

# Proposal to encode an Arabic tail character used for abbreviation

Roozbeh Pournader (WhatsApp) and Borna Izadpanah (University of Reading)  
February 21, 2020


## Background

One of the authors, Borna Izadpanah, is a researcher into the history of Arabic script printing and typography. The other author, Roozbeh Pournader, “discovered” the character proposed here in a presentation given by Borna Izadpanah at IStype 2019. Further research led to finding several more examples of the proposed character.

So far, this character has only been observed in several of the publications by the first printing press in Tehran, known as *Chap-e Mo'tamedi* or *Mo'tamedi Imprint*, established in 1820s. The metal types with which these publications were printed were most likely designed in Iran by Iranians. These types were based on the Persian flavor of the Naskh style. Encoding the character would help in the digitization efforts of early Persian movable type printing in Iran.

## Proposal

Encode the following character in the Unicode Standard:

Glyph	Codepoint	Name
	088E	ARABIC VERTICAL TAIL

The proposed character properties follow:

### UnicodeData.txt

```
088E; ARABIC VERTICAL TAIL;Lo;0;AL;;;;;N;;;;;
```

### ArabicShaping.txt

```
088E; VERTICAL TAIL; R; VERTICAL TAIL
```

All other properties should be similar to U+08AC ARABIC LETTER ROHINGYA YEH.

## Suggestion for NamesLists.txt

```
088E ARABIC VERTICAL TAIL
```

\* Early Persian movable type printing in Iran; only observed in final form

## Notes

1. The proposed character has only been observed in its final form. We don't know if it ever occurs in isolated form. In this way, it's similar to U+08AC ARABIC LETTER ROHINGYA YEH (see L2/10-288R). If a design for the not-yet-observed isolated form is desired, it could look like the final form with no right-side connection.
2. Architecturally, there are some similarities with U+FE73 ARABIC TAIL FRAGMENT (see L2/01-095). But that character came from legacy character sets that did not have automatic shaping and is only expected to occur next to other presentation forms. We believe our proposed character should instead be encoded as a normal shaping character, so that it can come after normal Arabic letters and affect their shapes.

## Samples of usage

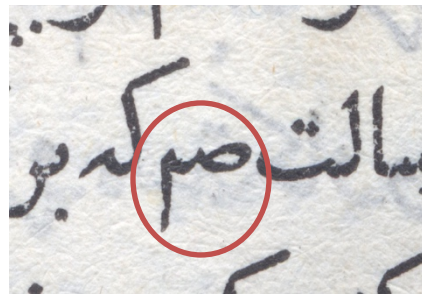


Figure 1. The proposed character occurring after *sad*, following a reference to the prophet Muhammad to represent the honorific *صلى الله عليه* or its variations. From Majlesi 1825, page 71.

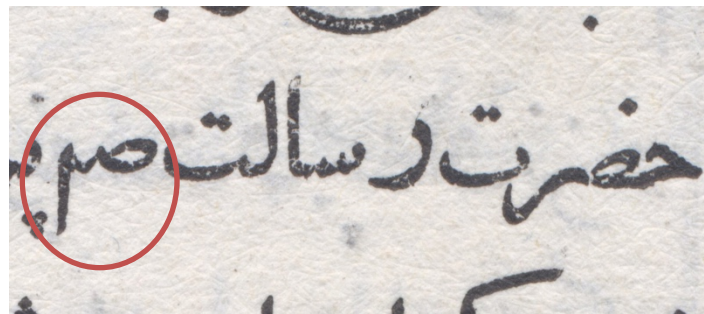


Figure 2. The proposed character occurring after *sad*, following a reference to the prophet Muhammad, to represent the honorific *صلى الله عليه* or its variations. From Majlesi 1825, page 71.

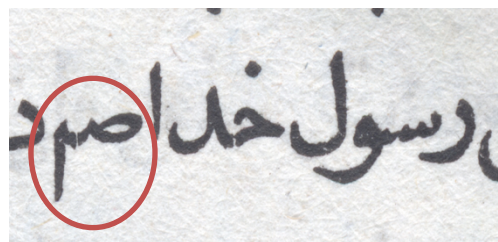


Figure 3. The proposed character occurring after *sad*, following a reference to the prophet Muhammad, to represent the honorific *صلى الله عليه* or its variations. From Majlesi 1825, page 56.

