Title: Proposal to Encode the Marwari Letter DDA for Devanagari
Source: Anshuman Pandey (pandey@umich.edu)
Status: Individual Contribution
Action: For consideration by UTC and WG2
Date: 2011-05-23

1   Introduction

This is a proposal to encode a new character in the Universal Character Set: ꞓ DEVANAGARI LETTER MARWARI DDA. It is used for writing /ɖ/ in the historical Devanagari orthography for Marwari (ISO 639: rwr), Kutchi (ISO 639: kfr), and other Rajasthani and Sindhi languages.

2   Character Details

Properties of the proposed character are:

GLYPH	CODE	CHARACTER NAME	GC	CCC	BIDI	MIRRORED
ꞓ	0978	DEVANAGARI LETTER MARWARI DDA	Lo	0	L	N

3   Background

In Devanagari, the phoneme /ɖ/ is represented using ड U+0921 DEVANAGARI LETTER DDA. In Hindi and related languages, /ɖ/ has an allophonic variant /ɽ/, which is written using ड़ U+095C DEVANAGARI LETTER DDDHA. This letter is produced by combining ड with ़ NUKTA. Historically, in Devanagari orthography for Marwari, /ɖ/ is represented by the letter ꞓ and its allophone /ɽ/ is written using ड.

The letter ꞓ is attested in printed (see figures 1 and 2) and written documents (see figures 3 and 4). It was used in printing until the early 20th century and may still be used in writing. The character belongs to orthographies for regional languages and because it is not used in standard Devanagari, it was not generally included in common Devanagari metal fonts. As a result, ꞓ was often inappropriately substituted with म MA when the letter was unavailable in a font (Grierson 1908: 20). The absence of support for typesetting the letter is a reason for its diminished use.

Information about the origins of the character is unavailable. But, based upon graphical typology, it is plausible that the character is derived from 𑆝 U+1119D SHARADA LETTER DDA. It may have entered into Devanagari through the Landa scripts used in Rajasthan and Sindh. The Landa scripts are related to Sharada and possess some variant of the form '𝟑' for writing /ɖ/. This form is the basis for ਡ U+0A21 GURMUKHI LETTER DDA, 𑚖 U+11696 TAKRI LETTER DDA, and 𑋈 U+112C8 KHUDAWADI LETTER DDA.

While ꞓ is used for writing /ɖ/, it cannot be treated as a glyphic variant of the existing Devanagari ड DDA. It is a distinct character that is used in conjunction with ड in older texts in Marwari, Kutchi, and other languages. An independent encoding for ꞓ will enable its representation in plain text.

4   Implementation Details

4.1   Name

The proposed character is assigned the name DEVANAGARI LETTER MARWARI DDA. This name seems more appropriate than other possible candidates, eg. DDDDA, etc.

4.2   Glyph Shape

The proposed glyph shape ड़ was drawn to conform to the design of Monotype Devanagari. There are two variant forms: ड़ (figure 1) and ड़ (figure 2).

4.3   Collation

The proposed character should be sorted before ड U+0921 DEVANAGARI LETTER DDA.

5   References

Bagster, Samuel. 1848. *The Bible of Every Land: A history of the Sacred Scriptures in every language and dialect into which translations have been made : illustrated with specimen portions in native characters, series of alphabets, coloured ethnographical maps, tables, indexes, etc.* London: Samuel Bagster and Sons.

Grierson, George A. 1908. *The Linguistic Survey of India*. Vol. IX. Indo-Aryan Family. Central Group. Part II. Specimens of the Rājasthānī and Gujarātī. Calcutta: Office of the Superintendent of Government Printing, India.

In writing Mārwāṛī the letters *ḍ* and *ṛ* are distinguished. This is not done as in Hindī by representing *ṛ* by the sign for *ḍ* with a dot placed under it, thus ड़. In Mārwāṛī there are two distinct characters, *viz.*, ॸ for *ḍ* and ड़ for *ṛ*. In printed Mārwāṛī, when the type for ॸ is not available, the Dēvanāgarī letter म is often substituted for it, which is very inconvenient ; as, in that case, for instance, only the context can tell whether बमो represents '*baḍō*' or '*bamō*.' In printing the Mārwāṛī specimens in the Dēvanāgarī character, I shall represent *ḍ* by ड़ and *ṛ* by ड़.

