Proposal to Encode the Kaithi Script in Plane 1 of ISO/IEC 10646

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

May 21, 2007

Contents

Proposal Summary Form i

1 Introduction 1

2 Characters Proposed 2

3 Technical Features 8

4 Background 12

5 Orthography 20

6 Regional Variants and Typeface Styles 32

7 Relationship to Other Scripts 33

8 References 42
List of Figures

1. Geo-political extent of the Kaithi script in South Asia.
2. Relationship of Kaithi to selected Nagari-based scripts.
3. A comparison of the three regional forms of Kaithi.
4. A list of Kaithi conjuncts used in the Maithili style of Kaithi.
5. Currency, weights, and measures signs used in Kaithi.
7. Specimen of hand-written Maithili style of Kaithi.
8. Specimen of hand-written Magahi style of Kaithi.
9. Excerpt from a specimen of Maithili written in the Magahi style of Kaithi.
10. Excerpt from a specimen of Awadhi written in Kaithi.
11. A specimen of Magahi printed in Kaithi type.
14. Table of the Kaithi script.
15. Inventory of Kaithi letters.
16. Comparison of numerals of Kaithi and other scripts.
17. Folios 1b and 2a from the Mahāganapatiṣṭotra in Devanagari and Kaithi.
18. Folios 1a and 4a from the Mahāganapatiṣṭotra in Devanagari and Kaithi.
19. Excerpt from a plaint from the district court of Patna, Bihar.
20. Excerpt from a plaint from the district court of Bhagalpur, Bihar.
21. Excerpt from a statement from the district court of Ranchi, Bihar.
22. Rent receipt from the former Principality of Seraikella.
23. Title and first pages of the Book of Genesis in Kaithi type.
26. A folio from the “Ekaḍālā” manuscript of Miragāvatī.
27. A folio from the Tale of Sudama.
28. A letter to the Supreme Civil Court of Appeals in Calcutta.
29. Comparison of Kaithi, Gujarati, and Devanagari types from the Linguistic Survey of India.
30. Comparison of hand-written Kaithi and Gujarati letters.
31. Comparison of Kaithi and Devanagari.
32. A comparison of the Kaithi script with the Devanagari and Mahajani.
33. Comparison of Kaithi drawn with the headstroke and Devanagari.
34. Comparison of writing techniques in Kaithi and Devanagari.
35. Comparison of scripts descended from proto-Bengali.
36. Comparison of Kaithi with other scripts used for writing Hindi.
37. Comparison of Kaithi with other Indic scripts.
38. A family tree of north Indian scripts showing Kaithi as a branch of Nagari.
39. The position of the Kaithi script with regard to others.

List of Tables

1. Glyph chart for Kaithi.
2. Character Names and Properties.
3. Comparison of metal and digitized Kaithi fonts.
A. Administrative

1. Title: Proposal to Encode the Kaithi Script in Plane 1 of ISO/IEC 10646
2. Requester’s name: University of California, Berkeley Script Encoding Initiative (Universal Scripts Project);
   author: Anshuman Pandey (pandey@umich.edu)
3. Requester type (Member Body/Liaison/Individual contribution): Liaison contribution
4. Submission date: May 21, 2007
5. Requester’s reference (if applicable): N/A
6. Choose one of the following:
   (a) This is a complete proposal: Yes
   (b) or, More information will be provided later: No

B. Technical - General

1. Choose one of the following:
   (a) This proposal is for a new script (set of characters): Yes
      i. Proposed name of script: Kaithi
   (b) The proposal is for addition of character(s) to an existing block: No
      i. Name of the existing block: N/A
2. Number of characters in proposal: 73
3. Proposed category: C - Major extinct
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. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?: Anshuman Pandey; True Type format
   (a) If available now, identify source(s) for the font and indicate the tools used: The font contains normalized forms of letters found in hand-written and printed Kaithi documents. It was drawn with Metafont and converted to True Type with FontForge.
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:
   (a) 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. **Character properties and numeric information are included.**

---

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?: **No**
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.)? **No**
   (a) If Yes, with whom?: **N/A**
      i. If Yes, available relevant documents: **N/A**
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? **Yes**
   (a) Reference: *Awadhi, Bhojpuri, Magahi, and Maithili speakers; as well as linguists, historians, legal scholars working with sources from colonial South Asia.*
4. The context of use for the proposed characters (type of use; common or rare): **Common**
   (a) Reference: *Court records from colonial India, pedagogical materials from north India, commercial and accounting records; religious and literary texts; bibles printed in north India during the 19th and early 20th century. Other contexts discussed at length in the text of the proposal.*
5. Are the proposed characters in current use by the user community?: **Yes, by scholars working in fields enumerated above. It is difficult to verify whether the script is presently in active use in India.**
   (a) If Yes, where? Reference: *In India, the United States, and other localities.*
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?: **No**
   (a) If Yes, is a rationale provided?: **N/A**
      i. If Yes, reference: **N/A**
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**
   (a) If Yes, is a rationale for its inclusion provided?: **N/A**
      i. If Yes, reference: **N/A**
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? **No**
   (a) If Yes, is a rationale provided?: **N/A**
      i. If Yes, reference: **N/A**
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? **Yes**
    (a) If Yes, is a rationale for its inclusion provided? **Yes**
       i. If Yes, reference: *See text of proposal*
11. Does the proposal include use of combining characters and/or use of composite sequences? **Yes**
    (a) If Yes, is a rationale for such use provided? **Yes**
       i. If Yes, reference: *See text of proposal*
    (b) Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? **Yes**
       i. If Yes, reference: *See text of proposal*
12. Does the proposal contain characters with any special properties such as control function or similar semantics? **Yes**
    (a) If Yes, describe in detail (include attachment if necessary): **Virama**
13. Does the proposal contain any Ideographic compatibility character(s)? **No**
    (a) If Yes, is the equivalent corresponding unified ideographic character(s) identified? **N/A**
       i. If Yes, reference: **N/A**
1 Introduction

This is a proposal to encode the Kaithi script in the Supplementary Multilingual Plane (Plane 1) of the Universal Character Set (ISO/IEC 10646). It is a revision of draft proposal L2/05-343 submitted to the Unicode Technical Committee on October 25, 2005. This revision incorporates some recommendations made by Michael Everson in L2/05-368 upon a review of L2/05-343. The major changes include the addition of new letters; the removal of dundas; the removal of fraction and accounting signs, which now appear in a separate proposal for “North Indian Accounting Signs” (L2/07-139); and the inclusion of additional specimens that further demonstrate the significance of Kaithi and its status as an independent, unique writing system.

1.1 Significance of the Script

Kaithi is a major historical script that was used throughout the British Indian provinces of Bihar and the North-Western Provinces & Oudh (hereafter, NWP&O); the region, roughly, of present-day Bihar and Uttar Pradesh. The script was used for writing Bhojpuri, Magahi, Urdu, and varieties of Hindi throughout the region. Kaithi was the common script of these languages until the early 20th century, at which time it was gradually replaced by Devanagari.

Kaithi is a historically significant and independent writing system that possessed a distinct scribal tradition and a print tradition based upon standardized metal fonts. Its importance in north Indian society can be measured by the activities for which it was employed and by the substantial number of materials written and printed in the script. Owing to its popularity, Kaithi was accorded status as the official script of the Government of Bihar in 1880. Official recognition contributed to the growth of the administrative and cultural importance of Kaithi. As the official script of Bihar, Kaithi was used in courts of law, where it replaced the Devanagari and Perso-Arabic scripts for writing both the formal Urdu of the courts and the local languages of the region. Also, on account of its use by large segments of the population, the governments of Bihar and NWP&O selected Kaithi as the medium of written instruction in the schools. The standardization of Kaithi by the British led to the development of printing in the script. The British produced Kaithi metal fonts, which were used for government documents and for printing textbooks. Printing in Kaithi was propagated by Western missionaries, who recognized Kaithi’s popularity and developed metal fonts for the script in order to print translations of Christian literature in regional north Indian languages.

Although primarily regarded as a secular script, Kaithi spread beyond use in administration, law, and education came to be used as a medium for writing religious and literary texts;

Kaithi was one of several scripts used throughout north India. As such, it appears in religious and Imanuscripts, commercial records, and other documents alongside Devanagari and other major scripts.

Its distinctiveness is evident from its use alongside several scripts, including Devanagari, in religious and literary manuscripts, commercial records, and court documents.

Beyond its administrative and commercial uses, Kaithi was used for informal correspondence and for the keeping of personal records. It was exported to the Caribbean, Africa, and other regions by north Indian diaspora communities and was maintained in these areas by their descendants.

1.2 Justification for Encoding

An encoding for Kaithi in the UCS is necessary for users engaged in the preservation, representation, and reproduction of written and printed Kaithi documents in digital media.

An encoding for Kaithi will provide users with the means to identify, represent, and process the script in electronic plain-text, not merely at the graphical presentation level. This is necessary for distinguishing between multiple scripts in single documents.

An encoding for Kaithi will also benefit non-specialist users seeking to access and preserve personal records written in the script. A digital standard for Kaithi will benefit individual researchers and preservation projects and will contribute to further study of the Kaithi script and documents written in it.

to support active research on Kaithi documents and efforts to preserve Kaithi manuscripts. Kaithi is used by specialists working with source materials in the regional languages of Uttar Pradesh and Bihar.

1.3 Acknowledgments

This project was made possible in part by a grant from the United States National Endowment for the Humanities (NEH), which funded the Universal Scripts Project (part of the Script Encoding Initiative at the University of California, Berkeley).

Digital reproductions of folios from a manuscript of the Mahāganapatiṣṭottara (shown in Figure 18 and Figure 19) are used here with permission from the University of Pennsylvania Libraries.

A digital reproduction of a folio from a manuscript of the Tale of Sudama (shown in Figure 28) is used here with permission from Sam Fogg, London.

2 Characters Proposed

The 73 letters in this proposal comprise the core set of Kaithi letters and signs. This set is sufficient for the general encoding and processing of Kaithi documents.

Consonants  There are 35 consonant letters:

<table>
<thead>
<tr>
<th>Kaithi Letter</th>
<th>Kaithi Letter</th>
<th>Kaithi Letter</th>
</tr>
</thead>
<tbody>
<tr>
<td>㏁</td>
<td>KAITHI LETTER KA</td>
<td>KAITHI LETTER DDA</td>
</tr>
<tr>
<td>㏂</td>
<td>KAITHI LETTER KHA</td>
<td>KAITHI LETTER DDDHA</td>
</tr>
<tr>
<td>㏃</td>
<td>KAITHI LETTER GA</td>
<td>KAITHI LETTER DDHA</td>
</tr>
<tr>
<td>㏄</td>
<td>KAITHI LETTER GHA</td>
<td>KAITHI LETTER RHA</td>
</tr>
<tr>
<td>㏅</td>
<td>KAITHI LETTER NGA</td>
<td>KAITHI LETTER NNA</td>
</tr>
<tr>
<td>㏆</td>
<td>KAITHI LETTER CA</td>
<td>KAITHI LETTER TA</td>
</tr>
<tr>
<td>㏇</td>
<td>KAITHI LETTER CHA</td>
<td>KAITHI LETTER THA</td>
</tr>
<tr>
<td>㏈</td>
<td>KAITHI LETTER JA</td>
<td>KAITHI LETTER DA</td>
</tr>
<tr>
<td>㏉</td>
<td>KAITHI LETTER JHA</td>
<td>KAITHI LETTER DHA</td>
</tr>
<tr>
<td>㏊</td>
<td>KAITHI LETTER NYA</td>
<td>KAITHI LETTER NA</td>
</tr>
<tr>
<td>㏋</td>
<td>KAITHI LETTER TTA</td>
<td>KAITHI LETTER PA</td>
</tr>
<tr>
<td>㏌</td>
<td>KAITHI LETTER TTBA</td>
<td>KAITHI LETTER PHA</td>
</tr>
</tbody>
</table>
Vowels  There are 10 independent vowels:

-  kaithi letter a
-  kaithi letter u
-  kaithi letter o
-  kaithi letter aa
-  kaithi letter uu
-  kaithi letter au
-  kaithi letter i
-  kaithi letter e
-  kaithi letter ii
-  kaithi letter ai

Vowel Signs  There are 9 dependent vowel signs:

-  kaithi vowel sign aa
-  kaithi vowel sign u
-  kaithi vowel sign ai
-  kaithi vowel sign i
-  kaithi vowel sign uu
-  kaithi vowel sign au
-  kaithi vowel sign ii
-  kaithi vowel sign e
-  kaithi vowel sign o

Special Signs  There are 5 special signs:

-  kaithi sign candrabindu
-  kaithi sign visarga
-  kaithi sign nukta
-  kaithi sign anusvara
-  kaithi sign virama

Digits  There are 10 digits:

-  kaithi digit zero
-  kaithi digit one
-  kaithi digit two
-  kaithi digit three
-  kaithi digit four
-  kaithi digit five
-  kaithi digit six
-  kaithi digit seven
-  kaithi digit eight
-  kaithi digit nine

Punctuation  There are 4 punctuation marks:

-  kaithi word separator
-  kaithi abbreviation sign
-  kaithi sentence separator
-  kaithi enumeration sign

2.1 Characters Not Proposed

The following characters are attested in printed and written Kaithi materials, but are not proposed for inclusion at the present because there is insufficient information regarding the character and its properties; the possibility of representing a character with another of similar or equal function; or a policy recommendation made by the Unicode Technical Committee. Space has been reserved in the Kaithi block to accommodate the possible inclusion of these characters in the future.

Vocalic R  The vowel sign ḍ for the Kaithi equivalent of द U+090B DEVANAGARI LETTER VOCALIC R (r) is shown in Figure 33 as a consonant-vowel ligature (r) with kaithi letter ka. The letter for this vowel is quite rare in Kaithi. Instead, the consonant-vowel combination ṓ for (kaithi letter ra + kaithi vowel sign ii) is used to transcribe the Sanskrit sound as it is pronounced in regional languages. The example given in Figure 33 is the only instance of this letter (as a vowel sign) found by the present author in the specimens examined. The independent letter for a Kaithi r has not been identified. There is, therefore, insufficient evidence to propose this character for inclusion in the Kaithi block.
DANDA and DOUBLE DANDA  The Unicode Standard currently recommends the use of U+0964 DEVANAGARI DANDA and U+0965 DEVANAGARI DOUBLE DANDA when these signs are to be used with other Indic scripts. The general thought is that introducing script-specific danās is similar to introducing distinct punctuation, as as commas and periods, for each script. For Kaithi, as for other scripts, the claim may be made that script-specific danās are necessary in order to ensure stylistic compatibility between danās and other characters. However, the Unicode Technical Committee has determined that unless evidence is presented to warrant the encoding of script-specific danās, the unification of these characters with those of Devanagari remains the recommendation. Although several specimens in the present proposal show distinctive Kaithi danās, they are not typologically different enough to justify a disunification with Devanagari danās. See section 5.11 for a detailed discussion of danās in Kaithi.

2.2 Basis for Character Shapes

The Kaithi script proposed here for encoding in the UCS is the Standard Kaithi developed and promoted by the British governments of Bihar and the NWP&O in the 19th century. The proposed script is an extension of Standard Kaithi, which includes letters not commonly found in written and printed documents, but that are considered traditional letters as attested in charts and other character inventories of the script. These rare letters are KAITHI LETTER NGA, KAITHI LETTER NYA, KAITHI LETTER NNA, and KAITHI LETTER SSA (see Figure 38 and Figure 39).

The letterforms of the Kaithi script proposed here are primarily normalized forms of the Kaithi metal fonts used in the Linguistic Survey of India, while the rarer letters absent from the former fonts are derived from those found in fonts of the Baptist Mission Press and in various charts illustrating the characters and other features of Kaithi. Digits and punctuation are derived from forms found in manuscripts and script charts.

The font for the proposed Kaithi script was drawn by Anshuman Pandey. The digitized letterforms were designed to express fidelity to the forms of Standard Kaithi found in the Linguistic Survey of India. The Kaithi font used to print specimens in the Linguistic Survey of India were most likely based on the Standard Kaithi of Bihar. Further discussion of Kaithi typefaces and regional variants is discussed in section 6.

The Standard Kaithi served as the basis for at least two sets of metal fonts. Both of these fonts were used by George A. Grierson for printing Kaithi in the Linguistic Survey of India. Other metal fonts exist, but their letterforms suggest influences drawn from other scripts. One of these is the metal font developed by the Baptist Mission Press of Calcutta. Table 3 compares letters from the two Kaithi fonts in the Linguistic Survey of India and the font of the Baptist Mission Press with the digitized font. The hand-written forms of the rare letters below serve as the basis for their counterparts in the proposed script.
<table>
<thead>
<tr>
<th>110A</th>
<th>110B</th>
<th>110C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1108</td>
<td>1109</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Glyph chart for Kaithi
<table>
<thead>
<tr>
<th>Character Name</th>
<th>Unicode Value</th>
<th>Category</th>
<th>Value</th>
<th>Width</th>
<th>Height</th>
<th>Decision Characters</th>
<th>Category Code</th>
<th>Decision Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>KAITHI SIGN CANDRABINDU</td>
<td>11080</td>
<td>Mn</td>
<td>0;N</td>
<td>0;N</td>
<td>N;N;</td>
<td></td>
<td>NSM;N;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI SIGN ANUSVARA</td>
<td>11081</td>
<td>Mn</td>
<td>0;N</td>
<td>0;N</td>
<td>N;N;</td>
<td></td>
<td>NSM;N;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI SIGN VISARGA</td>
<td>11082</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI LETTER A</td>
<td>11083</td>
<td>Lo</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI LETTER AA</td>
<td>11084</td>
<td>Lo</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI LETTER I</td>
<td>11085</td>
<td>Lo</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI LETTER II</td>
<td>11086</td>
<td>Lo</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI LETTER U</td>
<td>11087</td>
<td>Lo</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI LETTER UU</td>
<td>11088</td>
<td>Lo</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI LETTER E</td>
<td>11089</td>
<td>Lo</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI LETTER AI</td>
<td>1108A</td>
<td>Lo</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI LETTER O</td>
<td>1108B</td>
<td>Lo</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI LETTER AU</td>
<td>1108C</td>
<td>Lo</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI LETTER KA</td>
<td>1108D</td>
<td>Lo</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI LETTER KHA</td>
<td>1108E</td>
<td>Lo</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI LETTER GA</td>
<td>1108F</td>
<td>Lo</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN AA</td>
<td>11089</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN I</td>
<td>1108A</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN II</td>
<td>1108B</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN U</td>
<td>1108C</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN UU</td>
<td>1108D</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN E</td>
<td>1108E</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN AI</td>
<td>1108F</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN O</td>
<td>1109A</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN AU</td>
<td>1109B</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN KA</td>
<td>1109C</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN KHA</td>
<td>1109D</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN GA</td>
<td>1109E</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN TA</td>
<td>1109F</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN THA</td>
<td>1109A</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN DA</td>
<td>1109B</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN DA</td>
<td>1109C</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN DA</td>
<td>1109D</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN DA</td>
<td>1109E</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN DA</td>
<td>1109F</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN DA</td>
<td>1109A</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN DA</td>
<td>1109B</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN DA</td>
<td>1109C</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN DA</td>
<td>1109D</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN DA</td>
<td>1109E</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
<tr>
<td>KAITHI VOWEL SIGN DA</td>
<td>1109F</td>
<td>Mc</td>
<td>0;L</td>
<td>0;L</td>
<td>L;N;</td>
<td></td>
<td>NSM;L;</td>
<td>N;N;N;N;N</td>
</tr>
</tbody>
</table>

Table 2: Character Names and Properties
### Consonants

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ka</strong></td>
<td>ฤ</td>
<td>ฦ</td>
<td></td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
</tr>
<tr>
<td><strong>kha</strong></td>
<td>ฤ</td>
<td>ฦ</td>
<td></td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
</tr>
<tr>
<td><strong>ga</strong></td>
<td>ฤ</td>
<td>ฦ</td>
<td></td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
</tr>
<tr>
<td><strong>gha</strong></td>
<td>ฤ</td>
<td>ฦ</td>
<td></td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
</tr>
<tr>
<td><strong>ña</strong></td>
<td>—</td>
<td>—</td>
<td>ฅ</td>
<td>ฅ</td>
<td>—</td>
<td>—</td>
<td>ฅ</td>
<td>ฅ</td>
<td>—</td>
<td>—</td>
<td>ฅ</td>
</tr>
<tr>
<td><strong>ca</strong></td>
<td>ฤ</td>
<td>ฦ</td>
<td></td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>๚</td>
</tr>
<tr>
<td><strong>cha</strong></td>
<td>ฤ</td>
<td>ฦ</td>
<td></td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>๛</td>
</tr>
<tr>
<td><strong>ja</strong></td>
<td>ฤ</td>
<td>ฦ</td>
<td></td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>๛</td>
</tr>
<tr>
<td><strong>jha</strong></td>
<td>ฤ</td>
<td>ฦ</td>
<td></td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>ฦ</td>
<td>ฬ</td>
<td>๛</td>
</tr>
<tr>
<td><strong>ña</strong></td>
<td>—</td>
<td>—</td>
<td>ฅ</td>
<td>ฅ</td>
<td>—</td>
<td>—</td>
<td>ฅ</td>
<td>ฅ</td>
<td>—</td>
<td>—</td>
<td>ฅ</td>
</tr>
<tr>
<td><strong>ṭṭa</strong></td>
<td>セ</td>
<td>セ</td>
<td>セ</td>
<td>セ</td>
<td>セ</td>
<td>セ</td>
<td>セ</td>
<td>セ</td>
<td>セ</td>
<td>セ</td>
<td>セ</td>
</tr>
<tr>
<td><strong>ṭṭha</strong></td>
<td>セ</td>
<td>セ</td>
<td>セ</td>
<td>セ</td>
<td>セ</td>
<td>セ</td>
<td>セ</td>
<td>セ</td>
<td>セ</td>
<td>セ</td>
<td>セ</td>
</tr>
</tbody>
</table>

### Vowels

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a</strong></td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
</tr>
<tr>
<td><strong>ā</strong></td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
</tr>
<tr>
<td><strong>i</strong></td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
</tr>
<tr>
<td><strong>ī</strong></td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
<td>ฮ</td>
</tr>
</tbody>
</table>

Table 3: Comparison of Kaithi fonts used by Grierson in the *Linguistic Survey of India* (columns ‘A’ and ‘B’) and by the Baptist Mission Press (column ‘C’) with the digitized Kaithi font developed by Anshuman Pandey (column ‘D’).
3 Technical Features

3.1 Name

The name of the script in the UCS shall be Kaithi. The name of the script in the Latin transliteration recommended by ISO 15919 is Kaithi.\(^1\)

3.2 Classification

Kaithi is classified as a “Category C” (major extinct) as per the criteria specified in ISO/IEC JTC 1/SC 2/WG 2 N3002.\(^2\) The script is historically significant and there exists a substantial body of literature written in it.

3.3 Allocation

Kaithi is currently allocated in the Roadmap to the Supplementary Multilingual Plane (Plane 1) at the range U+11080..U+110CF.\(^3\) The five rows allocated for Kaithi in Plane 1 are sufficient to encode the script in the UCS.

The glyph chart in Table 1 and the list of character names in Table 2 show reserved code points. The positions U+110BA and U+110BB are reserved for the possible future inclusion of *danḍās*, which are attested, but not presently proposed for inclusion, as discussed in section 2.1. The range U+110CA..U+110CF following the digits is reserved for characters not yet identified.

3.4 Encoding Model

The Kaithi script is an abugida of the Brahmic type. It is written from left to right. The formation of syllables in Kaithi follows the pattern common to north Indian scripts. The encoding model for Kaithi may be based on the model implemented for Devanagari.

Consonant letters bear the inherent vowel *a* (Kaithi letter a) when unaccompanied by a vowel sign. The inherent vowel is suppressed by the *virāma* (Kaithi sign virama) to produce the bare consonant. The inherent vowel is changed by applying a vowel sign to the consonant. Vowel signs are placed above, below, and to the right of the consonant to which they are applied. The exception is *kaithi vowel sign i*, which is written to the left of the consonant.

Consecutive sequences of consonants are written as consonant conjuncts with *virāma*. Consonant conjuncts occur as true ligatures or as a sequence of consonant letters where an explicit *virāma* is written with each consonant except the final letter.

3.5 Character Properties

Vowels  All independent vowels have the following properties:

- General Category: Lo (Letter, Other)
- Combining Class: 0 (Spacing, split, enclosing, reordrant, and Tibetan subjoined)
- Bidirectional Class: L (Left-to-Right)

Vowel Signs  The dependent vowel signs are divided into two classes based upon their spacing attributes. The first class consists of kaithi vowel sign u, kaithi vowel sign uu, kaithi vowel sign e, and kaithi vowel sign ai, which have the following properties:

- General Category: Mn (Mark, Nonspaceing)
- Combining Class: 0 (Spacing, split, enclosing, reorderant, and Tibetan subjoined)
- Bidirectional Class: NSM (Non-Spacing Mark)

The second class consists of kaithi vowel sign aa, kaithi vowel sign i, kaithi vowel sign ii, kaithi vowel sign o, and kaithi vowel sign uu, which have the following properties:

- General Category: Mc (Mark, Spacing Combining)
- Combining Class: 0 (Spacing, split, enclosing, reorderant, and Tibetan subjoined)
- Bidirectional Class: L (Left-to-Right)

Consonants  All consonants have the following properties:

- General Category: Lo (Letter, Other)
- Combining Class: 0 (Spacing, split, enclosing, reorderant, and Tibetan subjoined)
- Bidirectional Class: L (Left-to-Right)

Special Signs  The kaithi sign candrabindu and kaithi sign anusvara belong to the “Mn” category, are of combining class “0”, and possess the bidirectional class value “NSM”.

The kaithi sign visarga belongs to the “Mc” category, is of combining class “0”, and possesses the bidirectional class value “NSM”.

The kaithi sign virama belongs to the the general category “Mn”, has a combining class value of “9” (Viramas), and is of the bidirectional class value “NSM”.

The kaithi sign nukta belongs to the the general category “Mn”, has a combining class value of “7” (Nuktas), and is of the bidirectional class value “NSM”. Combinations of consonants and kaithi sign nukta are to be treated as regular consonants.

Punctuation  All punctuation marks except for kaithi enumeration sign have the following properties:

- General Category: Po (Punctuation, Other)
- Combining Class: 0 (Spacing, split, enclosing, reorderant, and Tibetan subjoined)
- Bidirectional Class: L (Left-to-Right)

The kaithi enumeration sign possesses properties similar to the first class of dependent vowels. It belongs to the general category “Mn” and to the bidirectional class “NSM”.

Digits  All digits have the following properties:

- General Category: Nd (Number, Decimal Digit)
- Combining Class: 0 (Spacing, split, enclosing, reorderant, and Tibetan subjoined)
- Numerical Value: dependent upon on the respective digit: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9
- Bidirectional Class: L (Left-to-Right)
3.6 Collation

The collating order for Kaithi is dependent upon the language represented. Generally, languages written in Kaithi follow the sort order used for modern standard Hindi. Independent vowel letters are sorted before consonant letters.

Charts and tables of Kaithi and other north Indian scripts are inconsistent in their placement of anusvāra and visarga within the order of vowels. In some they appear at the head, in others at the back. The collation pattern used in dictionaries places anusvāra and visarga at the head of the vowel order before kaithi letter a and are written as combinations with the latter. This is the preferred order:

\[
\text{a, a, i, i, u, u, e, e, ai, o, au}
\]

Dependent vowels are sorted in the same position as their independent shape. Consonants with dependent vowels are sorted first by consonant letter and then by the vowel sign (including anusvāra and visarga) attached to the letter:

\[
\text{kañ, kah, ka, kā, ki, ku, ke, kai, ko, kau}
\]

The pattern for consonants is as follows:

\[
\text{ka, kha, ga, gha, na, ca, cha, ja, jha, ŋa, ŋa, ŋa, ra, ṍa, ṍa, ra, ṭha, ṭha, ṭha, ṇa}
\]

\[
\text{ta, tha, da, ṭha, na, pa, pha, ba, bha, ma, ya, ra, la, va, ſa, ſa, sa, ha}
\]

The \( \text{Kaithi letter dddha} \) is sorted in the same position as \( \text{Kaithi letter dda} \), and \( \text{Kaithi letter rha} \) is sorted with \( \text{Kaithi letter ddha} \). Cases in which the only difference between lexical forms is the unflapped and flapped retroflex stops (or nukta and non-nukta forms), eg. \( \text{padhanā} \) and \( \text{parhanā} \), the unflapped letter is sorted first. All letters written with Kaithi sign nukta are sorted by the same principle.

In some nasal-consonant conjuncts, the dental nasal, \( \text{Kaithi letter na} \), is used as the homorganic nasal letter for all articulation classes except for the labial class, where \( \text{Kaithi letter ma} \) is never substituted with \( \text{Kaithi letter na} \) (see Figure 37). When \( \text{Kaithi letter na} \) is used generically, it should be sorted as a member of the class to which the following consonant belongs (see section 5.7 for further details).

3.7 Character Classes

Based on structure, Kaithi consonant letters may be grouped into four different classes:

1. Class 1: Letters with full-height descenders:
2. Class 2: Letters with short descenders at the top:

\[ \text{ka kha ga gha ca ŋa na tha dha na} \]

\[ \text{pa ba bha ma ya ra va śa śa sa} \]

3. Class 3: Letters with rounded tops and no full-height descenders:

\[ \text{cha ja ta la} \]

4. Class 4: Letters with right-facing hooked tops and no full-height descenders:

\[ \text{jha ṭha ṭa ṭa ṭa ha} \]

The structure of letters influences the placement of vowel signs, *anusvāra*, *virāma*, and *nukta*:

- For Class 1 letters, above-base and below-base vowel signs are joined to the appropriate extremes of the descender. The *anusvāra* is centered above the top extreme of the descender. The *virāma* may be connected to the descender or positioned below it.

Examples: kku; cu; μu; šu; ke; ṇe; re; pām; n

- For Class 2 letters, above-base vowel signs are attached to the top of the descender and below-base vowel signs are centered below the letter. The *anusvāra* is positioned above the top extreme of the descender. The *virāma* is centered below the letter.

Examples: dhur; te; de; phe; ṭam; ph

- For Class 3 letters, above-base vowel signs are attached to the center of the top curve and below-base vowel signs are centered below the letter. The *anusvāra* is centered above the letter. The *virāma* is centered below the letter.

Examples: chur; lu; je; tam; l

- For Class 4 letters, above-base vowel signs are attached to the end of the hook and below-base vowel signs are centered below the letter. The *anusvāra* is centered above the letter. The *virāma* is centered below the letter.

Examples: jhu; the; he; dam; th
4 Background

4.1 Name of the Script

The name ‘Kaithi’ (कैथी) is derived from the term Sanskrit कायस्थ kāyastha, which refers to the name of the scribal community of north India. The term kaithi is the colloquial rendition of kāyasthi or kāyathī, which means “belonging to the scribe.” The script was also called Kaithināgari. During the British period, the name was Romanized as ‘Kayathi’. This was later simplified to ‘Kaithi’ and was adopted by the Government of Bihar as the official name and Latin spelling of the script. The name is transliterated in British books as both ‘Kaithi’ and ‘Kaithi’. This proposal uses the name ‘Kaithi’ without diacritics and recommends this form as the normalized name of the script in the UCS. It should be noted that while Kaithi refers to the formal name of a historical script of Bihar and the NWP&O, it also refers to a family of scripts and a style of writing (see section 7 for further details).

4.2 Languages Written in the Script

Kaithi is the traditional script of the Bhojpuri and Magahi languages. It was the popular script of the Awadhi and Maithili languages; the Devanagari and Tirhuta scripts being the more formal scripts, respectively. Presently, these languages are all written in Devanagari. Kaithi was also used to write Urdu or “Hindustani”, although now Urdu is associated with the Perso-Arabic script.

In several sources, Kaithi is associated with the ‘Bihari’ language. This association is misleading since there is no such language or language family called ‘Bihari’. The term ‘Bihari’ was adopted by Grierson to refer to the Eastern sub-family of the Indo-Aryan languages that included Bhojpuri, Magahi, and Maithili. He wrote “Bihārī means properly the language of Bihar” and it “occupies a middle place between Bengali and Eastern Hindi”. Other linguists propagated the concept of ‘Bihari’ as a language, following most likely from Grierson. Similarly, Hoernle wrote that the “Eastern Hindi” is commonly written in the Kaithi script. Hoernle’s “Eastern Hindi” includes Baiswari (a form of Awadhi) and Bhojpuri. It is for this reason that Kaithi is called a Bihari script; however, it’s geographic distribution and influence extended beyond the districts of Bihar (see section 7 for further discussion).

Awadhi
Kaithi was the common script for Awadhi, which is spoken predominantly in Uttar Pradesh and also in Nepal, Madhya Pradesh, and Bihar. Historically, Awadhi was written in both Kaithi and Devanagari. By the early 20th century, Devanagari had completely replaced Kaithi as the formal script for Awadhi, but its use in informal communication is attested through the middle of the century. A specimen of Awadhi written in Kaithi is given in Figure 10.

Bhojpuri
Kaithi was the traditional script of Bhojpuri, which is spoken predominantly in Uttar Pradesh and Bihar, with significant speaker populations in Madhya Pradesh. Bhojpuri speakers are also found outside of India in Nepal, Mauritius, Guyana, Suriname, Fiji, Trinidad and Tobago, and other parts of the Caribbean. The Devanagari script is used for literary and formal publication activity in Bhojpuri. Recent research indicates that Kaithi is still used for informal communication in Bhojpuri. Ethnologue also suggests currently activity by listing Kaithi and Devanagari as scripts used for the language. Kaithi was used in

---

5 Hoernle, 1880: 1. 6 Śākyavamsā, 1974: 64. 7 Grierson, 1903b: 1. 8 “Closely related to the Bengali script are the Gujarati and Kaithi scripts which were developed for writing the Gujarati and Bihari languages, respectively. Bihari is now written with Devanagari” (Coulmas, 1991: 190). Coulmas’s reference to a “Bihari” language is curious since no such language or language family exists. The “Bihari” languages referred to here are presumably Bhojpuri, Magahi, and Maithili. 9 Hoernle, 1880: 1. 10 Hoernle, 1880: viii. 11 Grierson, 1904a: 14. 12 Saxena, 1937: 20. 13 Grierson, 1903b: 48. 14 M. Verma, 2003: 519. 15 Gordon, 2005.
the Caribbean by Bhojpuri-speaking emigrants who carried the script with them. A specimen of Bhojpuri written in Kaithi is given in Figure 6 and a printed specimen in Figure 14.

**Magahi**  Kaithi was the traditional script of Magahi,\(^{16}\) which is spoken predominantly in Bihar, with significant populations of speakers in Jharkhand, and West Bengal. Devanagari is now preferred for formal use and literary activity, but research suggests that Kaithi may still be used for “personal communication and in semi-legal transactions”.\(^{17}\) A specimen of Magahi written in Kaithi is given in Figure 8 and a printed specimen in Figure 12.

**Maithili**  Maithili is spoken in Bihar and Nepal. In 2004, it was included in the Eighth Schedule of the Constitution of India and recognized as an official language of India. Maithili is traditionally written in the Mithilâksara (also known as Tirahuta, or Maithili), Kaithi, and Devanagari scripts. Brahmins use the Mithilâksara script and other communities use Kaithi.\(^{18}\) Devanagari began to place both Kaithi and the Maithili script at the turn of the twentieth century.\(^{19}\) A specimen of Maithili written in Kaithi is given in Figure 7 and a printed specimen in Figure 13.

**Urdu**  Kaithi was used for writing Urdu in the law courts of Bihar and Bengal when it replaced Perso-Arabic as the official script of these states during the British administration. The majority of extant legal documents from Bihar from the British period were in Urdu written in Kaithi. There is a substantial amount of such manuscripts, specimens of which are given in figures 20, 21, and 22.

**Other Languages**  At times, languages spoken in areas bordering the Kaithi region were written in the script. Kaithi was used for writing Bengali on the border of Bihar and Bengal.\(^{20}\) A specimen of Bengali in Kaithi is shown in Figure 11. On the Western periphery, in Rajasthan, the Marwari language was at times written in Kaithi; a specimen of which is given in Figure 28.

### 4.3 Brief Historical Overview

Kaithi is traditionally associated with the scribal community of north India and its literary practices. Through this association Kaithi is a secular script used for routine administrative purposes. Grierson wrote that “The Kaithi alphabet is, properly speaking, the alphabet used by the Kâyath or Kâyastha caste,—the writing caste of Northern India”.\(^{21}\) Although its origins are unclear, Kaithi became a significant writing system in the 17th century.\(^{22}\) From this time, the use of the script spread beyond the Kayastha community into the broader population. The *Education Commission Report* (1884) for Bihar states that “the Persian character was much affected by the higher classes of Muhammadans and learned Hindus; but the Kaithi character was known and used in every village in Behar by Hindus and Muhammadans alike”.\(^{23}\) The widespread use of Kaithi captured the attention of the British administrators of the Bengal Presidency. In 1880, Sir Ashley Eden, the Lieutenant-Governor of Bengal, issued a directive that established Kaithi as the official script of the Bihar government and ordered its exclusive use in the courts of the province.\(^{24}\) Eden’s goal was to replace the incumbent Persian script with a writing system more familiar to the population of Bihar.\(^{25}\) The official status conferred upon Kaithi was met with wide approval, as is evident from Grierson’s dedication in the *Handbook to the Kaithi Character*:

> To the Honorable Sir Ashley Eden, K.C.S.I., Lieutenant-Governor of Bengal, who, by first introducing Kaithi as the sole official character of our law courts, has done more for Bihâr than a decade of legislation\(^{26}\)

Eden’s adoption of Kaithi as the official script of Bihar was preceded by, and built upon, earlier British efforts to promote the script. Due to its importance, Kaithi was “raised to the dignity of type in Bihar” despite it being “primarily a cursive written hand”. In 1875, J. C. Nesfield, the Director of Public Instruction for Oudh, sought to develop a standardized Kaithi script in order to enlarge its utility for administrative and instructional purposes. Standardization was necessary due to variations in the style and manner of writing Kaithi across northern India. Nesfield assembled samples of Kaithi from all of the districts of NWP&O and devised a standard form of the script. Nesfield then produced the first set of Kaithi metal fonts based on this improved and standardized script. These fonts were used to print primers for elementary levels of the vernacular schools of Oudh. When the Government of Bihar recognized Kaithi as its official script in 1880, it also commissioned the creation of Kaithi metal fonts. These fonts were used for various administrative and other purposes as discussed in section 4.4.

The standardization and official recognition of Kaithi signifies the importance of Kaithi in north Indian society towards the end of the 19th century. In his preface to Edward H. Palmer’s *Oriental Penmanship*, a manual for the formal scripts of north India, Frederic Pincott writes that

> The derivative forms of Nāgarī,—such as the Kaithi, Mahājanī, Modī, Sarrāfī, Lundī, &c., &c.,—are not illustrated in this book, in that they are either local, or confined to special classes of people, and cannot, therefore, claim sufficient importance to lead to their general adoption in official documents. It may, however, be counted on as certain that, in the course of a few years, the Kaithi character will come into much more general use than is at present the case. It has been recognized officially in the province of Bihār, and will shortly spread to the North-West Provinces, where, indeed, it is almost universally

27 Grierson, 1903b: 11.
employed by Hindûs, although it has not yet been recognized in Courts and Kachharîs [courts].

Given its exclusive use in the law courts of Bihar, Grierson published his *A Handbook to the Kaithi Character* (1881) in order to introduce the script to aspiring British legal specialists in India. The association of Kaithi with the languages of Bihar led Grierson to use Kaithi metal fonts for all printed specimens of Bhojpuri, Magahi, and Maithili in volume five of the monumental *Linguistic Survey of India, Volume V* (1903). Grierson’s work remains the most comprehensive studies of Kaithi orthography and usage to date.

The factors that generated the official recognition of Kaithi also brought it into competition with Devanagari and Perso-Arabic and ultimately to its demise. In 1880, Rudolf Hoernle wrote that Devanagari “will probably in course of time entirely supersede the Kaithî; perhaps not altogether an advantage, as it can be written with less rapidity and ease than its rival”. The decline of Kaithi is generally attributed to the socio-political conditions of pre-Independence India. In the early 20th century, organizations such as the Nagari Pracharini Sabha (Society for the Promotion of Nagari) advocated the exclusive use of Hindi in Devanagari over regional languages and scripts. Thus, the linkage of Hindi exclusively with Devanagari “is a phenomenon that owes its origins primarily to the politics and sentiments of the past century.” The Nagari Pracharini Sabha was largely successful in popularizing Devanagari over regional scripts including Kaithi.

Kaithi continued as the official script in the courts of Bihar until at least 1913, at which time Devanagari became the preferred script owing to a momentous rise in its political and social currency. The accuracy of Hoernle’s forecast is evident in the contemporary dominance of Devanagari as the common script of Hindi and other north Indian languages. However, Grierson asserted in *A Handbook to the Kaithi Character* that “in the North-West Provinces, west of Benáras ... Hindi may fairly claim to be the vernacular of the country; but it is not, never was, and never can be, the vernacular of Bihar.” A similar claim can be made for the replacement of Kaithi by Devanagari.

### 4.4 Spheres of Use

The significance of Kaithi is evident from an examination of the contexts in which it was employed. The standardization of Kaithi led to the development of printing in Kaithi. However, the script was used in a broad spectrum of spheres, from routine administration to the annotation of Sanskrit manuscripts and from Christian proselytization to personal correspondence. These contexts are briefly discussed below.

There is also an effort initiated by the Government of India to catalogue and preserve Kaithi manuscripts. In February 2006, the National Mission for Manuscripts of the Government of India held a manuscriptology and palaeography workshop at the Khuda Baksh Oriental Public Library in Patna, Bihar, for the purpose of training researchers to read Kaithi and other scripts used in north Indian manuscripts.

**Court Records** Kaithi first rose to popularity by being established as the official script of the law courts of Bihar. Figure 20, Figure 21, and Figure 22 show samples of documents submitted to the courts of Bihar.


**Book Publishing** The standardization of Kaithi 1875 made it suitable for the printing of books. Kaithi entered the world of print through the major publishers, Khadgavilâs Press and Munshi Naval Kishor of Lucknow. The Khadgavilâs Press of Sâhib Prasâd Simha in Bankipur (near Patna) began printing Hindi grammars and primary-school readers in Kaithi script in 1886. After Kaithi was established as the official

---

script of Bihar, the Khadgavilās Press was the first printing house in Bihar to acquire Kaithi types and to begin printing Kaithi documents and books. Rāmadina Simha, the director of Khadgavilās Press, approached Grierson about Kaithi types and Grierson, through the assistance of the government, commissioned Kaithi types in Calcutta. Soon after, Khadgavilās Press began to publish Kaithi documents for the government, which ranged from agricultural and general administrative registers. In Awadh, the notable publisher of Urdu books, Munshi Naval Kishor of Lucknow, printed several professional texts and primers in Kaithi. Although the Kaithi printing tradition grew in strength, it did not become as influential as the Devanagari tradition. Writing on the condition of Kaithi publishing in 1893, Kellogg stated while books are printed in Kaithi, “it is by no means as common as the Devanāgarī”.

Hoernle corroborates Kellogg, stating that Kaithi “is used in printing as well as in writing; but owning to the preponderance of H. H. [High Hindi], which has adopted the Devanāgarī, the latter is much more common in books” and to justify the use of Devanagari in his grammar he writes that “I shall adopt it in this work also, as the more generally known of the two.”

Devanagari is the script for whatever is sought to be published today, or has been sought to be published since the early part of the twentieth century. Part of the reason may be simply the unavailability of typesetting for Kaithi. But more importantly, Devanagari in the greater Hindi area has a more scholarly image and is perceived as the right instrument for any kind of activity that has any claim to being literary. Devanagari as the ‘metropolitan script’ (as the name implied) has increasingly become the script of the new literatures of Bihari and other ‘regional’ languages and also replaced the earlier scripts.

Education  By 1881, the standard Kaithi had being prescribed for general use in the primary and middle vernacular schools of Bihar. “In Bihar, it [Kaithi] is used for teaching the lower classes, to whom a knowledge of Dēva-nāgarī is an unnecessary luxury, the elements of a primary education”. Used initially for printing elementary school primers, textbooks in standard Kaithi quickly appeared in primary and middle vernacular schools and were established for use in scribal examinations. It was known “from the commencement of education operations in Behar, that Kaithi was the popular character” of the province. As the British brought village schools of Bihar under government supervision, administrators learned that these schools taught their students to read and write in Kaithi. Therefore, Kaithi was retained in the schools of Bihar because the script “was the only one which could be employed with any hope of success, if the system of instruction was to be kept on the really broad and popular basis on which it rested”.

While Bihar retained Kaithi, the government of NWP&O ultimately adopted the reverse policy. The Report on Indigenous Education and Vernacular Schools of 1854 shows that 77,368 primers were printed in Kaithi, while 25,151 were printed in Devanagari in the province. Even though schools in which the Kaithi script was the prescribed written medium outnumbered those in which Devanagari was used, the government enacted a policy promoting Devanagari. Due to the prevalence of Hindi (as opposed to Avadhi, Bhojpuri, etc.) in the majority of the districts of the NWP&O motivated government officials to introduce Devanagari instead of Kaithi in scribal schools in which the vernacular language was the medium of instruction, not formal Hindi. The schools of Bihar continued to use the script until at least 1913.


Manuals and Script Primers

• Ambikāprasāda. 1880. *Kaithī-pratramālā*. Translated by Ambikāprasāda from the *Maktub-i Ahmadi*

---

Census Schedules  The popularity of Kaithi led other British administrators to employ the script in other bureaucratic endeavors, including census operations. Henry Beverly, the Inspector General of Registration in Bengal, listed Kaithi as one of the scripts in which census schedules and forms were printed:

The various forms required for the census were, with some few exceptions, printed at the larger Government Press ... near Calcutta. These forms had to be translated into several languages to suit the different nationalities to be found in Bengal. Thus, a Bengali translation was required for Bengal Proper; Hindustani in both the Persian and Kaithi character for Behar; Oorya for Orissa; Hindu in the Nagri character for Chota Nagpore and the Sonthal Pergunnahs, and Nepalese for some parts of Darjeeling.46

Commercial Transactions  Kaithi was used for recording commercial transactions. Receipts were the most common commercial documents written in Kaithi.47 Figure 23 shows a form printed in Devanagari and completed in Kaithi handwriting. Several official documents are similar to this receipt.

Literary and Religious Works  The use of Kaithi in administrative and legal affairs largely dissociated it from liturgical scribal practices. However, due very likely to its ubiquity, Kaithi was adapted for writing religious and literary works. It was used both as the sole orthographic vehicle in some manuscripts and in others it served a more annotative or interpretive function. Nonetheless, the presence of Kaithi and Devanagari on a single manuscript proves that Kaithi is distinct from Devanagari.

An impressive illuminated manuscript of the story of Sudama, taken from the Bhagavata Purana, is in the Marwari language written entirely in Kaithi. The manuscript is from Bikaner, Rajasthan and is dated to the early 19th century. Since the language is Marwari, one might expect the script to be the Mahajani script, which is more commonly associated with Rajasthani languages, or the Gujarati script. A comparison of the letterforms in the manuscript with Mahajani and Gujarati indicates that the script is distinctly Kaithi. The Miragavati of Qutban was originally written in 1503 in Avadh in the Persian script. Out of the five extant manuscripts of the Sufi romance, four are in Kaithi and one is in the Persian hand.48 Similarly, the 16th century Padmavat of Malik Muhammad Jayasi was also originally written in Avadh in the Persian script, but several derivative manuscripts are written in Kaithi.49

Although Devanagari was the preferred script for recording Hindu religious texts, Kaithi was also used for

the production of religious manuscripts. A manuscript of the Mahāgaṇapatisṭothra is remarkable for the fact that the content is written in both Devanagari and Kaithi (see figures 18 and 19). The stotra itself is Sanskrit written in Devanagari, but the commentary, which is also Sanskrit, is written in the Maithili style of Kaithi. Moreover, the last folio contains several annotations written in the Bhojpuri style of Kaithi. In other religious works, the sacred scribal importance of Devanagari influences Kaithi orthography. A manuscript of Jīvā Gosvāmi’s samkalpapatrī, or what Mukherjee and Wright call a “testamentary document”, was written in “Sanskritized Braj” in “Nagarized Kaithi”.50 Dated to 1665, the manuscript contains instructions and other matters related to the custody of the temples and libraries of the Caitanyapanthī Gauḍīya Vaiṣṇava sect founded by Jīvā Gosvāmi in Vṛndavana.

- Mahāgaṇapatisṭothra [anonymous]. [undated]. Miscellaneous stotra materials in the Devanagari and Kaithi scripts. Manuscript held at University of Pennsylvania (Poleman number 1876; University of Pennsylvania number 2584). Dimensions: 22.3cm × 11.9cm; 8-9 lines; stotra material in Kaithi script on ff.2-5. 5ff. 8.7a5 × 4.75. 8-9 lines.

- The Tale of Sudama. 1745-6 (Samvat 1802). Bikaner. Marwari language in Kaithi script. Ink, gold and opaque water-color on paper, 43 folios, 24 lines of text per page, 42 illustration running down the outer margin of the page, 1 full-page painting (f. 1v), 19th century pasteboard binding with marbled doublures, 29cm × 19.2cm.


Government Seals As an administrative script, Kaithi was used in government seals. It was one of three scripts used in the official seal of the Supreme Court of Appeals in Calcutta, the other two being Persian and Bengali (see Figure 29). The use of Kaithi in an official seal suggests the prominence of the script in government routine and further illustrates the formal acceptance of Kaithi as a script of state. One might expect such seals to contain text in Devanagari, but the absence of this script in official seals suggests that, by at least the 1850s, it had not acceded to its contemporary official position.

Missionary Operations The status of Kaithi as the common script of north India made it suitable to missionaries and bible societies who hoped to proselytize in the region. Once Kaithi was standardized by the government and metal types of the script were issued, Christian missionaries developed their own Kaithi metal fonts. Several bibles were printed in Kaithi (see Figure 24 and Figure 25). The importance of Kaithi in north India, and therefore, its value to missionaries, prompted the teaching of Kaithi in western universities. Kaithi was taught in the United States, most notably in the Department of Comparative Religion at the University of Chicago, where future missionaries to India were required to take a course in Hindi, which consisted of

[...] a careful study of the grammar of the Hindī language, both literary and provincial; the ordinary rules of syntax; exercises in Hindī composition and conversation; the writing of the language in the native character (both Nāgarī and Kaithī); [...]51

There was active communication about the publication of Christian literature in India. In particular, the annual report of The Baptist Missionary Society provided listing of languages and scripts in which such

50 Mukherjee and Wright, 1979: 298. 51 “Comparative Religion Notes,” 1894: 293.
literature appeared. Kaithi features prominently as the script preferred by the Society for publishing portions of the bible in Hindi.52


**Personal Records and Correspondence**  Kaithi was used for maintaining family records, private correspondence, and transactional accounts. Thomas Metcalf writes that the use “of distinctive scripts such as Kayathi and Mahajuni was common practice among Indian families, many of whom, especially among the mercantile community, wished in this way to preserve their records from prying eyes of uninitiated outsiders”.53 As large numbers of Bhojpuri-speaking inhabitants of north India migrated to Trinidad and other locations in the Caribbean during the 19th and 20th centuries, they carried the Kaithi script with them. The present author was contacted by two individuals who trace their families’ ancestry to north India and whose ancestors maintained the use of Kaithi outside of India, Mr. Nigel Ramoutar and Dr. Dipendra Sinha. Mr. Ramoutar’s family migrated from eastern Uttar Pradesh to Trinidad at the turn of the 20th century. His grandparents maintained family records and personal correspondence in Kaithi, which have been preserved by his family in Trinidad. Dr. Sinha, whose family hails ancestrally from Bihar, informed the author that Kaithi was used by migrant Indian communities in Jamaica as well. At present it is unknown exactly how prevalent the use of Kaithi was in Trinidad, Jamaica, and other locations in the Caribbean. Nonetheless, the preservation of the Kaithi script by Indian immigrants is evidence for the popular strength of the script in north India.

52 *The Journal of Sacred Literature* often published information about the progress of activity. In a section titled “Intelligence” the fifth volume, the journal reports that “From the 61st (1853) Report of the Baptist Missionary Society we learn that ... [the Hindooee Gospels, in the Kaithi character, have been undertaken and carried through the press to John vii., by the joint labours of Mr. Leslie and Mr. Parsons of Monghir”. 53 Metcalf, 1967: 679ff.
Epigraphical Records  Inscriptional records in Kaithi are extremely rare. However, the archaeology gallery at the Bharata Kala Bhavan at the Banaras Hindu University supposedly has in its holdings a copper plate bearing an inscription in Sanskrit in the Kaithi script.\(^{54}\) The text of the inscription is a land grant by Baj Bahadur Chandradeva (fl.1090), a ruler of the Gahadavala dynasty of Kanyakubja. If the script is in fact Kaithi, the Chandradeva’s inscription would be the earliest attested use of Kaithi.

5  Orthography

5.1  Distinguishing Features

Two of the most distinguishing features of Kaithi are the absence of the head-stroke and the presence of ‘serifs’ at the terminals of vertical strokes in metal fonts.

5.2  Vowels

In some cases no distinction is made between \textsc{kaithi letter i} and \textsc{kaithi letter ii}, or between \textsc{kaithi letter u} and \textsc{kaithi letter uu}.\(^{55}\) The tendency is to use the long vowels for writing both lengths in both the independent and dependent forms. However, the distinction between short and long forms are observed in print, primarily to preserve accuracy of pronunciation. This practice generally does not affect the other vowels.

5.3  Sibilant Consonants

In the languages of Uttar Pradesh and Bihar, there is a practice of assimilating retroflex and dental sibilants with the palatal sibilant. This is reflected in Kaithi orthography through the writing of the letters \textsc{kaithi letter ssa} and \textsc{kaithi letter sa} as \textsc{kaithi letter sha}. Both the palatal and dental sibilants are found in the Kaithi specimens in the \textit{Linguistic Survey of India}.

In a specimen of Maithili,\(^ {56}\) the word \textit{khusṭ} is written with the dental sibilant as \textit{khusṭ}:

\[
\text{कखुः स काभः काय सावः हस महा सा पा सा के } \text{कहुः}
\]

and in a specimen of Magahi\(^ {57}\) it is written with the palatal sibilant \textit{khusṭ}:

\[
\text{केण्या केण्या बालं जोकं में पुनः केण्या खसुंवः पुल्लकं कहे नास खसुंव}
\]

In some varieties of Hindi, the retroflex sibilant \textit{ṣa} is pronounced as the aspirated velar stop \textit{kha} and is written as \textsc{kaithi letter kha}. There are, however, no standard conventions regarding such practices and the correct spelling of words with the appropriate sibilant letter rests largely with the writer’s knowledge of lexical sources. For example, in Figure 3, Grierson shows the Kaithi counterpart of \textsc{U+0937 DEVANAGARI LETTER SSA} as \textsc{kaithi letter sha}, but in Figure 15, he shows \textsc{kaithi letter ssa}. Although rare in Kaithi documents, \textsc{kaithi letter ssa} is nevertheless attested and should be considered part of the character inventory. Its proposed form is based on the shape of the letter as found in Figure 15.

\(^{54}\) Bharat Kala Bhavan, Banaras Hindu University, 2001.  \(^{55}\) Hoernle, 1880: 2.  \(^{56}\) Grierson, 1903b: 74.  \(^{57}\) Grierson, 1903b: 74.
5.4 Nasal Consonants

Letters for the velar (KAITHI LETTER NGA), palatal (KAITHI LETTER NYA), and retroflex (KAITHI LETTER NNA) nasals are attested, but rarely found in use. They appear, however, in tables of the Kaithi script and are included here for completeness (see figures 3, 4, and 33).

The shapes of KAITHI LETTER NGA and KAITHI LETTER NYA resemble to some extent variant forms of KAITHI LETTER I and KAITHI LETTER U, respectively. It is possible that the these two vowel letters were used to represent the rare independent forms of NGA and NYA, but it is also possible that the resemblance is more likely attributable to the close structure of the characters.

The KAITHI LETTER NNA is used frequently in the bibles published by the Calcutta Bible Society. In the following they write नन्दननोकी भिनिन्ना (nandanokī bhinnā) using the letter:

5.5 The letters BA and VA

In the languages of Bihar there is no distinction between /b/ and /v/. A difference between the two sounds was made in writing by adding a dot to the letter for ba.\(^{58}\) Commonly, KAITHI LETTER BA is used for both /b/ and /v/, but in cases where phonetic accuracy is required, KAITHI LETTER VA is used to represent /v/. The following example\(^{59}\) shows a differentiation between ba and va through the use of the underdot to represent va.

5.6 The letter YA

The semi-vowel ya is typically written ୯, although in some documents it appears as ୯. Grierson uses a form without the nukta to write baniyā “merchant”:

However, the Kaithi font used by the Calcutta Bible Society (1851) uses the underdotted form of ya, as noted in the name Dāyūd:

The difference between ୯ and ୯ is a stylistic variation, not a phonological difference that is differentiated orthographically as in the case of Bengali ಡ U+094F BENGALI LETTER YA and ଡ U+09DF BENGALI LETTER YA.

Proposal to Encode the Kaithi Script in Plane 1 of ISO/IEC 10646

Anshuman Pandey

Yya. Presumably, the underdot was applied to the Kaithi yya in order to distinguish it from य KAITHI LETTER CA.

Generally, य is often written in place of र, as in the word योग jog (from Skt. योग्य yogya) “suitable”:

नं हन गोल्प मेला देवाया योग योग देव म म प म म गोम म ने

5.7 Nasalization

Anusvāra In Kaithi the anusvāra is used to represent true vowel nasalization. It is not used for indicating class nasals in nasal-consonant conjuncts, as is the practice in writing Hindi in Devanagari. Other signs for nasalization such as the Devanagari candrabindu are generally not used in Kaithi. A peculiar practice for marking nasalization is used in the bibles printed by the Calcutta Bible Society. In these bibles there are two marks for indicating nasalization. One mark resembles the regular anusvāra and the other is a small open-circle. The open-circle mark appears to be used to indicate class nasals, while the anusvāra represents true vowel nasalization:

Grierson shows the use of candrabindu to write the the Kaithi transcription of र:

Nasal-Consonant Conjuncts Nasal-consonant conjuncts are written either as ligatures with the half-form of the appropriate class nasal letter or with the full form of the nasal letter marked with an explicit virāma. A peculiar manner for writing such ligatures is used in the printed Kaithi specimens in the Linguistic Survey of India. For nasal-consonant conjuncts that are written as ligatures, the dental nasal न य KAITHI LETTER NA, is used as the generic nasal marker for all articulation classes (see Figure 37). For example, the conjunct न्ता is represented appropriately as न्ता. However, the conjunct न्दा is represented as न्दा. Unlike the case of न्ता, where the dental nasal letter ना is allied with the dental stop ता, in न्दा, the dental nasal is used to represent the the retroflex nasal ना since the following consonant of the following consonant, दा, is a retroflex stop. In such instances, KAITHI LETTER NA should be sorted as a member of the class to which the second element of the conjunct belongs.

The form of nasal-consonant conjuncts is not uniform in Grierson. In a specimen of Magahi,60, Grierson uses virāma to write the न्दा conjunct:

and in a specimen of Maithili,61 he uses a ligature to represent the न्दा conjunct:

60 Grierson, 1903b: 129. 61 Grierson, 1903b: 75.
5.8 Consonant-Vowel Ligatures

In writing, syllables composed of consonants with vertical descenders at the right edge and the vowels \( u \) and \( \bar{u} \) are represented as distinct consonant and vowel sign combinations or as ligatures. The latter is often a swash technique that results from rapid writing. In such cases, the vowels \( i \) and \( \bar{i} \) may at times be written in a manner that resembles the dependent forms of \( u \) and \( \bar{u} \). Therefore, these forms of \( u \) and \( \bar{u} \) are often curved leftward and over the consonant letter. Generally, the appropriate consonant-vowel combination is apparent from context.

In the excerpt below, the labeled portions indicate consonant-vowel ligatures. Portions ‘A’ and ‘B’ are the syllable \( k\bar{i} + \bar{\imath} = \bar{k\imath} \); portion ‘C’ is \( gu \bar{\imath} + \bar{\imath} = \bar{y} \); and ‘D’ is \( kh\bar{u} \bar{\imath} + \bar{\imath} = \bar{y} \):

In the following excerpt, portions ‘A’ and ‘D’ are the syllable \( tu \); portion ‘B’ is the syllable \( du \); and portion ‘C’ is the syllable \( ku \):

Such consonant-vowel ligatures are found only in written documents. In written documents \( \bar{k} + \bar{\imath} \) might appear as \( \bar{\beta} \), but in print the tendency is to use the mātrā form of vowels explicitly, such as \( \beta \). In the Kaithi font used by the Calcutta Bible Society, the mātrā form of \( u \) is a shallow arc and appears structurally related to the swash form found in written documents. In the excerpt below, portion ‘A’ is the syllable \( su \), ‘B’ is \( bhu \), and ‘C’ is \( mu \). This shape of the \( u \) mātrā may have been derived from the practice of writing \( u \) as ligatures:

The font used by Grierson in the Linguistic Survey of India has distinct mātrā forms for \( u \). The forms of \( u \) mātrā in the words marked ‘A,’ ‘B,’ and ‘D’ represent the typical shape of the dependent vowel; for example \( ku \bar{\imath} + \bar{\imath} \) as \( \bar{\beta} \) instead of \( \beta \). The shape of the syllable \( hu \) in portion ‘C’ is a ligature designed to accommodate the descending tail of \( ha \) with the \( u \) mātrā. The combination \( \bar{k} + \bar{\imath} \) is written as \( \bar{\beta} \) to avoid the appearance \( \bar{\beta} \).
Other forms include \( \tilde{k} + \tilde{\varsigma} \), which appear in print as \( \tilde{k} \), but in written documents as \( \tilde{\varsigma} \).

5.9 Consonant Conjuncts

There is substantial irregularity in the writing of conjuncts in Kaithi. This is due in part to the manner in which consonant clusters are handled in the languages that Kaithi was used to represent. In spoken Awadhi, Bhojpuri, and Magahi, there is a tendency to simplify consonant clusters through metathesis, or the insertion of a vowel between two consonants. This process is reflected in Kaithi orthography. For example: Sanskrit *karma* (Dev. कर्म्) becomes *karam* (Kai. कारम्); Sanskrit *pradeśa* प्रदेश becomes *pardes* पर्देस्; and Sanskrit *snāna* स्नान becomes *asanān* असानान्. Other practices of simplifying include: Sanskrit *vyavahāra* (Dev. व्यवहार) becomes *beohār* (Kai. बेोहार्); Sanskrit *jñāna* ज्ञान is simplified to *giān*. In instances where metathesis does not occur, the representation of the cluster was a conjunct depended upon the diligence of the scribe or in the case of printing, on the limitations of the font. Therefore, conjuncts may be written as ligatures, with half-forms, with explicit virama, or implied. For example, the conjunct *mba* may be written as \( \tilde{m} \) or \( \underline{m} \) or \( m \). When encoding Kaithi in Unicode, conjuncts should always be written with *virāma*. The conjunct *mba* should be expressed as

\[
\text{KAITHI LETTER MA + KAITHI VIRAMA + KAITHI LETTER BA}
\]

In instances where there is a requirement to encode conjuncts as they appear in a source document, then \( \text{U+200C ZERO WIDTH NON-JOINER} \) \( \text{and U+200D ZERO WIDTH JOINER} \) should be used. The sequence \( \text{mb} \) written with a half-form of *KAITHI LETTER MA* is expressed in Unicode as

\[
\text{KAITHI LETTER MA + KAITHI VIRAMA + U+200D ZERO WIDTH JOINER + KAITHI LETTER BA}
\]

The form \( \text{mb} \) is expressed as

\[
\text{KAITHI LETTER MA + KAITHI VIRAMA + U+200C ZERO WIDTH NON-JOINER + KAITHI LETTER BA}
\]

In Writing At times, the scribe would writing conjunct with an explicit *virāma* at other times he would produce a true conjunct form. Conjuncts, however, appear more often in Maithili documents (see Figure 4). The example below shows the two consonant conjunct *pra* in the word *pragana* marked ‘A’ and the three consonant conjunct with dependent vowel sign *stṛ* in the word *distṛkāt* marked ‘B’:

Figure 4 shows conjuncts that may be encountered in the Maithili style of written Kaithi. Some Kaithi documents also show ‘false’ conjuncts, especially when the second element of the conjunct is *ra*. In the following example the word *paraganāt* is written *praganāt*:
The example below illustrates a case where a ligature is used to write the conjunct mpu, but not rna in the word sampurna:

In writing, doubled consonants are written only once. A word like patta पत्ता is written pata पता. The use of virāma in printed Kaithi, as opposed to using a single character, may arise from the intention to represent phonological accuracy in published documents.

**In Print** In printed Kaithi, consonant clusters are represented both as ligatures and with virāma. It is unknown whether this is a reflection of actual practice or a limitation in the Kaithi fonts used for typesetting. In metal fonts, there existed a limited number of character primitives that could be used to produce conjuncts. The application of these primitives in the formation of conjuncts, however, does not appear to follow any patterns.

In some instances, consonant clusters are written using conjunct forms, as is done in the word acchā, where a half form of ca is attached to the full form of cha:

but in another specimen, a virāma is used to write the cluster cca:

Another example of inconsistent use of conjuncts is shown below. The word दोस्त dost is written in two ways in the Linguistic Survey of India. In the example below, the cluster sta is written with a ligature

but in another specimen, the conjunct is represented with a virāma form:

The example below shows the use of KAITHI LETTER SA to represent KAITHI LETTER SSA. Here is it used in a half-form to write the conjunct sta:

---

62 Kellogg, 1893: 23. 63 Grierson, 1903b: 75 64 Grierson, 1903b: 125
5.10 Word Boundaries

Punctuation is inconsistent in Kaithi. Grierson writes that “Kaithi has no stops except the full period” and “it is not customary to leave any space between the words.” Although lack of punctuation is not foreign to Indian scribal traditions, in Kaithi the lack of word boundaries is most likely a result of the need for writing rapidly in courts and other administrative arenas. Standardization began to change this. Grierson writes the “Standard Kaithi, however, used in Government offices, does separate its words.” The practice of marking word boundaries also depended upon the scribe’s regard for applying spaces or other delimiters between words. Some marked word boundaries consistently, others showed no such regard. Some manuscripts show the practice of identifying not only sentence boundaries, but boundaries between words, as well.

In printed Kaithi, word boundaries are generally marked by spaces and the end of sentences are distinguished using the danḍā or double danḍā. The example below shows the use of dashes to mark word boundaries:

The kaithi word separator is used to represent the word boundary marker.

5.11 Line and Paragraph Boundaries

There is no set convention for marking line boundaries in Kaithi. In many manuscripts the text is written continuously without delimitation of syntactic boundaries. Several scribes, however, employed various orthographic devices to mark line and paragraph terminations. The danḍā was introduced in printed Kaithi to mark phrase or sentence terminations.

Danḍā  As with other scripts derived from Brahmi, Kaithi uses the danḍā and double danḍā to indicate line endings and other terminations. The function of the Kaithi danḍā and double danḍā is similar to that of Devanagari and other Indic scripts, but should be encoded separately, not unified with Devanagari danḍā or double danḍā due to the distinct shape of the Kaithi forms.

Compare Kaithi danḍā l to Devanagari l, and Kaithi double danḍā ll to Devanagari ll.

The table below compares the forms of danḍā in the metal fonts used by Grierson to print Kaithi and Devanagari in the Linguistic Survey of India:

---

65 Grierson, 1899: 4.
The Kaithi forms of *danda* differ structurally from the Devanagari forms. The Kaithi forms possess a serif-like feature at the extremes of the *daṇḍā*. Devanagari does not have this feature.

The following example shows the use of dandas in written text. The dandas are written with serifs:

Swash Marks In addition to *daṇḍā* and double *daṇḍā*, lines and paragraphs are often terminated with horizontal lines. In the specimens in Figure 7 and Figure 21, the end of a paragraph is marked with two types of lines that fill the remainder of the line and extend to the margin of the text block.

The **Kaithi word separator** (–) can be used to represent both the dashes used for marking word boundaries and the horizontal line used for marking the end of sentences. The **Kaithi sentence separator** (=) can be used to represent the end of paragraph mark.

Other Marks The specimen below\(^\text{66}\) shows the use of a cross to mark phrase boundaries, while swashes are used to separate words:

5.12 Hyphenation

Hyphenation at line boundaries is rare in written Kaithi, but does appear in printed documents. In manuscripts, the scribe would simply break off writing anywhere in a word, and continue on the next line. If hyphenation were to be applied, it would occur within words at syllabic boundaries. The example below shows a hyphen splitting the word *कृषिका* at a line boundary:

\(^{66}\) Grierson, 1903b: 82.
An interesting example of hyphenation occurs in the example below. Here hyphenation splits a conjunct formed with virāma in the word barāmhan बाराम्हन. This example also illustrates metathesis in the consonant conjunct hma, which is rendered mha.

5.13 Abbreviation

Abbreviation is common in Kaithi documents, but there is no consistency in the manner in which it is indicated. Typically, abbreviation is indicated through the use of signs.

**Circle** The काithi abbreviation sign is found in written documents. It is similar in function to उ+0970 devanagari abbreviation sign. The Kaithi abbreviation sign may resemble the digit zero, however, the two are distinguishable through context. This sign is used to abbreviate common words and phrases, primarily at the beginning of legal documents. For example, in the following, the sequence ल to लि ('lit.'), using a variant form of क KAITHI LETTER LA, represents the Sanskrit likhitam लिखितम्, meaning “it is written” and is abbreviated due to its common use as an introductory element in written statements submitted to the courts.

**Colon** The following shows the use of a visarga-like character for marking abbreviations. In the example below, portion ‘A’ (से) is an abbreviation for the title से sekhd ‘sheikh’; portion ‘B’ ता is an abbreviation for the word तारिक्ष ‘date’.

**Below-Base Slash** Another method of abbreviation is evidenced in the following example. A virgule-like mark is written in the portion ‘B’ (नवम्बर) to abbreviate the word नवम्बर “November”. The text in portion ‘A’ is the same form shown in the example above.
**Swash** Another method, which is rare, is to indicate abbreviation through swash. The circled portion below shows the abbreviation नूठ (mudā) for the word नूठ (muda’an):

5.14 Enumeration

The काईंट enumeration sign is used in the writing of enumerated lists and numerical sequences. It is written either above or below a digit or sequence of digits. It is a stylized version of कैंट letter na and is an abbreviation of nambar, the transliteration of the English word ‘number’. Its basic function is identical to नूठ numero sign, the sign for indicating ordinal numbering in the Latin alphabet. The following specimen from a legal document illustrates the use of the enumeration sign to mark two distinct numerical contexts:

The circled portion labeled ‘A’ shows the enumeration sign written above the digit. It functions here to indicate the first item in an enumerated list. Portion ‘B’ shows the enumeration sign written under the numerical sequence. It is used here to indicate a court-case number. Portions ‘C’ and ‘D’ are identical in function to ‘B’. It appears as though the enumeration sign is written through the digit १ in ‘C’, but this is most likely the result of the descending stroke of the digit, as compared with ‘D’. Here the enumeration signs refer to plaintiffs (नूठ muda’an) ‘1’ and ‘2’ in the given case.

The function of the enumeration sign in Portion ‘A’ is different from its function in ‘B’, ‘C’, and ‘D’. In the former, it is used as a label in an enumerated list. In the latter three, it is used to mark specific numerical references. Nevertheless, there is no need to encode these forms separately as they may be considered positional variants of the same sign.

The following example shows the enumeration sign written in-line before the number being marked. Portion ‘E’ is the Kaithi text २।१।०.

5.15 Nukta

The कैंट sign nukta is commonly used to distinguish कैंट letter ba from कैंट letter va. It is also written under कैंट letter ya to produce the variant form २।१।०. In some cases, the nukta is
written below letters to represent sounds from Urdu. For example, the following specimen\textsuperscript{67} represents the sound of \texttt{U+0632 ARABIC LETTER ZAIN} by placing \textit{nukta} under \texttt{KAITHI LETTER JA} in the word \texttt{nazdik}:

\begin{verbatim}
\texttt{nazdik}
\end{verbatim}

The \textit{Kaithi o Hindi barnamala} also shows the use of \textit{nukta} under \texttt{KAITHI LETTER JA} to represent \texttt{U+0632 ARABIC LETTER ZAIN}. Similar to \texttt{U+095B DEVANAGARI LETTER ZA}, \texttt{KAITHI LETTER JA} with \textit{nukta} may also be used to transliterate \texttt{U+0630 ARABIC LETTER THAL}, \texttt{U+0636 ARABIC LETTER DAD}, \texttt{U+0638 ARABIC LETTER ZAH}, \texttt{U+0630 ARABIC LETTER ZAIN}, and \texttt{U+0698 ARABIC LETTER ZEH}. Also used in the text is a \textit{nukta} under \texttt{KAITHI LETTER KA}, similar to Devanagari \texttt{U+0958 DEVANAGARI LETTER QA} to represent \texttt{U+0642 ARABIC LETTER QAF}:

\begin{verbatim}
A B

\texttt{nukta}
\end{verbatim}

In a similar manner, the following shows the use of \textit{nukta} under \texttt{KAITHI LETTER KHA}, similar to Devanagari \texttt{U+0959 DEVANAGARI LETTER KHHA} to represent \texttt{U+062E ARABIC LETTER KHHA}:

\begin{verbatim}
(५) फल्कानकेचं सिंधुमैदं फिटितें हैं। सिम्यें
\end{verbatim}

The following example shows the use of \textit{nukta} under \texttt{KAITHI LETTER PHA}, similar to Devanagari \texttt{U+095E DEVANAGARI LETTER FA} to represent \texttt{U+0641 ARABIC LETTER FEH}:

\begin{verbatim}
(५) सुणाईं लुप्तशेष्याड़ों। काटन के साध्यशेषवय।
\end{verbatim}

The form of Kaithi letters in the \textit{Kaithi o Hindi barnamala} appear too heavily influenced by Devanagari, the Kaithi letters with \textit{nukta} are no less. Thus, although the letters with \textit{nukta} are attested in Kaithi documents, they are not recommended for encoding in the Kaithi block.

But the use of \textit{nukta} is inconsistent. The present author has not yet seen \texttt{KAITHI LETTER GA} with \textit{nukta}, although presumably it could be used like \texttt{U+095A DEVANAGARI LETTER GHHA} to represent \texttt{U+063A ARABIC LETTER GHAIN}. It might be possible to reproduce the repertoire of \textit{nukta} letters of the Devanagari block in the Kaithi block, but given that such letters are not generally listed in tables of Kaithi letters, it is better to encode the \texttt{KAITHI SIGN NUKTA} so that such letter can be created as needed.

\textsuperscript{67} Grierson, 1903b: 74.
5.16 Headstroke

Several manuscripts and books show Kaithi written and printed with a headstroke similar to that of Devanagari. In the majority of manuscripts, the line is not a headstroke, but a typographic device used for emphasis, titling, or sectioning. Grierson notes that “in many documents it is customary to rule only the first line, for show; and to leave the rest unruled, for comfort”. Hoernle adds

> Sometimes a series of lines is first ruled across the page, and the letters are afterwards hung on to them. These lines must not be confounded with the headstroke of the Devanagari, and in native writing the two are easy to distinguish.

The following example shows the use of ruled lines for written Kaithi:

![Ruled Kaithi Lines](image)

The following examples shows a line used only for the first line:

![First Line Kaithi](image)

Some Kaithi fonts were designed with headstroke, presumably to render similarity between Kaithi letters and Devanagari type:

![Kaithi with Headstroke](image)

Figure 34 shows a comparison of hand-written Devanagari and Kaithi. At first glance, it appears that the Kaithi is written with the head-stroke similar to the Devanagari. Comparing the two scripts, it becomes clear that the head-stroke accompanying the Kaithi letters is actually a ruled-line, which unlike the actual head-stroke of the Devanagari letters, is not broken between individual letters.

---

68 Grierson, 1899: 4. 69 Hoernle, 1975: 1f.
5.17 Accounting Notation

Accounting signs are commonly found in Kaithi documents. Among these are fractions (eg. \( \frac{1}{16} \), \( \frac{1}{2} \), etc.), currency (eg. rupayā mark), and other marks (eg. placeholder mark). A few of these are depicted in Figure 5. These signs are important for the complete representation of Kaithi in the UCS; however, the present author has determined that these signs are not exclusive to Kaithi. They are used across several north Indian scripts, including Devanagari, Gujarati, Gurmukhi, Mahajani, Maithili, and Modi. The author, therefore, recommends that these accounting signs be unified and encoded in a separate block in order to enable general use with other scripts. These accounting signs have been proposed for inclusion as the North Indian Accounting Signs block in L2/07-139R.

6 Regional Variants and Typeface Styles

6.1 Regional Variations

There are three distinct regional styles of Kaithi, each associated with the Bhojpuri, Magahi, and Maithili languages. In the Linguistic Survey of India, Grierson states:

So far as Bihārī is concerned, the Kaithī used changes slightly according to locality, and three varieties are recognised, viz., that of Mithilā, that of Magah, and that used in writing Bhojpuri.\(^70\)

These regional styles of Kaithi are compared in Figure 3. The differences between the three styles are due to regional scribal traditions. In his Handbook to the Kaithi Character, Grierson provides examples of model hand-written specimens in each of the three styles. He describes the Bhojpuri style as “the most legible” (Figure 6); the Maithili as “the most elegant” (Figure 7); and the Magahi as “a mean between the two” (Figure 8).\(^71\)

Although not mentioned in many sources, it is likely that Nesfield relied upon the Magahi style as the basis of his standard Kaithi. David Diringer alludes to this, writing that the “Magahi type had been adopted by the Bengal Government for official Bihari publications; books are printed in it in Patna, and the character has become more or less standardized”.\(^72\)

Encoding these styles separately is unnecessary. The primary differences are due to swash techniques that influence the writing style. A few variant shapes exist, but these can be unified with the prevailing form of the letter. Separate fonts may be produced for the regional styles and their variant letters.

<table>
<thead>
<tr>
<th>KAITHI LETTER A</th>
<th>STANDARD</th>
<th>VARIANT</th>
<th>KAITHI LETTER NYA</th>
<th>STANDARD</th>
<th>VARIANT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>أسل</td>
<td>ٱئ</td>
<td></td>
<td>ٱئ</td>
<td>ٱئ</td>
</tr>
<tr>
<td>KAITHI LETTER AA</td>
<td>ﭑئ</td>
<td>ٱئ</td>
<td>KAITHI LETTER PHA</td>
<td>ٱئ</td>
<td>ٱئ</td>
</tr>
<tr>
<td>KAITHI LETTER I</td>
<td>۱٢٣</td>
<td>۱٢٣</td>
<td>KAITHI LETTER YA</td>
<td>۱٢٣</td>
<td>۱٢٣</td>
</tr>
<tr>
<td>KAITHI LETTER II</td>
<td>۱٢٣</td>
<td>۱٢٣</td>
<td>KAITHI LETTER LA</td>
<td>۱٢٣</td>
<td>۱٢٣</td>
</tr>
<tr>
<td>KAITHI LETTER KA</td>
<td>۱٢٣</td>
<td>۱٢٣</td>
<td></td>
<td>۱٢٣</td>
<td>۱٢٣</td>
</tr>
</tbody>
</table>

6.2 Typefaces

There are at least three sets of Kaithi metal fonts known to the present author. Two of these sets were used by George A. Grierson for printing Kaithi specimens in the Linguistic Survey of India. Another Kaithi metal

---

\(^{70}\) Grierson, 1903b: 11. \(^{71}\) Grierson, 1899: 4. \(^{72}\) Diringer, 1968: 290.
font was developed by the Baptist Mission Press of Calcutta. The letters of this font share the fundamental structure of letters in Standard Kaithi, but the glyphs are cut in the style of Devanagari letters.

The fonts used by Grierson and by the Calcutta Bible Society possess radically different forms. Some Kaithi fonts were produced with the headstroke and others without. The Hindi bibles of the Baptist missionaries in Calcutta were generally printed in Kaithi fonts bearing the headstroke. While this practice may give the impression that the script is Devanagari, upon closer inspection it is quite clear that the letters belong distinctly to the Kaithi script. Other Kaithi fonts, such as those used by Grierson to print the specimens of the ‘Bihari’ languages in the Linguistic Survey of India, do not bear the headstroke.

The letters of this font share the fundamental structure of letters in Standard Kaithi, but the glyphs are cut in the style of Devanagari letters.

There is also stylistic variation in Kaithi metal types, of which three have been identified by the present author. Two different fonts were used by Grierson in the Linguistic Survey of India, as shown in the table of Kaithi letters in figure 15 and in the specimen of Maithili, shown in figure 13. Another used by the Calcutta Bible Society, shown in figure 24. Unfortunately, Grierson did not indicate the origin of the Kaithi font used to print specimens of the Bhojpuri, Magahi, and Maithili in the Linguistic Survey of India. It is also unclear if the types in the Linguistic Survey of India were based on those commissioned by Nesfield or if they were a new set produced by Grierson.

7 Relationship to Other Scripts

The Kaithi script is related to Devanagari, Gujarati, Bengali, and other major north Indian scripts in much the same way as the latter scripts are related to each other. It is no more similar in typology or genealogy to any other script based on northern Brahmi as are any of the other scripts derived from the same source. Modern Kaithi is descended from the Kutila script by way of the Gupta script and ultimately from Brahmi. Suniti Kumar Chatterji states that

the old Dēva-nāgarī style of the Indian alphabet which prevailed in Northern and Western India [which is the Gupta or ‘Proto-Nāgarī’ script] from the 7th century, namely, the «Kaiṭhī» script, came to Magadha by way of the Bhōjpuriyā tract; and this Kaiṭhī alphabet has held the ground till now. Kaiṭhī because of its simplicity has spread to Mithilā as well, where only the Brāhmans and other upper classes keep up the old Maithili character.\(^{73}\)

Based upon its structural characteristics and geographical location, Richard Salomon classifies Kaithi among the eastern group of scripts used for the New Indō-Aryan languages, which also includes Bengali, Maithili, and Oriya.\(^{74}\) This group is typically referred to as the Proto-Bengali or Gaudi family, which is a descendant of the Nagari form of Gupta. Genetically, Kaithi is related more to Bengali than to Devanagari. While general information about Kaithi’s genetic affiliation is ascertainable, details about its evolution from proto-Bengali to the modern standardized form are unclear.

Kaithi is considered to be the source of scripts such as Mahajani\(^{75}\) and Sylheti Nagari\(^{76}\) and quite possibly of the modern Gujarati script. Hoernle wrote that of the ‘four principal types of alphabet’ employed in northern India — “the Kāthī, the Bangālī, the Orā and the Gurmukhi” — “[t]he Kaithī is the most widely spread; it is used in writing not only in Eastern, but also, slightly modified, in Western Hindūstān, Marāṭhā and Gujarāt. In G. [Gujarati] and sometimes in E. H. [Eastern Hindustan] is it adopted also in print”.\(^{77}\) Hoernle suggests that other scripts found in the Hindi region are derived from Kaithi: “[T]here are two sub-types much in use in the area occupied by the Kaithī, to which they are the most nearly related. There are the

Kaithi is used to refer to a style of writing, similar to the terms ‘Mahājanī’, ‘Modī’, and ‘Laṇḍā’. These terms refer to particular styles of writing as well as to the formal names for distinct regional historical scripts, eg. Modi in Maharashtra, Mahajani in Rajasthan, Landa in Panjab. As terms for writing styles, these names refer to scripts used for routine purposes that were adapted for rapid writing without regard for accuracy or consistency. The terms themselves are adjectives that describe the perceived nature of the styles. As such, kaithī means ‘scribal’, mahājanī means ‘mercantile’, modī means ‘bent’, laṇḍā means ‘clipped’. Other terms such as sarrāfī ‘banker’ and vāniāī ‘mercantile’ are used to refer to further derivatives of these script styles.

These scripts are contrasted from those that are called nāgarī, which can be interpreted to mean ‘urbane’ or ‘refined’, and which are considered more formal styles. Nowadays, the term Nagari is almost synonymous with Devanagari, but Nandinagari and Jainanagari are historically distinct regional scripts that were also called Nagari. The term Nagari, however, is polysemic. Nagari also refers to the western script family that that evolved from the Kutila script, the other two being the northern-western Sharada and the eastern Gaudi, which is the ancestor of proto-Bengali, the parent of Kaithi. The scripts that descended from the Nagari script may be considered to be of the nāgarī class. Therefore, since Kaithi is descended from the Nagari branch of Kutila, it is accurate to state that the kaithī class of scripts are a subtype of the nāgarī class, or that Kaithi is a Nagari script just as Devanagari is a Nagari script.

The relationship of Kaithi to Gujarati and to Devanagari is best expressed within the framework of Kaithi as a class of scripts. Of the modern north Indian scripts descended from Brahmi, Kaithi bears the greatest resemblance to the Gujarati script. It is highly probable that Gujarati is descended from Kaithi. In fact, in the Linguistic Survey of India, Grierson writes that Kaithi “is in general use all over the north of India, from
the Gujerat coast to the river Kosi” in Bihar. To be sure, Kaithi and Gujarati are the only major north Indian scripts without the distinctive headstroke common to other scripts. They are also the major “running hands” or common scripts to be cast in metal type and used for the printing of books and other documents. It is the absence of the headstroke that is noted in The Unicode Standard, Version 5.0 as the historical link between Kaithi and Gujarati:

The Gujarati script is a North Indian script closely related to Devanagari. It is most obviously distinguished from Devanagari by not having a horizontal bar for its letterforms, a characteristic of the older Kaithi script to which Gujarati is related. The Gujarati script is used to write the Gujarati language of the Gujarat state in India.

However, Grierson’s statement that Kaithi “is in general use all over the north of India” complicates the relationship between Kaithi and Gujarati. The statement suggests that the script used in Gujarati is none other than Kaithi. To be sure, in some sources the names Kaithi and Gujarati are regarded as synonyms. The Book of a Thousand Tongues provides excerpts of bibles printed in the Gujarati language in the “Gujarati or Kaithi characters”, but also shows specimens of the “Bihar” languages as being in the “Kaithi character”. Despite the use of the same name, the actual metal types used for printing Gujarati and “Bihari” are different and are identical to the types used by Grierson for printing Gujarati and “Bihari” in the Linguistic Survey of India, as shown in Figure 30.

In Grierson’s descriptions of the relationship between Gujarati and other script scripts, he refers to different orthographic categories represented by generic script names. For instance, in 1903, Grierson wrote that Gujarati “is based on the same original as Dēva-nāgarī, and closely resembles the ordinary Kaithi character employed all over Northern India”. But in 1899 he wrote that the Gujarati script “corresponds to what is known as the Mahājani script in Upper India” and to what is “known as Vānāī or Şarrāĭ in Gujarat. However, the script formally known as ‘Mahajani’ in north India is typologically distinct from Gujarati, as shown in Figure 33. Grierson also writes that the “Moḍh, Gujarāṭí, and Kaithī alphabets ... possess a great similarity of character” while Mahajani is “still a further corruption, for cursive purposes” of Devanagari. In 1904, Grierson suggested that Mahajani and Landa are more likely derived from a different ancestor found in north-western India, related more directly to Sharada than to Devanagari.

As is evident from Grierson’s shifting classifications of Gujarati, generic terms should not be relied upon when determining the directionality or hierarchy of influence with regard to related scripts. Nor should formal names for historical scripts and names of script classes be conflated or considered as static. It is advisable to taken into consideration Hoernle’s admission that

The general likeness of these four types [the Kaithī, the Bangālī, the Oriā and the Gurmukhī] to one another as well as to the older Kutila and Gupta is unmistakeable, though their exact relation among themselves, their origin and age are matters not as yet fully elucidated.

The problem with the synonymity between Kaithi and Gujarati is determining which definition of ‘Kaithi’ is being refered to. Is it Kaithi, as the formal name of a historical script, or Kaithi, as class or family of script styles? Further analysis of Grierson proves that the term ‘Kaithi’ being used as a synonym of Gujarati is in fact Kaithi as the name of a family of scripts. Although Grierson cites the prevalence of Kaithi from Bihar to Gujarat, he states that “[t]hroughout this great tract it has of course many variations, some depending upon locality and others upon individual handwriting”. Thus, ‘Kaithi’ is a family of scripts whose regional forms developed into distinct scripts. The differentiation of the Kaithi class into regional writing systems accounts for the relationship between modern Kaithi and Gujarati and other scripts. In the eastern reaches of its geographical spread, ‘Kaithi’ retained its name; in the western periphery it assumed the name associated

---

79 Grierson, 1903b: 11.
81 American Bible Society, 1938: 159.
82 American Bible Society, 1938: 69.
83 Grierson, 1908: 338.
84 Grierson, 1899: 3.
85 Grierson, 1904b: 67.
86 Hoernle, 1975: 2.
87 Grierson, 1903b: 11.
with the language current in that region, Gujarati. Diringer concludes by stating that “Bihari Kaithi” and Gujarati “(if the latter is at all connected with the Kaithi) are essentially different”. \(^88\)

Within the Bihari sub-class are the Bhojpuri, Magahi, and Maithili forms. Of these, the “Magahi type had been adopted by the Bengal Government for official Bihari publications; books are printed in it in Patna, and the character has become more or less standardized”. \(^89\)

The differentiation of the Kaithi family into regional scripts explains the relationship between Kaithi and Syloti Nagri. James Lloyd-Williams, the author of the Syloti Nagri proposal, states that Syloti Nagri is “a form of Kaithi”. \(^90\) As such, Lloyd-Williams suggests that while Gujarati may be considered the western-most member of the Kaithi family, the distinction of the eastern-most member should go to Syloti Nagri, not the Bihari Kaithi. \(^91\) He writes that Syloti Nagri is most closely related to the Magahi style of Kaithi, however the features of the Syloti Nagri script, as well as distinct letterforms and orthographic devices, justify its status as an independent script separate from Kaithi. \(^92\)

Thus from the Kaithi family of scripts descended the Gujarati, Bihari, and Syloti regional forms. It is, therefore, reasonable to articulate that the ‘Kaithi’ script, which Grierson refered to as being “in general use” across north India, is the eponymous source whose regional styles developed into distinct historical regional scripts in Bihar, Gujrat, and Bengal. \(^93\) David Diringer supports this view when he refers to the specific development of regional styles of this the Kaithi family as “Bihari Kaithi” and Gujarati. \(^94\) He writes that “[t]he Gujarati character is essentially the literary, refined form of the script, now represented in its cursive form by the Kaithi type”. \(^95\) While Gujarati may be considered a script of the ‘Kaithi’ family or style, it is now formally known as Gujarati just as “Bihari Kaithi” is known formally as Kaithi.

But, genetic affinity is not the sole determinant of script classification and significance. Grierson writes that among this family of scripts, “Gujarati, the most western, differs little from Kaithi, the most eastern, and a Tirhutiia patwârî finds little difficulty in reading a Gujâraî book”. \(^96\) This reductive mutual intelligibility does not diminish the distiveness of a script that is determined by its cultural and sociological value. Despite the ability of Grierson’s village accountant in reading the Gujarati script with equal proficiency as Kaithi, the printing tradition associated with Gujarati embodies a particular characteristic associated within the regional Gujarati linguistic, scribal, and print traditions. \(^97\) Although Kaithi typefaces from Bihar could be used to typeset Gujarati books, the likelihood of such is parallel to the typesetting of Panjabi books using Sharada instead of Gurmukhi.

While there exists a closer relationship between Kaithi and Gujarati, many specialists associate Kaithi with Devanagari. Kaithi is often considered a corruption or cursive — parivarittra rüpa \(^98\) or ghasita śailt \(^99\) — form of Devanagari. Such characterizations are inaccurate. The relationship between the two scripts is more appropriately described as one of parallel development rather than linear descent. Grierson, the foremost authority on Kaithi, states that while Kaithi and Devanagari are related, “the two alphabets arose pari passu, from an older original still found existing in inscriptions and the like”. \(^100\) The emergence of Kaithi and De-

---

\(^88\) Diringer, 1968: 289. \(^89\) Diringer, 1968: 290. \(^90\) Lloyd-Williams, et al., 2003: 5. \(^91\) Lloyd-Williams, et al., 2003: 5. \(^92\) Lloyd-Williams, et al., 2003: 6. \(^93\) Others specialists agree with such classifications. In a table titled “Taxonomy of Alphabets and Scripts” in Thinking in Type, Alexander White poses a quaterinary hierarchy of writing systems. At the top of the hierarchy are “families of scripts” from which are successively descended “genera of scripts”, “scripts”, and finally “alphabets”. Kaithi is classified in this taxonomy as a “script” from which the Gujarati “alphabet” is derived. At the “script” level besides Kaithi are Devanagari, Bengali, and others. The Kaithi “script” belongs to the “genera” “Devanagaroid”, which belongs to the Brahmic “family” (White, 2005: viii). \(^94\) Diringer, 1968: 289–290. \(^95\) Diringer, 1968: 290. \(^96\) Grierson, 1899: 3. \(^97\) A point of interest is that while Grierson’s patwârî from Bihar could ably read Kaithi and Gujarati, a scribe from Vrindavan (in present-day western Uttar Pradesh) charged with copying into Devanagari the Kaithi manuscript of Jivâ Gosvâmî was unable to decipher the Kaithi (Mukherjee and Wright, 1979: 298ff4). Also, detailing his plan to travel to the Tirhut division of Bihar in 1915 for the purpose of taking up legal cases against the British government, Mahatma Gandhi, a native speaker of Gujarati, wrote in in his The Story of My Experiments with Truth, that “I find it difficult to understand the local dialect of Hindi, and I shall not be able to read papers written in Kaithi or Urdu.” (Gandhi, 1929: 367). \(^97\) Ojhâ, 1971: 130. \(^99\) Varmâ, 1972: 170. \(^100\) Grierson, 1899: 3.
vanagari from a common source explains the similarity of certain letter shapes, but the differences between them highlight the divergence in the development and use of the two scripts. In fact, as shown in Figure 2, Kaithi and Devanagari belong to different sub-families of Nagari.

Apart from typological differences, the fundamental distinction between Kaithi and Devanagari lies in the sphere of use. Grierson wrote that “[w]hile not so complete as the Dēva-nāgarī, for some of the rarer letters are altogether wanting, it [Kaithi] bears to that alphabet much the same relation that the English current written hand does to the printed character”.\(^{101}\) Grierson’s description suggests that Kaithi was the regular ‘cursive’ script used for routine purposes, while Devanagari was the ‘calligraphic’ script used for formal purposes. However, this does not mean that Kaithi was simply the cursive or hand-written form of Devanagari or that Devanagari is merely the formalized print version of Kaithi. The written form of Devanagari differs from Kaithi just as the printed form of Devanagari differs from the printed form of Kaithi. Through this orthographic division of labor, Kaithi was used to record the Bhojpuri, Magahi, and Maithili languages, while the Devanagari was used for Sanskrit and the formal styles of Hindi. Therefore, Kaithi is not as ‘complete’ a script as Devanagari because it was adapted for use with languages that did not possess the complex phonological features of Sanskrit and as such did not demand the preservation of such features in written form.

The scripts classified within the aforementioned categories may indeed possess similar features, such as the absence of the head-stroke, but the development of specific features among these regional styles resulted in modern writing systems that are not only typologically distinct from their historical siblings, but that are also tied to region-specific literary and cultural traditions. Grierson writes that “[t]he oldest books published in the Gujarāṭī language were printed in the Dēva-nāgarī type” and that the introduction of Gujarāṭī metal type “is a matter within the memory of the present generation”.\(^{102}\)

The Kaithi script proposed here for encoding in the UCS is the Standard Kaithi developed from the Magahi sub-type of Kaithi by the governments of Bihar and NWP&O in the 19th century. The standardization and official recognition of a script and the subsequent adoption of the script in print technology suggests that the script is an independent writing system with a distinct typology and scribal tradition. Although it was “primarily a cursive written hand”, Kaithi was “raised to the dignity of type in Bihar and Gujarat”\(^{103}\) In Gujarat, Kaithi was “elevated to the position of a national character”.\(^{104}\) In Bihar, Kaithi received the same status: “[a] fount of Kaithı type is adopted by the Bengal Government for official publications in the vernacular, intended for publication in Bihar, in which this deficiency has been supplied, and books are now printed in Patna in the same type, so that gradually the written character is becoming more correct in this respect”.\(^{105}\)

### 7.1 Comparison of Letters of Kaithi, Gujarati, and Devanagari

The differences between the standard Kaithi, Gujarati, and Devanagari scripts are evident in the typographic tradition that developed around the scripts. The differences between them are evident through a comparison of the Kaithi and Gujarati metal fonts used in the Linguistic Survey of India shown in Figure 30. Table 4 and Table 5 illustrates the differences between Kaithi, Gujarati, Devanagari, and Syloti Nagri through a comparison of the digitized fonts for each script. These comparison indicate that while several Kaithi, Gujarati, and Devanagari letterforms possess structural similarities, many are unique to the specific script. Apart from structure, the four scripts compared differ substantially in their representation and style. Thus, the similarities between the scripts owe more to reciprocal influences from contact than to unidirectionality.

Hoernle writes that “[t]hough [Kaithi] has a general resemblance to the modern Devanāgarī, there are but

---

\(^{101}\) Grierson, 1903b: 11.  \(^{102}\) Grierson, 1903b: 11.  \(^{103}\) Grierson, 1903b: 11.  \(^{104}\) Grierson, 1903b: 11.  \(^{105}\) Grierson, 1903b: 21.
few of its letters, which do not exhibit some points of difference.” He states that “all the vowels, and the consonants kh, ch, jh, bh, d, dh and r differ entirely in the two alphabets; and the horizontal top-line is omitted by the Kaithi in all letters alike”.

A comparison of the printed forms of 36 consonants of Kaithi, Gujarati, and Devanagari yields the following:

- 10 of 36 consonant letters between Kaithi and Gujarati are unique to Kaithi: ka, ja, jha, dddha, rha, nna, pha, ba, ra, la. Note: Gujarati does not have counterparts for Kaithi dddha and rha.
  8 are typologically similar, but graphically distinct: kha, nga, ca, ta, dha, na, bha, ha.
  17 are very similar: ga, gha, cha, nya, tta, tthha, dda, ddda, tha, da, pa, ma, ya, va, sha, ssa, sa.
  1 is not found in Kaithi: counterpart for Gujarati lla.

- 9 of 36 consonant letters between Kaithi and Devanagari are unique to Kaithi: kha, ja, jha, nna, da, dha, pha, ra, la.
  14 are typologically similar, but graphically distinct: ka, ga, nga, ca, ttha, ta, dha, na, ba, bha, ma, ya, va, ha.
  12 are very similar: gha, cha, nya, tta, dda, ddda, ddda, rha, pa, sha, ssa, sa.
  1 is not found in Kaithi: counterpart for Devanagari lla.

- 10 of 36 consonant letters between Gujarati and Devanagari are unique to Gujarati: ka, kha, ja, jha, nna, da, pha, ba, la, and lla.
  4 are typologically similar, but graphically distinct: ca, bha, ra, ha.
  20 are very similar: ga, gha, nga, cha, nya, tta, ttha, dda, ddda, ta, tha, dha, na, pa, ma, ya, va, sha, ssa, sa.
  2 are not found in Gujarati: counterparts for Devanagari ddda or rha.

A comparison of the printed forms of 11 independent vowels of the three scripts shows the following:

- 9 of 11 vowels between Kaithi and Gujarati are distinctive in Kaithi: a, aa, i, ii, uu, e, ai, o, and au.
  1 is typologically similar, but graphically distinct: u.
  1 is not found in Kaithi: counterpart for Gujarati vocalic r.

- 6 of 11 vowels between Kaithi and Devanagari are distinctive in Kaithi: a, aa, i, ii, o, and au.
  4 are very similar: u, uu, e, ai.
  1 is not found in Kaithi: counterpart for Devanagari vocalic r.

- 8 of 11 vowels between Gujarati and Devanagari are distinctive in Gujarati: a, aa, i, ii, e, ai, o, and au.

---

106 Hoernle, 1880: 1. 107 Hoernle, 1880: 1.
2 are typologically similar, but graphically distinct: \( \text{u} \) and \( \text{uu} \).

1 is very similar: vocalic \( \text{r} \).

Some letters in Kaithi and Gujarati have similar appearance, but different semantic value. For instance, \( \text{v} \) Kaithi letter \( \text{ja} \) resembles \( \text{v} \) U+0AB3 Gujarati letter \( \text{lla} \). Kaithi does not have a letter for \( \text{lla} \). Grierson\textsuperscript{108} shows a form of the consonant-vowel ligature for \( \text{hr} \) as \( \text{hr} \). This ligature is identical in shape to \( \text{hr} \) Kaithi letter \( \text{jha} \). This ligature would be written as \( \text{hr} \) \( \text{hari} \) in Kaithi.

\textsuperscript{108} Grierson, 1908: 338.
<table>
<thead>
<tr>
<th>KAITHI</th>
<th>GUJARATI</th>
<th>DEVANAGARI</th>
<th>SYLOTI</th>
<th>KAITHI</th>
<th>GUJARATI</th>
<th>DEVANAGARI</th>
<th>SYLOTI</th>
</tr>
</thead>
<tbody>
<tr>
<td>ka</td>
<td>क</td>
<td>क</td>
<td>पा</td>
<td>tha</td>
<td>ठ</td>
<td>ठ</td>
<td>न</td>
</tr>
<tr>
<td>kha</td>
<td>क्ष</td>
<td>ख</td>
<td>न</td>
<td>da</td>
<td>ध</td>
<td>ध</td>
<td>ज</td>
</tr>
<tr>
<td>ga</td>
<td>ग</td>
<td>ग</td>
<td>ग</td>
<td>dha</td>
<td>ध</td>
<td>ध</td>
<td>ज</td>
</tr>
<tr>
<td>gha</td>
<td>घ</td>
<td>घ</td>
<td>ब</td>
<td>na</td>
<td>न</td>
<td>न</td>
<td>ग</td>
</tr>
<tr>
<td>ŋa</td>
<td>ो</td>
<td>ड</td>
<td>—</td>
<td>pa</td>
<td>प</td>
<td>प</td>
<td>ज</td>
</tr>
<tr>
<td>ca</td>
<td>च</td>
<td>च</td>
<td>ब</td>
<td>pha</td>
<td>फ</td>
<td>फ</td>
<td>क</td>
</tr>
<tr>
<td>cha</td>
<td>छ</td>
<td>छ</td>
<td>फ़</td>
<td>ba</td>
<td>ब</td>
<td>ब</td>
<td>व</td>
</tr>
<tr>
<td>ja</td>
<td>ज</td>
<td>ज</td>
<td>ल</td>
<td>bha</td>
<td>भ</td>
<td>भ</td>
<td>र</td>
</tr>
<tr>
<td>jha</td>
<td>झ</td>
<td>झ</td>
<td>म़</td>
<td>ma</td>
<td>म</td>
<td>म</td>
<td>ल</td>
</tr>
<tr>
<td>ña</td>
<td>न</td>
<td>न</td>
<td>—</td>
<td>ya</td>
<td>य</td>
<td>य</td>
<td>—</td>
</tr>
<tr>
<td>ṭa</td>
<td>ट</td>
<td>ट</td>
<td>ठ</td>
<td>ra</td>
<td>र</td>
<td>र</td>
<td>न</td>
</tr>
<tr>
<td>֟ha</td>
<td>ۂ</td>
<td>ठ</td>
<td>छ</td>
<td>la</td>
<td>ल</td>
<td>ल</td>
<td>त</td>
</tr>
<tr>
<td>֠da</td>
<td>Ԁ</td>
<td>ड</td>
<td>उ</td>
<td>la</td>
<td>ल</td>
<td>ल</td>
<td>ठ</td>
</tr>
<tr>
<td>ṛa</td>
<td>ڑ</td>
<td>ढ</td>
<td>—</td>
<td>va</td>
<td>व</td>
<td>व</td>
<td>न</td>
</tr>
<tr>
<td>ḍha</td>
<td>ढ</td>
<td>ढ</td>
<td>छ</td>
<td>śa</td>
<td>झ</td>
<td>झ</td>
<td>झ</td>
</tr>
<tr>
<td>ṛha</td>
<td>ɽ</td>
<td>ढ</td>
<td>—</td>
<td>śa</td>
<td>झ</td>
<td>झ</td>
<td>झ</td>
</tr>
<tr>
<td>ṃa</td>
<td>$smarty�</td>
<td>ण</td>
<td>—</td>
<td>sa</td>
<td>स</td>
<td>स</td>
<td>—</td>
</tr>
<tr>
<td>ṭa</td>
<td>ठ</td>
<td>ठ</td>
<td>ठ</td>
<td>ha</td>
<td>ह</td>
<td>ह</td>
<td>प</td>
</tr>
</tbody>
</table>

Table 4: A comparison of the consonant letters of the Kaithi, Gujarati, Devanagari, and Syloti Nagri scripts
Table 5: A comparison of vowel letters and signs of the Kaithi, Gujarati, Devanagari, and Syloti Nagri scripts

<table>
<thead>
<tr>
<th>KAITHI</th>
<th>GUJARATI</th>
<th>DEVANAGARI</th>
<th>SYLOTI NAGRI</th>
<th>KAITHI</th>
<th>GUJARATI</th>
<th>DEVANAGARI</th>
<th>SYLOTI NAGRI</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>अ</td>
<td>आ</td>
<td>—</td>
<td>-a</td>
<td>ऑ</td>
<td>ओ</td>
<td>—</td>
</tr>
<tr>
<td>ñ</td>
<td>अं</td>
<td>आं</td>
<td>ता</td>
<td>-ña</td>
<td>ओं</td>
<td>ओं</td>
<td>ओं</td>
</tr>
<tr>
<td>i</td>
<td>ई</td>
<td>इ</td>
<td>इ</td>
<td>-i</td>
<td>ई</td>
<td>ई</td>
<td>ई</td>
</tr>
<tr>
<td>ñ</td>
<td>ईं</td>
<td>इं</td>
<td>इं</td>
<td>-iä</td>
<td>ओं</td>
<td>ओं</td>
<td>ओं</td>
</tr>
<tr>
<td>u</td>
<td>उ</td>
<td>ऊ</td>
<td>—</td>
<td>-u</td>
<td>उ</td>
<td>उ</td>
<td>—</td>
</tr>
<tr>
<td>ñu</td>
<td>उं</td>
<td>ऊं</td>
<td>—</td>
<td>-nü</td>
<td>उं</td>
<td>उं</td>
<td>—</td>
</tr>
<tr>
<td>r</td>
<td>र</td>
<td>र्ळ</td>
<td>—</td>
<td>-r</td>
<td>र</td>
<td>र</td>
<td>—</td>
</tr>
<tr>
<td>e</td>
<td>ए</td>
<td>एं</td>
<td>—</td>
<td>-e</td>
<td>ए</td>
<td>ए</td>
<td>—</td>
</tr>
<tr>
<td>ai</td>
<td>ऐ</td>
<td>ऐं</td>
<td>—</td>
<td>-ai</td>
<td>ऐ</td>
<td>ऐ</td>
<td>—</td>
</tr>
<tr>
<td>o</td>
<td>औ</td>
<td>औं</td>
<td>ऊं</td>
<td>-o</td>
<td>औ</td>
<td>औ</td>
<td>औ</td>
</tr>
<tr>
<td>au</td>
<td>ओू</td>
<td>ओूं</td>
<td>ऊूं</td>
<td>-au</td>
<td>ओू</td>
<td>ओू</td>
<td>ओू</td>
</tr>
</tbody>
</table>

Table 6: A comparison of digits of the Kaithi, Gujarati, Devanagari, and Syloti Nagri scripts. Note: Syloti Nagri uses Bengali digits.

<table>
<thead>
<tr>
<th>KAITHI</th>
<th>GUJARATI</th>
<th>DEVANAGARI</th>
<th>SYLOTI NAGRI</th>
<th>KAITHI</th>
<th>GUJARATI</th>
<th>DEVANAGARI</th>
<th>SYLOTI NAGRI</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>०</td>
<td>०</td>
<td>०</td>
<td>५</td>
<td>५</td>
<td>५</td>
<td>५</td>
</tr>
<tr>
<td>1</td>
<td>१</td>
<td>१</td>
<td>१</td>
<td>६</td>
<td>६</td>
<td>६</td>
<td>६</td>
</tr>
<tr>
<td>2</td>
<td>२</td>
<td>२</td>
<td>२</td>
<td>७</td>
<td>७</td>
<td>७</td>
<td>७</td>
</tr>
<tr>
<td>3</td>
<td>३</td>
<td>३</td>
<td>३</td>
<td>८</td>
<td>८</td>
<td>८</td>
<td>८</td>
</tr>
<tr>
<td>4</td>
<td>४</td>
<td>४</td>
<td>४</td>
<td>९</td>
<td>९</td>
<td>९</td>
<td>९</td>
</tr>
</tbody>
</table>
8 References


Proposal to Encode the Kaithi Script in Plane 1 of ISO/IEC 10646

Anshuman Pandey

Printing, India.


महागणपतिस्तोत्र [Mahāganapatistotra] (anonymous). (undated). Miscellaneous stotra materials in the Devanagari and Kaithi scripts. Manuscript held at University of Pennsylvania (Poleman number 1876; University of Pennsylvania number 2584).


that there is only one sign for each of the following groups of Nāgarī letters: 1) ण (properly = n) for the nasals ठ न, ढ न, ण न, तै न; 2) श (a combination of s and ꞧ) for the sibilants स s, झ ꞧ, झ ꞧ h; 3) थ for the labials थ b and ध v; 4) ठ (properly = j) for the palatals ठ j and ठ y; and also that of the two forms of चळ one is very much like to one of the two forms of दळ, the other to one of the two forms of य. For the vowels Kaithi has only four fundamental signs: ऐ ऐ, ए ए, उ उ, ओ ओ. The others
Table showing the Kaithi alphabet, as written by Tirhuti, Bhojpuri & Magahi Scribes.

<table>
<thead>
<tr>
<th>Tirhuti</th>
<th>Bhojpuri</th>
<th>Magahi</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>d</td>
<td>d</td>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>i</td>
<td>i</td>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td>u</td>
<td>u</td>
<td>u</td>
<td>u</td>
</tr>
<tr>
<td>r</td>
<td>r</td>
<td>r</td>
<td>r</td>
</tr>
<tr>
<td>ñ</td>
<td>ñ</td>
<td>ñ</td>
<td>ñ</td>
</tr>
<tr>
<td>g</td>
<td>g</td>
<td>g</td>
<td>g</td>
</tr>
<tr>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
</tr>
<tr>
<td>k</td>
<td>k</td>
<td>k</td>
<td>k</td>
</tr>
<tr>
<td>jh</td>
<td>jh</td>
<td>jh</td>
<td>jh</td>
</tr>
<tr>
<td>j</td>
<td>j</td>
<td>j</td>
<td>j</td>
</tr>
<tr>
<td>t</td>
<td>t</td>
<td>t</td>
<td>t</td>
</tr>
<tr>
<td>th</td>
<td>th</td>
<td>th</td>
<td>th</td>
</tr>
<tr>
<td>d</td>
<td>d</td>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>dh</td>
<td>dh</td>
<td>dh</td>
<td>dh</td>
</tr>
<tr>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td>p</td>
<td>p</td>
<td>p</td>
<td>p</td>
</tr>
<tr>
<td>ph</td>
<td>ph</td>
<td>ph</td>
<td>ph</td>
</tr>
<tr>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>bh</td>
<td>bh</td>
<td>bh</td>
<td>bh</td>
</tr>
<tr>
<td>m</td>
<td>m</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td>y</td>
<td>y</td>
<td>y</td>
<td>y</td>
</tr>
<tr>
<td>j</td>
<td>j</td>
<td>j</td>
<td>j</td>
</tr>
<tr>
<td>r</td>
<td>r</td>
<td>r</td>
<td>r</td>
</tr>
<tr>
<td>l</td>
<td>l</td>
<td>l</td>
<td>l</td>
</tr>
<tr>
<td>v</td>
<td>v</td>
<td>v</td>
<td>v</td>
</tr>
<tr>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>s</td>
<td>s</td>
<td>s</td>
<td>s</td>
</tr>
<tr>
<td>kh</td>
<td>kh</td>
<td>kh</td>
<td>kh</td>
</tr>
</tbody>
</table>

Figure 3: A comparison of the three regional forms of Kaithi, eg. the Tirhuti (Maithili), Magahi, and Bhojpuri (from Grierson, 1899: Plate II).
Combination of consonants

Strong conjuncts.

\[ \begin{align*}
\text{kk}, & \quad \text{kk}, \quad \text{kkh}, \quad \text{kk}, \quad \text{kt}, \quad \text{gdh}, \quad \text{ghch}, \quad \text{ghch}, \quad \text{ghch}, \quad \text{ij}, \quad \text{ij}, \quad \text{tt}, \quad \text{th}, \\
\text{dg}, & \quad \text{dd}, \quad \text{tk}, \quad \text{tt}, \quad \text{tt}, \quad \text{tp}, \quad \text{dg}, \quad \text{dd}, \quad \text{ddh}, \quad \text{ddh}, \quad \text{pt}, \quad \text{pp}, \quad \text{pph}, \\
\text{bj}, & \quad \text{bd}, \quad \text{bdh}, \quad \text{bd}, \quad \text{bb}.
\end{align*} \]

Weak conjuncts.

\[ \begin{align*}
\text{nn}, & \quad \text{ny}, \quad \text{nr}, \quad \text{nn}, \quad \text{nm}, \quad \text{ny}, \quad \text{nr}, \quad \text{m}, \quad \text{mn}, \quad \text{ny}, \quad \text{nr}, \\
\text{ml}, & \quad \text{mk}, \quad \text{yy}, \quad \text{rm}, \quad \text{ry}, \quad \text{f}, \quad \text{fr}, \quad \text{frh}, \quad \text{fh}, \quad \text{im}, \quad \text{ly}, \quad \text{li}, \quad \text{lh}, \quad \text{ly}, \\
\text{r}, & \quad \text{v}, \quad \text{r}, \quad \text{s}, \quad \text{r}, \quad \text{f}, \quad \text{fr}, \quad \text{frh}, \quad \text{frh}, \quad \text{sh}, \quad \text{sh}, \quad \text{sh}, \quad \text{sh}, \quad \text{s}, \quad \text{sh}, \quad \text{sh}, \\
\text{sr}, & \quad \text{sv}, \quad \text{ss}, \quad \text{hr}, \quad \text{hr}, \quad \text{hrh}, \quad \text{hr}, \quad \text{hrh}.
\end{align*} \]

Mixed conjuncts.

\[ \begin{align*}
\text{k}, & \quad \text{km}, \quad \text{ky}, \quad \text{kr}, \quad \text{kl}, \quad \text{kv}, \quad \text{kh}, \quad \text{akh}, \quad \text{ak}, \quad \text{gn}, \quad \text{gm}, \quad \text{gy}, \quad \text{gr}, \quad \text{gl}, \quad \text{gl}, \\
\text{g}, & \quad \text{gh}, \quad \text{g}, \quad \text{gh}, \quad \text{g}, \quad \text{gh}, \quad \text{g}, \quad \text{gh}, \quad \text{g}, \quad \text{gh}, \quad \text{g}, \quad \text{gh}, \quad \text{g}, \quad \text{gh}, \quad \text{g}, \quad \text{gh}, \\
\text{h}, & \quad \text{h}, \quad \text{h}, \quad \text{h}, \quad \text{h}, \quad \text{h}, \quad \text{h}, \quad \text{h}, \quad \text{h}, \quad \text{h}, \quad \text{h}, \quad \text{h}, \quad \text{h}, \quad \text{h}, \quad \text{h}, \quad \text{h}, \\
\text{y}, & \quad \text{y}, \quad \text{y}, \quad \text{y}, \quad \text{y}, \quad \text{y}, \quad \text{y}, \quad \text{y}, \quad \text{y}, \quad \text{y}, \quad \text{y}, \quad \text{y}, \quad \text{y}, \quad \text{y}, \quad \text{y}, \quad \text{y}, \\
\text{t}, & \quad \text{t}, \quad \text{t}, \quad \text{t}, \quad \text{t}, \quad \text{t}, \quad \text{t}, \quad \text{t}, \quad \text{t}, \quad \text{t}, \quad \text{t}, \quad \text{t}, \quad \text{t}, \quad \text{t}, \quad \text{t}, \quad \text{t}.
\end{align*} \]

*Pronounced like (y) with a nasal accent.*

Figure 4: A list of Kaithi conjuncts used in the Maithili (Tirhuti) style of Kaithi. These forms rarely appear in the Magahi or Bhojpuri styles (from Grierson, 1899: Plate III).
Figure 5: Currency, weights, and measures signs that appear in Kaithi documents (from Grierson, 1899: Plate IV). These signs are proposed for inclusion in the UCS in a separate proposal.
Figure 6: Specimen of hand-written Bhojpuri style of Kaithi (from Grierson, 1899: Plate XXVIII).
Figure 7: Specimen of hand-written Maithili style of Kaithi (from Grierson, 1899: Plate X).
Figure 8: Specimen of hand-written Magahi style of Kaithi (from Grierson, 1899: Plate XXVII).
Figure 9: Excerpt from a specimen of Maithili written in the Magahi style of Kaithi (from Grierson, 1903b: 82).
Figure 10: Specimen of Awadhi (from Grierson, 1904a: 51) written in what Grierson called “a sort of mixture of Dēvānāgārī and Kaithī,” which was “current in the District amongst the educated classes” (from Grierson, 1904a: 49)
Figure 11: A specimen of the form of Bengali spoken in the Purnea region of Bihar written in the Kaithi script (from Grierson, 1903a: 140).
Figure 12: A specimen of Magahi printed in Kaithi type (from Grierson, 1903b: 124).
BIHARI.

MAITHILI DIALECT (AS USED BY HINDUS OF THE LOWER CASTES).

(DARBHANGA DISTRICT.)

Figure 13: A specimen of Maithili printed in Kaithi type (from Grierson, 1903b: 74).
Proposal to Encode the Kaithi Script in Plane 1 of ISO/IEC 10646

Anshuman Pandey

BIHÁRÍ.

Bhojpúrí Dialect.

Western Sub-dialect. (District Azamgarh.)

SPECIMEN I.

(Babu Rama Smaran Lal, 1898.)

Figure 14: A specimen of Bhojpuri printed in Kaithi type (from Grierson, 1903b: 253).
KAITHI OR KAYATHI ALPHABET.

**Vowels.**

<table>
<thead>
<tr>
<th>a</th>
<th>A</th>
<th>Æ</th>
<th>i</th>
<th>ï</th>
<th>u</th>
<th>Ù</th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>E</td>
<td>Æ</td>
<td>ə</td>
<td>ə</td>
<td>a</td>
<td>əh</td>
</tr>
</tbody>
</table>

**Consonants.**

- **Gutturals.**  ka  kha  g  ga  gha
- **Palatals.**  cha  chha  ja  jha
- **Cerebrals.**  ta  tha  qa  qha
- **Dentals.**  da  tha  da  th̄a  na
- **Labials.**  pa  ph  ba  bha  ma
- **Semi-Vowels.**  ya  ra  la  va
- **Sibilants.**  sa  sha  sa
- **Aspirate.**  ha

In a, being inherent in each consonant, is only written when initial in a word or syllable; thus, we write आ ap, ओ tu, but अ pa, ए ta. The other vowels, when following a consonant, are substituted for the inherent ए a, and, in this case, they take the following forms:

- ए a (not expressed); ए̄ ə;
- ə ə;
- ə̄ ə;
- ai a;
- o o;
- au u.

Thus, the several vowel sounds, when they follow consonants, are written as follows:

<table>
<thead>
<tr>
<th>ka</th>
<th>kā, ki, kī, ku, kū, kē, kai</th>
</tr>
</thead>
</table>

The vowel mark ऊ is called anusvāra as in Deva-nagari. It denotes the nasalization of a preceding vowel, and can therefore never begin a syllable. It is written over, or to the right of the preceding vowel: as, आ am, आँ bāñ. It is used for both the anumāsik and the anusvāra of Deva-nagari. The mark is called bisārg, and indicates a weak aspiration. It is only found in pure Sanskrit words, and even then, though occurring in the original, is commonly omitted in Hindi; as, दुः dukkha written and pronounced दु दुः dukkha.

Figure 15: A table showing the characters of the Kaithi script (from Grierson, 1903b: 12).
Figure 16: Inventory of Kaithi letters (from Śākyavāṃśa, 1974: 64)

Figure 17: Comparison of numerals of Kaithi and other scripts (from Śākyavāṃśa, 1974: 76)
Folio 1b: Sanskrit in Devanagari script

Folio 2a: Maithili style of Kaithi

Figure 18: Folios 1b and 2a from the Mahāgaṇapatiṣṭotra written in Devanagari and Kaithi (continued in Figure 19). The reproductions of these folios are used with permission from the University of Pennsylvania.
Folio 1a: Invocatory text in Devanagari (lines 1-2) and Kaithi (lines 3-4).

Folio 4a: Text in Kaithi and Devanagari. This folio contains two styles of Kaithi. Lines 1 and 2 are written in the Maithili style; lines 3–7 are in the Bhojpuri style.

Figure 19: Folios 1a and 4a from the Mahāganapatiṣṭotra written in Devanagari and Kaithi (continued from Figure 18). The reproductions of these folios are used with permission from the University of Pennsylvania.
Figure 20: Excerpt from a plaint from the district court of Patna, Bihar hand-written in Kaithi (from Bihar High Court of Judicature, 1939).
Figure 21: Excerpt from a plaint from the district court of Bhagalpur, Bihar hand-written in Kaithi (from Bihar High Court of Judicature, 1939).
Figure 22: Excerpt from a statement from the district court of Ranchi, Bihar hand-written in Kaithi (from Bihar High Court of Judicature, 1939).
Figure 23: A rent receipt granted by the Pirpattidar of Dugni (Principality of Seraikella) written in Kaithi on a form printed in Devanagari (from Government of Bihar, 1954: plate following p.288).
Figure 24: The title, first, and second pages of the Book of Genesis printed in Kaithi type (from Calcutta Bible Society, 1851). The Kaithi font used here resembles Devanagari in the use of the headstroke, but distinct Kaithi letters can be identified.
Figure 25: The English title, Hindi title, and first page of the Hindi translation of the New Testament in Kaithi type (from Bible Translation Society, 1850). The Kaithi font used here is similar to that shown in Figure 24; it resembles Devanagari in the use of the headstroke, but distinct Kaithi letters can be identified. Note, in particular, the use of काइथी लेटर न्या in the word न्या, which appears in last word of the fifth sentence on the Kaithi title page.
Figure 26: Entries for the 'Bihari' languages in *The Book of a Thousand Tongues* showing specimens from bibles published in Kaithi and Devanagari type (from American Bible Society, 1938: 69). The Kaithi font used here is identical to that used in the *Linguistic Survey of India*. 

---

**BICOL**

18. Ḥuli ta an Dios namoot na gayo sa kinga-nan, na itinu na an saiyang Aking bogton, tagánaning an siisay man na minatubod sa siyarna, dai mapahanam, kundi magkaigwa nau bhay na dai na katapusan.

17. Ḥuli ta an Dios dai nagoso kon saiyang Aki-sa kingana na magpasol kon kingana; kundi tagánaning an kingana italigtaas huli siyarna.


9. Assin ini iyo an silcot, na an ilao napadigi sa kingana, asin an maga na tuwo namoot pang labi sa diikom kii sa liwanag; huli ta an saindang magña guibo marao.

20. Ḥuli ta an silcot man na nagguiguibo ni marao naogha sa liwanag, asin an minadokol sa ilao, tagánaning an saiyang magña guibo dai magkasaragad.


**BIHARI:** Magahi dialect

**BIHARI:** Bhojpuri dialect

**BIHARI:** Kortha dialect

**BIHARI:** Nagpuri dialect

---

**BIHARI:** Maithili dialect

---

**BIHARI:** Magahi dialect

**BIHARI:** Bhojpuri dialect

**BIHARI:** Kortha dialect

**BIHARI:** Nagpuri dialect
Figure 27: A folio from the "Ekadālā" manuscript of Miragāvati c.1828 (from Miśra, 1963: plate 2).
Figure 28: A folio from the Tale of Sudama, India, Bikaner, 1745-6 CE, No. 9028, Sam Fogg, London. Image © Sam Fogg, London
Figure 29: A letter to the Supreme Civil Court of Appeals in Calcutta. The letter is written in the high Persian idiom and šíkastā style common in courts of law. The seal at the top contains text in the Perso-Arabic (first two lines), Bengali (middle two lines), and Kaithi (bottom two lines) scripts. The Kaithi text reads “meye al adālat dīvānī šādar 1850” (mohar adālat dīvānī šādar 1850) is a transliteration of the Urdu “مہر عدالت دوےآئی صدر 1850,” meaning “the seal of the Supreme Civil Court of Appeals” (from Stewart, 1825: plate 12, p.54–55).
Figure 30: Comparison of Kaithi (from Grierson, 1903b: 124), Gujarati (from Grierson, 1908: 365), and Devanagari (from Grierson, 1916: 95) types from the Linguistic Survey of India.
Figure 31: Comparison of hand-written Kaithi (from Ojhā, 1971: Plate LXXVIII) and Gujarati letters (from Ojhā, 1971: Plate LXXIX).

| TABLE 13.1: KAITHI OR KĀYATHI SCRIPT WITH DEVANĀGARI EQUIVALENTS AND TRANSLITERATIONS |
|-----------------|-----------------|-----------------|-----------------|
| Vowels          | a               | o               | e               |
| ꤃               | ꤅               | ꤖ               | ꤘ               |
| ꤙ               | ꤛ               | ꤜ               | ꤝ               |
| ꤞ               | ꤟ               | ꤠ               | ꤡ               |
| ꤢ               | ꤣ               | ꤤ               | ꤥ               |
| ꤦ               | ꤧ               | ꤨ               | ꤩ               |
| ꤪ               | ꤫               | ꤬               | ꤭               |
| ꤮               | ꤯               | ꤰ               | ꤱ               |
| ꤲ               | ꤳ               | ꤴ               | ꤵ               |
| ꤶ               | ꤷ               | ꤸ               | ꤹ               |
| Consonants      | ꤾ               | ꤿ               | ꥀ               |
| Stops           | ꥁ               | ꥂ               | ꥃ               |
| Velar           | ꥄ               | ꥅ               | ꥆ               |
| Palatal         | ꥇ               | ꥈ               | ꥉ               |
| Cerebral        | ꥊ               | ꥋ               | ꥌ               |
| Dental          | ꥍ               | ꥎ               | ꥏ               |
| Labial          | ꥐ               | ꥑ               | ꥒ               |
| Semivowels      | ꥓               | ꥔               | ꥕               |
| Spirants        | ꥖               | ꥗               | ꥘               |
| Others          | ꥙               | ꥚               | ꥛               |
| Anusvāra        | ꥜               | ꥝               | ꥞               |
| Visarga         | ꥘               | ꥙               | ꥚               |

Figure 32: Comparison of Kaithi and Devanagari (from S. Verma, 2003: 502).
**Figure 33:** A comparison of the Kaithi script with the Devanagari and Mahajani (from Grierson, 1899: Plate I).
Figure 34: A table from *Kaithi vo hindi barnamālā*, a Kaithi script primer, showing the Kaithi and Devanagari scripts in parallel. The Kaithi letters in this primer are drawn in conformity to the Devanagari style, which is most noticeable in the presence of the headstroke. Note the difference in the headstroke in Devanagari and Kaithi (from *Kaithi vo hindi barnamālā*, 1882: 2).
Figure 35: Comparison of writing techniques in Kaithi and Devanagari (from *Kaiith vo hindt barnamālā*, 1882: 8).
Figure 36: Comparison of scripts descended from proto-Bengali (from Jensen, 1969: 370).
Figure 37: Comparison of Kaithi with other scripts used for writing Hindi (from Kellogg, 1893: 26–27).
Proposal to Encode the Kaithi Script in Plane 1 of ISO/IEC 10646

Anshuman Pande

Figure 38: Comparison of Kaithi with other Indic scripts (from Mule, 1974: 163-165).
Figure 39: Comparison of Kaithi with other Indic scripts (from Naik, 1971: Table 13).
Figure 40: A family tree of north Indian scripts showing Kaithi as a branch of Nagari (from Singh, 1991: 16).

Figure 41: The position of the Kaithi script with regard to others (from Naik, 1971: Plate 36).