Proposal to Encode Arabic Letter Teh with Small Teh Above for Bravanese (revision 1)

To: *UTC*
Date: *1 November 2014*
From: *Roozbeh Pournader (Google Inc.), Shervin Afshar (HighTech Passport, Ltd.)*
Live Doc: http://goo.gl/R3wtef

Background

Previously in 2013 a proposal by Banafunzi et al. (L2/13-178)[1] requested five Arabic script character to be encoded for Bravanese (Chimini). During UTC meeting #137 in November 2013, encoding of these five characters were discussed and four of them were accepted for encoding[2]:

```
08B6;ARABIC LETTER BEH WITH SMALL MEEM ABOVE
08B7;ARABIC LETTER PEH WITH SMALL MEEM ABOVE
08B9;ARABIC LETTER REH WITH SMALL NOON ABOVE
08BA;ARABIC LETTER YEH WITH TWO DOTS BELOW AND SMALL NOON ABOVE
```

Since information provided for a fifth character (ARABIC LETTER TEH WITH CURL) was not adequate for encoding purposes, the UTC decided to postpone encoding the character until receiving the needed information.

Objective

The objective of this document is to provide the missing information and propose this additional character to be incorporated into the Unicode Standard.

Discussion

The character originally proposed by Banafunzi et al. as ARABIC LETTER TEH WITH CURL appears as an Arabic letter ت with an added curl added by hand under the dots of typesetted letter:

---

[1] Banafunzi, Banafunzi, and Nuur. (L2/13-178)
[2] Pournader, "Proposal to encode four Arabic characters for Bravanese".

Research by Banafunzi and Vianello[3] notes this character among adaptations to Arabic letters to express sounds specific to Chimini language and describes it as "ت with another small ت above it".

In the introduction to *The Arabic Script in Africa*[4] (in which the work of Banafunzi and Vianello appears), Mumin and Versteegh name the character " *tā'* with superimposed *tā'* " and provide the following shape for it:

Considering these information, the authors propose this character to be encoded with properties as provided in the following sections.

Properties

UnicodeData.txt

```
08B8;ARABIC LETTER TEH WITH SMALL TEH ABOVE;Lo;0;AL;;;;;N;;;;;
```

ArabicShaping.txt

```
08B8; DOTLESS BEH WITH TEH ABOVE; D; BEH
```

---

[3] Banafunzi and Vianello, pp. 302, 309.
[4] Mumin and Versteegh, p. 15.

Acknowledgement

Bibliography

- Banafunzi, Bana, and Vianello, Alessandra. "Chimi:ni in Arabic Script: Examples from Brava Poetry". In Mumin and Versteegh (ed.), p. 293.
- Banafunzi, Hamid, and Banafunzi, Marghani, and Nuur, Maxamed. "Proposal to encode five Arabic script characters for the Bravanese (Chimiini)." UTC Document Register L2/13-178. 31 August 2013. http://www.unicode.org/L2/L2013/13178-bravanese.pdf.
- Pournader, Roozbeh. "Proposal to encode four Arabic characters for Bravanese". UTC Document Register L2/13-223R. 6 November 2013. http://www.unicode.org/L2/L2013/13223r-arabic-bravanese.pdf
- Mumin, Meikal, and Versteegh, Kees (ed.). *The Arabic Script in Africa*. Leiden: Koninklijke Brill NV, 2014. ISBN 978-90-04-25679-8.

ISO/IEC JTC1/SC2/WG2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[5]

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 Arabic Letter Teh with Small Teh Above for Bravanese
2. Requester's name: Roozbeh Pournader, Shervin Afshar
3. Requester type (Member body/Liaison/Individual contribution):   Individual contribution
4. Submission date: 30 Ocotber 2014
5. Requester's reference (if applicable):
6. Choose one of the following:

     This is a complete proposal: ■

     (or) More information will be provided later: ☐
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: ■
     Name of the existing block: Arabic Extended-A
2. Number of characters in proposal: One
3. Proposed category (select one from below - see section 2.2 of P&P document):

■A-Contemporary   ☐B.1-Specialized (small collection)   ☐B.2-Specialized (large collection)

☐C-Major extinct   ☐D-Attested extinct   ☐E-Minor extinct

☐F-Archaic Hieroglyphic or Ideographic   ☐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?   Michael Everson
     b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):   Michael Everson, everson@evertype.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)?

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?   Yes
   If YES explain   A proposal by Banafunzi et al. (L2/13-178) requested five Arabic script character to be encoded for Bravanese. Information provided for a fifth character (ARABIC LETTER TEH WITH CURL) was not adequate for encoding purposes, the UTC decided to postpone encoding the character until receiving the needed information.

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, available relevant documents:

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?   Yes
   Reference: See Banafunzi et al. (L2/13-178) and Pournader (L2/13-223R)

4. The context of use for the proposed characters type of use; common or rare)
   Reference: See Banafunzi et al. (L2/13-178)

5. Are the proposed characters in current use by the user community?   Yes
   If YES, where? Reference: See Banafunzi et al. (L2/13-178)

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?   Missing character from previously encoded set
   If Yes, reference: See Pournader (L2/13-223R)

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?
   If Yes, reference:

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?
   If Yes, reference:

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? No
    If YES, is a rationale for its inclusion provided?
    If Yes, reference:

11. Does the proposal include use of combining characters and/or use of composite sequences?
    If YES, is a rationale for such use provided?   No
    If Yes, reference:
    Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? No
    If Yes, reference:

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)

13. Does the proposal contain any Ideographic compatibility characters?   No
    If YES, are the equivalent corresponding unified ideographic characters identified?
    If Yes, reference: