L2/15-090R
2015-06-02

Proposal to Encode the Masaram Gondi Script in Unicode

Anshuman Pandey
Department of Linguistics
University of California, Berkeley
Berkeley, California, U.S.A.
anshuman.pandey@berkeley.edu

June 2, 2015

1 Introduction

This is a proposal to encode the Gondi script created by Mangal Singh Masaram in 1918 in Unicode. It replaces and supersedes the following documents:

- L2/10-207 "Preliminary Proposal to Encode the Gondi Script in the UCS"
- L2/12-235 "Revised Preliminary Proposal to Encode the Gondi Script"
- L2/15-005 "Proposal to Encode the Gondi Script"

This document provides a description of the writing system, a code chart and names list, character properties, and specimens that illustrate letterforms and usage. It is a revision of L2/15-005 and contains several changes to the encoding proposed in that document. The major changes are as follows:

- The block name has been changed from 'Gondi' to 'Masaram Gondi'
- Separate encoding of REPHA as cluster-initial and RA-KARA as cluster-final forms of RA
- Redefinition of VIRAMA specifically as a control character used only for producing conjuncts
- Addition of HALANTA as a combining sign used solely for silencing the inherent vowel
- Addition of CANDRA for transcribing foreign vowel sounds

These changes were introduced as a result of discussions with experts such as Mukund Gokhale, members of the user community, and with Unicode implementers at Google and Microsoft.

The 'Masaram Gondi' script is graphically and structurally distinct from another Gondi script known as 'Gunjala Gondi'. A preliminary proposal for 'Gunjala Gondi' has been submitted (see L2/15-086).

The symbol  appears in several sources containing Masaram's script and is a common Gond motif (see figures 10, 21). This symbol represents *persapen*, or the supreme spirit, in the indigenous Gond philosophical system known as *koya punem*. This symbol has been proposed for encoding in the 'Miscellaenous Symbols and Pictographs' block in Unicode, where several religious symbols are encoded (see L2/15-111).

2 Background

The script described here was invented by Munshi Mangal Singh Masaram of Kochewada, Balaghat District, Madhya Pradesh, India in 1918. It has no genetic relationship to other writing systems, but it is based upon the Brahmi model. The script was designed 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 both Devanagari and Telugu. Manuscripts containing yet another script, which appears to have been graphically inspired by these Modi writing system, was found in Gunjala in the Adilabad district of Telangana Masaram's Gondi script is actively used today for hand-written and printed materials. Fonts have been developed for the production of books. In 2011, the Akhil Gondvana Gondi Sahitya Parishad (Chandagadh, Maharashtra) passed a resolution adopting Masaram's script as the official script of the Gondi language.

Masaram's script has been slightly expanded and revised over the years in order to meet the needs and preferences of modern users. Innovations include the addition of new consonant letters, vowel signs, a vowel-sound modifier, and the adoption of a Devanagari-style *halanta* for indicating the absence of the inherent vowel. Some of these new characters are included in the proposed repertoire, while others are not (see section 4.12).

3 Script Details

3.1 Name

Earlier versions of this proposal referred to name of the script block as 'Gondi'. While the script is certainly used for writing Gondi, it is one of many scripts used for the language. Moreover, the 'Gunjala Gondi' script is also associated with the language and culture of the Gonds. Given this, it is appropriate to assign an identifier for the script block that precisely defines which 'Gondi' script is contained within that block. A designation that includes the name of the script's creator seems appropriate. For this reason, the name suggested for the script block in Unicode is 'Masaram Gondi'. The names of characters contain the block name. Users may refer to the script as 'Gondi' or by whatever name they prefer outside of a Unicode context.

3.2 Structure

Masaram's Gondi script is an alphasyllabary that is written from left to right. Consonant letters possess the inherent vowel *a*, which is graphically represented by a horizontal stroke that extends rightward from the right edge of each consonant letter. A bare consonant is represented by removing this stroke. Some modern users represent a word-final bare consonant by writing the *halanta* beneath the stroke of the consonant letter. Consonant clusters are represented as conjuncts, which are rendered as a linear sequence using bare forms for all consonants except for the final, which occurs in its regular form. There are some exceptions to this rule, namely the behavior of *ra* and the usage of distinctive letter ligatures for the conjuncts *kṣa*, *jña*, *tra*. Independent and initial vowels are written using vowel letters, while consonant-vowel combinations are expressed using dependent signs. There is no *mātrā* reordering.

3.3 Character Repertoire

A total of 75 characters are proposed for encoding in the 'Masaram Gondi' script block. A code chart and names list are attached. Names for characters follow the UCS convention for Brahmi-based scripts and align with the Latin transliteration of Devanagari analogues for Gondi letters given by B. S. Masaram (1951).

3.4 Glyphic Representations

The glyphic representations of some consonant letters and digits have changed since the invention of the script. These differences result from the simplification of glyphs for ease of writing, ie. sets of independent circles being joined into a single-stroked loop. Representative glyphs are based upon forms shown in published script primers and reflect modern preferences. The font was designed by the proposal author.

4 Proposed Encoding

4.1 Vowel Letters

Ten vowel letters are proposed for encoding:

𑵆	MASARAM GONDI LETTER A	𑵕	MASARAM GONDI LETTER UU
𑵇	MASARAM GONDI LETTER AA	𑵖	MASARAM GONDI LETTER E
𑵈	MASARAM GONDI LETTER I	𑵗	MASARAM GONDI LETTER AI
𑵉	MASARAM GONDI LETTER II	𑵘	MASARAM GONDI LETTER O
𑵊	MASARAM GONDI LETTER U	𑵙	MASARAM GONDI LETTER AU

Masaram's script does not have independent letters or dependent signs for the Dravidian long vowels /e:/ and /o:/, which correspond to Telugu ఏ *ē* (U+0C0F TELUGU LETTER EE) and ఓ *ō* (U+0C0F TELUGU LETTER OO). Space is reserved in the script block in the case that attestations for these characters are identified.

4.2 Vowel Signs

Ten dependent vowel signs are proposed for encoding:

◌𑵞	MASARAM GONDI VOWEL SIGN AA	◌𑵣	MASARAM GONDI VOWEL SIGN VOCALIC R
◌𑵟	MASARAM GONDI VOWEL SIGN I	◌𑵤	MASARAM GONDI VOWEL SIGN E
◌𑵠	MASARAM GONDI VOWEL SIGN II	◌𑵥	MASARAM GONDI VOWEL SIGN AI
◌𑵡	MASARAM GONDI VOWEL SIGN U	◌𑵦	MASARAM GONDI VOWEL SIGN O
◌𑵢	MASARAM GONDI VOWEL SIGN UU	◌𑵧	MASARAM GONDI VOWEL SIGN AU

Vowel signs are written above and below the horizontal stroke of a consonant letter:

𑵆	𑵆𑵞	𑵆𑵟	𑵆𑵠	𑵆𑵡	𑵆𑵢	𑵆𑵣	𑵆𑵦	𑵆𑵥	𑵆𑵤	𑵆𑵧
ka	*kā*	*ki*	*kī*	*ku*	*kū*	*kr̥*	*ke*	*kai*	*ko*	*kau*

These combinations would be represented in encoded text as follows:

kā 𑵆𑵞 <𑵆 KA, ◌𑵞 VOWEL SIGN AA>

ki 𑵆𑵟 <𑵆 KA, ◌𑵟 VOWEL SIGN I>

kī <KA, ◌ VOWEL SIGN II>

ku <KA, ◌ VOWEL SIGN U>

kū <KA, ◌ VOWEL SIGN UU>

kṛ <KA, ◌ VOWEL SIGN VOCALIC R>

ke <KA, ◌ VOWEL SIGN E>

kai <KA, ◌ VOWEL SIGN AI>

ko <KA, ◌ VOWEL SIGN O>

kau <KA, ◌ VOWEL SIGN AU>

There is no independent letter analogue for VOWEL SIGN VOCALIC R. The independent form of this vowel is represented using a consonant-vowel combination composed with the letter RA:

ṛ <RA, ◌ VOWEL SIGN VOCALIC R>

4.3 Consonants

There are 34 consonant letters:

	MASARAM GONDI LETTER KA		MASARAM GONDI LETTER DA
	MASARAM GONDI LETTER KHA		MASARAM GONDI LETTER DHA
	MASARAM GONDI LETTER GA		MASARAM GONDI LETTER NA
	MASARAM GONDI LETTER GHA		MASARAM GONDI LETTER PA
	MASARAM GONDI LETTER NGA		MASARAM GONDI LETTER PHA
	MASARAM GONDI LETTER CA		MASARAM GONDI LETTER BA
	MASARAM GONDI LETTER CHA		MASARAM GONDI LETTER BHA
	MASARAM GONDI LETTER JA		MASARAM GONDI LETTER MA
	MASARAM GONDI LETTER JHA		MASARAM GONDI LETTER YA
	MASARAM GONDI LETTER NYA		MASARAM GONDI LETTER RA
	MASARAM GONDI LETTER TTA		MASARAM GONDI LETTER LA
	MASARAM GONDI LETTER TTHA		MASARAM GONDI LETTER VA
	MASARAM GONDI LETTER DDA		MASARAM GONDI LETTER SHA
	MASARAM GONDI LETTER DDHA		MASARAM GONDI LETTER SSA
	MASARAM GONDI LETTER NNA		MASARAM GONDI LETTER SA
	MASARAM GONDI LETTER TA		MASARAM GONDI LETTER HA
	MASARAM GONDI LETTER THA		MASARAM GONDI LETTER LLA

The letter LLA is not part of Masaram's original script. It was introduced by modern users in order to properly represent Marathi ळ *ḷa* (see figure 19 for an example of LLA in usage).

The horizontal stroke of a letter may be extended in order to accommodate the placement of multiple combining signs. Examples are shown throughout the proposal in discussions of the relevant signs.

4.4 Vowel modifiers

Three vowel modifiers are proposed for encoding:

◌𑵠 MASARAM GONDI SIGN ANUSVARA

◌𑵡 MASARAM GONDI SIGN VISARGA

◌𑵢 MASARAM GONDI SIGN CANDRA

Anusvara The sign ◌𑵠 ANUSVARA is used for marking nasalization. It is placed above the horizontal stroke of a consonant. Its position differs slightly in different sources. In some documents the position is altered by the presence of an accompanying above-base vowel sign:

As shown above, the ANUSVARA is placed above the stroke when there is no vowel sign or the vowel sign is below the stroke. Its position is raised when ◌𑵐 VOWEL SIGN AA is present. It is placed to the right of ◌𑵑 VOWEL SIGN I, ◌𑵒 VOWEL SIGN II, ◌𑵓 VOWEL SIGN U, ◌𑵔 VOWEL SIGN UU, ◌𑵕 VOWEL SIGN E, ◌𑵖 VOWEL SIGN AI. It is placed to the right and raised higher with ◌𑵗 VOWEL SIGN O, ◌𑵘 VOWEL SIGN AU. Some modern users prefer a more stationary position for ANUSVARA and place it above the body of the consonant:

These positional preferences are to be managed in the font. The sign ANUSVARA is used in encoded text as shown below. It is always placed after a vowel sign in the encoded sequence:

kaṃ 𑵎𑵠 <𑵎 KA, ◌𑵠 ANUSVARA>

kāṃ 𑵎𑵑 <𑵎 KA, ◌𑵐 VOWEL SIGN AA, ◌𑵠 ANUSVARA>

kīṃ 𑵎𑵒𑵠 <𑵎 KA, ◌𑵒 VOWEL SIGN II, ◌𑵠 ANUSVARA>

kr̥ṃ 𑵎𑵣𑵠 <𑵎 KA, ◌𑵣 VOWEL SIGN VOCALIC R, ◌𑵠 ANUSVARA>

Visarga The ◌𑵡 VISARGA is used for the representation of Sanskrit words. It is written above the horizontal line of a consonant letter. When occurring with vowel signs its position is adjusted as follows:

Some modern sources show the ◌𑵡 VISARGA written as the glyphic variant ◌𑵞. It is placed after the stroke:

The VISARGA is used in encoded text as follows:

kaḥ 𑵐𑵉 <𑵐 KA, ◌𑵉 VISARGA>

kāḥ 𑵐𑵊𑵉 <𑵐 KA, ◌𑵊 VOWEL SIGN AA, ◌𑵉 VISARGA>

Candra The sign ◌𑵋 CANDRA is used for transcribing vowel sounds that do not occur natively in Gondi. Examples are given in figure 25. It is derived from the sign ◌ॅ used in Devanagari orthography for Marathi for representing the English vowel sounds /æ/ and /ɔ/. The CANDRA is written above the horizontal line of a consonant letter, and is used as follows:

/æ/ 𑵅𑵋 <𑵅 E, ◌𑵋 CANDRA>

/ɔ/ 𑵆𑵋 <𑵆 AA, ◌𑵋 CANDRA>

/kæ/ 𑵐𑵋 <𑵐 KA, ◌𑵋 CANDRA>

/kɔ/ 𑵐𑵊𑵋 <𑵐 KA, ◌𑵊 VOWEL SIGN AA, ◌𑵋 SIGN CANDRA>

In initial and independent contexts, CANDRA is used only with the vowel letters E and O for representing /æ/ and /ɔ/, respectively. The combination 𑵅𑵋 corresponds to ऍ U+090D DEVANAGARI LETTER CANDRA E, while 𑵆𑵋 corresponds to ऑ U+0911 DEVANAGARI LETTER CANDRA O. In dependent contexts, the CANDRA combines with a consonant letter for the sound /æ/, and with the VOWEL SIGN AA for the sound /ɔ/. The ◌𑵋 CANDRA corresponds to ◌ॅ U+0945 DEVANAGARI VOWEL SIGN CANDRA E, while the sequence <◌𑵊 GONDI VOWEL SIGN AA, ◌𑵋 CANDRA> corresponds to ◌ॉ U+0949 DEVANAGARI VOWEL SIGN CANDRA O. Although the Gondi sequences correspond to atomic characters in Devanagari, there is no need to encode such precomposed letters and signs with CANDRA for Gondi.

Although the contexts in which CANDRA is used are limited, the sign technically may be used with any letter and may occur with any vowel sign. It would be positioned with other vowel signs as follows:

𑵐𑵋 𑵐𑵊𑵋 𑵐𑵇𑵋 𑵐𑵈𑵋 𑵐𑵉𑵋 𑵐𑵌𑵋 𑵐𑵍𑵋 𑵐𑵎𑵋 𑵐𑵏𑵋 𑵐𑵐𑵋 𑵐𑵑𑵋

4.5 Nukta

The ◌𑵃 MASARAM GONDI SIGN NUKTA is used for representing sounds that are not native to the Gondi language. It is written beneath the horizontal stroke of a consonant:

𑵐𑵃 𑵐𑵊𑵃 𑵐𑵇𑵃 𑵐𑵈𑵃 𑵐𑵉𑵃 𑵐𑵌𑵃 𑵐𑵍𑵃 𑵐𑵎𑵃 𑵐𑵏𑵃 𑵐𑵐𑵃 𑵐𑵑𑵃

Some users prefer to position the NUKTA below the body of the consonant letter:

𑵐𑵃 𑵐𑵊𑵃 𑵐𑵇𑵃 𑵐𑵈𑵃 𑵐𑵉𑵃 𑵐𑵌𑵃 𑵐𑵍𑵃 𑵐𑵎𑵃 𑵐𑵏𑵃 𑵐𑵐𑵃 𑵐𑵑𑵃

These positional preferences are to be managed in the font. The NUKTA is used in encoded text as shown below. Note that the NUKTA is always placed after vowel signs in an encoded sequence:

ṛa 𑴘𑵂 <𑴘 DDA, ◌𑵂 NUKTA>

ṛā 𑴘𑵂𑴱 <𑴘 DDA, ◌𑵂 NUKTA, ◌𑴱 SIGN AA>

4.6 Halanta

The ◌𑵄 MASARAM GONDI SIGN HALANTA is proposed as a vowel silencer. It is used as follows:

k 𑴌𑵄 <𑴌 KA, ◌𑵄 HALANTA>

kh 𑴍𑵄 <𑴍 KHA, ◌𑵄 HALANTA>

The Gondi script as designed by Masaram does not have a native *halanta*, as the structure of the script does not require it. The embedding of the inherent vowel into the graphical structure of a consonant is a unique and innovative feature of the Gondi script. The horizontal stroke of each consonant letter represents the inherent vowel; removal of this stroke produces a bare consonant: 𑴌 *ka* → 𑴌𑵅 *k*, etc.

In most Indic scripts the inherent vowel is not part of the graphical structure of a consonant letter. As a result, these scripts require a mechanism for indicating the absence of the inherent vowel. In Devanagari this mechanism is a sign called ◌् U+094D DEVANAGARI SIGN VIRAMA: क *ka* + ◌् → क् *k*, etc.

However, modern users have adopted the Devanagari *halanta* (or VIRAMA) for marking a bare consonant at the end of a word (see figure 24). The GONDI SIGN HALANTA has been included in the repertoire in order to provide this functionality in the proposed encoding.

The proposed encoding for Gondi separates the two functions of the VIRAMA character as used in the models for most Indic scripts in Unicode. In Devanagari, for example, the VIRAMA functions both as a vowel silencer (*halanta*) and a control character for forming conjuncts. The default representation of VIRAMA in Devanagari is as a visible sign beneath the consonant with which it combines. If a consonant is placed after the VIRAMA it causes a conjunct to be formed from the two consonants around it. In order to display a visible VIRAMA between adjacent consonants, it is necessary to break conjunct formation. For this purpose, the control character [ZWNJ] U+200C ZERO WIDTH NON-JOINER is placed after VIRAMA. This approach was proposed for Gondi in L2/15-005. However, in the interest of developing a simple encoding model for Gondi, and to eliminate the need for usage of ZERO WIDTH NON-JOINER or other invisible control characters, the encoding of the vowel silencing feature of the Indic VIRAMA has been encapsulated into the character ◌𑵄 GONDI SIGN HALANTA. The conjunct forming function is retained in the GONDI VIRAMA, described below.

