

IRG N1014 Draft Agreement on Old Hanzi Encoding

Source: LI Guoying <bjlgy@hotmail.com> and Tom BISHOP <tbishop@wenlin.com>
Status: Expert Contribution
Date: 2003-11-19
Action: Consideration by IRG

On November 17, 2003, in room 353 of Guilin Park Hotel, the IRG Old Hanzi Encoding Interest Group held a discussion, led by Prof. LI Guoying, Vice Dean of Beijing Normal University. Also present were (in alphabetical order): Tom BISHOP of Wenlin Institute, USA; Kyeong Won LEE of Hanyang University, South Korea; LEE Kyoo Kap of Yonsei University, South Korea; LIU Zhijin of East China Normal University; LONG Yuchen of Macao; NGUYEN Quang Hong of Vietnam; WEI Li of The Commercial Press; YE Maofeng of China Electronic Technology Standardization Institute.

No formal vote was taken, but there appeared to be agreement on the following.

- (1) There is a pressing need for encoding Old (pre-Qin Dynasty) Hanzi to facilitate important historical and linguistic research, education, and publishing. The absence of a standard encoding is an obstacle to creating and exchanging Old Hanzi texts, fonts, and databases.
- (2) Three types of Old Hanzi that particularly need encoding are Jiaguwen (oracle bone inscriptions), Jinwen (bronze inscriptions), and Xiaozhuan (small seal). While other types (such as Zhangguo Wenzi, writing from the Warring States period) should eventually be considered, the three types are very important and have been most extensively studied, and are therefore emphasized at this stage.
- (3) There is generally no one-to-one correspondence between Old Hanzi and modern Hanzi, or between the main types of Old Hanzi, and the meanings and correspondences of many Old Hanzi are still unknown. Therefore, each type of Old Hanzi should be completely encoded in its own separate block (even when there appears to be a clear one-to-one correspondence with modern encoded Hanzi). This will enable the different types to be distinguished in plain text, and will facilitate research leading to mappings of correspondences.
- (4) Scholars commonly arrange Old Hanzi (not only Xiaozhuan but also Jiaguwen and Jinwen) according to the 540 radicals of Shuowen Jiezi. It is therefore proposed to use this order for assigning codepoints for all types of Old Hanzi. Old Hanzi that don't correspond to Shuowen characters may be positioned either according to their theoretical Shuowen radicals (when plausible), or in a section for characters without determined radicals. The exact ordering may necessarily be somewhat arbitrary, but once standard encodings exist, it will be possible to implement useful indexes and look-up methods according to various organizational principals.
- (5) For deciding between unification and separation of similar forms that are believed to be variants of each other, the unification principles for Old Hanzi should be established, resembling those for modern Hanzi, but with slightly more tendency toward separation, since the possible significances of some small differences are still uncertain, and Old Hanzi are primarily of interest to scholars who often need to make fine distinctions.
- (6) Only authentic Old Hanzi should be encoded in the new Old Hanzi blocks. Sometimes new imitation Old Hanzi, corresponding to modern Hanzi, are created for decorative purposes, but such usage is unimportant for research into authentic Old Hanzi, and in any case is already possible using special fonts with the same encodings as modern Hanzi.
- (7) The next steps to be taken include: (a) attempt to establish a consensus on the principles listed above; (b) prepare draft code charts for the three main types of Old Hanzi, along with detailed sources and references for each character; (c) make the draft code charts available for evaluation by the international community; (d) revise the draft code charts on the basis of this evaluation; (e) submit the revised code charts to IRG for approval.
- (8) The following persons have volunteered to take responsibility for initial preparation and distribution of draft code charts: LI Guoying for Xiaozhuan and Jiaguwen; LIU Zhijin for Jinwen.

(End of Document)