

Proposal to encode the TELUGU SIGN COMBINING ANUSVARA ABOVE

Srinidhi A and Sridatta A

Tumakuru, India

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

October 20, 2016

1 Introduction

This is a proposal to encode new character in the Telugu block. This is a second anusvara as attested in Prakrit texts to indicate linguistic anusvara.

Glyph Code Character name



0C04 TELUGU SIGN COMBINING ANUSVARA ABOVE

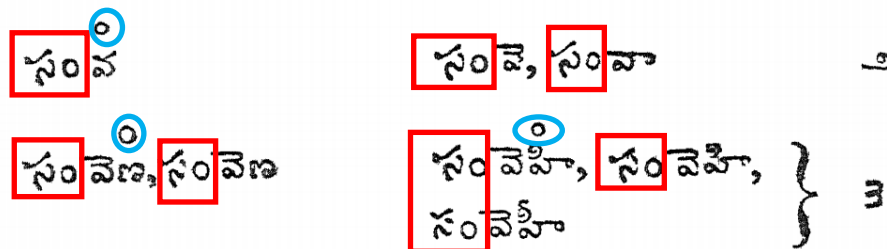
2 Background

The documents L2/13-061 and L2/14-003 proposed combining anusvara above for Grantha and Malayalam respectively. These were based upon contrastive use of regular anusvara and combining anusvara above in Prakrit language manuscripts and texts. The two characters were subsequently accepted for encoding. Now, a similar character is documented in the related Telugu script of South India.

3 Description

In Prakrit orthography, the regular anusvara sign is placed to the right is used as a gemination marker and the linguistic anusvara is marked by the anusvara sign placed above the consonant.

The following attestation explains the contrastive use of anusvara.



(Here the regular anusvara is denoted by red box, whereas the anusvara above is denoted by blue circle.)

sa^vvam̐
savvēṇam̐, savvēṇa

savvē, savvā
savvēhiṃ, savvēhi,
savvēhī

Here, the regular anusvara in words marked in red indicate the gemination. Whereas, the letters denoted by anusvara above marked in blue indicate the linguistic anusvara.

4 Character details

a. Character properties are given below:

0C04; TELUGU SIGN COMBINING ANUSVARA ABOVE; Mn; 0; NSM; ; ; ; N; ; ; ;

b. The following annotation is to be added to the names list:

0C02 TELUGU SIGN ANUSVARA

* Used in Prakrit language texts to indicate gemination of the following consonant

0C04 TELUGU SIGN COMBINING ANUSVARA ABOVE

* Used in Prakrit language texts for the linguistic anusvara

c. Collation and other properties are would be similar to that of combining anusvara above of Grantha and Malayalam.

The attestations are obtained from the book *Prākṛta śabdamanjarī* by Krishnaraya Pandita printed at Sarasvatilaya Press, Chennai, 1871. Also, Sanskrit texts in the same book use the regular anusvara to the right denoting the linguistic anusvara. The book is full of such usages and may be obtained from link. <http://www.dli.ernet.in/handle/2015/495555>

Hence, it is recommended to encode the above character based on the attestations of distinctive usage in Prakrit texts in Telugu script.

5 Attestations

Here the regular anusvara is denoted by red box, whereas the anusvara above is denoted by blue circle.

Figure 1. from page no.1

శ్రీరస్తు.

—ॐ—

సిరిజాణక్రవంశోహాఅణమో
 శ్రీజానకీవల్లభాయనమః

పాడజస్సంబమజరీ
 ప్రాకృతశబ్దమంజరీ
 అకారాంతఃపుంలింగోరామశబ్దః

ఏకవచనమ్.	బహువచనమ్	
రామో	రామా	౧
దేరామ్-దెరామో	దెరామా	౨
రామః	రామై-రామా	౩
రామణి-రామేణి	{ రామేహి-రామేహి రామేహీ }	౪
రామంస్-రామాః	రామాణి-రామాణి	౫
{ రామాహితో-రా మంతో-రామాదో- రామాదు-రామా ఓ-రామాడ - రా మా-రామాహి }	{ రామాహితో-రామేహి తో-రామానుతో-రామె సతో-రామాహి-రామె హి-రామంతో-రామా దో-రామాదు-రామాఓ రామాడ }	౬

Figure 2. from page no.14

౧౪ ప్రాకృతశబ్దమంజర్యాం

హలంతపులింగశబ్దాః

గోధుక్-ఆనడుహ-ఇత్యాదిహలంతధాతుశబ్దానాం
 హలోగిత్యగాగమె-గోధు-హలూఇత్యాదిరామవత్ *
 సుదిక్-శబ్దస్యఅధాతుత్వాంతహలోశధూదీభి- ఆంత్య
 హలోలోపెసుదిఇత్యాదిహరివత్ ॥

రెఫాంతఃపులింగశ్చత్వశబ్దః

చడరో-చంతారో-చంతారి-౧-౨-చడోహి - చడి
 హి-చడోహి ౩-చడణ్ణి-చడణ్ణి ౪- చడోహితో -చడం
 తో,చడోదో-చడోదు,చడోఓ-చడోఊ-చడోనుతో-చడోహి
 తో౫-చడణ్ణి-చడణ్ణి ౬-చడోసు,చడోసు,చడోసు,చడోసు

మకారాంతఃపులింగశ్చింత్యశబ్దః

కో	కె-కా	౧
కం	కా	౨
కీణా-కేణ-కేణ	కేహి-కేహి-కేహి	౩
కాస-కంస	కేసి-కాబి-కాణి	౪
కంణా-కేస-కేహో	కాహితో-కేహితో-కా	}౫
కాహితో-కంతో	సుతో-కేసుతో-కాహి	
కాదో-కాదు-కాఓ	కేహి-కంతో-కాదో-కా	
కాఊ-కా-కాహి	దు-కాఓ-కాఊ	

Figure 3. from page no.16

౧౬

ప్రాకృతశబ్దమంజర్యాం

చతుర్థీ-వల్ వస్త్రీ

అంస-ఇమంస-ఇమమి } ఏసు-ఏసు-ఇమేసు-ఇమేసు & ఇహ

నకారా న్తఃపుంలింగోరాజశబ్దః

రాఆ-రాఓ-రాఆ } రాఆ-రాఆణ-రాఇణో- }
 ణో-రాఆణ } రాఆణో- రాఆణాణో- } ౧

దెరాఅ- దెరాఓ- } దెరాఆ-దెరాఆణా, దెరా }
 దెరాఆణ- దెరాఆ } ఇణో-దెరాఆణో- దెరా }
 ణో } ఆణాణో } ౨

రాఅ-రాఇణ-రా } రాఆ-రాప-రాఆణ- రాఆ }
 ఆణ } ణే-కాఇణో-రాఆణో-రా }
 ఆణాణో } ౩

రంణా- రాఇణా - } రాఘో-రాఘో-రా }
 రాపణ- రాపణ-రా } ఘో-రాపణో-రాపణో }
 ఆణాణ-రాఆణాణ } రాపణో-రాఆణాణో-రా }
 ఆణాణో-రాఆణాణో } ౪

రంణో- రాఇణో- } రాఘో-రాఘో-రాఘో }
 రాఆణో- రాఆణా } రాఘో-రాఆణాణ-రా }
 ణో-రాఅంస -రా } ఆణాణ }
 ఆణంస } ౫

ISO/IEC JTC 1/SC 2/WG 2

PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS

FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹.

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from <http://std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

A. Administrative

1. Title:	Proposal to encode the TELUGU SIGN COMBINING ANUSVARA ABOVE
2. Requester's name:	Srinidhi A and Sridatta A
3. Requester type (Member body/Liaison/Individual contribution):	Individual contribution
4. Submission date:	October 20, 2016
5. Requester's reference (if applicable):	N/A
6. Choose one of the following:	
This is a complete proposal:	Yes
(or) More information will be provided later:	

B. Technical – General

1. Choose one of the following:	
a. This proposal is for a new script (set of characters):	No
Proposed name of script:	N/A
b. The proposal is for addition of character(s) to an existing block:	Yes
Name of the existing block:	Telugu
2. Number of characters in proposal:	1

¹ Form number: N4502-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

3. Proposed category (select one from below - see section 2.2 of P&P document):

A-Contemporary	<input type="checkbox"/>	B.1-Specialized (small collection)	<input checked="" type="checkbox"/>	B.2-Specialized (large collection)	<input type="checkbox"/>
C-Major extinct	<input type="checkbox"/>	D-Attested extinct	<input type="checkbox"/>	E-Minor extinct	<input type="checkbox"/>
F-Archaic Hieroglyphic or Ideographic	<input type="checkbox"/>	G-Obscure or questionable usage symbols	<input type="checkbox"/>		<input type="checkbox"/>

4. Is a repertoire including character names provided?

Yes

a. If YES, are the names in accordance with the “character naming guidelines”

in Annex L of P&P document?

Yes

b. 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?

Not necessary. The glyph may be borrowed from Grantha or Malayalam

b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):

Not necessary

6. References:

a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

Yes

b. Are published examples of use (such as samples from newspapers, magazines, or other sources)

of proposed characters attached?

Yes

7. Special encoding issues:

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

See text of proposal

8. Additional Information:

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. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard

at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database (<http://www.unicode.org/reports/tr44/>) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?	No
If YES explain	N/A
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?	No
If YES, with whom?	N/A
If YES, available relevant documents:	N/A
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?	Yes
Reference:	<i>Used by scholars desiring to document old Telugu books and manuscripts.</i>
4. The context of use for the proposed characters (type of use; common or rare)	rare
Reference:	<i>See text of proposal</i>
5. Are the proposed characters in current use by the user community?	Yes
If YES, where? Reference:	<i>Used by scholars desiring to document old Telugu books and manuscripts</i>
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	Yes
If YES, is a rationale provided?	Yes
If YES, reference:	<i>It belongs to Telugu block which is in BMP</i>
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	N/A
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	No
If YES, is a rationale for its inclusion provided?	N/A

If YES, reference:	N/A	

9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?		No
If YES, is a rationale for its inclusion provided?		N/A
If YES, reference:	N/A	

10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character?		Yes
If YES, is a rationale for its inclusion provided?		Yes
If YES, reference:	<i>Not in Telugu script. Identical characters are accepted in Grantha script property would be different.</i>	The

11. Does the proposal include use of combining characters and/or use of composite sequences?		Yes
If YES, is a rationale for such use provided?		Yes
If YES, reference:	<i>The character is an Indic non-spacing mark.</i>	

Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?		N/A
If YES, reference:	N/A	

12. Does the proposal contain characters with any special properties such as control function or similar semantics?		No
If YES, describe in detail (include attachment if necessary)		N/A

13. Does the proposal contain any Ideographic compatibility characters?		No
If YES, are the equivalent corresponding unified ideographic characters identified?		N/A
If YES, reference:	N/A	
