

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: 2010-12-08

1 Introduction

This is a proposal to encode a new character in the Universal Character Set: 𑖅 DEVANAGARI LETTER MARWARI DDA. It is used in Devanagari orthography for Marwari [rwr] and other ‘Rajasthani’ languages for representing [d].

2 Character Details

Properties of the proposed character are as follows:

GLYPH	CHARACTER NAME	GC	CCC	BIDI	MIRRORED
𑖅	DEVANAGARI LETTER MARWARI DDA	Lo	0	L	N

3 Background

In Devanagari, the phoneme [d] is represented using 𑖄 U+0921 DEVANAGARI LETTER DDA. In Hindi and its allied languages, [d] has an allophonic variant [ɖ], which is written using 𑖄 U+095C DEVANAGARI LETTER DDDHA. This letter is produced by writing 𑖄 with 𑖅 NUKTA.

Historically, in Devanagari orthography for Marwari in Rajasthan, [d] is represented by the letter 𑖅 and its allophone [ɖ] is written using 𑖄. The use of 𑖅 is shown in Figure 1 and Figure 2, and is attested in written documents, excerpts of which are shown in Figure 3 and Figure 4. Apparently, the character was also used in printed Marwari, at least in the early 20th century. In print, however, 𑖅 was often replaced with Devanagari 𑖄 MA when a metal punch for the letter was unavailable.¹

Information about the origins of the character is unavailable. But, based upon graphical typology, it is highly plausible that the character is derived ultimately from 𑖅 U+1119D SHARADA LETTER DDA. It probably entered into Devanagari through Landa scripts used in Rajasthan. The Landa scripts are related to Sharada and use some variant of 𑖅 for writing [d], eg. 𑖅 U+0A21 GURMUKHI LETTER DDA. This character also corresponds to 𑖅 U+11696 TAKRI LETTER DDA.

¹ Grierson 1908: 20.

4 Justification

It might be possible to identify ढ as a glyphic variant of ढ DDA. Its representation could then be managed at the font level: ढ would be assigned to U+0921 DDA and the glyph ढ would be shifted to U+095C DDDHA. This, however, would completely displace ढ. There are two additional problems with this approach.

The first is that it obscures character semantics. Devanagari ढ DDDHA has a compatibility decomposition of <ढ 095C, ढ 093C>. Therefore, replacing the glyph ढ with ढ for DDDHA would result in incorrect properties for the character.

The second is that it does not allow for the unique representation of characters in plain text. Treating ढ as a glyphic variant prevents its usage in text in which both ढ and ढ may occur. As shown in Figure 2, both ढ and ढ are used for representing [d]; ढ being a variant of ढ. In turn, both ढ and ढ are used for writing [ɽ]. Thus, ढ, ढ, ढ, and ढ may all appear together.

In order to maintain the stability of properties of existing characters and to ensure that Devanagari meets the orthographic requirements for various languages, an independent encoding for ढ is the most appropriate solution.

5 Notes

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

Glyph Shape The proposed glyph shape ढ is based upon the form given in Grierson (see Figure 1). Slightly different graphical forms, ie. ढ, are shown by Metzger (see Figure 2).

6 References

- 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.
- Metzger, Mathias. 2003. *Die Sprache der Vakīl-Briefe aus Rājasthān*. Beiträge zur Südasienforschung, Südasien-Institut, Universität Heidelberg, Band 193. Würzburg: Ergon Verlag.

In writing Mārwarī 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 ढ़. In Mārwarī there are two distinct characters, viz., ढ for *ḍ* and ढ़ for *ṛ*. In printed Mārwarī, when the type for ढ is not available, the Dēvanāgarī letter ढ 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 Mārwarī specimens in the Dēvanāgarī character, I shall represent *ḍ* by ढ and *ṛ* by ढ़.

Figure 1: Description of the letter ढ as used for writing Marwari (from Grierson 1908: 20).

Transliteration	Devanāgarī	Jagjivan Dās 1	Jagjivan Dās 2	Bhikhārī Dās
<i>t</i>	ट	ढ	ढ	ढ
<i>th</i>	फ	ढ	ढ	ढ
<i>ḍ</i>	ड	ढ	ढ	ढ
<i>ḍh</i>	ढ	ढ	ढ	ढ
<i>ṛ</i>	ढ़	ढ़	ढ़	
<i>ṇ</i>	ण	ण	ण	ण

Figure 2: Excerpt from Metzger (2003: 20) showing the use of ढ for representing *ḍa*.

जेठ जीरो रे होय ढावडा ला: छिवां मांयसू नैर्जा

Figure 3: Line written in Marwari style of Devanagari (from Grierson 1908: 63). The highlighted word is ढावडा [d̪a:v̪a:] ‘sons’.

अर मे करि मे' अमर सु नोम

Figure 4: Line written in Marwari style of Devanagari (from Grierson 1908: 131). The highlighted word is मंठोर [m̪əṇḍo:r] ‘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 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:	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:	2010-12-08
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 <input checked="" type="checkbox"/> B.1-Specialized (small collection) <input 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?	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.

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:	Yes See text of proposal.
4. The context of use for the proposed characters (type of use; common or rare) Reference:	Common See text of proposal.
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	Yes 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? 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)?	N/A
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:	Yes Yes See text of proposal.
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:	No
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)	No
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