4.7 Virama

The ◌𑵅 MASARAM GONDI VIRAMA is a control character that is used specifically for producing the bare form of a consonant letter. It is represented in the code chart as [◌𑵅] in order to indicate that it is a special character. Conceptually, VIRAMA produces a half-form by removing the horizontal stroke from the glyph of the letter after which it is placed.

k 𑴌𑵅 <𑴌 KA, ◌𑵅 VIRAMA>

kh 𑴍𑵅 <𑴍 KHA, ◌𑵅 VIRAMA>

It is used for producing conjuncts, similar to the control function of ◌् U+094D DEVANAGARI SIGN VIRAMA. The Gondi VIRAMA, however, is not designed for silencing the inherent vowel; the ◌𑵄 GONDI SIGN HALANTA is to be used for that purpose.

4.8 Consonant Conjuncts

Consonant clusters are represented as conjuncts and are rendered in a linear sequence using half-forms of all but the final letter in a cluster, which appears in its full form, eg. 𑴀𑴀 *kka*, 𑴀𑴅 *kta*, 𑴀𑴅𑴊 *ktva*, etc. Consonants are placed sequentially in the conjunct in the order that they occur in the cluster.

Conjuncts are represented in encoded text by placing the control character 𑵇 MASARAM GONDI VIRAMA after each non-initial consonant in a cluster. The sequence <(*C*, 𑵇 VIRAMA)*, *C*> produces a half-form of all *C* that precede the VIRAMA, while the *C* that immediately follows is rendered using its regular full form:

ka	𑴀	<𑴀 KA>
kta	𑴀𑴅	<𑴀 KA, 𑵇 VIRAMA, 𑴅 TA>
ktva	𑴀𑴅𑴊	<𑴀 KA, 𑵇 VIRAMA, 𑴅 TA, 𑵇 VIRAMA, 𑴊 VA>
ktvya	𑴀𑴅𑴊𑴎	<𑴀 KA, 𑵇 VIRAMA, 𑴅 TA, 𑵇 VIRAMA, 𑴊 VA, 𑵇 VIRAMA, 𑴎 YA>

There are some exceptions to the rule of conjunct formation. The following character are proposed in order to properly represent all Gondi conjuncts:

𑵆	MASARAM GONDI REPHA
◌𑵃	MASARAM GONDI RA-KARA
𑵐	MASARAM GONDI LETTER KSSA
𑵑	MASARAM GONDI LETTER JNYA
𑵒	MASARAM GONDI LETTER TRA

The behavior of RA is described in section 4.8.1, and the use of atomic ligatures for three conjuncts is discussed in section 4.8.2.

4.8.1 Forms of RA in conjuncts

Following the general rule of conjunct formation, the letter 𑴓 RA would occur as the half-form 𑵃 when it is initial or medial in a cluster and in its regular full-form when in cluster-final position. The Gondi RA, however, does not behave entirely as expected in conjuncts and is rendered in several ways. There are three ways of writing 𑴓 RA in conjuncts. It occurs in its half-form 𑵃 when cluster-initial or alternately as ◌𑵆 when cluster-initial and as ◌𑵃 when cluster-final. These are described below:

- *Half-form* The half-form of RA is used specifically for representing semantic distinctions of RA when the letter occurs at a morphological boundary. Such usage is influenced by Devanagari orthography for the Marathi language, in which र RA may occur as either the ◌् regular *repha* or the ◌= 'eyelash' *repha* when it is the initial consonant in a cluster; the 'eyelash' *repha* marks plural suffixes (दऱ्या *daryā* 'valleys' and दर्या *daryā* 'ocean') and inflectional suffixes (आचाऱ्यास *ācāryās* 'to the cook' and आचार्यास *ācāryās* 'to the teacher'). The half-form 𑵃 corresponds to the Devanagari ◌= 'eyelash' *repha*. It is also used when RA occurs in cluster-medial position.

- *Repha* When not used for marking morphological distinctions, cluster-initial RA is rendered as the ◌𑵆 *repha*. The logic of this character is based upon the Devanagari ◌् regular *repha*. The Gondi *repha*

attaches after the last letter in a conjunct, above the horizontal line or an extension of it, depending upon the presence of an above-stroke vowel sign, eg. 𑵆𑴌 *rka*, 𑵆𑴌𑴱 *rkā*, etc.

𑵆𑴌 𑵆𑴌𑴱 𑵆𑴌𑴲 𑵆𑴌𑴳 𑵆𑴌𑴴 𑵆𑴌𑴵 𑵆𑴌𑴶 𑵆𑴌𑴺 𑵆𑴌𑴼 𑵆𑴌𑴽 𑵆𑴌𑴿

Some modern sources show the ◌𑵆 *repha* represented using the form ◌𑵆, eg. 𑵆𑴌 *rka*. This form is simply the regular sign for ◌𑵆 *repha* with the left stroke drawn past the horizontal bar and curving to the right. It is a glyphic variant and is to be handled by the font.

- *Ra-kāra* In the current orthography, cluster-final RA is rendered as ◌𑵇 *ra-kāra* instead of as the full form 𑴦. The logic of the Gondi *ra-kāra* is based upon the Devanagari ◌्र *ra-kāra* and ◌्र *vattu*. The Gondi *ra-kāra* is positioned below the horizontal line of a consonant glyph or beneath an extension of the horizontal line:

  𑴌𑵇 𑴌𑵇𑴱 𑴌𑵇𑴲 𑴌𑵇𑴳 𑴌𑵇𑴴 𑴌𑵇𑴵 𑴌𑵇𑴶 𑴌𑵇𑴺 𑴌𑵇𑴼 𑴌𑵇𑴽 𑴌𑵇𑴿

  Some modern sources show the ◌𑵇 *ra-kāra* represented using the form ◌𑵇, eg. 𑴌𑵇 *kra*. This form is a glyphic variant and is to be handled by the font.

- *Full-form* The full-form of RA is rarely used at present when it is final in a cluster: 𑴌𑵄𑴦 *kra*. The preference is to use ◌𑵇 *ra-kāra*.

The representation of ◌𑵆 *repha* and ◌𑵇 *ra-kāra* requires an exception to the rule of conjunct formation in Gondi. The general rule states that the sequence <RA, VIRAMA, *C*> is rendered using the half-form 𑴦𑵄 of RA and the full-form of *C*. It also states that <*C*, VIRAMA, RA> would produce the half-form of *C* and the full-form 𑴦 of RA. For this reason another method is required for the encoded representation of *repha* and *ra-kāra*, for which the expected encoded sequences according to the general model of Indic scripts would also be <RA, VIRAMA, *C*> and <*C*, VIRAMA, RA>, respectively.

There are four possible models for accommodating encoded representations of RA in conjuncts. The first two are based upon the premise that the default behavior of RA in conjuncts is similar to that of all other consonants, ie. it is rendered using the half-form 𑴦𑵄 when cluster-initial and the full-form 𑴦 when cluster-final. These approaches treat *repha* and *ra-kāra* as exceptions. The third model deviates from this premise and establishes the *repha* and *ra-kāra* as default representations of RA in conjuncts, and the half-form and full-form as exceptions.

1. *Use the Zero-Width Joiner*

   In L2/15-005, it was suggested that the generic control character ZWJ U+200D ZERO WIDTH JOINER (ZWJ) be used for representing *repha* and *ra-kāra* in encoded text. The usage of ZWJ was chosen because the character is used in various Indic scripts for controlling different forms of letters in conjuncts. The same principle was applied to Gondi, such that various forms of RA in conjuncts would be produced as follows:

half-form	𑴦𑵄	<𑴦 RA, ◌𑵄 VIRAMA, *C*>
repha	◌𑵆	<𑴦 RA, ZWJ ZERO WIDTH JOINER, ◌𑵄 VIRAMA, *C*>
ra-kāra	◌𑵇	<*C*, ◌𑵄 VIRAMA, ZWJ ZERO WIDTH JOINER, 𑴦 RA>

Representation of approach #1 in encoded sequences would be as follows:

𑵇𑵊	*rka*	<𑵇 RA, 𑵅 VIRAMA, 𑵊 KA>
𑵊𑵇	*kra*	<𑵊 KA, 𑵅 VIRAMA, 𑵇 RA>
𑵊‍	*rka*	<𑵇 RA, ZWJ ZERO WIDTH JOINER, 𑵅 VIRAMA, 𑵊 KA>
𑵊‍	*kra*	<𑵊 KA, 𑵅 VIRAMA, ZWJ ZERO WIDTH JOINER, 𑵇 RA>

The Script Ad-hoc Committee recommended against the usage of ZWJ for such cases in Gondi. In L2/15-045, the Subcommittee suggested that an alternative model be developed and that "[f]or *repha* and *ra-kāra*, encoding separate characters may be useful, similar to the Malayalam *dot reph* and the *medial ra* in Myanmar and Tai Tham." The major concern is that usage of ZWJ introduces issues for both implementers and end users. For implementers, the use of ZWJ overloads the script-specific rules that must be tailored for each script. For end-users, ZWJ is problematic because it is an invisible control character and is neither readily available on keyboards or easily detectable in text sequence.

2. *Redefine the rule for rendering <RA, VIRAMA>*

The above model is based on the general rule of conjunct formation in Gondi: <*C*, VIRAMA> produces the half-form of *C* and therefore <RA, VIRAMA> should produce the half-form of RA. The *repha* and *ra-kāra* are treated as exceptions to the rule. A third approach for handling the different forms of RA in conjuncts is to redefine the default rendering for RA. With this approach the various forms of RA in conjuncts would be produced in a fashion similar to other Indic scripts:

repha	◌𑵊	<𑵇 RA, 𑵅 VIRAMA, *C*>
ra-kāra	◌𑵈	<*C*, 𑵅 VIRAMA, 𑵇 RA>

Representation of approach #2 in encoded sequences would be as follows:

𑵊‍	*rka*	<𑵇 RA, 𑵅 VIRAMA, 𑵊 KA>
𑵊‍	*kra*	<𑵊 KA, 𑵅 VIRAMA, 𑵇 RA>

This approach, however, does not provide a means for producing the half-form 𑵇 of RA, or even the cluster-final full-form if ever needed. Some mechanism would be required to produce these forms. Inevitably, a control character such as ZWJ would be required:

𑵇𑵊	*rka*	<𑵇 RA, ZWJ ZERO WIDTH JOINER, 𑵅 VIRAMA, 𑵊 KA>
𑵊𑵇	*kra*	<𑵊 KA, 𑵅 VIRAMA, ZWJ ZERO WIDTH JOINER, 𑵇 RA>

This approach is essentially the converse of what was proposed in L2/15-005. It poses the same issues as it requires usage of ZWJ.

3. *Encode a ligating form of RA*

Another option is to encode a dummy letter whose cluster-initial form is ◌𑵊 *repha* and whose non-initial form is ◌𑵈 *ra-kāra*. This letter might be called LIGATING RA and it would be defined for usage

only in conjuncts. Another possible name might be RRA, but this name is generally used in Indic scripts in Unicode for the Dravidian *ra* and as Gondi is a Dravidian language, this name is reserved for potentially encoding a such a character for the Gondi script. The representative glyph for this letter might be 𑵐, which is the letter 𑵋 RA placed within a dashed box that indicates that its function as a control character. The sequence <LIGATING RA, VIRAMA, *C*> would produce *repha*, while <*C*, VIRAMA, LIGATING RA> would produce *ra-kāra*. By extension, it does not have a half-form. The letter would be rendered as 𑵐 when used in any other context. Using this character the various forms of RA would be produced as follows:

𑵋𑵆𑵎	*rka*	<𑵋 RA, 𑵆 VIRAMA, 𑵎 KA>
𑵎𑵆𑵋	*kra*	<𑵎 KA, 𑵆 VIRAMA, 𑵋 RA>
𑵎𑵰	*rka*	<𑵐 LIGATING RA, 𑵆 VIRAMA, 𑵎 KA>
𑵎𑵱	*kra*	<𑵎 KA, 𑵆 VIRAMA, 𑵐 LIGATING RA>

Although LIGATING RA is an artificial letter and does not occur in the Gondi script, it facilitates the representation of *repha* and *ra-kāra* according to the general rules of conjunct formation, while also providing for the default rendering of RA as half-form and full-form in conjuncts:

𑵎𑵰	*rkra*	<𑵐 LIGATING RA, 𑵆 VIRAMA, 𑵎 KA, 𑵆 VIRAMA, 𑵐 LIGATING RA>
𑵋𑵎𑵆𑵋	*rkra*	<𑵋 RA, 𑵆 VIRAMA, 𑵎 KA, 𑵆 VIRAMA, 𑵋 RA>

This approach also aligns with the concept that *repha* and *ra-kāra* in Devanagari and other scripts are special forms of a RA character. The approach that utilizes LIGATING RA offers users with a clear method for producing normative and special forms of RA in conjuncts using VIRAMA, which follows the general rule for the encoded representation of conjuncts in Gondi.

4. *Encode 'repha' and 'ra-kāra' as separate characters*

In L2/15-045, the Script Ad-hoc Committee recommended that ◌𑵊 *repha* and ◌𑵌 *ra-kāra* be encoded as separate characters. While this approach provides the easiest means for encoding different forms of RA, it also raises several questions. Where should a *repha* character occur in an encoded sequence? Should it be handled logically and placed at the position where RA would normally occur, with the expectation that the rendering engine would reorder it to the end of the conjunct? Or should it be handled visually, and be placed manually at the end of the conjunct in an encoded sequence? For example, would the encoded representation for the hypothetical conjunct 𑵎𑵰 *rkra* be <◌𑵊 *repha*, 𑵎 KA, ◌𑵌 *ra-kāra*> or <𑵎 KA, ◌𑵌 *ra-kāra*, ◌𑵊 *repha*>? Another issue concerns the classification of *repha*: is it to be considered a letter (𑵊) or a sign (◌𑵊)? If *repha* is a letter, then is VIRAMA to be placed between it and the following consonant: <𑵊 *repha*, 𑵆 VIRAMA, 𑵎 KA, ◌𑵌 *ra-kāra*>?

This spproach requires the encoding of the following two characters:

𑵊	REPHA
◌𑵌	RA-KARA

The most feasible of the above four is approach #4. Representing cluster-specific forms of RA as separate characters offers a simpler implementation that does not require usage of control characters. For this reason,

the suggestion made by the Script Ad-hoc Committee in In L2/15-045 has been adopted. Furthermore, this approach requires the following considerations:

1. The REPHA is defined a letter, not a combining sign. In the code chart the ◌ REPHA is placed within a dashed box as ◌ in order to indicate that it requires special rendering. The glyph ◌ is used in the output.

2. The logical model for REPHA is preferred over the visual model, ie. typing the character at the beginning of a cluster instead of typing it at the end of a cluster and after any accompanying vowel signs. This approach provides the ability to type syllables according to the underlying phonology. The logical model requires that REPHA be placed at the same position in the encoded sequence that it occurs in phonetic expression. The rendering engine will re-order the REPHA to the end of the conjunct after any accompanying vowel signs.

3. The RA-KARA is a regular combining sign. It occurs logically in the encoded sequence before any vowel sign that may accompany the conjunct. When it occurs after NUKTA, the RA-KARA is written beneath an extension of the horizontal stroke. If HALANTA occurs after RA-KARA, then it is placed under an extension of the horizontal stroke of the base letter.

4. The REPHA or RA-KARA do not interact with VIRAMA Pairing these characters with VIRAMA will result in the visible display of VIRAMA with a dotted circle:

   <◌ REPHA, ◌ VIRAMA, 𑵐 KA> → ◌ 𑵐

   <𑵐 KA, ◌ VIRAMA, ◌ RA-KARA> → 𑵐 ◌ ◌

5. The REPHA does not interact with any combining signs. Pairing it with such signs will result in unrendered sequences:

   <◌ REPHA, ◌ VOWEL SIGN AA> → ◌ ◌

   <◌ REPHA, ◌ HALANTA> → ◌ ◌

   <◌ REPHA, ◌ NUKTA> → ◌ ◌

The REPHA is used in encoded sequences as shown below:

𑵐 *rka* <◌ REPHA, 𑵐 KA>

𑵐 *rkā* <◌ REPHA, 𑵐 KA, ◌ VOWEL SIGN AA>

𑵐 *rḳa* <◌ REPHA, 𑵐 KA, ◌ NUKTA>

The RA-KARA is used in encoded sequences as shown below:

𑵐 *kra* <𑵐 KA, ◌ RA-KARA>

𑵐 *krā* <𑵐 KA, ◌ RA-KARA, ◌ VOWEL SIGN AA>

𑵐 *ḳra* <𑵐 KA, ◌ NUKTA, ◌ RA-KARA>

𑵐 *kr* <𑵐 KA, ◌ RA-KARA, ◌ HALANTA>

The usage of VIRAMA after and before RA would produce the expected output:

𑴦𑵅𑴌 *rka* <𑴦 RA, ◌𑵅 VIRAMA, 𑴌 KA>

𑴌𑵅𑴦 *kra* <𑴌 KA, ◌𑵅 VIRAMA, 𑴦 RA>

4.8.2 Conjunct letters

The consonant clusters *kṣa*, *jña*, *tra* are represented not as regular conjuncts, but as distinctive letters. These are proposed for encoding as atomic letters:

𑴮 MASARAM GONDI LETTER KSSA

𑴯 MASARAM GONDI LETTER JNYA

𑴰 MASARAM GONDI LETTER TRA

Following the rules of conjunct formation, the expected representation of these three conjuncts would be:

kṣa 𑴌𑵅𑴪 <𑴌 KA, ◌𑵅 VIRAMA, 𑴪 SSA>

jña 𑴓𑵅𑴕 <𑴓 JA, ◌𑵅 VIRAMA, 𑴕 NYA>

tra 𑴛𑵅𑴦 <𑴛 TA, ◌𑵅 VIRAMA, 𑴦 RA>

In the Gondi script, each of these three letters phonetically represent a consonant cluster, but they all have the structure of an atomic letter. These forms are encoded as consonant letters because in all cases consonant conjuncts are written as linear sequences of half-forms, not as ligatures. While in most Indic scripts the written forms for *kṣa*, *jña*, *tra* have encoded representations as a character sequence, such an approach would not be consistent with this script.

It is evident that these ligatures were developed because distinctive forms exist in Devanagari. These three conjuncts are often shown at the end of Devanagari orthographies for various languages and are often interpreted by users as being distinctive letters that are fundamental elements of the script.

4.8.3 Rendering of conjuncts

As specified at the outset of this section, the sequence <(*C*, ◌𑵅 VIRAMA)*, *C*> produces a half-form of all *C* and the regular full-form of *C*. In order for the rendering of conjuncts to operate properly, the font must contain a complete set of consonants half-forms (glyphs without the horizontal stroke). The font should substitute each instance of <(*C*, ◌𑵅 VIRAMA)> with the appropriate half-form for *C*:

<𑴌 KA, ◌𑵅 VIRAMA, 𑴛 TA> → 𑴌𑵅𑴛

If this glyph is not available, then the output should show the full form of *C* followed by ◌𑵅 VIRAMA.

<𑴌 KA, ◌𑵅 VIRAMA, 𑴛 TA> → 𑴌◌𑵅𑴛

4.9 Digits

The script has a full set of digits:

𑶠	MASARAM GONDI DIGIT ZERO	𑶥	MASARAM GONDI DIGIT FIVE
𑶡	MASARAM GONDI DIGIT ONE	𑶦	MASARAM GONDI DIGIT SIX
𑶢	MASARAM GONDI DIGIT TWO	𑶧	MASARAM GONDI DIGIT SEVEN
𑶣	MASARAM GONDI DIGIT THREE	𑶨	MASARAM GONDI DIGIT EIGHT
𑶤	MASARAM GONDI DIGIT FOUR	𑶩	MASARAM GONDI DIGIT NINE

Variant forms of digits are shown in figure 18.

4.10 Punctuation

Script-specific punctuation is not attested in Masaram's Gondi script. The *daṇḍā* and double *daṇḍā* are commonly used. These signs of punctuation are not proposed for separate inclusion in the 'Masaram Gondi' block, but are to be unified with । U+0964 DEVANAGARI DANDA and ॥ U+0965 DEVANAGARI DOUBLE DANDA. Latin marks of punctuation, such as periods, are also used.

4.11 Collation

Gondi is sorted according to the general pattern for Brahmi-based scripts. Note that the following characters are sorted as:

𑶊 REPHA << 𑶓 RA << ◌𑶖 RA-KARA

4.12 Characters Not Proposed for Encoding

The following are newly-invented characters. Actual usage of these characters, apart from their inclusion in new charts of the script, is unknown. For this reason, they are not proposed for encoding at present.

Additional vocalic vowel signs The original script provided for the writing of only one Sanskrit vocalic vowel sign: ◌𑶊 VOWEL SIGN VOCALIC R. In order to accommodate anticipated usage of the script for linguistics research, Mukund Gokhale designed signs for the other three vocalic sounds:

◌𑶘 *GONDI VOWEL SIGN VOCALIC RR

◌𑶙 *GONDI VOWEL SIGN VOCALIC L

◌𑶚 *GONDI VOWEL SIGN VOCALIC LL

The sounds represented by these four vowel signs are not found natively in Gondi. They occur only in Sanskrit words, and only ◌𑶊 VOWEL SIGN VOCALIC R is attested in usage. There is no corresponding independent letter for these signs, and in such contexts they are represented using a consonant-vowel combination composed with the letters RA and LA and the dependent sign for the vocalic letter:

r̥ 𑶓𑶊 <𑶓 RA, ◌𑶊 VOWEL SIGN VOCALIC R>

*r̥̄ 𑴦◌ <𑴦 RA, ◌ *VOWEL SIGN VOCALIC RR>

*l̥ 𑴧◌ <𑴧 LA, ◌ *VOWEL SIGN VOCALIC L>

*l̥̄ 𑴧◌ <𑴧 LA, ◌ *VOWEL SIGN VOCALIC LL>

The VOWEL SIGN VOCALIC R is included in the proposed repertoire, but the other three signs are not proposed for inclusion at present because their actual usage is not attested. Space has been reserved for these characters in the block in the event that a justifiable case to encode them is made in the future.

5 Character Data

5.1 Character Properties

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

```
11D00;MASARAM GONDI LETTER A;Lo;0;L;;;;;N;;;;;
11D01;MASARAM GONDI LETTER AA;Lo;0;L;;;;;N;;;;;
11D02;MASARAM GONDI LETTER I;Lo;0;L;;;;;N;;;;;
11D03;MASARAM GONDI LETTER II;Lo;0;L;;;;;N;;;;;
11D04;MASARAM GONDI LETTER U;Lo;0;L;;;;;N;;;;;
11D05;MASARAM GONDI LETTER UU;Lo;0;L;;;;;N;;;;;
11D06;MASARAM GONDI LETTER E;Lo;0;L;;;;;N;;;;;
11D08;MASARAM GONDI LETTER AI;Lo;0;L;;;;;N;;;;;
11D09;MASARAM GONDI LETTER O;Lo;0;L;;;;;N;;;;;
11D0B;MASARAM GONDI LETTER AU;Lo;0;L;;;;;N;;;;;
11D0C;MASARAM GONDI LETTER KA;Lo;0;L;;;;;N;;;;;
11D0D;MASARAM GONDI LETTER KHA;Lo;0;L;;;;;N;;;;;
11D0E;MASARAM GONDI LETTER GA;Lo;0;L;;;;;N;;;;;
11D0F;MASARAM GONDI LETTER GHA;Lo;0;L;;;;;N;;;;;
11D10;MASARAM GONDI LETTER NGA;Lo;0;L;;;;;N;;;;;
11D11;MASARAM GONDI LETTER CA;Lo;0;L;;;;;N;;;;;
11D12;MASARAM GONDI LETTER CHA;Lo;0;L;;;;;N;;;;;
11D13;MASARAM GONDI LETTER JA;Lo;0;L;;;;;N;;;;;
11D14;MASARAM GONDI LETTER JHA;Lo;0;L;;;;;N;;;;;
11D15;MASARAM GONDI LETTER NYA;Lo;0;L;;;;;N;;;;;
11D16;MASARAM GONDI LETTER TTA;Lo;0;L;;;;;N;;;;;
11D17;MASARAM GONDI LETTER TTHA;Lo;0;L;;;;;N;;;;;
11D18;MASARAM GONDI LETTER DDA;Lo;0;L;;;;;N;;;;;
11D19;MASARAM GONDI LETTER DDHA;Lo;0;L;;;;;N;;;;;
11D1A;MASARAM GONDI LETTER NNA;Lo;0;L;;;;;N;;;;;
11D1B;MASARAM GONDI LETTER TA;Lo;0;L;;;;;N;;;;;
11D1C;MASARAM GONDI LETTER THA;Lo;0;L;;;;;N;;;;;
11D1D;MASARAM GONDI LETTER DA;Lo;0;L;;;;;N;;;;;
11D1E;MASARAM GONDI LETTER DHA;Lo;0;L;;;;;N;;;;;
11D1F;MASARAM GONDI LETTER NA;Lo;0;L;;;;;N;;;;;
11D20;MASARAM GONDI LETTER PA;Lo;0;L;;;;;N;;;;;
11D21;MASARAM GONDI LETTER PHA;Lo;0;L;;;;;N;;;;;
11D22;MASARAM GONDI LETTER BA;Lo;0;L;;;;;N;;;;;
11D23;MASARAM GONDI LETTER BHA;Lo;0;L;;;;;N;;;;;
11D24;MASARAM GONDI LETTER MA;Lo;0;L;;;;;N;;;;;
11D25;MASARAM GONDI LETTER YA;Lo;0;L;;;;;N;;;;;
11D26;MASARAM GONDI LETTER RA;Lo;0;L;;;;;N;;;;;
11D27;MASARAM GONDI LETTER LA;Lo;0;L;;;;;N;;;;;
11D28;MASARAM GONDI LETTER VA;Lo;0;L;;;;;N;;;;;
11D29;MASARAM GONDI LETTER SHA;Lo;0;L;;;;;N;;;;;
11D2A;MASARAM GONDI LETTER SSA;Lo;0;L;;;;;N;;;;;
```

```
11D2B;MASARAM GONDI LETTER SA;Lo;0;L;;;;;N;;;;;
11D2C;MASARAM GONDI LETTER HA;Lo;0;L;;;;;N;;;;;
11D2D;MASARAM GONDI LETTER LLA;Lo;0;L;;;;;N;;;;;
11D2E;MASARAM GONDI LETTER KSSA;Lo;0;L;;;;;N;;;;;
11D2F;MASARAM GONDI LETTER JNYA;Lo;0;L;;;;;N;;;;;
11D30;MASARAM GONDI LETTER TRA;Lo;0;L;;;;;N;;;;;
11D31;MASARAM GONDI VOWEL SIGN AA;Mn;0;NSM;;;;;N;;;;;
11D32;MASARAM GONDI VOWEL SIGN I;Mn;0;NSM;;;;;N;;;;;
11D33;MASARAM GONDI VOWEL SIGN II;Mn;0;NSM;;;;;N;;;;;
11D34;MASARAM GONDI VOWEL SIGN U;Mn;0;NSM;;;;;N;;;;;
11D35;MASARAM GONDI VOWEL SIGN UU;Mn;0;NSM;;;;;N;;;;;
11D36;MASARAM GONDI VOWEL SIGN VOCALIC R;Mn;0;NSM;;;;;N;;;;;
11D3A;MASARAM GONDI VOWEL SIGN E;Mn;0;NSM;;;;;N;;;;;
11D3C;MASARAM GONDI VOWEL SIGN AI;Mn;0;NSM;;;;;N;;;;;
11D3D;MASARAM GONDI VOWEL SIGN O;Mn;0;NSM;;;;;N;;;;;
11D3F;MASARAM GONDI VOWEL SIGN AU;Mn;0;NSM;;;;;N;;;;;
11D40;MASARAM GONDI SIGN ANUSVARA;Mn;0;NSM;;;;;N;;;;;
11D41;MASARAM GONDI SIGN VISARGA;Mn;0;NSM;;;;;N;;;;;
11D42;MASARAM GONDI SIGN NUKTA;Mn;7;NSM;;;;;N;;;;;
11D43;MASARAM GONDI SIGN CANDRA;Mn;0;NSM;;;;;N;;;;;
11D44;MASARAM GONDI SIGN HALANTA;Mn;9;NSM;;;;;N;;;;;
11D45;MASARAM GONDI VIRAMA;Mn;9;NSM;;;;;N;;;;;
11D46;MASARAM GONDI REPHA;Lo;0;L;;;;;N;;;;;
11D47;MASARAM GONDI RA-KARA;Mn;0;NSM;;;;;N;;;;;
11D50;MASARAM GONDI DIGIT ZERO;Nd;0;L;;0;0;0;N;;;;;
11D51;MASARAM GONDI DIGIT ONE;Nd;0;L;;1;1;1;N;;;;;
11D52;MASARAM GONDI DIGIT TWO;Nd;0;L;;2;2;2;N;;;;;
11D53;MASARAM GONDI DIGIT THREE;Nd;0;L;;3;3;3;N;;;;;
11D54;MASARAM GONDI DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;;
11D55;MASARAM GONDI DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;;
11D56;MASARAM GONDI DIGIT SIX;Nd;0;L;;6;6;6;N;;;;;
11D57;MASARAM GONDI DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;;
11D58;MASARAM GONDI DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;;
11D59;MASARAM GONDI DIGIT NINE;Nd;0;L;;9;9;9;N;;;;;
```

5.2 Linebreaking

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

```
11D00..11D06; AL   # MASARAM GONDI LETTER A .. MASARAM GONDI LETTER E
11D08..11D09; AL   # MASARAM GONDI LETTER AI .. MASARAM GONDI LETTER O
11D0B..11D30; AL   # MASARAM GONDI LETTER AU .. MASARAM GONDI LETTER TRA
11D31..11D36; CM   # MASARAM GONDI SIGN AA .. MASARAM GONDI SIGN VOCALIC R
11D3A..; CM        # MASARAM GONDI VOWEL SIGN E
11D3C..11D3D; CM   # MASARAM GONDI SIGN AI .. MASARAM GONDI VOWEL SIGN O
11D3F; CM          # MASARAM GONDI VOWEL SIGN AU
11D40..11D44; CM   # MASARAM GONDI SIGN ANUSVARA .. MASARAM GONDI SIGN HALANTA
11D45; CM          # MASARAM GONDI VIRAMA
11D36; AL          # MASARAM GONDI REPHA
11D47; CM          # MASARAM GONDI RA-KARA
11D50..11D59; NU   # MASARAM GONDI DIGIT ZERO .. MASARAM GONDI DIGIT NINE
```

5.3 Syllabic Categories

Syllabic categories given in the format of `IndicSyllabicCategory.txt`:

```
# Indic_Syllabic_Category=Bindu
11D40         ; Bindu               # Mn        MASARAM GONDI SIGN ANUSVARA
```

```
# Indic_Syllabic_Category=Visarga
11D41          ; Visarga              # Mc        MASARAM GONDI SIGN VISARGA

# Indic_Syllabic_Category=Nukta
11D42          ; Nukta                # Mn        MASARAM GONDI SIGN NUKTA

# Indic_Syllabic_Category=Virama
11D45          ; Virama               # Mn        MASARAM GONDI VIRAMA

# Indic_Syllabic_Category=Pure_Killer
11D44          ; Pure_Killer          # Mn        MASARAM GONDI SIGN HALANTA

# Indic_Syllabic_Category=Vowel_Independent
11D00..11D0B   ; Vowel_Independent    # Lo   [10] MASARAM GONDI LETTER A .. AU

# Indic_Syllabic_Category=Vowel_Dependent
11D31..11D36   ; Vowel_Dependent      # Mn    [6] MASARAM GONDI VOWEL SIGN AA .. VOCALIC R
11D3A          ; Vowel_Dependent      # Mn        MASARAM GONDI VOWEL SIGN E
11D3C..11D3D   ; Vowel_Dependent      # Mn    [2] MASARAM GONDI VOWEL SIGN AI .. O
11D3F          ; Vowel_Dependent      # Mn        MASARAM GONDI VOWEL SIGN AU
11D43          ; Vowel_Dependent      # Mn        MASARAM GONDI SIGN CANDRA

# Indic_Syllabic_Category=Consonant
11D0C..11D30   ; Consonant            # Lo   [40] MASARAM GONDI LETTER KA .. TRA

# Indic_Syllabic_Category=Consonant_Preceding_Repha
11D46          ; Consonant_Preceding_Repha  # Lo  MASARAM GONDI REPHA

# Indic_Syllabic_Category=Consonant_Medial
11D47          ; Consonant_Medial     # Mn       MASARAM GONDI RA-KARA
```

5.4 Positional Categories

Positional data for Masaram Gondi combining signs in the format of `IndicPositionalCategory.txt`:

```
# Indic_Positional_Category=Top
11D31..11D35   ; Top     # Mn    [5] MASARAM GONDI VOWEL SIGN AA .. UU
11D3A          ; Top     # Mn        MASARAM GONDI VOWEL SIGN E
11D3C..11D3D   ; Top     # Mn    [2] MASARAM GONDI VOWEL SIGN AI .. O
11D3F          ; Top     # Mn        MASARAM GONDI VOWEL SIGN AU
11D40          ; Top     # Mn        MASARAM GONDI SIGN ANUSVARA
11D41          ; Top     # Mn        MASARAM GONDI SIGN VISARGA
11D43          ; Top     # Mn        MASARAM GONDI SIGN CANDRA

# Indic_Positional_Category=Bottom
11D36          ; Bottom  # Mn        MASARAM GONDI VOWEL SIGN VOCALIC R
11D42          ; Bottom  # Mn        MASARAM GONDI SIGN NUKTA
11D44          ; Bottom  # Mn        MASARAM GONDI SIGN HALANTA
11D47          ; Bottom  # Mn        MASARAM GONDI RA-KARA
```

5.5 Script Extensions

The following characters should be extended to Masaram Gondi in `ScriptExtensions.txt`:

```
0964          ;  # Po       DEVANAGARI DANDA
0965          ;  # Po       DEVANAGARI DOUBLE DANDA
```

5.6 'Confusable' Characters

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

```
11D31 MASARAM GONDI VOWEL SIGN AA  ;  0304 COMBINING MACRON
11D21 MASARAM GONDI LETTER PHA     ;  1109D KAITHI LETTER NNA
11D2A MASARAM GONDI LETTER SSA     ;  0398 GREEK CAPITAL LETTER THETA
11D52 MASARAM GONDI DIGIT TWO      ;  0055 LATIN CAPITAL LETTER U
```

6 References

Anderson, Deborah; et al. 2015. "Recommendations to UTC #142 February 2015 on Script Proposals" (L2/15-045). http://www.unicode.org/L2/L2015/15045-script-rec.pdf

गुरूजी, मनीराम दुर्गा [Gurūjī, Manīrāma Durgā]. गोंदी लम्क पुंदान [*Goṃḍī lamka puṃdāna*].

मण्डाले, सीताराम [Maṇḍāle, Sītārām]. कोयाबोली [*Koyābolī*]. गोंडी शब्द संग्रह - गोंडी, मराठी, हिन्दी [Goṃḍī Śabda Saṃgraha - Goṃḍī, Marāṭhī, Hindī].

