Proposal to encode additional Arabic-script characters

Proposal:

Name of script: Arabic

Existing existing block: Arabic & Arabic Supplement

Number of characters in proposal: 5

Proposed Level of Implementation: A

Is a rationale provided for the choice? Yes

Is a repertoire including character names provided? Yes

If YES, are the names in accordance with the 'character naming guidelines in Annex L of ISO/IEC 10646-1: 2000'? Yes

Are the character shapes attached in a legible form suitable for review? Yes

Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard? Jonathan Kew, SIL International

If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

jonathan_kew@sil.org

TrueType font generated with FontLab 4.6

Are references (to other character sets, dictionaries, descriptive texts etc.) provided? Yes

Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? Yes

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 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/UnicodeCharacterDatabase.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? Not to WG2
   If YES explain (but characters approved for encoding at UTC #98, February 2004)

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? Linguists and NGOs working in S. Asia, N. Africa
   If YES, available relevant documents: See §4 below

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 §3 below

4. The context of use for the proposed characters (type of use; common or rare) Rare
   Reference: See §3 below, rare only because of limited literacy in user communities

5. Are the proposed characters in current use by the user community? Yes
   If YES, where? Reference: Northern Areas of Pakistan; Mali

6. After giving due considerations to the principles in Principles and Procedures document (a WG 2 standing document) must the proposed characters be entirely in the BMP? Yes
   If YES, is a rationale provided? Yes
   If YES, reference: Extensions to BMP Arabic repertoire, see §3 below

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? No

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? Possibly
   If YES, is a rationale for its inclusion provided? Yes
   If YES, reference: See §3.2.1 below

10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to 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 (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)? Yes
    If YES, is a rationale for such use provided? Yes
    If YES, reference: Arabic-script vowel marks are inherently combining characters
    Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? No
    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:

1. Proposed character additions (shaded cells)

<table>
<thead>
<tr>
<th>Row</th>
<th>Column 060</th>
<th>061</th>
<th>062</th>
<th>063</th>
<th>064</th>
<th>065</th>
<th>066</th>
<th>067</th>
<th>075</th>
<th>076</th>
<th>077</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>ٍ</td>
<td>ﯲ</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td>١</td>
<td></td>
</tr>
</tbody>
</table>

| Column A | ٣ | ٤ | ٥ | ٦ | ٧ | ٨ | ٩ |
| Column B | ٣ | ٤ | ٥ | ٦ | ٧ | ٨ | ٩ |
| Column C | ٣ | ٤ | ٥ | ٦ | ٧ | ٨ | ٩ |
| Column D | ٣ | ٤ | ٥ | ٦ | ٧ | ٨ | ٩ |
| Column E | ٣ | ٤ | ٥ | ٦ | ٧ | ٨ | ٩ |
| Column F | ٣ | ٤ | ٥ | ٦ | ٧ | ٨ | ٩ |
2. Names list for character additions

<table>
<thead>
<tr>
<th>Listing of file ArabicUTC98-names.txt</th>
</tr>
</thead>
<tbody>
<tr>
<td>:</td>
</tr>
<tr>
<td>: Additions to Arabic block</td>
</tr>
<tr>
<td>@@ 0600 Arabic 06FF</td>
</tr>
<tr>
<td>: 065D ARABIC REVERSED DAMMA</td>
</tr>
<tr>
<td>* Ormuri, African languages</td>
</tr>
<tr>
<td>065E ARABIC FATHA WITH TWO DOTS</td>
</tr>
<tr>
<td>* Kalami</td>
</tr>
<tr>
<td>: Additions to Arabic Supplement block</td>
</tr>
<tr>
<td>@@ 0750 Arabic Supplement 077F</td>
</tr>
<tr>
<td>: 076B ARABIC LETTER REH WITH TWO DOTS VERTICALLY ABOVE</td>
</tr>
<tr>
<td>* Torwali, Ormuri</td>
</tr>
<tr>
<td>076C ARABIC LETTER REH WITH HAMZA ABOVE</td>
</tr>
<tr>
<td>* Ormuri</td>
</tr>
<tr>
<td>076D ARABIC LETTER SEEN WITH TWO DOTS VERTICALLY ABOVE</td>
</tr>
<tr>
<td>* Kalami, Ormuri</td>
</tr>
</tbody>
</table>
3. Discussion of the proposed additions

In many minority languages of Pakistan, there is no long literary tradition, but in recent years there have been moves to develop suitable orthographies and an indigenous literature. Orthographies are normally based on those of neighboring major languages, such as Urdu and Pashto, but additional letters may be needed for sounds not found in those languages.

The present proposal is for three new Arabic-script characters that have been used in writing several such languages. Examples are shown from published materials in Kalami, Ormuri, and Torwali, although these are not necessarily the only language communities where the characters may be used. In some languages of this region, orthographies are not yet standardized, and it is likely that as conventions are established in one language community, they may well be adopted among neighboring communities as well.

3.1 Background on the language communities

Basic information about the languages discussed here is taken from SIL’s Ethnologue (see http://www.ethnologue.com/). Further comments are based on communication with linguists who have studied in the region and have contact with local language communities and writers.

3.1.1 Kalami (Gawri)

Region: Upper Swat Kohistan from between Peshmal and Kalam north to upper valleys above Kalam, also in Dir Kohistan, in Thal, Lamuti (Kinolam), Biar (Jiar), and Rajkot (Patrak) villages. People at Khata Khotan, China, are reported to be related, recognized by their clothing and language.
Alternate names: GARWI, GAWRI, GOWRI, GARWA, GAAWRO, KALAMI KOHISTANI, KOHISTANI, KOHISTANA, BASHKARIK, BASHGHARIK, DIR KOHISTANI, DIRI, DIRWALI
Dialects: KALAM, USHU, THAL, LAMUTI (LAMTI), RAJKOTI (PATRAK), DASHWA.
Comments: Dialect differences do not hinder communication, except that speakers of other dialects have difficulty with Rajkot. 90% to 93% lexical similarity among the main dialects; Rajkoti has 75% with Kalami; Dashwa has 77% with Kalami, and 74% with Rajkoti. The most widely understood indigenous language in northern Swat and Dir Kohistan. Men are fairly bilingual in Pashto; women are more limited. Rajkoti men have high bilingualism in Pashto. Uneducated men and women are limited in Urdu. There appear to be few active speakers of Dashwa. Kalami and Ushu speakers indicate some negative attitudes toward each other’s speech. Dashwa is a clan name of people originally from around Rajkot; little information available. About one-third migrate in winter to Mingora, Mardan, Peshawar, or the Punjab in search of work. Speakers of Pashto, Gujar, Khowar, and other Kohistani languages live among them, but they are generally in the majority. Patrilineal descent groups are: Drekhel, Nilo (Niliyor), Jaflor (Jaflor). The Drekhel are divided into the Kalamkhel, Akarkhel, and Chinorkhel. The Mullakhel are Pashtoons from Lower Swat who now speak Pashto as first language, but speak, understand, and identify with Kalami. Muslim.

Baart & Sagar (2002) give further background on the Kalami or Gawri language and orthography. Although a few people have been writing Kalami in the past, little has been published in the language and until recently there was no standardization of the writing system. This is now changing, however:

In the summer of 1995, seven educated native Gawri speakers from Kalam formed a spelling committee and discussed a proposal for a writing system. They discussed the question as to which symbols should be used for representing the sounds and tones of Gawri... [I]t was felt that there should be maximum conformity of the Gawri writing system with that of Urdu. ... On the other hand, the committee recognized the uniqueness of the Gawri language and chose to preserve this uniqueness by designing a writing system that can accurately represent all the distinctive sounds of the language.

Since 1995, there have been a number of literary publications in the Kalami language, making use of the characters proposed here. Baart & Sagar include bibliographic references for several works by local authors published by the Kalam Cultural Society.
3.1.2 Ormuri

Population: 3,000 or more in Pakistan (1992). Population total both countries 3,050 or more.

Region: Kaniguram, a pocket in Mohsud Pashto area northwest of Dera Ismail Khan, Wazirstan. Also spoken in Afghanistan.

Alternate names: URMURI, ORMUR, ORMUI, BARGISTA, BARAKS, BARAKI

Dialects: KANIGURAMI, LOGAR.

Classification: Indo-European, Indo-Iranian, Iranian, Western, Northwestern, Ormuri-Parachi.

Comments: 27% lexical similarity with Waneci, 25% to 33% with Pashto dialects. The Kanigurami retain the language.

As indicated by the Ethnologue statistics, Ormuri is a small language group, and as one might expect there has been little literary activity. However, in recent years at least one book has been published in Ormuri, and there is work in progress to compile a dictionary. Burki (2001) mentions the alphabet developed for these purposes, including a description of the phonemes for which new letters were devised.

3.1.3 Torwali


Region: Swat Kohistan, on both sides of Swat River from just beyond Madyan north to Asrit (between Mankjal and Peshmal), and in Chail Valley east of Madyan, Bahrain and Chail are centers.

Alternate names: TURVALI

Dialects: BAHRAIN, CHAIL.


Comments: 44% lexical similarity with Kalkoti and Kalami, 89% between Behrain and Chail. Men are fairly bilingual in Pashto, more limited in Urdu. Women are limited in use of Pashto, and know almost no Urdu. Sunni Muslim.

Although Torwali is a much larger language community than Ormuri, the situation with regard to writing is similar. As yet there has been very little published, and no effort by a spelling committee or other body to standardize orthography. However, the examples shown in figures 13–14, from an indigenous publication, show that similar approaches are being used to extend the script.

3.1.4 North African languages

One of the proposed vowel marks has been used in writing a number of North African languages, including Songhoy, Fulfulde/Pulaar, Zarma, and Hausa. Background information can be found in document L2/03-223 (N2598), based largely on Chtatou (1992).

3.2 The proposed characters

3.2.1 Base characters

Three of the proposed characters are new Arabic-script letters, and should be most appropriately encoded in the Arabic Supplement block at U+0750. All three are of General Category Lo; Combining Class 0; Bidi Type AL.

<table>
<thead>
<tr>
<th>Glyph</th>
<th>Code</th>
<th>Character name</th>
<th>Shaping</th>
<th>See figures</th>
</tr>
</thead>
<tbody>
<tr>
<td>ﺽ</td>
<td>076B</td>
<td>ARABIC LETTER REH WITH TWO DOTS VERTICALLY ABOVE</td>
<td>REH</td>
<td>2, 4, 5, 6, 13, 14</td>
</tr>
<tr>
<td>ﺽ</td>
<td>076C</td>
<td>ARABIC LETTER REH WITH HAMZA ABOVE</td>
<td>REH</td>
<td>1, 4, 5</td>
</tr>
<tr>
<td>ﺼ</td>
<td>076D</td>
<td>ARABIC LETTER SEEN WITH TWO DOTS VERTICALLY ABOVE</td>
<td>SEEN</td>
<td>3, 4, 5, 6, 7, 9, 10, 11, 12</td>
</tr>
</tbody>
</table>

The alphabet chart in Burki (2001), shown in figure 5 below, implies that these characters should be collated as follows:

- ﺽ U+076B somewhere between ﺽ U+0632 and ﺽ U+0698;
- ﺽ U+076C between ﺽ U+0631 and ﺽ U+0632;
- ﺼ U+076D between ﺼ U+0633 and ﺼ U+0634.

However, these are merely suggestions for defaults; the exact values chosen for the UCA DUCET are not critical, as language-specific tailorings are likely to be needed for all of these languages in any case.
The proposed character \( \text{ظهور} \) \( \text{U+076C} \) deserves special mention, in that a similar graphic appearance could be produced by the sequence \( \text{ الرحمن} \) \( \text{U+0631} \), \( \text{U+0654} \), applying the HAMZA ABOVE mark to the standard Arabic letter REH. However, this is not an appropriate encoding, as this Ormuri letter does not have any association with the HAMZA; it has merely borrowed its graphic form as a means to create a new consonant.

In this context, we should note the existence of \( \text{ الرحمن} \) \( \text{U+0681} \) HAH WITH HAMZA ABOVE. This character is well established in Pashto orthography, and probably provided the inspiration for the Ormuri extension of REH with a HAMZA-shaped mark. There is thus a precedent for the use of this mark as a consonant modification, independent of its conventional Arabic meaning.

We may also note that \( \text{ الرحمن} \) \( \text{U+0681} \) does not decompose to a sequence \( \text{ الرحمن} \) \( \text{U+062D} \), \( \text{U+0654} \), as would be expected if the mark here were in fact an occurrence of HAMZA. This is in contrast to \( \text{ الرحمن} \) \( \text{U+0623} \), \( \text{U+0624} \), etc., which do decompose. This supports the view that \( \text{ الرحمن} \) \( \text{U+0681} \) is a single, indivisible letter, not a combination of an existing letter with HAMZA added. The same would be true of the proposed \( \text{ظهور} \) \( \text{U+076C} \).

Thus, although the mark seen on \( \text{ الرحمن} \) \( \text{U+0681} \) and \( \text{ظهور} \) \( \text{U+076C} \) is visually based on HAMZA, it is not in fact HAMZA but an integral part of a new letter. It should not be encoded as \( \text{ الرحمن} \) \( \text{U+0654} \), especially given the inconsistency this would introduce between these two letters.

### 3.2.2 Vowel diacritics

The remaining two proposed characters are combining marks used to indicate vowels in extended Arabic-based writing systems. These should both have General Category Mn; Combining Class 30; Bidi type NSM. (The combining class value is somewhat arbitrary; ideally, all Arabic vowel marks written above the base letter would have the same CC value, but the already-defined fixed-position classes make this impossible. Class 30, originally assigned to ARABIC FATHA, is arbitrarily chosen for these new vowel marks.)

<table>
<thead>
<tr>
<th>Glyph</th>
<th>Code</th>
<th>Character name</th>
<th>See figures</th>
</tr>
</thead>
<tbody>
<tr>
<td>ḫ</td>
<td>065D</td>
<td>ARABIC REVERSED DAMMA</td>
<td>15, 16, 17, 18, 19, 20</td>
</tr>
<tr>
<td>ḫ</td>
<td>065E</td>
<td>ARABIC FATHA WITH TWO DOTS</td>
<td>8, 9, 10, 11, 12</td>
</tr>
</tbody>
</table>

In the examples below, the REVERSED DAMMA is written with an “open” glyph, something like \( \text{ภาพย} \), rather than \( \text{ภาพย} \). Note, however, that in such hand-written sources from north African languages, it is also normal to see a form such as \( \text{ภาพย} \) for DAMMA, rather than the \( \text{ภาพย} \) of traditional Naskh-style typography. These “open” forms are glyph variants particularly typical of hand-written text in Africa. However, it is clear that the sign being used in these languages for the /o/ vowel is derived from DAMMA by reversing the orientation of the shape, and thus \( \text{ภาพย} \) would be an appropriate representative glyph in a typical Naskh typeface.

### 3.3 Examples of use

![Figure 1](image)

Figure 1: Burki (1999), page 7: showing proposed \( \text{ظهور} \) \( \text{U+076C} \).
Figure 2: Burki (1999), page 8: showing proposed $\dot{\jmath}$ U+076B.

Figure 3: Burki (1999), page 10: showing proposed $\check{\jmath}$ U+076D.

[iteralsound [r] (voiceless trilled-r); It is an Ormuri-specific sound, which is not used in any other language of the subcontinent and other Indo-Iranian languages. It is a voiceless trilled r, which has also been explained by Joan Baart in the January, 1998 issue of North Pakistan Newsletter as follows: “Among other things, it was interesting to learn that the language has a voiceless trilled r of Czech (as in the name of the composer Antonin Dvorak). There is also a hard to hear phonemic contrast between two Kinds of esh sound, probably palato-alveolar vs. alveolo-palatal.”

[criptorsound [z] (voiced alveolo-palatal grooved fricative).
This sound is though also used in Waziri dialect of Pashto but Pashto has no symbol for its expression and the regular symbol of (z) is used for this sound despite the difference. It sounds like [S] in pleasure, which is different from [S] of vision.

[criptorsound [c] (voiceless palato-alveolar grooved fricative).
It is also a specific Ormuri sound though also used in Waziri dialect. It sounds like [c] in pissing in Urdu or Punjabi. In Pashto no specific symbol is used for this sound and is written with a regular [c]. In Ormuri if the difference of both the sounds i.e. [z] and [c] is not indicated, many words and their meaning would cause confusion.

Figure 4: Burki (2001), page 62: showing proposed $\dot{\jmath}$ U+076B, $\check{\jmath}$ U+076C, $\check{\jmath}$ U+076D.
arranged nuqtas under the character (ح). The shape of Sin with two vertically arranged nuqtas above the character (س) is used for the retroflex fricative /s/.

Retroflex fricatives and affricates occur in many languages of northern Pakistan, and many different symbols have been proposed for writing them. The symbols chosen by the Gawri spelling committee (ش and ح) were adapted from work by Karimi (1982/1995) on the neighbouring Torwali language.

Figure 5: Burki (2001), page 64: showing proposed ٧ U+076B, ٨ U+076C, ٩ U+076D.

Figure 6: Burki (2001), page 65: showing proposed ٧ U+076B, ٩ U+076D.

Figure 7: Baart & Sagar (2002), page 9: showing proposed ٩ U+076D.
The writing of vowels poses a special challenge in Gawri, as the language has six basic vowel qualities, and a length distinction (short vs. long) is applied to all of them, giving a total of twelve pure vowels (not counting nasalized vowels). This contrasts with Urdu, which has a total of only eight such pure vowels. A way had to be designed, then, to mark those extra vowels of Gawri in the writing system.

As Gawri has two open, that is a-like, vowels, it was decided to introduce a new vowel diacritic for the front open vowel. This new vowel diacritic has the shape of a Zabar with a little dot on each side, as in بَت/‘stones’. A regular Zabar is used to indicate the back open vowel, as in بَت/‘stone’.

**Figure 8:** Baart & Sagar (2002), page 9: showing proposed ◄ U+065E.

**Figure 9:** Baart & Sagar (2002), page 21: showing proposed ◄ U+065E, ◄ U+076D.

**Figure 10:** Sagar (n.d.), part 2, page 2: showing proposed ◄ U+065E, ◄ U+076D.
Figure 11: Ulfat (2000), alphabet chart: showing proposed $\U+065E$, $\U+076D$.

Figure 12: Ulfat (2000), page 1: showing proposed $\U+065E$, $\U+076D$.

Figure 13: Karimi (1995), page 7: showing proposed $\U+076B$.

Figure 14: Karimi (1995), part 2, page 7: showing proposed $\U+076B$. 
and by a diacritic mark \( \underline{\text{ٴ}} \) followed by a \( \text{wāw} /\text{o}/ \) in the case of \( /\text{oo}/ \):

<table>
<thead>
<tr>
<th>(9) Songhoy</th>
<th>Arabic</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>kooro</td>
<td>كورو</td>
<td>hyena</td>
</tr>
<tr>
<td>tooru</td>
<td>تور</td>
<td>fetish</td>
</tr>
<tr>
<td>yoobu</td>
<td>يوب</td>
<td>market</td>
</tr>
</tbody>
</table>

**Figure 15:** Chtatou (1992), page 31: showing proposed \( \text{U+065D} \) used in Songhoy.

3.2.2.1. *Vowels*

The transcription of short vowels of this language is based on *alif* /\( \text{l}/\) with various vowel marks to give the different vowel qualities without the use of *hamza*, glottal stop, as it is the case.

(22)

<table>
<thead>
<tr>
<th></th>
<th>َ</th>
<th>َا</th>
<th>َي</th>
<th>َو</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>ٍ</td>
<td>ٍ</td>
<td>۸</td>
<td>۸</td>
</tr>
<tr>
<td>i</td>
<td>َ</td>
<td>َا</td>
<td>َي</td>
<td>َو</td>
</tr>
<tr>
<td>u</td>
<td>ٍ</td>
<td>ٍ</td>
<td>۸</td>
<td>۸</td>
</tr>
<tr>
<td>e</td>
<td>َ</td>
<td>َا</td>
<td>َي</td>
<td>َو</td>
</tr>
<tr>
<td>o</td>
<td>َ</td>
<td>َا</td>
<td>َي</td>
<td>َو</td>
</tr>
</tbody>
</table>

**Figure 16:** Chtatou (1992), page 38: showing proposed \( \text{U+065D} \) used in Pulaar.

3.3.2. *Zarma*

3.3.2.1. *Vowels*

The sign \( \underline{\text{ٴ}} \) is being used to represent the vowel \( o \) which does not exist in Arabic:

(29) boro \( \text{بورو} \) man

**Figure 17:** Chtatou (1992), page 43: showing proposed \( \text{U+065D} \) used in Zarma.

**Figure 18:** Chtatou (1992), page 51: showing proposed \( \text{U+065D} \) used in Fulfulde.
Figure 19: Centre Littérature Evangelique (1992), page 'F': showing proposed \U065D. Compare shape used for \U064F, highlighted in blue.

Figure 20: Mission Evangelique Luthérienne (1996), page 1: showing proposed \U065D. Compare shape used for \U064F, highlighted in blue.
4. References


Burki, Rozi Khan. 1999. [Should we leave our language on its death-bed?]. Tank, Pakistan: Muslim Book Agency.


Sagar, Muhammad Zaman (ed). 2002(?). Chunjaal [Wake up!]: a collection of Gawri stories by different writers of Dir Kohistan. 