ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]
Please fill all the sections A, B and C below.
Please read Principles and Procedures Document (P & P) from http://std.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://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html**.**
See also http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

A. Administrative

1. **Title:** *BENGALI LETTER UA USED IN KOKBOROK*
2. Requester's name: *Sanghmitra Sahu (sanghmitra.sahu@gmail.com)*
3. Requester type (Member body/Liaison/Individual contribution): *Individual*
4. Submission date: *12 February 2015*
5. Requester's reference (if applicable):
6. Choose one of the following:
    This is a complete proposal: *complete*
    (or) More information will be provided later: *-*

B. Technical – General

1. Choose one of the following:
    a. This proposal is for a new script (set of characters): *No*
        Proposed name of script: *-*
    b. The proposal is for addition of character(s) to an existing block: *Yes*
        Name of the existing block: *BMP, Bengali (U+0980 to U+09FF)*
2. Number of characters in proposal: *one*
3. Proposed category (select one from below - see section 2.2 of P&P document):
    A-Contemporary *Y*   B.1-Specialized (small collection)   B.2-Specialized (large collection)
    C-Major extinct   D-Attested extinct   E-Minor extinct
    F-Archaic Hieroglyphic or Ideographic   G-Obscure or questionable usage symbols
4. Is a repertoire including character names provided? *Yes*
    a. If YES, are the names in accordance with the "character naming guidelines"
        in Annex L of P&P document? *Y*
    b. Are the character shapes attached in a legible form suitable for review? *Yes*
5. Fonts related:
    a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?
        Samples in the enclosure
    b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
        Samples in the enclosure
6. References:
    a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? *Yes*
    b. Are published examples of use (such as samples from newspapers, magazines, or other sources)
    of proposed characters attached? *Yes*
7. Special encoding issues:
    Does the proposal address other aspects of character data processing (if applicable) such as input,
    presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)? *Yes*
    Enclosed

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.

---

[1] Form number: N4502-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?   *No*
    If YES explain   *-*

2. Has contact been made to members of the user community (for example: National Body,
    user groups of the script or characters, other experts, etc.)?   *yes*
      If YES, with whom?   *www.tritripura.in , www.ttaadc.gov.in, www.kohm.twipra.com/*
      If YES, available relevant documents: *Kokborok books, writings in Bengali script*

3. Information on the user community for the proposed characters (for example:
    size, demographics, information technology use, or publishing use) is included?   *Tripura etc*
    Reference: *http://en.wikipedia.org/wiki/Kokborok*

4. The context of use for the proposed characters (type of use; common or rare)   *common*
    Reference: *See enclosure*

5. Are the proposed characters in current use by the user community?   *yes*
    If YES, where? Reference: *Tripura (India) and nearby areas (including Bangladesh)*

6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely
    in the BMP?   *Yes*
      If YES, is a rationale provided?   *Yes*
      If YES, reference: *Enclosed*

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?   *Sometimes yes*
      If YES, is a rationale for its inclusion provided?   *Yes*
      If YES, reference: *Enclosed*

9. Can any of the proposed characters be encoded using a composed character sequence of either
    existing characters or other proposed characters?   *Yes*
      If YES, is a rationale for its inclusion provided?   *Yes*
      If YES, reference: *Enclosed*

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?   *No*
      If YES, is a rationale for its inclusion provided?   *-*
      If YES, reference: *-*

11. Does the proposal include use of combining characters and/or use of composite sequences?   *No*
    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: *-*

12. Does the proposal contain characters with any special properties such as
    control function or similar semantics?   *no*
      If YES, describe in detail (include attachment if necessary)   *-*

13. Does the proposal contain any Ideographic compatibility characters?   *No*
    If YES, are the equivalent corresponding unified ideographic characters identified?   *-*
      If YES, reference: *-*

ENCLOSURE

KOKBOROK LANGUAGE:

Kokborok is a modern and living language spoken by a number of tribes in Tripura (a state in India) and by members of these tribes living in neighbouring areas (Assam, Mizoram and parts of Bangladesh). The language is spoken by Tripura, Reang/Bru, Jamatia, Rupini etc tribes. It is part of Bodo group of languages.

OFFICIAL STATUS OF KOKBOROK LANGUAGE:

Kokborok is the second official language of Tripura since 1979. There currently is a debate over giving the language recognition as a National language of India.

SCRIPT OF KOKBOROK:

At present Kokborok is officially written in a modified Bengali script. The modifications are on the sound "wa" as in "water", use of apostrophe to denote a light "a" sound at the end of the word (sometimes, despite the implicit/explicit hasant U+09CD), and a special vowel and vowel sign ("w"). It is also written in Latin script by some groups.

