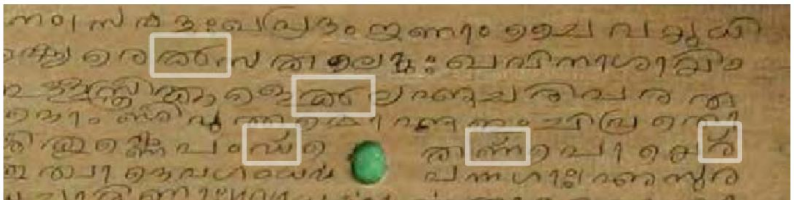
The proposal raised some questions

1. In Tulu, semi-circular Virama mark represents an equivalent of Samrutokaram (Front un-rounded and back retroflex vowels for A and E sounds) present in Malayalam. Should they be encoded separately?

Our opinion is it should be represented using Virama. It should not be encoded separately as in Malayalam.

2. Do chillaksharas need separate encoding? Since Tigalari is used primarily for Sanskrit, its standard behaviour is different to how it'll behave for Tulu. The chillus do not carry the un-rounded sounds when used for Tulu. and only the context of Tulu is a Char+Semi-Cir.Virama ≠ ChilluCharacter.

In Sanskrit both forms (ligated and unligated) are semantically equal. Even in Tulu the ligated forms also carry un-rounded sounds. The unrounded sound can represented with either Char+Semi-Cir.Virama or ligated virama (Chillu like characters). Hence are both forms whether in Sanskrit/Tulu are semantically equal, separate encoding is inessential. It is seen in Figure 14 of the proposal



Here the dative suffix kụ as in [...]ṣyere*k*, strīkuḷe*k* is represented using ligated Virama.

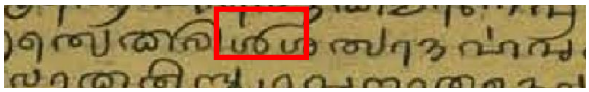
In above Tulu text both forms ligated and semi circular virama are used for un-rounded sounds.

The script ad hoc report says “Provide a full list of the ligature forms occurring with virama, and identify the default shape. Are there only four (K, T, TT, and N), or is there evidence for more (as suggested in footnote 18 on page 13)?”

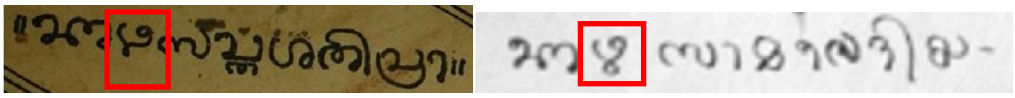
In addition to K, T, TT and N, ligated forms are also seen in G and sometimes for M. The footnote 18 on page 13 says

ൻ[NNA+*Virama*], ർ [RA+*Virama*], ൽ [SHA+*Virama*] might be present as-well. Needs further examination.

As analyzed by authors NN, R and SH are represented by regular virama. Ligated viramas are not found in manuscripts for these consonants. The below folio shows ligature form of Virama for GA ഗ.



The form of M similar to Grantha 𑌂 is seen in few manuscripts. It is not commonly seen in manuscripts.



4 Characters whose encoding is not required

Some of the proposed characters need additional research and their encoding is not required at present. The details are given as follows.

4.1 Short (hrasva) E and O

As said in proposal TIGALARI LETTER O & TIGALARI LETTER E are not present in the traditional Tigalari orthography. These two characters are essential for writing Tulu and Kannada languages. As there are no attestations available and other scholars and users may suggest different shapes for them. We request to reserve the code points for these vowels and their vowel signs. They may be proposed in future after having consensus on their encoding from native users and scholars.

4.2 Digits

In general Tigalari script uses Kannada digits. The proposed digits are similar to Malayalam and manuscripts provided in Figures 20 & 21 appear to be in Malayalam script rather than Tigalari. The letter forms, style and orthography closely resemble Malayalam script. Other sources provided are based analysis of these manuscripts. If these manuscripts belong to Malayalam script, independent encoding of digits is not required.

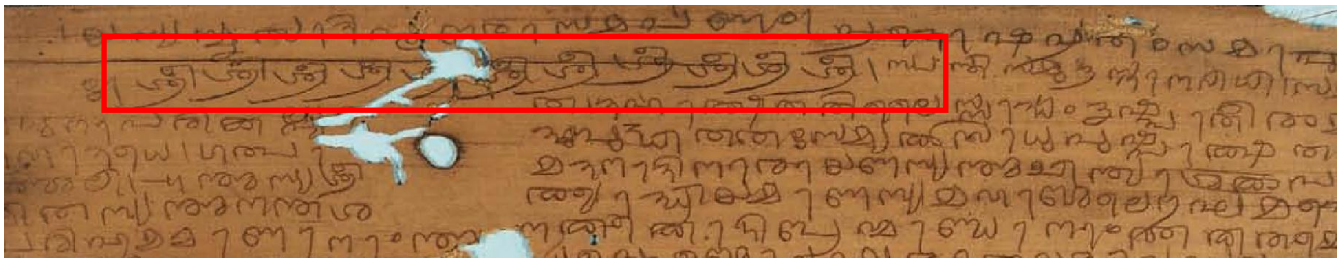
4.3 TIDDU SIGN

This character is not unique to Tigalari; it is also attested in other Indic scripts. It is reasonable to encode in a generic block instead in Tigalari.

4.4 SHRII

 appears to be joined form of  . Both forms may be 1. read as a ligature 2. a symbol indicating beginnings, pauses, endings or space fillers. There is no need to distinguish them in plain text. Its independent encoding needs further examination and it is not required to encode it at present.

For example in below folio  is used multiple times a space filler.

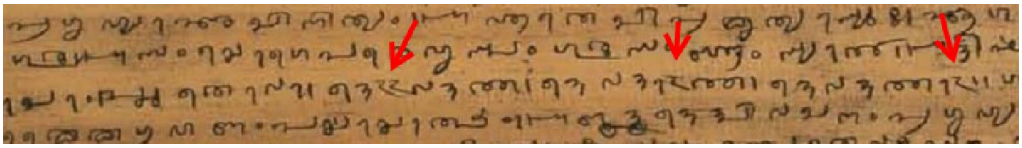


5 New characters which require encoding

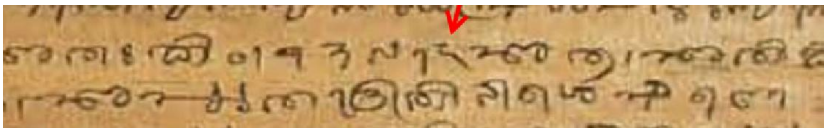
The following characters require encoding which are not proposed in L2/16-241.

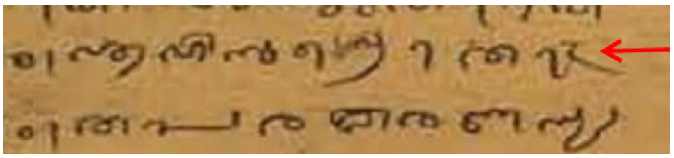
5.1 PLUTA

A pluta sign  is attested for Tigalari as denoting 'extra-normal' length for a vowel. It corresponds to GRANTHA SIGN PLUTA ૣ. The attestations are from the Sanskrit Grammatical work Madhyasiddhāntakaumudī.



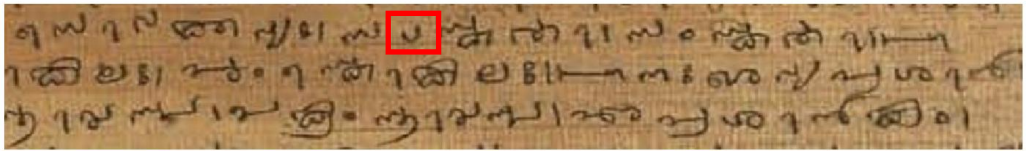
Here pluta is by 3 in transliteration. The pluta is seen in the phrases de3vadatta । devadā3tta । devadattā3 ॥. Other instances of usage of pluta.





5.2 ANUNASIKA

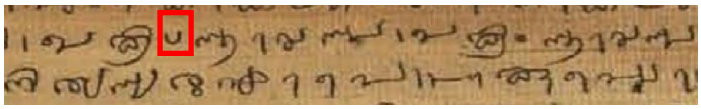
Anunasika indicates nasalization of a vowel or a semivowel. It is widely used in Tigalari. The function is similar to candrabindu of other scripts. It is different from Anusvara which indicates pure nasal as seen in below folios.

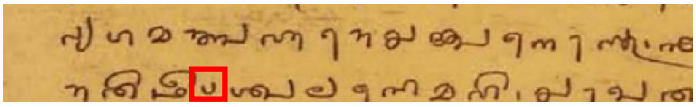


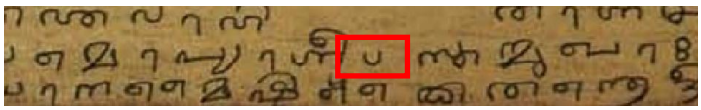
Corresponding text in Devanagari clarifies its use.

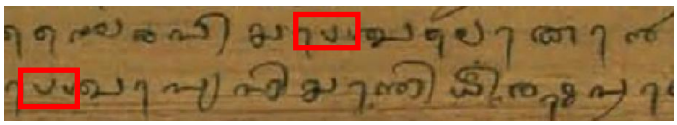
पदान्तस्य रेफस्य विसर्गः स्यात् । इति प्राप्ते—*संपुंकानां सो
वक्तव्यः । सँस्स्कर्ता संस्स्कर्त्ता ॥ पुमः खय्यम्परे ।८।३।६।

Other examples of Anunasika



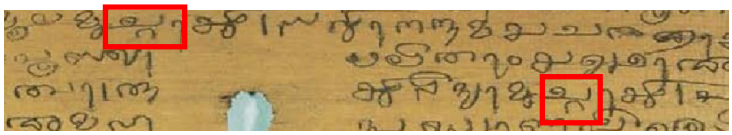


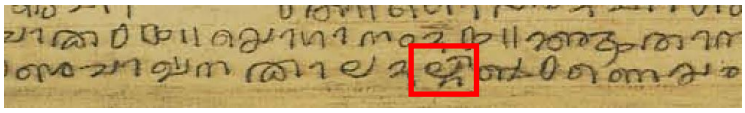




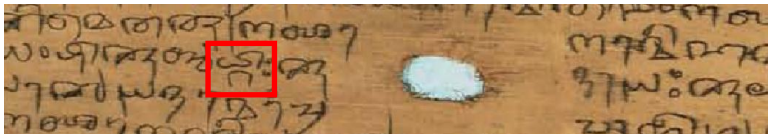
5.3 GEMINATION MARK

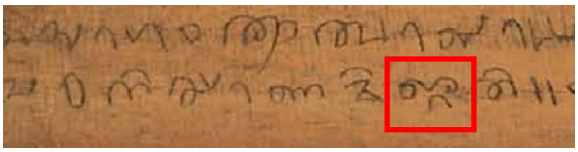
A gemination mark is also sometimes used in Tigalari. It denotes doubling of the consonant. It is represented by ◌ന. It is similar to GURMUKHI ADDAK and KHOJKI SIGN SHADDA. It may be encoded on same lines. It can be seen in the words vidyāmuccārya and kāladalli





It is also be used for doubling of non-aspirated sound followed by aspirated of aspirated consonant. In such cases gemination mark is written below the aspirated consonant as seen below in buddhiḥ and niryāṇamicchati.



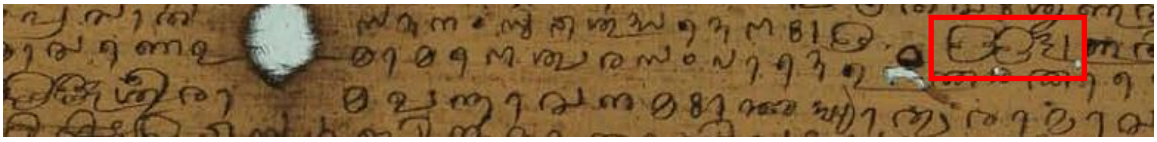


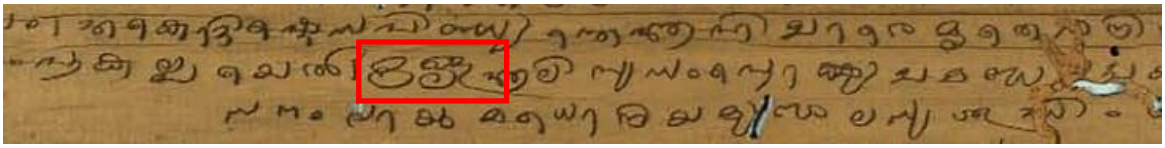
5.4 DOUBLE END OF TEXT

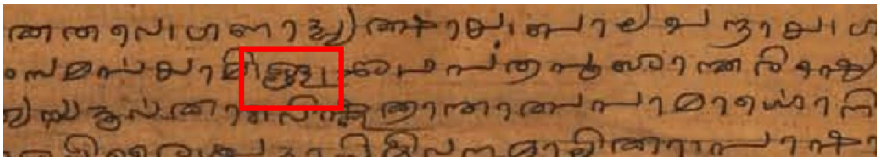
The proposal recommends the name PUSHPIKA which means flower for 113D2. It is mainly used in end of texts and sections. The symbol does not seem to be derived from Flower. PUSHPIKA generally refers to A8F8 ꣸ DEVANAGARI SIGN PUSHPIKA which also used in Tigalari. Instead of the name PUSHPIKA we suggest the name of like PUNCTUATION END OF TEXT or END OF TEXT MARK.

Section 8.2 of proposal says *"In this interval (pause), one could also choose to chant Om (from which Pushpika symbol seems to be derived) 23 or Shrii."* .The annotation says *also used indicate Om and Period*. Our opinion is that this character is not derived from Om and is not used to indicate Om.

In addition to 113D2, a double end of text  is also used in Tigalari. We recommend its encoding it along with 113D2. The folios given below shows the double end of text.

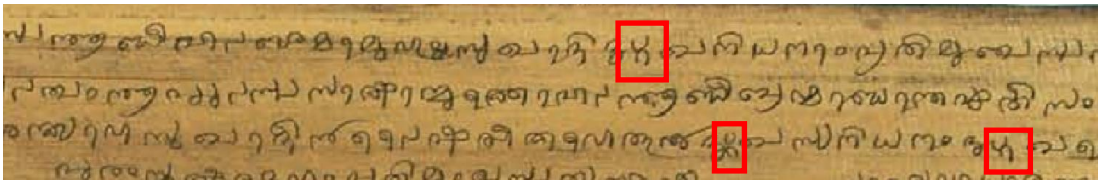


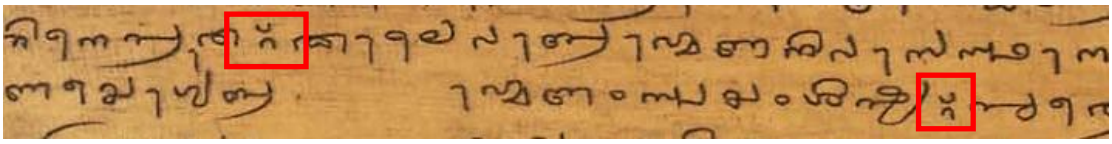




6 Script Extensions

Kannada digits, Devanagari Dandas & Pushpika, Indic fractions and Ardhavisarga are used in Tigalari.





The above folios show Ardhavisarga used for jihvamuliya and upadhmaniya. It should be represented using 1CF2 × VEDIC SIGN ARDHAVISARGA.

Fraction signs occur in Tigalari sources (see page 23 of L2/15-243). These may be represented using characters already encoded in the 'Common Indic Number Forms' block. A8F8 DEVANAGARI SIGN PUSHPIKA is employed in Tigalari script (see page 2 of L2/17-098). These characters should be specified as script extensions for Tigalari.

7 Note on the some of the attestations used in the proposal

As noted in footnote 11 and Figure 45 some images produced by Tulu academy and others contain several errors. Figures 29, 30, 42, 43, 44 and 45 are based on Burnell's chart. Hence they carry same errors in all these figures.

Figure 32 has lot of Malayalam influences in the letter forms. For instance the characters EE, AI,OO, JA, JHA,NYA,TTA ,BHA, VOWEL SIGNS EE and OO appears to be taken from Malayalam.

Figure 42. The proposal says *Additional Tigalari vowels E and O as proposed by Dr Devarakonda Reddy. The need for these two characters have been clearly stated here.* But we observe Additional Tigalari vowels E and O are not proposed here. The glyphs are taken from Burnell's chart. The letters ḷ, ḹ, ē, ai are misread as e, ē, ai and o.

The above figures can be analysed as charts showing use of script,but cannot be considred as genuine attestations or reference.

Figure 38 appears to be tabulated by proposal author. This also contains some incorrect forms. For example kṭa , tkha, śta, śṭṭa are not wriiten as ligatures but as stack. Bh does not have ligated form of Virama. Vowel signs U and UU are used in place for Vocalic R, RR for kr̥, kr̥̄, bhr̥ bhr̥̄ etc. We request to modify Figure 38 to correct forms.

8 Consonant clusters

Section 5.5 Ligatures of proposal explains formation of consonant clusters. The document says

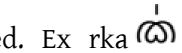
"It is therefore hard to come to a consensus and identify a definitive conjunct/ligature set or form for this script. Further study into this subject is required. As of now, the alternate forms of conjuncts can be handled using the opentype feature—stylistic sets."

According to our understanding its formations are similar to Grantha and pre-reformed Malayalam. We conclude following rules after studying manuscripts.

1. In general consonant clusters are preferably written as ligatures like kka  .

2. If a particular consonant cluster has no ligated form, or if the writer does not wish to form a ligature, then consonant clusters form stacks. Ex ppa  . Both ligatures and Stacks are equivalent as in case of jja which can be as both ligature and stack.

3. The post base ya, ra, la and va have special rendering which is explained in page 15 of the proposal.

4. For RA+Virama+Consonant repha/arkavottu is used. Ex rka  . In addition to repha some consonants like ya and va have special form which is not mentioned in proposal. Rya can be written either as  or  .Rva can be written either as  or  . These forms are also seen in Malayalam mss like  and  and Grantha mss like 

5. Unlike Kannada where anusvara is preferred for combination of nasal with its respective consonant of the varga. In Tigalari consonant cluster is preferred over anusvara. The ligatures of nasal and its respective consonant of KA, CA and PA varga have special forms that are unrelated to their parent shapes.

The clusters ṅka and ṅga may also be written as  and , but their instances are very rare.

6. Consonant clusters involving three or more consonants are may be written as ligatures, stacks or combinations of all above forms depending on the nature of consonants. Maximum of three elements in stack is attested. Due to use of large number of ligatures, three-level stack itself is rare in Tigalari. Instances requiring more than three elements in a stack are not found in Sanskrit texts of Tigalari.

Other comments on consonant clusters

1. The kva  on page 15 requires some modification. It should be written as  .

2. The triangular form of gemination of consonant of CA, BA, YA and VA like  are not commonly seen in Standard Tigalari. If they exist, the appearance of these forms in some manuscripts is likely due to influence of Malayalam script.

3. The conjuncts similar to ṅka  should not be analyzed as NA+VIRAMA+KA. Correct the sequence of ṅkta in page 15 of the proposal.

4. The script ad hoc report raises the *question "Section 5.5 mentions that conjuncts can be formed horizontally or vertically. Do the two orientations need to be differentiated? If so, sequences with ZWJ or ZWNJ could be used, as in Malayalam."*

As said earlier both ligatures and Stacks are equivalent in Tigalari. To handle Virama and consonant conjuncts the Virama model used for Grantha can also be applied to Tigalari.

5. Modifying conjunct behavior

The proposal recommends the following.

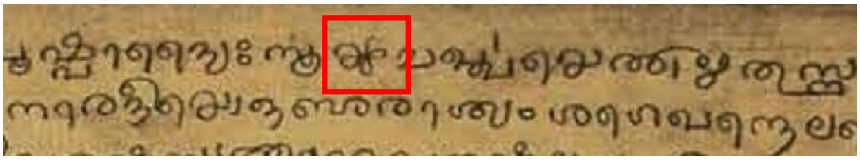
ർക്ക	=	ര	+	‍	+	്	+	ക
RKA		RA		ZWNJ		virama		KA

But this is different from the model adapted in Grantha, Devanagari etc. The ZWNJ should be used after the Virama. Its details are given below:

Default conjunct formation may be modified using ‍ U+200C ZERO WIDTH NON-JOINER:

C_1, ് VIRAMA, ‍ ZWNJ, C_2

ർക്ക	=	ര	+	്	+	‍	+	ക
RKA		RA		virama		ZWNJ		KA



Folios showing the use of ്യ rya and ്വ rva.

9 Other comments

1. Names of svarita and anudatta. The proposal suggests the names as TIGALARI SVARITA and TIGALARI ANUDATTA. As in Vedic extensions and Devanagari TIGALARI SIGN/TONE SVARITA and TIGALARI SIGN/ TONE ANUDATTA is recommended to be used.

2. The proposal questions *"Two part vowels / reordering / encoding: U+0D57 (Malayalam AU length mark) is provided as an encoding for the right side of the two part vowel U+0D4C (Malayalam vowel sign AU). I'm proposing to follow the same for Tigalari. Is this a good practice? In the context of Tigalari, this length mark is never used alone. (It might be a redundant practice—as advised by Cibu CJ)"*

In Grantha and Malayalam AU length mark was encoded due its use in modern texts, whereas vowel sign AU was generally used in older texts. In case of Tigalari only vowel sign AU is used. Encoding of AU length mark may not required as it is never used alone.

4. NNNA is used in Tamil and Malayalam languages. As Tigalari is not used to write these languages reserving space for NNA is inessential. Also, there are several gaps in the Code chart and Code points like 113A7, 113B8, 113C9, 113CC, 113CD, 113D3. It is appropriate to keep the characters continuous instead of having gaps in the middle.

5. Page 5 says *"Total number of characters: 63"*. But 90 characters are proposed. Update this section.

6. Page 12 says *"Tigalari has two characters that represent the Dravidian sounds in Tigalari: LLLA and RRA (Shakata repha). These two characters are rare and are mostly found in Kannada (language) Manuscripts"*. It should be noted here that LLLA is found in Tulu and RRA in Kannada manuscripts.

The following figures provides Character inventory (Aksharamala) seen in many manuscripts.