Proposal to encode yet two more Arabic honorifics

Roozbeh Pournader (WhatsApp), Soheil Hooshdaran, and Mustafa Jibaly
January 15, 2020

Background
The Unicode Technical Committee, in its meeting #160 in July 2019, accepted fourteen Arabic honorifics proposed by the authors. After we found evidence for two more, they were accepted in UTC meeting #161 in November 2019. Only characters for which we could find examples in running text were proposed in those documents. Since then, we have found usage of two more honorifics in published material.

Proposal
Encode the following characters in the Unicode Standard (the figures column lists figures with examples):

<table>
<thead>
<tr>
<th>Glyph</th>
<th>Codepoint</th>
<th>Name</th>
<th>Figures</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Glyph" /></td>
<td>FD4E</td>
<td>ARABIC LIGATURE TABAARAKA WA-TAAALAA</td>
<td>1</td>
</tr>
<tr>
<td><img src="image2" alt="Glyph" /></td>
<td>FD4F</td>
<td>ARABIC LIGATURE RAHIMAHUM ALLAAH</td>
<td>2</td>
</tr>
</tbody>
</table>

The main character properties should be as follows:
FD4E;ARABIC LIGATURE TABAARAKA WA-TAAALAA;So;0;ON;;;;N;;;;;
FD4F;ARABIC LIGATURE RAHIMAHUM ALLAAH;So;0;ON;;;;N;;;;;

The script property for both characters should be Arabic. All other properties should be similar to U+FDFD ARABIC LIGATURE BISMILLAH AR-RAHMAN AR-RAHEEM.

Notes
1. The character names match the existing patterns in Unicode character names. The Arabic phrases have been transcribed with a key similar to that used in U+FDFD ARABIC LIGATURE BISMILLAH AR-RAHMAN AR-RAHEEM and the Quranic characters accepted at
U+08D5..U+08D2 (see L2/14-105), as opposed to U+FDF0..U+FDFB and U+0610..U+0613. This is because the former pattern is more common among Arabic speakers.

2. The following twenty-three forms exist in some honorific fonts, but we haven’t yet found evidence for their usage in published material. They may be proposed later:

<table>
<thead>
<tr>
<th>Glyph</th>
<th>Spelled out form</th>
<th>Glyph</th>
<th>Spelled out form</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Glyph" /></td>
<td>سلام الله عليهم اجمعين</td>
<td><img src="image2.png" alt="Glyph" /></td>
<td>رحمها الله</td>
</tr>
<tr>
<td><img src="image3.png" alt="Glyph" /></td>
<td>رضوان الله عليهم اجمعين</td>
<td><img src="image4.png" alt="Glyph" /></td>
<td>رحمهما الله</td>
</tr>
<tr>
<td><img src="image5.png" alt="Glyph" /></td>
<td>صلوات الله عليه</td>
<td><img src="image6.png" alt="Glyph" /></td>
<td>رحمهن الله</td>
</tr>
<tr>
<td><img src="image7.png" alt="Glyph" /></td>
<td>صلوات الله عليها</td>
<td><img src="image8.png" alt=" Glyph" /></td>
<td>قدس سرهما</td>
</tr>
<tr>
<td><img src="image9.png" alt="Glyph" /></td>
<td>صلوات الله عليهما</td>
<td><img src="image10.png" alt="Glyph" /></td>
<td>قدس سرهما</td>
</tr>
<tr>
<td><img src="image11.png" alt="Glyph" /></td>
<td>صلوات الله عليهم</td>
<td><img src="image12.png" alt="Glyph" /></td>
<td>قدس سره الشريف</td>
</tr>
<tr>
<td>عجل الله تعال</td>
<td>عجل الله تعال</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>---------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>رحمة الله عليه</td>
<td>حفظه الله</td>
<td></td>
<td></td>
</tr>
<tr>
<td>رحمة الله عليها</td>
<td>حفظها الله</td>
<td></td>
<td></td>
</tr>
<tr>
<td>رحمة الله عليهم</td>
<td>حفظهم الله</td>
<td></td>
<td></td>
</tr>
<tr>
<td>رحمة الله عليهن</td>
<td>جل وعلا</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sample of usage


قال: (والعنان) قال أبو بكر: قالوا: والعنان، قال: (كم ترون بينكم وبين السماء؟) قالوا: لا ندري، قال: فإن بينكم وبينها إما واحداً، أو اثنين، أو ثلاثةً وسبعين سنة، والسماء فوقها كذلك، حتى عد سبع سماوات، ثم فوق السماء السابعة ببحر، بين أعلاه وأسفله كما بين سماء إلى سماء، ثم فوق ذلك ثمانية أو عاشر، بين أطلافن وركهن كما بين سماء إلى سماء، ثم على ظهورهن العرش، بين أعلاه وأسفله كما بين سماء إلى سماء، ثم الله ﷺ.

Figure 1. Sample for TABAARAKA WA-TAAALAA, in red oval, from ash-Shâyi’ 2019, p. 131. In line 4, you can also see U+F DFA ARABIC LIGATURE SALLALLAHOU ALAYHE WASALLAM used, in blue oval.
Figure 2. Sample for RAHIMAHUM ALLAAH, in red oval, from as-Sadugh 2010. In blue ovals you can also see U+FD41 ARABIC LIGATURE RADI ALLAAHU ANH and U+FD47 ARABIC LIGATURE ALAYHI AS-SALAAM, proposed in Pournader and Jibaly 2019a.

Bibliography

ISO/IEC JTC 1/SC 2/WG 2

PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646

A. Administrative

1. Title: Proposal to encode yet two more Arabic honorifics
2. Requester's name: Roozbeh Pournader, Soheil Hooshdaran, and Mustafa Jibaly
3. Requester Type: Expert Contribution
4. Submission date: January 10, 2020
5. Requester's reference, if applicable: N/A
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: N/A
   b. The proposal is for addition of character(s) to an existing block: Yes
   Name of existing block: Arabic Presentation Forms-A
2. Number of characters in proposal: 2
3. Proposed category: A-Contemporary
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? Roozbeh Pournader
   b. Identify the party granting a license for use of the font by the editors (include address,
      e-mail, ftp-site, etc.):
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 endorse
   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 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.

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? Yes
   If YES explain: **Examples for the proposed characters are now provided.**
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? **The authors have seen frequent examples of usage in publications around the Muslim world. There are many requests by the user community on social media.**
   If YES, available relevant documents: N/A
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 above
4. The context of use for the proposed characters (type of use; common or rare): **Frequent in religious text, although rarer in comparison to some other honorifics proposed before**
   Reference: N/A
5. Are the proposed characters in current use by the user community? Yes
   If YES, where? Reference: See above
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. **Need to be next to similar characters.**
   If YES, reference: N/A
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? Yes
   If YES, is a rationale for its inclusion provided? Yes
   If YES, reference: **Similarity to already-encoded characters and different usage.**
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: See above
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? Yes
    If YES, is a rationale for its inclusion provided? Yes. **The proposed characters have different identities.**
    If YES, reference: See above
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? N/A
    If YES, reference: N/A
    Is a list of composite sequences and their corresponding glyph images (graphic symbols)
<table>
<thead>
<tr>
<th>Question</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>provided?</td>
<td>N/A</td>
</tr>
<tr>
<td>If YES, reference:</td>
<td>N/A</td>
</tr>
<tr>
<td>12. Does the proposal contain characters with any special properties</td>
<td>No</td>
</tr>
<tr>
<td>such as control function or similar semantics?</td>
<td></td>
</tr>
<tr>
<td>If YES, describe in detail (include attachment if necessary):</td>
<td>N/A</td>
</tr>
<tr>
<td>13. Does the proposal contain any Ideographic compatibility characters?</td>
<td>No</td>
</tr>
<tr>
<td>If YES, are the equivalent corresponding unified ideographic characters</td>
<td>N/A</td>
</tr>
<tr>
<td>identified?</td>
<td></td>
</tr>
<tr>
<td>If YES, reference:</td>
<td>N/A</td>
</tr>
</tbody>
</table>