

To: Unicode Technical Committee
From: Azzeddine Lazrek (via Deborah Anderson)
Contact: lazrek@uca.ac.ma
Subject: More precisions about Hamza
Date: 27 December 2017

Some remarks about comments on page 4-6 of the script ad hoc recommendations:
Recommendations to UTC #153 November 2017 on Script Proposals L2/17-384
<http://www.unicode.org/L2/L2017/17384-script-ad-hoc-recs.pdf>
 against [L2/17-252](#) Proposal to encode some Hamza Quranic marks.

Ad hoc recommendations :

- Character #1 ◌ ARABIC HIGH WASLA

The examples show this combining character only over *alef*.

The proposal mentions "Moreover, we could need to use Wasla alone, to explain its use for example." This indeed happens, and is used in pedagogical text. We do not recommend encoding a separate combining mark due to destabilization of encoding, but a spacing mark for wasla, similar to the characters in the U+FBB2..U+FBC1 range could be encoded.

Note that a visually similar character placed over other characters was earlier proposed by Mussa Abudena (figure 1 in L2/16-056), but that is a *waqf* sign, which is graphically much larger. The script ad hoc had earlier recommended this character for encoding in L2/16-342. After discussion with Arabic experts, the suggested name put forward was ARABIC SMALL HIGH WORD SAH (with an annotation "sign of waqf").

Comments:

This proposed character has no semantic relation with the earlier proposed one with the suggested name ARABIC SMALL HIGH WORD SAH (with an annotation "sign of waqf") ; even if they have some visually similarities.

Effectively, the first one could be only over Alef. However, the second one is over the end of a sentence.

We don't need to encode this character only to be able to use Wasla alone for pedagogical uses. Mainly, we need this character to be able to add it if it appears, as it have already explained at L2/17-068, (see Table 1 – Variations of an encoding word Page 3).

For example, how to codify the following word in different situations:

اعبدوا	ا 0627	X	ع 0639	ب 0628	د 062F	و 0648	ا 0627
--------	-----------	---	-----------	-----------	-----------	-----------	-----------

Where X could be:

	اعبدوا	أعبدوا	آعبدُوا	اَعْبَدُوا	اَعْبَدُوا
X	Nothing	◌	◌	◌ and ◌	◌ and ◌

- The variant  must be encoded, by the sequence:  U+0627 and  0654, but not by:  U+0623, even if it exists. It will be recommended to not use  U+0623 anymore.
- The variant  must be encoded, by the sequence:  U+0627 and  , but not by:  U+0671, even if it exists. It will be recommended to not use  U+0671 anymore.
- The variant  must be encoded by the sequence:  U+0627 and  and  , but not only by one code.
- The variant  must be encoded by the sequence:  U+0627 and  and  , but not only by one code, as it is proposed by Mussa Abudena L2/15-329 and L2/16-044.

By , any research engine will be able to find any occurrence of the preceding word at all its situations.

Regarding the ad hoc recommendations' comment:

We do not recommend encoding a separate combining mark due to destabilization of encoding, but a spacing mark for wasla, similar to the characters in the U+FBB2..U+FBC1 range could be encoded.

...

We recommend ... a revised proposal should be prepared, with references to the appropriate examples:

a. ARABIC SYMBOL WASLA ABOVE (at U+FBC2)

???

Ad hoc recommendations :

We recommend ... a revised proposal should be prepared, with references to the appropriate examples:

[...]

c. If the size of the stroke can be resolved (and code points to be determined)

ALEF WITH ATTACHED FATHA (whose glyph has a stroke at the top)

ALEF WITH MIDDLE STROKE

ALEF WITH ATTACHED KASRA (whose glyph has a stroke on the bottom)

Comments:

In our opinion, for the same reason, I do not agree to encode ALEF WITH some symbols. However, I have proposed to encode those symbols ALONE.

(The numbers in the far left column below refer to proposed characters from [L2/17-252](#).)

- Characters #3 and #5 have same semantic with different shapes, we find both in practice.

3	08C2		ARABIC SMALL DOT LEFT • Starting Hamza Wasl with Damma → 302E • 
11	08CA		ARABIC BIG DOT LEFT • Facilitate Hamza with Damma → 302E • 

- Characters #4 and #6 have same semantic with different shapes, we find both in practice.

4	08C3		ARABIC SMALL DOT BELOW • Starting Hamza Wasl with Kasra
---	------	---	--

6	08C5		ARABIC SMALL RING BELOW • Starting Hamza Wasl with Kasra used at some Mushaf Qaloon
---	------	---	--

- Characters #11 and #12 have same semantic with different positions, we find both in practice.

11	08CA		ARABIC BIG DOT LEFT • Facilitate Hamza with Damma → 302E • 
12	08CB		ARABIC BIG DOT RIGHT • Facilitate Hamza with Damma used at some Mushaf(s) Qaloon → 08CA • 

- Characters #2 and #10 have similar shape with variant sizes (the first one is smaller than the second one) but differ semantically.

2	08C1		ARABIC SMALL DOT ABOVE • Starting Hamza Wasl with Fatha
---	------	---	--

10	08C9		ARABIC BIG DOT ABOVE • Facilitate Hamza with Fatha • Used for Replace Hamza with Fatha
----	------	---	--

- Characters #3 and #11 have similar shape with variant sizes (the first one is smaller than the second one) but differ semantically.

3	08C2		ARABIC SMALL DOT LEFT • Starting Hamza Wasl with Damma → 302E • 
11	08CA		ARABIC BIG DOT LEFT • Facilitate Hamza with Damma → 302E • 

- Characters #4 and #13 have similar shape with variant sizes (the first one is smaller than the second one) but differ semantically.

4	08C3		ARABIC SMALL DOT BELOW • Starting Hamza Wasl with Kasra
---	------	---	--

13	08CC		ARABIC BIG DOT BELOW • Facilitate Hamza with Kasra • Used also for Replace Hamza with Kasra
----	------	---	---

- Characters #7 and #16 have similar shape with variant sizes (the first one is smaller than the second one) but differ semantically.

7	08C6		ARABIC SMALL STROKE ABOVE • Used for Hamza Wasl after a letter with Fatha
---	------	---	--

16	08CF		ARABIC BIG STROKE ABOVE • Transport Hamza with origin Fatha • Used also for Replace Hamza with Fatha
----	------	---	--

- Characters #8 and #18 have similar shape with variant sizes (the first one is smaller than the second one) but differ semantically.

8	08C7		ARABIC SMALL STROKE RIGHT • Used for Hamza Wasl after a letter with Damma
---	------	---	--

18	08D1		ARABIC BIG STROKE RIGHT • Transport Hamza with origin Damma in some printed Mushaf(s)
----	------	---	--

- Characters #9 and #19 have similar shape with variant sizes (the first one is smaller than the second one) but differ semantically.

9	08C8		ARABIC SMALL STROKE BELOW • Used for Hamza Wasl after a letter with Kasra → U+061A  Arabic small Kasra → U+0650  Arabic Kasra
19	08D2		ARABIC BIG STROKE BELOW • Transport Hamza with origin Kasra • Used also for Replace Hamza with Kasra → U+061A  Arabic small Kasra → U+0650  Arabic Kasra

- Characters #17 and #18 have same semantic with different positions, we find both in practice.

17	08D0		ARABIC BIG STROKE LEFT • Transport Hamza with origin Damma • Used also for Replace Hamza with Damma
18	08D1		ARABIC BIG STROKE RIGHT • Transport Hamza with origin Damma in some printed Mushaf(s)

Ad hoc recommendations :

Character #5  ARABIC SMALL RING LEFT and Character #6  ARABIC SMALL RING BELOW

What is the character in Figure 36.b (see below, from L2/17-252)? Is it #6?



Comments:

Characters #6 and one used in Figure 36.b have similar shape with variant sizes (the first one is smaller than the second one) but differ semantically.

The first one is used for Starting Hamza Wasl with Kasra (find at some Mushaf Qaloon). However, the second one is used for Al-Imala (find at some Mushaf).

The first one is used only under Alef. However, the second one is used under some other letters.