

Title: Preliminary Proposal to Encode the Balti ‘B’ Script in ISO/IEC 10646
Source: Script Encoding Initiative (SEI)
Author: Anshuman Pandey (pandey@umich.edu)
Status: Liaison Contribution
Action: For consideration by UTC and WG2
Date: 2011-04-13

1 Introduction

This is a preliminary proposal to encode a writing system that is tentatively named Balti ‘B’. It was initially described in “Introducing Another Script for Writing Balti” (N3842 L2/10-231). This document provides additional information to assist in encoding the script.

2 Background

Balti (ISO 639: bft) (Tibetan སུ་བོ་ཏི་ *sBal-ti*; Urdu بَلْتِي *baltī*) is Tibeto-Burman language that belongs to the Western Tibetan sub-family, which also includes Ladakhi (lbj), Purik (prx), and Zangskari (zau). It is spoken primarily in the Gilgit-Baltistan region of northern Pakistan, which is also called Baltiyul or Balawaristan, and in neighboring Ladakh in India. There is no official script for Balti. The Tibetan script was introduced in the 8th century when the region was brought under Tibetan control.¹ The Arabic script replaced the Tibetan in the 17th century as the influence of Islam grew in the area. An extended form of Devanagari was developed for Balti by the Central Institute of Indian Languages in the 1970s.² In recent years Balti speakers have tried to revive usage of the Tibetan script, which they call ‘Yige’. In addition to these major scripts, at least two indigenous scripts were devised for writing Balti. These are labeled here as Balti ‘A’ and Balti ‘B’ for ease of reference. Balti ‘A’ is described by Rick McGowan in “Unicode Technical Report #3” (N2042), where it is simply named ‘Balti’ (see Figure 6). Both scripts are likely extinct. Formal names for these scripts have not yet been identified. A comparison of Balti ‘B’, Tibetan, and Arabic is shown in tables 2 and 3.

Balti ‘B’ is attested in a manuscript containing the text of ‘The Story of the Creation of Shigar’ by Wazir Ahmad. This appears to be the only existing record in the script. Photographs of the manuscript were published by Sydney Schuler in the *Central Asiatic Journal* (1978) and the text given in Latin transliteration and English translation. The manuscript consists of six folios, two of which are shown here in figures 4 and 5. It may be the only surviving example of a historical literary tradition in the Balti language. Two charts of the script have been published; the first by Schuler (Figure 2) and the other by Banat Gul Afridi (Figure 3). The versions of Balti ‘B’ shown by Schuler and Afridi display great similarity in character repertoire and glyph shapes. There are a few variations in the graphical structure of letters, but these are negligible. The letters shown by Schuler are derived from the Wazir Ahmad manuscript; the source for Afridi’s characters is unknown (see Table 1 for a comparison).

It is reported that the script was created by the grandfather of Wazir Ahmad, a man named Farhat.³ Additional information about its origins is unavailable. However, it is clear that Balti ‘A’ and Balti ‘B’ are unrelated. The two scripts differ considerably in structure, directionality, and appearance.

¹ Hasnain 2002.

² Rangan 1975.

³ Schuler 1978: 103.

3 Script Details

The script is tentatively named Balti ‘B’ in order to distinguish it from another Balti script described in UTR #3. Research is being done to identify an indigenous name. Balti ‘B’ consists of 46 characters: 30 consonant letters, 1 vowel-carrier letter, 4 vowel signs, a VIRAMA, and 10 digits. This repertoire is based upon an analysis of the available script charts and the folios of the Wazir Ahmad manuscript. A code chart and names list is provided in Figure 1. The traditional ordering for the script is unknown. Schuler shows an arrangement based upon the Latin order. Afridi shows an order that bears some resemblance to the Arabic order. The proposed encoding order is based upon neither of these, but is aligned with the order for Tibetan in the UCS. The font, which was designed by the author, is a work in progress.

4 Writing System

4.1 Structure

Balti ‘B’ is an alphasyllabic script based upon the Brahmi model. It is written left to right.

4.2 Vowels

The letter ࠁ A is a vowel carrier. It is used for writing independent forms of vowels, ie. when they appear in word-initial or isolated positions. The value of the vowel carrier is changed using combining vowel signs.

| | | | | | | | | |
|---|---|-----|-----|---------|---------------|-----|--------|-----|
| ࠁ | A | /a/ | ◌̇ | SIGN AA | /a:/ | ◌̈ | SIGN U | /u/ |
| | | | ◌̈̇ | SIGN I | /i/, /e/, /ɛ/ | ◌̈̈ | SIGN O | /o/ |

4.3 Virama

The sign ࠇ VIRAMA is used for silencing the inherent vowel of a consonant.

