TOWARDS MODI SCRIPT PRESERVATION: TOOLS FOR DIGITIZATION

Kishor Patil, Neha Gupta, Damodar M and Ajai Kumar

Centre for Development of Advanced Computing (C-DAC), Pune, India

ABSTRACT

Modi (मोडी, modī) is a heritage script belonging to Brahmi family, which is used mainly for writing Marathi, an Indo-Aryan language spoken in western and central India, mostly in the state of Maharashtra. "Modi-manuscript "written from the past, reveals the history of the Maratha Empire from its inception under Chhatrapati Shivaji Maharaj; to the creation of movable metal type when Modi was slowly relegated to an inferior position, unfolds perspectives and reflects the social, political and cultural sense of his time." Today it is very important for historians, researchers and students to understand this script and use it for historical heritage. Other regional languages such as Hindi, Gujarati, Kannada, Konkani and Telugu were also using Modi. This paper presents our contribution in helping the community for preserving the script, by way of using various tools, which will facilitate the collection, analysis, and digitization of the Modi script.

KEYWORDS

Language preservation, language development, Modi, Language Analysis, Heritage script.

1. INTRODUCTION

The origin of Modi script is a debatable issue. According to certain historians, Modi can be dated back to the Maurya Dynasty (322–185 BC) and hence the name: 'Modi'. However, the most credible account of the origin of Modi is that Hemadripant, is credited with the invention of the Modi script and the date assigned to its 'birth' is the year 1260.

The invention of Modi can be termed as an act of genius. Modi was invented as a cursive 'shorthand' or speed writing to note down the royal edicts. Traditional Devanagari was found to be excessively time-consuming since each character required as many as 3 to 5 strokes and the lifting of the hand, each time the stroke was completed. Modi got round this obstacle by 'bending' the letters thereby doing away with the need of lifting the hand. This invention thus allowed for a continuous writing which could be used by court scribes to note the edicts. As an example: the handwritten Marathi letter 'has seven 'hand-lifts' whereas' in Modi because of continuous flow requires not a single hand lift. Termed as 'Lapetdar' or Cursive, Modi became extremely popular and rivaled the 'Shikasta' script of Persian. The introduction and history about the Modi script has been covered by the [1][2] and [3] references.

2. LITERATURE SURVEY

Modi is a heritage script and not much of the content is available on the internet. However, the proposal by Anshuman Pandey, "Proposal to Encode the Modi Script in ISO/IEC 10646" covers

the details about the Modi Script [5]. The proposal covered the background and writing system of the Modi script which is a great starting point to understand the script. The proposal also covers characters and character combinations in detail, which is very helpful to understand the orthography of the script. The article covered all the aspects of the Modi script including conjuncts formations, head strokes, word and section boundaries and collating order. Comparison of Modi script and Devanagari has also been covered in the article. With the Author's efforts, Modi is now a part of the Unicode Standard [6]

Figure 1. Modi Unicode code chart

3. THE EVOLUTION OF MODI

Modi script remained in practice from the 13th century to the first half of 20th century. During the seven odd centuries, Modi calligraphy underwent several transitions. The official correspondence of Maharashtra as well as the regions like Madras, Mysore, Bundelkhand, Gujarat, and Rajasthan was written in Modi for a long time. Transitions in Modi calligraphy fall in four periods Bahamani, Shivkalin, Shahukalin and Anglakalin. In each of these periods the 'style' and 'lapeti' of Modi calligraphy underwent a considerable change.

3.1. Bahamani kalin Modi

Modi script in the Bahamani period was impacted by Perso-Arabic script. The letter of Shahaji Maharaj (18 March 1594 - 23 January 1664) shows script structure of Bahamani era (Figure 2).
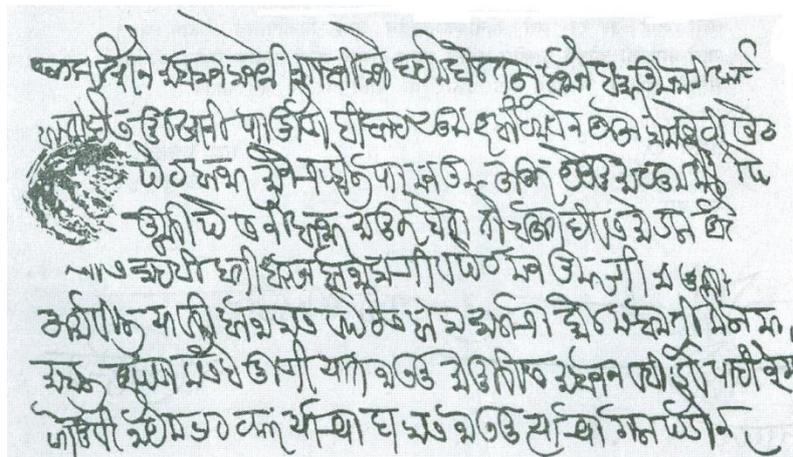

Figure 2. Letter of Bahamani kalin Modi

3.2. Shivkalin Modi

Modi script in the period of Chhatrapati Shivaji Maharaj (19 February 1630 - 3 April 1680) period was also impacted by Perso- Arabic script. The letter of Chhatrapati Sambhaji Maharaj shows the beauty of Shivkalin era (Figure 3, Figure 4).
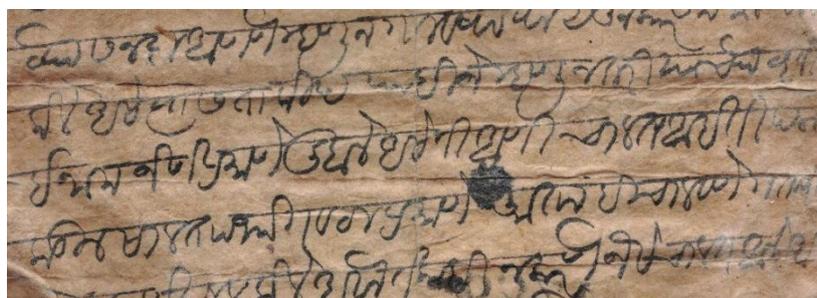

Figure 3. Letter of Shivkalin Modi
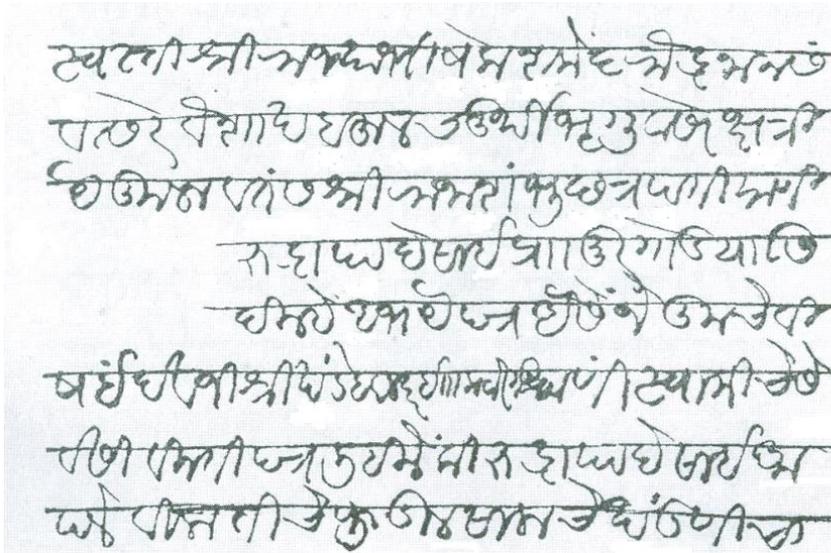
Figure 4. Letter of Shivkalin Modi

3.3. Shahukalin Modi

Modi script in the period of Chhatrapati Shahu Maharaj (18 May 1682 - 15 December 1749) and Peshwa's is more cursive. The curves are more marked and elongated. The curves in fact gave rise to various styles such as Mahadajipanti, Biwalkari and Ranadi (Figure 5).
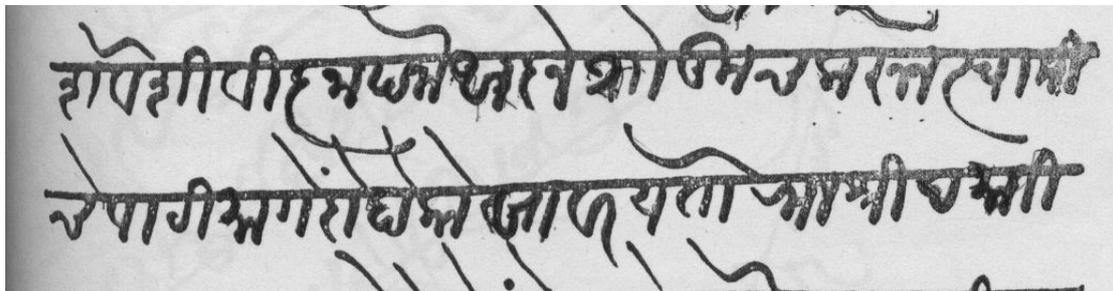

Figure 5. Letter of Shahukalim Modi

3.4. Anglakalin Modi

Under the first half of the British period (1818-1947) English and Modi co-existed together. Most of the correspondence in the Deccan was in Modi and advertisements in English newspapers were also in Modi. However with the innovation of Moving Metal-type and the spread of English language the downfall of Modi script became certain. The Modi of the British period is influenced by the use of the pen to write letters. Thick-thin variants show their presence for the first time as in the following specimen of a letter of General Grant Duff (Figure 6).
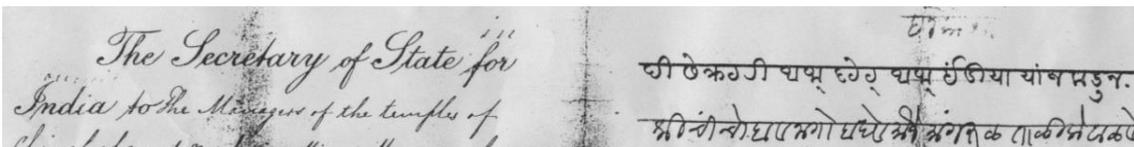
Figure 6. Letter of Anglakalin Modi

4. MODI WRITING SYSTEM DETAILS

Although Modi is inherited the model as Devanagari, it differs considerably from the Devanagari in terms of letter forming, rendering behaviours, and orthography. Modi was invented as a cursive "shorthand" or speed writing to note down the royal edicts. The traditional Devanagari turned out to be too laborious since each character required up to 3-5 strokes and the raising of the hand, each time the stroke was completed. Modi got found this obstacle by "bending" the letters thereby doing away with the need of lifting the hand. This invention thus allowed for a continuous writing which could be used by court scribes to note the edicts (Figure 7, Figure 8, Figure 9, Figure 10, Figure 11, Figure 12). However, this 'speed-writing' led to certain modifications of which the most notable features are as under:

- The total absence of short and long vowel forms as well as vowel modifier forms
- Use of certain specific markers used in Modi as prefixes for numerical notation
- Practically no derived ligature forms: conjuncts being marked either by use of the virama or by use of half characters

More details about the Modi alphabets is covered in the link [4]

ꦈꦉꦊ ꦋꦌꦍꦎꦏꦐꦑꦒꦓꦔꦕ

Figure 7. Modi Vowels
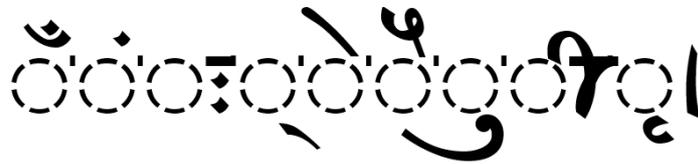
Figure 8. Modi Vowel Signs

꧑꧒꧓꧔꧕꧖꧗꧘꧙꧐

Figure 9. Modi Numerals
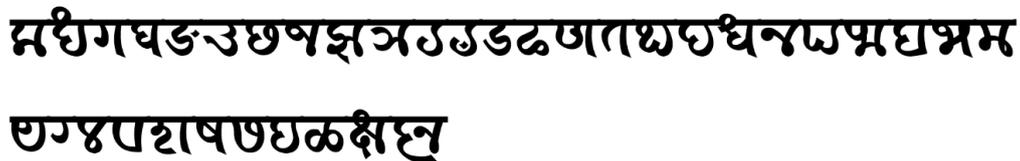
Figure 10. Modi Consonants

The shapes of some consonants, vowels, and vowel signs are similar. The actual differences are visible when characters are part of consonant-vowel combinations / consonant conjuncts.

ळ्धिगाघाडाउाछाऋझसालबडाढाणाताध्रघाधाभपाभ्रघझम
म्नाभ्रहाषाज्यघळकिज़ख

Figure 11. Modi Consonants-aa matra combinations

Figure 12. Modi Conjuncts

5. COMPARISON BETWEEN MODI AND OTHER SCRIPTS

This section compares Modi and Other scripts. Modi and various other scripts, as shown below, have some similarities, as they are either variant or influences from Devanagari. The below table (figure 13, figure 14) shows the consonants comparison between Modi, Mahajani [7], Landa [8], Kaithi [9], and Devanagari [10].

	Modi	Mahajani	Landa	Kaithi	Devanagari
KA					क
KHA					ख
GA					ग
GHA					घ
NGA		—			ङ
CA					च
CHA					छ
JA					ज
JHA					झ
NYA					ञ

Figure 13. Consonants comparison (ka-nya)

	Modi	Mahajani	Landa	Kaithi	Devanagari
TTA	उ	४	८	८	ट
TTHA	उ	६	८	ऽ	ठ
DDA	ड	ऽ	३	७	ड
DDHA	ढ	७	८	७	ढ
NNA	ज	२॥	ट	॥	ण
TA	त	२	�	∩	त
THA	ध्य	ॼ	व	थ	थ
DA	ठ	६	४	६	द
DHA	ध	२	व	य	ध
NA	न	ॳ	⅄	⪑	न
PA	ध	५	५	प	प
PHA	फ्र	३	८	ॐ	फ
BA	घ	५	ॶ	व	ब
BHA	झ	९	∄	ᶍ	भ
MA	म	ᴎ	४	न	म
YA	ठ	—	२	४	य
RA	ᒎ	०	)	⅂	र
LA	४	ॼ	௸	⊖	ल
VA	ढ	५	८	व	व
SHA	श	—	—	ॴ	श
SA	ठ	६	३	स	स
HA	ठ	॰	५	६	ह

Figure 14. Consonants comparison (tta-ha)

6. COLLECTION OF TOOLS

This section explains existing work that are related to the Modi Puratan Dastavej Jatan Pranali - Digital Annotation & Archiving System

6.1. MODI-SHAHU (मोडी–शाहू) Font

MODI-SHAHU font is specially designed font as per the letter forms, rendering behaviour and orthography. This font fulfils all the basic requirements such as Structure and form of the character for better digitization (Figure 15).

Figure 15. Modi-Shahu Font

6.2. MODI Typing Tool

Modi Typing tool is developed for inputting the content /text in Modi script. The keyboard is designed using the INSCRIPT principle. The keyboard is based on Unicode; hence, the documents thus created will be easily be viewed properly on any Unicode enabled operating systems such as Windows 10. On-screen keyboard for Modi Script is also provided in this tool to make typing more easily. The tool uses the font "MODI-SHAHU" for rendering the character on the Keyboard (Figure 17).

Figure 16. Modi Typing Tool UI

Figure 17. Modi Keyboard

6.3. MODI - DEVANAGARI Converter

The Modi-Devanagari converter converts data from Modi script to Devanagari script. The converter is designed based on research and study of Modi and Devanagari scripts written in Marathi language (Figure 18).
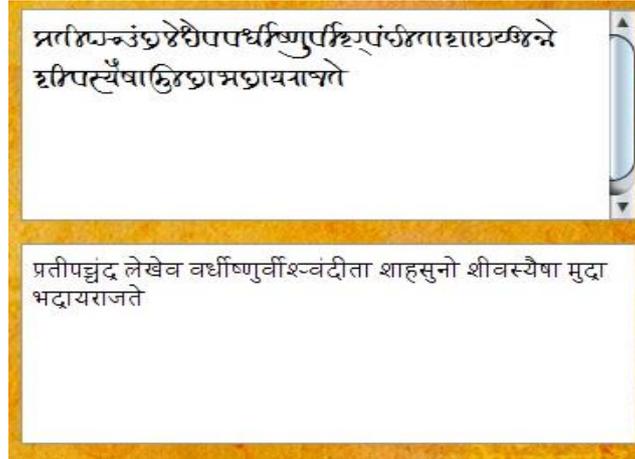
Figure 18. Modi-Devanagari Converter

6.4. Modi Puratan Dastavej Jatan Pranali - Digital Annotation & Archiving System

The "Modi Puratan Dastavej Jatan Pranali (मोडी पुरातन दस्तावेज जतन प्रणाली) Annotation and Digital Archiving System" is a web application system for the digital preservation of cultural heritage resources and manuscripts for Modi. The system takes historical Modi documents available as images and allows the user to annotate and type in the Modi text. A virtual keyboard is provided for easy typing of Modi script. A transliteration system is also provided to convert Modi text to Devanagari. System uses the specially designed font MODI-SHAHU for proper display of content in the Modi Script. (Figure 19, Figure 20,Figure 21,Figure 22).

Figure 19. Modi Puratan Dastavej Jatan Pranali - Start Screen
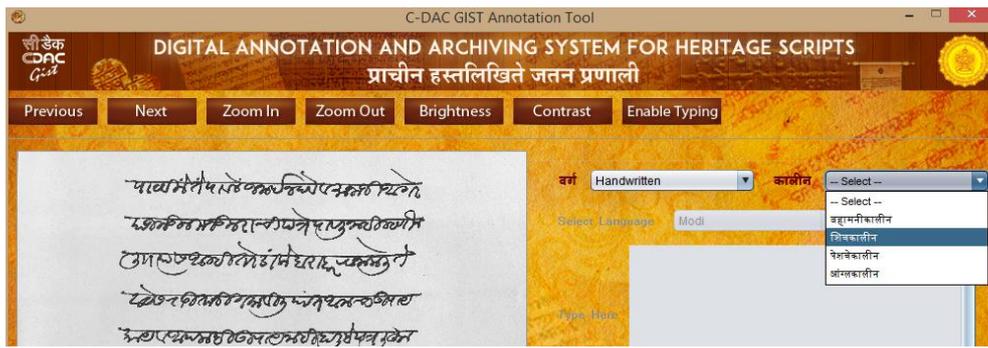
Figure 20. Modi Puratan Dastavej Jatan Pranali - Selection of Era
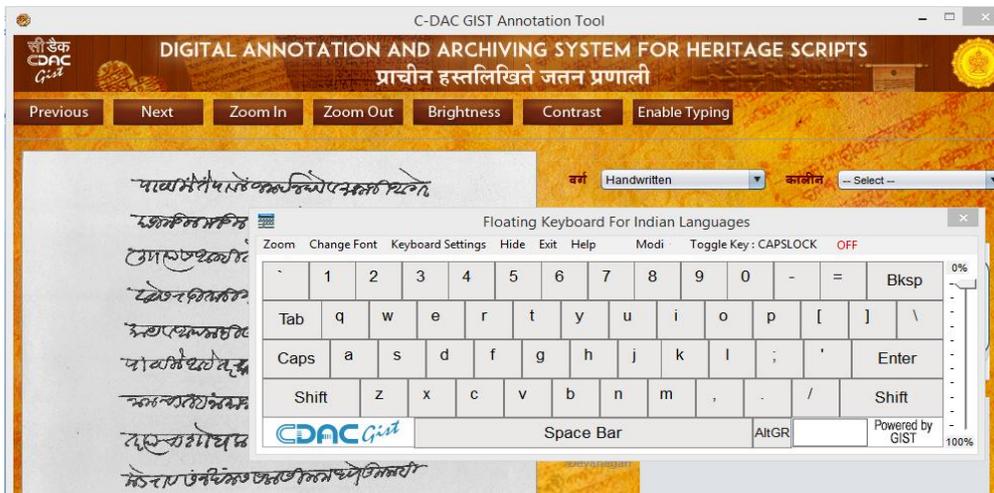
Figure 21. Modi Puratan Dastavej Jatan Pranali - Keyboard
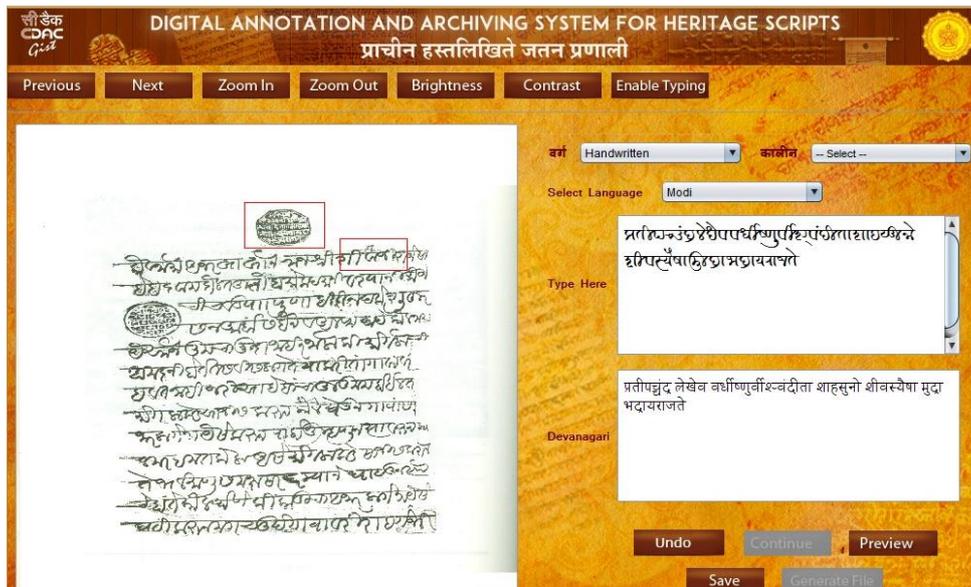
Figure 22. Modi Puratan Dastavej Jatan Pranali - Annotation and converter screen

7. CONCLUSION

The development of "Modi Puratan Dastavej Jatan Pranali" is a functional tool for preserving the script from extinction. It also facilitates various tools and technologies for the community user to better understand the script and digitize it. "Modi Puratan Dastavej Jatan Pranali" is a collection of various tools by which data available in the form of images will be easily inputted, displayed, converted, and stored in a database. However if further applications such as OCR, the search engine could be developed and integrated with it, it can increase the usability of the existing systems, which is a future proposition.

ACKNOWLEDGEMENTS

REFERENCES

[1]   "Modi script," https://en.wikipedia.org/wiki/modi_script.

[2]   "History of modi lipi," http://www.modilipi.in/2011/02/modi-script-of-maharashtra-script-which.html.

[3]   "The origin and development of indian writing system," https://narendranath.webs.com/.

[4]   "Modi alphabet," https://omniglot.com/writing/modi.htm.

[5]   "Proposal to Encode the Modi Script in ISO/IEC 10646", http://unicode.org/L2/L2011/11212r-n4034-modi.pdf

[6]   "Modi Unicode Code Chart", https://www.unicode.org/charts/PDF/U11600.pdf

[7]   "Mahajani Script", https://omniglot.com/writing/mahajani.htm

[8]   "Landa Script", http://std.dkuug.dk/JTC1/SC2/WG2/docs/n3766.pdf

[9]   "Kaithi Script", https://omniglot.com/writing/kaithi.htm

[10] "Devanagari Script", https://omniglot.com/writing/devanagari.htm

AUTHORS

Mr. Kishore Patil has worked as a typographer, designer and visualizer for over 20 years. He has worked in various fields of fine arts, such as advertising, newspapers, agricultural research, medical research and the IT sector. His interest in designing and developing Open Type fonts for all Indian languages and heritage scripts.

Ms. Neha Gupta has more than 14 years of experience in projects related to research and development of local language technology and standardization initiatives. Her specialization includes developing language processing tools/systems viz. Internationalized Domain Names implementation at India Level, Neural Machine Translation (NMT), NER, POS to name a few specifically with AI/ML/DLand rule based methodologies.

Dr. Damodar Magdum has been working in Indian language technology for over 15 years. His specialization in the creation of text to speech (TTS) corpus in Indian language. He has special interest in the proliferation and preservation of heritage language in India.

Dr. Ajai Kumar is Senior Director and Head of the Applied Artificial Intelligence & GIST Group and have more than 20 years of experience working in Natural Language Processing including Machine Translation, Speech Technology and Information Extraction & Retrieval and E-learning systems. His key role is in initiating mission mode consortium projects in the area of Natural Language Processing, language learning, Video Surveillance etc.