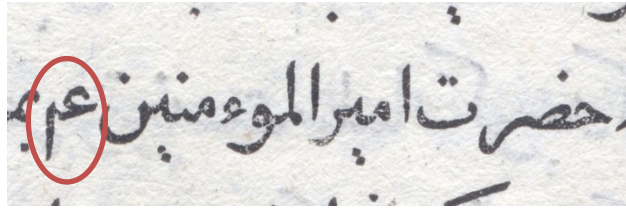


Figure 4. The proposed character occurring after *ain*, following a reference to Ali, to represent the honorific *عليه السلام*. From Majlesi 1825, page 71.

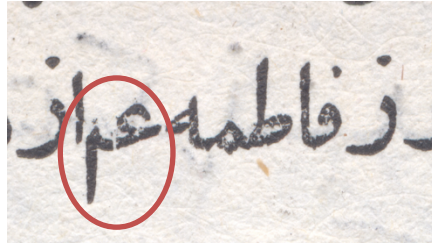


Figure 5. The proposed character occurring after *ain*, following the name of Fatimah, to represent the honorific *عليها السلام*. From Majlesi 1825, page 72. Note the gap between *ain* and the tail, showing that this was most probably a different piece of type.

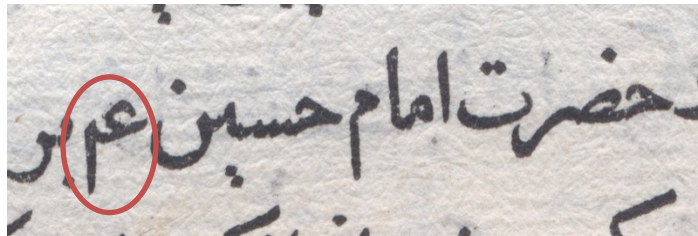


Figure 6. The proposed character occurring after *ain*, following the name of Husayn ibn Ali, to represent the honorific *عليه السلام*. From Majlesi 1825, page 53.

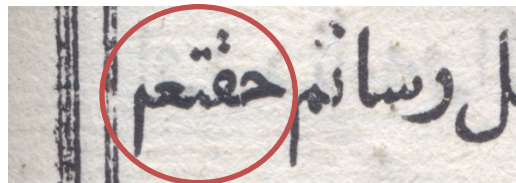


Figure 7. The proposed character occurring at the end of a word-like phrase, *hah-qaf-teh-ain*, as an abbreviation for the phrase *حق تعالی*. From Majlesi 1825, page 72.

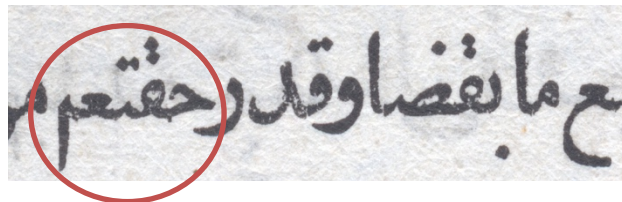
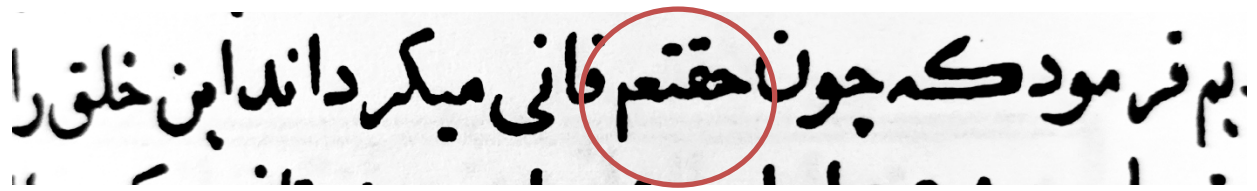
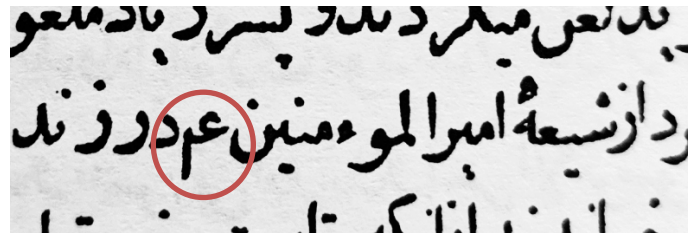


Figure 8. The proposed character occurring at the end of a word-like phrase, *hah-qaf-teh-ain*, as an abbreviation for the phrase *حق تعالی*. From Majlesi 1825, page 56.



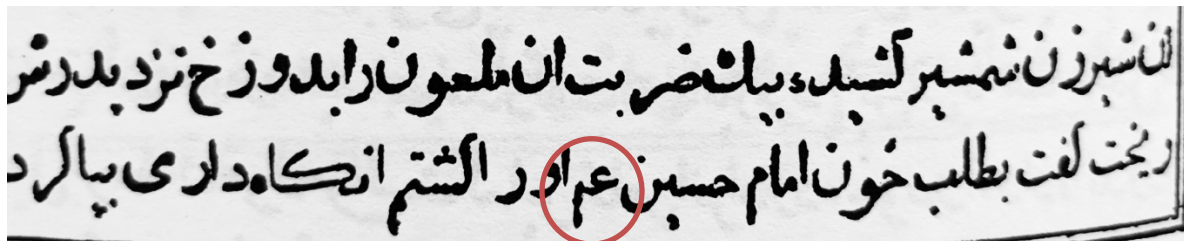
بم فرمود که چون حقتعم فانی میکرد اند این خلق را

Figure 9. The proposed character occurring at the end of a word-like phrase, *hah-qaf-teh-ain*, as an abbreviation for the phrase حق تعالی. From Majlesi 1843, last page. Reproduced in Babazadeh 1999, page 215.



بدل عن میکردند و پسرد باد معو  
رد از شیعۀ امیرالمؤمنین عم در زند

Figure 10. The proposed character occurring after *ain*, following a reference to Ali, to represent the honorific علیه السلام. From Va'ez Heravi 1845, page 37. Reproduced in Babazadeh 1999, page 242.



لنا شیرین شمشیر کشید و یک ضربت آن ملعون را بدوزخ نزد پدرش  
ریخت گفت بطلب خون امام حسین عم او را کشته از کاه داری بیا کرد

Figure 11. The proposed character occurring after *ain*, following the name of Husayn ibn Ali, to represent the honorific علیه السلام. From Va'ez Heravi 1845, page 168. Reproduced in Babazadeh 1999, page 243.

در نظر آید باب دانش مطبوع ابد از اجمله چون کربستن بر مصائب و  
 محمد مصطفی ص و سایر ائمه هدی علیه التحیه و الثناء ثواب بسیار  
 مخرق القلوب که در تعزیت داری و مصائب جناب شجر کابنار  
 سایر ائمه هدی علیه ما التحیه و الثناء تالیف شده مطبوع میشود امین  
 بیت رسول خدا از آن منتفع شده شاید بدین وسیله جناب اقدس  
 عریسات رابع و البدین رحمت کند و برحمت الله عبد اقل امین  
 یقین که مطبوع طماع شیعیان به سبب از آن خاندان  
 رسول ص خواهد آمد و معمول به ایشان خواهد  
 کرد بدین شهره شوال المکرم

Figure 12. The proposed character occurring twice after *sad*, following references to the prophet Muhammad, to represent the honorific *صلى الله عليه* or its variations. From Naraghi 1832, last page. Reproduced in Babazadeh 1999, page 231. This may be a ligated form, since the *sads* appear to have lost their teeth.

میباید پس فرموده سرچشمه  
 خواهد بود در زمین و خواهند بود در دنیا بمنزله پیغمبران و انما عم و ابان از انبیا و انما  
 بناحق و زنا کردن صادر شود با نیتانی که خدا هرگز زمین را از پیغمبری با امامی از فرزندان آدم خالی  
 نداشته است با نیتانی که خدا میفرماید که نفرستادیم پیش از تو یعنی بسوی خلق مگر مردانی چند  
 که وحی میفرستادیم بسوی ایشان از اهل شهرها پس این دلیلست بر آنکه ملئکه را بسوی زمین  
 نفرستاده است که پیشوایان و حکام باشند بلکه ایشان را بسوی پیغمبران خود فرستاده است پس راویان  
 عرض کردند که پس بنا بر این شیطان نیز میباید ملک نباشد فرمود که او نیز ملک نبود بلکه از جن بود چنانچه  
 حقه فرموده است که \* انه كان من الجن \* و باز فرموده است که و انجا خلقناه من قبل من نار السموم \*  
 و بدستیکه خبر داد مرا پدرم از جدم از حضرت امام رضا (ع) از پدرانش از رسول خدا صلی الله علیه  
 و آله که آنحضرت فرمود که حقه اختیار کرد از جمیع علیان محمد و آل محمد و اختیار کرد پیغمبران را و  
 اختیار کرد ملئکه مقربان را و اختیار نکرد ایشان را مگر برای آنکه میداند است که کاری نخواهند کرد که  
 از ولایت و دوستی خدا بیرون روند و از عصمت الهی ببری شوند و ختم شوند با کسرویی که مستحق  
 عذاب خدا گردیده اند پس راویان گفتند که بیمار و ابت رسیده است که چون حضرت رسول صلی الله علیه  
 و آله فرمود بر حضرت امیرالمؤمنین (ع) امامت و عرض کرد خداوند عالمیان و ولایت آنحضرت را بر ملئکه پس  
 کرده بسیاری قبول ولایت آنحضرت نکردند و خدا ایشان را مسخ کرد بصورت و ذغابی فرمود  
 معاذ الله این حدیث را بر ملا روغ بسته اند و ملئکه رسولان خدا بند و چنانچه بر پیغمبران  
 خدا انفرادیست بر ایشان نیز و انست و شان ملئکه عظیم است و مرتبه ایشان جلیل  
 است و از امثال این امور متوجهند و باین جا منتهی شد آنچه از تفسیر امام عم (ع) نقل  
 کردیم و ساد احوال ملئکه و بیان عصمت ایشان از کاتب روح الارواح

