TO: UTC   L2/13-165
FROM: Deborah Anderson, Ken Whistler, and Roozbeh Pournader
SUBJECT: Recommendations to UTC on Script Proposals
DATE: 25 July 2013

As in the past, we reviewed the following script/character encoding proposals and provide recommendations.

Latin
1.   **L2/13-129** **Proposal for the addition of Latin characters for Marshallese – Everson**

Discussion: We reviewed this proposal, which proposes two options to handle Marshallese, Latvian and Livonian cedillas/commas below:

(a) the four characters proposed and approved by WG2 in Vilnius, Lithuania (LATIN CAPITAL and SMALL LETTER MARSHALLESE L and LATIN CAPITAL and SMALL MARSHALLESE N WITH CEDILLA)

(b) twelve characters, which include upper- and lowercase Latin D, G, K and R "WITH INVARIANT CEDILLA," in addition to Latin upper- and lowercase Latin L and N mentioned in (a), but with a name change ("LATIN LETTER [x] WITH INVARIANT CEDILLA"). All characters have no decompositions.

Recommendation: The proposal involves an architectural issue, and hence should go to the UTC for discussion and a decision.

Arabic
2.   **L2/13-130** **Proposal to encode three Arabic characters for Arwi – Pournader**

Discussion: We reviewed this proposal, which requests three Arabic letters for Arwi (Arabo-Tamil). The characters were earlier proposed by Seyed Mohamed Buhari in 2009 (**L2/09-143**) (with documented evidence), but the original proposal did not include the discrete glyphs with their contextual forms. The proposed code points are acceptable.

The name for the combining character ARABIC INVERTED DAMMA BELOW corresponds to 0657 ARABIC INVERTED DAMMA, which likewise has a "6" shape. However, the shape is properly "turned," that is, the DAMMA's shape is turned 180 degrees:

064F ARABIC DAMMA

0657 ARABIC INVERTED DAMMA

(proposed) ARABIC INVERTED DAMMA BELOW

Recommendation: We recommend the UTC approve the characters, but discuss the name for ARABIC INVERTED DAMMA BELOW.

Cyrillic
3. [L2/13-153](#) **Proposal to Use Standardized Variation Sequences to Encode Church Slavonic Glyph Variants in Unicode – Andreev, et al.**

Discussion: We need to study the proposal further, but note comments sent in by Ralph Cleminson and David Birnbaum (L2/13-164).

Recommendation: We will relay our recommendation to the UTC orally.

4. [L2/13-139](#) **Proposal to Encode Combining Half Marks Used for Cyrillic Supralineation – Andreev, et al.**

Discussion: We reviewed this proposal for two characters, U+FE2E COMBINING CYRILLIC TITLO LEFT HALF and U+FE2F COMBINING CYRILLIC TITLO RIGHT, and note additional feedback from Ralph Cleminson and David Birnbaum (L2/13-164).

Recommendation: We recommend the UTC approve these two characters, after reviewing the comments from two other experts.

Supplemental Punctuation
5. [L2/13-140](#) **Proposal to Encode a Slavonic Punctuation Mark – Andreev, et al**

Discussion: We reviewed this proposal for one character, 2E43 SLAVONIC SPEAR, and feedback from experts, confirming its existence (L2/13-164).

As noted by Ralph Cleminson, the name SLAVONIC SPEAR might need to be re-considered, as the character is proposed in the Supplemental Punctuation block and might be used by different scripts outside of Cyrillic and Glagolitic. The line-breaking property should probably be BA, based on characters that function in the same way in other scripts, with the script property "Common." The UTC should discuss the Script Extensions property for this character.

Recommendation: We recommend this punctuation character be approved, but with a discussion on its name and properties.

Book Pahlavi
6. [L2/13-141](#) **Preliminary Proposal to Encode the Book Pahlavi Script – Pournader**

Discussion: We reviewed this preliminary proposal, which builds off of [L2/07-234](#) Preliminary proposal to encode the Book Pahlavi script.

Recommendation: We recommend the UTC review this proposal, discuss it, and send the author feedback.

South Asian: Gujarati

[L2/13-143](#) **Proposal to encode Gujarati Letter ZHA – Rajan**

Discussion: We reviewed this proposal, which is a revision of a proposal discussed at UTC meeting #135 in May 2013 ([L2/13-066](#) Proposal to encode Gujarati Sign Triple Nukta). The author has revised the proposal in accordance with the UTC discussion.

Recommendation: We recommend the UTC accept the proposed character, 0AF9 GUJARATI LETTER ZHA.

Siddham

7. [L2/13-130](#) **Proposal to Encode Variants for Siddham Variants – Kawabata et al.**
    [L2/13-136](#) **Additional Siddham Variants – Pandey**

Discussion: We reviewed the proposal by Kawabata et al. for six characters (approved by WG2), and the response from Pandey.

Recommendation: We recommend the UTC discuss the proposed six characters in L2/13-130, taking into consideration Pandey's points in L2/13-136, and see whether the committee can come to a decision on how best to proceed.

8. [L2/13-135](#) **Proposal to Encode a Set of Digits for Siddham   – Pandey L2/13-135**

Discussion: We reviewed this proposal, which asks for the encoding of 10 digits designed by contemporary users. (The original Siddham proposal, L2/12-234R, did not propose these modern innovations, until additional research could be done.)

The proposal has one image, but no additional evidence. It is not clear whether the characters are used for entertainment or other purposes.

Recommendation: We recommend the UTC ask the author to provide additional evidence, then re-submit the proposal.

Bhaiksuki

9. [L2/13-167](#) **Proposal to Encode the Bhaiksuki Script in ISO/IEC 10646 – Pandey and Dimitrov**

Discussion: We reviewed this proposal, which shows significant work and careful research beyond the preliminary proposal (L2/11-259).

One feature of the script, discussed in §4.6, involves sequences of virama and certain consonants, which can be represented by (a) a consonant and visible virama or (b) a special ligature. Rather than encode a new "KHANDA MARK" to represent the forms with a special ligature, we recommend joiners be considered, particularly if the special forms are just for presentation. However, input from those members with Indic implementations should be sought out, to verify if joiners would work in this case (and/or if there would be technical problems in implementing a second virama-type character).

The properties of a few characters need correction: the AVAGRAHA has "Mn;9;…", but this should be changed to "…Mc;0;…". If there is adequate justification for the KHANDA MARK, its properties should be changed from "Mc;0…" to "Mc;9…", since it is a virama-like character.

Recommendation: We recommend the UTC review the proposal, in particular discussing the KHANDA MARK, and request the author make changes to the properties, as mentioned above.