Proposal to encode seventeen Arabic honorifics

Roozbeh Pournader, Google Inc.
July 27, 2014

Background

Unicode includes two sets of Arabic honorifics, one as combining marks at U+0610..0614 (mostly for use in Pakistan, proposed in L2/01-425), and another at U+FDF0..FDF2 (all around the Muslim world). But the encoded set is far from complete.

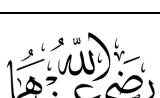
The characters proposed here are commonly used by both Sunni and Shi'a Muslims around the world. The last four are mostly used by the Persian- and Arabic-speaking Shi'a Muslims in Iran and Iraq. The user community has been using hacked fonts or images in order to represent the missing ones, or when missing fonts, using parenthesized abbreviations.

Several variations in glyph shape exist among the proposed characters, but they are interchangeable and would all refer to the same semantics when the underlying phrase is the same.

Proposal

Encode the following characters in the Unicode Standard:

Glyph	Codepoint	Name and notes
	FD40	ARABIC LIGATURE RAHIMAHU ALLAAH
	FD41	ARABIC LIGATURE RAHIMAHUM ALLAAH
	FD42	ARABIC LIGATURE RAHIMAHUMAA ALLAAH
	FD43	ARABIC LIGATURE RADI ALLAAHU ANH

	FD44	ARABIC LIGATURE RADI ALLAAHU ANHAA
	FD45	ARABIC LIGATURE RADI ALLAAHU ANHUM
	FD46	ARABIC LIGATURE RADI ALLAAHU ANHUMAA
	FD47	ARABIC LIGATURE SALLALLAAHU ALAYHI WA-AALIH
	FD48	ARABIC LIGATURE SALLALLAAHU ALAYHI WA-AALIHEE WA-SALLAM
	FD49	ARABIC LIGATURE ALAYHI AS-SALAAM
	FD4A	ARABIC LIGATURE ALAYHAA AS-SALAAM
	FD4B	ARABIC LIGATURE ALAYHIM AS-SALAAM
	FD4C	ARABIC LIGATURE ALAYHIMAA AS-SALAAM
	FD4D	ARABIC LIGATURE QUDDISA SIRRAH
	FD4E	ARABIC LIGATURE QUDDISA SIRRAHUM
	FD4F	ARABIC LIGATURE QUDDISA SIRRAHUMAA
	FDFE	ARABIC LIGATURE AJJALA ALLAAHU TAAALAA FARAJAHU ASH-SHARIF

The main character properties should be as follows:

```
FD40;ARABIC LIGATURE RAHIMAHU ALLAAH;So;0;ON;;;;;N;;;;;
FD41;ARABIC LIGATURE RAHIMAHUM ALLAAH;So;0;ON;;;;;N;;;;;
FD42;ARABIC LIGATURE RAHIMAHUMAA ALLAAH;So;0;ON;;;;;N;;;;;
FD43;ARABIC LIGATURE RADI ALLAAHU ANH;So;0;ON;;;;;N;;;;;
FD44;ARABIC LIGATURE RADI ALLAAHU ANHAA;So;0;ON;;;;;N;;;;;
FD45;ARABIC LIGATURE RADI ALLAAHU ANHUM;So;0;ON;;;;;N;;;;;
FD46;ARABIC LIGATURE RADI ALLAAHU ANHUMAA;So;0;ON;;;;;N;;;;;
FD47;ARABIC LIGATURE SALLALLAAHU ALAYHI WA-AALIH;So;0;ON;;;;;N;;;;;
FD48;ARABIC LIGATURE SALLALLAAHU ALAYHI WA-AALIHEE
        WA-SALLAM;So;0;ON;;;;;N;;;;;
FD49;ARABIC LIGATURE ALAYHI AS-SALAAM;So;0;ON;;;;;N;;;;;
FD4A;ARABIC LIGATURE ALAYHAA AS-SALAAM;So;0;ON;;;;;N;;;;;
FD4B;ARABIC LIGATURE ALAYHIM AS-SALAAM;So;0;ON;;;;;N;;;;;
FD4C;ARABIC LIGATURE ALAYHIMAA AS-SALAAM;So;0;ON;;;;;N;;;;;
FD4D;ARABIC LIGATURE QUDDISA SIRRAH;So;0;ON;;;;;N;;;;;
FD4E;ARABIC LIGATURE QUDDISA SIRRAHUM;So;0;ON;;;;;N;;;;;
FD4F;ARABIC LIGATURE QUDDISA SIRRAHUMAA;So;0;ON;;;;;N;;;;;
FDFE;ARABIC LIGATURE AJJALA ALLAAHU TAAALAA FARAJAHU
        ASH-SHARIF;So;0;ON;;;;;N;;;;;
```

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

Suggestion for NamesLists.txt

```
@           Honorific word ligatures
FD40   ARABIC LIGATURE RAHIMAHU ALLAAH
       x (arabic sign rahmatullah alayhe - 0612)
FD41   ARABIC LIGATURE RAHIMAHUM ALLAAH
FD42   ARABIC LIGATURE RAHIMAHUMAA ALLAAH
FD43   ARABIC LIGATURE RADI ALLAAHU ANH
       x (arabic sign radi allahou anhu - 0613)
FD44   ARABIC LIGATURE RADI ALLAAHU ANHAA
FD45   ARABIC LIGATURE RADI ALLAAHU ANHUM
FD46   ARABIC LIGATURE RADI ALLAAHU ANHUMAA
FD47   ARABIC LIGATURE SALLALLAAHU ALAYHI WA-AALIH
       x (arabic ligature sallallahou alayhe wasallam - FDFA)
FD48   ARABIC LIGATURE SALLALLAAHU ALAYHI WA-AALIHEE WA-SALLAM
FD49   ARABIC LIGATURE ALAYHI AS-SALAAM
       x (arabic sign alayhe assallam - 0611)
FD4A   ARABIC LIGATURE ALAYHA AS-SALAAM
```

FD4B ARABIC LIGATURE ALAYHIM AS-SALAAM
FD4C ARABIC LIGATURE ALAYHIMAA AS-SALAAM
FD4D ARABIC LIGATURE QUDDISA SIRRAH
FD4E ARABIC LIGATURE QUDDISA SIRRAHUM
FD4F ARABIC LIGATURE QUDDISA SIRRAHUMAA

FDFE ARABIC LIGATURE AJJALA ALLAAHU TAAALAA FARAJAHU ASH-SHARIF

Notes

1. The two characters proposed at FD47 ARABIC LIGATURE SALLALLAAHU ALAYHI WA-AALIH and FD48 ARABIC LIGATURE SALLALLAAHU ALAYHI WA-AALIHEE WA-SALLAM are alternate versions of U+FDFA ﷺ ARABIC LIGATURE SALLALLAHOU ALAYHE WASALLAM. All three are used as an honorific for the prophet Muhammad. The two new proposed characters add the word و آله "and his family", which, although used by both the sunni and the shia'a, is used more by the shia's since the prophet's family has a special importance among the shi'a.
2. The character names match the existing patterns in Unicode character names. The following patterns have been used:
    a. The 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..08D2 (see L2/14-105), as opposed to U+FDF0..FDFB and U+0610..0613, since the former pattern is more common among Arabic speakers.
3. Characters have been arranged in groups of two, three, or four. In each group, they are sorted alphabetically based on the part of the word that changes (male, female, plural, and dual, in that order).
4. Evidence of usage in fonts or text was not found for the female forms RAHIMAHAA ALLAAH and QUDDISA SIRRAHAA. They are not proposed here.

Samples of usage

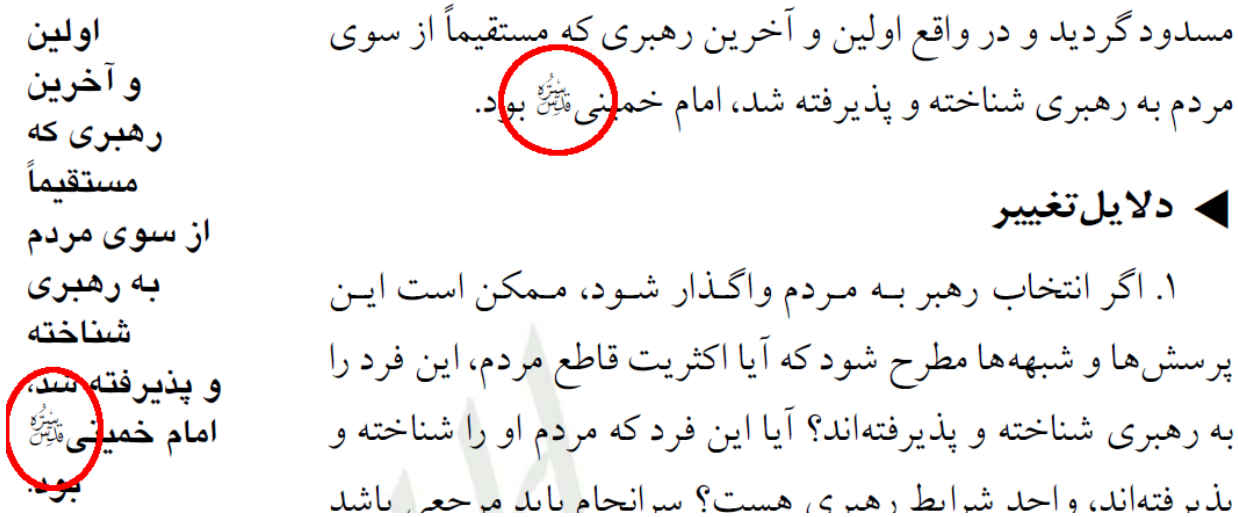


Figure 1. Sample of usage of proposed FD4D ARABIC LIGATURE QUDDISA SIRRAH from Var'i 2006. Note that both usages come from the same font, although they are surrounded by fonts of different styles and weight.

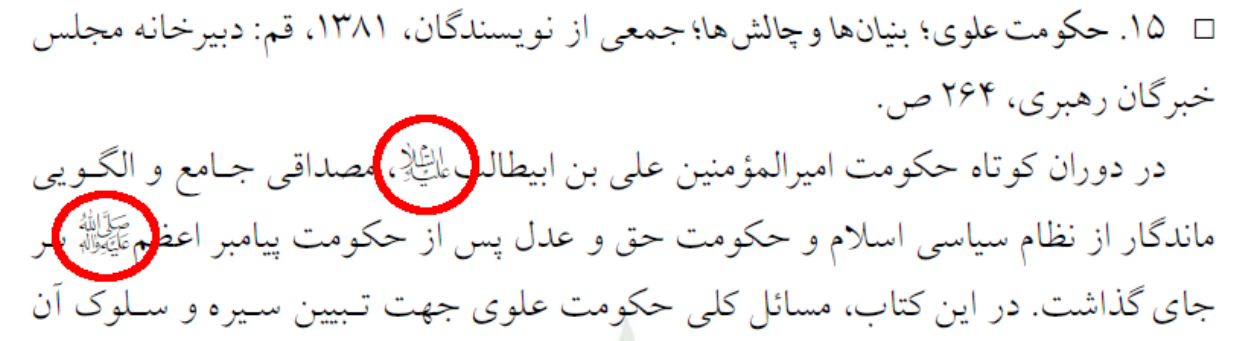


Figure 2. Sample of usage of proposed FD49 ARABIC LIGATURE ALAYHI AS-SALAAM and FD47 ARABIC LIGATURE SALLALLAAHU ALAYHI WA-AALIH from Var'i 2006.

اشخاص بدهد. ادلهٔ بیعت بالاتفاق در مورد معصوم علیه‌السلام وارد شده که ولایتشان منوط به بیعت نیست. نقش بیعت همین مقدار است «قیام الحجة بوجود الناصر» یعنی بیعت حجت را بر آنها تمام می‌کرد که یار و یاور به مقدار کافی دارند برای ایفای مسؤولیت.[۲]

● **۲.منافات توقیت با مقام الهیِ رهبری**

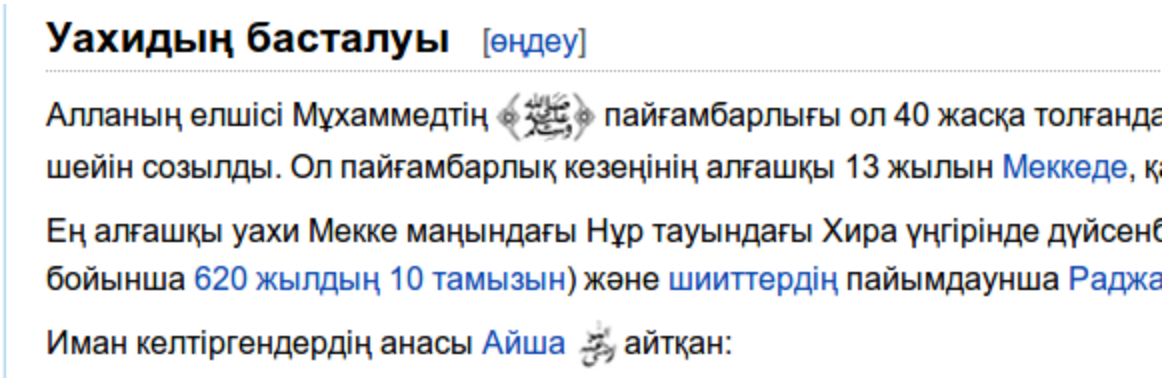
جامعهٔ مدرسین حوزهٔ علمیهٔ قم، پس از اطلاع از مشروح مذاکرات شورای بازنگری قانون اساسی، با ارسال نامه‌ای اظهار امیدواری کردند که: «اعضای محترم شورا، مسألهٔ ولایت فقیه را که منقبتی الهی است و ادامهٔ ولایت ائمه معصومین علیهم‌السلام است، محدود به زمان نکنند، که بدون تردید، محدودیت، موجب تضعیف مقام ولایت فقیه خواهد بود.»[۳]

این گروه بر این نظر بودند که:

اگر توقیت شرعاً هم منعی نداشته باشد، چون ولایت و رهبری یک مقام الهی و معنوی است و مردم رهبر را نایب حضرت ولی عصر(عج) می‌دانند و با عشق و علاقه به او می‌نگرند

سال یازدهم / شمارهٔ ۳

Figure 3. Sample of usage of proposed FD49 ARABIC LIGATURE ALAYHI AS-SALAAM and FD4B ARABIC LIGATURE ALAYHIM AS-SALAAM from Var'i 2006. Note the circled phrase on the last line, the parenthesized (عج), which is a replacement for U+FDFE ARABIC LIGATURE AJJALA ALLAAHU TAAALAA FARAJAHU ASH-SHARIF that was probably not available in the production environment.