THE PRESENT PROPOSAL:

Spoken Kokborok has a sound similar to Bengali word ওয়া (U+0993 U+09DF U+09BE) or Hindi word वा (U+0935 and U+093E). It is written in Kokborok modified Bengali script as looking like উা or U+0989 and U+09BE. It has been called here "*ua*". It is pronounced similar to "wa" in English word "water." It is almost equivalent to the sound वा (U+0935 and U+093E) in Hindi. It could be written as Assamese ৱা (U+09F1 and U+09BE). But all Kokborok books use the letter উা.

In Kokborok grammar, this letter is a vowel. However, sound-wise it is closer to a consonant. This letter (উা) does not take any vowel sign as উাা, উাু etc (at least in the current Kokborok vocabulary), whereas consonants (ক, খ etc) may take a vowel sign like া, ি, ু etc (as কা, খু etc).

HOW THE LETTER WRITTEN IN LATIN SCRIPT:

Even though Kokborok is officially written in modified Bengali script, some people also write the language in Latin script. For example, English-Kokborok dictionary at http://kokrwbam.twipra.com is in Latin script.

In Latin script, the sound is written in Kokborok as "*ua*". For example, search Kokborok word for English word "pig" at website http://kokrwbam.twipra.com. It is "*uak*", i.e., উাক. ("*Uaksa*" means "one pig", "*uak bwsa*" means "piglet ('pig cub')" etc.)

EXAMPLES:

Examples from Kokborok textbooks are attached separately in PDF format. These are textbooks published by government agencies. These books are recent publications, published in December 2014 CE. English version of the attached Kokborok text is also separately attached in PDF format. The translation has been done by Translation Cell of General Administration Department (Government of Tripura, Agartala, India).

Examples of "Bengali letter Ua" as used in Kokborok.

The examples (circled or underlined) below are from the textbooks of class I and class II standards, prescribed and published by government agencies.

Figures 1-3 are from *Puila Parima - Class I Kokborok* textbook
Reference: *Puila Parima (First Reading) Class One. A textbook for Class- I* by Tripura Tribal Areas Autonomous District Countil Laxmidhan Murasing. Ninth edition. December, 2014. Publisher: Director, State Council of Educational Research & Training, Tripura. Printer: Hooghly Printing Company Limited (Government of India enterprise), 41, Chowringhee Road, Kolkata. (In Kokborok)
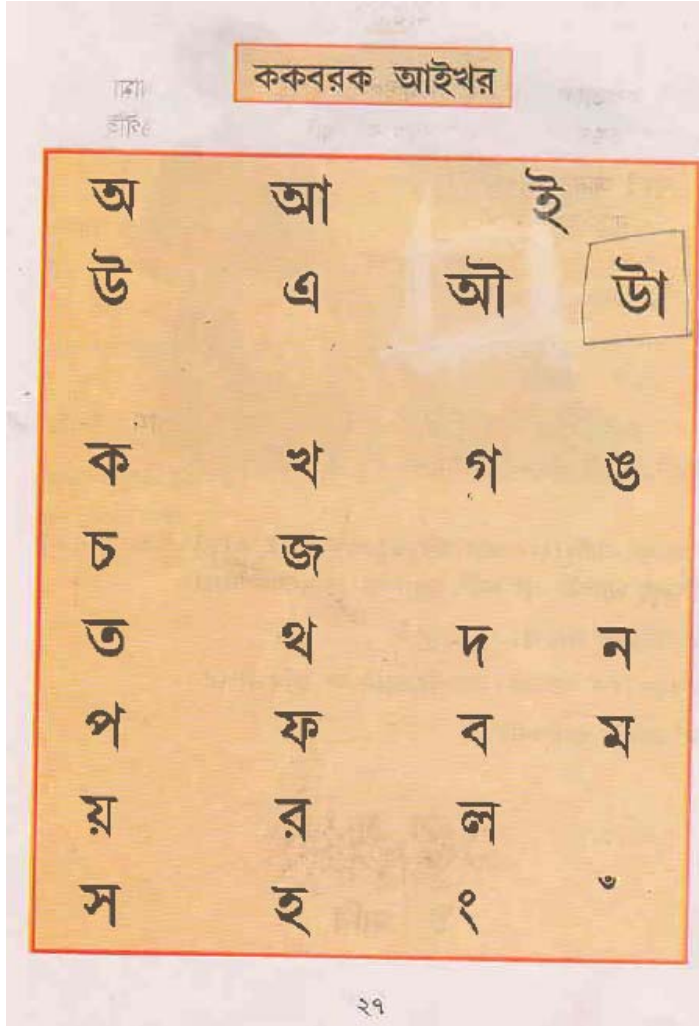

Fig. 1: page 27

ডাক

ডাক হা উর'।
হা উরাই কেনজুডা চাঅ।

Fig. 2: page 21

তখা পুঙগ কা — কা।
তখা মগদাম চাখা।

| উা | ৹ | প |

পরিদি ঃ

ডাক, কেনজুডা, পুমা, পুঙগ, বাসা।

রীঙরীকমুঙ

১। তলাঅ ককরগনি বিসিংনি উা আইঝরনি তলাঅ রিকদি।
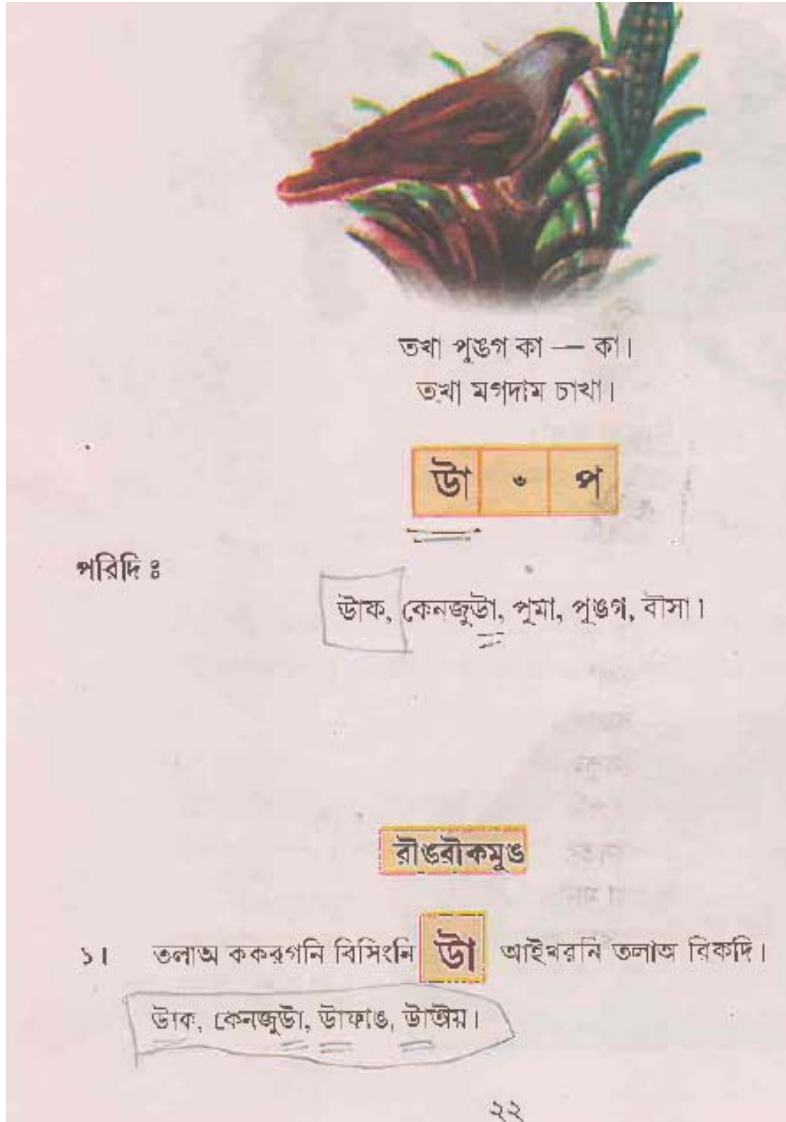
ডাক, কেনজুডা, উাফাঙ, উাঞ্জম।

২২

Fig. 3: page 22

Figure 4 is from *Larima - Class II Kokborok* textbook.

Reference: *Larima (Butterfly) Class Two. A textbook for class- II* by Tripura Tribal Areas Autonomous District Council—Laxmidhan Murasing, Shyamlal Debbarma. Ninth edition, December 2014.
Publisher: Director, SCERT. Printer: Srishti Graphics, 3/2 Matheshwartala Road, Kolkata-700046.
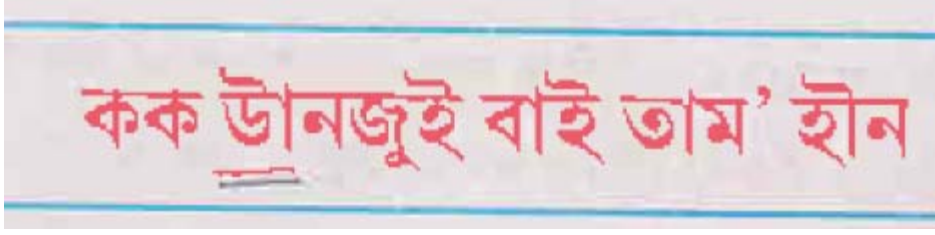(In Kokborok)

কক ডাংজুই বাই তাম' হীন

Fig. 4: page 47