# Proposal to encode 0C34 TELUGU LETTER LLLA

Shriramana Sharma, Suresh Kolichala, Nagarjuna Venna, Vinodh Rajan

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

2012-Jan-17

## §1. Character to be encoded

ఴ

0C34    TELUGU LETTER LLLA

## §2. Background

LLLA is the Unicode term for the native form *in South Indian scripts* of the voiced retroflex approximant. It is an ill-informed notion that it is limited or unique to one particular language or script. Such written forms are already known to exist in Tamil, Malayalam and Kannada and characters for the same are already encoded in Unicode. Evidence for the native presence of corresponding written forms of LLLA in other South Indian scripts, which may or may not be identical to characters already encoded, is forthcoming. This document provides evidence for Telugu LLLA identical to Kannada LLLA and requests its encoding. Similar evidence is to be expected for other South Indian scripts in future as well.

## §3. Attestation

Currently the Telugu script does not use LLLA. However historic usage before the 10th century is undeniably attested. When the literary Telugu language was effectively standardized by the time of the Mahābhārata of Nannayya in the 11th century, this letter had disappeared from use. It is clear that this disappeared from the Telugu language long before it disappeared from Kannada. However, it was definitely used as part of the Telugu script and language and is considered a Telugu character by Telugu epigraphist scholars.

Even though the pre-10th-century Telugu writing had much (more) in common with the Kannada writing of those days and is perhaps better analysed as a single proto-Telugu-Kannada script, and hence the older written form would not be a part of the modern Telugu

script currently encoded in the block 0C00-0C7F, the modern written form used when transcribing the older form is most definitely a part of the modern Telugu script and hence should be encoded as part of it.

The authoritative text *Telugu Śāsanālu* (ref 1) speaks on this (on p 3) as follows:

సుమారు క్రీస్తు పదవశతాబ్ద్యంతము వఱకు అనగా నన్నయ భట్టారకుడు గ్రాంథిక భాషను శాసించుషఱకు శాసనములందు ('ఴ') అను రూపమున వ్రాయబడు ఆధరముంఙెడిది. ఇది బండి 'ఋ' లోని అడ్డు గీటును తొలగించి వ్రాసినట్లు శాసనములందు కనపించును. దీనిని గూర్చి కీర్తిశేషులు జయంతి రామయ్య పంతులుగారు, వేటూరి ప్రభాకర శాస్త్రిగారు, మల్లంపల్లి సోమశేఖర శర్మ గారు మున్నగు పలువురు పరిశోధనలు జరిపి అది మనకిప్పుడు వాడుకలో లేని వేఱొక ఆఴరమని నిర్ధారణము గావించిరి. అది క్రమముగా కొన్నిచోట్ల 'డ'

A free (but true) translation is provided:

> *Until about the tenth century after Christ, before the time Nannayya Bhaṭṭākaraka defined the literary language, there existed in epigraphs the letter written in the form of: ఴ. It is observed in these epigraphs that this letter was written by removing the horizontal stroke of ఋ. Many scholars … have researched this written form and concluded that it is a letter which is lost from usage to us today.*

The text further discusses the usage pattern of the letter and concludes (on p 4):

కనుక 'ఴ' అనేది తెలుగు భాషకి చెందిన ఆఴరమే;

*Therefore ఴ is indeed a letter belonging to Telugu.*

It is to be noted that the relationship of the loss of the horizontal stroke in LLLA compared to RRA has been maintained from old proto-Telugu-Kannaḍa into modern Telugu, whence we have in the old writing: ⊕ RRA and ⊙ LLLA whereas in the modern script ఋ RRA and ఴ LLLA. This is seen in the Telugu script evolution chart (ref 2 p 80):

| Roman transliteration | … | r̠ | ẕ |
|---|---|---|---|
| … | | … | |
| Śālamkāyana Telugu–Kannaḍa script 5th cent CE | | ⊕ | ⊙ |
| Pallava Tamil Grantha 7th cent CE | | 𑀭 | 𑀴 |
| Telugu–Kannaḍa 7th cent AD | … | ఴఴ | ౧౦ |
| Eastern Cāḷukyan 10th cent CE | | ఆఴ | ౮ |
| Rājarājanarendra Telugu–Kannaḍa 11th cent CE | | టఴ | ౮ |
| Kākati Gaṇapati 13th cent CE | | ఴ | |

Further samples from *Telugu Śāsanālu* (ref 1):

ఈయతర మొకటి పూర్వము తెలుగు ఖాషలో గలదని నిశ్చయముగ కాసనము లను బట్టి తెలియుచున్నది. అది 'చోటి' పదములలో 'చోడ' లేక 'చోళ' అనియు ; 'నొంటంబ' పదములో 'నోళంబ' అనియు, తెందులూరు అనుచోట దెందులూరు గాను, క్టిన్ద అనునది క్రిన్ద (క్రింద) గాను 'క్రొట్టె' అనువదము 'క్రొచ్చె' ; 'వ్రటచె' అమునది 'వ్రచ్చె' అనియు మార్పునొందెను. ఈ యతరము

క్టిన్దన్ = క్రిందన్          కొట్రొచె = క్రొచ్చె (శిలపై చెక్కుట) ....పటికి = వ్రక్కి

From South Indian Inscriptions Vol X (ref 3 inscription #29):

22 దీని నటెసినవాణ్డు శ్రీవ[వ్వఁ]తమును ఖారనా

28 .   .   [వి]లలాను క్రబాహ్మణులను అటెసిన

An estampage of an inscription (color-inverted for better clarity) in the old Telugu(-Kannada) script with its transcription into the modern Telugu script (ref 4 inscription #3):

స్వస్తి శ్రీ చో టి మహా
రాజు ల్లేఖన్ ఎఱగల్
దుగరాజుల్ ఇచ్చిన ప-
న్నన కొచ్చియ పాటి [రే]-
వ సమ్మార్గఔకిన్ తేని టి _
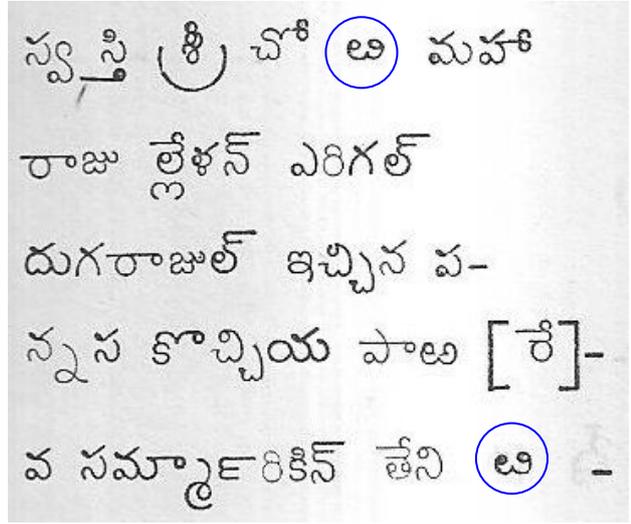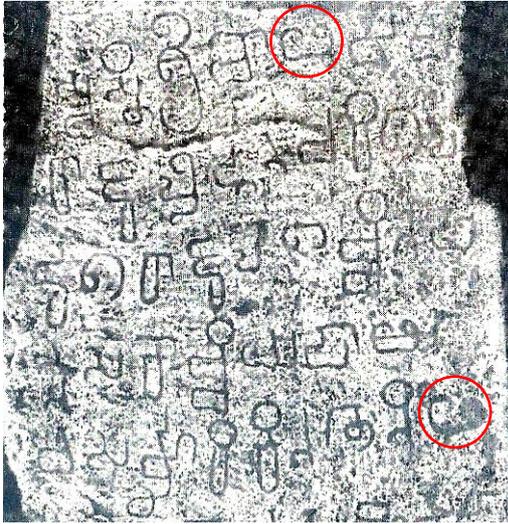
It is worthy noting that the GOI TDIL document on Telugu (ref 5) has mentioned this character (on p 87) even though no effort was later made to encode it:

**Special/obsolete graphemes**

obsolete:

| ఋ | ఌ | ౡ | ఱ | ఴ |
|----|----|----|----|----|
| r̄ | ļ | ī | ṛa | z̧ a |

# §4. 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 RRRA. 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

# §5. Unicode Character Properties

`0C34;TELUGU LETTER LLLA;Lo;0;L;;;;;N;;;;;`

The codepoint 0C34 is chosen as it is already reserved for this in the ISCII pattern.

Other properties like linebreaking are as for standard Indic consonants.

# §6. References

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

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

3) South Indian Inscriptions Vol X, Ed. J Ramayya Pantulu, Delhi, 1948

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

5) Vishvabharat Issue 5, 2002 Apr, GOI TDIL Programme, http://tdil.mit.gov.in/pdf/Vishwabharat/tdil-april-2002.zip from http://tdil.mit.gov.in/Publications/Vishwabharat.aspx

# §7. Official Proposal Summary Form

(Based on N3902-F)

## A. Administrative

1. Title
**Proposal to encode 0C34 Telugu Letter LLLA**
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-17**
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. Dr Vijayavenugopal and Dr Sathyanarayanan, epigraphists, EFEO, Pondicherry.**

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 should be placed in the codepoint reserved for it.**

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?

**Yes. It is glyphically (and in sound value) identical to 0CDE KANNADA LETTER LLLA.**

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

**The script property would be different, i.e. script=telugu.**

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–