

To: UTC and script ad-hoc
 Title: Clarification on spelling of lam-alef ligatures for al-Dani
 From: Lorna Priest Evans (SIL International)
 Date: 19 October 2021
 Reference: L2/16-044, L2/19-306, L2/21-130

Introduction

[L2/16-044](#) and [L2/19-306](#) were the documents proposing the new alef characters in Unicode 14.0 in the U+0870..U+0882 range.

Public Review feedback from Patrik Sjöwall (see [L2/21-130](#) page 23) expresses the need for further information in implementing the lam/alef ligature. “These letters require more shaping information. It is not clear how the attached *fatha* or dot will behave in an obligatory LAM-ALEF ligature.”

The original proposal document (L2/16-044) clearly says these are used for “Aldani” style Qurans (in contrast to the more common “Hafs” style. In researching how to implement these *alef* characters in a lam/alef ligature, I came to the realization that in Aldani, the stroke representing the lam and alef are opposite of the stroke in Hafs.

If we look at two characters in the presentation forms (FEF5, FEF6) we can see that, at least for the representative glyph, it is the right part of the stroke that is the *lam* (blue) and it is the left part (red) of the stroke that it is the *alef*. This is evidenced in two ways, one the glyph name itself (which says “alef with madda above”) and also the decompositions which are to isolate and final forms of “0644 0622”.

FEF5 ARABIC LIGATURE LAM WITH ALEF WITH MADDA ABOVE ISOLATED FORM

FEF6 ARABIC LIGATURE LAM WITH ALEF WITH MADDA ABOVE FINAL FORM



Now, if we take a look at L2/16-044, page 4, we can see a chart (excerpted below) which illustrates Hafs (Khrraz) and Aldani *lam-alef* ligatures. These were proposed as new characters in L2/16-044. At the far right (second to last column) we see the Hafs (Khrraz) column first and Aldani is the last column. The “name” column was the suggested name for the characters proposed in the Aldani column. It is clear from this example that the *alef* is the right stroke in Aldani and the *lam* is the left part of the stroke. Thus, the issue is actually bigger than the new *alef* characters added to Unicode 14.0. There are implementation ramifications for all *lam-alef* ligatures.

Comparison between the writing of Quran in narration Hafs (Khrraz) and narration Qaloon (Aldani)

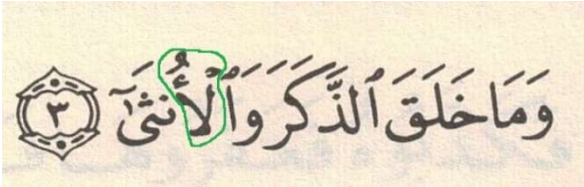



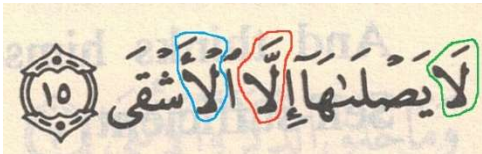
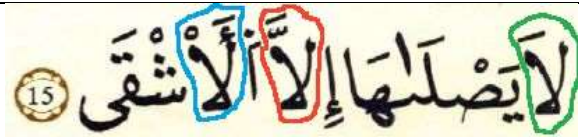
No.	Name	Khrraz	Aldani
31	ARABIC LAM ALEF WITH SUKUN ABOVE LAM AND HAMZA JOINED WITH ALEF AND DAMMA ABOVE		
32	ARABIC LAM ALEF WITH FATHA ABOVE		
33	ARABIC LAM ALEF WITH FATHA ABOVE LAM AND MADDAH ABOVE		

This has clear spelling ramifications. If we take line 33, what is the encoding? For Hafs, the encoding is U+0644 U+064E U+0622 (lam, fatha, alef with madda). Is the encoding the same for Aldani? Or does the al-Dani (we will use al-Dani from here onwards as that is the term L2/19-306 uses) display require the encoding to be U+0644 U+0653 U+0627 U+064E (lam, madda, alef, fatha)? Or even U+0622, U+0644, U+064E (alef-madda, lam, fatha)?

If both al-Dani and Hafs are spelled the same, then font implementations for al-Dani will need to consider the right stroke to be *alef* and the left stroke to be *lam*. In fonts that try to support both readings, a feature would need to be added which changes the position of where combining marks attach (depending on whether the mark is on the *alef* or the *lam*).

Taking this situation a step further, let us look at line 31 in the table above. The spelling for the Hafs glyph is *lam, sukun, alef, hamza, damma*. Because of the new U+0881 character, the al-Dani spelling must be *lam, sukun, alef with attached right hamza, damma*.

Another example, using Hafs and al-Dani (example from L2/16-044, Figure 9), demonstrates that the al-Dani reading also has the *alef* as the right stroke.

Hafs (alef is left side of ligature)	al-Dani (alef is right side of ligature)
 <p>lam, sukun, alef, hamza, damma</p>	 <p>lam, sukun, alef with attached right hamza, damma</p>
 <p>1: lam, fatha, lam, sukun, alef, hamza, fatha 2: lam, sukun, alef, hamza, damma</p>	 <p>1: lam, fatha, kashida, hamza, fatha, lam, sukun, alef 2: lam, sukun, alef with attached right hamza, damma</p>
 <p>1: lam, fatha, alef 2: lam, shadda, fatha, alef 3: lam, sukun, alef, hamza, fatha</p>	 <p>1: lam, fatha, alef 2: lam, shadda, fatha, alef 3: lam, sukun, alef, hamza, fatha</p>

Decisions and Implications

This document attempts to demonstrate the different orthographic interpretations of the classical lam-alef ligature form when viewed with Hafs conventions vs those of al-Dani and the difficulties these interpretations raise in determining correct encoding (spelling). In the absence of guidance from Unicode there is a risk of text authors and font developers choosing different and incompatible approaches to meet al-Dani needs.

The key guidance needed is whether texts, for example as shown in the above charts and samples, should, where feasible, be encoded (i.e., spelled) the same regardless of whether Hafs or al-Dani rendering is used, or whether they should be spelled differently for Hafs and for al-Dani. (Note that in some cases different characters will be used, e.g., U+0623 ARABIC LETTER ALEF WITH HAMZA ABOVE in Hafs vs U+0881 ARABIC LETTER ALEF WITH ATTACHED RIGHT HAMZA in al-Dani, but the relative ordering of the various base and mark characters would be the same.)

If the decision is to spell them the same, then most texts will be interchangeable and the rendering differences achieved by font choice. (It is technically possible to implement a font feature to switch conventions but the font logic is more complex and probably unlikely to be widely implemented.)

If the decision is to spell them differently, then texts are not interchangeable but, depending on exact spelling details, it might be relatively easy for a single font to support both. In this case specific recommendations for the correct sequence of letters (both base letters and marks) for al-Dani conventions is needed.