

DATE: 19 September 2014

DOC TYPE: Working Group Document

TITLE: Response to PDAM2 ballot comments from Great Britain on Chinese Characters needed for Slavonic

SOURCE: Deborah Anderson, Script Encoding Initiative, UC Berkeley

STATUS: Liaison contribution

ACTION: For consideration by WG2

DISTRIBUTION: ISO/IEC JTC1/SC2/WG2

In WG2 N4614 Summary of Voting/Table of Replies for the PDAM2 ballot, Great Britain expressed concerns about the 18 characters located at U+9FCD...U+9FE9 that it felt needed to be addressed before the characters could be accepted for encoding. This document addresses those concerns.

1. *[Ballot comments] The characters at 9FCD through 9FE9 have not been reviewed by WG2, and were not added to the repertoire of Amendment 2 by WG2 resolution, but were added at the discretion of the project editor. We approve of the inclusion of 9FCD through 9FD7 in this amendment, but have concerns about the appropriateness of encoding the eighteen characters at 9FD8 through 9FE9.*

The 18 characters from U+9FCD...U+9FE9 were discussed at both IRG #41 and IRG #42, as well as WG2 #62 in San Jose in February 2014.

At the WG2 meeting, the (unconfirmed) minutes from the meeting (SC2 N4359) record the following:

[Dr. Lu Qin] 6.3 – Slavonic transcriptions - the requested characters look legitimate. But their use is different from the current CJK characters. Some of these are already in CJK - ideographic than being just phonetic. One of the concerns is how to sequence these phonetically – since they are used in transcriptions. Some being already encoded in CJK, it would be better to order them with Radicals - not useful from transcription point of view. Submitters have been given the feedback.

At the IRG meetings, the peculiarities of these 18 characters were discussed, and the only objection from the IRG was about the phonetic ordering that was originally requested (as noted in the WG2 minutes, above).

However, the phonetic ordering was subsequently dropped in the revised proposal, WG2 N4583. In May 2014, IRG #42 passed the following resolution (IRG M42.9):

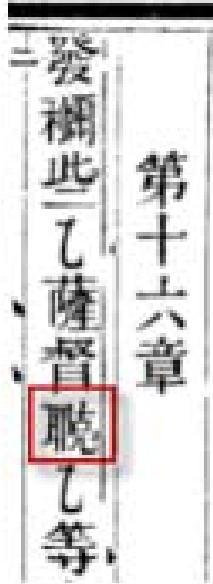
IRG resolves to accept the proposals for 20 UNCs used for transcribing Slavonic (IRGN1954 and IRGN1954R) and the UTC proposal for 5 UNCs (IRGN2005) and instructs the IRG Rapporteur to send the proposal documents to WG2 before WG2 Meeting No 63 for processing.

Hence the characters have had critical review by the IRG (and WG2 meeting members were apprised of the discussion at the February meeting).

2. *[Ballot comments] In the case of 9FD8 through 9FE9, it is quite possible that these are also pairs of fanqie characters squeezed into a single character space rather than individual characters, in which case displaying the pairs of fanqie characters in a single character space is a presentation issue, which can easily be solved by word processors or in html/css.We are concerned that if the eighteen characters at 9FD8 through 9FE9 are encoded, it will set a precedent for encoding an open-ended set of hundreds of fanqie character pairs used for phonetic notation of other languages. We believe that the wider issue of encoding such phonetic character pairs should be addressed before accepting 9FD8 through 9FE9 for encoding.*

While the characters located at U+9FCD...U+9FE9 may appear to be *fanqie* characters, they are, in fact, not *fanqie*.

Compare Figure 13 of N4583, from the translation of Matthew 16.1:



The English text of that verse is, roughly, “The Pharisees and Sadducees came, and to test Jesus they asked him to show them a sign from heaven.”

Compare the OCS text of the same version (from http://www.biblical-data.org/OCS/Slavic_NT.pdf):

1 ѿ прѣстѣпѣша (къ немѹ) фарїсеѣ ѿ саддукѣѣ, ѿскѣшающе
просѣша єго знаменїѣ съ нѣсѣ показати ѿмъ.

The reading of the Chinese in Figure 13 shows two runs of four transliteration syllables each, followed by the pluralizer, namely:

fa1-ri-xie1-yi3 sa4-du1-ke-yi3 deng3, ...

In other words, “The Pharisees and Sadducees...”

Very clearly that “ke” is not an inline *fanqie* annotating a preceding character. It is nothing other than another transliteration syllable, used systematically just like the other transliteration syllables in these words, and in fact systematically through this translation of the text. (Including the “ri” example also shown here.)

While these are bizarre creations, and certainly do not follow the canon for the normal types of radical + phonetic extensions used over the centuries to add new Chinese characters, claiming that in this context these are not characters is bogus. They are systematically distinct from true *fanqie* usage.

The question boils down to whether these additions are or are not a fixed set of coinages of single characters to fill out a particular transliteration set used in these OCS materials. The fact that the same forms are used over and over again for the same phonetic representations also suggests their characterhood, rather than being nonce *fanqie* forms made up on the spot.