

# Revised Proposal to Encode the Bhaiksuki Script in ISO/IEC 10646

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

Dragomir Dimitrov  
Fachgebiet Indologie und Tibetologie  
Philipps-Universität Marburg  
Marburg, Germany  
dimitrov@staff.uni-marburg.de

October 27, 2013

## 1 Introduction

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

- N4121 L2/11-259 “Preliminary Proposal to Encode the Bhaiksuki Script in ISO/IEC 10646”
- N4469 L2/13-167 “Proposal to Encode the Bhaiksuki Script in ISO/IEC 10646”

This document is a revision of N4469 L2/13-167, which was submitted to the Unicode Technical Committee for review at the August 2013 meeting. The major change is a revision of the encoding model for consonant-virama ligatures (see Section 3.7). Minor changes include corrections to character properties.

## 2 Background

Bhaiksuki (𑂔𑂗𑂢𑂰 *bhaikṣukī*; Devanagari भैक्षुकी) is a Brahmi-based script that was used around the turn of the first millenium CE mainly in the present-day states of Bihar and West Bengal in India, as well as in regions that are now part of Bangladesh. Records have been also located in Tibet, Nepal, and Burma. The script is known variously as the ‘Arrow-Headed Script’ or ‘Point-Headed Script’ in English, ‘Pfeilspitzenschrift’ in German, and ‘Śaramāṭṛkā Lipī’ in Hindi and modern Sanskrit. An older designation, ‘Sindhura’, has been used in Tibet for at least three centuries.

The script is attested exclusively in Buddhist textual materials. Only eleven inscriptions and four manuscripts written in this script are presently known to exist. These are the Bhaiksuki manuscripts of the *Abhidharmasamuccayakārikā*, *Mañicūḍajātaka*, *Candrālamkāra*, and at least one more Buddhist canonical text. The codex of the *Abhidharmasamuccayakārikā* was once kept in Tibet, but it is now inaccessible and its exact place of preservation is unknown. The fourth codex was discovered in Tibet and was recently shown in a Chinese documentary; however, information about this manuscript is limited. It is likely that additional materials in Bhaiksuki may become available in the future.

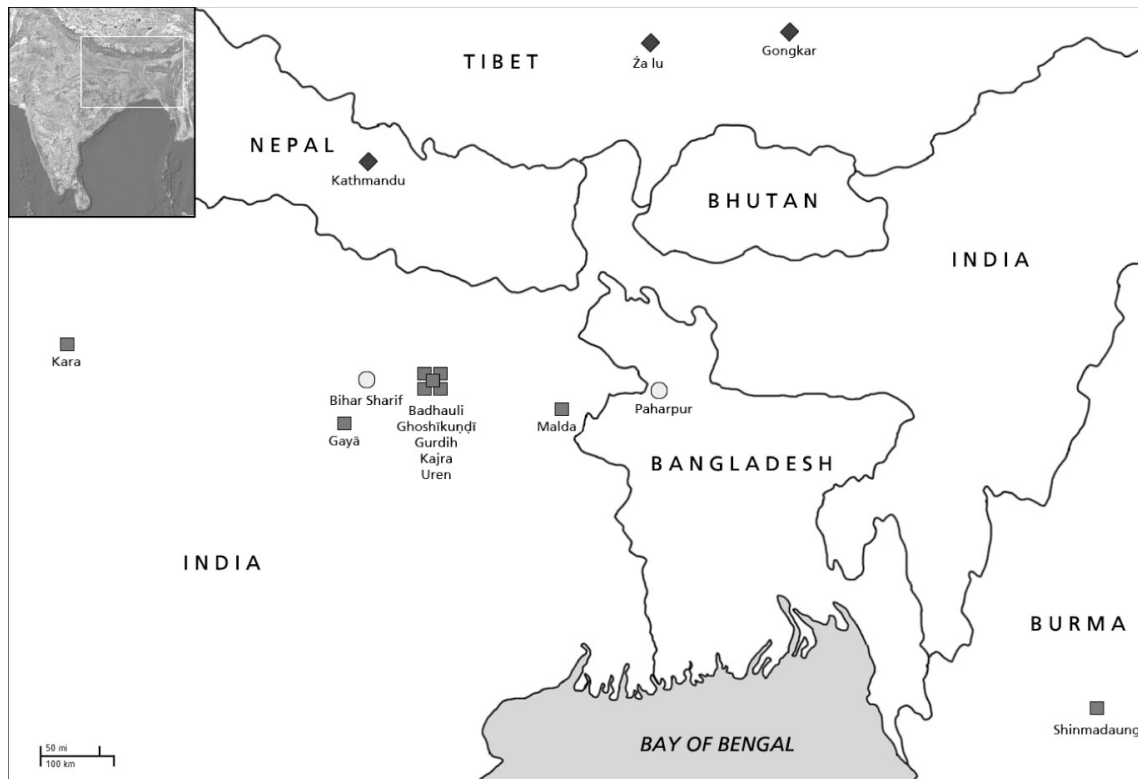


Figure 1: Sites of inscriptions (■) and manuscripts (◆) and other places (○) where Bhaiksuki has reportedly been used (from Dimitrov 2010: 52).

There has been scholarly interest in Bhaiksuki from the time that Cecil Bendall presented the script to Western academic communities in the 1880s. In the 1890s, Bruno Liebich made further advances through his study of the materials available at the time. More recently, the Bhaiksuki manuscript of the *Maṇicūḍajātaka* was studied by Albrecht Hanisch in 2008 and the manuscript of the *Candrālamkāra* was presented by Dragomir Dimitrov in 2010. Since 2004 the Arrow-headed Script Project at Philipps-Universität, Marburg, Germany has been engaged in research on Bhaiksuki and in developing resources for the script.

### 3 Script Details

#### 3.1 Name

The transliterated name of the script, *bhaikṣukī*, is normalized as ‘Bhaiksuki’. The proposed block consists of 94 characters, which is tentatively allocated to the SMP at the range U+11C00..11C6F. Character names are aligned with those used in other script encoded in the UCS. A code chart and names list is attached.

#### 3.2 Structure

The general structure (phonetic order, *mātrā* reordering, use of *virāma*, etc.) of Bhaiksuki is similar to that of Devanagari. Some dependent vowel signs consist of two or three parts and behave like those in Bengali. Several vowel signs have contextual forms when they combine with certain consonants, and certain

consonant-vowel sequences are written as ligatures. In some cases, consonant + *virāma* pairs are rendered using both a visible *virāma* and a special ligature. Consonant clusters are represented as conjuncts.

### 3.3 Virāma

The ◌ VIRAMA is used for indicating the absence of the inherent vowel in a consonant letter. It is identical in function to the VIRAMA in Devanagari. Certain pairs of consonant + VIRAMA are rendered as both with visible VIRAMA and as a special ligature (see Section 3.7).

### 3.4 Vowels

There are 13 vowel letters:

अ	A	उ	U	ए	VOCALIC L	ऊ	AU
आ	AA	ऊ	UU	इ	E		
इ	I	उ	VOCALIC R	अ	AI		
ई	II	उ	VOCALIC RR	उ	O		

The vowel letter \*VOCALIC LL is not attested, but space has been reserved for it.

### 3.5 Vowel Signs

There are 12 dependent vowel signs:

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

The \*VOWEL SIGN VOCALIC LL is not attested, but space has been reserved for it. Several signs are written using alternate forms when they occur with certain consonants (see Section 3.8).

### 3.6 Consonants

There are 33 consonant letters:

𑂀	KA	𑂁	NYA	𑂂	DHA	𑂃	LA
𑂄	KHA	𑂅	TTA	𑂆	NA	𑂇	VA
𑂈	GA	𑂉	TTHA	𑂊	PA	𑂋	SHA
𑂌	GHA	𑂍	DDA	𑂎	PHA	𑂏	SSA
𑂐	NGA	𑂑	DDHA	𑂒	BA	𑂓	SA
𑂔	CA	𑂕	NNA	𑂖	BHA	𑂗	HA
𑂘	CHA	𑂙	TA	𑂚	MA		
𑂛	JA	𑂜	THA	𑂝	YA		
𑂞	JHA	𑂟	DA	𑂠	RA		

Each consonant bears the inherent vowel /a/, represented by 𑂡 A, which is silenced using 𑂢 VIRAMA. Consonant clusters are written as conjuncts (see Section 3.9).

### 3.7 Consonant-Virama Ligatures

Sequences of <C, VIRAMA> are rendered by default using a visible VIRAMA. However, three combinations are also represented using a special ligature, which is referred to here as a “*khaṇḍa*” form:

	combining <i>virāma</i>	<i>khaṇḍa</i> ligature
<𑂙 TA, 𑂢 VIRAMA>	𑂙	𑂙
<𑂛 NA, 𑂢 VIRAMA>	𑂛	𑂛
<𑂚 MA, 𑂢 VIRAMA>	𑂚	𑂚

An analysis of the available manuscripts indicates that there is no semantic distinction between the visible *virāma* forms and *khaṇḍa* ligatures. The two forms of <TA, VIRAMA> and <NA, VIRAMA> are used alternately in the same context. The *khaṇḍa* form of TA is used inconsistently and when it does occur, its use is functionally identical to 𑂡 U+09CE BENGALI LETTER KHANDA TA. The combination <MA, VIRAMA> occurs only as the *khaṇḍa* ligature in the available sources.

It is necessary to represent both the *virāma* and *khaṇḍa* forms because they occur simultaneously. Moreover, they must be distinguished for analytical purposes. For instance, in a pedagogical text on Bhaiksuki palaeography that describes the formation of the conjunct 𑂛 *mpa*, it is necessary to render <MA, VIRAMA> using a visible VIRAMA, eg. the statement “𑂚 *m* + 𑂛 *pa* → 𑂛 *mpa*” properly expresses the rule that a conjoining form of PA is written beneath a full-form of MA in order to produce the conjunct *mpa*, while the statement “𑂚 *m* + 𑂛 *pa* → 𑂛 *mpa*” does not correctly do so.

Given that there are only three attested *khaṇḍa* forms, it may be possible to encode each as independent characters. The limitation of this approach is that the discovery of other *khaṇḍa* forms would require encoding

each as distinct characters. Another option is to encode a combining character with VIRAMA-like behavior and properties, which would control the representation of *khaṇḍa* forms. Such a control character would allow for a generic way of representing both existing and other potential *khaṇḍa* forms; however, encoding such a character may increase the complexities of implementing support for the script in rendering engines.

As there is no known semantic distinction between the *khaṇḍa* forms and their visible *virāma* representations, the three ligatures are to be considered contextual variants and will not be represented in plain text. The display of *khaṇḍa* forms is to be controlled using smart font ligature features, such as those available in Graphite and OpenType. If a requirement to represent these *khaṇḍa* ligatures at the character level arises as a result of new information, then the matter may be discussed at that time.

### 3.8 Consonant-Vowel Combinations

Several vowel signs are written using alternate forms when they combine with certain consonants. This is standard behavior for Bhaiksuki. These alternate forms are contextual variants and are not proposed for separate encoding. The font is responsible for producing the appropriate alternate form of a vowel sign based upon the presence of the base consonant.

**VOWEL SIGN AA** The ◌<sup>◌</sup> VOWEL SIGN AA takes the shape ◌<sup>◌</sup> when it attaches to the letter ◌<sup>◌</sup> KHA, eg. ◌<sup>◌</sup> *khā*. This form resembles the corresponding sign in Devanagari.















**VOWEL SIGN I** If the letter to which ◌<sup>◌</sup> VOWEL SIGN I is attached has two arrow heads, the sign is written in place of the right-hand arrow head. In the case of KHA, which has no arrow head, the sign attaches to the upper right side of the letter, eg. ◌<sup>◌</sup> *khi*.



**VOWEL SIGN U** The ◌<sup>◌</sup> VOWEL SIGN U takes the shape ◌<sup>◌</sup> when written with KA, GA, TA, BHA, SHA, eg. ◌<sup>◌</sup> *ku*, ◌<sup>◌</sup> *gu*, ◌<sup>◌</sup> *tu*, ◌<sup>◌</sup> *bhu*, ◌<sup>◌</sup> *śu*. When VOWEL SIGN U is attached to an independent RA the combination is rendered using the special ligature ◌<sup>◌</sup> *ru*. However, the form ◌<sup>◌</sup> is used with RA whenever the pair occurs in conjuncts, eg. ◌<sup>◌</sup> *bru* and ◌<sup>◌</sup> *śru*.










**VOWEL SIGN UU** The ◌<sup>◌</sup> VOWEL SIGN UU takes the alternate form ◌<sup>◌</sup> when it occurs with KA, GA, TA, BHA, and SHA, eg. ◌<sup>◌</sup> *gū*, ◌<sup>◌</sup> *tū*, ◌<sup>◌</sup> *bhū*, ◌<sup>◌</sup> *śū*. When VOWEL SIGN UU is attached to an independent RA the combination is written using the special ligature ◌<sup>◌</sup> *rū*. However, the form ◌<sup>◌</sup> is used with RA whenever the letter appears in conjuncts, eg. ◌<sup>◌</sup> *bhrū*, ◌<sup>◌</sup> *śrū*.


**VOWEL SIGN E** The ◌<sup>◌</sup> VOWEL SIGN E consists of two parts and is written to the right and left of a consonant letter, eg. ◌<sup>◌</sup> *ke*, ◌<sup>◌</sup> *ne*. It takes the alternate shape ◌<sup>◌</sup> with the following letters, whose arrow-head is placed to the right of the letter body: GA, NGA, NYA, TTA, TTHA, NNA, THA, DHA, SHA, eg. ◌<sup>◌</sup> *ge*, ◌<sup>◌</sup> *te*, ◌<sup>◌</sup> *the*, ◌<sup>◌</sup> *ne*, ◌<sup>◌</sup> *the*, ◌<sup>◌</sup> *ne*, ◌<sup>◌</sup> *śe*. In the case of KHA, which has no arrow-head, the sign takes the shape ◌<sup>◌</sup>, eg. ◌<sup>◌</sup> *khe*.

**VOWEL SIGN AI** The ◌<sup>◌</sup> VOWEL SIGN AI consists of two parts and is written to the right and left of a consonant letter, eg. ◌<sup>◌</sup> *jai*, ◌<sup>◌</sup> *nai*, ◌<sup>◌</sup> *pai*. As with VOWEL SIGN E, it takes the shape ◌<sup>◌</sup> when it occurs with GA, NGA, NYA, TTA, TTHA, NNA, THA, DHA, SHA, eg. ◌<sup>◌</sup> *nai*, ◌<sup>◌</sup> *śai*. For the letter KHA, which does not have an arrow-headed top, the VOWEL SIGN AI is written as ◌<sup>◌</sup>, eg. ◌<sup>◌</sup> *khyai*.










**VOWEL SIGN O** The ◌◌ VOWEL SIGN O consists of two parts and attaches to the right and left of a consonant letter, eg.  *ko*,  *co*,  *to*. It takes the alternate shape ◌◌ with the letters GA, NGA, NYA, TTA, TTHA, NNA, THA, DHA, SHA:  *go*,  *ño*,  *ño*,  *to*,  *tho*,  *ño*,  *tho*,  *dho*,  *so*. When the sign is written with DDA, NA, and RA it takes the shape ◌◌, eg.  *no*,  *ro*. Similarly, in the case of KHA the sign is an elongated half downward stroke. In the case of NYA, the vertical stroke on the right-hand side is shortened.





The letters KHA, NYA, DDA, NA, YA, RA are slightly modified when they occur with VOWEL SIGN O, eg.  *yo*,  *ro* (see specific details on each letter below).




**VOWEL SIGN AU** The ◌◌ VOWEL SIGN AU consists of two parts and is written to the right and left of a consonant, eg.  *cau*,  *dhau*,  *dau*,  *pau*,  *bhau*. It takes the alternate shape ◌◌ when it is written with GA, NGA, NYA, TTA, TTHA, NNA, THA, DHA, SHA:  *gau*,  *nau*,  *sau*. As with VOWEL SIGN O, when VOWEL SIGN AU attaches to NA and RA, it takes the shape ◌◌, eg.  *nau*, in which the right-hand vertical stroke is elongated.

The letters NA, YA, RA are slightly modified when they occur with VOWEL SIGN AU, eg.  *yau* (see specific details on each letter below).

Although combinations of KHA, NYA, DDA with VOWEL SIGN AU are not attested, the rendering of such pairs should be similar to that of VOWEL SIGN O.












































**PA** The letter  PA takes the contextual form  when written with certain vowel signs. The regular form is used in  *pu*,  *pū*,  *po*, but the form  is used for writing  *pā*,  *pi*,  *pī*. This alternate shape also occurs in consonant conjuncts when PA is a non-initial consonant.

**YA** The letter  YA takes a contextual form when it combines with all vowel signs, eg. VOWEL SIGN E ( *ye*), VOWEL SIGN AI ( *yai*), VOWEL SIGN AU ( *yau*).

**RA** The letter  RA takes a contextual form when it combines with all vowel signs, eg. VOWEL SIGN E ( *re*) and VOWEL SIGN AI ( *rai*).














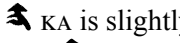














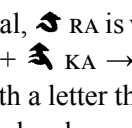









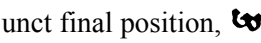






### 3.9 Consonant Conjuncts

Consonant clusters are written as conjuncts, which are generally rendered as vertically stacked ligatures, but in some few cases may also appear as independent ligatures. The encoded representation for conjuncts is <C<sub>1</sub>, ◌ VIRAMA, C<sub>2</sub>>. Examples of conjuncts are:



 <i>kṣa</i>	 KA +  SSA	 <i>tpa</i>	 TA +  PA
 <i>ggra</i>	 GA +  GA +  RA	 <i>dda</i>	 DA +  DA
 <i>gna</i>	 GA +  NA	 <i>ddha</i>	 DA +  DHA
 <i>nga</i>	 NGA +  GA	 <i>nthā</i>	 NA +  THA
 <i>ṅśa</i>	 NGA +  ŚHA	 <i>ndha</i>	 NA +  DHA
 <i>cca</i>	 CA +  CA	 <i>nna</i>	 NA +  NA
 <i>jña</i>	 JA +  NYA	 <i>bdha</i>	 BA +  DHA



Generally in writing conjuncts, the arrow-heads of non-initial consonants are removed, eg. see *cca*, *tta*, *dda* above. In some cases, however, the consonant retains its arrow-head, eg. see *nga*, *nśa*, *jña*, *ṣṭa* above.


Several letters take special forms when they occur in conjuncts:



- KA When initial,  KA is slightly modified to , eg.  KA +  RA →  *kra*;  KA +  SA →  *ssa*;  KA +  SA →  *sa*;  KA +  YA →  *kya*.
- PA The letter  PA takes the contextual form  when written non-initial in a cluster, eg.  MA +  PA →  *mpa*
- RA When cluster initial,  RA is written as an arrow-head  on top of the arrow-head of the following consonant, eg.  RA +  KA →  *rka*. This initial form is identical to the *repha* in Devanagari. When *repha* occurs with a letter that has two arrow-heads, eg.  SA,  YA, etc., it is written on top of the right-hand arrow head, eg.  RA +  YA →  *rya*. Above-base vowel signs attach to the *repha*. When RA occurs final in a cluster it is written using the subjoined form , eg.  TA +  RA →  *tra*. This form is to be differentiated from the  VOWEL SIGN VOCALIC R.
- YA When in conjunct final position,  YA is written as the subjoined form , eg.  KA +  YA →  *kya*;  JA +  YA →  *jya*;  YA +  YA →  *yya*;

### 3.10 Various Signs

**Candrabindu** The  SIGN CANDRABINDU is used for nasalization. The variant form  is written with a dot instead of a ring. This form is a glyphic variant and is semantically identical to the regular form.



**Anusvara** The  SIGN ANUSVARA is used for nasalization. It has the variant form , written as a dot instead of as a ring. The dotted form is a glyphic variant and is semantically identical to the regular form.




**Visarga** The  SIGN VISARGA represents post-vocalic aspiration (/h/) in Sanskrit.

**Avagraha** The  SIGN AVAGRAHA marks the elision of word-initial  *a* in Sanskrit as a result of sandhi.

### 3.11 Punctuation

**Daṇḍā-s** The  DANDA and  DOUBLE DANDA are marks of general punctuation.

**Word Separator** The  WORD SEPARATOR is used for demarcating lexical boundaries. It is written at the head-height. It also appears as a  dot instead of a vertical bar.

**Gap Filler** The  GAP FILLER is used generally as a spacing or completion mark, especially for justifying a text block, both at the end of the line before the binding area of the palm-leaf, as well as at the absolute end of a line. In other cases, the mark is used for indicating a gap after a deletion or to indicate a lacuna. The GAP FILLER is commonly written as , but the shape  is also attested. In some cases, this latter shape is used as mark of deletion and is written over the text to be erased. These two characters may be considered glyphic variants of a single gap, completion, or deletion mark.

### 3.12 Digits

The following decimal digits are attested: ONE, TWO, FOUR, FIVE, SIX, SEVEN, EIGHT, NINE. The digits \*ZERO and \*THREE are not attested, but space has been reserved for them.

### 3.13 Numbers

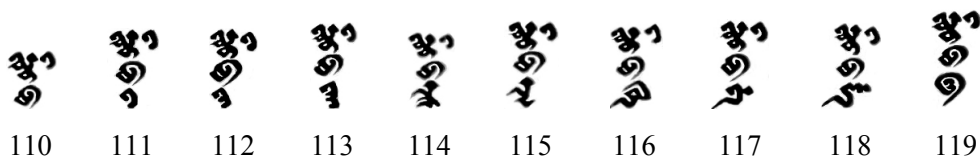
In addition to digits, there is a position numerical system used in Bhaiksuki records. The system contains numbers (referred to as ‘letter-numerals’ in the scholarly literature) for the primary and tens units and a generic unit mark for the hundreds:

୧	NUMBER ONE	୫	NUMBER EIGHT	୫	NUMBER SIXTY
୩	NUMBER TWO	୭	NUMBER NINE	୫	NUMBER SEVENTY
୩	NUMBER THREE	୭	NUMBER TEN	୫	NUMBER EIGHTY
୫	NUMBER FOUR	୫	NUMBER TWENTY	୫	NUMBER NINETY
୫	NUMBER FIVE	୫	NUMBER THIRTY	୫	NUMBER HUNDREDS
୫	NUMBER SIX	୫	NUMBER FORTY		
୫	NUMBER SEVEN	୫	NUMBER FIFTY		

Numbers are written vertically with each unit occupying a separate line. The largest unit is written first and all other units are written below. For instance, in writing the number 11, the NUMBER TEN is written first and NUMBER ONE is written below it. The same pattern is followed for all numbers 10–99. Shown below are the numbers 10–19:



The hundreds are written using the unit mark NUMBER HUNDREDS followed by a number that represents the primary numbers 1–9. For instance, 100 is ୫୧, 200 is ୫୩, and so forth. Given below are the numbers 110–119:



The system theoretically provides for all numbers in the range 1–999 to be written. However, numbers beyond 264 are not attested in the available manuscripts.



### 3.14 Editorial marks

Editorial marks, as the insertion marks ୫ and ୫, are used commonly in Bhaiksuki manuscripts. These are presently not proposed for encoding. The deletion and insertion marks resemble editorial marks used in































































other Indic scripts. It may be practical to unify these characters in a separate block of pan-Indic editorial marks.

### 3.15 Ornaments

In the Bhaiksuki manuscript shown recently in a Chinese documentary two elaborately ornamented *cakra*-s in different colors are visible. These may be classified as proper ornaments. However, analysis of these ornaments requires access to the newly revealed manuscript. In the *Maṇicūḍajātaka* manuscript the sign  is used in one instance apparently as an ornamental device. It is graphically identical to  NUMBER NINE. Until additional information is available, this sign is to be represented using NUMBER NINE.

### 3.16 Collation

The primary collating order for Bhaiksuki is as follows:

 A <  AA <  I <  II <  U <  UU <  VOCALIC R <  
 VOCALIC RR <  VOCALIC L <  E <  AI <  O <  AU <  KA <  
 KHA <  GA <  GHA <  NGA <  CA <  CHA <  JA <  
 JHA <  NYA <  TTA <  TTHA <  DDA <  DDHA <  NNA <  
 TA <  THA <  DA <  DHA <  NA <  PA <  PHA <  BA <  
 BHA <  MA <  YA <  RA <  LA <  VA <  SHA <  SSA <  
 SA <  HA <  AVAGRAHA <  SIGN AA <  SIGN I <  SIGN II <  
 SIGN U <  SIGN UU <  SIGN VOCALIC R <  SIGN VOCALIC RR <  
 SIGN VOCALIC L <  SIGN E <  SIGN AI <  SIGN O <  SIGN AU <  
 VIRAMA

The following characters have secondary weights:

 CANDRABINDU,  ANUSVARA,  VISARGA,

## 4 Character Data

### 4.1 Character Properties

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

```

11C00;BHAIKSUKI LETTER A;Lo;0;L;;;;N;;;;;
11C01;BHAIKSUKI LETTER AA;Lo;0;L;;;;N;;;;;
11C02;BHAIKSUKI LETTER I;Lo;0;L;;;;N;;;;;
11C03;BHAIKSUKI LETTER II;Lo;0;L;;;;N;;;;;
  
```

11C04;BHAIKSUKI LETTER U;Lo;0;L;;;;N;;;;;  
 11C05;BHAIKSUKI LETTER UU;Lo;0;L;;;;N;;;;;  
 11C06;BHAIKSUKI LETTER VOCALIC R;Lo;0;L;;;;N;;;;;  
 11C07;BHAIKSUKI LETTER VOCALIC RR;Lo;0;L;;;;N;;;;;  
 11C08;BHAIKSUKI LETTER VOCALIC L;Lo;0;L;;;;N;;;;;  
 11C0A;BHAIKSUKI LETTER E;Lo;0;L;;;;N;;;;;  
 11C0B;BHAIKSUKI LETTER AI;Lo;0;L;;;;N;;;;;  
 11C0C;BHAIKSUKI LETTER O;Lo;0;L;;;;N;;;;;  
 11C0D;BHAIKSUKI LETTER AU;Lo;0;L;;;;N;;;;;  
 11C0E;BHAIKSUKI LETTER KA;Lo;0;L;;;;N;;;;;  
 11C0F;BHAIKSUKI LETTER KHA;Lo;0;L;;;;N;;;;;  
 11C10;BHAIKSUKI LETTER GA;Lo;0;L;;;;N;;;;;  
 11C11;BHAIKSUKI LETTER GHA;Lo;0;L;;;;N;;;;;  
 11C12;BHAIKSUKI LETTER NGA;Lo;0;L;;;;N;;;;;  
 11C13;BHAIKSUKI LETTER CA;Lo;0;L;;;;N;;;;;  
 11C14;BHAIKSUKI LETTER CHA;Lo;0;L;;;;N;;;;;  
 11C15;BHAIKSUKI LETTER JA;Lo;0;L;;;;N;;;;;  
 11C16;BHAIKSUKI LETTER JHA;Lo;0;L;;;;N;;;;;  
 11C17;BHAIKSUKI LETTER NYA;Lo;0;L;;;;N;;;;;  
 11C18;BHAIKSUKI LETTER TTA;Lo;0;L;;;;N;;;;;  
 11C19;BHAIKSUKI LETTER TTHA;Lo;0;L;;;;N;;;;;  
 11C1A;BHAIKSUKI LETTER DDA;Lo;0;L;;;;N;;;;;  
 11C1B;BHAIKSUKI LETTER DDHA;Lo;0;L;;;;N;;;;;  
 11C1C;BHAIKSUKI LETTER NNA;Lo;0;L;;;;N;;;;;  
 11C1D;BHAIKSUKI LETTER TA;Lo;0;L;;;;N;;;;;  
 11C1E;BHAIKSUKI LETTER THA;Lo;0;L;;;;N;;;;;  
 11C1F;BHAIKSUKI LETTER DA;Lo;0;L;;;;N;;;;;  
 11C20;BHAIKSUKI LETTER DHA;Lo;0;L;;;;N;;;;;  
 11C21;BHAIKSUKI LETTER NA;Lo;0;L;;;;N;;;;;  
 11C22;BHAIKSUKI LETTER PA;Lo;0;L;;;;N;;;;;  
 11C23;BHAIKSUKI LETTER PHA;Lo;0;L;;;;N;;;;;  
 11C24;BHAIKSUKI LETTER BA;Lo;0;L;;;;N;;;;;  
 11C25;BHAIKSUKI LETTER BHA;Lo;0;L;;;;N;;;;;  
 11C26;BHAIKSUKI LETTER MA;Lo;0;L;;;;N;;;;;  
 11C27;BHAIKSUKI LETTER YA;Lo;0;L;;;;N;;;;;  
 11C28;BHAIKSUKI LETTER RA;Lo;0;L;;;;N;;;;;  
 11C29;BHAIKSUKI LETTER LA;Lo;0;L;;;;N;;;;;  
 11C2A;BHAIKSUKI LETTER VA;Lo;0;L;;;;N;;;;;  
 11C2B;BHAIKSUKI LETTER SHA;Lo;0;L;;;;N;;;;;  
 11C2C;BHAIKSUKI LETTER SSA;Lo;0;L;;;;N;;;;;  
 11C2D;BHAIKSUKI LETTER SA;Lo;0;L;;;;N;;;;;  
 11C2E;BHAIKSUKI LETTER HA;Lo;0;L;;;;N;;;;;  
 11C2F;BHAIKSUKI VOWEL SIGN AA;Mc;0;L;;;;N;;;;;  
 11C30;BHAIKSUKI VOWEL SIGN I;Mn;0;NSM;;;;N;;;;;  
 11C31;BHAIKSUKI VOWEL SIGN II;Mn;0;NSM;;;;N;;;;;  
 11C32;BHAIKSUKI VOWEL SIGN U;Mn;0;NSM;;;;N;;;;;  
 11C33;BHAIKSUKI VOWEL SIGN UU;Mn;0;NSM;;;;N;;;;;  
 11C34;BHAIKSUKI VOWEL SIGN VOCALIC R;Mn;0;NSM;;;;N;;;;;  
 11C35;BHAIKSUKI VOWEL SIGN VOCALIC RR;Mn;0;NSM;;;;N;;;;;  
 11C36;BHAIKSUKI VOWEL SIGN VOCALIC L;Mn;0;NSM;;;;N;;;;;  
 11C38;BHAIKSUKI VOWEL SIGN E;Mn;0;NSM;;;;N;;;;;  
 11C39;BHAIKSUKI VOWEL SIGN AI;Mn;0;NSM;;;;N;;;;;  
 11C3A;BHAIKSUKI VOWEL SIGN O;Mc;0;L;;;;N;;;;;  
 11C3B;BHAIKSUKI VOWEL SIGN AU;Mc;0;L;;;;N;;;;;  
 11C3C;BHAIKSUKI SIGN CANDRABINDU;Mn;0;NSM;;;;N;;;;;  
 11C3D;BHAIKSUKI SIGN ANUSVARA;Mn;0;NSM;;;;N;;;;;  
 11C3E;BHAIKSUKI SIGN VISARGA;Mc;0;L;;;;N;;;;;  
 11C3F;BHAIKSUKI SIGN VIRAMA;Mn;9;L;;;;N;;;;;  
 11C40;BHAIKSUKI SIGN AVAGRAHA;Lo;0;L;;;;N;;;;;  
 11C41;BHAIKSUKI DANDA;Po;0;L;;;;N;;;;;  
 11C42;BHAIKSUKI DOUBLE DANDA;Po;0;L;;;;N;;;;;

```

11C43;BHAIKSUKI WORD SEPARATOR;Po;0;L;;;;N;;;;;
11C44;BHAIKSUKI GAP FILLER;Po;0;L;;;;N;;;;;
11C51;BHAIKSUKI DIGIT ONE;Nd;0;L;;1;1;1;N;;;;;
11C52;BHAIKSUKI DIGIT TWO;Nd;0;L;;2;2;2;N;;;;;
11C54;BHAIKSUKI DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;;
11C55;BHAIKSUKI DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;;
11C56;BHAIKSUKI DIGIT SIX;Nd;0;L;;6;6;6;N;;;;;
11C57;BHAIKSUKI DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;;
11C58;BHAIKSUKI DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;;
11C59;BHAIKSUKI DIGIT NINE;Nd;0;L;;9;9;9;N;;;;;
11C5A;BHAIKSUKI NUMBER ONE;No;0;L;;;1;N;;;;;
11C5B;BHAIKSUKI NUMBER TWO;No;0;L;;;2;N;;;;;
11C5C;BHAIKSUKI NUMBER THREE;No;0;L;;;3;N;;;;;
11C5D;BHAIKSUKI NUMBER FOUR;No;0;L;;;4;N;;;;;
11C5E;BHAIKSUKI NUMBER FIVE;No;0;L;;;5;N;;;;;
11C5F;BHAIKSUKI NUMBER SIX;No;0;L;;;6;N;;;;;
11C60;BHAIKSUKI NUMBER SEVEN;No;0;L;;;7;N;;;;;
11C61;BHAIKSUKI NUMBER EIGHT;No;0;L;;;8;N;;;;;
11C62;BHAIKSUKI NUMBER NINE;No;0;L;;;9;N;;;;;
11C63;BHAIKSUKI NUMBER TEN;No;0;L;;;10;N;;;;;
11C64;BHAIKSUKI NUMBER TWENTY;No;0;L;;;20;N;;;;;
11C65;BHAIKSUKI NUMBER THIRTY;No;0;L;;;30;N;;;;;
11C66;BHAIKSUKI NUMBER FORTY;No;0;L;;;40;N;;;;;
11C67;BHAIKSUKI NUMBER FIFTY;No;0;L;;;50;N;;;;;
11C68;BHAIKSUKI NUMBER SIXTY;No;0;L;;;60;N;;;;;
11C69;BHAIKSUKI NUMBER SEVENTY;No;0;L;;;70;N;;;;;
11C6A;BHAIKSUKI NUMBER EIGHTY;No;0;L;;;80;N;;;;;
11C6B;BHAIKSUKI NUMBER NINETY;No;0;L;;;90;N;;;;;
11C6C;BHAIKSUKI NUMBER HUNDREDS;No;0;L;;;100;N;;;;;

```

## 4.2 Linebreaking

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

```

11C00..11C08; AL # LETTER A .. LETTER VOCALIC L
11C0A..11C2E; AL # LETTER E .. LETTER HA
11C2F..11C36; CM # VOWEL SIGN AA .. VOWEL SIGN VOCALIC L
11C38..11C3F; CM # VOWEL SIGN E .. VIRAMA
11C40; AL # SIGN AVAGRAHA
11C41..11C43; BA # DANDA .. GAP FILLER
11C51..11C52; NU # DIGIT ONE .. DIGIT TWO
11C54..11C59; NU # DIGIT FOUR .. DIGIT NINE
11C5A..11C6C; AL # NUMBER ONE .. NUMBER HUNDREDS

```

## 4.3 Syllabic Categories

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

```

# Indic_Syllabic_Category=Bindu
11C3C ; Bindu # Mn SIGN CANDRABINDU
11C3D ; Bindu # Mn SIGN ANUSVARA

# Indic_Syllabic_Category=Visarga
11C3E ; Visarga # Mc SIGN VISARGA

# Indic_Syllabic_Category=Virama
11C3F ; Virama # Mn SIGN VIRAMA

```

```
# Indic_Syllabic_Category=Vowel_Independent
11C00..11C0D ; Vowel_Independent # Lo [14] LETTER A .. LETTER AU

# Indic_Syllabic_Category=Vowel_Dependent
11C2F ; Vowel_Dependent # Mc [3] VOWEL SIGN AA
11C30..11C36 ; Vowel_Dependent # Mn [8] VOWEL SIGN I .. VOWEL SIGN VOCALIC L
11C38..11C3B ; Vowel_Dependent # Mc [2] VOWEL SIGN E .. VOWEL SIGN AU

# Indic_Syllabic_Category=Consonant
11C0E..11C2E ; Consonant # Lo [34] LETTER KA .. LETTER HA
```

#### 4.4 Matra Categories

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

```
# Indic_Matra_Category=Right
11C2F ; Right # Mc VOWEL SIGN AA

# Indic_Matra_Category=Left_And_Right
11C38 ; Left_And_Right # Mc VOWEL SIGN E
11C3A ; Left_And_Right # Mc VOWEL SIGN O

# Indic_Matra_Category=Top
11C30..11C31 ; Top # Mn [2] VOWEL SIGN I .. VOWEL SIGN II

# Indic_Matra_Category=Top_And_Left_And_Right
11C39 ; Top_And_Left_And_Right # Mc VOWEL SIGN AI
11C3B ; Top_And_Left_And_Right # Mc VOWEL SIGN AU

# Indic_Matra_Category=Bottom
11C32..11C36 ; Bottom # Mn [5] VOWEL SIGN U .. VOWEL SIGN VOCALIC L
11C3F ; Bottom # Mn SIGN VIRAMA
```

















































































































## 5 References

- Arrow-headed Script Project. <http://www.uni-marburg.de/fb10/iksl/indologie/forschung/drittmittelprojekte/bhaiksuki>
- Dimitrov, Dragomir. 2010. *The Bhaiksukī Manuscript of the Candrālaṅkāra: Study, Script Tables, and Facsimile Edition*. Harvard Oriental Series, 72. Cambridge: Department of Sanskrit and Indian Studies, Harvard University.
- Hanisch, Albrecht. 2008. “Sarvarakṣita’s *Maṇicūḍajātaka*. Reproduction of the Codex Unicus with Diplomatic Transcript and Palaeographic Introduction to the Bhaiksukī Script.” In *Sanskrit Texts from Giuseppe Tucci’s Collection*, Part I. Manuscripta Buddhica, I, p. 195–342. Edited by Harunaga Isaacson and Francesco Sferra. Rome: Istituto Italiano per l’Africa e l’Oriente.
- Pandey, Anshuman. 2011. “Preliminary Proposal to Encode the Bhaiksuki Script in ISO/IEC 10646”. N4121 L2/11-259 (July 11, 2011). <http://std.dkuug.dk/jtc1/sc2/wg2/docs/n4121.pdf>
- Pandey, Anshuman and Dragomir Dimitrov. 2013. “Proposal to Encode the Bhaiksuki Script in ISO/IEC 10646”. N4469 L2/13-167 (July 22, 2013). <http://std.dkuug.dk/jtc1/sc2/wg2/docs/n4469.pdf>

## **6 Acknowledgments**

The authors would like to express their gratitude to Shriramana Sharma for discussions regarding the representation of consonant-virama ligatures.

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

	11C0	11C1	11C2	11C3	11C4	11C5	11C6
0	 11C00	 11C10	 11C20	 11C30	 11C40		 11C60
1	 11C01	 11C11	 11C21	 11C31	 11C41	 11C51	 11C61
2	 11C02	 11C12	 11C22	 11C32	 11C42	 11C52	 11C62
3	 11C03	 11C13	 11C23	 11C33	 11C43		 11C63
4	 11C04	 11C14	 11C24	 11C34	 11C44	 11C54	 11C64
5	 11C05	 11C15	 11C25	 11C35		 11C55	 11C65
6	 11C06	 11C16	 11C26	 11C36		 11C56	 11C66
7	 11C07	 11C17	 11C27			 11C57	 11C67
8	 11C08	 11C18	 11C28	 11C38		 11C58	 11C68
9		 11C19	 11C29	 11C39		 11C59	 11C69
A	 11C0A	 11C1A	 11C2A	 11C3A		 11C5A	 11C6A
B	 11C0B	 11C1B	 11C2B	 11C3B		 11C5B	 11C6B
C	 11C0C	 11C1C	 11C2C	 11C3C		 11C5C	 11C6C
D	 11C0D	 11C1D	 11C2D	 11C3D		 11C5D	
E	 11C0E	 11C1E	 11C2E	 11C3E		 11C5E	
F	 11C0F	 11C1F	 11C2F	 11C3F		 11C5F	

**Independent vowels**

11C00	𑄀	BHAIKSUKI LETTER A
11C01	𑄁	BHAIKSUKI LETTER AA
11C02	𑄂	BHAIKSUKI LETTER I
11C03	𑄃	BHAIKSUKI LETTER II
11C04	𑄄	BHAIKSUKI LETTER U
11C05	𑄅	BHAIKSUKI LETTER UU
11C06	𑄆	BHAIKSUKI LETTER VOCALIC R
11C07	𑄇	BHAIKSUKI LETTER VOCALIC RR
11C08	𑄈	BHAIKSUKI LETTER VOCALIC L
11C09	𑄉	<reserved>
11C0A	𑄊	BHAIKSUKI LETTER E
11C0B	𑄋	BHAIKSUKI LETTER AI
11C0C	𑄌	BHAIKSUKI LETTER O
11C0D	𑄍	BHAIKSUKI LETTER AU

**Consonants**

11C0E	𑄎	BHAIKSUKI LETTER KA
11C0F	𑄏	BHAIKSUKI LETTER KHA
11C10	𑄐	BHAIKSUKI LETTER GA
11C11	𑄑	BHAIKSUKI LETTER GHA
11C12	𑄒	BHAIKSUKI LETTER NGA
11C13	𑄓	BHAIKSUKI LETTER CA
11C14	𑄔	BHAIKSUKI LETTER CHA
11C15	𑄕	BHAIKSUKI LETTER JA
11C16	𑄖	BHAIKSUKI LETTER JHA
11C17	𑄗	BHAIKSUKI LETTER NYA
11C18	𑄘	BHAIKSUKI LETTER TTA
11C19	𑄙	BHAIKSUKI LETTER TTHA
11C1A	𑄚	BHAIKSUKI LETTER DDA
11C1B	𑄛	BHAIKSUKI LETTER DDHA
11C1C	𑄜	BHAIKSUKI LETTER NNA
11C1D	𑄝	BHAIKSUKI LETTER TA
11C1E	𑄞	BHAIKSUKI LETTER THA
11C1F	𑄟	BHAIKSUKI LETTER DA
11C20	𑄠	BHAIKSUKI LETTER DHA
11C21	𑄡	BHAIKSUKI LETTER NA
11C22	𑄢	BHAIKSUKI LETTER PA
11C23	𑄣	BHAIKSUKI LETTER PHA
11C24	𑄤	BHAIKSUKI LETTER BA
11C25	𑄥	BHAIKSUKI LETTER BHA
11C26	𑄦	BHAIKSUKI LETTER MA
11C27	𑄧	BHAIKSUKI LETTER YA
11C28	𑄨	BHAIKSUKI LETTER RA
11C29	𑄩	BHAIKSUKI LETTER LA
11C2A	𑄪	BHAIKSUKI LETTER VA
11C2B	𑄫	BHAIKSUKI LETTER SHA
11C2C	𑄬	BHAIKSUKI LETTER SSA
11C2D	𑄭	BHAIKSUKI LETTER SA
11C2E	𑄮	BHAIKSUKI LETTER HA

**Dependent vowel signs**

11C2F	𑄯	BHAIKSUKI VOWEL SIGN AA
11C30	𑄰	BHAIKSUKI VOWEL SIGN I
11C31	𑄱	BHAIKSUKI VOWEL SIGN II
11C32	𑄲	BHAIKSUKI VOWEL SIGN U
11C33	𑄳	BHAIKSUKI VOWEL SIGN UU
11C34	𑄴	BHAIKSUKI SIGN VOCALIC R
11C35	𑄵	BHAIKSUKI SIGN VOCALIC RR
11C36	𑄶	BHAIKSUKI SIGN VOCALIC L
11C37	𑄷	<reserved>
11C38	𑄸	BHAIKSUKI VOWEL SIGN E
11C39	𑄹	BHAIKSUKI VOWEL SIGN AI
11C3A	𑄺	BHAIKSUKI VOWEL SIGN O
11C3B	𑄻	BHAIKSUKI VOWEL SIGN AU

**Various signs**

11C3C	𑄼	BHAIKSUKI SIGN CANDRABINDU
11C3D	𑄽	BHAIKSUKI SIGN ANUSVARA
11C3E	𑄾	BHAIKSUKI SIGN VISARGA
11C3F	𑄿	BHAIKSUKI SIGN VIRAMA
11C40	𑅀	BHAIKSUKI SIGN AVAGRAHA

**Punctuation**

11C41	𑅁	BHAIKSUKI DANDA
11C42	𑅂	BHAIKSUKI DOUBLE DANDA
11C43	𑅃	BHAIKSUKI WORD SEPARATOR
11C44	𑅄	BHAIKSUKI GAP FILLER

**Digits**

11C50	𑅅	<reserved>
11C51	𑅆	BHAIKSUKI DIGIT ONE
11C52	𑅇	BHAIKSUKI DIGIT TWO
11C53	𑅈	<reserved>
11C54	𑅉	BHAIKSUKI DIGIT FOUR
11C55	𑅊	BHAIKSUKI DIGIT FIVE
11C56	𑅋	BHAIKSUKI DIGIT SIX
11C57	𑅌	BHAIKSUKI DIGIT SEVEN
11C58	𑅍	BHAIKSUKI DIGIT EIGHT
11C59	𑅎	BHAIKSUKI DIGIT NINE

**Numbers***Also known as letter-numerals*

11C5A	𑅏	BHAIKSUKI NUMBER ONE
11C5B	𑅐	BHAIKSUKI NUMBER TWO
11C5C	𑅑	BHAIKSUKI NUMBER THREE
11C5D	𑅒	BHAIKSUKI NUMBER FOUR
11C5E	𑅓	BHAIKSUKI NUMBER FIVE
11C5F	𑅔	BHAIKSUKI NUMBER SIX
11C60	𑅕	BHAIKSUKI NUMBER SEVEN
11C61	𑅖	BHAIKSUKI NUMBER EIGHT
11C62	𑅗	BHAIKSUKI NUMBER NINE
11C63	𑅘	BHAIKSUKI NUMBER TEN
11C64	𑅙	BHAIKSUKI NUMBER TWENTY
11C65	𑅚	BHAIKSUKI NUMBER THIRTY
11C66	𑅛	BHAIKSUKI NUMBER FORTY
11C67	𑅜	BHAIKSUKI NUMBER FIFTY
11C68	𑅝	BHAIKSUKI NUMBER SIXTY
11C69	𑅞	BHAIKSUKI NUMBER SEVENTY
11C6A	𑅟	BHAIKSUKI NUMBER EIGHTY
11C6B	𑅠	BHAIKSUKI NUMBER NINETY
11C6C	𑅡	BHAIKSUKI NUMBER HUNDREDS



Figure 2: Folios 2a and 2b of the *Candrālamkāra* in Bhaiksuki (from Dimitrov 2010). Transliteration given in figure 3.



[fol. 2<sup>a</sup>] śeṣā hiṃsyā doṣā rāgādayo yasya sa ○ **pradhvasta** iti *KARTTARI CĀRAMBHE KTAḤ* (1.2.68) <<|>> pradhvastavāṃś cāsau āśeṣadoṣas ceti **pradhva:**<sub>(2)</sub>**stāśeṣadoṣaḥ** | itas tataḥ kṣeptuṃ : ○ prārabdhavān<sup>1</sup> hiṃsyān<sup>1</sup> rāgādīn ity arthaḥ | yadvā pradhvaṃsitum ārabdho aśeṣa:<sub>(3)</sub>doṣo yeneti karmmaṇi kte aśeṣaśa○bdaḥ sāvaśeṣārthaḥ <<|>> etenotpādita-vodhicitta ity uktaṃ | nanv evaṃbhūtaḥ pṛthagja[no]<sub>(4)</sub>pi bhavatīti viśi-naṣṭi | samasyante : ○ satvasantāne prakṣipyante iti **samastās** ca te **guṇās** ca [da]yāmudito pekṣā<sub>(5)</sub>kṣamādaya:ḥ | teṣāṃ **śalanam** śālah ○ | *śala gatau* (Dhātup. 1.572)<sup>72</sup> ghañ{a} | jñānam śobhā vā prāptir vāsyāstīti | athavā taiḥ śali(tap)[i](tā)<sub>(6)</sub> jñātum | śobhitum āptum vā śīlam asye○ti | etena svārthanirapekṣāvyāhataparārthakāritayā ārya tam uktaṃ | a[ta e]<sub>(7)</sub>vāha | **parasmai anugrahas** tena **dakṣo** va:○rddh[i]to bodhisatvatvena lokair jñāto vā | *dakṣa vṛddhau gatau ceti* (Dhātup. 1.446, 1.518)<sup>73</sup> Dhā[itu](pāṭh). .. .. [a]<sub>(8)</sub>ta evaitasmād guṇatrayayogād **Bu:○ddhaḥ** | budhadhātor dantyauşṭhavakārapakṣe vaśāditvābhāvād .(ṣa) .. .. + + + + .. .. .. .. <sub>(9)</sub> jñānam tām jihite uttarottaram adhiga○cchati pūrvvapūrvvaṃ jahātīti | *JHALOJAŚ* iti (6.3.67) dhasya [da]ḥ (l) .. .. .

[fol. 2<sup>b</sup>] [kr̥]te Buddhaḥ | athavā *KARTTARI CĀRAMBHE KTAḤ* (1.2.68) | svayaṃ ○ boddhuṃ pravṛttaḥ parañ ca vedayitum iti | yadvā budhyate sa [tair iti b]u[d]dhaḥ | [evam](bhūtā)<sub>(2)</sub>ya [na]amaḥ sadeti śeṣaḥ <<|>> bhinnakleśatvena *grhī* : ○ vā *navako pi vā vandyo vratadharair* iti (*Gurupañcāśikā* 4bc)<sup>74</sup> vacanāt | dvāv eva vandyau .. ++ .. (vṛddha)[ś] (cet)[i] (asarvva)<sub>(3)</sub>janaviṣayatvāc cācāryasya tannamaskāra evā○śaṃsyata ity **astu** ity uktaṃ | Bhagavatas tu trailokyagurutvāt svata e[va s]iddha[tvād] .. .. . <sub>(4)</sub>[t]enāpi tadviśayaḥ śiṣyāṇam āśaṃsyate | a○[śr]utavāt | **sid-dham ityādi** vākyasya tadarthatvāc ca | bhikṣutvena la[bdhatvān n(ā)]śa-s(o) nāpi <sub>(5)</sub>sadā sarvvakālam astv iti vidhau | ā vodheḥ śara○ṇagamane-na vidhyabhāvāt <<|>> ggranthasyādau namo stu pratipannaśāstraketa-robhayeṣā:<sub>(6)</sub>m ity abhisandhinā *praiśānujñāyor lloḍ* ity anye Su○gate namaskāram ādbhaḥ<sup>75</sup> <<|>> nanv ācāryasyānekaggranthakartṛtvāt [k]im iti na Lokānandādau : <sub>(7)</sub>ṭikā kriyata iti stūyamānaśālitayā ○ śāstraṃ stauti **jayātīyādi** | Candranāmnā praṇītas **Candraḥ** prabhākaravata upa[ca]-<sub>(8)</sub>ryamāṇasya dhvanayo bhavanti yathā yaṣṭī<<ḥ>> *pra○veśayeti*<sup>76</sup> puṃliṅgaḥ śāstravācī | ata eva pūrvvaśloke nāmānabhidhāne [pi] viś[e]<sub>(9)</sub>ṣa-[par]iggrahaḥ | athavā sakalalakṣāvabhāsa○katvāc candayatīti candraḥ | *cadi āhlādane* | *dīptau ceti* (*Dhātupradīpa* 1.55) Raki<sup>77</sup> | katham avabhāsayā

Figure 3: Transliteration of Folio 2 of the *Candrālaṃkāra* from figure 2.



Figure 4: A statue of Buddha from Gaya, Bihar with a Bhaiksuki inscription (enlarged) on the underside of the base.

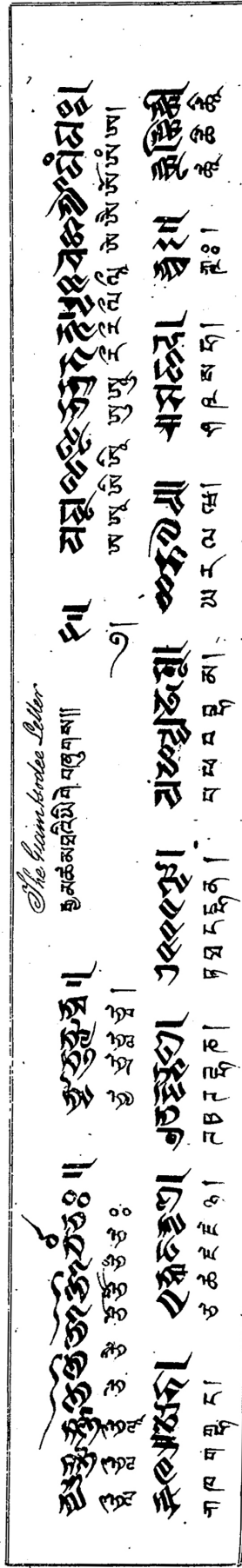
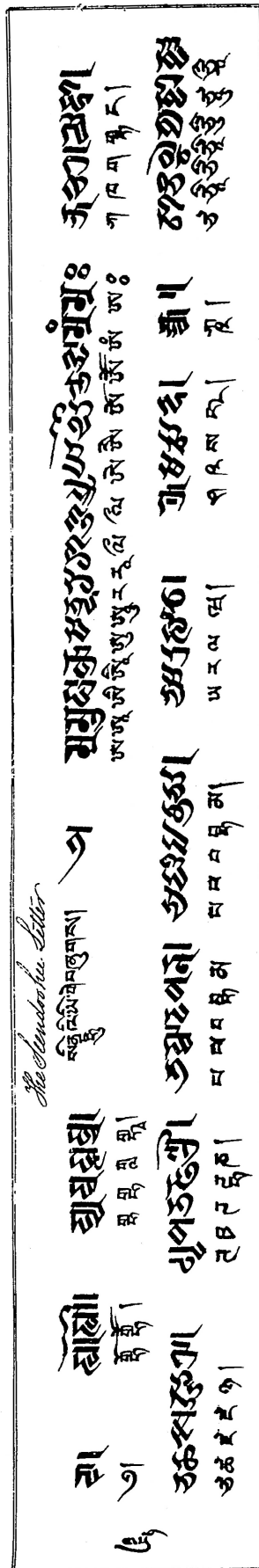


Figure 5: Folios showing the ‘Sindhura’ script with Tibetan correspondences.



  
 a ā i ī u ū ṛ ṝ ḷ e ai o au

Figure 6: Bhaiksuki vowels (from Dimitrov 2010: 75).































				
ka	kha	ga	gha	ṅa
				
ca	cha	ja	jha	ña
				
ṭa	ṭha	ḍa	ḍha	ṇa
				
ta	tha	da	dha	na
				
pa	pha	ba	bha	ma
				
ya	ra	la	va	
				
śa	ṣa	sa	ha	

Figure 7: Bhaiksuki consonants (from Dimitrov 2010: 75).



Figure 8: Consonant-vowel combinations with VOWEL SIGN AA (from Dimitrov 2010: 85).

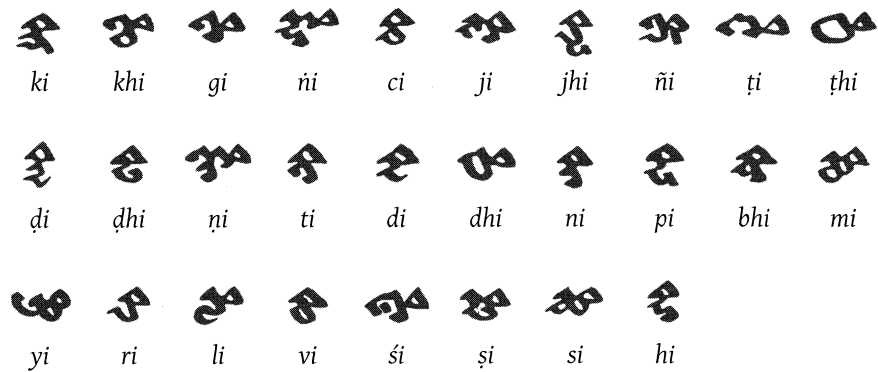


Figure 9: Consonant-vowel combinations with VOWEL SIGN I (from Dimitrov 2010: 86).

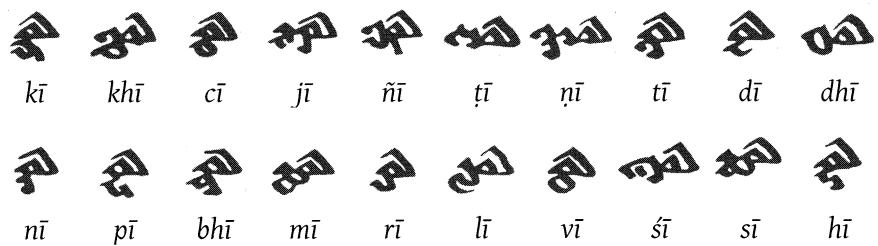


Figure 10: Consonant-vowel combinations with VOWEL SIGN II (from Dimitrov 2010: 86).

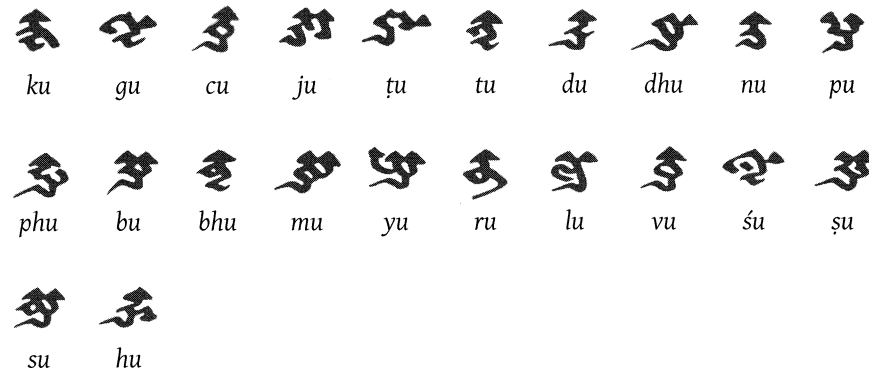


Figure 11: Consonant-vowel combinations with VOWEL SIGN U (from Dimitrov 2010: 86).

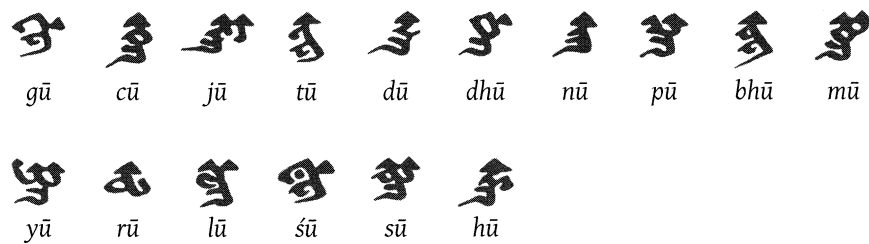


Figure 12: Consonant-vowel combinations with VOWEL SIGN UU (from Dimitrov 2010: 87).

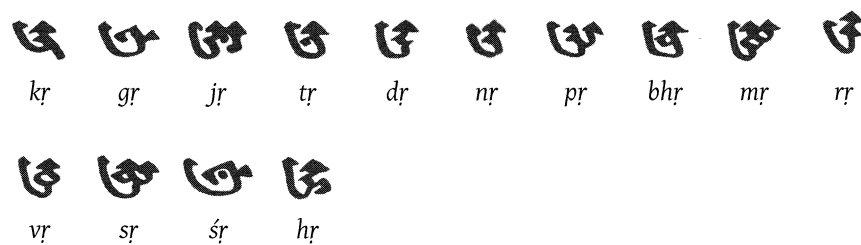


Figure 13: Consonant-vowel combinations with VOWEL SIGN VOCALIC R (from Dimitrov 2010: 87).

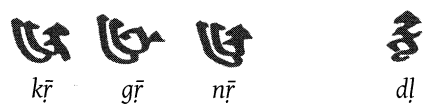


Figure 14: Consonant-vowel combinations with VOWEL SIGN VOCALIC RR and VOWEL SIGN VOCALIC L (from Dimitrov 2010: 87).

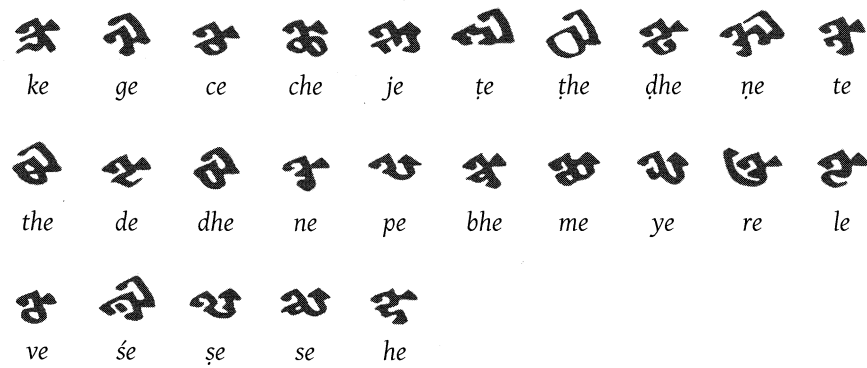


Figure 15: Consonant-vowel combinations with VOWEL SIGN E (from Dimitrov 2010: 87).



Figure 16: Consonant-vowel combinations with VOWEL SIGN AI (from Dimitrov 2010: 88).

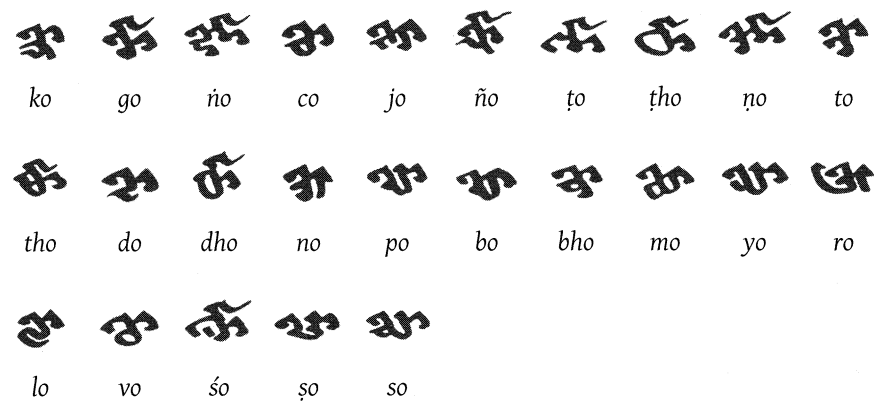


Figure 17: Consonant-vowel combinations with VOWEL SIGN O (from Dimitrov 2010: 88).

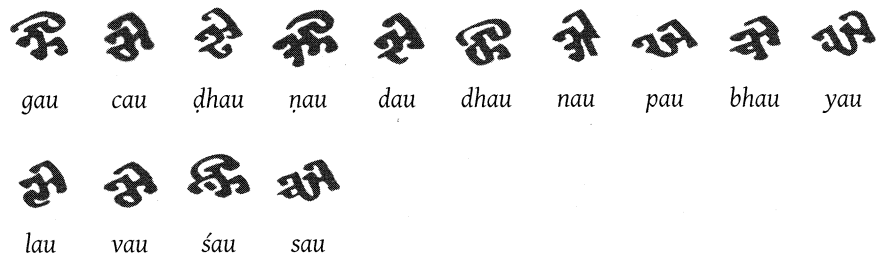


Figure 18: Consonant-vowel combinations with VOWEL SIGN AU (from Dimitrov 2010: 88).

3.1.4 Anusvāra, visarga, avagraha, virāma, and other symbols

kaṃ	ṇaṃ	taṃ	daṃ	naṃ	paṃ	yaṃ	raṃ	laṃ	vaṃ
śaṃ	saṃ	kāṃ	thāṃ	pāṃ	yāṃ	kiṃ	hiṃ	puṃ	kaḥ
gaḥ	ṅaḥ	caḥ	ṅāḥ	ṭhaḥ	ḍaḥ	ṇaḥ	taḥ	daḥ	dhaḥ
naḥ	paḥ	maḥ	yaḥ	raḥ	vaḥ	śaḥ	ṣaḥ	saḥ	k
j	t	ṭ	n	ṅ	p	m	ṃ	ṁ	ṡ
ṁ	ḥ	'	†	∴	∷				^
×	+	\	/	4	5	6	7	8	9

Figure 19: Bhaiksuki digits and various signs (from Dimitrov 2010: 89).



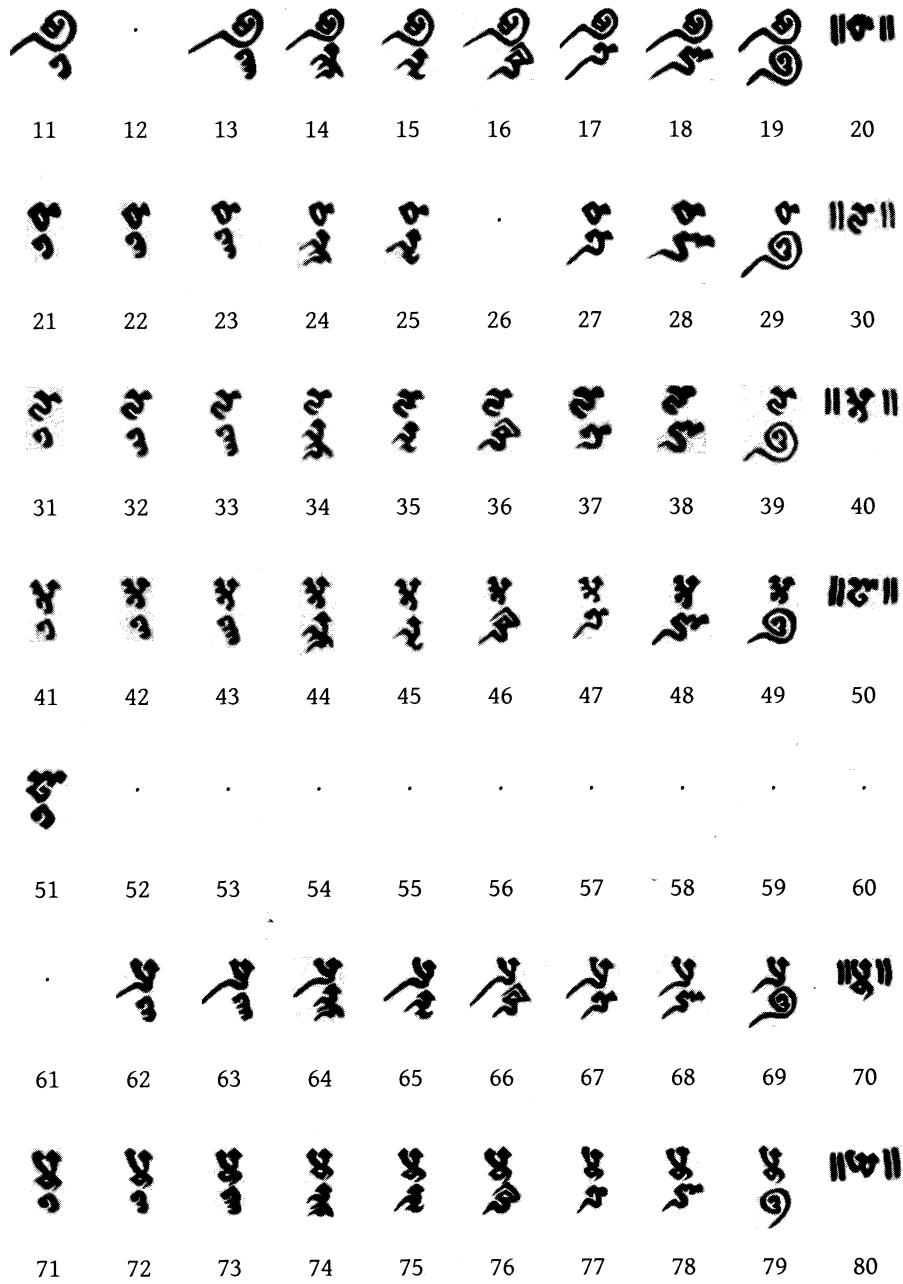


Figure 20: Some Bhaiksuki numbers (from Dimitrov 2010: 54).

	Bhai.	Sid.	Sha.	Dev.		Bhai.	Sid.	Sha.	Dev.
A	अ	𑀅	𑀆	अ					
AA	आ	𑀇	𑀈	आ	SIGN AA	𑀇	𑀇	𑀇	𑀇
I	इ	𑀉	𑀊	इ	SIGN I	𑀉	𑀉	𑀉	𑀉
II	ई	𑀋	𑀌	ई	SIGN II	𑀋	𑀋	𑀋	𑀋
U	उ	𑀍	𑀎	उ	SIGN U	𑀍	𑀍	𑀍	𑀍
UU	ऊ	𑀏	𑀐	ऊ	SIGN UU	𑀏	𑀏	𑀏	𑀏
VOC.R	ऋ	𑀑	𑀒	ऋ	SIGN R	𑀑	𑀑	𑀑	𑀑
VOC.RR	ॠ	𑀓	𑀔	ॠ	SIGN RR	𑀓	𑀓	𑀓	𑀓
VOC.L	ऌ	𑀕	𑀖	ऌ	SIGN L	𑀕	—	𑀕	𑀕
VOC.LL	—	𑀗	𑀘	ऌ	SIGN LL	—	—	𑀕	𑀕
E	ए	𑀙	𑀚	ए	SIGN E	𑀙	𑀙	𑀙	𑀙
AI	ऐ	𑀛	𑀜	ऐ	SIGN AI	𑀛	𑀛	𑀛	𑀛
O	ओ	𑀝	𑀞	ओ	SIGN O	𑀝	𑀝	𑀝	𑀝
AU	औ	𑀟	𑀠	औ	SIGN AU	𑀟	𑀟	𑀟	𑀟

Table 1: Comparison of vowel letters and signs of Bhaiksuki, Siddham, Sharada, and Devanagari.

	Bhai.	Sid.	Sha.	Dev.		Bhai.	Sid.	Sha.	Dev.
KA	𑂔	𑂒	क	क	DA	𑂔	𑂒	द	द
KHA	𑂕	𑂓	ख	ख	DHA	𑂕	𑂓	ध	ध
GA	𑂖	𑂑	ग	ग	NA	𑂖	𑂑	न	न
GHA	𑂗	𑂒	घ	घ	PA	𑂗	𑂒	प	प
NGA	𑂘	𑂒	ङ	ङ	PHA	𑂘	𑂒	फ	फ
CA	𑂙	𑂒	च	च	BA	𑂙	𑂒	ब	ब
CHA	𑂚	𑂒	छ	छ	BHA	𑂚	𑂒	भ	भ
JA	𑂛	𑂒	ज	ज	MA	𑂛	𑂒	म	म
JHA	𑂜	𑂒	झ	झ	YA	𑂜	𑂒	य	य
NYA	𑂝	𑂒	ञ	ञ	RA	𑂝	𑂒	र	र
TTA	𑂞	𑂒	ट	ट	LA	𑂞	𑂒	ल	ल
TTHA	𑂟	𑂒	ठ	ठ	VA	𑂟	𑂒	व	व
DDA	𑂠	𑂒	ड	ड	SHA	𑂠	𑂒	श	श
DDHA	𑂡	𑂒	ढ	ढ	SSA	𑂡	𑂒	ष	ष
NNA	𑂢	𑂒	ण	ण	SA	𑂢	𑂒	स	स
TA	𑂣	𑂒	त	त	HA	𑂣	𑂒	ह	ह
THA	𑂤	𑂒	थ	थ					

Table 2: Comparison of consonant letters of Bhaiksuki, Siddham, Sharada, and Devanagari.

**ISO/IEC JTC 1/SC 2/WG 2  
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS  
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646<sup>1</sup>**

**Please fill all the sections A, B and C below.**

Please read Principles and Procedures Document (P & P) from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

**A. Administrative**

1. Title:	<b>Proposal to Encode the Bhaiksuki Script in ISO/IEC 10646</b>
2. Requester's name:	<i>Script Encoding Initiative (SEI) / Anshuman Pandey (pandey@umich.edu) / Dragomir Dimitrov (dimitrov@staff.uni-marburg.de)</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>Liaison contribution</i>
4. Submission date:	<i>2013-10-27</i>
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	<input checked="" type="checkbox"/> <b>Yes</b>
(or) More information will be provided later:	<input type="checkbox"/>

**B. Technical – General**

1. Choose one of the following:	
a. This proposal is for a new script (set of characters):	<input checked="" type="checkbox"/> <b>Yes</b>
Proposed name of script:	<i>Bhaiksuki</i>
b. The proposal is for addition of character(s) to an existing block:	
Name of the existing block:	
2. Number of characters in proposal:	<i>94</i>
3. Proposed category (select one from below - see section 2.2 of P&P document):	
A-Contemporary <input type="checkbox"/> B.1-Specialized (small collection) <input checked="" type="checkbox"/> B.2-Specialized (large collection) <input type="checkbox"/>	
C-Major extinct <input type="checkbox"/> D-Attested extinct <input type="checkbox"/> E-Minor extinct <input type="checkbox"/>	
F-Archaic Hieroglyphic or Ideographic <input type="checkbox"/> G-Obscure or questionable usage symbols <input type="checkbox"/>	
4. Is a repertoire including character names provided?	<input checked="" type="checkbox"/> <b>Yes</b>
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<input checked="" type="checkbox"/> <b>Yes</b>
b. Are the character shapes attached in a legible form suitable for review?	<input checked="" type="checkbox"/> <b>Yes</b>
5. Fonts related:	
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	<i>Anshuman Pandey</i>
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	<i>Anshuman Pandey (pandey@umich.edu)</i>
6. References:	
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<input checked="" type="checkbox"/> <b>Yes</b>
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<input checked="" type="checkbox"/> <b>Yes</b>
7. Special encoding issues:	
Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?	<input checked="" type="checkbox"/> <b>Yes</b>

**8. Additional Information:**

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database ( <http://www.unicode.org/reports/tr44/> ) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

<sup>1</sup> Form number: N4102-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

**C. Technical - Justification**

1. Has this proposal for addition of character(s) been submitted before? If YES explain	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.)? If YES, with whom? If YES, available relevant documents:	Yes
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	Yes
4. The context of use for the proposed characters (type of use; common or rare) Reference:	Rare
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	Yes
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? 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? If YES, is a rationale for its inclusion provided? If YES, reference:	No
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference:	No
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character? If YES, is a rationale for its inclusion provided? If YES, reference:	No
11. Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a rationale for such use provided? If YES, reference: Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If YES, reference:	Yes Yes <i>Combining signs</i>
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary)	Yes <i>Virama;</i> <i>see text of the proposal</i>
13. Does the proposal contain any Ideographic compatibility characters? If YES, are the equivalent corresponding unified ideographic characters identified? If YES, reference:	No