Proposal to encode 0C5A TELUGU LETTER RRRA

Shriramana Sharma, Suresh Kolichala, Nagarjuna Venna, Vinodh Rajan

jamadagni, suresh.kolichala, vnagarjuna and vinodh.vinodh: *-at-gmail.com

2012-Jan-18

§1. Character to be encoded

ఴ

0C5A   TELUGU LETTER RRRA

§2. Background

Apart from the archaic Telugu LLLA, which is graphically related to the old Telugu RRA, there is also evidence for another related character in inscriptions written as ఴ or ఴ. The Telugu epigraphical publication *Telugu Śāsanālu* (ref 1) speaks on this (on p 4) as follows:

ఇంక్ోక యఉరమ(ఴ)ేక ద* ఇది దంత్య 'థ' కారముగాదు పై చెప్పిన 'ట' కంఠు రెండు మూడు శాఉ్దులు పూర్వమే యిది కాసనములందు నై తము ఉపయోగమునుండి తొలగెను. నన్నయఖట్టు నాటికే యిది ఖాషలో లేదు. సంయుక్తాఉరములలో 'న' 'ణ' తో కలిసి యుంఢెడిది న్ఴ, ణ్ఴ, మూస్ఴ: చాఢ్ఴా ఢ్లాకు, చ ద (మూడు, చాఢ్ఴాలురకు, చెడు) మున్నగు పదములలో కాననగును. కలమ్ఴ్ల కాసనములో 'ధనంజయుదు రేనాఢ్ఴ, ఏశ్న్ చిఉంఖూరి రేవణకాలు....' అని దు, ఴులు రెండు [వాయబడి యుండు టను బట్ట(ఴ)ఉలు ఖిన్నాఉరములనవలెను అఴ్లే రామేశ్వరము ([ఛొద్దుటూరు) కాసన మొకదానిలో ఇ, ఆ. ఴ, డ, శ, లు స్పఴ్టముగా [వాయబడి యున్నవి గనుక నవి యన్నియ ఖిన్నాఉరములే యగును. పై చెప్పిన 'ధనంజయుదు'

Another letter ఴ (or ఴ) is (in inscriptions similar to but) not the dental థ. This went entirely out of use from inscriptions two or three centuries before ఴ became outmoded, as explained above. This is also not present (as a distinct character) in the language of Nannayya Bhaṭṭa. It is seen to form

1

consonant clusters with na and ṇa. … In the inscriptions of Kalamaḷḷa … ꖉ and ꖊ are both distinctly seen to be used. Further in the Rāmeśvara (Produṭūru) inscription, ꖊ, ꖊ, ꖉ, ꖌ and ꖎ are all seen to be clearly used, indicating that they are all distinct characters.

Thus ꖉ/ꖊ is a distinct character worthy of encoding. It is however clearly phonologically related to ꖊ RRA/ṟa as seen in its usage in words like mūnṟu (old Dravidian, "three").

It has been suggested (ref 2) that in the early Telugu language, the alveolar stop ṯ/ḏ was phonemically distinct from the alveolar trill ṟ (at least for a brief period of time):

A careful analysis of the reflexes of Proto-Dravidian *ṯ in Central Dravidian languages shows that there is sufficient evidence for reconstructing not one, but two separate phonemes in its place — a stop (*ṯ) and a resonant (*ṟ).

It is possible that the representation ꖉ/ꖊ in contrast with its relative ꖋ/ꖊ was used to indicate the alveolar stop in contrast to the alveolar trill.

It should be noted that ꖉ/ꖊ is also related to ꖌ DDA, as it is the character which it mostly degrades to or merges with later on, as in mūḍu (modern Telugu, "three").

§3. Attestation

The attestation is of two kinds. One is from epigraphs, as for example:

(ref 3, inscription #15, previously mentioned as that of Rāmeśvara/Produṭūru; occurrences of ꖉ/ꖊ are circled in red, with one of the contrasting ꖊ in pink)

However this only attests the character as part of the old Telugu(-Kannada) script which probably merits its own separate encoding. The actual attestation for the present proposal is hence from modern epigraphical publications which transcribe such epigraphs into modern Telugu writing, as in the transcript of the above inscription:
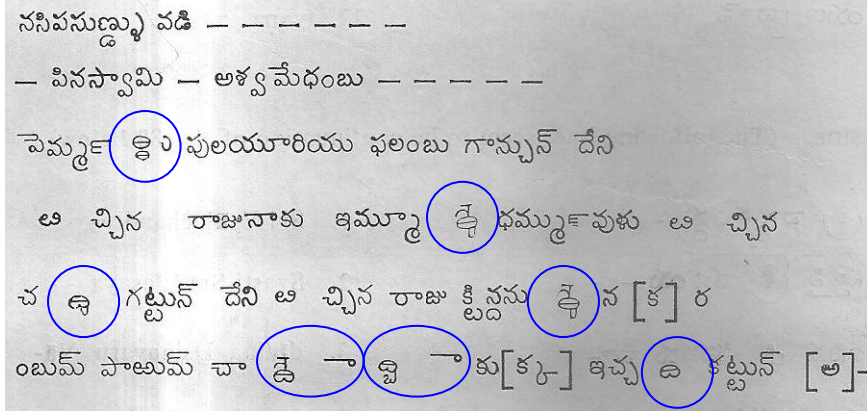
One however observes that this publication has resorted to a mixture of merely copying the old epigraphical glyph by hand and only occasionally but even then only brokenly combining it with modern Telugu vowel signs. The obvious reason is the lack of printer's glyphs of this old character, which was not used for a millennium.

Another publication has presented this character as ⊖ where it occurs with the inherent vowel 'a' intact or as a sub-base form (i.e. where no vowel signs are attached in the modern script) and elsewhere replaced it by DA-2 i.e. దే. It says in a footnote (ref 1 p 4):

✱   తెలుగు లిపిలో ఇపుడు కనబడని ఈ అక్షరము, ఇకముందు, ఆకారము తప్ప తక్కిన అచ్చులు కలపవలసివచ్చినపుడు ద అని ము[దింపబడ గలదు;   అటులకాక ఆకారాంత అక్షరముగ వచ్చినపుడును సంయుక్తమున వత్తుగ వచ్చినపుడును ⊖ అని ము[దింప బడగలదు. పాఠకులు దీనిని గమనింపవలెను.

> *This letter which is not seen in the Telugu script today is printed as DA-2 దే*
>
> *when it occurs in combination with vowels other than 'a'. On the other hand,*
>
> *it is printed as ⊖ when it occurs in combination with 'a' or when it occurs as*
>
> *the vattu (sub-base conjoining form). Readers should keep this in mind.*

Since DA with all the vowel signs was available in the printer's repertoire, by placing a two ఎ above it దే was created as a replacement for ⊖/ఆ with the ఎ serving as a nukta-like consonant modifier as seen in the case of ౘ TSA and ౙ DZA.

However, with today's font technology there is no difficulty in proper typesetting in the case of the vowel signs:

<p align="center">ఎ ఔా ఏ ఈ ఉ ఊా ఎ ఏ ఐ ఓ ఔో ఔౌ</p>

... or in producing conjoining forms, so as to render a more appropriate transcription, as in the following rendering of the Rāmeśvara epigraph:

నసిపసుణ్డ్క వడి ...

... పినస్వామి ... అశ్వమేధంబు ...

పెమ్మ ఱఁడ్డు పులయూరియు ఫలంబు గాన్చున్ దేని

ఇచ్చిన రాజనాకు ఇమ్మూఁన్డు ధర్మఁవులు ఇచ్చిన

ఫఁడు గఁట్టున్ దేని ఇచ్చిన రాజ క్తిదన్ న్డున[క]ర

ంబుం పాఁయుం చాన్డఁడ్బకు[క్క] ఇచ్చ ఒకఁట్టున్ [అ]

Therefore there is no need to support the DA-2 written forms as they are merely a stopgap measure for lack of glyphs. It is sufficient to encode the attested character ⊖/ఴ.

§4. Choice of representative glyph

It has been noted that the character under discussion has two forms:

<p align="center">⊖ ఴ</p>

The form on the left is the older one after the simpler glyphs descended from Brahmi, and the form on the right is the more cursive one of later writings, as seen in the evolution chart on the right: (ref 4 p 80)

Of course, one may possibly rarely see both forms in the same writing sample but there is no doubt that they are equivalent presentations of the same character.

Roman transliteration	...	r̠	ẕ
...		...	
Śālamkāyana Telugu–Kannaḍa script 5th cent CE		⊕	Φ
Pallava Tamil Grantha 7th cent CE		⊕	⚇
Telugu–Kannaḍa 7th cent AD	...	⊖ఴ	∞
Eastern Cāḷukyan 10th cent CE		౪ఴ	౬
Rājarājanarendra Telugu–Kannaḍa 11th cent CE		౪ఴ	౬
Kākati Gaṇapati 13th cent CE		ఴ	

This proposal recommends that the form on the right be used as the representative glyph because the proposal is as part of the modern Telugu script and this is the form chronologically more recent and hence stylistically coherent with modern Telugu writing.

§5. Collation

As this character was never part of literary Telugu, it has never been collated. As such it is advisable to not disturb the existing order of collation and place this character at the end. However, we are also submitting a separate proposal for Telugu LLLA. which is also absent from literary Telugu. RRRA is however unique to old Telugu and not found in any other languages unlike LLLA, and hence it is advisable to place it *after* LLLA to isolate it. Hence:

YA < RA < LA < VA < SHA < SSA < SA < HA < LLA < K·SSA < RRA < LLLA < RRRA

§6. Unicode Character Properties

```
0C5A;TELUGU LETTER RRRA;Lo;0;L;;;;;;N;;;;;
```

The codepoint 0C5A is chosen to place this along with the other historic Telugu characters TSA and DZA at 0C58 and 0C59. The name RRRA is chosen as this character is most closely related to the existing RRA character and since 'RRRA' has not been used so far in Indic.

Other properties like linebreaking are as for standard Indic consonants.

§7. References

1) Telugu Śāsanālu, Ed. Dr P V Parabrahma Shastri, Andhra Pradesh Sahitya Akademi, Hyderabad, 1975

2) "The three-way phonological distinction in Dravidian coronal consonants", George Starostin, Aspects of Comparative Linguistics, v. 4. Moscow: RSUH Publishers, 2009, pp. 243-260

3) Inscriptions of Andhra Pradesh: Cuddapah District - I, Ed. Dr P V Parabrahma Shastri, Dept of Archaeology and Museums, Govt of Andhra Pradesh, 1977

4) The Dravidian Languages, Bhadriraju Krishnamurthy, Cambridge University Press, Cambridge, 1st South Asian Edition, 2003, ISBN: 978-0-521-77111-5

§8. Official Proposal Summary Form

(Based on N3902-F)

A. Administrative

1. Title
Proposal to encode 0C5A Telugu Letter RRRA
2. Requester's name
Shriramana Sharma, Suresh Kolichala, Nagarjuna Venna, Vinodh Rajan
3. Requester type (Member body/Liaison/Individual contribution)
Individual contribution
4. Submission date
2012-Jan-18
5. Requester's reference (if applicable)
6. Choose one of the following: This is a complete proposal (or) More information will be provided later
This is a complete proposal.

B. Technical – General

1. Choose one of the following:
1a. This proposal is for a new script (set of characters), Proposed name of script
No
1b. The proposal is for addition of character(s) to an existing block, Name of the existing block
Yes, Telugu
2. Number of characters in proposal
1 (one)
3. Proposed category
Category B1, specialized small (for this character)
4. Is a repertoire including character names provided?
Yes
4a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?
Yes
4b. Are the character shapes attached in a legible form suitable for review?
Yes
5. Fonts related:
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?
Shriramana Sharma
b. Identify the party granting a license for use of the font by the editors (include address, e-mail etc.)
Shriramana Sharma, based on the GPL-ed Pothana font © K Desikachary
6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?
Yes
6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?
Yes
7. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?
Yes
8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script.
See detailed proposal.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.
No
2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?
Yes
2b. If YES, with whom?

Dr Bhadriraju Krishnamurti, retd from Osmania University, Hyderabad. Dr N S Ramachandra Murthy, Government Oriental Manuscripts Library & Research Centre, Hyderabad.

2c. If YES, available relevant documents

None specifically. The matter was discussed in person and via email/phone.

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

Epigraphists who desire to transcribe ancient Telugu inscriptions etc into the modern script

4a. The context of use for the proposed characters (type of use; common or rare)

Rare

4b. Reference

See detailed proposal.

5a. Are the proposed characters in current use by the user community?

The character is used by epigraphists.

5b. If YES, where?

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

Yes

6b. If YES, is a rationale provided?

It belongs in the Telugu block which is in the BMP.

6c. If YES, reference

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

Only one character is proposed. It is to be placed contiguous with other historic characters.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

No.

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences?

No

11b. If YES, is a rationale for such use provided?

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?

13c. If YES, reference:

–o–o–o–