Уахидың басталуы [өңдеу]

Алланың елшісі Мұхаммедтің ﷺ пайғамбарлығы ол 40 жасқа толғанда
шейін созылды. Ол пайғамбарлық кезеңінің алғашқы 13 жылын Меккеде, қ

Ең алғашқы уахи Мекке маңындағы Нұр тауындағы Хира үңгірінде дүйсенб
бойынша 620 жылдың 10 тамызын) және шиіттердің пайымдауынша Раджа

Иман келтіргендердің анасы Айша رضي الله عنها айтқан:

Figure 4. Sample of usage of U+FDFA ﷺ ARABIC LIGATURE SALLALLAHOU ALAYHE WASALLAM and the proposed FD43 ARABIC LIGATURE RADI ALLAAHU ANH from the Kazakh Wikipedia. The male honorific is mistakenly used after the name of Aisha, the youngest wife of the prophet. Correct grammatical usage would have been FD44 ARABIC LIGATURE RADI ALLAAHU ANHA.

AH/1135-1199 AD) and Saint Nat-har *Walī* رضي الله عنه (d.417 AH/1026 AD), and Saint Shāhul Ḥamīd *Walī* رضي الله عنه (910-978 AH/1504-1570 AD), Nadwi's book does not mention even a single person of the region. Foreign scholars living farther from the Arwi Muslims appear to know more than

Figure 5. Sample of usage of the proposed FD43 ARABIC LIGATURE RADI ALLAAHU ANH in the English language from Alim 1993, page XLI.

in the place called Periapattinam - an area in the ancient city called Pauttira Mānikka Pattinam (page 53).

(c) The tomb of *Hābīl*, son of *Ādam* عليه الصلاة والسلام, is in Rameswaram Island. The ruler of Ramnad, Muthukumara Vijaya Ragunatha Sethupathi, endowed a village called Pudukulam in 1745 to the said shrine (page 43). (When this writer visited this shrine, he

Figure 6. The honorific *alayhi as-salaat was-salaam* (عليه الصلاة والسلام) in English text, from Alim 1993, page 5. The character is not proposed here as the author has not seen it is used in other texts.

Acknowledgments

The author is grateful to Karim Senussi who helped him with information about the Sunni usage of the honorifics and proofread the character names.

Bibliography

1. Tayka Shu'ayb Alim. 1993. *Arabic, Arwi and Persian in Sarandib and Tamil Nadu- Parts I & II*. Imamul Arus Trust, Madras.
2. Wikipedia Contributors. 2014. "Құран." The Kazakh Wikipedia. Accessed June 12, 2014. http://kk.wikipedia.org/w/index.php?title=%D2%9A%D2%B1%D1%80%D0%B0%D0%BD&oldid=2088749
3. Jonathan Kew. 2001. "Proposal to add Arabic-script honorifics and other marks." UTC Document Register L2/01-425, The Unicode Consortium. http://www.unicode.org/L2/L2001/01425-arabic_marks.pdf
4. Roozbeh Pournader. 2014. "Proposal to encode fourteen Pakistani Quranic marks." UTC Document Register L2/14-105, The Unicode Consortium. http://www.unicode.org/L2/L2014/14105-pakistani-koranic-marks.pdf

5. Seyyed Javad Var'i (ed.). 2006. *Islamic Government: A Journal of Islamic Political Thought and Jurisprudence.* Vol. 11, No. 3, Oct-Dec 2006. Assembly of Experts of Iran. ISSN 1561-008X. Accessed June 12, 2014. http://www.rcipt.ir/pdfmagazines/041/041.pdf

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 seventeen Arabic honorifics***
2. Requester's name: **Roozbeh Pournader**
3. Requester Type (Member body/Liaison/Individual contribution): **Individual Contribution**
4. Submission date: **June 12, 2014**
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: **17**
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? **Probably Michael Everson**
   b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.): **N/A**
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? **No**
   If YES explain: **N/A**
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 author used to live in Iran and has seen frequent examples of usage.**
   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): **Common. Some of them comparatively rare.**
   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) provided? **N/A**

If YES, reference: **N/A**

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): **B/A**

13. Does the proposal contain any Ideographic compatibility characters? **No**

If YES, are the equivalent corresponding unified ideographic characters identified? **N/A**

If YES, reference: **N/A**