4.4 Consonants

Consonants possess the inherent vowel /a/. The attested letters are:

| | | | | | | | | |
|---|-----|------------|---|------|---------------|----|------|---------------|
| ཀ | KA | /k/ | ཁ | THA | /tʰ/ | ཙ | ZA | /z/ |
| ཁ | KHA | /kʰ/ | ད | DA | /d/ | ལ | YA | /j/, /i/, /e/ |
| ག | GA | /g/ | ཎ | NA | /n/, /ɲ/, /ŋ/ | ར | RA | /r/ |
| ཅ | CA | /tʃ/ | པ | PA | /p/ | ལ | LA | /l/ |
| ཆ | CHA | /tʃʰ/ | ཕ | PHA | /pʰ/, /f/ | ཤ | SHA | /ʃ/ |
| ཇ | JA | /dʒ/, /ts/ | བ | BA | /b/ | ས | SA | /s/ |
| མ | TTA | /t/ | མ | MA | /m/ | ཨ | HA | /h/ |
| ཙ | DDA | /d/ | ཅ | TSA | /ts/, /z/ | ཀྲ | QA | /q/ |
| ཛ | RRA | /r/ | ཅ | TSHA | /tsʰ/ | ཀླ | KHHA | /x/ |
| ཌ | TA | /t/ | ཇ | WA | /w/, /u/, /o/ | ཇ | GHHA | /ɣ/ |

4.5 Consonant Conjuncts

Consonant conjuncts are produced by writing a visible VIRAMA.

4.6 ‘Headline’

Each letter possesses a ‘headline’ or top-stroke. Unlike in Devanagari, the headlines of letter do not connect with those of surrounding letters. Some exception to this is to be expected in hand-written documents.

4.7 Digits

The digits represent decimal numbers.

4.8 Punctuation

Script-specific punctuation is not attested; marks similar to Latin punctuation are used.

4.9 Variant Forms

There are some variant forms for certain letters: ࠠ for ࠡ JA; ࠢ for ࠣ YA; ࠤ for ࠥ LA. Afridi shows ࠠ as the regular glyph for JA and assigns ࠡ as the independent letter for writing /tsʰ/. For the present, ࠠ is treated as a variant of ࠡ JA. These variant forms should be managed at the font level.

5 Implementation Details

5.1 Collation

The sort order follows the encoded order.

