

To: UTC and ISO/IEC JTC1/SC2 WG2  
Title: Proposal to add ARABIC LETTER ZAIN WITH INVERTED V ABOVE (L2/12-181R)  
From: Lorna A. Priest (SIL International)  
Date: 9 May 2012

## Introduction

This is a proposal to add one character for Berber to the Arabic Extended-A block.

There is not a whole lot written in Berber. People do write it by hand and in some books. The dictionary referenced in this proposal is the most official publication on the Berber language in Morocco. It includes lexical items from the three main Berber languages in Morocco: Tachelhit (shi), Tamazight (tmz) and Tarifit (rif).

Figure 1 is from volume 1 of the 3 volume Arabic-Berber Dictionary by Mohamed Shafiq. It shows the three special characters the author uses to represent sounds in Berber which do not have a normal character in Arabic script: 1)  $\underset{\cdot}{z}$  (IPA  $z^{\text{c}}$ ) 2)  $\underset{\cdot}{g}$  3)  $\underset{\cdot}{k}$  (IPA  $\text{ç}$ ). The 2<sup>nd</sup> and 3<sup>rd</sup> characters are already in Unicode. The character we are proposing is number 1, the pharyngealized z.

## Proposed character

There seems to be an inconsistency in naming with “inverted small v” and “inverted v” both being used. This proposal uses “inverted v”. The character could also be named ARABIC LETTER REH WITH DOT AND INVERTED V ABOVE, however, that is not our preference as the character is related to “z”.

Proposed codepoint	Glyph	Final Glyph	Proposed Unicode Name
08B2	ز̣	ز̣	ARABIC LETTER ZAIN WITH INVERTED V ABOVE

## Unicode character Properties

Character properties should be the same as for 0632:

```
08B2;ARABIC LETTER ZAIN WITH INVERTED V ABOVE;Lo;0;AL;;;;;N;;;;;
```

## Joining type and group for ArabicShaping.txt

```
08B2; REH WITH DOT AND INVERTED V ABOVE; R; REH
```

## Suggested Collation

Since tailoring will be required, ARABIC LETTER ZAIN WITH INVERTED V ABOVE can sort with other reh-like characters such as U+0699 or U+076B.

## Normalization Issues

ARABIC LETTER ZAIN WITH INVERTED V ABOVE should never be decomposed.

## Confusability

This character should be added to the list of confusables.



**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:** Proposal to add ARABIC LETTER ZAIN WITH INVERTED V ABOVE

2. Requester's name: Lorna Priest

3. Requester type (Member body/Liaison/Individual contribution): Individual Contribution

4. Submission date: 4 May 2012 (revised 9 May 2012)

5. Requester's reference (if applicable): L2/12-181R

6. Choose one of the following:

This is a complete proposal: Yes

(or) More information will be provided later: No

**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: Arabic Extended-A

2. Number of characters in proposal: 1

3. Proposed category (select one from below - see section 2.2 of P&P document):

A-Contemporary  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? SIL

b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.): SIL, lorna\_priest@sil.org

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

See proposal

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:	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
4. The context of use for the proposed characters (type of use; common or rare) Reference:	Somewhat common
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:	Yes
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	na
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:	Perhaps
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:	Perhaps
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:	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 characters? If YES, are the equivalent corresponding unified ideographic characters identified? If YES, reference:	No