

PROPOSAL TO ENCODE THREE QURANIC ARABIC CHARACTERS

Rikza F. Sh.

ريكزا ف. ص.

rikzafsh@gmail.com

7 July 2022

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

1. Introduction

The Quran is a holy book for Moslems that was revealed by Allah ﷻ to the Prophet Muhammad ﷺ through the intercession of the angel Jibrīl عليه السلام. At first, the Quran was only memorized by the companions of the Prophet ﷺ and was not written into a complete book. Until the caliph Uthman bin ‘Affan عليه السلام codified and established a standard version, which becomes the basis for writing most of the Qurans that exist today.

At that time, the Quran text was only written with the body letter without any dots or other diacritical marks as in today’s Quran. This is because the Arabic Moslems can easily read it depend on the context. The additional dots and other diacritical marks are later added to ease the Moslems who can’t speak Arabic to recite the Quran properly.

However, this additional diacritical marks may be different across different Qurans. For example, the Madinah mushaf has a different style of additional diacritical marks compared to the Indonesian mushaf. Although this style difference is only orthographic and not affected on recitation nor the meaning of the Quran.

Currently, most of the characters needed to type the Quran text are well-encoded in Unicode. However, there are several special characters used in various mushaf that not yet encoded.

2. Character to be encoded

2.1. ARABIC RAISED SMALL ALEF

This Character has similar shape and function with U+0670 ARABIC LETTER SUPERScript ALEF. The difference is that U+0670 is a non-spacing combining mark, while ARABIC RAISED SMALL ALEF is a spacing mark like U+06E5 ARABIC SMALL WAW and U+06E6 ARABIC SMALL YEH.

These SMALL ALEF characters, either U+0670 and ARABIC RAISED SMALL ALEF have the following function:

- To mark unwritten alef in Uthmanic rasm

In Madinah mushaf, the unwritten alef is marked by a SMALL ALEF after the previous letter. If the previous letter is in initial or medial form, the sequences <[letter], U+064E ARABIC FATHA, U+0640 ARABIC TATWEEL, U+0670 ARABIC LETTER SUPERScript ALEF> should be used. If the previous letter is in isolated or final form, the sequences <[letter], U+064E ARABIC FATHA, ARABIC RAISED SMALL ALEF> should be used.

In Iran mushaf, the unwritten alef is marked with the same way as in Madinah mushaf, except without U+064E ARABIC FATHA after the previous letter.

In Indonesian mushaf (also in Indian and Pakistan mushaf), the unwritten alef is marked with SMALL ALEF placed above the previous letter without U+064E ARABIC FATHA regardless of what form the previous letter is.

Verse	Madinah Mushaf [2]	Iran Mushaf [3]	Indonesian Mushaf [4]	Explanation
2:2	اذلِكَ الْكِتَابُ	ذَلِكَ الْكِتَابُ	ذَلِكَ الْكِتَابُ	a)
2:29	سَمَوَاتٍ	سَمَوَاتٍ	سَمَوَاتٍ	b)
2:35	يَقَادِمُ	اِيَادِمُ	اِيَادِمُ	c)
2:40	يَبْنِي إِسْرَائِيلَ	يَبْنِي إِسْرَائِيلَ	يَبْنِي إِسْرَائِيلَ	d)
2:43	الرَّكْعَيْنِ	الرَّكْعَيْنِ	الرَّكْعَيْنِ	e)

- The spacing small alef after isolated letter ذ and combining small alef over tatweel after medial letter ذ in Madinah and Iran mushafs. In Indonesian mushaf the combining small alef is used over both letters.
- The combining small alef over tatweel after medial letter س and spacing small alef after final letter ت in Madinah and Iran mushafs. In Indonesian mushaf the combining small alef is used over both letters.
- The combining small alef over tatweel after initial letter ا in Madinah and Iran mushafs, and above the initial letter ا in Indonesian mushaf.
- The combining small alef over tatweel after initial letter ب in Madinah and Iran mushafs, and above the initial letter ب in Indonesian mushaf. The spacing small alef after final letter ل in Madinah mushaf. Iran and Indonesian mushafs use the regular alef instead.
- The spacing small alef after final letter ن in Madinah and Iran Mushaf, the position of small alef is in the left of shadda. The combining small alef above the shadda on the final letter ن in Indonesian mushaf.

- To mark the و U+0648 ARABIC LETTER WAW, ى U+0649 ARABIC LETTER ALEF MAKSURA, ي U+064A ARABIC LETTER YEH or ي U+06CC ARABIC LETTER YEH FARSI to be threaten as letter alef (to mark long /a:/ vowel) in Madinah and Iran mushaf. In Indonesian mushaf the small alef is placed above the previous letter and the letter waw or yeh is ignored and no diacritic is placed.

Verse	Madinah Mushaf [2]	Iran Mushaf [3]	Indonesian Mushaf [4]	Explanation
2:2	الصَّلَاةِ	الصَّلَاةِ	الصَّلَاةِ	a)

2:7	وَعَلَىٰ	وَعَلَىٰ	وَعَلَىٰ	b)
2:29	أَسْتَوَىٰ	استوى	استوى	c)

- The combining small alef above final letter و in Madinah and Iran mushafs, and above medial letter ل in Indonesian mushaf.
- The combining small alef above final letter ع in Madinah and Iran mushafs, and above medial letter ل in Indonesian mushaf.
- The combining small alef above isolated letter ع in Madinah and Iran mushafs, and above final letter و in Indonesian mushaf.

In Madinah mushaf, the spacing SMALL ALEF that marks the unwritten alef after an isolated or final letter may be handled by a Quran specific font that automatically adjusts the position of U+0670 ARABIC LETTER SUPERSCRIPT ALEF if it inserted after U+064E ARABIC FATHA. But most of common Arabic font can't handle this case, they will stack the SMALL ALEF over FATHA like in some places in Mushaf Tunis Qaloon:



Figure 1. Showing SMALL ALEF stacked over FATHA in 19:96, [7] p.312

Unifying a combining and spacing SMALL ALEF could be ambiguous in the case like in Iran Mushaf which has contrast use between those two characters. If the SMALL ALEF comes after LETTER WAW, it could be a confusion, is it marking the unwritten alef so the LETTER WAW is pronounced as /wa:/, or is it marking the LETTER WAW to be threatened as a LETTER ALEF so the letter before waw is read with long vowel /a:/?.

The sequence <00A0 NO BREAK SPACE, 0670 ARABIC LETTER SUBSCRIPT ALEF> may be used to produce spacing SMALL ALEF. But most of Arabic font may be not give the right position for the spacing SMALL ALEF. And since the combining and spacing modifier mark are usually separated as different character in Unicode (like U+06E7 ARABIC SMALL HIGH YEH with U+06E6 ARABIC SMALL YEH, U+06D3 ARABIC SMALL LOW WAW with U+06E5 ARABIC SMALL WAW, and also Combining Diacritical Marks block and Spacing Modifier Letters block), I think it's better to encode ARABIC RAISED SMALL ALEF as a separate character with U+0670 ARABIC LETTER SUPERSCRIPT ALEF.

I named "raised small alef" because the its position is raised from the baseline, unlike U+06E6 ARABIC SMALL YEH that usually appear exactly above the baseline. Also, the character U+0888 which named ARABIC RAISED ROUND DOT compared to U+0887 ARABIC BASELINE ROUND DOT.

The character ARABIC RAISED SMALL ALEF can also written twice to mark longer vowel.

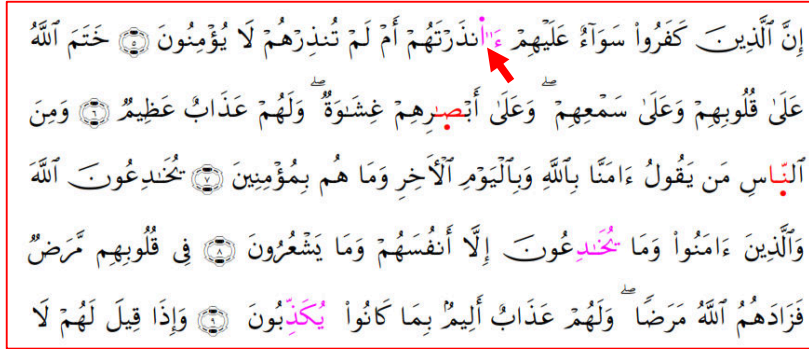


Figure 2. Showing ARABIC RAISED SMALL ALEF written twice. From [5] p.3

The character ARABIC RAISED SMALL ALEF can also combined with combining marks, like U+06E4 ARABIC SMALL HIGH MADDA, and U+08CE ARABIC LARGE ROUND DOT ABOVE.

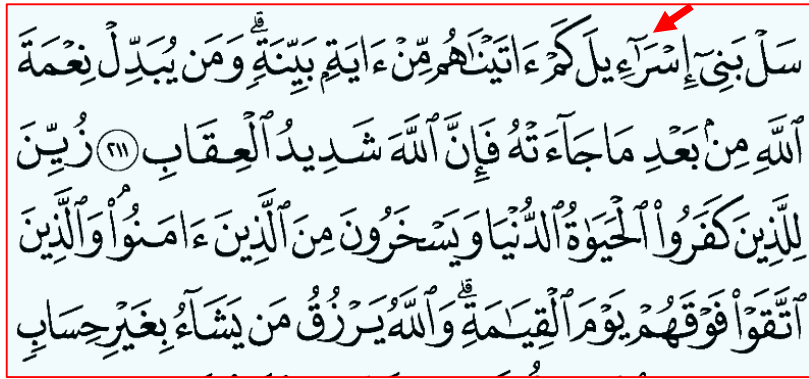


Figure 3. Showing ARABIC RAISED SMALL ALEF combined with ARABIC SMALL HIGH MADDA. From [2] p.33

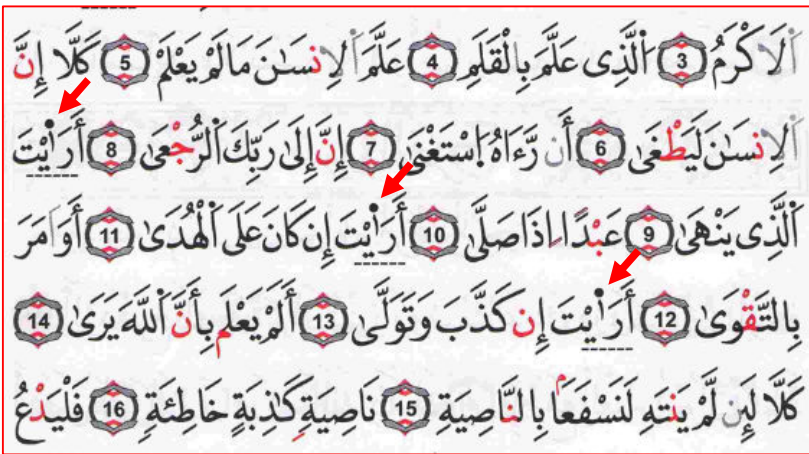


Figure 4. Showing ARABIC RAISED SMALL ALEF combined with ARABIC LARGE ROUND DOT ABOVE. From [6] p.597.

The character ARABIC RAISED SMALL ALEF may be confusable with U+0670 ARABIC LETTER SUPERSCRIPIT ALEF and U+08AD ARABIC LETTER LOW ALEF since their shapes are very similar (or even identical), only their positions are different. The character ARABIC LETTER SUPERSCRIPIT ALEF is used when the “small alef” is placed exactly above the letter or tatweel. The character ARABIC

LETTER LOW ALEF is a spacing character, and its position is lowered to the baseline. The character ARABIC RAISED SMALL ALEF is a spacing character like ARABIC LETTER LOW ALEF but the position is raised. The character ARABIC RAISED SMALL ALEF is used after a final or isolated letter and it's placed to the left of the letter, instead of placed exactly above the letter like ARABIC LETTER SUPERSCRIPT ALEF does.

We can find the contrast use of those character on Quran surah Al-Baqarah 2:71-74 as in image from various mushafs below. The character ARABIC RAISED SMALL ALEF is marked with red arrow, the character ARABIC LETTER LOW ALEF is marked with green arrow, and the character ARABIC LETTER SUPERSCRIPT ALEF is marked with blue arrow.

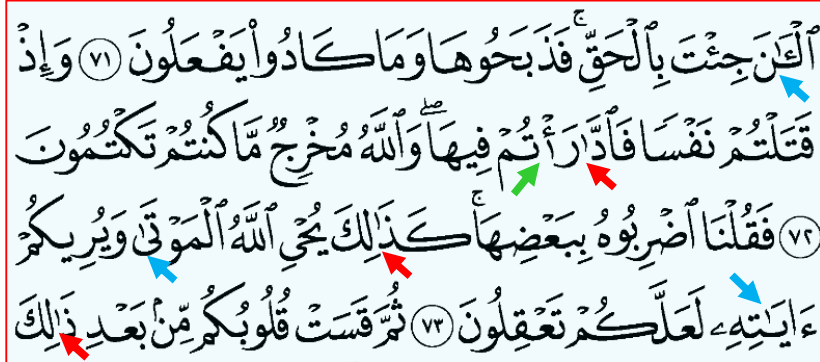


Figure 5. From [2] p.11.

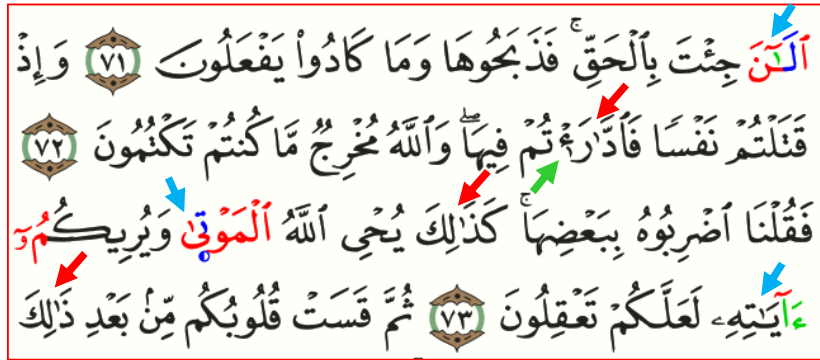


Figure 6. From [8] p.10.

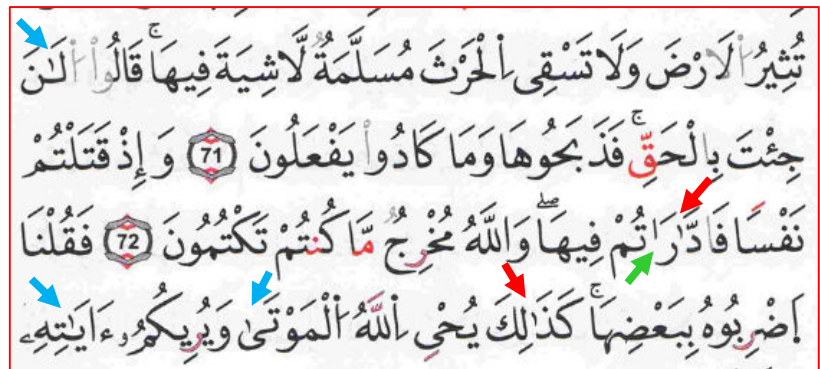


Figure 7. Showing two types of small alef. Note that in the word فادراً, the second small alef is appeared as a straight top-half alef. From [6] p.11.

تُثِيرُ الْأَرْضَ وَلَا تَسْقِي الْحَرْثَ مُسَلِّمَةً لَأَشِيَةِ فِيهَا قَالُوا أَلَكُنْ
 جِئْتَ بِالْحَقِّ فَذَبْحُوهَا وَمَا كَادُوا يَفْعَلُونَ ﴿٧١﴾ وَإِذْ قَتَلْتُمْ
 نَفْسًا فَادْرَأْتُمْ فِيهَا وَاللَّهُ مُخْرِجٌ مَّا كُنْتُمْ تَكْتُمُونَ ﴿٧٢﴾ فَقُلْنَا
 أَضْرِبُوهُ بِبَعْضِهَا كَذَلِكَ يُحْيِي اللَّهُ الْمَوْتَى وَيُرِيكُمْ آيَاتِهِ

Figure 8. Showing two types of small alef. Note that in the word فادرم, the second small alef is appeared as a straight top-half alef. From [7] p.11.

أَلَكُنْ جِئْتَ بِالْحَقِّ فَذَبْحُوهَا وَمَا كَادُوا يَفْعَلُونَ ﴿٧٠﴾ وَإِذْ
 قَتَلْتُمْ نَفْسًا فَادْرَأْتُمْ فِيهَا وَاللَّهُ مُخْرِجٌ مَّا كُنْتُمْ تَكْتُمُونَ ﴿٧١﴾
 فَقُلْنَا أَضْرِبُوهُ بِبَعْضِهَا كَذَلِكَ يُحْيِي اللَّهُ الْمَوْتَى وَيُرِيكُمْ
 آيَاتِهِ لَعَلَّكُمْ تَعْقِلُونَ ﴿٧٢﴾ ثُمَّ قَسَتْ قُلُوبُكُم مِّنْ بَعْدِ ذَلِكَ

Figure 9. From [9] p.11

أَلَكُنْ جِئْتَ بِالْحَقِّ فَذَبْحُوهَا وَمَا كَادُوا يَفْعَلُونَ ﴿٧١﴾ وَإِذْ
 قَتَلْتُمْ نَفْسًا فَادْرَأْتُمْ فِيهَا وَاللَّهُ مُخْرِجٌ مَّا كُنْتُمْ تَكْتُمُونَ ﴿٧٢﴾
 فَقُلْنَا أَضْرِبُوهُ بِبَعْضِهَا كَذَلِكَ يُحْيِي اللَّهُ الْمَوْتَى وَيُرِيكُمْ
 آيَاتِهِ لَعَلَّكُمْ تَعْقِلُونَ ﴿٧٣﴾ ثُمَّ قَسَتْ قُلُوبُكُم مِّنْ بَعْدِ ذَلِكَ

Figure 10. From [10] p.11

- ١- وَإِذْ قَتَلْتُمْ نَفْسًا فَادْرَأْتُمْ فِيهَا: الكل عدا المذكورين لاحقًا.
- ٢- فَادْرَأْتُمْ فِيهَا: السوسى.
- ٣- قَتَلْتُمْ نَفْسًا فَادْرَأْتُمْ فِيهَا: قالون والمكى.
- ٤- فَادْرَأْتُمْ فِيهَا: أبوجعفر.

Figure 11. From [13] p.135

Note that the second small alef in the word فَأَذَرْتُمْ is should be produced with sequences <08AD, 0654, 06E1> ARABIC LETTER LOW ALEF, ARABIC HAMZA ABOVE, ARABIC SMALL HIGH DOTLESS HEAD OF KHAH, and not <0621, 0656, 06E1> ARABIC LETTER HAMZA, ARABIC SUBSCRIPT ALEF, ARABIC SMALL HIGH DOTLESS HEAD OF KHAH. This is because the character U+0656 ARABIC SUBSCRIPT ALEF was intended to be used as vowel sign long /i:/. If we use the sequences <0621, 0656, 06E1>, the position of letter hamza will be placed above the baseline, and the “subscript alef” will be placed below the baseline فَأَذَرْتُمْ.

Other examples of the uses of ARABIC RAISED SMALL ALEF:

- ١- تسهيل الهمزة المفتوحة بعد همزة استفهام وإدخال أَلِفٍ بَيْنَهُمَا، في نحو: ﴿أَنْذَرْتَهُمْ﴾ [البقرة: ٦].^(١)
- ٢- تحقيق الهمزة المضمومة بعد همزة استفهام وإدخال أَلِفٍ بَيْنَهُمَا، في نحو: ﴿أَنْزَلْنَاهُ﴾ [القرن: ٢٥].^(٢)
- ٣- تحقيق الهمزة المكسورة بعد همزة استفهام وإدخال أَلِفٍ بَيْنَهُمَا، في نحو: ﴿أَنْزَلْنَاكَ﴾ [يوسف: ٩٠].^(٣)
وَيُسْتَثْنَى لَهُ سِتَّةُ مَوَاضِعَ لَيْسَ لَهَا فِيهَا إِلَّا التَّحْقِيقُ مَعَ الإِدْخَالِ، وَهِيَ ﴿أَنْزَلْنَاكُمْ﴾ [الأعراف: ٨١]، ﴿أَنْزَلْنَا﴾ [الأعراف: ١١٣]، ﴿أَنْزَلْنَا﴾ [مريم: ٦٦]، ﴿أَنْزَلْنَا﴾ [الشعراء: ٤١]، ﴿أَنْزَلْنَاكَ﴾ [الصافات: ٥٢]، ﴿أَنْزَلْنَاكَ﴾ [الصافات: ٨٦]، بالإضافة إلى مَوْضِعٍ سَابِعٍ فِي سُورَةِ فَصَّلَتْ، لَهُ فِيهِ التَّسْهِيلُ أَيْضًا، وَهُوَ التَّالِي ذِكْرُهُ.
- ٤- تسهيل الهمزة المكسورة بعد همزة استفهام وإدخال أَلِفٍ بَيْنَهُمَا، في ﴿أَبْتَكُمُ﴾ [فصلت: ٩].^(٤)

Figure 12. Showing ARABIC RAISED SMALL ALEF after letter hamza or alef hamza. From [13] p.75

2.2. ARABIC SMALL LOW NOON

This character has similar shape and function with U+08D9 ARABIC SMALL LOW NOON WITH KASRA, but without the kasra below the letter noon. This character is used in Mushaf Standar Indonesia (The standard mushaf published in Indonesia) contrast with U+08D9. The small low noon character is used to mark noon wasl when there is tanwin (U+064B..064D) before the letter alef wasla¹. Here are the rules of adding noon wasl in Indonesian standard mushaf:

- If the harakat tanwin is DAMMATAN (U+064C) or KASRATAN (U+064D), then those tanwin become regular DAMMA (U+064F) and KASRA (U+0650), and the low noon is written with kasra (U+08D9).



Figure 13. Showing with U+08D9 ARABIC SMALL LOW NOON WITH KASRA preceded with U+064F ARABIC DAMMA. From [4] p.333

¹ The wasla symbol is not written in Indonesian mushaf, so it will be appeared as regular alef

- If the harakat tanwin is fathatan (U+064B), the tanwin is not changed into regular fatha (U+064E), and the low noon is written without kasra. Except when fathatan is placed over the letter teh marbuta (U+0629), fathatan becomes regular fatha, and the low noon is written with kasra.



Figure 14. Showing ARABIC SMALL LOW NOON without kasra preceded with U+064B ARABIC FATHATAN. From [4] p.335

- If the harakat tanwin is in the end of ayah with waqf sign ﷞ (U+06D6) or ﷞ (U+06D9), then the low noon is written without kasra.



Figure 15. Showing ARABIC SMALL LOW NOON without kasra preceded with sign the end of ayah with waqf sign ﷞ (U+06D9) ARABIC SMALL HIGH LAM ALEF. From [4] p.601

The small low noon without kasra character is also attested in other mushafs, but it usually doesn't have contrast use with the low noon with kasra.



Figure 16. Showing ARABIC SMALL LOW NOON without kasra. From [14] p.33

2.3. ARABIC SMALL YEH BARREE WITH TWO DOTS BELOW

This character has similar shape with U+08CB ARABIC SMALL HIGH YEH BARREE WITH TWO DOTS BELOW. However, U+08CB is a non-spacing mark, and the ARABIC SMALL YEH BARREE WITH TWO DOTS BELOW is a spacing mark like U+06E6 ARABIC SMALL YEH. This character is used in Mushaf Standar Indonesia to mark unwritten yeh in Uthmanic rasm.



Figure 17. Showing ARABIC SMALL YEH BARREE WITH TWO DOTS BELOW and with fatha above. From [4] p.380

The unwritten yeh in Uthmanic rasm is usually written with small yeh in modern mushafs. However, the shape and position of small yeh could be different across different mushafs. Here's the comparison:

Verse	Madinah Mushaf [2]	Indo-Pak Mushaf [11]	Indonesian Mushaf [4]	Tripoli Mushaf [12]
7:196	إِنَّ وَلِيِّ اللَّهِ	إِنَّ وَلِيِّ اللَّهِ	إِنَّ وَلِيِّ اللَّهِ	إِنَّ وَلِيِّ اللَّهِ
25:49	لِنُحِّيَ بِهِ	لِنُحِّيَ بِهِ	لِنُحِّيَ بِهِ	لِنُحِّيَ بِهِ
46:33	يُحْيِي الْمَوْتَى	يُحْيِي الْمَوْتَى	يُحْيِي الْمَوْتَى	يُحْيِي الْمَوْتَى
27:36	أَتَأْتِنِي اللَّهُ	فَمَا آتَنِي اللَّهُ	أَتَأْتِنِي اللَّهُ	أَتَأْتِنِي اللَّهُ

As we see in the table, The Madinah Mushaf used the SMALL YEH with the shape like ے U+06D2 ARABIC YEH BARREE. But the position of SMALL YEH is depend on the word's final letter. If the word ends with ى U+06CC ARABIC LETTER FARSI YEH, the SMALL YEH inserted before ى and it's placed above the base letter. The tatweel (U+0640) could be used to stretch character and U+06E7 ARABIC SMALL HIGH YEH is inserted above the tatweel. If the word ends with any letter except ى, the position of small yeh is inserted after the word's final letter as a spacing mark, so U+06E6 ARABIC SMALL YEH should be used.

In the Indo-Pak Mushaf, the shape of SMALL YEH is same with Madinah Mushaf. But the position of SMALL YEH is always above the word's final letter regardless of what final letter it is. So, U+06E7 ARABIC SMALL HIGH YEH should be used.

The Indonesian Mushaf used the SMALL YEH with the shape like ﺀ U+06D2 ARABIC YEH BARREE but it has two dots below. The SMALL YEH is always inserted after the word's final letter as a spacing mark regardless of what final letter it is. Currently, this character is not encoded in Unicode. So, I think we need to encode this character.

The Tripoli Mushaf used two different shapes of SMALL YEH. If the word ends with ﻯ U+06CC ARABIC LETTER FARSI YEH, the shape of SMALL YEH is like ﺀ U+06D2 ARABIC YEH BARREE but it has two dots below the base line and inserted before ﻯ. So, ﺀ U+0885 ARABIC TATWEEL WITH TWO DOTS BELOW and U+06E7 ARABIC SMALL HIGH YEH could be used. If the word ends with any letter except ﻯ, the shape of SMALL YEH is like ﻯ and it's placed above the word's final letter. So, U+08CA ARABIC SMALL HIGH FARSI YEH should be used.

As ﺀ U+0885 ARABIC TATWEEL WITH TWO DOTS BELOW encoded as an atomic character, I think ARABIC SMALL YEH BARREE WITH TWO DOTS BELOW should be encoded as an atomic character too, and not represented by sequence U+06E6 ARABIC SMALL YEH and combining two dots below. And although it is a "small" yeh, the size of those two dots are as big as dots in regular letter yeh ﻯ.

3. Character Data

Codepoint

Regarding L2/21-181, suggesting that historical and technical character should be encoded in SMP, I propose these three Quranic Arabic characters in Arabic Extended-C block, specifically in codepoint U+10EF9..10EFB along with other Quranic characters².

Unicode character properties

```
10EF9; ARABIC RAISED SMALL ALEF;Lm;0;AL;;;;N;;;;;
10EFA; ARABIC SMALL LOW NOON;Mn;230;NSM;;;;N;;;;;
10EFB; ARABIC SMALL YEH BARREE WITH TWO DOTS BELOW;Lm;0;AL;;;;N;;;;;
```

Annotation

The following annotation could be updated in Namelist.txt

```
08AD; ARABIC LETTER LOW ALEF
  * Bashkir, Tatar
  * also used in Quranic orthography
10EF9; ARABIC RAISED SMALL ALEF
  x (arabic letter superscript alef - 0670)
  x (arabic letter low alef - 08AD)
10EFA; ARABIC SMALL LOW NOON
  x (arabic small low noon with kasra - 08D9)
10EFB; ARABIC SMALL YEH BARREE WITH TWO DOTS BELOW
  x (arabic small yeh - 06E6)
  x (arabic small high yeh barree with two dots below - 08CB)
```

² U+10EFC which has been accepted for future version of Unicode, and U+10EFD..10EFF which have been accepted for Unicode 15.0

4. Other unencoded characters.

I have found about 300 unencoded Arabic characters. Most of them are Quranic characters that I found in various mushafs. Some of them can be observed in some figures in this proposal, like ARABIC LEFT-HALF-FILLED CIRCLE BELOW in figure 6, ARABIC LETTER HALF ALEF in figure 7-8, ARABIC FILLED SQUARE BELOW and ARABIC SMALL LOW UPRIGHT RECTANGULAR ZERO in figure 9, also ARABIC SMALL HIGH SEEN INITIAL FORM in figure 10. Those characters are not proposed in this document. I still need more information about those characters and selecting which characters should be encoded. I recommend Unicode to allocate new Arabic Extended-D block in SMP to the roadmap for those unencoded characters because I think that Arabic Extended-C block is not enough.

5. Reference

- [1] Pedoman Pentashihan Mushaf Al-Qur'an, Lajnah Pentashihan Mushaf Al-Qur'an Badan Litbang dan Diklat Kementerian Agama Republik Indonesia.
- [2] Mushaf Madinah
المصحف الشريف بالمدينة المنورة
- [3] Mushaf Iran
مصحف جمهورية اسلامى ايران به شيوه نسخ عربي
- [4] Qur'an Kemenag v2.3 beta, Mushaf Standar Indonesia. Published by Ministry of Religion of the Republic of Indonesia c.q. Lajnah Pentashihan Mushaf Al-Qur'an.
- [5] Mushaf Ad-Duuri 'an Abi 'Amr. Published by www.islamweb.net
المصحف الشريف برواية الدوري عن أبي عمر
- [6] Mushaf Tunis Warsh Asbahani
المصحف المعلم برواية ورش عن الإمام نافع المدني من طريق الأصهباني مع إبرازهم قواعد الترتيل
- [7] Mushaf Tunis Qaloon
المصحف المعلم برواية قالون عن نافع المدني مع إبرازهم قواعد الترتيل والأداء
- [8] Mushaf Warsh al-Azraq. Published by www.bhr-q.com
المصحف المعلم برواية ورش لقراءة نافع طريق الأزرق من الشاطبية
- [9] Mushaf Tayseer Dory Basry. Published by www.alwa7y.com
مصحف التيسير برواية الدوري عن أبي عمرو من طريق الشاطبية
- [10] Mushaf al-Syarif Khalaf an Hamzah
المصحف الشريف برواية خلف عن حمزة بطريقة الشاطبية
- [11] Qurān Majīd. Published by Taj Company Limited.
قران مجيد تاج كيني لميئد
- [12] Mushaf Tripoli Qaloon
مصحف الجماهيرية برواية الإمام قالون والرسم العثماني على ما اختاره الحافظ أبو عمرو الداني
- [13] Aš-Šumūsu-n-Nayyirāt
الشموس التيرتات في جمع القراءات العشر المتواترات من طريقي الشاطبية والدرة
- [14] Al Quran Al Kareem (Noori Usmaan) Qaish Zaadah

**ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646³**

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from <http://std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.
See also <http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

A. Administrative

1. Title:	Proposal to Encode Three Quranic Arabic Characters
2. Requester's name:	<i>Rikza F. Sh.</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>Individual contribution</i>
4. Submission date:	<i>7 July 2022</i>
5. Requester's reference (if applicable):	<i>N/A</i>
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:	<i>N/A</i>
b. The proposal is for addition of character(s) to an existing block:	YES
Name of the existing block:	<i>Arabic Extended-C</i>
2. Number of characters in proposal:	3
3. Proposed category (select one from below - see section 2.2 of P&P document):	
A-Contemporary <input checked="" type="checkbox"/>	B.1-Specialized (small collection) <input type="checkbox"/>
B.2-Specialized (large collection) <input type="checkbox"/>	C-Major extinct <input type="checkbox"/>
D-Attested extinct <input type="checkbox"/>	E-Minor extinct <input type="checkbox"/>
F-Archaic Hieroglyphic or Ideographic <input type="checkbox"/>	G-Obscure or questionable usage symbols <input type="checkbox"/>
4. Is a repertoire including character names provided?	
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?	<i>Rikza F. Sh.</i>
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	<i>Rikza F. Sh. (rikzafsh@gmail.com)</i>
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 enclose information)?	YES
<i>See proposal</i>	
8. Additional Information:	
<p>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.</p>	

³ Form number: N4502-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?	YES
If YES explain	<i>The character "small low noon" has been proposed by Lateef Sagar Shaikh in L2/14-095 to be used with small kasra. But later, Unicode encoded them as an atomic character U+08D9 ARABIC LOW NOON WITH KASRA.</i>
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.)?	NO
If YES, with whom?	N/A
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:	<i>Moslems around the world</i>
4. The context of use for the proposed characters (type of use; common or rare)	Common
Reference:	<i>See proposal</i>
5. Are the proposed characters in current use by the user community?	YES
If YES, where? Reference:	<i>Moslems around the world</i>
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	NO
If YES, is a rationale provided?	N/A
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:	<i>See proposal</i>
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:	<i>See proposal</i>
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
If YES, reference:	<i>See proposal</i>
11. Does the proposal include use of combining characters and/or use of composite sequences?	YES
If YES, is a rationale for such use provided?	YES
If YES, reference:	<i>See proposal</i>
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)	N/A
	N/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