5.2 Character Properties

```
11CC0;BALTI-B LETTER A;Lo;0;L;;;;N;;;;;
11CC1;BALTI-B LETTER KA;Lo;0;L;;;;N;;;;;
11CC2;BALTI-B LETTER KHA;Lo;0;L;;;;N;;;;;
11CC3;BALTI-B LETTER GA;Lo;0;L;;;;N;;;;;
11CC4;BALTI-B LETTER CA;Lo;0;L;;;;N;;;;;
11CC5;BALTI-B LETTER CHA;Lo;0;L;;;;N;;;;;
11CC6;BALTI-B LETTER JA;Lo;0;L;;;;N;;;;;
11CC7;BALTI-B LETTER TTA;Lo;0;L;;;;N;;;;;
11CC8;BALTI-B LETTER DDA;Lo;0;L;;;;N;;;;;
11CC9;BALTI-B LETTER RRA;Lo;0;L;;;;N;;;;;
11CCA;BALTI-B LETTER TA;Lo;0;L;;;;N;;;;;
11CCB;BALTI-B LETTER THA;Lo;0;L;;;;N;;;;;
11CCC;BALTI-B LETTER DA;Lo;0;L;;;;N;;;;;
11CCD;BALTI-B LETTER NA;Lo;0;L;;;;N;;;;;
11CCE;BALTI-B LETTER PA;Lo;0;L;;;;N;;;;;
11CCF;BALTI-B LETTER PHA;Lo;0;L;;;;N;;;;;
11CD0;BALTI-B LETTER BA;Lo;0;L;;;;N;;;;;
11CD1;BALTI-B LETTER MA;Lo;0;L;;;;N;;;;;
11CD2;BALTI-B LETTER TSA;Lo;0;L;;;;N;;;;;
11CD3;BALTI-B LETTER TSHA;Lo;0;L;;;;N;;;;;
11CD4;BALTI-B LETTER WA;Lo;0;L;;;;N;;;;;
11CD5;BALTI-B LETTER ZA;Lo;0;L;;;;N;;;;;
11CD6;BALTI-B LETTER YA;Lo;0;L;;;;N;;;;;
11CD7;BALTI-B LETTER RA;Lo;0;L;;;;N;;;;;
11CD8;BALTI-B LETTER LA;Lo;0;L;;;;N;;;;;
```

```

11CD9;BALTI-B LETTER SHA;Lo;0;L;;;;N;;;;;
11CDA;BALTI-B LETTER SA;Lo;0;L;;;;N;;;;;
11CDB;BALTI-B LETTER HA;Lo;0;L;;;;N;;;;;
11CDC;BALTI-B LETTER QA;Lo;0;L;;;;N;;;;;
11CDD;BALTI-B LETTER KHHA;Lo;0;L;;;;N;;;;;
11CDE;BALTI-B LETTER GHHA;Lo;0;L;;;;N;;;;;
11CDF;BALTI-B VOWEL SIGN AA;Mn;0;NSM;;;;N;;;;;
11CE0;BALTI-B VOWEL SIGN I;Mn;0;NSM;;;;N;;;;;
11CE1;BALTI-B VOWEL SIGN U;Mn;0;NSM;;;;N;;;;;
11CE2;BALTI-B VOWEL SIGN O;Mn;0;NSM;;;;N;;;;;
11CE3;BALTI-B SIGN VIRAMA;Mn;9;NSM;;;;N;;;;;
11CE6;BALTI-B DIGIT ZERO;Nd;0;L;0;0;0;N;;;;;
11CE7;BALTI-B DIGIT ONE;Nd;0;L;1;1;1;N;;;;;
11CE8;BALTI-B DIGIT TWO;Nd;0;L;2;2;2;N;;;;;
11CE9;BALTI-B DIGIT THREE;Nd;0;L;3;3;3;N;;;;;
11CEA;BALTI-B DIGIT FOUR;Nd;0;L;4;4;4;N;;;;;
11CEB;BALTI-B DIGIT FIVE;Nd;0;L;5;5;5;N;;;;;
11CEC;BALTI-B DIGIT SIX;Nd;0;L;6;6;6;N;;;;;
11CEF;BALTI-B DIGIT SEVEN;Nd;0;L;7;7;7;N;;;;;
11CEE;BALTI-B DIGIT EIGHT;Nd;0;L;8;8;8;N;;;;;
11CEF;BALTI-B DIGIT NINE;Nd;0;L;9;9;9;N;;;;;

```

6 References

- Afridi, Banat Gul. 1988. *Baltistan in History*. Peshawar: Emjay Books International.
- (M.) Hasnain Sengge Thsering. 2002. “Key Issues in Contemporary Balti Language and Script.” Paper presented at the Eighth Himalayan Languages Symposium, University of Bern - Institute for Linguistics, September 2002. Electronic version at http://india_resource.tripod.com/baltiyul.html. Accessed July 2010.
- McGowan, Rick. 1997. “Unicode Technical Report #3: Early Aramaic, Balti, Kirat (Limbu), Manipuri (Meitei), and Tai Lü scripts.” (N2042 1997-07-20). <http://std.dkuug.dk/jtc1/SC2/wg2/docs/n2042.pdf>
- Pandey, Anshuman. 2010. “Introducing Another Script for Writing Balti” (N3842 L2/10-231). <http://std.dkuug.dk/jtc1/SC2/wg2/docs/n3842.pdf>
- Rangan. K. 1975. *Balti Phonetic Reader*. CIIL Phonetic Reader Series, no. 17. Mysore: Central Institute of Indian Languages.
- Schuler, Sidney. 1978. “The ‘Story of the Creation of Shigar’ of Wazir Ahmad.” *Central Asiatic Journal*, vol. XXII, no. 1–2, pp. 102–120.

7 Acknowledgments

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 of the Humanities.

| | 11CC | 11CD | 11CE |
|---|------------|------------|------------|
| 0 | 𑆀 11CC0 | 𑆁 11CD0 | 𑆂 11CE0 |
| 1 | 𑆃 11CC1 | 𑆄 11CD1 | 𑆅 11CE1 |
| 2 | 𑆆 11CC2 | 𑆇 11CD2 | 𑆈 11CE2 |
| 3 | 𑆉 11CC3 | 𑆊 11CD3 | 𑆋 11CE3 |
| 4 | 𑆌 11CC4 | 𑆍 11CD4 | |
| 5 | 𑆎 11CC5 | 𑆏 11CD5 | |
| 6 | 𑆐 11CC6 | 𑆑 11CD6 | 𑆒 11CE6 |
| 7 | 𑆓 11CC7 | 𑆔 11CD7 | 𑆕 11CE7 |
| 8 | 𑆖 11CC8 | 𑆗 11CD8 | 𑆘 11CE8 |
| 9 | 𑆙 11CC9 | 𑆚 11CD9 | 𑆛 11CE9 |
| A | 𑆜 11CCA | 𑆝 11CDA | 𑆞 11CEA |
| B | 𑆟 11CCB | 𑆠 11CDB | 𑆡 11CEB |
| C | 𑆣 11CCC | 𑆤 11CDC | 𑆥 11CEC |
| D | 𑆧 11CCD | 𑆨 11CDD | 𑆩 11CED |
| E | 𑆪 11CCE | 𑆫 11CDE | 𑆬 11CEE |
| F | 𑆮 11CCF | 𑆯 11CDF | 𑆰 11CEF |

Vowel

11CC0 𑆀 BALTI-B LETTER A

Consonants

11CC1 𑆃 BALTI-B LETTER KA
11CC2 𑆆 BALTI-B LETTER KHA
11CC3 𑆉 BALTI-B LETTER GA
11CC4 𑆊 BALTI-B LETTER CA
11CC5 𑆌 BALTI-B LETTER CHA
11CC6 𑆐 BALTI-B LETTER JA
11CC7 𑆓 BALTI-B LETTER TTA
11CC8 𑆖 BALTI-B LETTER DDA
11CC9 𑆙 BALTI-B LETTER RRA
11CCA 𑆜 BALTI-B LETTER TA
11CCB 𑆟 BALTI-B LETTER THA
11CCC 𑆣 BALTI-B LETTER DA
11CCD 𑆧 BALTI-B LETTER NA
11CCE 𑆪 BALTI-B LETTER PA
11CCF 𑆮 BALTI-B LETTER PHA
11CD0 𑆁 BALTI-B LETTER BA
11CD1 𑆄 BALTI-B LETTER MA
11CD2 𑆇 BALTI-B LETTER TSA
11CD3 𑆉 BALTI-B LETTER TSHA
11CD4 𑆍 BALTI-B LETTER WA
11CD5 𑆏 BALTI-B LETTER ZA
11CD6 𑆑 BALTI-B LETTER YA
11CD7 𑆔 BALTI-B LETTER RA
11CD8 𑆗 BALTI-B LETTER LA
11CD9 𑆚 BALTI-B LETTER SHA
11CDA 𑆝 BALTI-B LETTER SA
11CDB 𑆠 BALTI-B LETTER HA
11CDC 𑆤 BALTI-B LETTER QA
11CDD 𑆨 BALTI-B LETTER KHHA
11CDE 𑆫 BALTI-B LETTER GHHA

Vowel signs

11CDF 𑆯 BALTI-B SIGN AA
11CE0 𑆂 BALTI-B SIGN I
11CE1 𑆅 BALTI-B SIGN U
11CE2 𑆈 BALTI-B SIGN O

Various signs

11CE3 𑆋 BALTI-B SIGN VIRAMA

Digits

11CE6 𑆒 BALTI-B DIGIT ZERO
11CE7 𑆕 BALTI-B DIGIT ONE
11CE8 𑆘 BALTI-B DIGIT TWO
11CE9 𑆛 BALTI-B DIGIT THREE
11CEA 𑆞 BALTI-B DIGIT FOUR
11CEB 𑆡 BALTI-B DIGIT FIVE
11CEC 𑆥 BALTI-B DIGIT SIX
11CED 𑆩 BALTI-B DIGIT SEVEN
11CEE 𑆬 BALTI-B DIGIT EIGHT
11CEF 𑆰 BALTI-B DIGIT NINE

Figure 1: Proposed code chart and names list for Balti-B

| | | | | | |
|---------|----|-------|---------|-----|----------|
| a | ا | آ (ہ) | l | ل | ل |
| b | ب | ب | m | م | م |
| c | چ | چ | n | ن | ن |
| ch | چھ | چھ | o, u, w | و | و (و, و) |
| d | د | د | p | پ | پ |
| D | ڈ | ڈ | r | ر | ر |
| e, i, y | ی | ی (ہ) | R | ڑ | ڑ |
| f, ph | ف | ف | s | س | س |
| g | گ | گ | š | ش | ش |
| γ | غ | غ | t | ت | ت |
| h | ھ | ھ | th | تھ | تھ |
| j | ج | ج | T | ٹ | ٹ |
| k | ک | ک | tsh | تسھ | تسھ |
| kh | کھ | کھ | z | ز | ز, ر |
| x | خ | خ | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|----|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| ✓ | ۲ | ۳ | ۴ | ۵ | ۶ | ۷ | ۸ | ۹ | ۱۰ |

Figure 2: Chart of Balti 'B' (from Schuler 1978: 120).

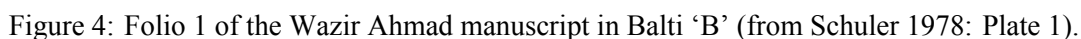
جا خا نا پا با آ
چا خر ٹا ڈا ڈا
را رازا زارزا
خا فا یا نا لا وا شا
ٹا ڈا ڈا ڈا ڈا

پیش زبر زیر

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 |
|---|---|---|---|---|---|---|---|---|---|
| م | و | ح | ن | ج | خ | ع | ح | س | |

ਗੁਰੇਹ ਕ ਰਾਜੇ ਜੋ ਲਗ ਸਿੰਧ ਤੇ ਮਿਲਾਤ
 ਵੇਖੇਯਾ ਤੇ ਘਟੇ, ਜੇਘਰੇ ਕੇਦਰ, ਤੇ ਮੇਰੇ
 ਫਿਰੇ ਫਿਰੇਯਾਯੋ ਭੇਦਰ.

One school of thought is of the view that only Hindi is best suited for the writing of Balti language, while the other prefers Roman script. It is for the experts in languages to evolve a suitable script.





| | AFRIDI | SCHULER | | AFRIDI | SCHULER | | AFRIDI | SCHULER |
|--------------------|--------|---------|--------------------|--------|---------|------|--------|---------|
| /a/ | ⬤ | ? | /b/ | ⬤ | ⬤ | /a:/ | ⬤ | □ |
| /k/ | ⬤ | ⬤ | /m/ | ⬤ | ⬤ | /i/ | ⬤ | □ |
| /k ^h / | ⬤ | ⬤ | /ts/ | ⬤ | — | /u/ | ⬤ | □ |
| /g/ | ⬤ | ⬤ | /ts ^h / | ⬤ | ⬤ | /o/ | — | □ |
| /tʃ/ | ⬤ | ⬤ | /w/ | ⬤ | ⬤ | ∅ | ⬤ | — |
| /tʃ ^h / | ⬤ | ⬤ | /z/ | ⬤ | ⬤, ⬤ | | | |
| /dʒ/ | ⬤ | ⬤ | /j/ | ⬤ | ⬤ | 0 | — | ⬤ |
| /t/ | ⬤ | ⬤ | /r/ | ⬤ | ⬤ | 1 | ⬤ | ⬤ |
| /d/ | ⬤ | ⬤ | /l/ | ⬤ | ⬤ | 2 | ⬤ | ⬤ |
| /t̪/ | ⬤ | ⬤ | /ʃ/ | ⬤ | ⬤ | 3 | ⬤ | ⬤ |
| /t/ | ⬤ | ⬤ | /s/ | ⬤ | ⬤ | 4 | ⬤ | ⬤ |
| /t ^h / | ⬤ | ⬤ | /h/ | ⬤ | ⬤ | 5 | ⬤ | ⬤ |
| /d/ | ⬤ | ⬤ | /q/ | ⬤ | — | 6 | ⬤ | ⬤ |
| /n/ | ⬤ | ⬤ | /x/ | ⬤ | ⬤ | 7 | ⬤ | ⬤ |
| /p/ | ⬤ | ⬤ | /ɣ/ | ⬤ | ⬤ | 8 | ⬤ | ⬤ |
| /p ^h / | ⬤ | ⬤ | | | | 9 | ⬤ | ⬤ |

Table 1: Comparison of Balti ‘B’ characters shown by Afridi and Schuler. The ● used by Afridi and the □ used by Schuler for showing vowel signs are dummy characters akin to ○ dotted circle.

| | BALTI B | TIBETAN | ARABIC | | BALTI B | TIBETAN | ARABIC |
|-----|---------|---------|--------|------|---------|---------|--------|
| A | ཨ | ཨ | ا | BA | བ | བ | ب |
| KA | ཀ | ཀ | ك | MA | མ | མ | م |
| KHA | ཁ | ཁ | ڪ | TSA | ཅ | ཅ | ث |
| GA | ཁ | ག | گ | TSHA | ཇ | ཇ | ظ |
| CA | ཅ | ཅ | چ | WA | ཊ | ཊ | و |
| CHA | ཆ | ཆ | چھ | ZA | ཐ | ཐ | ز |
| JA | ཇ | ཇ | ج | YA | ཌ | ཌ | ي |
| TTA | མ | མ | ط | RA | ར | ར | ر |
| DDA | ཙ | ཙ | ڌ | LA | ལ | ལ | ل |
| RRA | ཛ | ཛ | ڙ | SHA | ཤ | ཤ | ش |
| TA | ཏ | ཏ | ت | SA | ས | ས | س |
| THA | ཐ | ཐ | تھ | HA | ཏ | ཏ | ه |
| DA | ཌ | ཌ | د | QA | འ | འ | ق |
| NA | ཎ | ཎ | ن | KHHA | ཁ | ཁ | خ |
| PA | པ | པ | پ | GHHA | ག | ག | غ |
| PHA | ཕ | ཕ | پھ | | | | |

Table 2: Correspondences between consonant letters of Balti ‘B’, Tibetan, and Arabic.

| | BALTI B | TIBETAN | ARABIC | | BALTI B | TIBETAN | ARABIC |
|--------|---------|---------|--------|-------|---------|---------|--------|
| SIGN A | ◌́ | ◌ྐ | ◌َ | ZERO | • | 0 | ◌◌ |
| SIGN I | ◌̇ | ◌ྑ | ◌ِ | ONE | ◌ | ◌ | ◌ |
| SIGN U | ◌̈ | ◌ྒ | ◌ُ | TWO | ◌ | ◌ | ◌ |
| SIGN O | ◌̊ | ◌ྒྷ | — | THREE | ◌ | ◌ | ◌ |
| VIRAMA | ◌̣ | ◌ྔ | ◌ْ | FOUR | ◌ | ◌ | ◌ |
| | | | | FIVE | ◌ | ◌ | ◌ |
| | | | | SIX | ◌ | ◌ | ◌ |
| | | | | SEVEN | ◌ | ◌ | ◌ |
| | | | | EIGHT | ◌ | ◌ | ◌ |
| | | | | NINE | ◌ | ◌ | ◌ |

Table 3: Correspondences between vowel signs and digits of Balti ‘B’, Tibetan, and Arabic.

Balti

The Balti script is now extinct, but was formerly used to write the Balti language of Baltistan, in what is now part of Ladakh in Northern Kashmir. The script was apparently introduced in about the fifteenth century CE when the people converted to Islam. It is related to the Arabic script.

In contrast to many other Brahmic scripts, Balti is written from *right to left* horizontally, in the Arabic manner. All of the vowel signs except long *a* are integrated into the glyphs used for consonants, becoming projections from the consonants rather than being separate marks as in most of the modern Brahmic scripts. The consonants apparently have an inherent *a* vowel (or an explicit vowel sign *a* may appear; there may not be a distinction between long and short *a*). There appears to be a sign (overdot) used to indicate the end of a word, but no interword spacing seems to be used.

The base form of *b* is the same as *p* and *t*; only the dots distinguish these. There are two other similar pairs. These appear to approximately parallel similar dotted versus dotless letters in Arabic.

Issues: The set of Balti consonants is too small to make it worth encoding parallel to any of the other Brahmic scripts, or to Arabic. Not enough information is available at this time to determine the completeness of the accompanying chart. The digits, if any, are unknown. It is unknown how much literature is available in the old Balti script, or what the level of scholarly interest in it is. The function of the character listed in the names list as “Balti null vowel or word ending” is uncertain.

Grierson, G. A. *Linguistic Survey of India*, Vol. 3.

One photocopy of 2 pages (326 and 327) from an unknown volume in German.

Balti Names, draft 1999-07-20

| | |
|---------------------|-----------------------------|
| 00 BALTI LETTER A | 10 BALTI LETTER NA |
| 01 BALTI LETTER BA | 11 BALTI LETTER HA |
| 02 BALTI LETTER PA | 12 BALTI LETTER JA |
| 03 BALTI LETTER TA | 13 BALTI LETTER KHA |
| 04 BALTI LETTER GA | 14 BALTI LETTER THA |
| 05 BALTI LETTER HHA | 15 BALTI LETTER TSA |
| 06 BALTI LETTER CA | 16 BALTI LETTER NGA |
| 07 BALTI LETTER CHA | 17 BALTI VOWEL SIGN A |
| 08 BALTI LETTER DA | 18 BALTI VOWEL SIGN AA |
| 09 BALTI LETTER RA | 19 BALTI VOWEL SIGN E |
| 0A BALTI LETTER ZA | 1A BALTI VOWEL SIGN I |
| 0B BALTI LETTER SA | 1B BALTI VOWEL SIGN O |
| 0C BALTI LETTER SHA | 1C BALTI VOWEL SIGN U |
| 0D BALTI LETTER KA | 1D BALTI NULL VOWEL OR WORD |
| 0E BALTI LETTER LA | ENDING? |
| 0F BALTI LETTER MA | |

| | 0 | 1 |
|---|---|---|
| 0 | 𑂀 | 𑂁 |
| 1 | 𑂂 | 𑂃 |
| 2 | 𑂄 | 𑂅 |
| 3 | 𑂆 | 𑂇 |
| 4 | 𑂈 | 𑂉 |
| 5 | 𑂊 | 𑂋 |
| 6 | 𑂌 | 𑂍 |
| 7 | 𑂎 | 𑂏 |
| 8 | 𑂐 | 𑂑 |
| 9 | 𑂒 | 𑂓 |
| A | 𑂔 | 𑂕 |
| B | 𑂖 | 𑂗 |
| C | 𑂘 | 𑂙 |
| D | 𑂚 | 𑂛 |
| E | 𑂜 | |
| F | 𑂞 | |

Page 3

Figure 6: Report on ‘Balti A’ as given in N2042 (from McGowan 1997: 3).

**ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹**

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: | Preliminary Proposal to Encode the Balti 'B' Script in ISO/IEC 10646 |
| 2. Requester's name: | <i>Script Encoding Initiative (SEI) / Anshuman Pandey <pandey@umich.edu></i> |
| 3. Requester type (Member body/Liaison/Individual contribution): | <i>Liaison contribution</i> |
| 4. Submission date: | <i>2011-04-09</i> |
| 5. Requester's reference (if applicable): | |
| 6. Choose one of the following: | |
| This is a complete proposal: | |
| (or) More information will be provided later: | <i>Yes</i> |

B. Technical – General

| | |
|---|--|
| 1. Choose one of the following: | |
| a. This proposal is for a new script (set of characters): | <i>Yes</i> |
| Proposed name of script: | <i>Balti 'B'</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>46</i> |
| 3. Proposed category (select one from below - see section 2.2 of P&P document): | |
| A-Contemporary | <input checked="" type="checkbox"/> B.1-Specialized (small collection) |
| C-Major extinct | <input checked="" type="checkbox"/> B.2-Specialized (large collection) |
| D-Attested extinct | <input type="checkbox"/> E-Minor extinct |
| F-Archaic Hieroglyphic or Ideographic | <input type="checkbox"/> G-Obscure or questionable usage symbols |
| 4. Is a repertoire including character names provided? | <i>Yes</i> |
| a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document? | <i>Yes</i> |
| b. Are the character shapes attached in a legible form suitable for review? | <i>Yes</i> |
| 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? | <i>Yes</i> |
| b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? | <i>Yes</i> |
| 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)? | <i>Yes,</i> |

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.

¹ Form number: N3702-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)

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: | No |
| 3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference: | Unknown |
| 4. The context of use for the proposed characters (type of use; common or rare) Reference: | Rare See text of proposal. |
| 5. Are the proposed characters in current use by the user community? If YES, where? Reference: | Unknown |
| 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: | No |
| 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 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 Combining vowel signs |
| 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 Virama; |
| 13. Does the proposal contain any Ideographic compatibility character(s)? If YES, is the equivalent corresponding unified ideographic character(s) identified? If YES, reference: | No |