

Telugu Script LGR Proposal

Introduction, Current Analysis and Next Steps

Prof. Uma Maheshwar Rao
Gangadhar Pandey

NBGP F2F Meeting, Colombo
15th December 2017



Agenda

1

Introduction to
Telugu Script

2

Classification of
Characters

3

Repertoire Analysis

4

Within & Cross-
Script Variants

5

WLE Rules

6

Current Status and
Next Steps for
Completion

Introduction to Telugu Script

- ⦿ **Population:** The Telugu languages is a Dravidian language spoken by about 75 million (ca. 2011) people mainly in the southern Indian states of Andhra Pradesh and Telangana where it is the official language.
- ⦿ **Geographical area:** It is also spoken in such neighboring states as Karnataka, Tamil Nadu, Orissa, Maharashtra and Chattisgarh, and is one of the 22 scheduled languages of India. There are also quite a few Telugu speakers in Canada, the USA, South Africa, Malaysia, Mauritius, Myanmar, Sri Lanka and Réunion.
- ⦿ **Languages written in Telugu script:** Telugu (EGIDS 2) , Gondi (EGIDS 5)

Classification of Characters: Consonants

The Telugu consonants have an implicit vowel /a/ included in them. As per the traditional classification they are categorized according to their phonetic properties.

Non	య	ర	ల	ళ	వ
Varga	శ	ష	స	-	హ

Varga	Unvoiced		Voiced		Nasal
	_Asp	+Asp	_Asp	+Asp	
Velar	క	ఖ	గ	ఘ	ఙ
Palatal	చ	ఛ	జ	ఝ	ఞ
Retroflex	ట	ఠ	డ	ఢ	ణ
Dental	త	థ	ద	ధ	న
Bi-labial	ప	ఫ	బ	భ	మ

Table 2: Varga classification of consonants

Classification of Characters: Vowels

Repertoire Included - 1

Sr. No	Unicode Code Point	Glyph	Character Name	Unicode General Category (gc)	Indic Syllabic Category	Reference	Y/N
1	0C00	◌̣	TELUGU SIGN combining CHANDRABINDU	Mn	combining CHANDRABINDU		
2	0C01	◌̣̣	TELUGU SIGN CHANDRABINDU OR ARASUNNA	Mc	ARA-SUNNA		
3	0C02	◌̣̣̣	TELUGU SIGN ANUSVARA	Mc	NIMDU SUNNA		
4	0C03	◌̣̣̣̣	TELUGU SIGN VISARGA	Mc	VISARGA		
5	0C05	అ	TELUGU LETTER A	Lo	Vowel		
6	0C06	ఆ	TELUGU LETTER AA	Lo	Vowel		
7	0C07	ఇ	TELUGU LETTER I	Lo	Vowel		
8	0C08	ఈ	TELUGU LETTER II	Lo	Vowel		
9	0C09	ఉ	TELUGU LETTER U	Lo	Vowel		
10	0C0A	ఊ	TELUGU LETTER UU	Lo	Vowel		

Repertoire Included - 2

Sr. No.	Unicode Code Point	Glyph	Character Name	Unicode General Category (gc)	Indic Syllabic Category	Reference	Y/N
11	0C0B	ఱ	TELUGU LETTER VOCALIC R	Lo	Vowel		
12	0C0C	ౠ	TELUGU LETTER VOCALIC L	Lo	Vowel		
13	0C0E	ం	TELUGU LETTER E	Lo	Vowel		
14	0C0F	ః	TELUGU LETTER EE	Lo	Vowel		
15	0C10	ఄ	TELUGU LETTER AI	Lo	Vowel		
16	0C12	ఋ	TELUGU LETTER O	Lo	Vowel		
17	0C13	ౠ	TELUGU LETTER OO	Lo	Vowel		
18	0C14	ౡ	TELUGU LETTER AU	Lo	Vowel		
19	0C15	క	TELUGU LETTER KA	Lo	Consonant		
20	0C16	ఖ	TELUGU LETTER KHA	Lo	Consonant		

Repertoire Included - 3

Sr. No	Unicode Code Point	Glyph	Character Name	Unicode General Category (gc)	Indic Syllabic Category	Reference	Y/N
21	0C17	గ	TELUGU LETTER GA	Lo	Consonant	[Omniglot]	
22	0C18	ఘ	TELUGU LETTER GHA	Lo	Consonant	[Omniglot]	
23	0C19	ఙ	TELUGU LETTER NGA	Lo	Consonant	[Omniglot]	
24	0C1A	చ	TELUGU LETTER CA	Lo	Consonant	[Omniglot]	
25	0C1B	ఛ	TELUGU LETTER CHA	Lo	Consonant	[Omniglot]	
26	0C1C	జ	TELUGU LETTER JA	Lo	Consonant	[Omniglot]	
27	0C1D	ఝ	TELUGU LETTER JHA	Lo	Consonant	[Omniglot]	
28	0C1E	ఞ	TELUGU LETTER NYA	Lo	Consonant	[Omniglot]	
29	0C1F	ట	TELUGU LETTER TTA	Lo	Consonant	[Omniglot]	
30	0C20	ఠ	TELUGU LETTER TTHA	Lo	Consonant	[Omniglot]	

Repertoire Included - 4

Sr. No.	Unicode Code Point	Glyph	Character Name	Unicode General Category (gc)	Indic Syllabic Category	Reference	Y/N
31.	0C22	ఢ	TELUGU LETTER DDHA	Lo	Consonant		
32.	0C23	ణ	TELUGU LETTER NNA	Lo	Consonant		
33.	0C24	త	TELUGU LETTER TA	Lo	Consonant		
34.	0C25	థ	TELUGU LETTER THA	Lo	Consonant		
35.	0C26	ద	TELUGU LETTER DA	Lo	Consonant		
36.	0C27	ధ	TELUGU LETTER DHA	Lo	Consonant		
37.	0C28	న	TELUGU LETTER NA	Lo	Consonant		
38.	0C2A	ప	TELUGU LETTER PA	Lo	Consonant		
39.	0C2B	ఫ	TELUGU LETTER PHA	Lo	Consonant		
40.	0C22	ఢ	TELUGU LETTER DDHA	Lo	Consonant		

Repertoire Included - 5

Sr. No.	Unicode Code Point	Glyph	Character Name	Unicode General Category (gc)	Indic Syllabic Category	Reference	Y/N
41.	0C2C	బ	TELUGU LETTER BA	Lo	Consonant		
42.	0C2D	భ	TELUGU LETTER BHA	Lo	Consonant		
43.	0C2E	మ	TELUGU LETTER MA	Lo	Consonant		
44.	0C2F	య	TELUGU LETTER YA	Lo	Consonant		
45.	0C30	ర	TELUGU LETTER RA	Lo	Consonant		
46.	0C31	రి	TELUGU LETTER RRA	Lo	Consonant		
47.	0C32	ల	TELUGU LETTER LA	Lo	Consonant		
48.	0C33	ల్ల	TELUGU LETTER LLA	Lo	Consonant		
49.	0C34	ల్ల	TELUGU LETTER LLLA	Lo	Consonant		
50.	0C35	వ	TELUGU LETTER VA	Lo	Consonant		

Repertoire Included - 6

Sr. No.	Unicode Code Point	Glyph	Character Name	Unicode General Category (gc)	Indic Syllabic Category	Reference	Y/N
51.	0C36	శ	TELUGU LETTER SHA	Lo	Consonant		
52.	0C37	ష	TELUGU LETTER SSA	Lo	Consonant		
53.	0C38	స	TELUGU LETTER SA	Lo	Consonant		
54.	0C39	హ	TELUGU LETTER HA	Lo	Consonant		
55.	0C3E	ఱ	TELUGU VOWEL SIGN AA	Mn	Matra		
56.	0C3F	ి	TELUGU VOWEL SIGN I	Mn	Matra		
57.	0C40	ీ	TELUGU VOWEL SIGN II	Mn	Matra		
58.	0C41	ు	TELUGU VOWEL SIGN U	Mc	Matra		
59.	0C42	ూ	TELUGU VOWEL SIGN UU	Mc	Matra		
60.	0C43	ృ	TELUGU VOWEL SIGN VOCALIC R	Mc	Matra		

Repertoire Included - 7

Sr. No.	Unicode Code Point	Glyph	Character Name	Unicode General Category (gc)	Indic Syllabic Category	Reference	Y/N
61.	0C44	ౠ	TELUGU VOWEL SIGN VOCALIC RR	Mc	Matra		
62.	0C46	ౡ	TELUGU VOWEL SIGN E	Mn	Matra		
63.	0C47	ౢ	TELUGU VOWEL SIGN EE	Mn	Matra		
64.	0C48	ౣ	TELUGU VOWEL SIGN AI	Mn	Matra		
65.	0C4A	౤	TELUGU VOWEL SIGN O	Mn	Matra		
66.	0C4B	౥	TELUGU VOWEL SIGN OO	Mn	Matra		
67.	0C4C	౦౦	TELUGU VOWEL SIGN AU	Mn	Matra		
68.	0C4D	ౠ	TELUGU SIGN VIRAMA	Mn	Matra		
69.	0C55	ౠౠ	TELUGU LENGTH MARK	Mn	Matra		
70.	0C56	ౠౡ	TELUGU AI LENGTH MARK	Mn	Matra		

Repertoire Included - 8

Sr. No.	Unicode Code Point	Glyph	Character Name	Unicode General Category (gc)	Indic Syllabic Category	Reference	Y/N
71.	0C58	త	TELUGU LETTER TSA	Lo	Consonant		
72.	0C59	డ	TELUGU LETTER DZA	Lo	Consonant		
73.	0C5A	ఱ	TELUGU LETTER RRA	Lo	Consonant		
74.	0C60	ఱ	TELUGU LETTER VOCALIC RR	Lo	Vowel		
75.	0C61	ౠ	TELUGU LETTER VOCALIC LL	Lo	Vowel		
76.	0C62	ౡ	TELUGU Vowel sign VOCALIC L	Lo	Matra		
77.	0C63	ౢ	TELUGU Vowel SIGN VOCALIC LL	Lo	Matra		

Repertoire Excluded

- ⦿ Code Points in MSR-2 but excluded because they are either not in common use or used for special purpose only (historic, religious, etc.):

Code Point	Glyph	Type	Notes

Within Script Variants for Telugu - 1

List of code points which are variants

Group 1: Visually similar Telugu characters

Code Point 1 + Glyph 1	Code Point 2 + Glyph 2	Type (allocatable or blocked)

Within Script Variants for Telugu - 2

- List of code points which are variants:
Group 2: visually similar Telugu character combinations, due to presence dots and other characters

Code Point 1 + Glyph 1	Code Point 2 + Glyph 2	Type (allocatable or blocked)

Cross Script Variants for Telugu and <script>

- List of code points which are variants because Telugu and <script> scripts are closely related to each other:

Code Point 1 + Glyph 1	Code Point 2 + Glyph 2	Type (blocked)

Whole Label Evaluation Rules (in plain English) - 1

- ⊙ C → Consonant
- ⊙ M → Matra
- ⊙ V → Vowel
- ⊙ B → Anusvara (Bindu)
- ⊙ D → Candrabindu
- ⊙ X → Visarga
- ⊙ H → Halanta / Virama

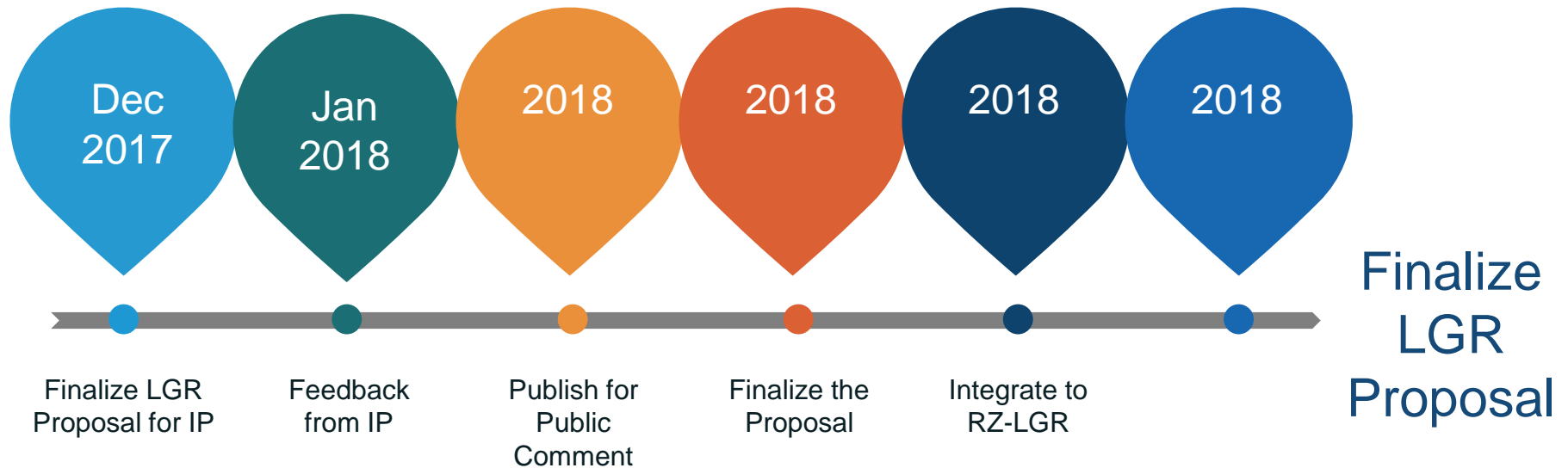
Whole Label Evaluation Rules (in plain English)

- ⦿ Rule 1. The sign for H i.e. the vowel deleting halanta must be preceded only by the C and not any thing.
- ⦿ Rule 2. The sign for M, i.e. the secondary vowel signs must always be preceded only by the C and not any thing.
- ⦿ Rule 3. X, the visarga sign must always be preceded only by the V or M.
- ⦿ Rule 4. B, the anusvara sign must always be preceded only by the V or M.
- ⦿ Rule 5. D, the chandrabindu sign must always be preceded only by the V or M.
- ⦿ Rule 6. The vowel V shall not be preceded by the sign H.

Current Status

Section in Google Doc	Status (To be initiated / In Progress / Completed / Reviewed)	Additional Notes
Description of Script and use		
Languages using Script		
Classification of characters		
Repertoire Analysis		
Within-script Variants		
Cross-script Variants		
WLE Rules		
References		

Plan Till Completion



To Summarize

The LGR proposal for Telugu script is progressing as most of the components have been completed, particularly just the variants part is still being worked upon. Ideally, we are expecting a final proposal by end of December for feedback from the Integration Panel.

Thank You and Questions

Contributors:

Name	Address	Email
Prof. Uma Masheshwar Rao		
Gangadhar Pandey		