Masaram, Bhava Singh. 1951. "गोंडी लिपि" [Goṃḍī lipi]. Central Institute of Indian Languages, Multimedia library, photograph no. 64.

Pandey, Anshuman. 2010. "Preliminary Proposal to Encode the Gondi Script in the UCS" (L2/10-207). http://www.unicode.org/L2/L2010/10207-gondi.pdf

———. 2012. "Revised Preliminary Proposal to Encode the Gondi Script" (L2/12-235). http://www.unicode.org/L2/L2012/12235-n4291-gondi.pdf

———. 2015a. "Proposal to Encode the Gondi Script" (L2/15-005). http://www.unicode.org/L2/L2015/15005-gondi.pdf

———. 2015b. "Preliminary Proposal to Encode the Gunjala Gondi Script in Unicode" (L2/15-086). http://www.unicode.org/L2/L2015/15086-gunjala-gondi.pdf

———. 2015c. "Proposal to Encode the 'Parsapen' Symbol in Unicode" (L2/15-111) http://www.unicode.org/L2/L2015/15111-parsapen-symbol.pdf

रामानन्द [Rāmānanda]. गोंडी अक्षर ज्ञान [Goṃḍī Akṣara Jñāna].

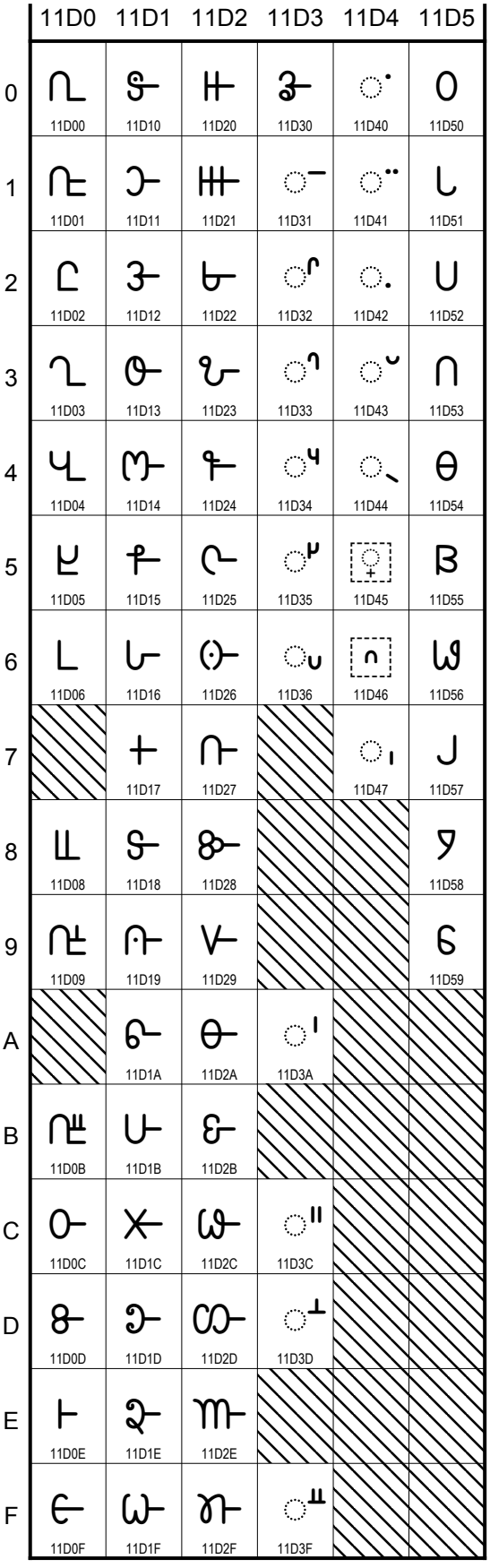
Vahia, M. N.; Ganesh Halkare. 2013. "Aspects of Gondi Astronomy". *Journal of Astronomical History and Heritage*, vol. 16, no. 1, pp. 29–44. http://www.tifr.res.in/~archaeo/VahiaGondTrip/JAHHGondsfinal.pdf

7 Acknowledgments

I am deeply indebted to Mukund Gokhale (Script Research Institute, Pune) for providing me with numerous materials in Masaram's Gondi script and for patiently answering my numerous questions about the script over the past few years. Gokhale's efforts in developing typefaces for this Gondi script have helped to address issues regarding the interaction of multiple glyphs with a base letter. I also appreciate him for contacting Motiravan Kangle (Akhil-Gondwana Gondi Sahitya Parishad, Nagpur) with my questions. I thank Kangle for indulging my inquiries and for providing examples of current orthography of *ra* in conjuncts, and the usage of *halanta* and *candra*. The Gondi chart shown in figure 1 was provided by B. A. Sharada and Suman

Kumari of the Central Institute of Indian Languages (Mysore). Mark Penny provided the chart shown in figure 6. I am also thankful to Roozbeh Pournader (Google) and Andrew Glass (Microsoft) for sharing their insights regarding effective models for representing the Gondi RA in conjuncts.

This project was made possible in part through a Google Research Award, granted to Deborah Anderson for the Script Encoding Initiative, and a grant from the United States National Endowment for the Humanities (PR-50205-15), which funds 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 Google or the National Endowment for the Humanities.


11D0 11D1 11D2 11D3 11D4 11D5
0 11D00 11D10 11D20 11D30 11D40 11D50
1 11D01 11D11 11D21 11D31 11D41 11D51
2 11D02 11D12 11D22 11D32 11D42 11D52
3 11D03 11D13 11D23 11D33 11D43 11D53
4 11D04 11D14 11D24 11D34 11D44 11D54
5 11D05 11D15 11D25 11D35 11D45 11D55
6 11D06 11D16 11D26 11D36 11D46 11D56
7 11D17 11D27 11D47 11D57
8 11D08 11D18 11D28 11D58
9 11D09 11D19 11D29 11D59
A 11D1A 11D2A 11D3A
B 11D0B 11D1B 11D2B
C 11D0C 11D1C 11D2C 11D3C
D 11D0D 11D1D 11D2D 11D3D
E 11D0E 11D1E 11D2E
F 11D0F 11D1F 11D2F 11D3F

Vowels

11D00 MASARAM GONDI LETTER A
11D01 MASARAM GONDI LETTER AA
11D02 MASARAM GONDI LETTER I
11D03 MASARAM GONDI LETTER II
11D04 MASARAM GONDI LETTER U
11D05 MASARAM GONDI LETTER UU
11D06 MASARAM GONDI LETTER E
11D07 <reserved>
11D08 MASARAM GONDI LETTER AI
11D09 MASARAM GONDI LETTER O
11D0A <reserved>
11D0B MASARAM GONDI LETTER AU

Consonants

11D0C MASARAM GONDI LETTER KA
11D0D MASARAM GONDI LETTER KHA
11D0E MASARAM GONDI LETTER GA
11D0F MASARAM GONDI LETTER GHA
11D10 MASARAM GONDI LETTER NGA
11D11 MASARAM GONDI LETTER CA
11D12 MASARAM GONDI LETTER CHA
11D13 MASARAM GONDI LETTER JA
11D14 MASARAM GONDI LETTER JHA
11D15 MASARAM GONDI LETTER NYA
11D16 MASARAM GONDI LETTER TTA
11D17 MASARAM GONDI LETTER TTHA
11D18 MASARAM GONDI LETTER DDA
11D19 MASARAM GONDI LETTER DDHA
11D1A MASARAM GONDI LETTER NNA
11D1B MASARAM GONDI LETTER TA
11D1C MASARAM GONDI LETTER THA
11D1D MASARAM GONDI LETTER DA
11D1E MASARAM GONDI LETTER DHA
11D1F MASARAM GONDI LETTER NA
11D20 MASARAM GONDI LETTER PA
11D21 MASARAM GONDI LETTER PHA
11D22 MASARAM GONDI LETTER BA
11D23 MASARAM GONDI LETTER BHA
11D24 MASARAM GONDI LETTER MA
11D25 MASARAM GONDI LETTER YA
11D26 MASARAM GONDI LETTER RA
11D27 MASARAM GONDI LETTER LA
11D28 MASARAM GONDI LETTER VA
11D29 MASARAM GONDI LETTER SHA
11D2A MASARAM GONDI LETTER SSA
11D2B MASARAM GONDI LETTER SA
11D2C MASARAM GONDI LETTER HA
11D2D MASARAM GONDI LETTER LLA

Conjunct letters

11D2E MASARAM GONDI LETTER KSSA
11D2F MASARAM GONDI LETTER JNYA
11D30 MASARAM GONDI LETTER TRA

Dependent vowel signs

11D31 MASARAM GONDI VOWEL SIGN AA
11D32 MASARAM GONDI VOWEL SIGN I
11D33 MASARAM GONDI VOWEL SIGN II
11D34 MASARAM GONDI VOWEL SIGN U
11D35 MASARAM GONDI VOWEL SIGN UU
11D36 MASARAM GONDI VOWEL SIGN VOCALIC R
11D37 <reserved>
11D38 <reserved>
11D39 <reserved>
11D3A MASARAM GONDI VOWEL SIGN E
11D3B <reserved>
11D3C MASARAM GONDI VOWEL SIGN AI
11D3D MASARAM GONDI VOWEL SIGN O
11D3E <reserved>
11D3F MASARAM GONDI VOWEL SIGN AU

Various signs

11D40 MASARAM GONDI SIGN ANUSVARA
11D41 MASARAM GONDI SIGN VISARGA
11D42 MASARAM GONDI SIGN NUKTA
11D43 MASARAM GONDI SIGN CANDRA
• used for transcribing foreign vowels
11D44 MASARAM GONDI SIGN HALANTA
• used for silencing the inherent vowel

Virama

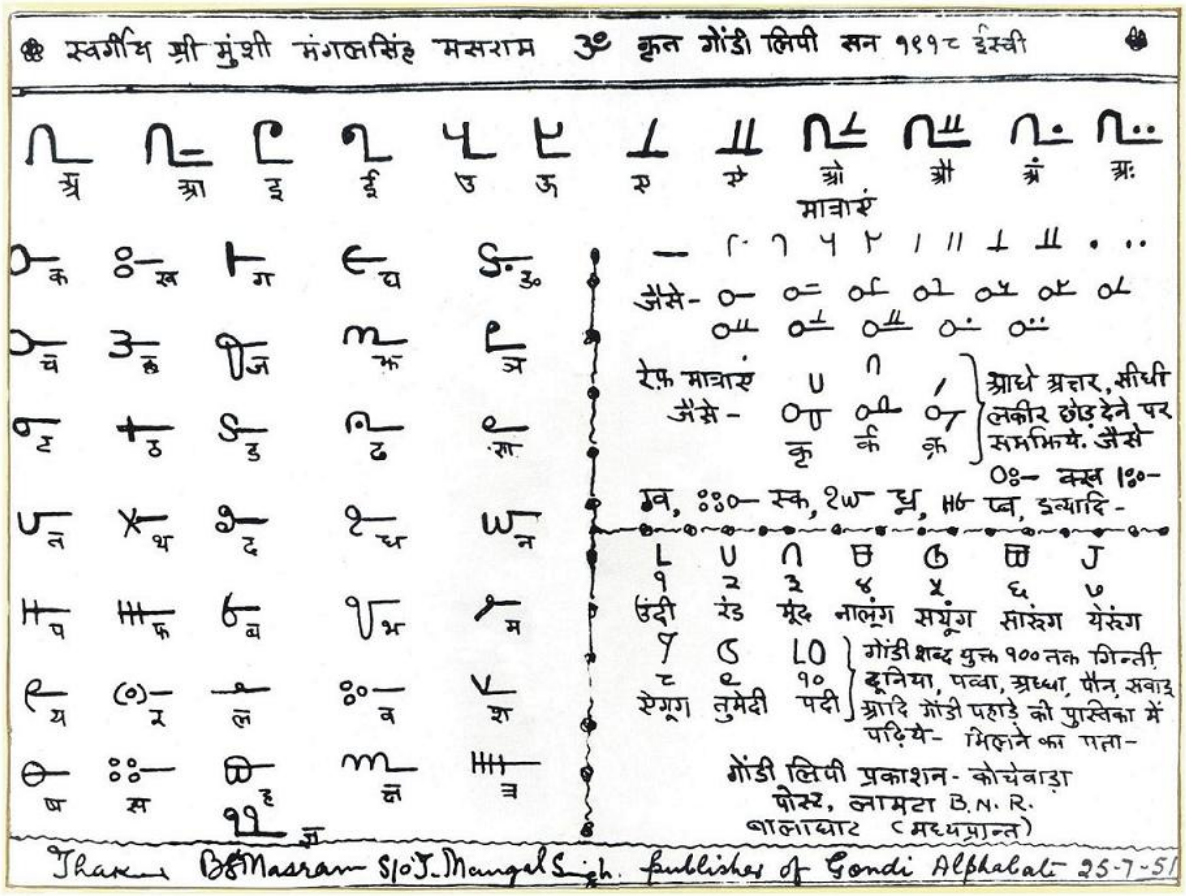
11D45 MASARAM GONDI VIRAMA
• used for producing conjuncts

Cluster-specific consonant forms

11D46 MASARAM GONDI REPHA
• cluster-initial form of RA
11D47 MASARAM GONDI RA-KARA
• cluster-final form of RA

Digits

11D50 MASARAM GONDI DIGIT ZERO
11D51 MASARAM GONDI DIGIT ONE
11D52 MASARAM GONDI DIGIT TWO
11D53 MASARAM GONDI DIGIT THREE
11D54 MASARAM GONDI DIGIT FOUR
11D55 MASARAM GONDI DIGIT FIVE
11D56 MASARAM GONDI DIGIT SIX
11D57 MASARAM GONDI DIGIT SEVEN
11D58 MASARAM GONDI DIGIT EIGHT
11D59 MASARAM GONDI DIGIT NINE



SRI MUNSHI MANGALASIMHA MASARAM KRIT GONDI LIPI - 25.7.1951

Figure 1: A document illustrating the characters and basic principles of Mangal Singh Masaram's Gondi script. It was created by Mangal Singh's son Bhava Singh in 1951. Image courtesy of the Central Institute of Indian Languages, Mysore.

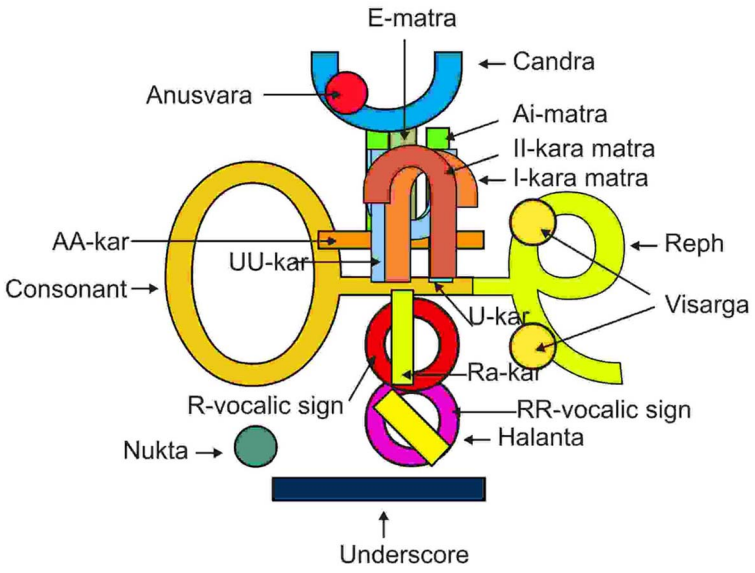


Figure 2: A schematic designed by Mukund Gokhale showing the position of all combining signs on a base letter. The diagram shows a sign labelled "RR-vocalic sign", which is a length variant of ◌ु VOWEL SIGN VOCALIC R introduced by Gokhale (see section 4.12). The 'underscore' represents the depth at which underlining should occur.

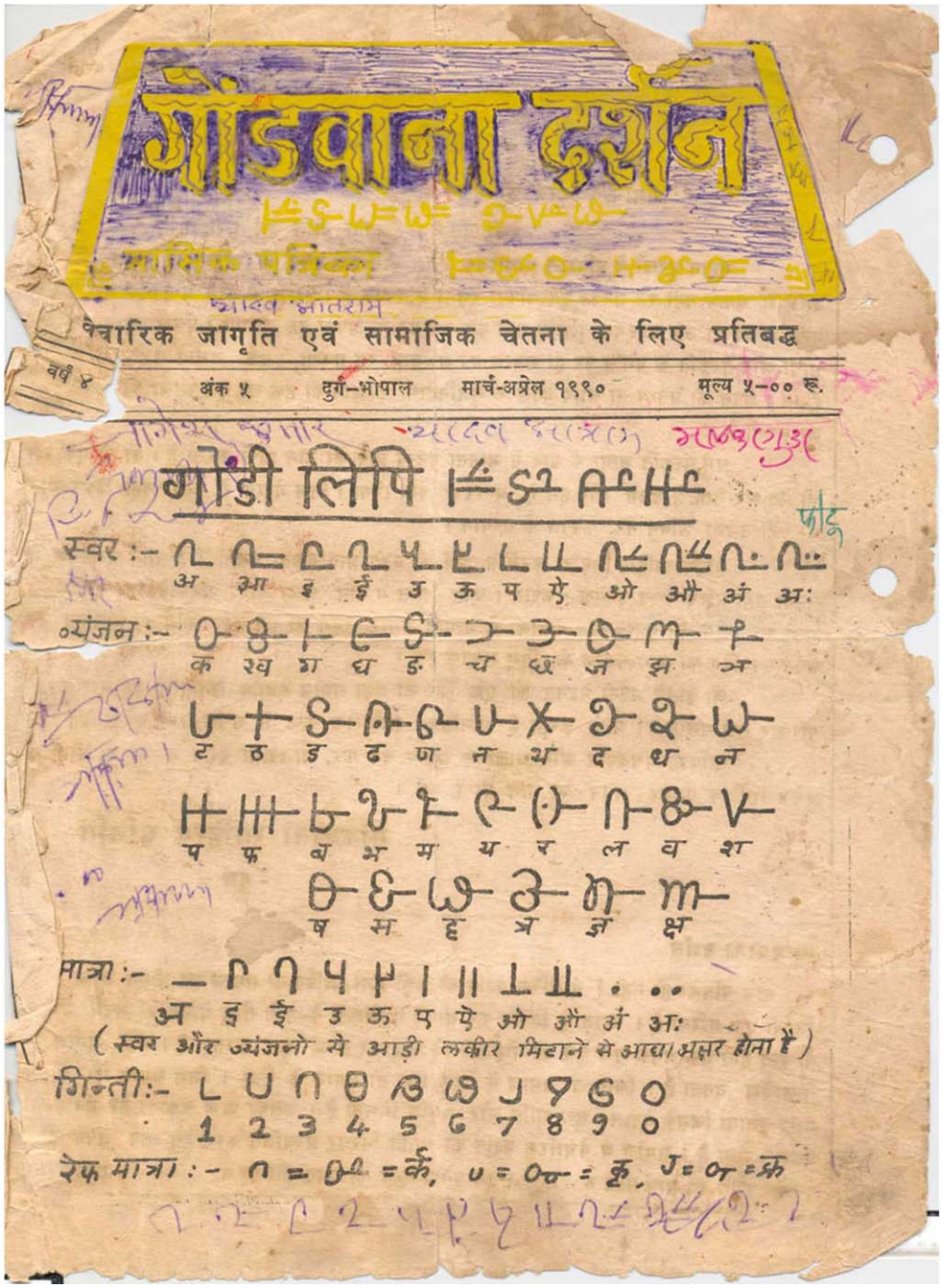


Figure 3: Cover of *Gondwana Darshan* (March-April 1990, vol. 5).

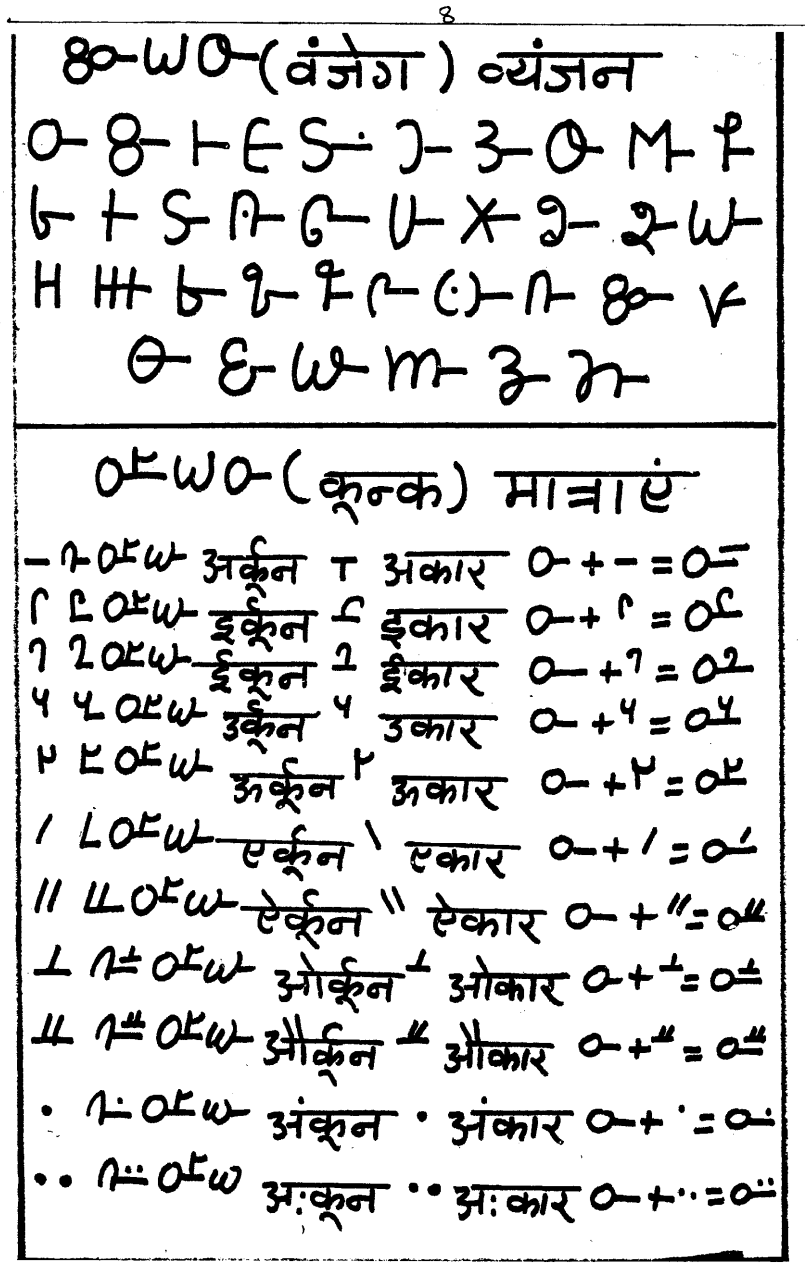


Figure 4: Page from *Goṃḍī Akṣara Jñāna* showing consonants and vowel signs, and consonant-vowel combinations (from Rāmānanda: 8).

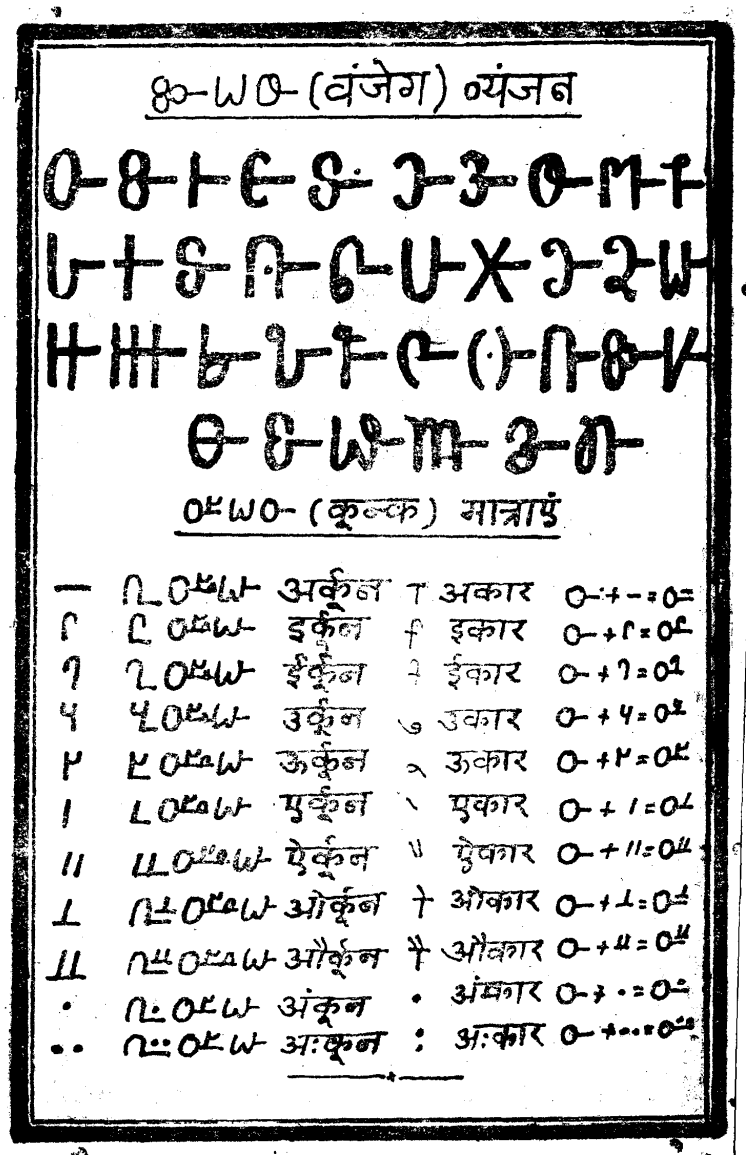
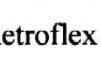


Figure 5: Page from *Goṃḍī Lamk Pundan* showing consonants and vowel signs, and consonant-vowel combinations (from Guruji: 11).

Consonants व्यञ्जन

	Cons. only	– [voice]		+ [voice]		Nasals	Cons. with word
Aspiration		– [asp]	+ [asp]	– [asp]	+ [asp]		
Velar							
Palatal							
Retroflex							
Dental							
Bilabial							
Liquids and Semi-Vowels							
Fricatives							
Affricates							

Vowels स्वर

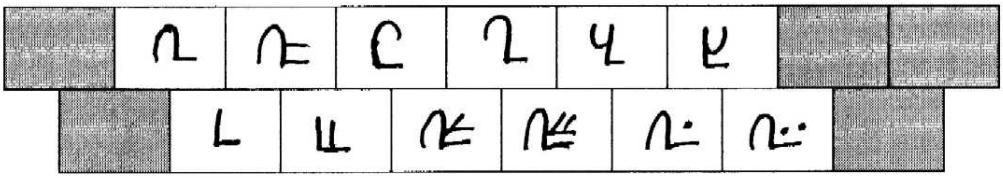


Figure 6: A handwritten chart of the Gondi script. Source: Ramesh Gedam and Mark Penny (2001).

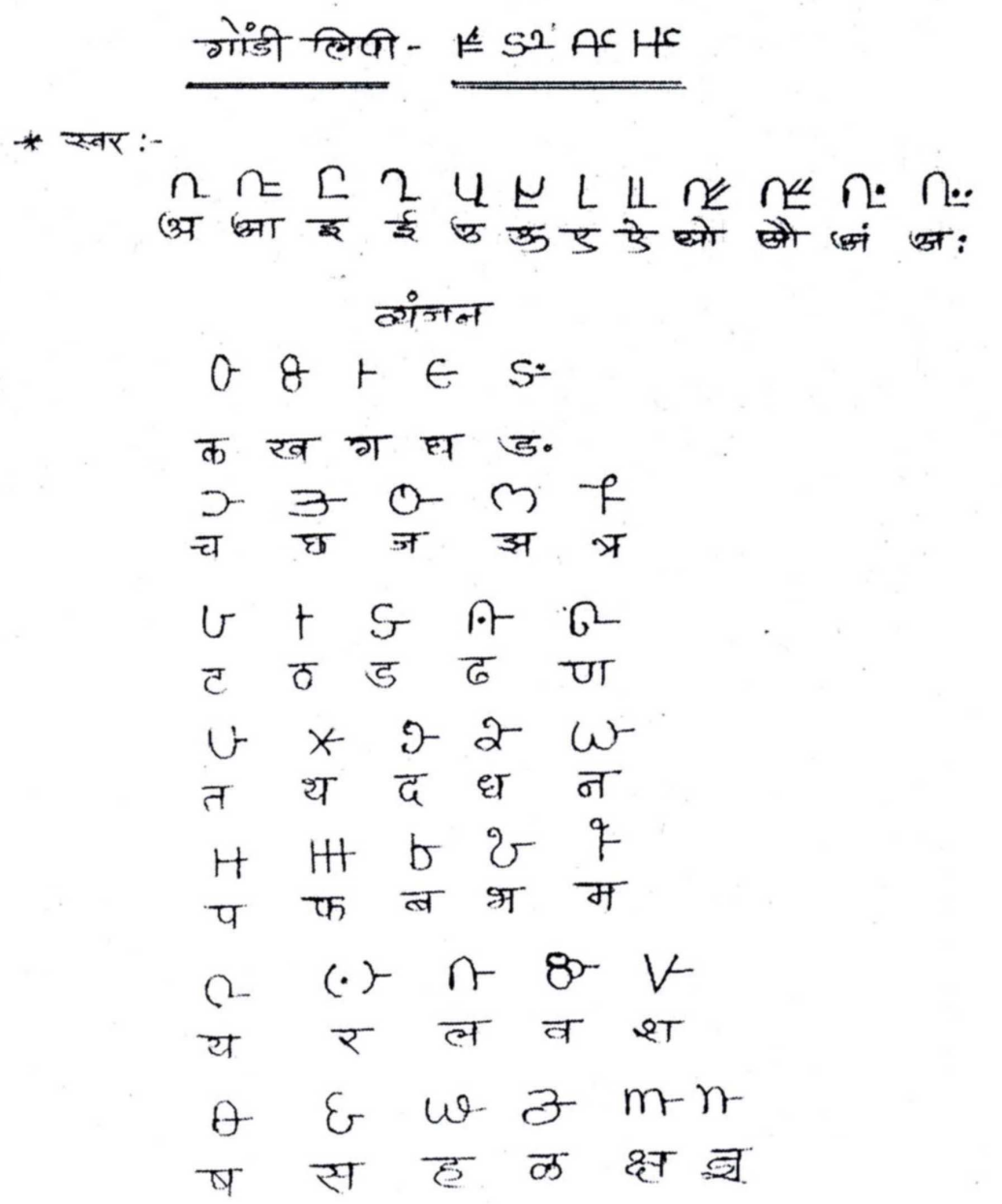


Figure 7: A handwritten chart of the Gondi script (Maṇḍāle 2008: 8).

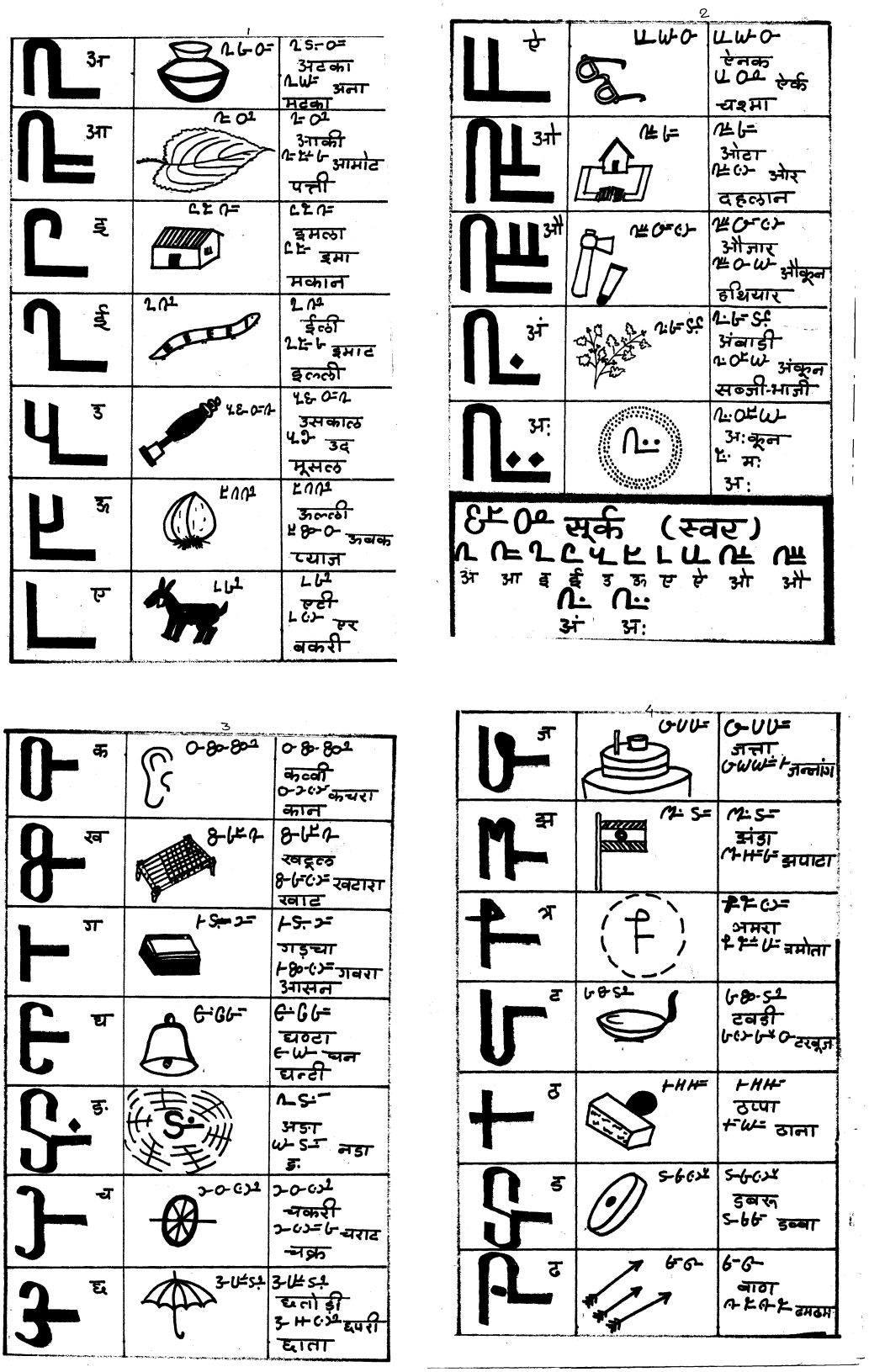


Figure 8: Page from *Goṃḍī Akṣara Jñāna* showing vowel letters (from Rāmānanda: 1–4).