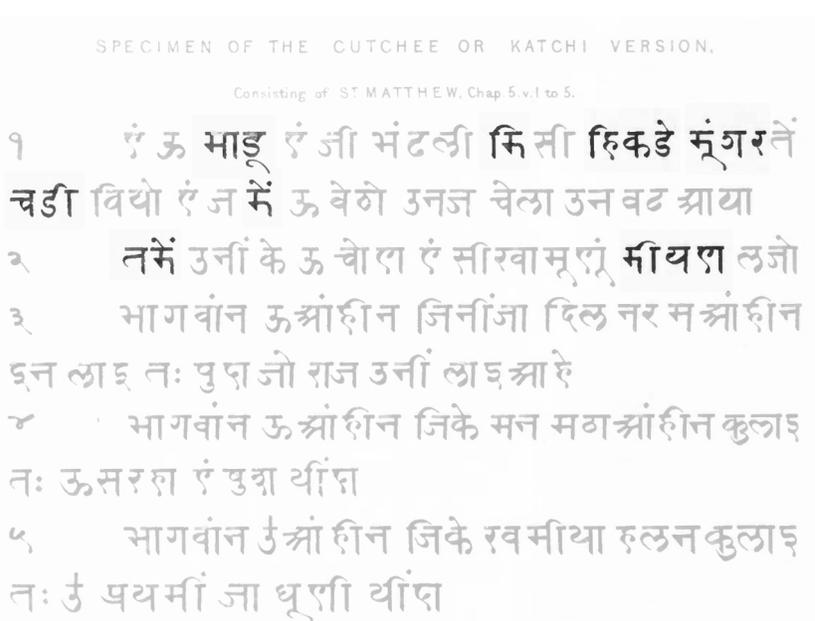Figure 1: Description of the letter ॸ as used for writing Marwari (from Grierson 1908: 20). A variant form ॸ is shown here.

Figure 2: Excerpt from Bagster (1848: Plate V) showing the use of ड़ DDA for /ɽ/ and ॸ MARWARI DDA for /ḍ/ in printed Kutchi text; eg. the words हिकडे /hɪkɽe/ 'one' and डूंगर /ḍuŋɡər/ 'mountain' in the first line of the text. The angular form ॸ is used in the original source.

Figure 3: Text written in the Marwari style of Devanagari (from Grierson 1908: 63). The highlighted word is ॸावडा /ḍavɽa/ 'sons'.

Figure 4: Text written in the Marwari style of Devanagari (from Grierson 1908: 131). The highlighted word is मंॸोर /məndor/ 'Mandor', a city in the state of Rajasthan, India.

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://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:** *Proposal to Encode the Marwari Letter DDA for Devanagari*
2. Requester's name: *Anshuman Pandey <pandey@umich.edu>*
3. Requester type (Member body/Liaison/Individual contribution): *Individual contribution*
4. Submission date: *2011-05-23*
5. Requester's reference (if applicable):
6. Choose one of the following:
   This is a complete proposal: *Yes*
   (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):
      Proposed name of script:
   b. The proposal is for addition of character(s) to an existing block: *Yes*
      Name of the existing block: *Devanagari*
2. Number of characters in proposal: *1*
3. Proposed category (select one from below - see section 2.2 of P&P document):
   A-Contemporary   **X**   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? *Yes*
   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?
      Anshuman Pandey
   b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
      Anshuman Pandey <pandey@umich.edu>
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,*

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.

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.)?   *No*
   - If YES, with whom?
   - If YES, available relevant documents:
3. Information on the user community for the proposed characters (for example:
   size, demographics, information technology use, or publishing use) is included?   *Yes*
   - Reference: *See text of proposal.*
4. The context of use for the proposed characters (type of use; common or rare)   *Common*
   - Reference: *See text of proposal.*
5. Are the proposed characters in current use by the user community?   *Yes*
   - If YES, where? Reference: *See text of proposal.*
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely
   in the BMP?   *No*
   - If YES, is a rationale provided?
   - If YES, reference:
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?   *N/A*
8. Can any of the proposed characters be considered a presentation form of an existing
   character or character sequence?   *No*
   - If YES, is a rationale for its inclusion provided?
   - If YES, reference:
9. Can any of the proposed characters be encoded using a composed character sequence of either
   existing characters or other proposed characters?   *No*
   - If YES, is a rationale for its inclusion provided?
   - If YES, reference:
10. Can any of the proposed character(s) be considered to be similar (in appearance or function)
    to an existing character?   *Yes*
    - If YES, is a rationale for its inclusion provided?   *Yes*
    - If YES, reference: *See text of proposal.*
11. Does the proposal include use of combining characters and/or 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 character(s)?   *No*
    - If YES, is the equivalent corresponding unified ideographic character(s) identified?
    - If YES, reference: