

Proposal to add Arabic letter for Swahili

Denis Moyogo Jacquerye <moyogo@gmail.com>
2024-02-11

This document requests the addition of an Arabic letter character (ARABIC LETTER YEH WITH FOUR DOTS BELOW) used in 20th century manuscripts of Muhamadi Kijuma in Swahili (ISO 639-3: swa). The letter was used for the palatal nasal, ⟨ny⟩ in Latin orthography.

Muhammad bin Abubakar bin Umar al-Bakriy (1885-1845), known as Muhamadi Kijuma (or Kijumwa), was a writer, poet, translator in the Swahili language. His manuscripts are collected in Kenya, Tanzania and in Europe. In 2022, professor of Literatures African Languages at University of Bayreuth, Clarissa Vierke, co-author of *Muhamadi Kijuma – Texts from the Dammann Papers and other Collections* 2010, said "[Muhamadi Kijuma] supplied many European researchers with manuscripts of Swahili poetry and introduced also them to the manuscript culture and Islamic poetry. He truly shaped our understanding of Swahili in Arabic script and most of the manuscripts stored in London, Hamburg, or Berlin nowadays are in his handwriting, including glossaries and explanations."

Character to be added

| Codepoint | Isolated | Final | Medial | Initial | Name |
|-----------|----------|-------|--------|---------|--|
| U+10EC7 | ﻲ̣ | ﻲ̣ | ﻲ̣ | ﻲ̣ | ARABIC LETTER YEH WITH FOUR DOTS BELOW |

Annotations and properties

NamesLists.txt

The following annotations are recommended for NamesList.txt:

10EC7 ARABIC LETTER YEH WITH FOUR DOTS BELOW
* Swahili manuscripts

UnicodeData.txt

10EC7;ARABIC LETTER YEH WITH FOUR DOTS BELOW;Lo;0;AL;;;;;N;;;;;

ArabicShaping.txt

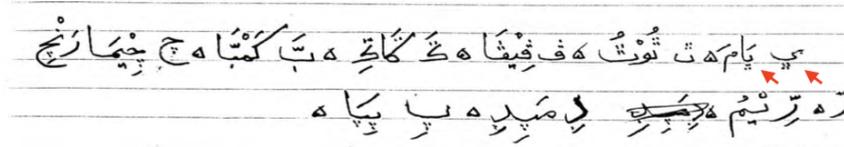
10EC7; DOTLESS YEH WITH 4 DOTS BELOW; D; YEH

Confusables

The ARABIC LETTER YEH WITH FOUR DOTS BELOW has forms confusable with forms of ﻲ U+0680 ARABIC LETTER BEHEH.

Examples

- ny* palatal /n/ *nyama* ‘meat’
- t* cerebral [retroflex] /t/ *toṭo* ‘small’ [= /ch/ Kisiu, marked by Dammann, but written in the texts only as /t/ without diacritic]
- v* dentilabial voiced /v/ *viva* ‘ripen’ (Kiunguja: *iva*)
- c* *chachache* ‘few’ (Kiunguja: *chache*)
- mb* *k^hamba* ‘crab’
- ranji* *ranji* ‘colour’. M.K. probably also pronounced the adjective *-ema* in class 5 ‘good’ as *jema*
- nd^ɾ* with cerebral pronunciation *nd^ɾimo* ‘it is inside’. In Lamu one most commonly hears /nd^ɾ/.
- ɖ* with interdental or postdental pronunciation. The only example given by M.K. is the nasal compound in *m(u)pendi* ‘lover’.
- p* *pipa* ‘barrel’



A note on Arabic script: SBB Hs. or. 9910

Figure 1. ARABIC LETTER YEH WITH FOUR DOTS BELOW and its transliteration ⟨ny⟩ in Mieke & Vierke 2010, p. 169 in “ny” and “nyama”.



بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

مَدَّوْخٌ لَكُتَّبٌ ۚ اَيْنَ لَمَلٌ وَهَبٌ ۚ حَدِيثٌ يَّعْتَوِبُ ۚ غَيْدٌ كَوْمِيَا ۚ
 لِبِسْمِ اللَّهِ اَوَّلٌ ۚ بِكَ اِسْمِي مِثَالِ ۚ بَانَ اَمْرُ تَوَالٍ ۚ وَفَلِمَ وَرَدْنِيَا ۚ
 بَانَ هَيَّ بِكُتَّوْ ۚ رِي وَكُ عَبْدِو ۚ لَالَهَ الْاَهُو ۚ بِبِ اَمْرٍ اِنِّيَا
 يَعْقُبُ ثُمَّ وَكِل ۚ نَمَكُو مِرْسَالٍ ۚ وَبُرْ حَقِيمٌ خَلِيلٍ ۚ سَاسٌ تَوْحِدِ شِيَا
 يَعْقُبُ كُورَانٍ ۚ يَلِنَ مَثِ يَفِيْنِ ۚ اَكْرَا فَهْمُوْنِ ۚ نُوْتَدُ هَتْكِيَا
 اَكْرَا هَتْكِيَا ۚ مَسُو يَكْتِمِيَا ۚ مَتَدُ كِمَ نَمِيَا ۚ حَسْبُ يُوْنِ بِبِيَا
 مَتَدُ كُورِيَا ۚ كُلُّهُتْ مِيَا مِيَا ۚ بِدَعَشْرُوْتِ بِبِيَا ۚ كُوْتِ وَكْتِمِيَا
 كُرُو يَسْفِ بِبِيَا ۚ نُوْتَدُ وَسْتِكِيَا ۚ حَتَّى نَتِ اِسْلِكِيَا ۚ كَتَّتِ اَكْتَحِيَا
 هُنَّ

Figure 2. ARABIC LETTER YEH WITH FOUR DOTS BELOW in *Qisṣati Yūsufu* (Hs. or. 9893), p. 1.

هَنْ يَوْسُفَ كَلَوُ ۚ الْوَوْنَ وِرْو ۚ وَنَسَمَبُ زَسَوُ ۚ مَيْنُ وِقْ كَغِيَا ۙ
 اَكْمَتْ وَاكْ بَابُ ۚ بَابُ مِيمِ سَمَبُ سِنْ ۚ نَبْلُغُ مَعْنُ ۚ سَمَنْ هَلْ نَهْمِيَا
 كَلَّ سِدْ كَمُوزِ ۚ كُورِيَا يِكْ كِرْ ۚ بَانَ هُنْلِي ۚ سَبَابُ رِغْ نَهْمِيَا
 كَجِبُ رِغْ مَنَعُ ۚ نَمُوتُ مَلْ وُخْ ۚ اَلِيَا رِغْنُ بِنَعُ ۚ رِي مَنِي كُنْتِيَا
 اَكُومَبُ كُورْمَبُ ۚ بَانَ وِوْمِنِي مَبُ ۚ مَمْلُ هَنْ سَمَبُ ۚ يَوْسُفُ نَمُومَبِيَا
 رِغْ اِكْتِدِكِيَا ۚ جَبْرِيَلُ اَمِكِيَا ۚ جَمِ سَمَبُ كَمْتِيَا ۚ بَيْنُ مِمَكْتِيَا
 مَكْتِيَا يَبِيوُنِ ۚ هَبِ سَمَبُ جَمْنُ ۚ كَيْكِيَا مَوْمَيْنِ ۚ مَكْتُكْرُ جَلِيَا
 جَبْرِيَلُ مَعْلَمُ ۚ اَكَاوُ سَلَامُ ۚ اِلَهَ اَكْسَلَمُ ۚ مَلْ وَاكْ مَنَمِيَا
 پَايَنْخِي يَعْقَبُ ۚ جَبْرِيَلُ مَكْجِبُ ۚ مَمْتَا كَا اَرْبُ ۚ نَمُ مَفْرَحِيَا
 اَكِيَا تَحِيَا ۚ كَا اَرْبُ مِيَا مِيَا ۚ جَبْرِيَلُ كَرَجِيَا ۚ كَبَانَ مَرَانِيَا
 نِيَعَقَبُ شَرِيْفُ ۚ اَتْمُ اِسْخَلِفُ ۚ سَمَبُ كَمِيَا يَوْسُفُ ۚ حَبُ اَلْمَزْدِيَا
 بَبَاكُ اَكْمِيْدُ ۚ كَطَا كَمْتُدُ ۚ زَمُ كُوزُ كُ يَدُ ۚ مَكْجِبُ كَمَغِيَا

Figure 3. ARABIC LETTER YEH WITH FOUR DOTS BELOW in *Qişşati Yūsufu* (Hs. or. 9893), p. 2.

قصّة قاضي

زَمَانٍ بَرَّعَرَبٍ بَلَيْنَ مَثْ أَيْنَ لَاحِ (حَطَّابِي) كَزِيَا كَهْرٍ
 مَثُونِ هِ كَزَغِيَا كُونِ اَلْكَومِ مَيْنِ اَلْفِيَا سِيَدِ يَنْحِ كِيَاثِ هِ
 مَيْسِ مَنِيغِ كَسْفِيَرِي اَلْجِيَا هِ اَكْنِيَرِي مَغْنِيَا هِ اَلْغِيَا كَثْرُجِ
 اَلْجِيَا كِيَاثِ يَمْبِ مَرْيَا نَدِيَمِي اَلْكَيْتِ مَرْيَا مَلِيْدِ سَانَ
 اَلْجِيَا يَمْبِ مَرْيَا هِ وَكِ كِيَانِ مَن مَكِ اَلْجِيَا حَطَّابِي كَمُوْرَا هِ
 حَطَّابِ اَلْكَرُوْرِ مَثُونِ اَلْجِيَا مَدِ وَاكِ هَمْدِ كِيَا مِي يَمُوْتِ كُوَا
 كَيْسِ كَمْبِ وِي كَعْلِ هِ سِيَدِ يَنْحِ مَن هِي هِ حَتْمِي سِيَدِ مِيَا
 قَاضِي وُجِ اَلْكَيْتِ اَلْكَوْنِ اَلْجِيَا مَلْعُونِ مِسْمَامِ قَاضِي مَلِيْدِ
 سَانَ هِ كَمِيَا تَغِي رَانَ نَتَكِ كَدَا لِيْزِ مَنِيْنِ هِ وَكَغِيَاثِ قَاضِي

قاضي كَمُوْرِيْزِ

Figure 4. ARABIC LETTER YEH WITH FOUR DOTS BELOW in *Qiṣṣat Qādhī* (Hs. or. 9893), p. 73. Note: the final form has a complete shape in the next example from p. 76.

حَتَّى كَايَ قَرِيبٍ إِلَيَّ مَثُ كَمِيَا حَطَابِي مَسِكِيَا مَدْرَاكُ أَمَمُوتُ
 كَمِيَا نَقِيلُ هَ إِلَيَّ مَثُ مَلِيَا نَشْطَانِ هَ كَمِيَا سِ شَطَانِ نَبِيَا مِيَا إِلَيَّ
 مَلِيَا أَحْسَنُ وَتُ كَمِيَا نَشْطَانِ وَوَهْنُ سِ شَطَانِ هَ بَاسِ مِمِ
 رِي شَطَانِ هَ سَاسِ مِمِ تَا أَفْنِي پُورَا كَيْسِ وَوَنَتُكُوا أَنْوَرُ هَ
 لَكِنِ سَمَلِرِ مَثُ إِلَّا قَاضِي حَطَابِي كُوتُ كَارِ نِيلُوتُ نَشْطَانِ مَكِسِ
 كَيْعُوزُ كَمِ پُرَا حَطَابِي كَزِيغِ نِيُورُ كَمُوزُ قَاضِي كَمُوتُ
 پُرَا مَلِيدُ كَمُنُو كَرِيَا لِمِيَا حَطَابِي كَيْكِيَا فَضْ رَاكِ كِنِرَا رَاكِ
 قَاضِي كَمَتُكُو پُرَا هَ كُلُّ سِيكُ قَاضِي هَمِيَرُ پُورَا هَ أَكِنِرَا شَمْبُ
 حَتَّى سِيكُ مَوِيَا قَاضِي كِصَالِ شَمْبُ نِيُورُ مَمُوكُ مَمِيَا يَاكِ
 اِي پُرِي كِيَا جَا مَمِيَا يِقَاضِي مَرَّ قَاضِي كَمُوتُ پُورَا أَمِكُو
 كَمِ كُنْدُرَا كِسِ مَلُوكُمِ كُوكُ هَ كِسِ مَلُوكُمِ كُوكُ لُوكُ هَ كِسِ مَلُوكُ
 كَمِ پِنْدِ كِسِ كَرُوكُ هَ كَغِيَا رَانِ يَمِيمُ يِرُوكُ هَ قَاضِي كِيَا جَا يُو

Figure 5. ARABIC LETTER YEH WITH FOUR DOTS BELOW in *Qiṣṣat Qādhi* (Hs. or. 9893), p. 76.



تَنْتَ يَخُ حَدِيثَ • نَيْغُونِ شَلُورِثَ • كَحُوْ بِمَبْعَثِ • مَعُ ثَتْ سَهْلِيَا •
 تَارِيحُ مَبُّ هِيَا • اَلْفُ ثَلَاثَ مِيَا • خَمْسِيْنَ تَنْ ثِيَا • بَانَ رَمَانَ اَكِيَا •
 تَكْفِيْ اَرْ فَيْعِ • كِنْيَا صَفِ يَحِكِ • كَا اَمْرٍ خَلَاكِ • نِيْرَ كُوْمِيَا •
 تَكُو كَنْدَ مَانَ • كَكَنْدِ كَنْلَنْتَ • رِي جِم كُوْنِيَانِ • نَسْتِ ثَكِ تَمِيْتِيَا •
 كَدَا لِكْرَ شِيْلَا • مَلْ كَتَّيْ سَهَا لَ • تَشْكَرِي كَكَلْ • نَسَبِ يَنْقَهْوَا •
 تَكْحِيْتِ كَحِشْمِ • سَلَا تَنْ يَمْلِيَا • تَكَلْ كَكَلْ كَحِمَ • نَكْسِ تَكْتَمِيْتِيَا •
 يَاطُ تَكْرَ مَيِّنِ • كَنْدَ مَسِكِيْتِ • نَانَ رِيْتِ هَا كَحِنِ • يَكْدَا كِيْنِيْجِيَا •
 تَكْرُخُ مَبُّ يِيَا • نَبِيْ وَثْ بِمِيَا • اَلِيْ تَشْكَرِيَا • هَبِلِ اَيْنِ سَكِيَا •
 تَكْشَكُ سَتِ يِيَا • يَاطُ تَكْشَكُ رِي • اَمُّ تَكْ سَكْلِيَا • كَبْرُكَ نَعْفِيَا •
 بُعْرِي

Figure 6. ARABIC LETTER YEH WITH FOUR DOTS BELOW in *Utendi wa Safāri* (SBB Hs. or. 9896) in “Tukafanya” and “kuonyana”. Note: the final form has a complete shape in the next example from the following page.

تَبَعْدِي وَيُرِي ۞ كَلِّ وَسِدْوَمِي ۞ كِرْكُرُوا كِرِي ۞ مَبِكْ نَصْ يَدِيَا
 تُكِنِرَ مَثْرَب ۞ تُكِبَاتْ مَنِرَب ۞ تُكَغِيَا شَمْبَان ۞ تُكُكْتِ سَتِيَا
 وَكِيَا وَمَثْرَب ۞ هَكْسَلِيَا مِين ۞ كُكُلْ تُكْتَمَان ۞ كَلْ تَلْ كِتْكُوا
 مَنِي شَمْبَ مَعْرِف ۞ كُكْتِ كُتْ ظِف ۞ كُرْتَالِ زَمْدَا ۞ سِبَسِ نَبِي يَمِيَا
 يَكُو هِبَكْتِيَا ۞ نَكُكُلْ هَكِلِدَه ۞ كُرْزِرِ مَخْلِك ۞ خِرَامْ كُرْجِيَا
 يَبْ تُكُنِكْ دِيَا ۞ اِلِ اِمْ كُرْجِيَا ۞ كِبْتِ نَصْ يَدِيَا ۞ كِي تُكَا اِنِيَا
 تُكِنِرَ كُتْ اَم ۞ نُؤِ يَبْ وَنَسِيم ۞ تُكُوغَنْ كَتَام ۞ كِسْ تُكَعْرَمِيَا
 تُكَا اِنِرْ تَكُوا ۞ كِتْمِ يَوْمَتَكُو ۞ تُكُنُولِ مَشُو ۞ تُكُوَا تَكْ سَكَلِيَا
 هَبِلْ نَهْرِيَا ۞ كُوَا وَكُنْدِيَا ۞ وَالِ كُشِيَا ۞ تُكِنِرَ كِتْمَبِيَا
 تُكُشَكْ مَفُون ۞ بَحْرُ كِبِلْ يِيَان ۞ تُكْتَمَبِيَا كُشَان ۞ فُرَاخْ زَكْتَمِيَا
 تُكَا اِنِ قَبْرِ ۞ يَشِيخْ مَنُصُور ۞ مَنْ نَخَطْ نَزْر ۞ عَجَبْ وَكُغَلِيَا
 يَانْ جُو يُولِك ۞ يَدِ شُو خَطِ يَاك ۞ هِيَا نَمْنِ يَاك ۞ مَسْمِيَا اَغَلِيَا
 نَشْرِيخْ يَدِ شُو ۞ تِنِ يَخَطِ بَجُو ۞ مَنَارَاوْ شُو ۞ اِلْ بَانْ وَوُ يِيَا

Figure 7. ARABIC LETTER YEH WITH FOUR DOTS BELOW in *Utendi wa Safāri* (SBB Hs. or. 9896) in “kinyezi” and “Mwenye”.

References by Muhamadi Kijuma

- *Qiṣṣati Yūsufu* (Hs. or. 9893). 1913. <http://resolver.staatsbibliothek-berlin.de/SBB00015E2700000000>
- *Qiṣṣat Qādhi* (Hs. or. 9893). 1937. <http://resolver.staatsbibliothek-berlin.de/SBB00015E2700000000>
- *Swahili Notes and Glosses (MS 47796a)*. 1935. <http://digital.soas.ac.uk/LSMD000192/00001>
- *Utendi wa Safāri* (SBB Hs. or. 9896). 1937.

Other references

- Mohammad Ibrahim Mohammad Abou Egl, *The life and works of Muhamadi Kijuma*, University of London, 1983. <https://eprints.soas.ac.uk/29683/>
- Afrikanists Assemble, *Africanist Hall of Fame | Afrikanists Assemble | Episode 15*, on YouTube, 2022-07-01. <https://www.youtube.com/watch?v=qIzcVJDJokU>
- Gudrun Mieke & Clarissa Vierke, *Muhamadi Kijuma – Texts from the Dammann Papers and other Collections*, Köln: Rüdiger Köppe verlag, 2010.

**ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646**

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.

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)

A. Administrative

| | | |
|--|--|--------------------------------|
| 1. Title: | Proposal to add Arabic letter for Swahili | |
| 2. Requester's name: | Denis Moyogo Jacquerye <moyogo@gmail.com> | |
| 3. Requester type (Member body/Liaison/Individual contribution): | | Individual contribution |
| 4. Submission date: | | 2024-02-11 |
| 5. Requester's reference (if applicable): | | |
| 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): | | |
| Proposed name of script: | | |
| b. The proposal is for addition of character(s) to an existing block: | | YES |
| Name of the existing block: | Arabic Extended-C | |
| 2. Number of characters in proposal: | | 1 |
| 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) |
| C-Major extinct | <input type="checkbox"/> D-Attested extinct | <input type="checkbox"/> E-Minor extinct |
| F-Archaic Hieroglyphic or Ideographic | <input type="checkbox"/> | <input type="checkbox"/> G-Obscure or questionable usage symbols |
| 4. Is a repertoire including character names provided? | | YES |
| a. If YES, are the names in accordance with the "character naming guidelines" | | 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? | | |
| | Denis Moyogo Jacquerye | |
| b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.): | | |
| | <moyogo@gmail.com> | |
| 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 issue | | |
| 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 (presentation, sorting) |
| 8. 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 UAX#44: 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? If YES explain | NO |
| 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.)? If YES, available relevant documents: | YES |
| 3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference: | YES |
| 4. The context of use for the proposed characters type of use; common or rare) Reference: | YES |
| 5. Are the proposed characters in current use by the user community? If YES, where? Reference: | YES |
| 6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If Yes, reference: | YES Yes See proposal |
| 7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? | YES |
| 8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If Yes, reference: | NO |
| 9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If Yes, reference: | NO |
| 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? If YES, is a rationale for its inclusion provided? If Yes, reference: | NO |
| 11. Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a rationale for such use provided? If Yes, reference: Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If Yes, reference: | NO |
| 12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary) | NO |
| 13. Does the proposal contain any Ideographic compatibility characters? If YES, are the equivalent corresponding unified ideographic characters identified? If Yes, reference: | NO |