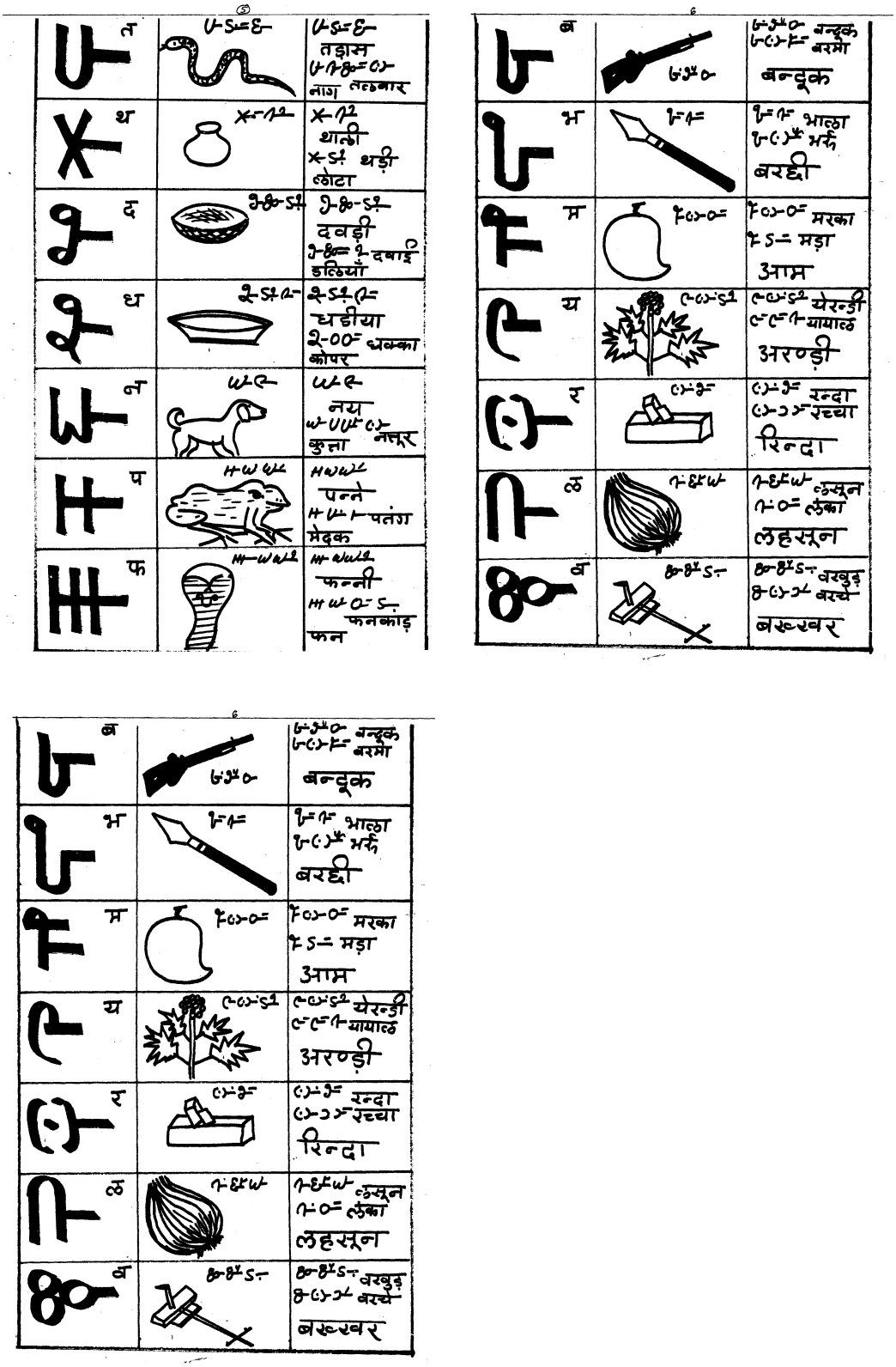


Figure 9: Page from *Goṃḍī Akṣara Jñāna* showing the letters TA .. VA (from Rāmānanda: 5–7).

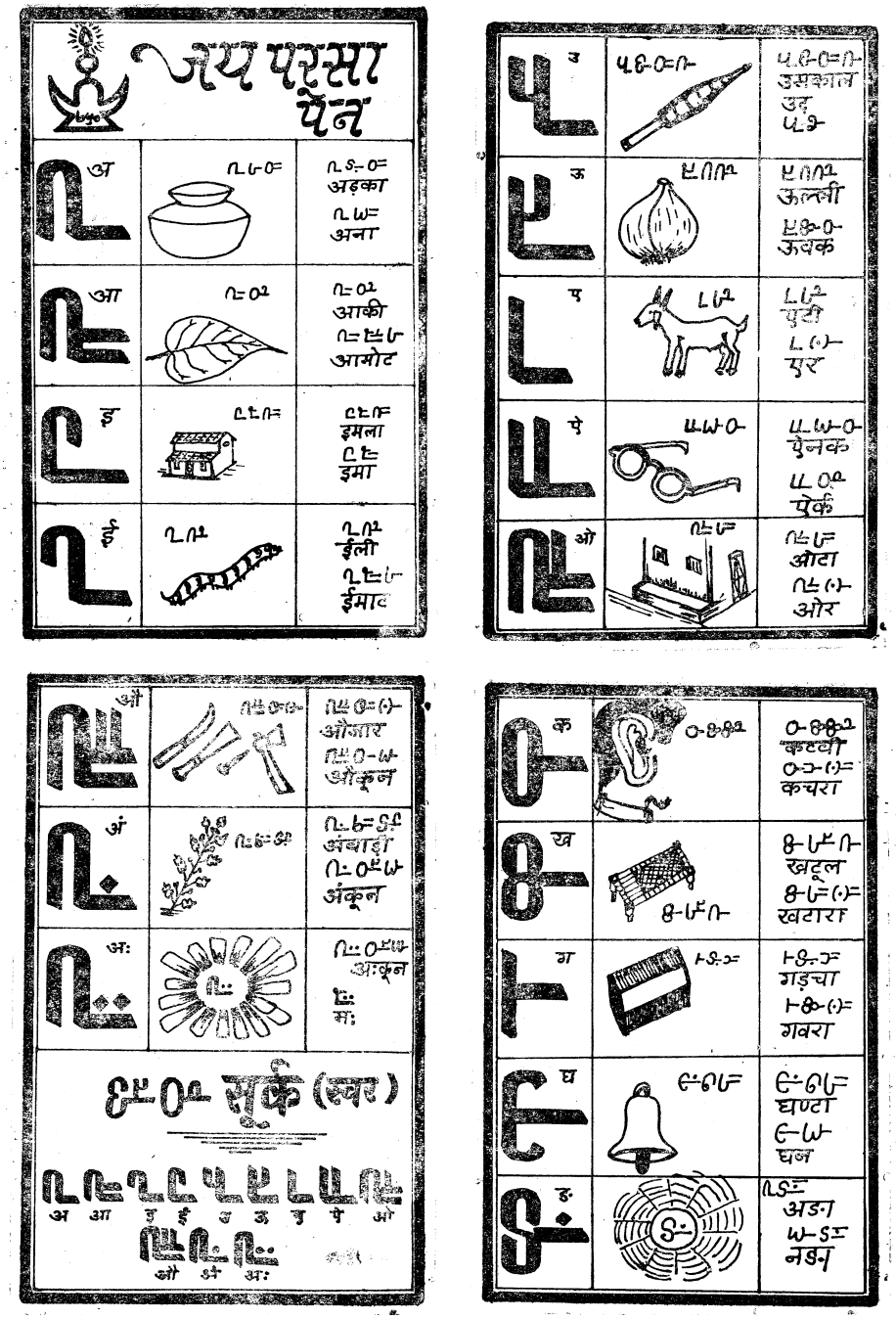


Figure 10: Page from *Gomḍī Lamk Pundan* showing vowel letters (from Guruji: 1–4). The first page (top right corner) shows the *persapen* symbol.

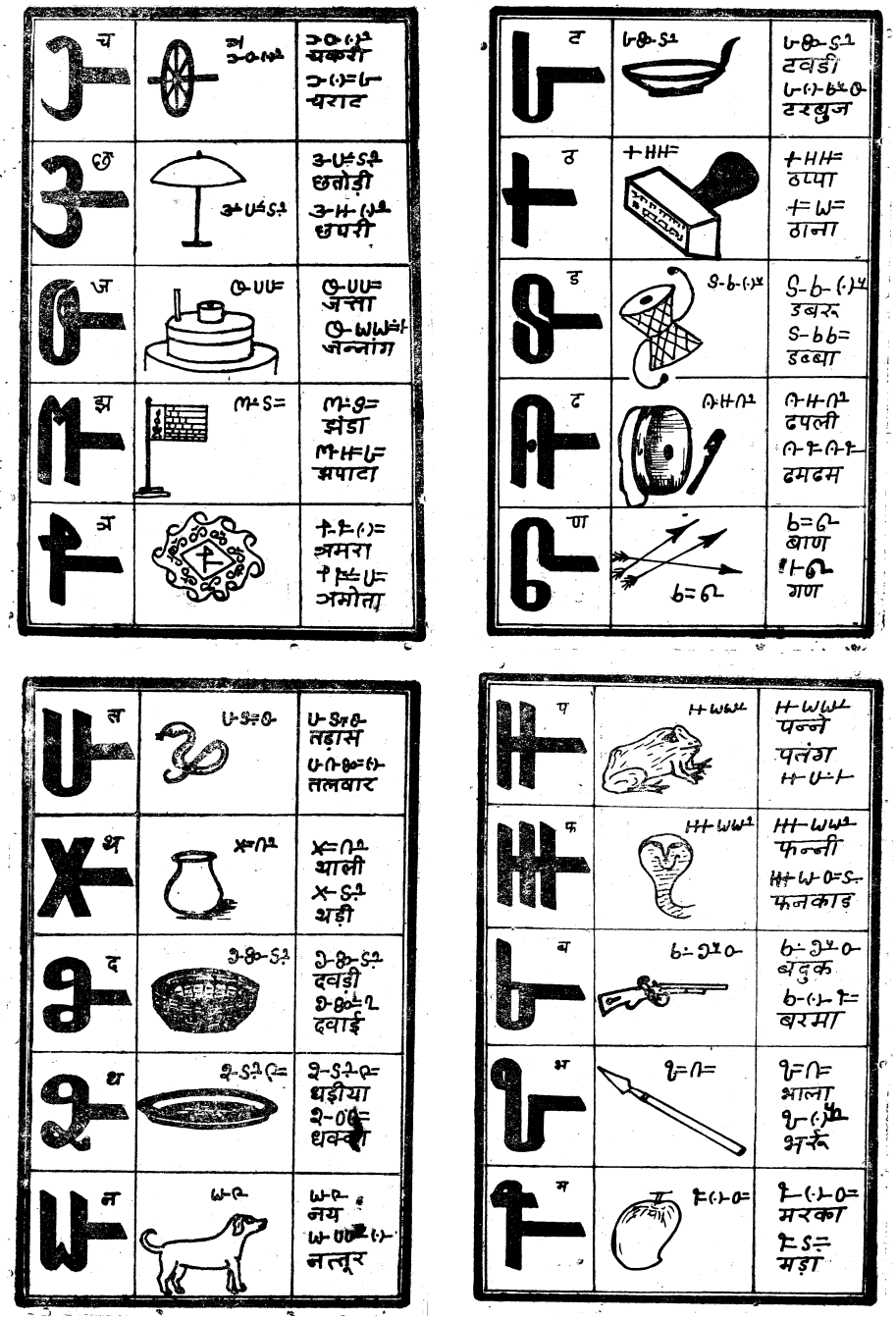


Figure 11: Page from *Goṃḍī Lamk Pundan* showing vowel letters (from Guruji: 5–8).

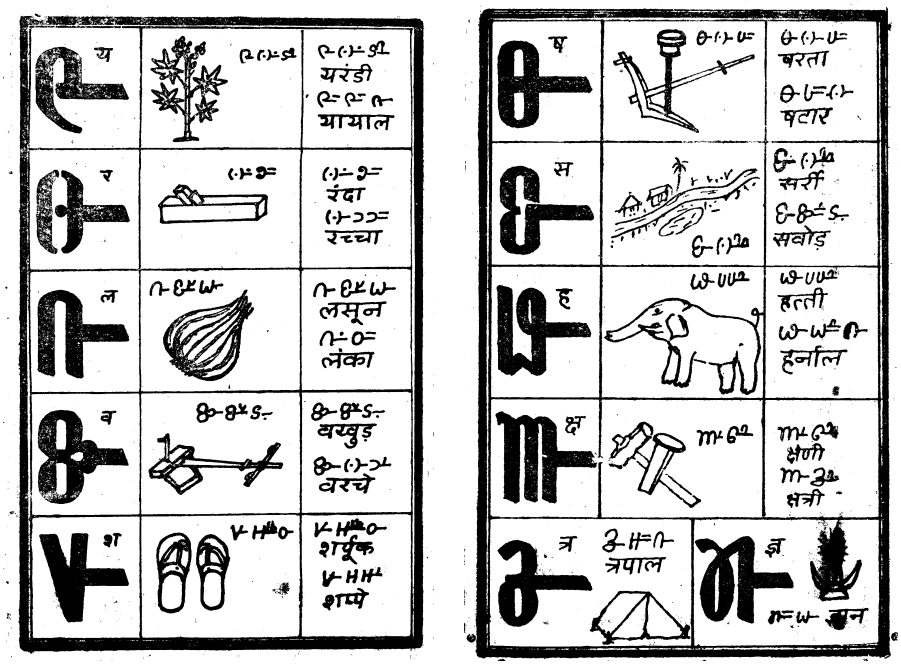


Figure 12: Page from *Goṃḍī Lamk Pundan* showing vowel letters (from Guruji: 9–10).

Figure 13: Page from *Goṃḍī Akṣara Jñāna* showing consonant-vowel combinations for KA .. DA (from Rāmānanda: 9, 10).

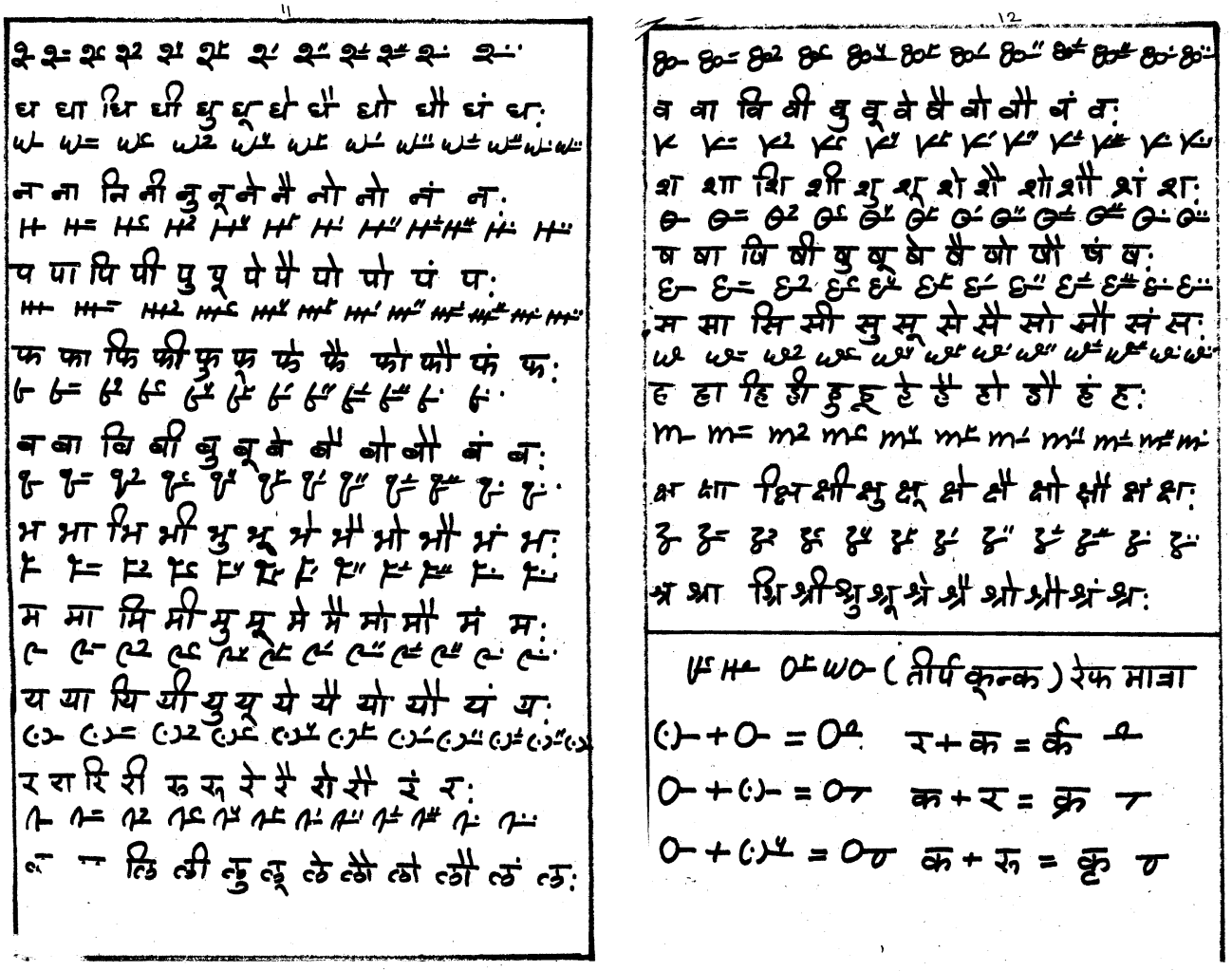


Figure 14: Page from *Goṃḍī Akṣara Jñāna* showing consonant-vowel combinations for DHA .. TRA (from Rāmānanda: 11, 12).

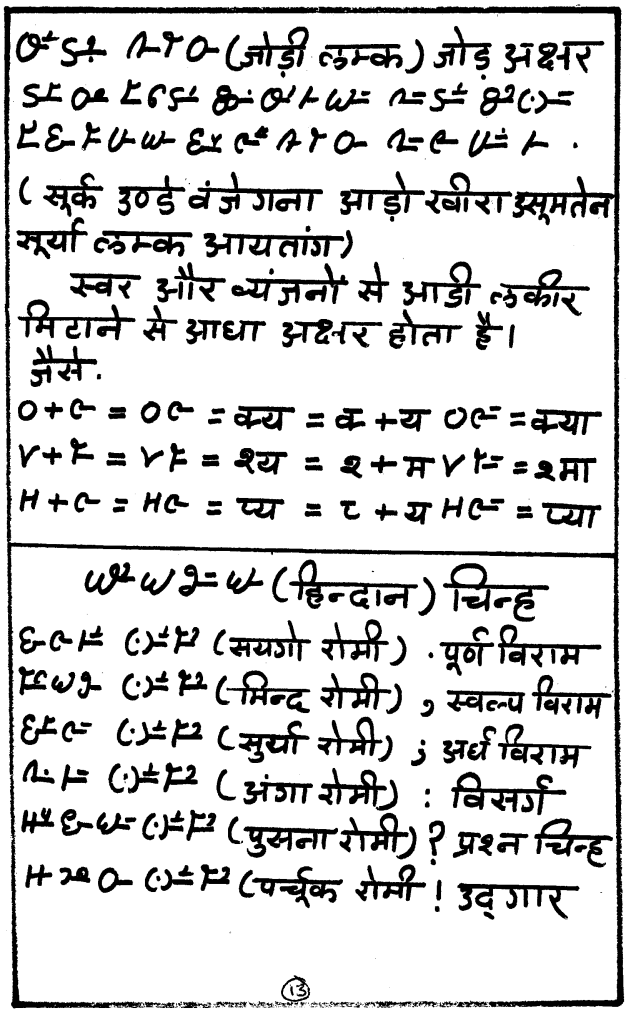


Figure 15: Page from *Goṃḍī Akṣara Jñāna* explaining conjunct formation (from Rāmānanda: 13). Bottom half describes the usage of Latin marks of punctuation.

Figure 16: Page from *Goṃḍī Akṣara Jñāna* showing Gondi digits (from Rāmānanda: 14).

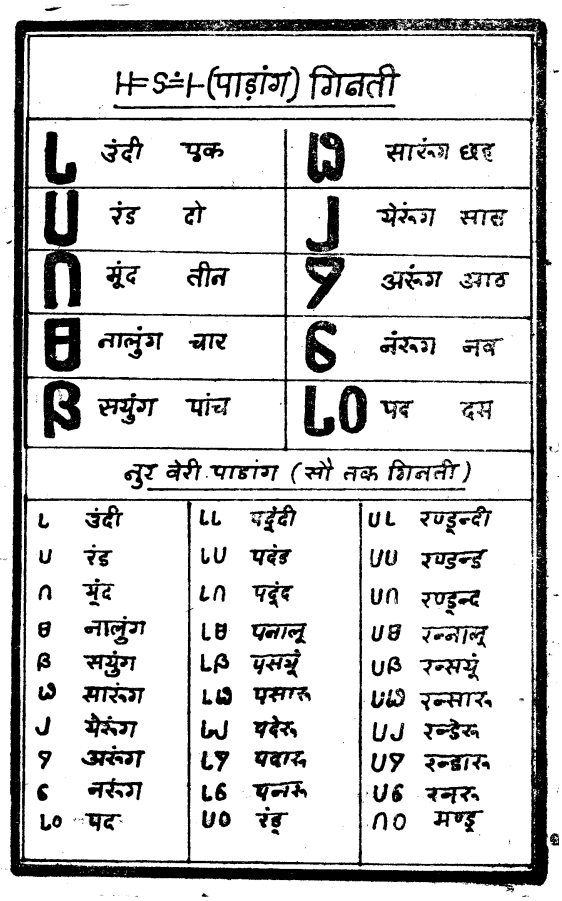


Figure 17: Page from *Goṃḍī Lamk Pundan* showing Gondi digits (from Guruji: 14).

Number	Name	Style in Maharashtra	Style in Madhya Pradesh
1	Undi		
2	Rand		
3	Munda		
4	Nalung		
5	Sayyung		
6	Sarung		
7	Yerung		
8	Arung		
9	Narung		
10	Pad		

Figure 18: Comparison of regional variants of Gondi digits (from Vahia and Halkare 2013: 33).

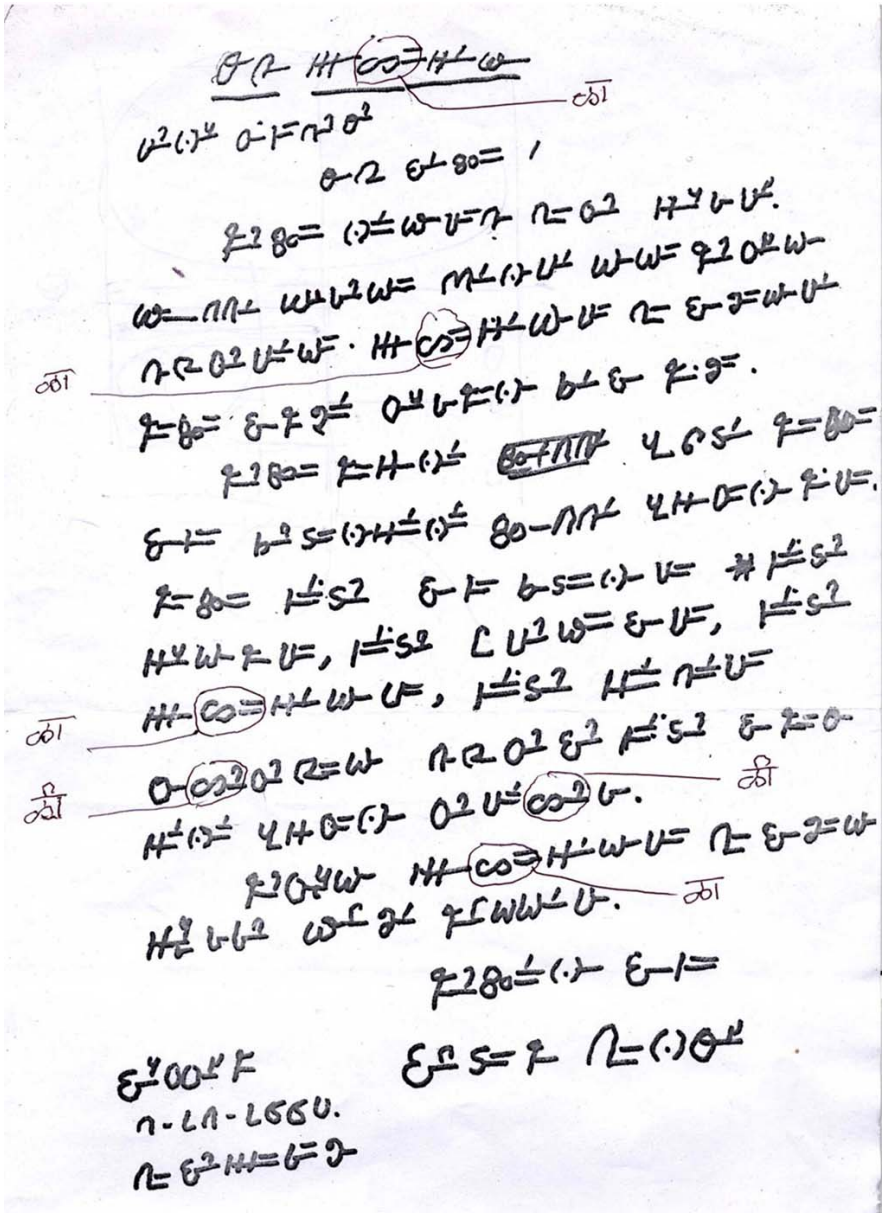


Figure 19: A document showing usage of 00- LLA.

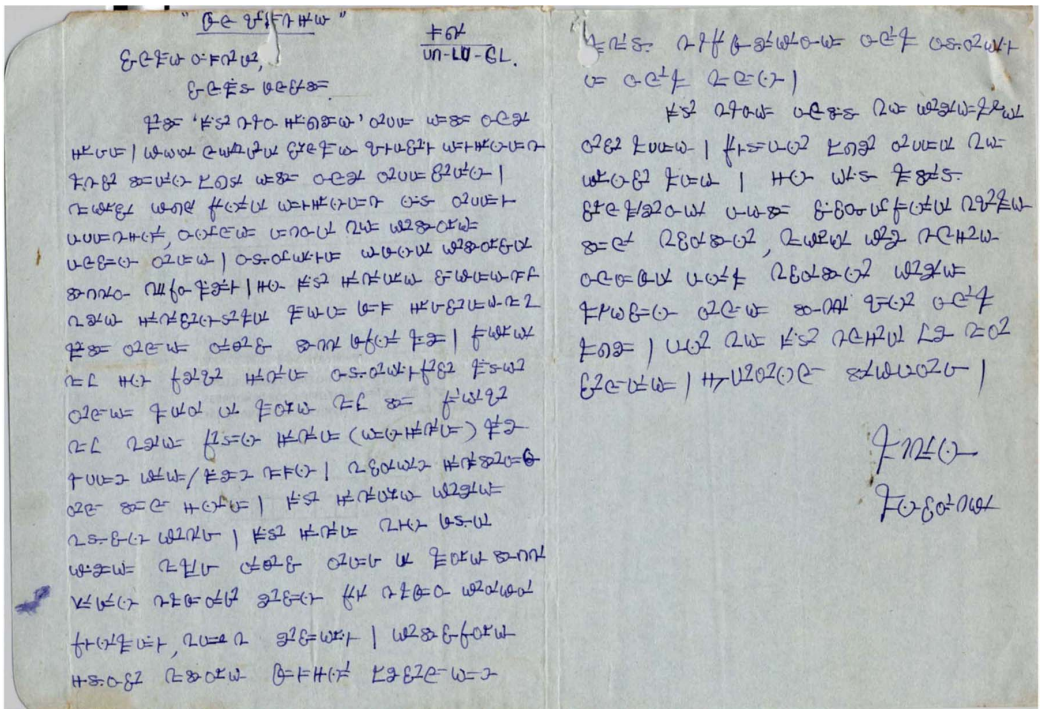


Figure 20: A letter written in Gondi.

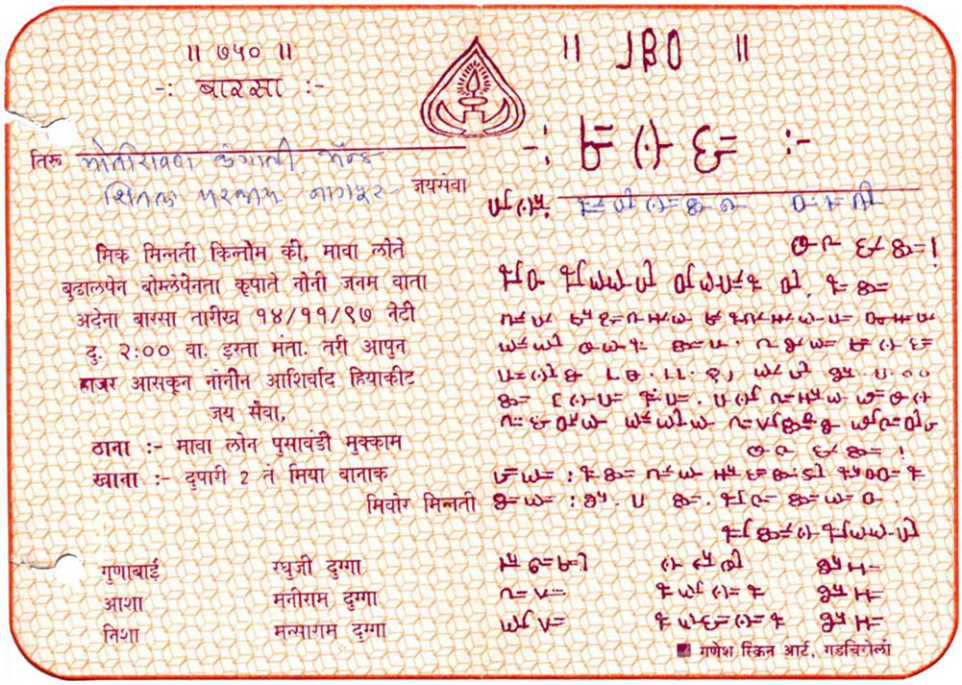


Figure 21: An invitation card written in the Devanagari and Gondi scripts. The *persapen* symbol appears in the top center of the card.

पेरसापेन पुयंनीना पाटा

ओ सल्ला शक्ति इमा, गांगरा शक्ति इमा ।
परसापेन मावा ओ परसापेन मावा ॥
जीवा शक्ति इमा, जायी शक्ति इमा ।
सजोर शक्ति मावा, ओ परसापेन मावा ॥
कयमो शक्ति मावा, करूमो शक्ति मावा ।
कामो शक्ति मावा, ओ परसापेन मावा ॥
मेंदोल शक्ति इमा, मती शक्ति इमा ।
मोह शक्ति मावा, ओ परसापेन मावा ॥
पुनो शक्ति इता, उनो शक्ति इमा ।
गुणो शक्ति मावा, ओ परसापेन मावा ॥
मान शक्ति इमा, माई शक्ति इमा ।
पुट्से शक्ति मावा, ओ परसापेन मावा ॥
कासे शक्ति इमा, मासे शक्ति इमा ।
गार शक्ति मावा, ओ परसापेन मावा ॥
तयमो शक्ति इमा, रयमो शक्ति इमा ।
सयमो शक्ति मावा, ओ परसापेन मावा ॥
गुर्रे शक्ति इमा, बुर्रे शक्ति इमा ।
यायाल शक्ति मावा, ओ परसापेन मावा ॥
मंत्रे शक्ति इमा, जंत्रे शक्ति इमा ।
बाबाल शक्ति मावा, ओ परसापेन मावा ॥
पुंगार शक्ति इमा, सिंगार शक्ति इमा ।
नींगार शक्ति मावा, ओ परसापेन मावा ॥
केंसार शक्ति इमा, वेंजार शक्ति इमा ।
लेंगार शक्ति मावा, ओ परसापेन मावा ॥
अगास शक्ति मावा, पताल शक्ति इमा ।
पोकराल शक्ति मावा, ओ परसापेन मावा ॥
विन्नासी शक्ति इमा, इवौसि शक्ति इमा ।
सल्ला शक्ति मावा, ओ परसापेन मावा ॥
ओ बुढालपेन मावा ॥ ओ परसापेन मावा ॥
॥ सेवा ॥ सेवा ॥ सेवा ॥

सीताराम मंडाले (सचीव)
आदिवासी गोंड समाज मंडळ, पुणे.

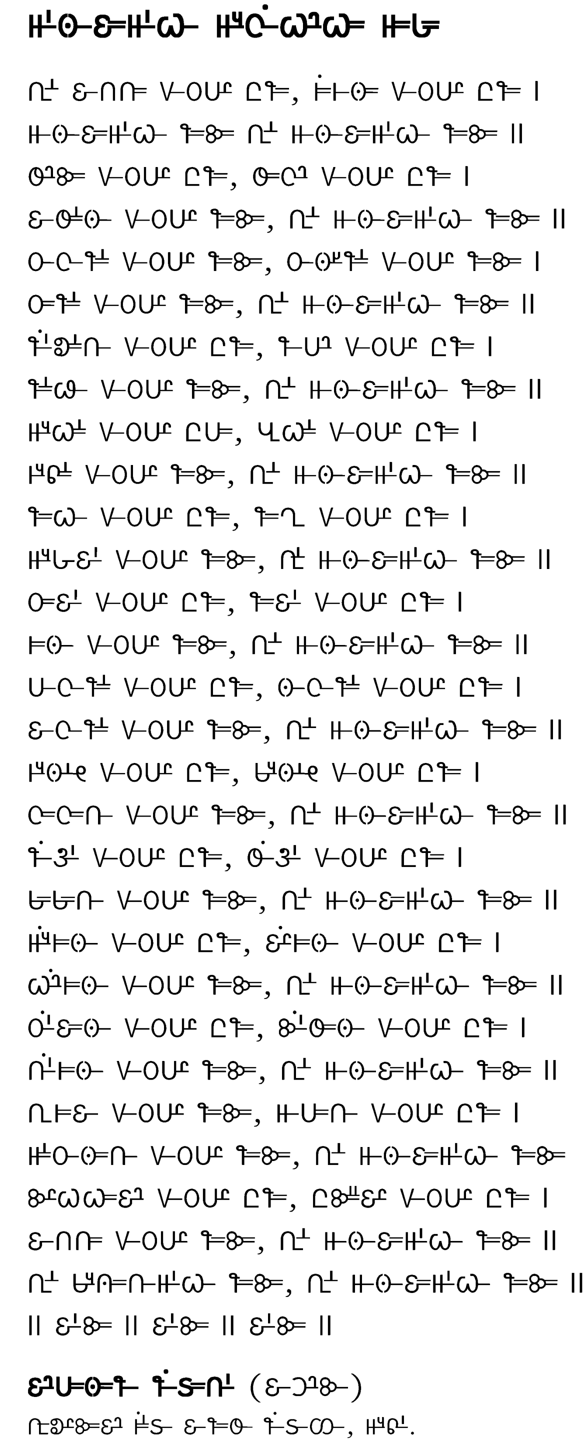


Figure 22: A Christian prayer typeset using digitized Gondi and Devanagari fonts. The Gondi font used in this specimen was designed by Mukund Gokhale.

