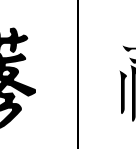JTC1/SC2/WG2 N4650
2014-12-01

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
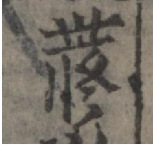Международная организация по стандартизации

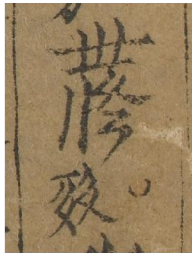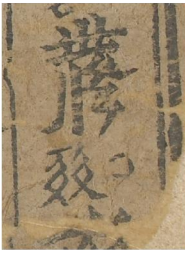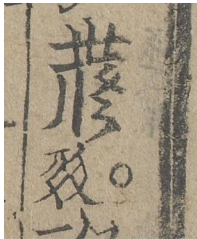|	
Doc Type:	**Working Group Document**
Title:	**Discussion of Tangut character L2008-4148**
Source:	**Andrew West, Viacheslav Zaytsev, Michael Everson**
Status:	**Individual Contribution**
Action:	**For consideration by JTC1/SC2/WG2 and UTC**
Date:	**2014-12-01**

Discussion

A review of the Tangut repertoire in PDAM2.2 has identified one character that has a problematic glyph form. The character in question is 176CE 蓚 (L2008-4148), meaning "common people", which in some modern sources is written with a dot stroke under the 卄 component, whereas in some other modern sources it is written without this dot stroke. The form of the character in PDAM2.2 follows Li Fanwen 2008 in having a dot stroke.

Nevsky 1960	Sofronov 1968	Shi Jinbo et al. 1983	Han Xiaomang 2004	Li Fanwen 1986	Li Fanwen 1997	Kychanov 2006	Li Fanwen 2008
蓚	蓚	蓚	蓚	蓚	蓚	蓚	蓚
dot	no dot	dot	dot	no dot	dot	no dot	dot

In primary Tangut sources this character is only attested in dictionaries and other linguistic works, so there are only a very limited number of known examples of its use. Nevertheless, every single legible example of this character that we have been able to identify in primary Tangut texts is written with no dot stroke under the 卄 component, as shown below.

Wenhai 1-172.152*	*Wenhai* 1-88.222	*Wenhai* 3-15.222	*Tongyin Wenhai Baoyun Hebian* A 23.124

Tongyin A 17A51	*Tongyin* B2 18A17	*Tongyin* B3 18A17	*Tongyin* D 18A17

*In *Wenhai* 1-172.152 the middle vertical stroke of the ⧸ component is broken (a defect in the wooden printing block) so that in a poor quality reproduction of the text it might have been misinterpreted as a dot stroke.

The reason why some modern sources write the character with a dot stroke which is not present in original Tangut sources is that the component beneath ⧸ is expected to be the semantically related character 㒼 (L2008-3818) "person". We cannot be certain of how L2008-4148 was constructed because the original character construction analysis for this character (given for all head entries in the *Wenhai* "Sea of Characters" dictionary) has not survived. However, according to the analysis by Jia Changye 贾常业, the character L2008-4818 "common people" is derived from the top of L2008-4172 㦴 "people" and the whole of L2008-3818 㒼 "person" (see Jia Changye, *Xīnbiān Xīxiàwén Zìdiǎn* 新编西夏文字典 "New Dictionary of the Xixia Script" [Gansu Wenhua Chubanshe, 2013] p. 482). Although this analysis is plausible, it does not accord with evidence of the dotless glyph forms found in primary Tangut sources.

An alternative explanation is that although the bottom part of L2008-4818 was derived from L2008-3818, the dot stroke was deliberately omitted to avoid interfering with the vertical strokes of the top component. In this case the analysis of the construction of L2008-4818 would be that it was created from the top of L2008-4172 and the abbreviated whole of L2008-3818.

Recommendation

We believe that it would be wrong to use an artificial modern glyph form for L2008-4818, based on what some modern scholars think the character should look like, rather than what the character actually looks like in authentic Tangut sources. We therefore propose modifying the glyph for L2008-4818 to follow the models given in primary Tangut sources, as shown below.

PDAM2.2 Glyph	Proposed New Glyph
蓡	蓡

As the new glyph has one fewer stroke (11 totals strokes instead of 12), we also propose moving 176CE to after 17A60 (L2008-4098), and reordering 17A61..176CD to 17A62..176CE.