ISO/IEC JTC1/SC2/WG2 N3539

Date: 13-10-2008

Title: Response from Tangut scholars of China on the Tangut Unicode proposal

Source: China

Action: To be considered by experts of WG2

A. On character unification

The following are opinions from China on the unification of Tangut characters. These opinions are all based on academic consideration.

- 1. A certain degree of character standardization is necessary, because:
 - a. In Li Fanwen's Tangut Dictionary (1997), there are some very obvious reduplications of single character, e.g. 0042 and 4537. They don't have difference in any sense.
 - b. It is unpractical to include all forms of characters that are actually used in Tangut texts. For example, Fig. 1 and 2 are both the radical for "human", which composes the largest set of Tangut character among all radicals. Fig. 1 normally occurs in earlier manuscripts, while Fig. 2 occurs in later manuscripts, with a stroke slightly different from Fig. 1 (in red circle). However, all contemporary Tangut dictionaries treat them to be the same, and in fact most Tangutologists are not aware of this difference. There are a lot of similar cases in Tangut manuscripts. If all these "allographs" which don't really distinguish characters have to be represented descriptively, then the table will finally be expanded unlimitedly (esp. when many Tangut manuscripts are hand-written), whether put it into variation sectors or not.





Fig.1 Fig.2

- 2. After careful examination, we agree that there are some mistakes in Han Xiaomang's unification of Tangut characters. For example, 0541 and 0542 are different in all Tangut sources. It seems to us that, while Han Xiaomang corrected most mistakes in Li's dictionary (1997), he made a few new mistakes at the same time.
- 3. We still acknowledge Han's work to be the most in depth research on Tangut characters. Yet we are open to further academic debate on the standardization of Tangut characters.

B. On Tangut Unicode proposals

The following are our opinions on Tangut Unicode proposals after discussion on recent documents:

- 1. The Tangut Unicode proposal should adopt the font developed by Jing Yongshi. (To put it into a new "G" column is fine).
- The proposal should include all characters (esp the "new characters") as listed in Jing Yongshi's work.

- 3. The proposal should include all radicals as listed in Jing Yongshi's work.
- 4. If there are disagreements on whether a character (or a variation of character, or a radical) should be included or not, it will be better to include it rather then to omit it.
- 5. We are fine with all technical solutions which can address all the concerns above.