Figure 13. The proposed character occurring eight times on a single page, after *sad* and *ain*, as well as the word-like phrase, *hah-qaf-teh-ain*. From Majlesi 1845, last page. Reproduced in Babazadeh 1999, page 246.

**Bibliography**

1. Shahlā Babazadeh. 1999. *The History of Printing in Iran (Taarih-e Chaap dar Iraan)*. Tehran: Tahoori. ISBN 964-6414-08-7.
2. IBM Egypt and the Unicode Consortium. 2001. "Proposal to add "Arabic Tail Fragment" character." UTC Document Register L2/01-095, The Unicode Consortium. <https://www.unicode.org/L2/L2001/01095-N2322-ArabTail.pdf>

3. Borna Izadpanah. 2019. "The emergence of printing in Qajar Iran." Presented June 15, 2019 at Istanbul Typography Conference (ISType). Video posted to YouTube on Sep 8, 2019. <https://youtu.be/n67bRO3PYBs>
4. Mohammad-Bagher Majlesi. 1825 (1240 AH). *Jalaa' al-'Ayun*. Tehran: Mo'tamedi Imprint. From the personal collection of Borna Izadpanah.
5. Mohammad-Bagher Majlesi. 1843 (1259 AH). *Haqq ol-Yaqin*. Tehran: Zeynolabedin Tabrizi. Reproduced in Babazadeh 1999.
6. Mohammad-Bagher Majlesi. 1845 (1261 AH). *Hayaat ol-Qolub*. Tehran: Mohammad Esmail, Mo'tamedi Imprint. Reproduced in Babazadeh 1999.
7. Mehdi Naraghi. 1832 (1247 AH). *Moharreq al-Qolub*. Tehran: Zeynolabedin Tabrizi of Mo'tamedi Imprint. Reproduced in Babazadeh 1999.
8. Lorna A. Priest and Martin Hosken. 2010. "Proposal to add Arabic script characters for African and Asian languages." UTC Document Register L2/10-288R, The Unicode Consortium. <https://unicode.org/L2/L2010/10288r-arabic-proposal.pdf>
9. Ata'ollah Va'ez Heravi. 1845 (1261 AH). *Mokhtaarnaameh = Rowzat ol-Mojaahedin*. Tehran: Abdolkarim. Reproduced in Babazadeh 1999.

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 an Arabic tail character used for abbreviation***
2. Requester's name: **Roozbeh Pournader and Borna Izadpanah**
3. Requester Type: **Expert Contribution**
4. Submission date: **February 21, 2020**
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 Extended-B**
2. Number of characters in proposal: **1**
3. Proposed category: **D-Attested Extinct**
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? **Borna Izadpanah**
  - b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.): **Borna Izadpanah**
6. References:
  - a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? **No**
  - 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 second author is part of the user community, an expert researcher studying the history of the Arabic script printing and typography**  
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:
4. The context of use for the proposed characters (type of use; common or rare): **Rare**  
Reference:
5. Are the proposed characters in current use by the user community? **No**  
If YES, where? **N/A**  
Reference:
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. Needs to be next to similar characters.**  
If YES, reference:
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? **N/A**
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? **No**  
If YES, is a rationale for its inclusion provided? **N/A**  
If YES, reference: **N/A**
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? **No**  
If YES, is a rationale for its inclusion provided? **N/A**  
If YES, reference: **N/A**
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. See proposal.**  
If YES, reference: **See proposal.**
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): **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**