

DATE: 18 September 2017
DOC TYPE: Working Group Document
TITLE: Comments on WG2 #66 (Sept. 2017) documents
SOURCE: Deborah Anderson, Ken Whistler, Roozbeh Pournader, and Lisa Moore
STATUS: Expert contribution
ACTION: For consideration by WG2
DISTRIBUTION: ISO/IEC JTC1/SC2/WG2

The following are comments by a group of experts on WG2 #66 documents.

1. Latin

a. Document: **N4842 Proposal to add two Sinological Latin letters (L2/17-299)** -- Michael Everson and Andrew West

In the view of the experts, the evidence is strong for the two proposed characters in the Latin Extended-E block, but the name of U+AB67 LATIN SMALL LETTER TS DIGRAPH WITH HOOK should be modified to LATIN SMALL LETTER TS DIGRAPH WITH *RETROFLEX* HOOK.

This change will make it consistent with U+AB66 LATIN SMALL LETTER DZ DIGRAPH WITH RETROFLEX HOOK as well as the set from U+1D8F LATIN SMALL LETTER A WITH RETROFLEX HOOK to U+1D9A LATIN SMALL LETTER EZH WITH RETROFLEX HOOK.

b. Document: **N4841 Proposal to add five Latin Tironian letters (L2/17-300)** -- Michael Everson and Andrew West

This proposal requested U+ 204A TIRONIAN SIGN ET, a character in the General Punctuation block, have its property changed from “Punctuation other” to “Letter lower” in order to add a corresponding uppercase LATIN CAPITAL LETTER TIRONIAN ET.

The proposal also requested four additional characters be added in the Latin Extended-D block: upper- and lowercase LETTER TIRONIAN ET WITH HOOK and upper- and lowercase LETTER TIRONIAN ET WITH HOOK AND STROKE.

The experts had the following comments:

- Because the TIRONIAN SIGN ET character is not used in modern texts, and no evidence was given to require a case pairing (such as need for automatic capitalization), no change should be made to the property of this character, in our opinion.

- We suggest TIRONIAN SIGN CAPITAL ET be located in the Supplemental Punctuation block at U+2E4F, with properties as given for U+204A TIRONIAN SIGN ET (Po;0;ON;;;;;N;;;;;), instead of in the Latin Extended-D block.
- The case pairs LATIN CAPITAL LETTER TIRONIAN ET WITH HOOK and LATIN LETTER TIRONIAN ET WITH HOOK AND STROKE appear to be variants. More evidence should be provided to demonstrate the need for these in plaintext.

2. Tangut

a. Document: N [4850](#) Glyph Corrections for 31 Tangut ideographs and one Tangut component ([L2/17-313](#)) -- Andrew West and Viacheslav Zaytsev

In §4 Tangut Source (page 25), the experts had the following comments:

- In the first entry, correct the second line from “U+17D36” to “U+17378”
- Remove “MergedSrc” data, and show instead both the current and suggested new values for kRSTUnicode

We also recommend experts review the proposed changes.

b. Document: N [4851](#) Proposal to encode six additional Tangut ideographs ([L2/17-314](#)) -- Andrew West and Viacheslav Zaytsev

In §4 Source References on page 22, we noted the proposal cites itself as a source. We recommend the authors publish their analysis, such as in a Unicode Technical Note, and use the UTN as a source in TangutSrc.txt (and TangutSources.txt for Unicode). Such an article need not include all the elements of a proposal, but an analysis of the 6 Tangut ideographs, with discussion of the shapes and the characters’ meanings.

3. SW Chinese Minority Hieroglyphs

Document: N [4856](#) Preliminary Proposal on Encoding Southwest China Minority Hieroglyphs, Part 1 Muya, Part 2 Namuz

This preliminary proposal provides a catalog of instances of the glyphs for two different systems, “Muya” and “Namuz”. In our view, instead of using SC2 and WG2 as the publisher of such catalogs, we recommend the authors focus on analyzing and identifying the units, and then, after publication of their work and peer review, preliminary proposals could be put forward, based on such analysis.

Other comments:

- The names as suggested are not usable, but should refer to the analysis. (It was also noted that multiple citations in the list have the same name, and are clearly not distinct enough to be separately encoded as characters. For example, 111 to 117 all have the name MUYA HIEROGLYPH TOR MA FOR MALE MISFORTUNE.)

- Once the script has been shown to be eligible for encoding, the glyphs should be re-thought: at present some glyphs appear to be rock rubbings, which is not appropriate for standardization in a font. This is true for both the Muya and Namuz sets (for the latter, cf. 37-47). In the Namuz section, color seems to be distinguishing different “characters”, such as 64 and 65, both named “MI35”, but color is not shown in the code charts.

4. Iteration/Punctuation marks

Document: N4847 Proposal to encode two marks for ancient Chinese texts (L2/17-310) -- Andrew West, Eiso Chan

This proposal requested two characters be added to the Ideographic Symbols and Punctuation block, U+16FE2 OLD CHINESE HOOK MARK and U+16FE3 OLD CHINESE ITERATION MARK. The evidence presented is, in our view, strong, but it is unclear whether the representative glyph for the HOOK MARK is the best choice. We suggest the authors provide further documentation on the representative glyph for this character.

5. Musical notation

a. Document: N4848 Proposal to encode old Chinese lute notation (L2/17-311) -- Andrew West and Eiso Chan

The following comments were raised during discussion:

- The proposal needs evidence the community of experts agrees on how to represent this material.
- The research as presented should be peer reviewed
- It is not clear the collection of combining marks (and for flute notation, subjoining marks) is necessarily the right way to handle encoding.
- Chording needs a scoping mechanism, but it is not clear it is plaintext.

b. Document: N4849 Proposal to encode old Chinese flute notation (L2/17-312) -- Andrew West and Eiso Chan

Although the details vary from Old Chinese lute notation, many of the same comments apply. (See comments above for N4848.)