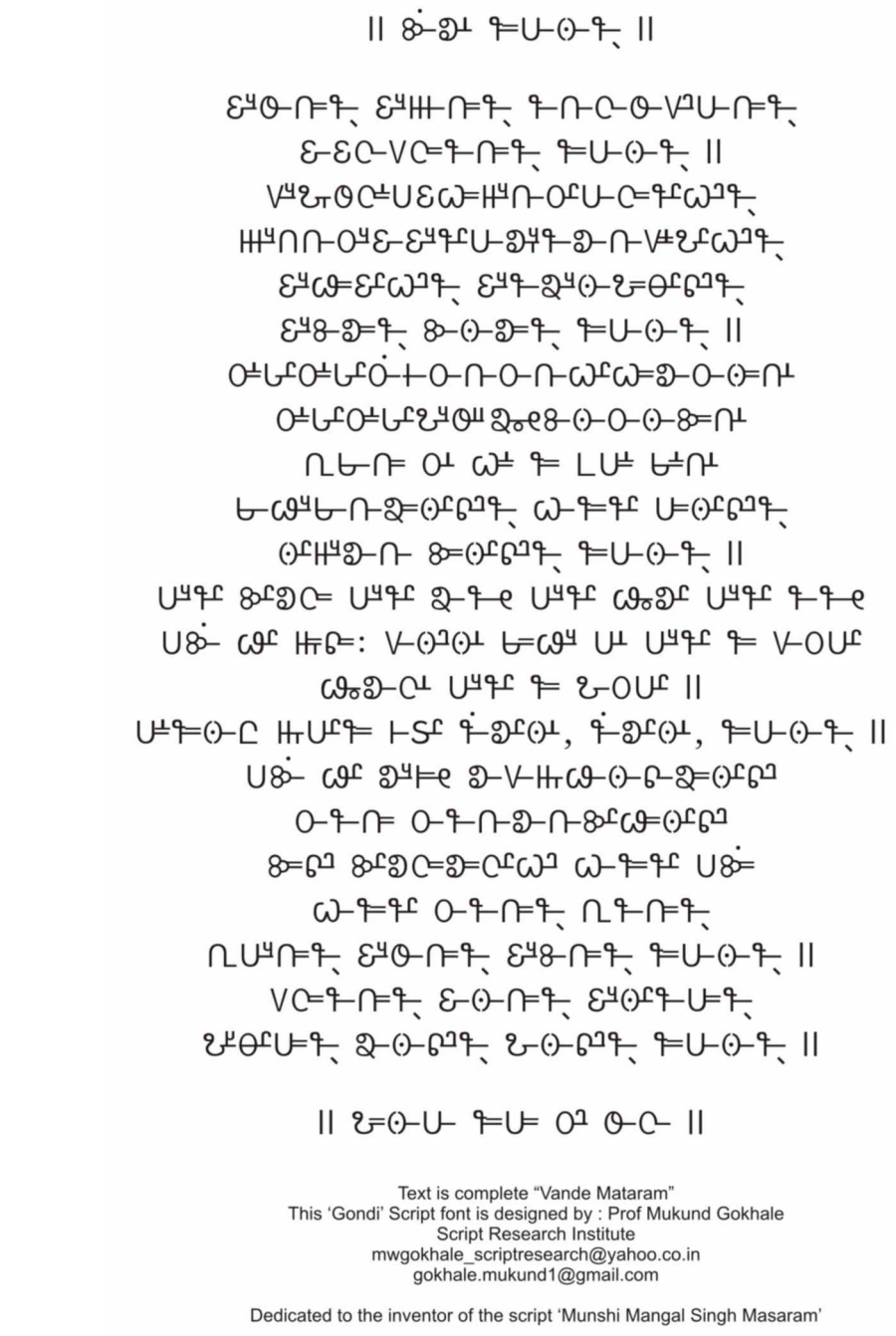
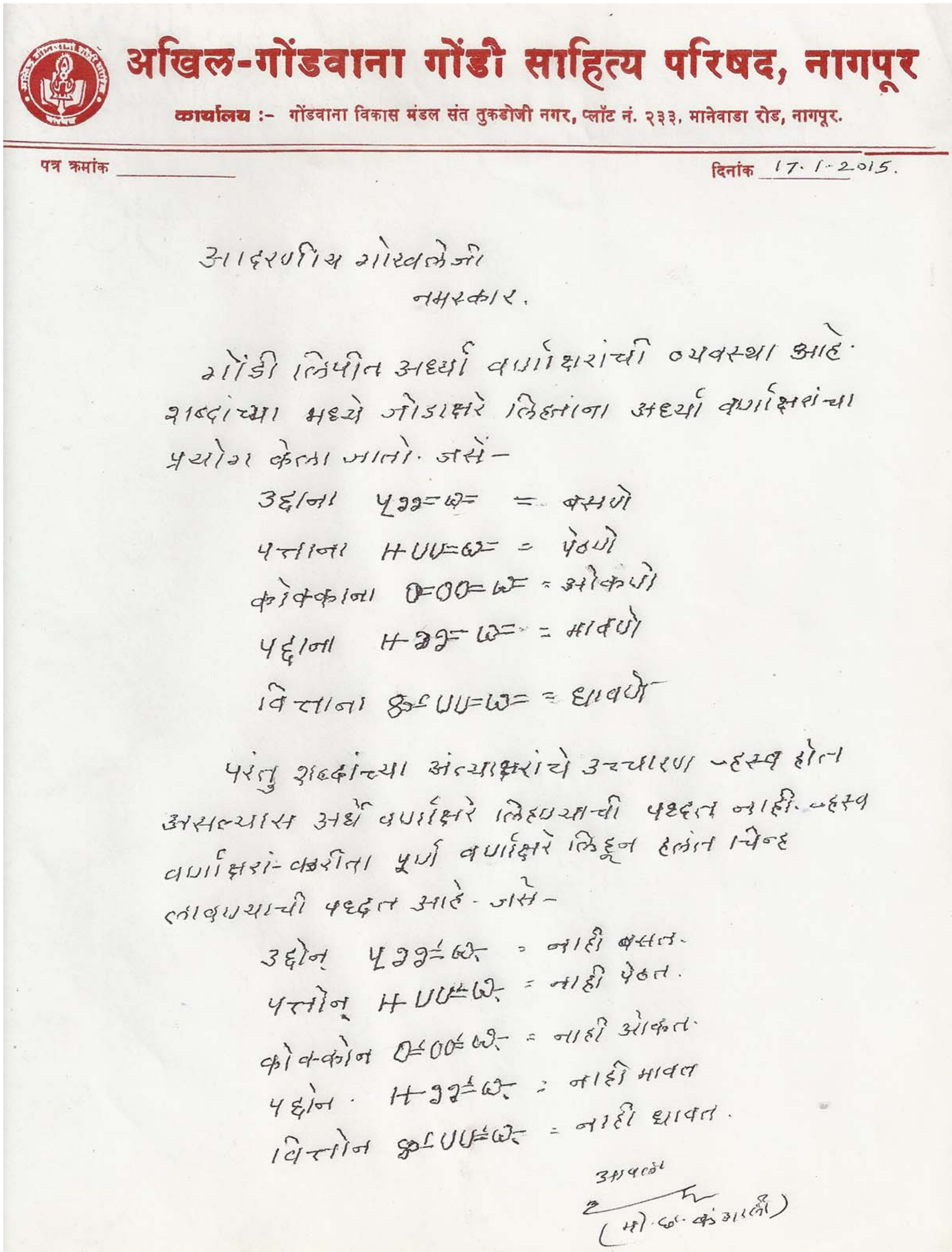


Figure 23: The lyrics to “Vande Mataram” transliterated into Gondi.



अखिल-गोंडवाना गोंडी साहित्य परिषद, नागपूर

कार्यालय :- गोंडवाना विकास मंडल संत तुकडोजी नगर, प्लॉट नं. २३३, मानेवाडा रोड, नागपूर.

पत्र क्रमांक ____________ दिनांक 17.1.2015.

आदरणीय गोखलेजी

नमस्कार.

गोंडी लिपीत अर्ध्या वर्णाक्षरांची व्यवस्था आहे. शब्दांच्या मध्ये जोडाक्षरे लिहतांना अर्ध्या वर्णाक्षरांचा प्रयोग केला जातो. जसे-

उद्दाना = बसणे
पत्ताना = पेठणे
कोक्काना = ओकणे
पद्दाना = मावणे
वित्ताना = धावणे

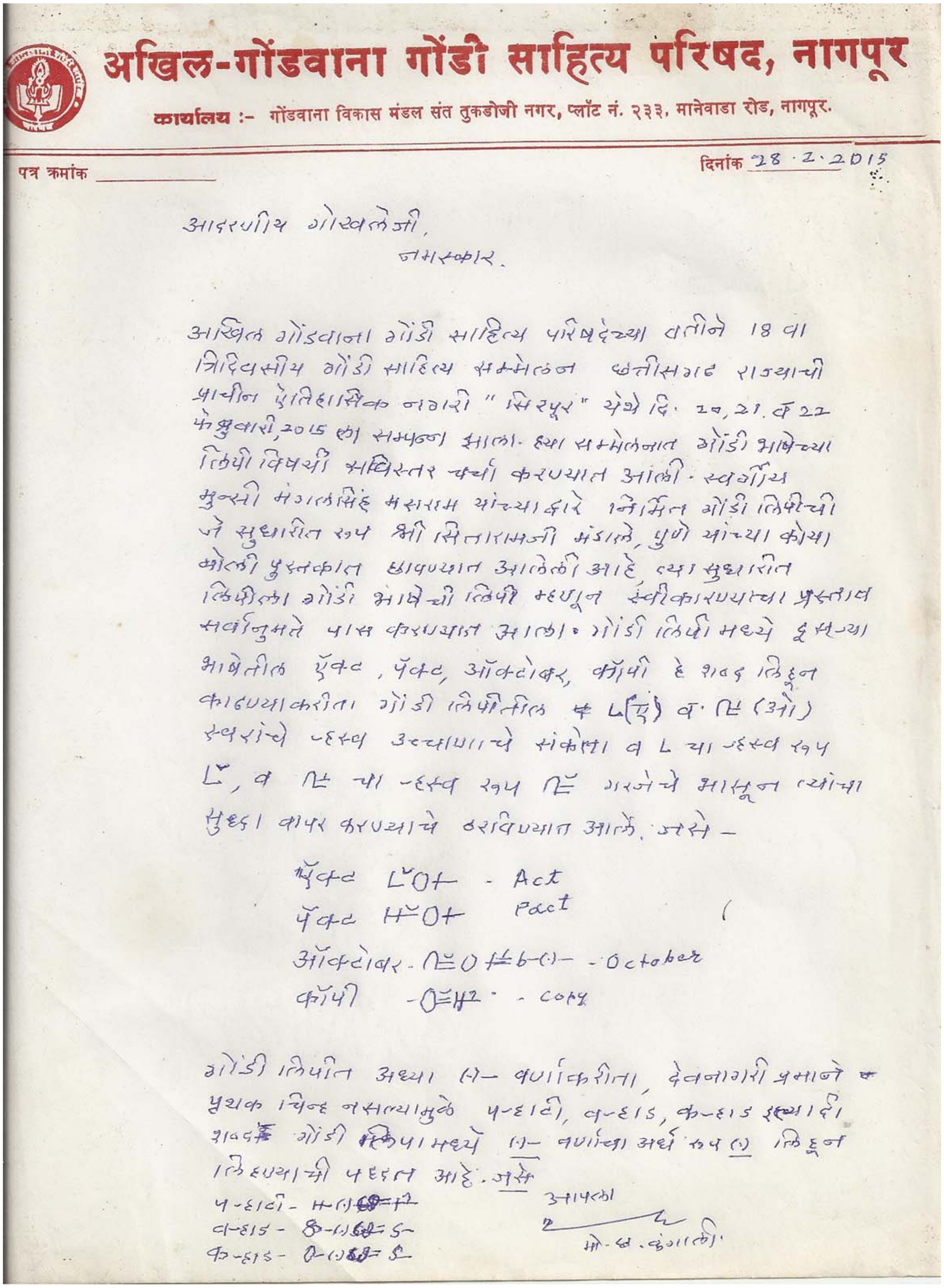
परंतु शब्दांच्या अंत्याक्षरांचे उच्चारण र्‍हस्व होत असल्यास अर्धे वर्णाक्षरे लिहण्याची पद्धत नाही. र्‍हस्व वर्णाक्षरां-करीता पूर्ण वर्णाक्षरे लिहून हलंत चिन्ह लावण्याची पद्धत आहे. जसे-

उद्दोन् = नाही बसत.
पत्तोन् = नाही पेठत.
कोक्कोन = नाही ओकत.
पद्दोन = नाही मावत
वित्तोन = नाही धावत.

आपला

(मो. छ. कंगाली)

Figure 24: Usage of the sign HALANTA with regular forms of consonants for representation of bare consonants. Contents of letter described in section 4.6.



अखिल-गोंडवाना गोंडी साहित्य परिषद, नागपूर

कार्यालय :- गोंडवाना विकास मंडल संत तुकडोजी नगर, प्लॉट नं. २३३, मानेवाडा रोड, नागपूर.

पत्र क्रमांक ________ दिनांक 28 · 2 · 2015

आदरणीय गोस्वलजी,

नमस्कार.

अखिल गोंडवाना गोंडी साहित्य परिषदेच्या वतीने 18 वा त्रिदिवसीय गोंडी साहित्य सम्मेलन छत्तीसगढ राज्याची प्राचीन ऐतिहासिक नगरी "सिरपूर" येथे दि. 20, 21 व 22 फेब्रुवारी, 2015 ला सम्पन्न झाला. ह्या सम्मेलनात गोंडी भाषेच्या लिपी विषयी सविस्तर चर्चा करण्यात आली. स्वर्गीय मुन्सी मंगलसिंह मसराम यांच्याद्वारे निर्मित गोंडी लिपीची जे सुधारीत रूप श्री सितारामजी मंडाळे, पुणे यांच्या कोया कोली पुस्तकात छापण्यात आलेली आहे, त्या सुधारीत लिपीला गोंडी भाषेची लिपी म्हणून स्वीकारण्याचा प्रस्ताव सर्वानुमते पास करण्यात आला. गोंडी लिपी मध्ये दुसऱ्या भाषेतील ॲक्ट, पॅक्ट, ऑक्टोबर, कॉपी हे शब्द लिहून काढण्याकरीता गोंडी लिपीतील [illegible] (ऍ) व [illegible] (ओ) स्वरांचे ऱ्हस्व उच्चाराचे संकेत व [illegible] चा ऱ्हस्व रूप [illegible], व [illegible] चा ऱ्हस्व रूप [illegible] गरजेचे मासून त्यांचा सुद्धा वापर करण्याचे ठरविण्यात आले. जसे –

ॲक्ट [illegible] - Act
पॅक्ट [illegible] Pact
ऑक्टोबर - [illegible] - October
कॉपी - [illegible] - copy

गोंडी लिपीत अर्ध्या [illegible] वर्णाकरीता, देवनागरी प्रमाणे पृथक चिन्ह नसल्यामुळे प-हाडी, व-हाड, क-हाड इत्यादी शब्द गोंडी लिपी मध्ये [illegible] वर्णाचा अर्ध रूप [illegible] लिहून लिहण्याची पद्धत आहे. जसे

प-हाडी - [illegible]
व-हाड - [illegible]
क-हाड - [illegible]

आपला
[illegible]
मो. ल. कंगाली

Figure 25: Usage of the sign CANDRA. Contents of letter described in section 4.4.

Figure 26: A calendar showing names for days of the week in Gondi and dates in Gondi digits.

Figure 27: Screen-shot of adivasiswara.org showing content in Gondi script. The text in the banner is valid Gondi content. The text in the main frame of the site, however, is invalid: the script representing meaningless sequences of Gondi letters as the underlying text is Latin-script content in English. The Gondi text is represented using a server-side font based upon the Latin encoding that contains Gondi glyphs. The site appears to be a work-in-progress and is awaiting proper support for the Gondi script in Unicode in order to deliver proper content.

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]
Please fill all the sections A, B and C below.
Please read Principles and Procedures Document (P & P) from http://www.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://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html**.**
See also http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

A. Administrative

1. **Title:** *Proposal to Encode the Masaram Gondi Script in Unicode*
2. Requester's name: *Script Encoding Initiative (SEI) / Anshuman Pandey (anshuman.pandey@berkeley.edu)*
3. Requester type (Member body/Liaison/Individual contribution): *Liaison contribution*
4. Submission date: *2015-06-02*
5. Requester's reference (if applicable):
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): *Yes*
      Proposed name of script: *Masaram Gondi*
   b. The proposal is for addition of character(s) to an existing block:
      Name of the existing block:
2. Number of characters in proposal: *75*
3. Proposed category (select one from below - see section 2.2 of P&P document):
   A-Contemporary X B.1-Specialized (small collection) B.2-Specialized (large collection)
   C-Major extinct D-Attested extinct E-Minor extinct
   F-Archaic Hieroglyphic or Ideographic G-Obscure or questionable usage symbols
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?
   Anshuman Pandey
   b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
   Anshuman Pandey (anshuman.pandey@berkeley.edu)
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*

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.

[1] Form number: N4102-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)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? *Yes*
   If YES explain *Replaces L2/10-207, L2/12-235, L2/15-005*
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.)? *Yes*
   If YES, with whom? *Mukund Gokhale (Script Research Institute, Pune)*
   Motiravan Kangle (Akhil Gondwana Gondi Sahitya Parishad, Nagpur)
   If YES, available relevant documents:
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? *Yes*
   Reference:
4. The context of use for the proposed characters (type of use; common or rare) *Common*
   Reference: *Used for writing the Gondi language in India*
5. Are the proposed characters in current use by the user community? *Yes*
   If YES, where? Reference:
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? *N/A*
   If YES, is a rationale provided?
   If YES, reference:
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? *Yes*
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?
   If YES, reference:
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?
   If YES, reference:
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? *No*
    If YES, is a rationale for its inclusion provided?
    If YES, reference:
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: *Combining signs*
    Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
    If YES, reference:
12. Does the proposal contain characters with any special properties such as control function or similar semantics? *Yes*
    If YES, describe in detail (include attachment if necessary) *Virama;*
    see text of the proposal
13. Does the proposal contain any Ideographic compatibility characters? *No*
    If YES, are the equivalent corresponding unified ideographic characters identified?
    If YES, reference: