Proposal to Encode the Modi Script in ISO/IEC 10646

Anshuman Pandey
Department of History
University of Michigan
Ann Arbor, Michigan, U.S.A.
pandey@umich.edu

November 5, 2011

1   Introduction

This is a proposal to encode the Modi script in the Universal Character Set (ISO/IEC 10646). It replaces the following documents:

- L2/09-249R "Preliminary Proposal to Encode the Modi Script in ISO/IEC 10646"
- N3780 L2/10-084 "Revised Code Chart and Names List for the Modi Script"

2   Background

Modi (मोडी, ꢀꢵꢜ *moḍī*) is a Brahmi-based script used mainly for writing Marathi (ISO 639-3: mar), an Indo-Aryan language spoken in western and central India, predominantly in the state of Maharashtra. Modi was also used for writing various other regional languages such as Hindi, Gujarati, Kannada, Konkani, Persian, Tamil, and Telugu. According to an old legend, the Modi script was brought to India from Sri Lanka by Hemādrī Paṇḍit, known also as Hemāḍpant, who was the chief minister of Rāmacandra (r. 1271–1309), the last king of the Yadava dynasty. Another tradition credits the creation of the script to Bāḷājī Avajī, the secretary of state to the Maratha king Shivaji Raje Bhonsle (r. 1642–1680), also known as Chhatrapati Shivaji Maharaj. While the veracity of such accounts are difficult to ascertain, it is clear that Modi derives from the Nagari family of scripts and is a modification of the Nagari model intended for continous writing. More historically, Modi emerged as an administrative writing system in the 16th century before the rise of the Maratha dynasties. It was adopted by the Marathas as an official script beginning in the 17th century and it was used in such a capacity in Maharashtra until the middle of the 20th century. In the 1950s the use of Modi was formally discontinued and the Devanagari script, known as 'Balbodh', was promoted as the standard writing system for Marathi. A revival of Modi has occurred over the past decade and its user community continues to grow.

There are thousands of Modi documents preserved in South Asia and Europe. The majority of these are in various archives in Maharashtra, while smaller collections are kept in Denmark and other countries, owing to the presence of these Europeans in Tanjore, Pondicherry, and other regions in South Asia through the 19th century. The earliest extant Modi document is dated to the early 17th century. While the majority of Modi documents are official letters, land records, and other administrative documents, the script was also

used in education, journalism, and other routine activities before the 1950s. Printing in Modi began in the early 19th century after Charles Wilkins cut the first metal fonts for the script in Calcutta (see figures 36 and 37). Newspapers were published in Modi (see figure 38); primers were produced for teaching the script in schools (see figure 30); and various personal papers and diaries were kept in the script.

At present, Modi users regularly publish books and magazines in the script (see figure 39). Several institutions now offer tutorials for learning Modi, which range from weekend workshops organized by the Maharashtra State Department of Archives (see *Daily News and Analysis* 2011) to formal courses offered by Bharat Itihas Samshodhak Mandal (BISM) in Pune. Some colleges also offer certifications in Modi, which are recognized by the Government of Maharashtra. The revival of Modi runs parallel to a growing need for Modi experts to catalogue and manage the vast archives of Modi documents at BISM and other locations. The Government of India has allocated funds for the cataloguing of these archives, which is being done by teams recently trained as Modi experts (see *Times of India* 2011).

Modi users have developed support for using the script on computers, mainly in the form of digitized fonts. But given the lack of a character-encoding standard for the script, users these fonts are based upon legacy encodings or are mapped to Unicode blocks such as Devanagari. Electronic materials are produced as images or in Portable Document Format (PDF). But, these solutions do not allow for the representation of Modi in plain text. The lack of a Unicode standard for Modi impedes its use in basic computing activities, such as writing e-mails, creating websites, and in digital publishing, as well as in larger scale efforts such as localization and internationalization. An encoding for Modi in the UCS will permit its representation in plain text, which will meet the needs of the growing Modi user community. It will also provide a standard foundation upon which the community may develop new resources for the script.

2.1 Relationship to Devanagari

Although Modi is based upon the same model as Devanagari, it differs considerably from the latter in terms of letterforms, rendering behaviors, and orthography. A comparison of the two scripts is shown in figure 3 and 4. The shapes of some consonants, vowels, and vowel signs are similar. The actual differences are visible in the behaviors of these characters in certain environments, such as consonant-vowel combinations and in consonant conjuncts, that are standard features of Modi orthography. Proper representation of Modi in plain-text requires the preservation of its distinct rendering behaviors. This can only be accomplished at the character level, through character content that is independent of font changes or other formatting.

3 Proposal Details

3.1 Script Name and Allocation

The characters are proposed for encoding in a new script block to be named 'Modi'.

3.2 Allocation

The block is allocated to the Supplementary Multilingual Plane (SMP) at the range U+11600..1165F.

3.3 Character Repertoire

A total of 79 characters is required to encode a basic character set for Modi. The character names follow the UCS convention for Indic scripts. The proposed code chart and names list are shown in figures 1 and 2.

3.4   Font

The Modi font used in this document was designed by the proposal author. The letterforms are inspired by the Chitnisi style of Modi used during the Peshve era.

3.5   Encoding Principles

The proposed encoding for Modi seeks to represent the script in plain text. It is understood that Modi is a highly stylistic and cursive script. The proposed Modi glyphs shown in the code chart are an attempt to represent the basic graphical structure of Modi letters, as is done in traditional *bārākhaḍī*, or charts showing the consonant letters written with each vowel sign. The joining features and other stylistic elements of the script should be considered typographic matters and managed through the design of fonts and other features at the presentation level.

4   Writing System Details

4.1   Structure

The general structure (phonetic order, *mātrā* reordering, use of *virāma*, etc.) of Modi is similar to that of Devanagari. Several consonant-vowel combinations are written as ligatures. Consonant clusters are represented as conjuncts. Some consonants have special behaviors when they occur in certain environments. These are described below.

4.2   Styles

There are several styles of Modi. The earliest is the 'proto-Modi' of the 12th century, known as Adyakalin आद्यकालीन. A distinct Modi form emerges during the 13th century and is known as Yadavakalin यादव कालीन. The next stage of development is the Bahamanikalin बहमनी कालीन of the 14th–16th century, followed by the Shivakalin शिव कालीन of the 17th century. The well-known Chitnisi form develops during this period. In the 18th century, various Modi styles began to proliferate. This era is known as Peshvekalin पेश्वे कालीन, which lasted until 1818. The distinct styles of Modi used during this period are known as Chitnisi, Bilavalkari, Mahadevapanti, and Ranadi. The final stage of Modi is associated with English rule and is called Anglakalin आंगल कालीन. This form were used from 1818 until 1952.

Four of the most well-known historical forms are Bahamani (figure 26), Chitnisi (figure 27), Peshve (figure 28), and Anglakal (figure 29). Another style of Modi was used in the primary school books produced during the 19th and 20th centuries (figures 30 and 31). This form was not written in the typical cursive style, a feature that was consciously avoided in order to ensure legibility.

4.3   Virāma

The Modi ◌् VIRAMA is identical in function to the corresponding character in other Indic scripts. The Marathi name for this sign is *halant*.

4.4 Vowels

There are 14 vowel letters:

ꣳ	A	ꣴ	U	ꣵ	VOCALIC L	ꣶ	O
	AA		UU		VOCALIC LL		AU
	I		VOCALIC R		E		
	II		VOCALIC RR		AI		

Generally, the distinction between regular and long forms of *i* and *u* are not preserved in Modi. The letter ꣳ II may represent both *i* and *ī*; ꣴ U may be used for writing both *u* and *ū*. However, both regular and long forms are proposed for encoding because they are attested in script charts as distinct letters. The vocalic letters are shown in Modi charts, but as is the case in other Indic scripts, their inclusion is more a matter of tradition than usage. Modi ꣵ VOCALIC R may be transcribed as <ꣳ RA, ꣴ VOWEL SIGN II> *rī*.

Some vowel letters may be represented using a sequence of a base vowel-letter and a combining vowel sign. This practice is not recommended. The independent vowel letter is always to be used:

RECOMMENDED	NOT RECOMMENDED
ꣳ VOWEL LETTER E	ꣳ VOWEL LETTER A + ꣴ VOWEL SIGN E
ꣳ VOWEL LETTER AI	ꣳ VOWEL LETTER A + ꣴ VOWEL SIGN AI
ꣳ VOWEL LETTER O	ꣳ VOWEL LETTER AA + ꣴ VOWEL SIGN E
ꣳ VOWEL LETTER AU	ꣳ VOWEL LETTER AA + ꣴ VOWEL SIGN AI

4.5 Vowel Signs

There are 13 combining vowel signs:

ꣿ	SIGN AA	ꣿ	SIGN UU	ꣿ	SIGN VOCALIC LL	ꣿ	SIGN AU
	SIGN I		SIGN VOCALIC R		SIGN E		
	SIGN II		SIGN VOCALIC RR		SIGN AI		
	SIGN U		SIGN VOCALIC L		SIGN O		

The descriptor 'VOWEL SIGN' used in the names of these characters is abbreviated as '-' throughout this document, eg. 'VOWEL SIGN AA' is referred to as '-AA'.

The sign ꣿ -II is used for writing both *i* and *ī*; ꣿ -U is used for both *u* and *ū*. As is the case for encoding the independent letters I and UU, the dependent vowel signs corresponding to these letters are also proposed for encoding because they are attested as distinct characters in script charts and other materials.

Certain consonant-vowel combinations are represented as ligatures. These are described in Section 4.7 and illustrated in Table 1.

4.6 Consonants

There are 34 consonant letters:

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

Each consonant letter bears an inherent vowel, represented by /a/ (ᱫ A). This inherent vowel is silenced using ᱫ VIRAMA. When certain consonants appear in combination with certain dependent vowel signs, they are normally represented as ligatures (see Section 4.7). Similarly, when certain consonant letters appear in consonant conjuncts or in sequences with other consonants, there are distinct rules that govern their representation (see Section 4.8).

4.7 Consonant-Vowel Combinations

Consonant-vowel (CV) combinations are represented by attaching vowel signs to a consonant letter or as ligatures.

1. **Method 1: CV as base letter + sign**

   (a) **Combination with no shaping changes**

   Only 8 consonants are represented in all CV combinations with no change to glyph shapes: ᱫ GHA, ᱫ NGA, ᱫ CHA, ᱫ DDHA, ᱫ NNA, ᱫ SHA, ᱫ SSA, ᱫ LLA.

   (b) **Combination with contextual form of consonant letter**

   There are 7 consonant letters that take a contextual form when they combine with certain vowel signs: ᱫ THA, ᱫ DA, ᱫ DHA, ᱫ PA, ᱫ MA, ᱫ YA, ᱫ RA. The contextual forms never occur independently and occur only in CV combinations with the vowel signs below:

	regular	contextual	occurs with
THA	ᱫ	ᱫ	ᱫ -AA, ᱫ -U, ᱫ -UU, ᱫ -O, ᱫ -AU
DA	ᱫ	ᱫ	ᱫ -AA, ᱫ -O, ᱫ -AU
DHA	ᱫ	ᱫ	ᱫ -AA, ᱫ -UU, ᱫ -O, ᱫ -AU
PA	ᱫ	ᱫ	ᱫ -AA, ᱫ -U, ᱫ -UU, ᱫ -O, ᱫ -AU
MA	ᱫ	ᱫ	ᱫ -U, ᱫ -UU
YA	ᱫ	ᱫ	ᱫ -U, ᱫ -UU
RA	ᱫ	ᱫ	ᱫ -AA, ᱫ -U, ᱫ -UU, ᱫ -O, ᱫ -AU

2. **Method 2: CV as ligature**

   Certain CV combinations with ◌ा -AA, ◌ु -U, ◌ू -UU, ◌ो -O, ◌ौ -AU are represented as ligatures. Generally, the ligature consists of a reduced or half-form of the consonant letter and a contextual form of the vowel sign. These ligatures are illustrated in table 1.

   (a) **Ligatures with -AA, -O, -AU**

       When the vowel signs ◌ा -AA, ◌ो -O, ◌ौ -AU, combine with certain consonant letters, the vertical descender ◌ा in the signs changes to the contextual form ◌ऽ. Ligatures formed with these vowel signs can be divided into two classes based upon the structure of the consonant:

       i. *Consonants with looped right-hand terminals*   The final upwards stroke is directed downwards and looped once at the baseline before being joined to the loop of the contextual vowel sign. This behavior is exhibited by the following letters: क KA, झ JHA, ञ NYA, फ PHA. For example, ‹क KA, ◌ा -AA› → का.

       ii. *Consonants without looped right-hand terminals*   The final upwards stroke is truncated and directed rightwards in order to join to the contextual vowel sign: This behavior is exhibited by the following letters: ख KHA, ज JA, ट TTA, ठ TTHA, न NA, ब BA, ल LA, व VA. For example, ‹ख KHA, ◌ा -AA› → खा.

   (b) **Ligatures with -U, -UU**

       The vowel signs ◌ु -U and ◌ू -UU take the shapes ◌ी and ◌ीू, respectively, and are writing as ligatures when combined with च CA, झ JHA, ञ NYA, ध DHA, ब BA, ल LA, व VA. For example, ‹च CA, ◌ु -U› → ची and ‹च CA, ◌ू -UU› → चीू.

   (c) **Atomic Ligatures**

       Certain CV combinations are written as atomic ligatures. For example, ‹च CA, ◌ा -AA› → चा; ‹स SA, ◌ा -AA› → सा; ‹स SA, ◌ु -U› → सी. A complete list of such ligatures is given in table 1.

4.8   Consonant Conjuncts

Consonant clusters are generally represented as conjuncts. Conjuncts are formed by writing VIRAMA after each bare consonant in a cluster. Certain consonants have contextual shapes and rendering behaviors when they are written as part of a conjunct.

1. **Ligatures**   Two conjuncts are written as atomic ligatures: ‹क KA, ◌् VIRAMA, ष SSA› → क्ष *ksa* ; and ‹त TA, ◌् VIRAMA, र RA› → त्र *tra*.

2. **Half Forms** C_1 is generally represented using a half-form. Some examples are: TA त → त्; SA स → स्; MA म → म्; DA द → द्. The half-form of द DA is used for writing द्न ‹द DA, ◌् VIRAMA, न NA›, the Modi equivalent of Devanagari ज्ञ ‹ज JA, ◌् VIRAMA, ञ NYA› *jña*.

3. **Contextual Forms**   Some consonants have a contextual form when they appear in a cluster: THA थ → थ; PA प → प; MA म → म; YA य → य. For example, त्य ‹त TA, ◌् VIRAMA, य YA›. The list of consonants that exhibit this behavior is given in Section 4.7, item 1(b).

4.9   Forms of RA

When written independently, the letter ꌃ RA retains its basic shape:

When RA combines with vowel signs, occurs in conjuncts, or in other environments, it is written using contextual forms. These cases described below; see also figures 18–21.

1. **Conjunct Initial**   When RA appears as C_1 in a consonant cluster it takes the following shapes:

   (a) *Repha*   The ◌̊ *repha* is produced as <ꌃ RA, ◌ VIRAMA>:

   (b) *Eyelash RA*   The form ꟷ◌ *eyelash RA* is used in place of *repha* in certain contexts. In some instances, it represents a palatalized /r/ and is transliterated as *ř a*, as per ISO 15919. The *eyelash RA* is used only in native Marathi words and occurs in clusters involving ꌥ YA and ꌅ HA. It is to be produced using the sequence <ꌃ RA, ◌ VIRAMA, [zwj] ZWJ>, as is the convention in Devanagari.

   Compare *ryā* in the word *asaṇāryā*

   with the *rya* in *cāturya*

   The distinction between *eyelash RA* and *repha* is both morphological and phonemic. The *eyelash RA* is used in place of *repha* at morphological boundaries, eg. the creation of a plural form by addition of a suffix: compare ꌥ *daryā* 'valleys' (ꌥ *darī* 'valley' [f.s.] + ꌥ *yā* [f.pl.suf]) and ꌥ *daryā* 'ocean'. It is also used with inflectional suffixes: compare ꌥ *ācāryās* 'to the cook' (ꌥ *ācārī* 'cook' + ꌥ *-as* [dat.suf]) and ꌥ *ācāryās* 'to the teacher' (ꌥ *ācārya* 'teacher' + ꌥ *-as* [dat.suf]). The distinction is also phonemic: the syllabic division of ꌥ *daryā* 'valleys' is /da.ryā/, that of ꌥ *daryā* 'ocean' is /dar.yā/. The /ryā/ in ꌥ is pronounced with a palatalized /r/, which is represented by the *eyelash RA*.

   Additionally, in prosody, as Madhav Deshpande (2011, private correspondence) explains, the *repha* in a non-Marathi RA-initial cluster turns the preceding short vowel into a *guru* 'heavy' syllable, while this change does not occur in a Marathi cluster. Accordingly, the first syllable of ꌥ /da.ryā/ is *laghu* 'light', while the first syllable of ꌥ /dar.yā/ is *guru* 'heavy'.

   These features of *eyelash RA* suggest that it may be encoded as an independent character, distinct from ꌃ RA; see Muller 2004 for additional justification for encoding the character separately in Devanagari. However, an *EYELASH RA character is *not* proposed for independent encoding in Modi for the following reasons:

i. The ⁼○ *eyelash RA* is a half-form of a letter that does not exist. Hypothetically, it could be considered a half-form of र, a contextual form of RA that is used in certain consonant-vowel ligatures (see Section 4.7). Its graphical structure supports this claim: <र, ○ VIRAMA> → ⁼○. However, this र form is never written independently.

ii. It is unclear how an *EYELASH RA would be used in writing a conjunct. Would it be treated as a combining sign or would it be treated as a letter that would be followed by VIRAMA?

iii. The writing of *eyelash RA* using 🆉🆆 ZWJ is an established convention in Devanagari. While precedent may not offer the most suitable rationale for retaining a particular implementation, especially when there are reasonable grounds for a different approach, in this case it will be beneficial to users of Modi, who are Marathi speakers and existing users of Devanagari familiar with the representation of *eyelash RA*.

It is recommended that the encoded representation of *eyelash RA* in Modi be identical to that in Devanagari: < J RA, ○ VIRAMA, 🆉🆆 ZWJ>.

2. **Conjunct Non-Initial**   When J RA is a non-initial element in a conjunct, it is represented in the following ways. The encoded representation of all these forms is <C, ○ VIRAMA, RA>. Shaping and glyph substitutions are managed in the font.

   (a) *Rakār*   The subjoined form ⌐○ *rakār*. For instance *pra*:

   (b) *Vattu*   The subjoined form ○ *vattu*. For instance *ṣṭra*:

   (c) Special ligatures   Some consonant clusters with RA are represented as distinct ligatures. The most common of these is the conjunct <TA, VIRAMA, RA> *tra*:

3. **Word Level**   The form of RA is also influenced at the word level depending upon the characters in the preceding syllable. The forms described below are to be managed in the font using substitutions.

   (a) र KA   Sequences of <KA, RA> are represented using distinct ligatures: As shown below, the sequence <KA, RA> *kara* is written using the ligature 5̣:

   Contrast *kara* with <KA, VIRAMA, RA> *kra*:

   The sequence <KA, -AA, RA> *kāra* is written as the ligature 5̣̄:

[Modi script text]

Vowel signs that modify RA in such ligatures are written normally, eg. <KA, RA, -AA> *karā*:

[Modi script text]

(b) ꢸ SA   When RA occurs after SA (<SA, RA>) it takes the shape ◌꣄꣄:

[Modi script text]

(c) *Other letters*   In sequences with consonant letters that have a right descender, RA takes the shape ◌꣄. For instance <TA, RA> *tara*:

[Modi script text]

Contrast the above with <TA, VIRAMA, RA> *tra*:

[Modi script text]

(d) *Rounded letters*   After rounded letters such as ꢦ KHA, ꢥ DHA, ꢲ HA, the letter RA takes the shape ◌꣄. Shown below are the sequences <KHA, RA> and <HA, -E, RA>, respectively:

[Modi script text]

As indicated above, the change in the shape of RA in these environments is not suppressed by the presence of a vowel sign attached to the preceding consonant.

(e) *Vowel signs*   When RA follows vowels signs whose graphical structure contains vertical descenders, such as -AA, -II, and -O, it takes the shape ◌꣄. For example, observe its form in the word <TA, -AA, RA, NNA, -O, RA> *tāraṇora*:

[Modi script text]

When it occurs after -II, it attaches to the bottom of the vowel sign. The word *jahāgīrī* below contains the sequence <-II, RA, -II>:

[Modi script text]

When RA occurs after the vowel sign -U, the two characters form the ligature ◌꣄. The sequence <PA, -U, RA> *pura* appears below:

Vowel signs that modify RA in this ligature are written in the normal practice. For instance, <PA, -U, RA, -E> *pure*:

(f) *Punctuation*   When written after DANDA or DOUBLE DANDA in certain environments, RA is written as ◌ and attached to the terminal of the DANDA. See Section 4.12 for additional details.

4.9.1 Nasalization

Nasalization is indicated using ◌̇ ANUSVARA.

4.10 Ruled Lines and Headstrokes

Several manuscripts and books show Modi with a head-line similar to that of Devanagari. The line is not a headstroke, but a typographic device. In general practice, lines are drawn upon a page prior to writing. The text is then written with the Modi letters hanging from the line. In some cases, lines are used in Modi text for emphasis, titling, or sectioning.

The specimen below shows the indiscriminate use of ruled lines. The text is written upon ruled lines, but numbers are written freely, except for the last value, which is a total of the preceding figures:

Although the line is not part of a character's graphical structure it was incorporated into the design of glyphs for the first Modi metal fonts. This was likely an attempt to represent the appearance of hand-written Modi documents. The practice of designing Modi glyphs with headstrokes continues in digital typography.

If there is a requirement to represent Modi without the headstroke, font developers should produce sets of Modi fonts with head-strokes and without. This will facilitate the representation of Modi text that is not written under ruled lines.

4.11   Word and Section Boundaries

Traditionally, word boundaries are generally not marked in Modi. This is on account of Modi being an administrative script, which is characterized by the practice of rapid writing without lifting the pen. Paragraph and other section boundaries are indicated in some Modi documents through the use of whitespace. Modern practice uses spaces and various punctuation conventions, including DANDA and Western marks.

Spacing in a typical Modi document looks like:

The practice of marking such boundaries depended upon the scribe or the occasion. The following specimen is written with word-spacing; it is not known if spacing was introduced intentionally, perhaps in order to enhance legibility:

The manner of marking word boundaries changes between printed and written Modi. In some printed books, word and sentence boundaries are marked. In the specimen below, word boundaries are indicated with spaces and the ends of sentences are marked using a period:

4.12   Punctuation

Modi books printed by the American Tract Society use Latin punctuation, such as comma and semi-colons:

More commonly used are the | DANDA and || DOUBLE DANDA.
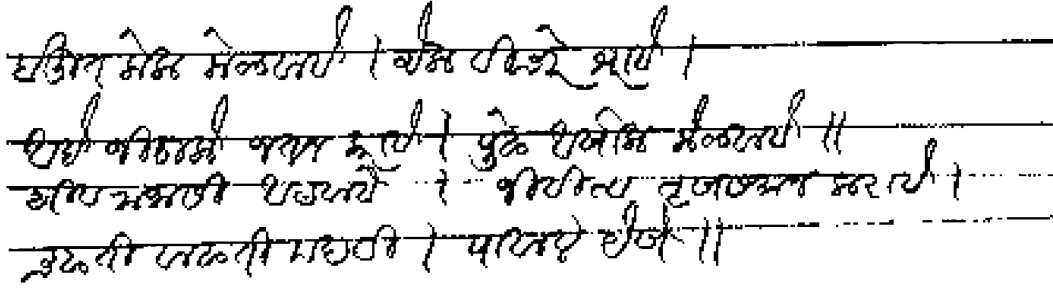
These characters exhibit special behavior in Modi. For example, in the figure below the sequence <DOUBLE DANDA, RA> results in the ligation of the two characters using a contextual form of RA:

The usage of double *daṇḍā* in such cases requires that distinctive *daṇḍā*-s be encoded for Modi. Therefore, the Modi | DANDA and || DOUBLE DANDA are proposed for independent encoding and disunification from the corresponding Devanagari characters on account of the special behavior required of these characters. The ligating behavior (eg. <|| DOUBLE DANDA, ꝩ RA> → ||) is to be handled at the font level.

4.13   Abbreviations

Abbreviations are indicated using the ◦ ABBREVIATION SIGN:
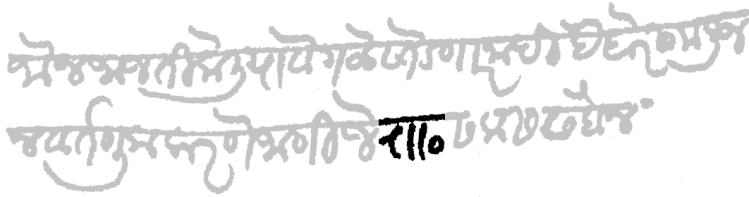
4.14   Various Signs

1. ̆ ARDHACANDRA   This sign is used for transcribing sounds used in English names and loanwords.

2. ◌: VISARGA   This sign represents an allophone of ꝩ RA or ꝩ SA at word-final position in Sanskrit orthography.

4.15   Special Signs

Several special signs are used in Modi documents for invocations. The Devanagari श्री *śrī* is written at the top of letters and other correspondence. There are also several Modi-specific signs, such as ꝩ *īṃ*. Some

of these signs are shown in figures 23 and 24. Additional research is required in order to determine the meaning, usage, and suitablility of encoding these characters. Of these, only the following character is currently proposed for encoding.

1. ﻝ HUVA   This sign is written as an invocation in several Modi documents. It is derived from the Arabic هوا *huva*.

ﻝ

प्राप्तन्न

एकमछत्तरमछघयतचश्व
गाघघनछफार्घेठऔनिफ़्ली
मयांरबेनरा्ठगेरघनछ्घ

4.16   Digits

Modi has a full set of decimal digits: ० ZERO, १ ONE, २ TWO, ३ THREE, ४ FOUR, ५ FIVE, ६ SIX, ७ SEVEN, ८ EIGHT, ९ NINE.

4.17   Number Forms and Unit Marks

Several number forms and unit marks are used for writing Modi. The majority of these are to be represented using characters already encoded in the "Common Indic Number Forms" block (U+A830 .. U+A83F):[1]

1. *Currency signs*   Currency values are written using ﻝ U+A838 NORTH INDIC RUPEE MARK.

2. *Fraction signs*   Basic fraction signs used in Modi are identical to those used in other northern Indic scripts. The signs are part of a base-16 system:

   I   U+A830 NORTH INDIC FRACTION ONE QUARTER

   II   U+A831 NORTH INDIC FRACTION ONE HALF

   III   U+A832 NORTH INDIC FRACTION THREE QUARTERS

   ⌐   U+A833 NORTH INDIC FRACTION ONE SIXTEENTH

   =   U+A834 NORTH INDIC FRACTION ONE EIGHTH

   ≡   U+A835 NORTH INDIC FRACTION THREE SIXTEENTHS

   The absence of intermediate units is indicated by the ⌅ U+A837 NORTH INDIC PLACEHOLDER MARK, which is called आळी *āḷī* or अळी *āḷī* in Marathi.

   The ० U+A836 NORTH INDIC QUARTER MARK is used for representing *ānnā* values.

3. *Other signs*   There are other number forms in Modi documents that are not yet encoded in the UCS. One of these is '÷', which appears to be associated with fraction notation:

---

[1] See Pandey 2007 for detailed descriptions of these characters.

The following figure shows the use of ⟨ NORTH INDIC PLACEHOLDER MARK as well as '÷':

The Modi '÷' is not currently proposed for encoding. Additional research is required in order to determine its meaning.

4.18   Collation

The primary collating order for Modi is as follows:

ॲ A < ॐ AA < ऎ I < ईं II < ॉ U < ॐ UU < ॠ VOCALIC R < ॠ VOCALIC RR < ॡ VOCALIC L < ॡ VOCALIC LL < ऎ E < ऎं AI < पे O < पं AU < क KA < ख KHA < ग GA < घ GHA < ङ NGA < च CA < छ CHA < ज JA < झ JHA < ञ NYA < ट TTA < ठ TTHA < ड DDA < ढ DDHA < ण NNA < त TA < थ THA < द DA < ध DHA < न NA < प PA < फ PHA < ब BA < भ BHA < म MA < य YA < र RA < ल LA < व VA < श SHA < ष SSA < स SA < ह HA < ळ LLA < ा SIGN AA < ि SIGN I < ी SIGN II < ु SIGN U < ू SIGN UU < ृ SIGN VOCALIC R < ॄ SIGN VOCALIC RR < ॢ SIGN VOCALIC L < ॣ SIGN VOCALIC LL < े SIGN E < ै SIGN AI < ो SIGN O < ौ SIGN AU

The following characters have secondary weights: ं ANUSVARA, ः VISARGA, ् VIRAMA, ̆ ARDHACANDRA.

5   Character Data

5.1   Character Properties

The properties for Modi characters in the Unicode Character Database format are:

```
11600;MODI LETTER A;Lo;0;L;;;;;N;;;;;
```

14

```
11601;MODI LETTER AA;Lo;0;L;;;;;N;;;;;
11602;MODI LETTER I;Lo;0;L;;;;;N;;;;;
11603;MODI LETTER II;Lo;0;L;;;;;N;;;;;
11604;MODI LETTER U;Lo;0;L;;;;;N;;;;;
11605;MODI LETTER UU;Lo;0;L;;;;;N;;;;;
11606;MODI LETTER VOCALIC R;Lo;0;L;;;;;N;;;;;
11607;MODI LETTER VOCALIC RR;Lo;0;L;;;;;N;;;;;
11608;MODI LETTER VOCALIC L;Lo;0;L;;;;;N;;;;;
11609;MODI LETTER VOCALIC LL;Lo;0;L;;;;;N;;;;;
1160A;MODI LETTER E;Lo;0;L;;;;;N;;;;;
1160B;MODI LETTER AI;Lo;0;L;;;;;N;;;;;
1160C;MODI LETTER O;Lo;0;L;;;;;N;;;;;
1160D;MODI LETTER AU;Lo;0;L;;;;;N;;;;;
1160E;MODI LETTER KA;Lo;0;L;;;;;N;;;;;
1160F;MODI LETTER KHA;Lo;0;L;;;;;N;;;;;
11610;MODI LETTER GA;Lo;0;L;;;;;N;;;;;
11611;MODI LETTER GHA;Lo;0;L;;;;;N;;;;;
11612;MODI LETTER NGA;Lo;0;L;;;;;N;;;;;
11613;MODI LETTER CA;Lo;0;L;;;;;N;;;;;
11614;MODI LETTER CHA;Lo;0;L;;;;;N;;;;;
11615;MODI LETTER JA;Lo;0;L;;;;;N;;;;;
11616;MODI LETTER JHA;Lo;0;L;;;;;N;;;;;
11617;MODI LETTER NYA;Lo;0;L;;;;;N;;;;;
11618;MODI LETTER TTA;Lo;0;L;;;;;N;;;;;
11619;MODI LETTER TTHA;Lo;0;L;;;;;N;;;;;
1161A;MODI LETTER DDA;Lo;0;L;;;;;N;;;;;
1161B;MODI LETTER DDHA;Lo;0;L;;;;;N;;;;;
1161C;MODI LETTER NNA;Lo;0;L;;;;;N;;;;;
1161D;MODI LETTER TA;Lo;0;L;;;;;N;;;;;
1161E;MODI LETTER THA;Lo;0;L;;;;;N;;;;;
1161F;MODI LETTER DA;Lo;0;L;;;;;N;;;;;
11620;MODI LETTER DHA;Lo;0;L;;;;;N;;;;;
11621;MODI LETTER NA;Lo;0;L;;;;;N;;;;;
11622;MODI LETTER PA;Lo;0;L;;;;;N;;;;;
11623;MODI LETTER PHA;Lo;0;L;;;;;N;;;;;
11624;MODI LETTER BA;Lo;0;L;;;;;N;;;;;
11625;MODI LETTER BHA;Lo;0;L;;;;;N;;;;;
11626;MODI LETTER MA;Lo;0;L;;;;;N;;;;;
11627;MODI LETTER YA;Lo;0;L;;;;;N;;;;;
11628;MODI LETTER RA;Lo;0;L;;;;;N;;;;;
11629;MODI LETTER LA;Lo;0;L;;;;;N;;;;;
1162A;MODI LETTER VA;Lo;0;L;;;;;N;;;;;
1162B;MODI LETTER SHA;Lo;0;L;;;;;N;;;;;
1162C;MODI LETTER SSA;Lo;0;L;;;;;N;;;;;
1162D;MODI LETTER SA;Lo;0;L;;;;;N;;;;;
1162E;MODI LETTER HA;Lo;0;L;;;;;N;;;;;
1162F;MODI LETTER LLA;Lo;0;L;;;;;N;;;;;
11630;MODI VOWEL SIGN AA;Mc;0;L;;;;;N;;;;;
11631;MODI VOWEL SIGN I;Mc;0;L;;;;;N;;;;;
11632;MODI VOWEL SIGN II;Mc;0;L;;;;;N;;;;;
11633;MODI VOWEL SIGN U;Mn;0;NSM;;;;;N;;;;;
11634;MODI VOWEL SIGN UU;Mn;0;NSM;;;;;N;;;;;
11635;MODI VOWEL SIGN VOCALIC R;Mn;0;NSM;;;;;N;;;;;
11636;MODI VOWEL SIGN VOCALIC RR;Mn;0;NSM;;;;;N;;;;;
11637;MODI VOWEL SIGN VOCALIC L;Mn;0;NSM;;;;;N;;;;;
11638;MODI VOWEL SIGN VOCALIC LL;Mn;0;NSM;;;;;N;;;;;
11639;MODI VOWEL SIGN E;Mn;0;NSM;;;;;N;;;;;
1163A;MODI VOWEL SIGN AI;Mn;0;NSM;;;;;N;;;;;
1163B;MODI VOWEL SIGN O;Mc;0;L;;;;;N;;;;;
1163C;MODI VOWEL SIGN AU;Mc;0;L;;;;;N;;;;;
1163D;MODI SIGN ANUSVARA;Mn;0;NSM;;;;;N;;;;;
```

```
1163E;MODI SIGN VISARGA;Mc;0;L;;;;;;N;;;;;
1163F;MODI SIGN VIRAMA;Mn;9;L;;;;;N;;;;;
11640;MODI SIGN ARDHACANDRA;Mn;0;NSM;;;;;N;;;;;
11641;MODI DANDA;Po;0;L;;;;;N;;;;;
11642;MODI DOUBLE DANDA;Po;0;L;;;;;N;;;;;
11643;MODI ABBREVIATION SIGN;Po;0;L;;;;;N;;;;;
11644;MODI SIGN HUVA;So;0;L;;;;;N;;;;;
11650;MODI DIGIT ZERO;Nd;0;L;;0;0;0;N;;;;;
11651;MODI DIGIT ONE;Nd;0;L;;1;1;1;N;;;;;
11652;MODI DIGIT TWO;Nd;0;L;;2;2;2;N;;;;;
11653;MODI DIGIT THREE;Nd;0;L;;3;3;3;N;;;;;
11654;MODI DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;;
11655;MODI DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;;
11656;MODI DIGIT SIX;Nd;0;L;;6;6;6;N;;;;;
11657;MODI DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;;
11658;MODI DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;;
11659;MODI DIGIT NINE;Nd;0;L;;9;9;9;N;;;;;
```

5.2   Linebreaking

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

```
11600..1162F; AL   # LETTER A .. LETTER LLA
11630..11640; CM   # SIGN AA .. SIGN ARDHACANDRA
11641..11642; BA   # DANDA .. DOUBLE DANDA
11643; BA          # ABBREVIATION SIGN
11644; AL          # SIGN HUVA
11650..11659; NU   # DIGIT ZERO .. DIGIT NINE
```

5.3   Syllabic Categories

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

```
# Indic_Syllabic_Category=Bindu
1163D        ; Bindu              # Mn        SIGN ANUSVARA

# Indic_Syllabic_Category=Visarga
1163E        ; Visarga            # Mc        SIGN VISARGA

# Indic_Syllabic_Category=Virama
1163F        ; Virama             # Mn        SIGN VIRAMA

# Indic_Syllabic_Category=Vowel_Independent
11600..1160D ; Vowel_Independent # Lo   [14] LETTER A .. LETTER AU

# Indic_Syllabic_Category=Vowel_Dependent
11630..11632 ; Vowel_Dependent   # Mc    [3] VOWEL SIGN AA .. VOWEL SIGN II
11633..1163A ; Vowel_Dependent   # Mn    [8] VOWEL SIGN U .. VOWEL SIGN AI
1163B..1163C ; Vowel_Dependent   # Mc    [2] VOWEL SIGN O .. VOWEL SIGN AU

# Indic_Syllabic_Category=Consonant
1160E..1162F ; Consonant          # Lo   [34] LETTER KA .. LETTER LLA
```

5.4   Matra Categories

Matra categories given in the data format of `IndicMatraCategory.txt`:

```
# Indic_Matra_Category=Right
11630          ; Right  # Mc       VOWEL SIGN AA
11632          ; Right  # Mc       VOWEL SIGN II
1163B..1163C  ; Right  # Mc   [2] VOWEL SIGN O .. VOWEL SIGN AU

# Indic_Matra_Category=Left
11631          ; Left   # Mc       VOWEL SIGN I

# Indic_Matra_Category=Top
11639..1163A   ; Top    # Mn   [2] VOWEL SIGN E .. VOWEL SIGN AI

# Indic_Matra_Category=Bottom
11633..11638   ; Bottom # Mn   [6] VOWEL SIGN U .. VOWEL SIGN VOCALIC LL
1163F          ; Bottom # Mn       SIGN VIRAMA
```

5.5 'Confusable' Characters

Below is a list of Modi characters that may resemble others encoded in the UCS:

```
11608 MODI LETTER VOCALIC L        ; 090C DEVANAGARI LETTER VOCALIC L
11609 MODI LETTER VOCALIC LL       ; 0961 DEVANAGARI LETTER VOCALIC LL
11610 MODI LETTER GA               ; 0917 DEVANAGARI LETTER GA
11611 MODI LETTER GHA              ; 0918 DEVANAGARI LETTER GHA
11612 MODI LETTER NGA              ; 0919 DEVANAGARI LETTER NGA
11614 MODI LETTER CHA              ; 091B DEVANAGARI LETTER CHA
11616 MODI LETTER JHA              ; 091D DEVANAGARI LETTER JHA
11617 MODI LETTER NYA              ; 091E DEVANAGARI LETTER NYA
11625 MODI LETTER BHA              ; 092D DEVANAGARI LETTER BHA
1162C MODI LETTER SSA              ; 0937 DEVANAGARI LETTER SSA
1162F MODI LETTER LLA              ; 0933 DEVANAGARI LETTER LLA
11631 MODI VOWEL SIGN I            ; 0940 DEVANAGARI VOWEL SIGN II
11633 MODI VOWEL SIGN U            ; 0941 DEVANAGARI VOWEL SIGN U
11634 MODI VOWEL SIGN UU           ; 0942 DEVANAGARI VOWEL SIGN UU
11635 MODI VOWEL SIGN VOCALIC R    ; 0943 DEVANAGARI VOWEL SIGN VOCALIC R
11636 MODI VOWEL SIGN VOCALIC RR   ; 0944 DEVANAGARI VOWEL SIGN VOCALIC RR
11637 MODI VOWEL SIGN VOCALIC L    ; 0962 DEVANAGARI VOWEL SIGN VOCALIC L
11638 MODI VOWEL SIGN VOCALIC LL   ; 0963 DEVANAGARI VOWEL SIGN VOCALIC LL
11639 MODI VOWEL SIGN E            ; 0947 DEVANAGARI VOWEL SIGN E
1163D MODI SIGN ANUSVARA           ; 0902 DEVANAGARI SIGN ANUSVARA
1163E MODI SIGN VISARGA            ; 0903 DEVANAGARI SIGN VISARGA
1163F MODI SIGN VIRAMA             ; 094D DEVANAGARI SIGN VIRAMA
11643 MODI ABBREVIATION SIGN       ; 0970 DEVANAGARI ABBREVIATION SIGN
11645 MODI SIGN HUVA               ; 0644 ARABIC LETTER LAM
11650 MODI DIGIT ZERO              ; 0966 DEVANAGARI DIGIT ZERO
```

Modi digits (U+11650..U+11659) may be considered 'confusable' with Devanagari digits (U+0966..U+096F).

6   References

American Tract Society. 1837. "First Book". Bombay.

———. 1837. *Yacavicara*. Bombay.

Carey, William. 1810. *A dictionary of the Mahratta language*. Serampore: Serampore Missionary Press.

Daily News and Analysis. 2011. "With 4,000 followers, Modi script is going strong" (May 6, 2011). http://www.dnaindia.com/india/report_with-4000-followers-modi-script-is-going-strong_1540025

Dīkṣit, Moreśvar G. and V. G. Khobarekar. 1969. मराठेकालीन प्रसिद्ध व्यक्तींची हस्ताक्षरयुक्त पत्र [*Marāṭhekālīna prasiddha vyaktīṃcī hastākṣarayukta patra*]. पुराभिलेख विभाग — ऐतिहासिक प्रकाशन क्र॰ १. मुंबई: शासकीय मध्यवर्ती मुद्रणालय.

Faulmann, Carl. 1880. *Das Buch der Schrift: Enthaltend die Schriftzeichen und Alphabete aller Zeiten und aller Völker der Erdkreises*. Zweite Vermehrte und verbesserte Auflage. Wein: Der Kaiserlich-Königlichen Hof- und Staatsdruckerei.

Grierson, George A. 1905. *The Linguistic Survey of India*. Vol. VII. Indo-Aryan Family. Southern Group. Specimens of the Marathi language. Calcutta: Office of the Superintendent of Government Printing, India.

Gupte, B. A. 1905. "The Modi Character". *Indian Antiquary*, vol. XXXIV, p. 27–30.

Kulkarṇī, Madhukar. 1999. अक्षर लेणी [*Akṣara leṇī*]. मराठ्यांच्या इतिहासातील ओजस्वी पत्रांचा मोडी-मराठी लिप्यंतर संग्रह. Puṇe: Mānasanmāna Prakāśana.

Kunte, B. G. [ed]. 1987. *Illustrative Modi Documents*. Maharashtra State Archives, Government of Maharashtra. Bombay: Government Central Press.

म्हात्रे, कृष्णाजी [Mhātre, Kṛṣṇājī]. 2009. मोडी लिप्यंतर कौशल्य (*Moḍī Lipyaṃtar Kauśalya*). मुंबई: शील प्रकाशन.

Muller, Eric. 2004. "Devanagari Eyelash Ra". `http://unicode.org/~emuller/iwg/p8/utcdoc.html`

Naik, Bapurao S. 1971. *Typography of Devanagari*. 1st rev. ed. Vol. 1. Bombay: Directorate of Languages, Government of Maharashtra.

Navalkar, Ganpatrao R. 1925. *The Student's Maráṭhí Grammar*. 4th ed. Poona: Scottish Mission Press.

Ojhā, Gaurīśaṅkara Hīrācanda. 1971. भारतिय प्राचीन लिपिमाला [*Bhāratīya prācīna lipimālā* = The Palæography of India]. Reprint of the rev. and enl. 2nd ed., 1918; first ed. published in 1894 under the title *Prācīna lipimālā*. New Delhi: Munshiram Manoharlal.

Pandey, Anshuman. 2007. "Proposal to Encode North Indic Number Forms in ISO/IEC 10646." ISO/IEC JTC1/SC2/WG2 N3367 L2/07-354. October 7, 2007. `http://std.dkuug.dk/jtc1/sc2/wg2/docs/n3367.pdf`

———. 2009. "Preliminary Proposal to Encode the Modi Script in ISO/IEC 10646." L2/09-249R. `http://www.umich.edu/~pandey/09249-modi.pdf`

———. 2010. "Revised Code Chart and Names List for the Modi Script." ISO/IEC JTC1/SC2/WG2 N3780 L2/10-084. `http://std.dkuug.dk/jtc1/sc2/wg2/docs/n3780.pdf`

पठवर्धन [Paṭhavardhana]. nd. 𑘦𑘻𑘓𑘱 𑘩𑘱𑘢𑘿𑘧𑘽𑘝𑘨 पुस्तक (*Moḍī Pahile Pustak*) [= First Book for Modi]. पुणे: दत्तत्रेय परशुराम जोशी.

Ross, Fiona. 1999. *The Printed Bengali Character and its Evolution*. Richmond: Curzon.

Strandberg, Elisabeth [ed]. 1983. *The Moḍī Documents from Tanjore in Danish Collections*. Beiträge zur Südasienforschung, Südasien-Institut, Universität Heidelberg, Band 81. Wiesbaden: Franz Steiner Verlag.

Times of India. 2011. "25 women help decipher Modi scripts" (March 9, 2011). `http://articles.timesofindia.indiatimes.com/2011-03-09/pune/28671782_1_modi-script-peshwa-era`

7 Acknowledgments

	1160	1161	1162	1163	1164	1165
0	ए 11600	ग 11610	ध 11620	◌ा 11630	◌̆ 11640	० 11650
1	ध 11601	घ 11611	४ 11621	◌ी 11631	◌। 11641	१ 11651
2	ध 11602	ङ 11612	ण 11622	◌ी 11632	॥ 11642	२ 11652
3	ध 11603	उ 11613	भ्र 11623	◌ु 11633	० 11643	३ 11653
4	ओ 11604	छ 11614	ध 11624	◌ू 11634	J 11644	४ 11654
5	ओ 11605	ख 11615	म 11625	◌ृ 11635		५ 11655
6	ऋ 11606	झ 11616	म 11626	◌ॄ 11636		६ 11656
7	ॠ 11607	ञ 11617	ए 11627	◌ॢ 11637		७ 11657
8	ऌ 11608	ट 11618	ज 11628	◌ॣ 11638		८ 11658
9	ॡ 11609	ठ 11619	४ 11629	◌े 11639		९ 11659
A	ऐ 1160A	ड 1161A	प 1162A	◌ो 1163A		
B	ऐ 1160B	ढ 1161B	श 1162B	◌ो 1163B		
C	ऐ 1160C	ज 1161C	ष 1162C	◌ौ 1163C		
D	ऐ 1160D	त 1161D	४ 1162D	◌ं 1163D		
E	ऱ 1160E	थ 1161E	ठ 1162E	◌ः 1163E		
F	ऐ 1160F	द 1161F	क 1162F	◌ॗ 1163F		

Figure 1: Proposed code chart for Modi.

Independent vowels

11600	ॲ	MODI LETTER A
11601	ॲ	MODI LETTER AA
11602	ॲ	MODI LETTER I
11603	ॲ	MODI LETTER II
11604	ॲ	MODI LETTER U
11605	ॲ	MODI LETTER UU
11606	ॠ	MODI LETTER VOCALIC R
11607	ॠ	MODI LETTER VOCALIC RR
11608	ॡ	MODI LETTER VOCALIC L
11609	ॡ	MODI LETTER VOCALIC LL
1160A	ॲ	MODI LETTER E
1160B	ॲ	MODI LETTER AI
1160C	ॲ	MODI LETTER O
1160D	ॲ	MODI LETTER AU

Consonants

1160E	क	MODI LETTER KA
1160F	ख	MODI LETTER KHA
11610	ग	MODI LETTER GA
11611	घ	MODI LETTER GHA
11612	ङ	MODI LETTER NGA
11613	च	MODI LETTER CA
11614	छ	MODI LETTER CHA
11615	ज	MODI LETTER JA
11616	झ	MODI LETTER JHA
11617	ञ	MODI LETTER NYA
11618	ट	MODI LETTER TTA
11619	ठ	MODI LETTER TTHA
1161A	ड	MODI LETTER DDA
1161B	ढ	MODI LETTER DDHA
1161C	ण	MODI LETTER NNA
1161D	त	MODI LETTER TA
1161E	थ	MODI LETTER THA
1161F	द	MODI LETTER DA
11620	ध	MODI LETTER DHA
11621	न	MODI LETTER NA
11622	प	MODI LETTER PA
11623	फ	MODI LETTER PHA
11624	ब	MODI LETTER BA
11625	भ	MODI LETTER BHA
11626	म	MODI LETTER MA
11627	य	MODI LETTER YA
11628	र	MODI LETTER RA
11629	ल	MODI LETTER LA
1162A	व	MODI LETTER VA
1162B	श	MODI LETTER SHA
1162C	ष	MODI LETTER SSA
1162D	स	MODI LETTER SA
1162E	ह	MODI LETTER HA
1162F	ळ	MODI LETTER LLA

Dependent vowel signs

11630	ॊ	MODI VOWEL SIGN AA
11631	ॊ	MODI VOWEL SIGN I
11632	ॊ	MODI VOWEL SIGN II
11633	ॊ	MODI VOWEL SIGN U
11634	ॊ	MODI VOWEL SIGN UU
11635	ॊ	MODI VOWEL SIGN VOCALIC R
11636	ॊ	MODI VOWEL SIGN VOCALIC RR
11637	ॊ	MODI VOWEL SIGN VOCALIC L
11638	ॊ	MODI VOWEL SIGN VOCALIC LL
11639	ॊ	MODI VOWEL SIGN E
1163A	ॊ	MODI VOWEL SIGN AI
1163B	ॊ	MODI VOWEL SIGN O
1163C	ॊ	MODI VOWEL SIGN AU

Various signs

1163D	◌	MODI SIGN ANUSVARA

1163E	◌ः	MODI SIGN VISARGA
1163F	◌	MODI SIGN VIRAMA
		= halant
11640	◌	MODI SIGN ARDHACANDRA
		• used for transcribing English

Punctuation

11641	।	MODI DANDA
11642	॥	MODI DOUBLE DANDA

Sign

11643	◦	MODI ABBREVIATION SIGN

Invocation sign

11644	॒	MODI SIGN HUVA

Digits

11650	०	MODI DIGIT ZERO
11651	१	MODI DIGIT ONE
11652	२	MODI DIGIT TWO
11653	३	MODI DIGIT THREE
11654	४	MODI DIGIT FOUR
11655	५	MODI DIGIT FIVE
11656	६	MODI DIGIT SIX
11657	७	MODI DIGIT SEVEN
11658	८	MODI DIGIT EIGHT
11659	९	MODI DIGIT NINE

Figure 2: Proposed names list for Modi.

The following table lists regular consonant-vowel combinations in the Modi script. The individual Modi glyphs are shown in the source; their presence is indicated below (✓ = form shown, (✓) = form shown in parentheses in the source).

	A	AA	I	II	U	UU	R	O	AU
KA	✓	✓			✓	(✓)	✓	✓	✓
KHA	✓	✓			✓	(✓)		✓	✓
GA	✓					✓			
CA	✓	✓			✓	✓		✓	✓
JA	✓	✓			✓	(✓)		✓	✓
JHA	✓	✓			✓	✓		✓	✓
NYA	✓	✓			✓	✓		✓	✓
TTA	✓	✓						✓	✓
TTHA	✓	✓						✓	✓
DDA	✓		✓	✓					
TA	✓				✓	(✓)			
THA	✓	✓			✓	✓		✓	✓
DA	✓	✓			✓	✓		✓	✓
DHA	✓	✓			✓	✓		✓	✓
NA	✓	✓			✓	(✓)		✓	✓
PA	✓	✓			✓	(✓)		✓	✓
PHA	✓	✓			✓	(✓)		✓	✓
BA	✓	✓			✓	✓		✓	✓
BHA	✓	✓						✓	✓
MA	✓	✓			✓	✓		✓	✓
YA	✓	✓			✓	(✓)		✓	✓
RA	✓	✓			✓	(✓)		✓	✓
LA	✓	✓	✓	(✓)	✓	✓		✓	✓
VA	✓	✓			✓	✓		✓	✓
SA	✓	✓			✓	✓		✓	✓
HA	✓	✓			✓	✓		✓	✓

Table 1: Regular consonant-vowel combinations in Modi.

22

	A	B		A	B		A	B		A	B
A	अ	अ	KA	क	क	DA	द	द	ZERO	०	०
AA	आ	आ	KHA	ख	ख	DHA	ध	ध	ONE	१	१
I	इ	इ	GA	ग	ग	NA	न	न	TWO	२	२
II	ई	ई	GHA	घ	घ	PA	प	प	THREE	३	३
U	उ	उ	NGA	—	ङ	PHA	फ	फ	FOUR	४	४
UU	ऊ	ऊ	CA	च	च	BA	ब	ब	FIVE	५	५
R	—	ऋ	CHA	छ	छ	BHA	भ	भ	SIX	६	६
RR	—	ॠ	JA	ज	ज	MA	म	म	SEVEN	७	७
L	—	ऌ	JHA	झ	झ	YA	य	य	EIGHT	८	८
LL	—	ॡ	NYA	ञ	ञ	RA	र	र	NINE	९	९
E	ए	ए	TTA	ट	ट	LA	ल	ल			
AI	ऐ	ऐ	TTHA	ठ	ठ	VA	व	व			
O	ओ	ओ	DDA	ड	ड	SHA	श	श			
AU	औ	औ	DDHA	ढ	ढ	SSA	ष	ष			
			NNA	ण	ण	SA	स	स			
			TA	त	त	HA	ह	ह			
			THA	थ	थ	LLA	ळ	ळ			

Table 2: Comparison of basic Modi characters in ('A') the 'Hemadree' font by Somesh Bartakke and ('B') the font designed by Anshuman Pandey.

Figure 3: Comparison of consonant-vowel combinations in Modi and Devanagari (from Strandberg 1983: 32, 33).

24
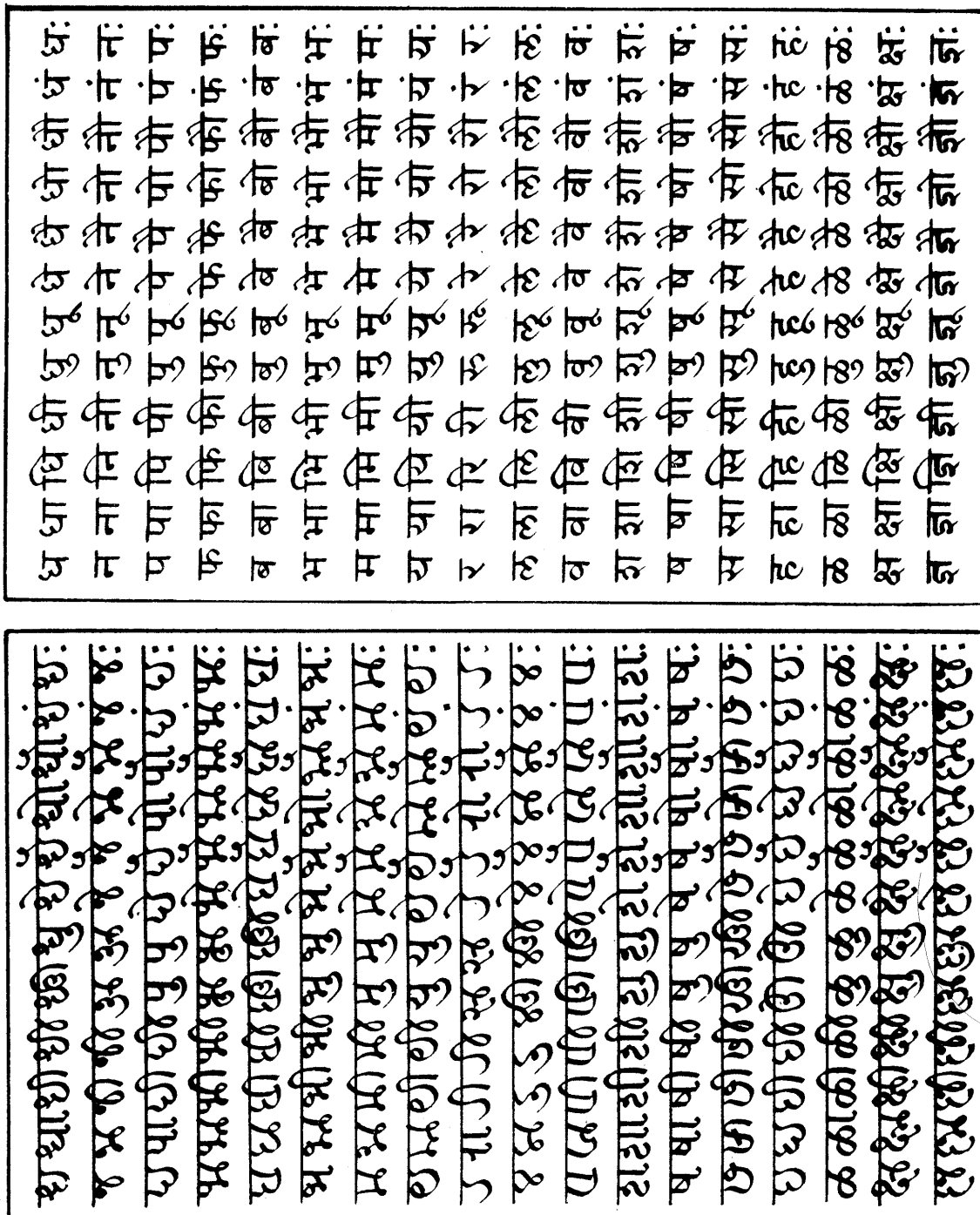
Figure 4: Comparison of consonant-vowel combinations in Modi and Devanagari (from Strandberg 1983: 34, 35).
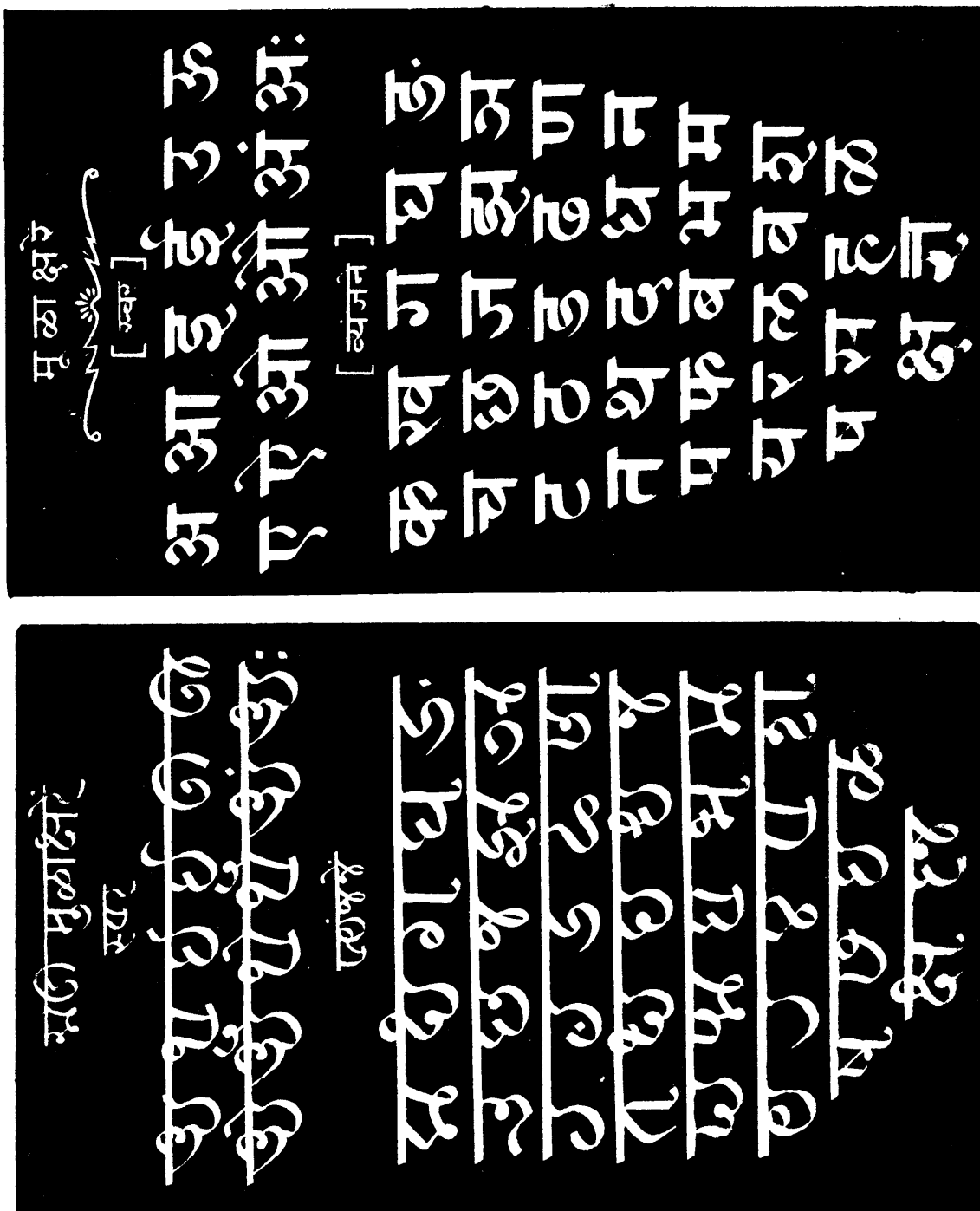
Figure 5: Comparison of consonant-vowel combinations in Modi and Devanagari (from Strandberg 1983: 30, 31).

Figure 6: Modi table (from Naik 1971: 130, 131).

27

Figure 7: Modi table (from Naik 1971: 132, 133).

Figure 8: Modi table (Naik 1971: 134).

Figure 9: Chart of Modi (from Ojhā 1999: 79).

1. MODI CHARACTERS.

1. The Modí letters are thus written and pronounced :—

Vowels.

या *a*, या *a*, ई *i*, ॳ *u*, ॉ *e*, ॉ *ai*, ॊ *o*, ॊ *au*, ॐ *am*, यः *ah*

Consonants.

ॡ *ka*	ॐ *da*	म *ma*
ॐ *kha*	ॐ *dha*	य *ya*
ग *ga*	ॐ *na*	ॐ *ra*
घ *gha*	ॐ *ta*	ॐ *la*
ॐ *na*	ॐ *tha*	व *va*
ॐ *cha*	ॐ *da*	श *sa*
ॐ *chha*	ॐ *dha*	ष *sha*
ॐ *ja*	न *na*	स *sa*
ॐ *jha*	ॐ *pa*	ॐ *ha*
ॐ *ña*	ॐ *pha*	ॐ *la*
ॐ *ta*	ब *ba*	क्ष *ksha*
ॐ *tha*	भ *bha*	ॐ *dnya*

2. The following are three Modí *Bárákhadís*, or modes of syllabification :—

ka ká ki kí ku ku ké kai ko kau kam kah.

या ॡ ॡ ॡ ॡ ॡ ॡ ॡ ॡ ॡ ॡ याः

ga gá gi gí gu gu ge gai go gau gam gah.

ग गा गी गी ॐ ॐ गे गे गो गो गं गः

ta tá tí tí tu tu te tai to tau tam tah.

ॐ ॐ ॐ ॐ ॐ ॐ ॐ ॐ ॐ ॐ ॐ ॐ

Figure 10: Chart of Modi printed in metal types (from Navalkar 1925: 374).

Marāṭhī is usually written in the so-called Bālbōdh or in the so-called Mōḍi character.[1]

Written character.   Bālbōdh, *lit.* 'teachable to children' is identical with Dēva-nāgarī, and has been described in Vol. v. Part ii, pp. 7 and ff. It is used in almost all printed books, and also, to a great extent, in private transactions and letters. The Mōḍī character is almost totally confined to the latter kind of writing. An example of its use will be found on pp. 259 and ff. It consists of the following signs :—

VOWELS.

ॐ *a*   ॠ *ā*   ई *i, ī,*   उ *u, ū,*

ए *ē*   ऐ *ai*   ओ *ō*   औ *au*   अं *aṁ*   अः *aḥ*

CONSONANTS.

ka	*kha*	*ya*	*gha*	*ṅa*
cha	*chha*	*ja*	*jha*	*ña*
ṭa	*ṭha*	*ḍa*	*ḍha*	*ṇa*
ta	*tha*	*da*	*dha*	*na*
pa	*pha*	*ba*	*bha*	*ma*
ya	*ra*	*la*	*va*	
śa	*sha*	*sa*	*ha*	
ḷa	*ksha*	*dnya*		

The forms of the vowels given above are only used at the beginning of words or syllables. When the vowels follow a consonant they are expressed by means of secondary signs in the same way as in the Dēvanāgarī alphabet. For the sake of teaching these signs the alphabet is disposed in *Bārākhᵃḍis*, or series of twelve letters, each containing a consonant combined with all possible vocalic sounds. Such Bārākhᵃḍis are :—

ka	*kā*	*ki*	*kī*	*ku*	*kū*	*kē*	*kai*	*kō*	*kau*	*kaṁ*	*kah*
ga	*gā*	*gi*	*gī*	*gu*	*gū*	*gē*	*gai*	*go*	*gau*	*gaṁ*	*gah*
ṭa	*ṭā*	*ṭi*	*ṭī*	*ṭu*	*ṭū*	*ṭē*	*ṭai*	*ṭō*	*ṭau*	*ṭaṁ*	*ṭah*
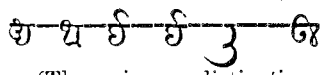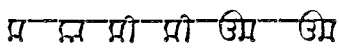
A short note on *the Modi Character* by B. A. Gupte will be found in the *Indian Antiquary*, Vol. xxxiv, 1905, pp. 27 and ff.

Figure 11: Chart of Modi from the *Linguistic Survey of India* (from Grierson 1905: 20).

COMPARATIVE MODIFICATIONS OF THE MODI CHARACTER.

Bâlbôdh. (Head-lines to Characters separately.)	*Môdi.* (Single head-line to all the Characters.)	*Bâlbôdh.* (Head-lines to Characters separately.)	*Môdi.* (Single head-line to all the Characters.)

The chief merit of the Môdi Character consists in the addition of **vowel signs**, thus:—

Bâlbôdh.

अ आ इ ई उ ऊ
a   â   i   î   u   û

Môdi.

(There is no distinction between long and short *i* and *u*.)

क का कि की कु कू
K   Kâ   Ki   Kî   Ku   Kû

(The whole line could be written without lifting the pen for the addition of the ↑ or ⌐.)

1 No change.   2 No change.   3 Change in the angular shape.
4 Angular.   5 Note *supra*; the change is in the position of the loop.
6 The loop is from left to right, and not from right to left as in.   7 Mark the loop.
8 Note (d and h); the difference is in the rounding of the right-hand upper corner in the first.
9 Interchangeable with above.

Figure 12: Chart of Modi from the *Indian Antiquary* (from Gupte 1905: 29).

MARAṬHISCH.

Zeichen	Wert	Zeichen	Wert	Zeichen	Wert	Zeichen	Wert	Zeichen	Wert
౮	a	‍	ka	‍	ńa	‍	na	‍	ḷa
‍	ā	‍	kha	‍	ṭa	‍	pa	‍	va
‍	i	‍	ga	‍	ḍa	‍	pha	‍	sa
‍	u	‍	gha	‍	ḍha	‍	ba	‍	śa
‍	ṛ	‍	ṅa	‍	ṇa	‍	bha	‍	ša
‍	e	‍	tša, tsa	‍	ta	‍	ma	‍	ha
‍	ai	‍	tšha,tsha	‍	tha	‍	ya		
‍	o	‍	dža, dza	‍	da	‍	ra		
‍	au	‍	džha, dzha	‍	dha	‍	la		

Ligaturen.

‍	swa	‍	pra	‍	wā	‍	kša	‍	nta
‍	stha	‍	pr	‍	hā	‍	kšya	‍	ddha
‍	tra	‍	šḍha	‍	tu	‍	mha	‍	kta
‍	džya	‍	kra	‍	tya	‍	dya		

Vokalverbindung: ‍ kā, ‍ ki, ‍ ku, ‍ ke, ‍ kai, ‍ ko, ‍ kau, ‍ rk, ‍ kā.

Die Maraṭhen, welche einen Theil des Dekhan erobert haben, sind die einzigen Bewohner dieses Theiles von Indien, welche sich der Sanskritsprache bedienen, sie schreiben dieselbe entweder mit *Devanagari* und nennen diese Schrift *Bālbodh* oder mit ihrer eigenen Schrift, die *Moḍ* heisst. Diese Schrift schwankt zwischen *Gudžaratisch* und *Devanagari*, die Form ‍ *ńa* scheint eine Zusammenziehung von *Sindh* ‍ *ka* und *Devanagari* ‍ zu sein. Mit der Devanagari hat sie die Verbindung der Zeichen sowohl durch die Schreiblinie (welche aber nicht wie bei uns unter, sondern über den Buchstaben hinläuft) als auch die Verbindungen der Konsonanten gemein, doch ist der herabreichende Strich noch nicht so konsequent ausgebildet und selbst bei manchen *ā* - Formen nicht isolirt, während andere mit der Devanagari übereinstimmen, z. B. ‍ *yā*. Die Unterscheidung zwischen *i* und *ī*, *u* und *ū* ist nicht zu bemerken, auch scheint kein eigenes Zeichen für *tha* vorhanden zu sein, denn dieses wird durch *ta* oder auch durch ‍ wiedergegeben.

Die Sanskritlaute *tš*, *dž* lauten nur vor *e*, *i* und *ī* in gleicher Weise, vor *a, o, u, ṛ, ḷ* wie *ts* und *dz*; um *tša* und *dža* zu schreiben wird von manchen Schreibern ein *ya* eingeschoben oder ein Punkt unter die Zeichen gesetzt.

Figure 13: Specimens of glyphs from a Modi metal font (from Faulmann 1880: 124).

In all cases the *akṣara* with short *i* and *u* may also be used for long *i* and *u* and vice versa. Among the variants occurring in our material we may list those which are not so easily recognisable:

Figure 14: Variant forms of consonant letters and consonant-vowel combinations (from Strandberg 1983: 36).

Moreover the following less recognisable ligatures, which are not represented on the charts at all, occur in our material.
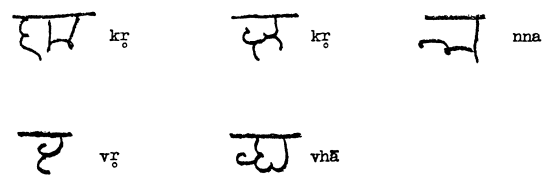
Figure 15: More variant forms of consonant-vowel combinations (from Strandberg 1983: 37).

Figure 16: Variant forms of consonant letters and consonant-vowel combinations. Courtesy of Rajesh Khilari.

समान दिसणारे शब्द
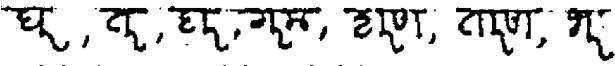
Figure 17: Modi letters and combinations that are graphically similar. Courtesy of Rajesh Khilari.
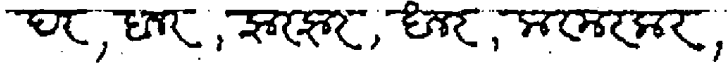
७. ` र ' ची करामत

` र ' हे मराठी व्यंजनमालेतील सत्ताविसावे व दंड (।) नसलेले एकमेव व्यंजन आहे. ` र ' चा पाढा वाचणे म्हणजे रडगाणे सांगणे किंवा दुबळेपणा दाखविणे पण मोडी लिपीत मात्र र चे कर्तृत्व अफाट आहे. या र मुळेच मोडी लिपीला एक वैभव प्राप्त होऊन तिच्यात ऐटबाजपणा व डौलदारपणा आला आहे. कसे ते पाहू —
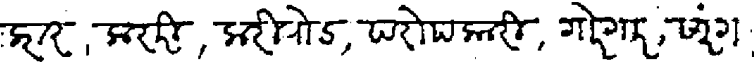
१) ज्या मोडी अक्षराला बालबोधप्रमाणे काना असतो, त्या कान्याला "⟨Modi⟩" असा ` र ' जोडणे.

उदा. : ⟨Modi examples⟩
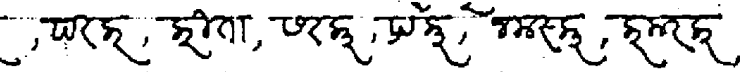
२) काना नसलेले मोडी अक्षर जेथे संपते तेथेच त्या अक्षराला ⟨Modi⟩ असा ` र ' जोडणे

उदा. : ⟨Modi examples⟩
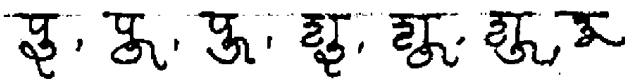
३) वरील १ व २ प्रकारे जोडून काढलेल्या ` र ' ला काना, वेलांटी, मात्रा, अनुस्वार देणे.

उदा. ⟨Modi examples⟩
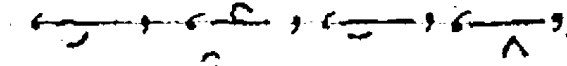
४) क किंवा का ह्या अक्षरांच्या पूर्ततेपूर्वी त्याला ⟨Modi⟩ असा ` र ' जोडणे.

⟨Modi examples⟩

५) बालबोध काना असलेल्या मोडी अक्षराला ` उ ' कार जोडून त्या ` उ ' काराला पुढे ` र ' जोडणे

उदा. : ⟨Modi examples⟩

६) ` र ' हे व्यंजन जोडाक्षरात आल्यास ते चार प्रकारांनी लिहिले जाते. त्यांची ⟨Modi marks⟩

अशी चिन्हे आहेत.

उदा. : ⟨Modi examples⟩
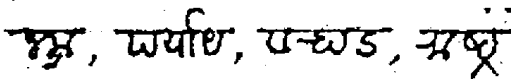
Figure 18: Description of the forms of ꣳ RA (from Mhātre 2009: 25). Continued in figure 19.

37

चार प्रकार खालीलप्रमाणे —

६अ) काही व्यंजनानंतर जोडाक्षरात 'र' आल्यास तो ꞉ असा लिहितात. तेथे एक गोष्ट लक्षात घ्यायला हवी की 'र' हे जोडाक्षरातील पुढे येणारे व्यंजन असेल तर व्र , ꞉ असे लेखन. म्हणजे या जोडाक्षरात 'र' चा उच्चार पूर्ण आहे.

यांत

या व्यंजनानंतर 'र' जोडून येणारी जोडाक्षरे मोडी लिपीत नाहीत.

 असे जोडाक्षर मोडी लिपीत आढळत नाही. मात्र

 असे इंग्रजी शब्द मोडी लिपीत आले आहेत.

६आ) 'रेफ' / 'रफार' 'र' या व्यंजनानंतर येणाऱ्या अक्षरावर ꞈ असे चिन्ह दिले जाते.

उदा.

यांत 'र' चा उच्चार 'र' च्या मागील अक्षराबरोबर होतो व त्या अक्षरावर आघात (जोर) येतो. अशा वेळेला जोडाक्षरात येणारे र हे व्यंजन पुढील अक्षरावर ꞈ रेफ हे चिन्ह देउन दर्शविले जाते. याला रफार असेही म्हणतात.

मोडी शब्दातील रफार (रेफ) कोणत्या अक्षरावर येईल यासंबंधी मोडी लिपीत गोंधळ आढळून येतो. उदा. : ऐवजी

 ऐवजी वजी

अशी चुकीची रुपे लिहिलेली आढळतात.

Figure 19: Description of the forms of ꙋ RA (from Mhātre 2009: 26). Continued in figure 20.

६इ) 'र' हे व्यंजन जोडाक्षरात आधी आले तर '॒' या चिन्हाने दाखविले जाते. हे चिन्ह 'र' या व्यंजनानंतर येणाऱ्या अक्षरात जोडतात.

उदा. झ्न्या , स्र्ळड, छीकड, न्ऱस्प, खऱ्या

सुन्या या शब्दात 'र' नंतर य हे अक्षर आले आहे. ते उच्चारताना सु + ॒ +या असे उच्चारले जाते. मागील अक्षरावर आघात (ज़ोर) मुळीच दिला जात नाही. उलट सूर्य या शब्दाचा उच्चार करताना सू वर आघात येतो व 'र' चा उच्चार सू बरोबर (सूर् + य) होतो.

एकंदरीत 'र' या व्यंजनाचा उच्चार त्याच्यापुढील अक्षरासह होतो व 'र' च्या मागील अक्षरावर आघात किंवा जोर दिला जात नाही. तेव्हा 'र' व्यंजनाऐवजी 'र' असेल तर ⌐ हे चिन्ह केव्हा वापरायचे व '॒' हे चिन्ह केव्हा वापरायचे ते लक्षात येईल.

मोडी लिपीतील सर्व 'उ' कार न्हस्व असतात, 'उ' काराबाबत मोडी लिपीत न्हस्व दीर्घ भेदाभेद नाही. त्यामुळे खउर्ट, खऱ्या

यातील 'खउ' सारखाच काढतात.

६ई) 'ट्' आणि 'ड्' नंतर 'र' आल्यास हे चिन्ह ट किंवा ड च्या खाली येते.

उदा. : ट्राम, ड्रम, षड्रस, राष्ट्र : ट्राम ड्रम षड्रस नाष्ट्र

ट आणि ड या व्यंजनानंतर 'र' जोडून येणारी जोडाक्षरे मोडी लिपीत नाहीत.

'र' बाबत जोडाक्षरात वारंवार होणाऱ्या चुका

स् + र् – स्र : स्रोण, स्राळ, छछकाऽधी, षैऱ्ऱल

स् + त् + र् – स्त्र : स्त्री, पस्ल, ट्टिरस्राण, ट्टास्त्री

७) ज्या शब्दाचे पहिले अक्षर 'र, री, रे, रै, र' ने सुरू होते अशा वेळी 'उ॒' या मुळाक्षरानेच शब्द सुरू होतो.

उदा. ट्टमगीरी, रीतीने, रेपछंडा, रैम, रंझर.

रा : 'ना' किंवा रु : 'स' असे काढतात.

उदा. न्ऱरराम, नास्ऱंध्स, नास्नाऽ, स्रष्णे, स्रअष्ठ

Figure 20: Description of the forms of उ ʀᴀ (from Mhātre 2009: 27). Continued in figure 21.

८) मोडी लिपीमध्ये *ररारी* (र रा री) अशी बालबोधसदृश अक्षरे काढलेली आढळतात.

उदा. [handwritten examples]

९) मोडी ही मराठीची जलद लिहिण्यासाठी वापरली जाणारी लपेटीदार लिपी त्या लपेटीदार वळणात 'र' पूर्णपणे आडवा काढलेला आढळतो.

उदा. [handwritten] (नंबर), [handwritten] (हुजूर), [handwritten] (लवकर)

[handwritten] (हजारी) [handwritten] (सावंनूर)

[handwritten] (बरोबर) [handwritten] (हेरलेकर)

'र' हे अक्षर मोडी लिपीत मुळाक्षरांच्या पायाशी, कमरेत, कुशीत, पोटात आणि डोक्यावर कोठेही कसेही घुसून आपले अस्तित्व दाखवून देते.

'र' हे मोडी लिपीत जास्तीत जास्त व्यंजनांना निरनिराळ्या प्रकारे जोडण्यात येणारे एकमेव अक्षर आहे.

१०) 'र' शी संबंधित अक्षरे व जोडाक्षरे :

Figure 21: Description of the forms of उ RA (from Mhātre 2009: 28).

40

Figure 22: Document showing the invocation sign ⏌ ʜᴜᴠᴀ (from Strandberg 1983: 192).

Figure 23: Modi signs used as invocations in correspondence. Examples of some signs are given in figure 24.

Variant form of ꔥ huva

The invocation ī

The invocation īṃ

The invocation īṃ

The invocation īṃ

The invocation īṃ

Figure 24: Some invocations used in Modi documents.

Figure 25: Document showing Modi written upon ruled lines and in free form (from Dīkṣit and Khobarekar 1969: 209).
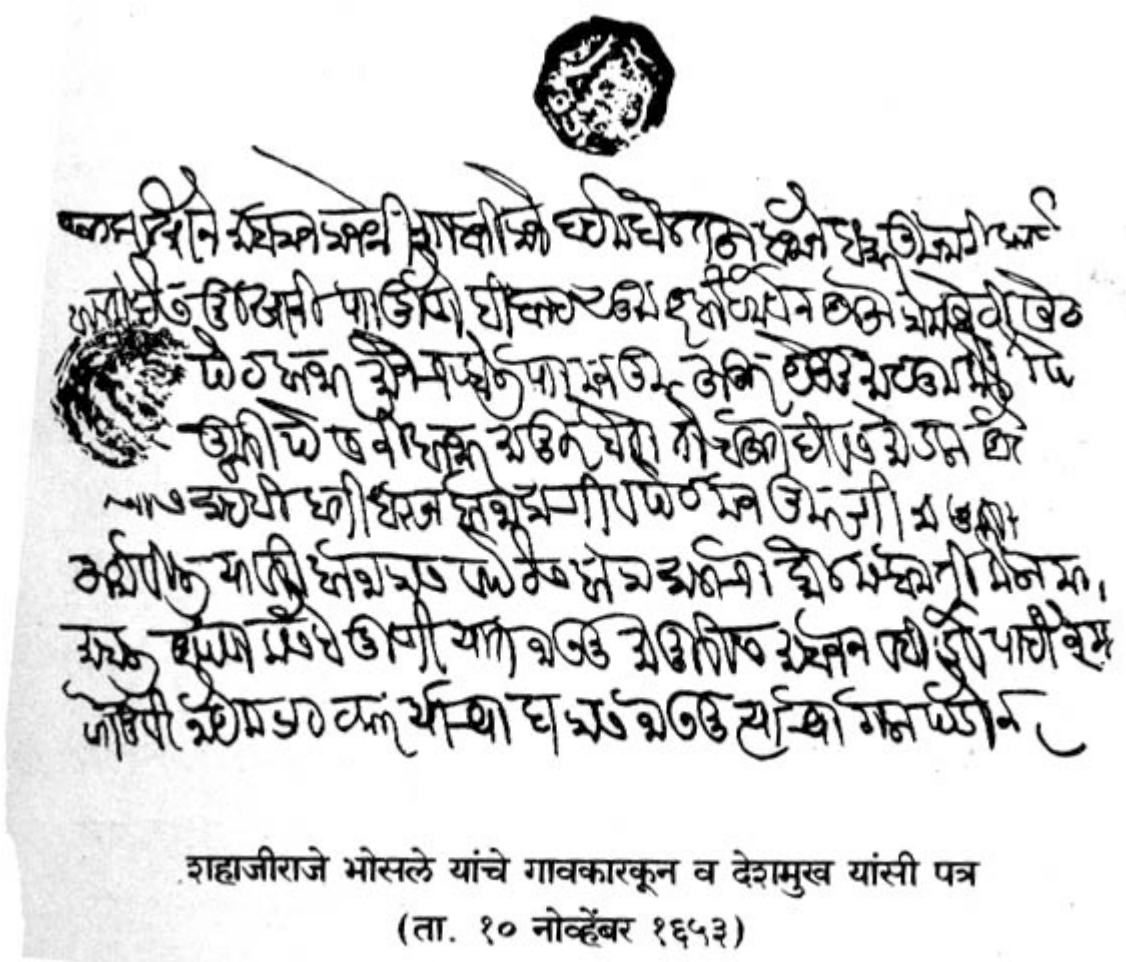
Figure 26: A letter written by Shahaji Raje Bhonsle in the Bahamani style of Modi.
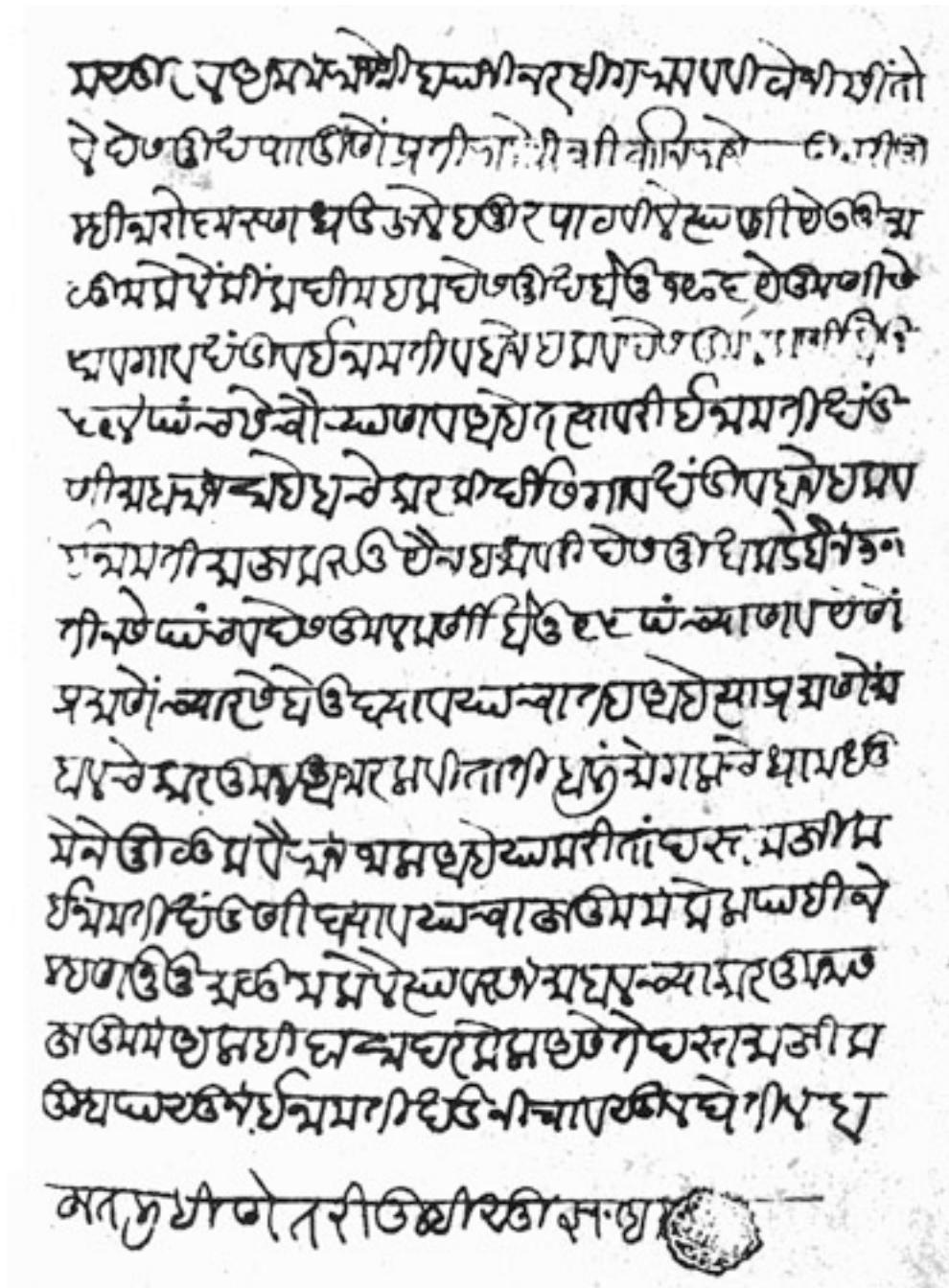
Figure 27: A letter written in the Chitnisi style of Modi during the reign of Shivaji.

Figure 28: A letter written by Madhavrao in the Peshve style of Modi.
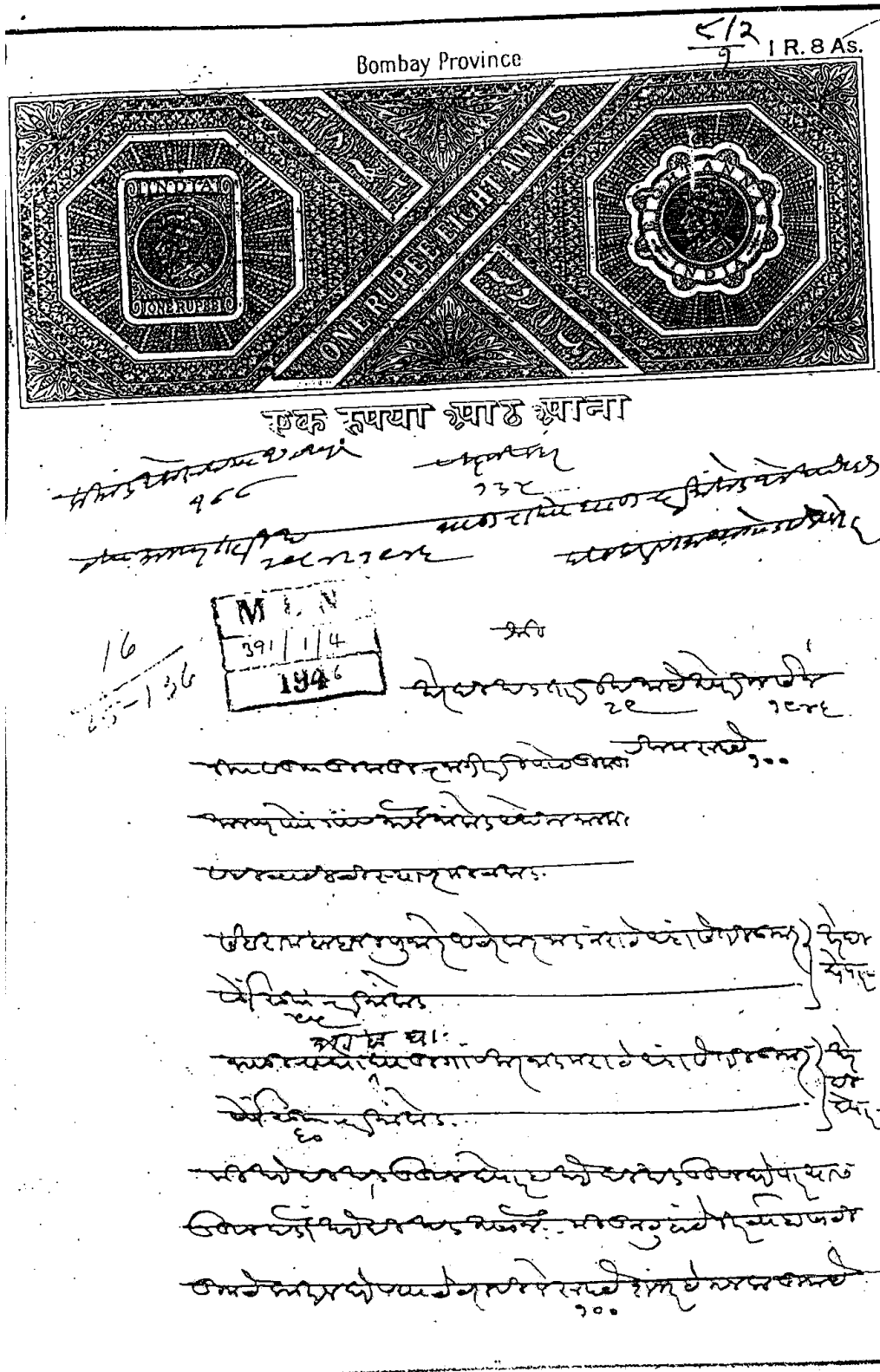
Figure 29: A letter written in the Anglakal style of Modi.
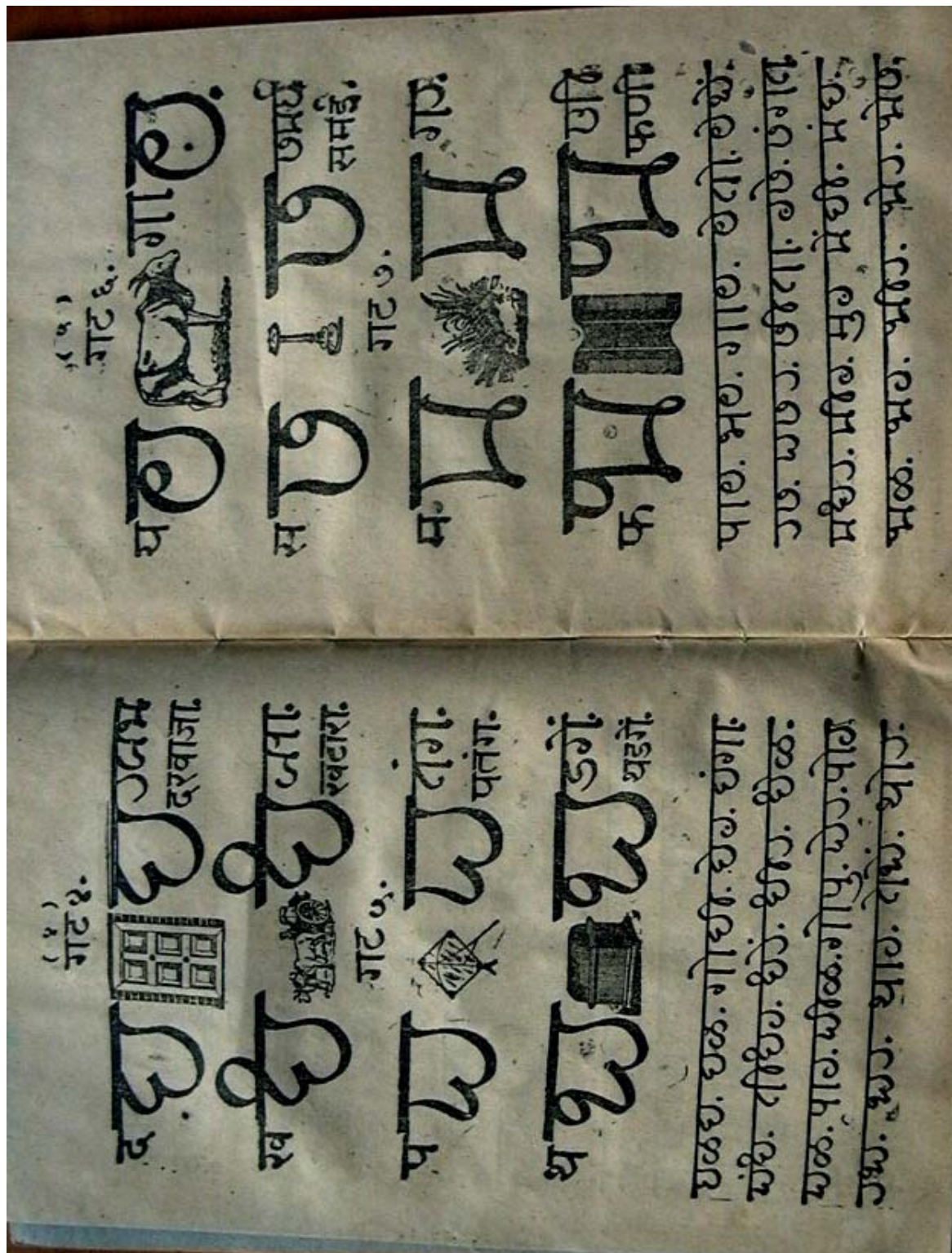
Figure 30: Excerpt from a 20th century Modi school book (from Pathavardhana: 4–5).

Figure 31: Excerpt from *Yacavicara* (from American Tract Society 1837: 18–19).

श्री

Figure 32: A letter from the office of Bājīrāva Ballāḷa Peśave, dated December 1721 (from Dīkṣit and Khobarekar 1969: 71–73). The Modi is written in the clear Peshve style.
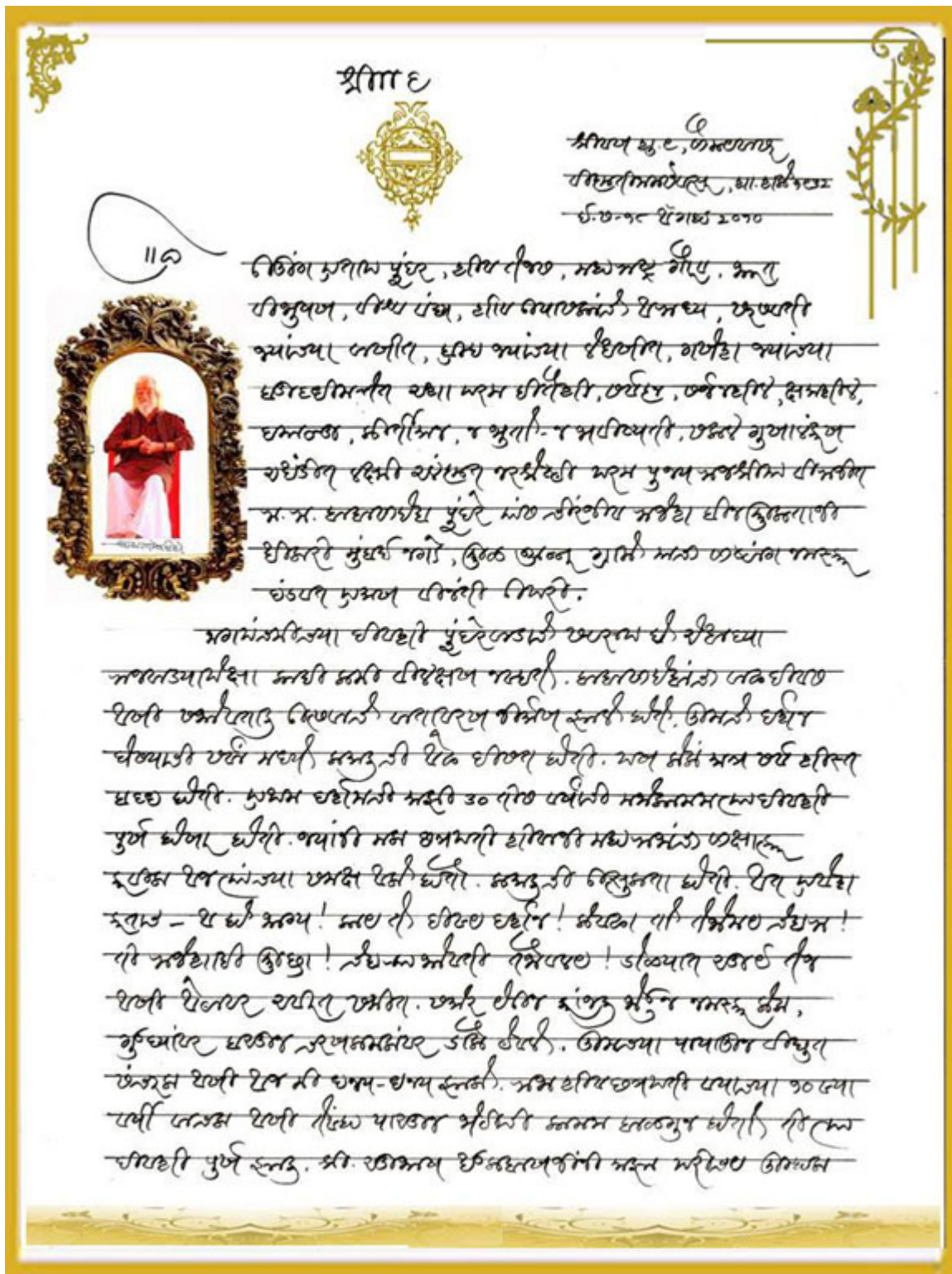
Figure 33: A contemporary letter written in modern Modi. Courtesy of Rajesh Khilari.

MAHRATTA DICTIONARY.

अ.

अऋणी, *a.* out of debt, not in debt.

अकच्छ, *a.* not having a piece of the cloth which is worn round the loins tucked in behind.

अकथरूप, *a.* of an unascertained terrific form.

अकथ्य, *a.* unspeakable, unfit to be spoken, obscene.

अकथनीय, *a.* unspeakable, unutterable, unfit to be spoken, obscene.

अकपट, *a.* ingenuous, frank, honest, sincere.

अकरिय, *a.* improper to be done, improper.

अकर्णवेध, *a.* without bored ears, (an epithet of reproach to a Musulman.)

अकर्मी, *a.* useless.

अकर्मी, *a.* useless, unfit for any purpose, indolent.

अकामांपुरेता, *a.* useless, unfit for any purpose, indolent.

अकर्मण, *a.* useless, not fit for any purpose.

अकरणीय, *a.* unfit to be done.

[२]

अका.

अकरूण, *a.* unfeeling, unmerciful.

अकस्मीरा, *a.* without tergiversation, sincere, undissembled, unadulterated.

अकस्मात, *s.* bad fortune, an unfortunate thing.

अकस्मात्, *ad.* suddenly, unexpectedly.

अकसूर, *s.* malice, revenge in the mind.

अकातर, *a.* composed, undisturbed, unanxious.

अकार्य, *s.* useless, improper.

अकारणा, *a.* causeless, needless, useless.

अकाल, *a.* out of season, unseasonable. *s.* a famine.

अकालकन, *a.* born out of due time, abortive.

अकालजरा, *a.* born out of the proper time.

अकाली, *ad.* prematurely.

अकिंचन, *a.* destitute, not possessing any thing.

अकुदग्रारा, *a.* unshared, undivided.

अकृत्रीम, *a.* unartificial, not counterfeited, genuine, pure.

अक्का, *s.* an elder sister.

अकरा, *a.* eleven.

अकरावद्, *a.* the eleventh.

अज्ञात्पीनाळ, *a.* unascertained, uncertain.

Figure 34: Excerpts from Carey's *A dictionary of the Mahratta language* (from Carey 1810: 1–2).

Figure 35: A Modi document from the *Linguistic Survey of India* showing Modi written with word spacing (from from Grierson 1905: 259–260).
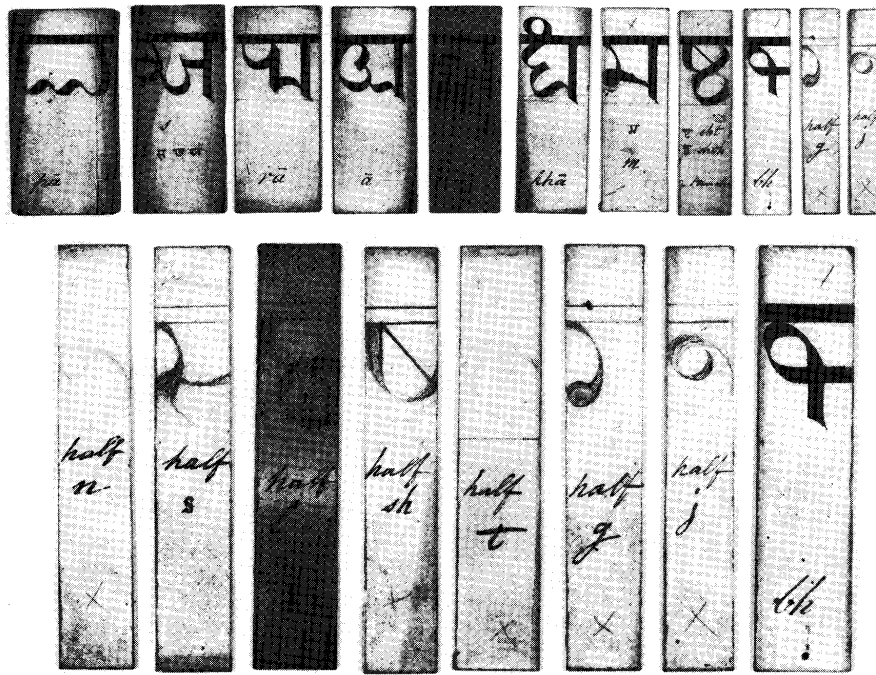
Figure 36: Proofs of Charles Wilkin's Modi font (from Ross 1999: 29).

Figure 37: Specimen of Charles Wilkin's Modi font (from Ross 1999: 29).

55

Figure 38: First page of the *Khabardar*, the first newspaper printed in Modi. Courtesy of Bharat Itihas Sanshodhak Mandal (BISM) and Suresh Pethe.

Figure 39: Cover of *Modi Vaibhav*, a monthly magazine published in Modi. Courtesy of Rajesh Khilari.

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.
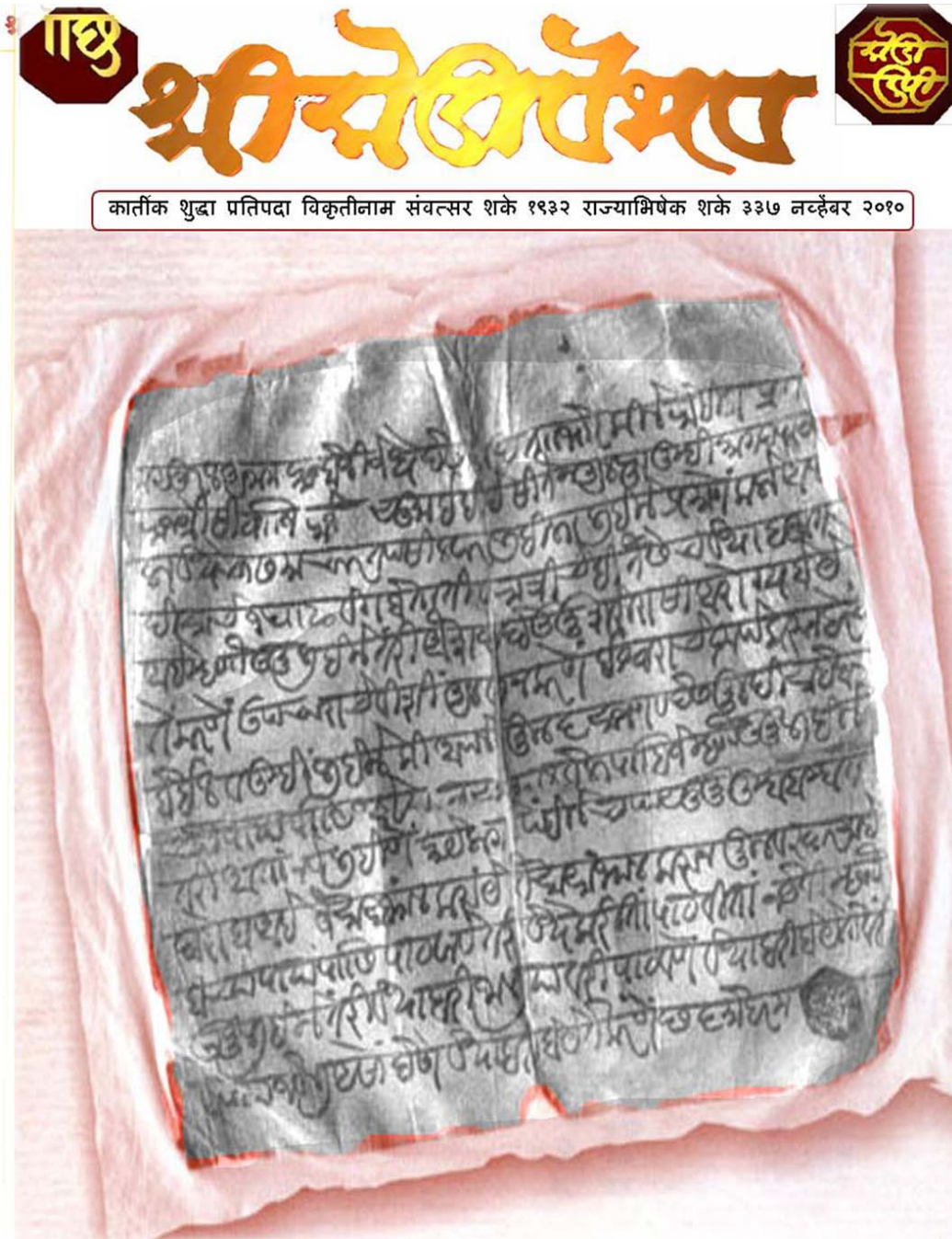**Please read Principles and Procedures Document (P & P) from** http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html **for guidelines and details before filling this form.**
Please ensure you are using the latest Form from http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html**.**
See also http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

A. Administrative

1. **Title:** *Proposal to Encode the Modi Script in ISO/IEC 10646*
2. Requester's name: *Script Encoding Initiative (SEI) / Anshuman Pandey <pandey@umich.edu>*
3. Requester type (Member body/Liaison/Individual contribution): *Liaison contribution*
4. Submission date: *2011-11-05*
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: *Modi*
   b. The proposal is for addition of character(s) to an existing block:
      Name of the existing block:
2. Number of characters in proposal: *79*
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 <pandey@umich.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*
   See proposal for additional details

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 http://www.unicode.org/Public/UNIDATA/UCD.html 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

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? — *Rajesh Khilari <rajeshkhilari @gmail.com> Madhav Deshpande <mmdesh @umich.edu>*
   If YES, available relevant documents: — *See citations in text of proposal*

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? — *Yes*
   Reference: — *Marathi-speaking communities in India and elsewhere; institutions exist for the promotion and use of Modi; historians, linguists, and other specialists in India and Europe. See text of proposal for additional details.*

4. The context of use for the proposed characters (type of use; common or rare) — *Common*
   Reference: — *Administrative documents from the Maratha period and from European countries with former presence in Marathi-speaking regions of South Asia. Also, personal correspondence and private documents of Marathi speakers See text of proposal for additional details.*

5. Are the proposed characters in current use by the user community? — *Yes*
   If YES, where? Reference: — *There is an active scholarly community, primarily in Maharashtra, India and in parts of Europe, engaged in the study of Modi and documents written in the script. See text of proposal for further details.*

6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? — *No*
   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 an existing character? — *Yes*
    If YES, is a rationale for its inclusion provided? — *Yes*
    If YES, reference: — *See text of proposal.*

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 vowel signs; see text of proposal*
    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 proposal*

13. Does the proposal contain any Ideographic compatibility character(s)? — *No*
    If YES, is the equivalent corresponding unified ideographic character(s) identified?
    If YES, reference: