

Title: Revised Preliminary Proposal to Encode the Gondi Script
Source: Script Encoding Initiative (SEI)
Author: Anshuman Pandey (pandey@umich.edu)
Status: Liaison Contribution
Action: For consideration by UTC and WG2
Date: 2012-07-23

1 Introduction

This is a revised proposal to encode the Gondi script in the Universal Character Set (ISO/IEC 10646). It supersedes “Preliminary Proposal to Encode the Gondi Script in the UCS” (N3841 L2/10-207). This document provides a description of the writing system, a code chart and names list, character properties, and a few specimens. Some issues requiring further attention are specified in Section 5. The present author is in contact with the user community and will provide additional details in the future.

2 Background

The Gondi script was designed by Munshi Mangal Singh Masaram of Balaghat district, Madhya Pradesh, India in 1928. The writing system is based upon the Brahmi model. It is used for writing Gondi (ISO 639-3: gon), a Dravidian language spoken by 2.6 million people, primarily in Madhya Pradesh and Maharashtra, with some speakers in Andhra Pradesh and Chhattisgarh. The language is generally written in Devanagari and Telugu. The Gondi script has no genetic relationship to other writing systems. The script appears to be actively used and fonts have been developed for it. The available materials indicate that Masaram’s original script has been modified over the years.

3 Script Details

3.1 Structure

The script is based upon the Brahmi model. It is written from left to right. Consonant letters possess the inherent vowel *a*, which is graphically represented by the horizontal line at the right edge of each consonant glyph. There is no *virāma*; the inherent vowel is silenced by removing the horizontal line. Independent and initial vowels are written using letters, while dependent signs are used for medial and final vowels. All vowel signs are written either above or below the horizontal stroke. There is no *mātrā* reordering.

3.2 Character Repertoire

Gondi consists of 67 characters, shown in the code chart and names list (figures 1 and 2). Names for characters follow the UCS convention for Brahmi-based scripts and align with names given by Masaram (1951). An analysis of the available materials indicates several variations in the glyphs for characters. These differences may be attributed to simplification for ease of writing, eg. independent circles being joined as loops with a single stroke. Normative glyphs will be determined through communication with users.

3.3 Virama

In the absence of a native *virāma*, it is necessary to encode a control character in order to write conjuncts according to the Unicode model for Brahmi-based scripts. This character is VIRAMA. It is not rendered visibly. The dotted box indicates that the character has special properties.

3.4 Vowels

There are 10 vowel letters:

𑌒	A	𑌓	II	𑌔	E	𑌕	AU
𑌖	AA	𑌗	U	𑌘	AI		
𑌙	I	𑌚	UU	𑌛	O		

3.5 Vowel Signs

There are 10 dependent vowel signs:

◌𑌐	VOWEL SIGN AA	◌𑌑	VOWEL SIGN UU	◌𑌒	VOWEL SIGN O
◌𑌓	VOWEL SIGN I	◌𑌔	VOWEL SIGN VOCALIC R	◌𑌕	VOWEL SIGN AU
◌𑌖	VOWEL SIGN II	◌𑌗	VOWEL SIGN E		
◌𑌘	VOWEL SIGN U	◌𑌙	VOWEL SIGN AI		



An independent letter for ◌𑌔 VOWEL SIGN VOCALIC R is not attested.

3.6 Consonants



There are 34 consonant letters:

𑌐	KA	𑌑	NYA	𑌒	DHA	𑌓	LA
𑌔	KHA	𑌕	TTA	𑌖	NA	𑌗	VA
𑌘	GA	𑌙	TTHA	𑌚	PA	𑌛	SHA
𑌜	GHA	𑌝	DDA	𑌞	PHA	𑌟	SSA
𑌠	NGA	𑌡	DDHA	𑌢	BA	𑌣	SA
𑌤	CA	𑌥	NNA	𑌦	BHA	𑌧	HA
𑌨	CHA	𑌩	TA	𑌪	MA	𑌫	LLA
𑌭	JA	𑌮	THA	𑌯	YA		
𑌱	JHA	𑌲	DA	𑌳	RA		




3.7 Conjuncts

Consonant clusters are generally represented using half-forms of consonant letters. A half-form is rendered by removing the horizontal stroke at the right of each letter. Following the UCS *virāma* model, half-forms are represented in encoded text as <C,  VIRAMA>, for example: <𑌒 KA,  VIRAMA> = 𑌒 . Figure 5 shows several half-forms of consonants in actual text. Conjuncts are written sequentially, but there are four exceptions: RA and three atomic ligatures.

Conjuncts with RA The Gondi RA behaves similar to Devanagari RA:

1. *repha* When RA is the first consonant in a cluster, it is represented as the combining mark \circ^a , which is written above the horizontal line of a consonant glyph. Some current users of the script have modified the shape and placement of the original *repha*, which they now write as \circ^b . This new sign is placed linearly after the non-initial consonant. The change was supposedly required in order to rectify the issues posed by the placement of multiple combining marks on the horizontal link of a consonant letter. Ideally, the encoding will support both the original and modern forms of *repha*, which should be considered glyphic variants despite their different positions. The encoded representation of the *repha* is <𑌒 RA,  VIRAMA>.
2. *ra-kāra* When RA is a non-initial consonant in a cluster, it is represented as the combining mark \circ_c , which is written below the horizontal line of a consonant glyph. The encoded representation of *ra-kāra* is <C,  VIRAMA, 𑌒 RA>.

Atomic ligatures The conjuncts *kṣa*, *jña*, and *tra* are written using independent ligatures. These are to be encoded using the following sequences:

1. *kṣa* <𑌒 KA,  VIRAMA, 𑌒 SSA> = 𑌒
2. *jña* <𑌒 JA,  VIRAMA, 𑌒 NYA> = 𑌒
3. *tra* <𑌒 TA,  VIRAMA, 𑌒 RA> = 𑌒

3.8 Various Signs

1. \circ^a SIGN ANUSVARA This sign indicates nasalization. It combines to the right of the accompanying letter; in some cases it is written above the letter.
2. \circ_c SIGN VISARGA This sign indicates post-vocalic aspiration and is used primarily for writing Sanskrit. It combines to the right of the accompanying letter.

3.9 Digits

There is a full set of decimal digits: 0 ZERO, 1 ONE, 2 TWO, 3 THREE, 4 FOUR, 5 FIVE, 6 SIX, 7 SEVEN, 8 EIGHT, 9 NINE.

3.10 Punctuation

Script-specific punctuation is not attested. The use of *daṇḍā*-s are attested (see figure 5), but these are to be unified with 𑌒 U+0964 DEVANAGARI DANDA and 𑌒 U+0965 DEVANAGARI DOUBLE DANDA. Latin punctuation, such as periods, are also used.

4 Character Data

4.1 Character Properties

The properties for Gondi in the Unicode Character Database format are:

```

11B90;GONDI LETTER A;Lo;0;L;;;;;N;;;;;
11B91;GONDI LETTER AA;Lo;0;L;;;;;N;;;;;
11B92;GONDI LETTER I;Lo;0;L;;;;;N;;;;;
11B93;GONDI LETTER II;Lo;0;L;;;;;N;;;;;
11B94;GONDI LETTER U;Lo;0;L;;;;;N;;;;;
11B95;GONDI LETTER UU;Lo;0;L;;;;;N;;;;;
11B96;GONDI LETTER E;Lo;0;L;;;;;N;;;;;
11B97;<reserved>
11B98;GONDI LETTER AI;Lo;0;L;;;;;N;;;;;
11B99;GONDI LETTER O;Lo;0;L;;;;;N;;;;;
11B9A;<reserved>
11B9B;GONDI LETTER AU;Lo;0;L;;;;;N;;;;;
11B9C;GONDI LETTER KA;Lo;0;L;;;;;N;;;;;
11B9D;GONDI LETTER KHA;Lo;0;L;;;;;N;;;;;
11B9E;GONDI LETTER GA;Lo;0;L;;;;;N;;;;;
11B9F;GONDI LETTER GHA;Lo;0;L;;;;;N;;;;;
11BA0;GONDI LETTER NGA;Lo;0;L;;;;;N;;;;;
11BA1;GONDI LETTER CA;Lo;0;L;;;;;N;;;;;
11BA2;GONDI LETTER CHA;Lo;0;L;;;;;N;;;;;
11BA3;GONDI LETTER JA;Lo;0;L;;;;;N;;;;;
11BA4;GONDI LETTER JHA;Lo;0;L;;;;;N;;;;;
11BA5;GONDI LETTER NYA;Lo;0;L;;;;;N;;;;;
11BA6;GONDI LETTER TTA;Lo;0;L;;;;;N;;;;;
11BA7;GONDI LETTER TTHA;Lo;0;L;;;;;N;;;;;
11BA8;GONDI LETTER DDA;Lo;0;L;;;;;N;;;;;
11BA9;GONDI LETTER DDHA;Lo;0;L;;;;;N;;;;;
11BAA;GONDI LETTER NNA;Lo;0;L;;;;;N;;;;;
11BAB;GONDI LETTER TA;Lo;0;L;;;;;N;;;;;
11BAC;GONDI LETTER THA;Lo;0;L;;;;;N;;;;;
11BAD;GONDI LETTER DA;Lo;0;L;;;;;N;;;;;
11BAE;GONDI LETTER DHA;Lo;0;L;;;;;N;;;;;
11BAF;GONDI LETTER NA;Lo;0;L;;;;;N;;;;;
11BB0;GONDI LETTER PA;Lo;0;L;;;;;N;;;;;
11BB1;GONDI LETTER PHA;Lo;0;L;;;;;N;;;;;
11BB2;GONDI LETTER BA;Lo;0;L;;;;;N;;;;;
11BB3;GONDI LETTER BHA;Lo;0;L;;;;;N;;;;;
11BB4;GONDI LETTER MA;Lo;0;L;;;;;N;;;;;
11BB5;GONDI LETTER YA;Lo;0;L;;;;;N;;;;;
11BB6;GONDI LETTER RA;Lo;0;L;;;;;N;;;;;
11BB7;GONDI LETTER LA;Lo;0;L;;;;;N;;;;;
11BB8;GONDI LETTER VA;Lo;0;L;;;;;N;;;;;
11BB9;GONDI LETTER SHA;Lo;0;L;;;;;N;;;;;
11BBA;GONDI LETTER SSA;Lo;0;L;;;;;N;;;;;
11BBB;GONDI LETTER SA;Lo;0;L;;;;;N;;;;;
11BBC;GONDI LETTER HA;Lo;0;L;;;;;N;;;;;
11BBD;GONDI LETTER LLA;Lo;0;L;;;;;N;;;;;
11BBE;GONDI VOWEL SIGN AA;Mn;0;NSM;;;;;N;;;;;
11BBF;GONDI VOWEL SIGN I;Mn;0;NSM;;;;;N;;;;;
11BC0;GONDI VOWEL SIGN II;Mn;0;NSM;;;;;N;;;;;
11BC1;GONDI VOWEL SIGN U;Mn;0;NSM;;;;;N;;;;;
11BC2;GONDI VOWEL SIGN UU;Mn;0;NSM;;;;;N;;;;;
11BC3;GONDI VOWEL SIGN VOCALIC R;Mn;0;NSM;;;;;N;;;;;
11BC4;GONDI VOWEL SIGN E;Mn;0;NSM;;;;;N;;;;;
11BC5;<reserved>

```

```

11BC6;GONDI VOWEL SIGN AI;Mn;0;NSM;;;;;N;;;;;
11BC7;GONDI VOWEL SIGN O;Mn;0;NSM;;;;;N;;;;;
11BC8;<reserved>
11BC9;GONDI VOWEL SIGN AU;Mn;0;NSM;;;;;N;;;;;
11BCA;GONDI SIGN ANUSVARA;Mn;0;NSM;;;;;N;;;;;
11BCB;GONDI SIGN VISARGA;Mn;0;NSM;;;;;N;;;;;
11BCC;GONDI SIGN VIRAMA;Mn;9;NSM;;;;;N;;;;;
11BD0;GONDI DIGIT ZERO;Nd;0;L;;0;0;0;N;;;;;
11BD1;GONDI DIGIT ONE;Nd;0;L;;1;1;1;N;;;;;
11BD2;GONDI DIGIT TWO;Nd;0;L;;2;2;2;N;;;;;
11BD3;GONDI DIGIT THREE;Nd;0;L;;3;3;3;N;;;;;
11BD4;GONDI DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;;
11BD5;GONDI DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;;
11BD6;GONDI DIGIT SIX;Nd;0;L;;6;6;6;N;;;;;
11BD7;GONDI DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;;
11BD8;GONDI DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;;
11BD9;GONDI DIGIT NINE;Nd;0;L;;9;9;9;N;;;;;

```

4.2 Linebreaking

Linebreaking properties given in the data format of LineBreak.txt:

```

11B90..11BBD; AL # LETTER A .. LETTER LLA
11BBE..11BCC; CM # SIGN AA .. SIGN VIRAMA
11BD0..11BD9; NU # DIGIT ZERO .. DIGIT NINE

```

4.3 ‘Confusable’ Characters

Gondi characters that bear resemblances to those of other scripts are listed below:

```

11BB1 GONDI LETTER PHA ; 1109D KAITHI LETTER NNA
11BBA GONDI LETTER SSA ; 0398 GREEK CAPITAL LETTER THETA
11BD2 GONDI DIGIT TWO ; 0055 LATIN CAPITAL LETTER U

```

5 Issues

- *Additional vowels* Does Gondi have distinct vowel letters and signs for the Dravidian /e:/ and /o:/, corresponding to U+0C0F TELUGU LETTER EE and U+0C13 TELUGU LETTER OO? Space has been reserved for these letters and their dependent signs in the code chart in the case that such characters are attested.
- *Virāma* Masaram’s original script lacks a VIRAMA. The structure of the script does not require the visible representation of such a character. However, as shown in figure 6, a Devanagari-like Q VIRAMA is used in Gondi text for representing half-forms of consonants. Such usage is superfluous, given that a half-form is written by eliminating the horizontal line that accompanies each consonant letter. Is the use of the Devanagari-like *virāma* in figure 6 common or idiosyncratic?
- *Repha* Will the requirement to support both forms of *repha* present any implementation issues?

6 References

Mandavi, Ashutosh. 2008. “घोटुल [Ghoṭul]: Tribal Arts and Cultural Initiative”. http://ashutoshmandavi.blogspot.com/2008/11/blog-post_07.html

Maṇḍāle, Sītārām. कोयाबोली [*Koyābolī*]. गोंडी शब्द संग्रह - गोंडी, मराठी, हिन्दी [= *Gomḍī Śabda Saṁgraha - Gomḍī, Marāṭhī, Hindī*].

Masaram, Mangalasinha. 1951. “गोंडी लिपि” [*Gomḍī lipi*]. Central Institute of Indian Languages, Multimedia library, photograph no. 64.

National Folklore Support Center and Jatan Trust. n.d. “The Gonds of Madhya Pradesh.” <http://www.slideshare.net/nfsc/the-gonds-of-madhya-pradesh>

Pandey, Anshuman. 2010. “Preliminary Proposal to Encode the Gondi Script in the UCS”. ISO/IEC JTC1/SC2/WG2 N3841 L2/10-207. May 20, 2010. <http://std.dkuug.dk/jtc1/sc2/wg2/docs/n3841.pdf>

Ramakrishna, G., N. Gayathri, Debiprasad Chattopadhyaya. 1983. *An Encyclopaedia of South Indian Culture*. Calcutta: K. P. Bagchi & Co.

7 Acknowledgments

I would like to extend my gratitude to B. A. Sharada and Suman Kumari of the Central Institute of Indian Languages (Mysore) for providing a copy of the Gondi chart shown in Figure 3. I am also grateful to Mukund Gokhale, Raymond Doctor, and Mark Penny for providing me with information regarding the current status of the script and with specimens of the script.

This project was made possible in part by a grant from the United States National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at the University of California, Berkeley). Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment for the Humanities.




































































	11B9	11BA	11BB	11BC	11BD
0	 11B90	 11BA0	 11BB0	 11BC0	 11BD0
1	 11B91	 11BA1	 11BB1	 11BC1	 11BD1
2	 11B92	 11BA2	 11BB2	 11BC2	 11BD2
3	 11B93	 11BA3	 11BB3	 11BC3	 11BD3
4	 11B94	 11BA4	 11BB4	 11BC4	 11BD4
5	 11B95	 11BA5	 11BB5		 11BD5
6	 11B96	 11BA6	 11BB6	 11BC6	 11BD6
7		 11BA7	 11BB7	 11BC7	 11BD7
8	 11B98	 11BA8	 11BB8		 11BD8
9	 11B99	 11BA9	 11BB9	 11BC9	 11BD9
A		 11BAA	 11BBA	 11BCA	
B	 11B9B	 11BAB	 11BBB	 11BCB	
C	 11B9C	 11BAC	 11BBC	 11BCC	
D	 11B9D	 11BAD	 11BBD		
E	 11B9E	 11BAE	 11BBE		
F	 11B9F	 11BAF	 11BBF		

Figure 1: Proposed code chart for Gondi.

Vowels

11B90	᱀	GONDI LETTER A
11B91	᱁	GONDI LETTER AA
11B92	᱂	GONDI LETTER I
11B93	᱃	GONDI LETTER II
11B94	᱄	GONDI LETTER U
11B95	᱅	GONDI LETTER UU
11B96	᱆	GONDI LETTER E
11B97	᱇	<reserved>
11B98	᱈	GONDI LETTER AI
11B99	᱉	GONDI LETTER O
11B9A	᱊	<reserved>
11B9B	᱋	GONDI LETTER AU

Consonants

11B9C	᱌	GONDI LETTER KA
11B9D	ᱍ	GONDI LETTER KHA
11B9E	ᱎ	GONDI LETTER GA
11B9F	ᱏ	GONDI LETTER GHA
11BA0	᱐	GONDI LETTER NGA
11BA1	᱑	GONDI LETTER CA
11BA2	᱒	GONDI LETTER CHA
11BA3	᱓	GONDI LETTER JA
11BA4	᱔	GONDI LETTER JHA
11BA5	᱕	GONDI LETTER NYA
11BA6	᱖	GONDI LETTER TTA
11BA7	᱗	GONDI LETTER TTHA
11BA8	᱘	GONDI LETTER DDA
11BA9	᱙	GONDI LETTER DDHA
11BAA	ᱚ	GONDI LETTER NNA
11BAB	ᱛ	GONDI LETTER TA
11BAC	ᱜ	GONDI LETTER THA
11BAD	ᱝ	GONDI LETTER DA
11BAE	ᱞ	GONDI LETTER DHA
11BAF	ᱟ	GONDI LETTER NA
11BB0	ᱠ	GONDI LETTER PA
11BB1	ᱡ	GONDI LETTER PHA
11BB2	ᱢ	GONDI LETTER BA
11BB3	ᱣ	GONDI LETTER BHA
11BB4	ᱤ	GONDI LETTER MA
11BB5	ᱥ	GONDI LETTER YA
11BB6	ᱦ	GONDI LETTER RA
11BB7	ᱧ	GONDI LETTER LA
11BB8	ᱨ	GONDI LETTER VA
11BB9	ᱩ	GONDI LETTER SHA
11BBA	ᱪ	GONDI LETTER SSA
11BBB	ᱫ	GONDI LETTER SA
11BBC	ᱬ	GONDI LETTER HA
11BBD	ᱭ	GONDI LETTER LLA

Dependent vowel signs

11BBE	ᱮ	GONDI VOWEL SIGN AA
11BBF	ᱯ	GONDI VOWEL SIGN I
11BC0	ᱰ	GONDI VOWEL SIGN II
11BC1	ᱱ	GONDI VOWEL SIGN U
11BC2	ᱲ	GONDI VOWEL SIGN UU
11BC3	ᱳ	GONDI VOWEL SIGN VOCALIC R
11BC4	ᱴ	GONDI VOWEL SIGN E
11BC5	ᱵ	<reserved>
11BC6	ᱶ	GONDI VOWEL SIGN AI
11BC7	ᱷ	GONDI VOWEL SIGN O
11BC8	ᱸ	<reserved>
11BC9	ᱹ	GONDI VOWEL SIGN AU

Various signs

11BCA	ᱺ	GONDI SIGN ANUSVARA
11BCB	ᱻ	GONDI SIGN VISARGA

Virama

11BCC	ᱼ	GONDI SIGN VIRAMA
-------	---	-------------------

Digits

11BD0	ᱽ	GONDI DIGIT ZERO
11BD1	᱾	GONDI DIGIT ONE
11BD2	᱿	GONDI DIGIT TWO
11BD3	᱀	GONDI DIGIT THREE
11BD4	᱁	GONDI DIGIT FOUR
11BD5	᱂	GONDI DIGIT FIVE
11BD6	᱃	GONDI DIGIT SIX
11BD7	᱄	GONDI DIGIT SEVEN
11BD8	᱅	GONDI DIGIT EIGHT
11BD9	᱆	GONDI DIGIT NINE

Figure 2: Proposed names list for Gondi.

Consonants व्यञ्जन

	Cons. only	- [voice]		+ [voice]		Nasals	Cons. with word
		- [asp]	+ [asp]	- [asp]	+ [asp]		
<i>Aspiration</i>							
Velar		ᱠ	ᱡ	ᱢ	ᱣ	ᱤ	
Palatal		ᱥ	ᱦ	ᱧ	ᱨ	ᱩ	
Retroflex		ᱪ	ᱫ	ᱬ	ᱭ	ᱮ	
Dental		ᱰ	ᱱ	ᱲ	ᱳ	ᱴ	
Bilabial		ᱵ	ᱶ	ᱷ	ᱸ	ᱹ	
Liquids and Semi-Vowels		ᱺ	ᱻ	ᱼ	ᱽ		
Fricatives		᱾	᱿	ᱠ	ᱡ		
Affricates			ᱢ	ᱣ	ᱤ		

Vowels स्वर

	ᱠ	ᱡ	ᱢ	ᱣ	ᱤ	ᱥ		
	ᱦ	ᱧ	ᱨ	ᱩ	ᱪ	ᱫ		

Figure 4: A handwritten chart of the Gondi script. Source: Ramesh Gedam and Mark Penny (2001).



Figure 6: Cover of a book on Gondi language (from Mandale). Content printed in the Gondi script contains a Devanagari-like ्र VIRAMA. Image courtesy of Mukund Gokhale

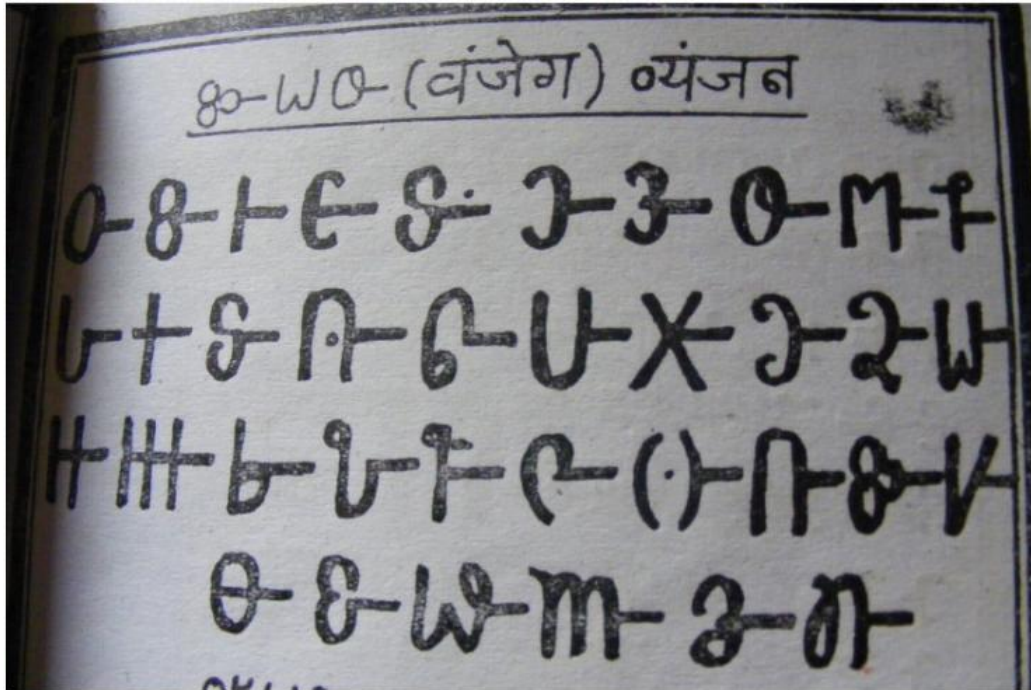


Figure 7: Excerpt from a photograph of a chart of the Gondi script (from National Folklore Support Center and Jatan Trust: 13).

HINDI KRUTI DEV AND GONDI FONT KEYBOARD

!!	/	रु	++	ˆ	:	'	-	'	;	;	द्ध	.
८	७	७२	७३	७४	७५	७६	७७	७८	७९	८०	८१	८२
८३	८४	८५	८६	८७	८८	८९	९०	९१	९२	९३	९४	९५
९६	९७	९८	९९	१००	१०१	१०२	१०३	१०४	१०५	१०६	१०७	१०८
१०९	११०	१११	११२	११३	११४	११५	११६	११७	११८	११९	१२०	१२१
१२२	१२३	१२४	१२५	१२६	१२७	१२८	१२९	१३०	१३१	१३२	१३३	१३४
१३५	१३६	१३७	१३८	१३९	१४०	१४१	१४२	१४३	१४४	१४५	१४६	१४७

Figure 8: Keyboard map for a non-Unicode Gondi font (from Mandavi 2008).