Title: **Report of Shuishu discussion in WG2 #68 (2019/June/19)**
Date: **2019/July/01**
Place: Microsoft Building 40 (Redmond, US)
Participants: Pan Zhongxi, Chingju Lai (Daisy), Chen Zhuang, Andrew West, Michael Everson, Deborah Anderson, Lisa Moore, Liang Hai, Ken Whistler, suzuki toshiya.
Source: suzuki toshiya
Relevant docs: N4922 Additional repertoire for PDAM 2.2 (Suignard)
N4758 Updated proposal for encoding Shuishu (China)
N4839 Towards the ordering of the Shuishu script (Everson)
N4956 Analysis of Shuishu Character Repertoire (West / Chan)
N4942 Comments on Shuishu in PDAM2.2 text (suzuki)
N4946 Additional Comments on Shuishu in PDAM2.2 text (suzuki)

Because no Shuishu experts had participated in the last WG2 meeting at London, Andrew explained his submission WG2 N4956 to Shuishu experts (Pan Zhongxi and Daisy) briefly. The discussed remarks are following (ordered from the remak with higher priority):

a) Unification and how to choose the representative glyph

In PDAM 2.2 Shuishu character set, the identity of the characters are defined by their pronunciations and meaning (specified by corresponding Han words). There are no multiple characters sharing same pronunciation and same meaning. To choose single representative character, the statistic survey for Shuishu materials preserved in Tsinghua University was done. But, some experts concerned that the rules to pick the representative glyphs from the statistical survey result are sometimes unclear, and some arbitrary decisions might be applied, as suzuki toshiya asked in WG2 N4942 and N4946. The experts recommended to avoid making a character set as a bunch of arbitrary decisions, and show clear reasoning behind every character decision.

In WG2 N4956 p.3, the summary of recommendation, Andrew suggested to set some unification criteria, and encode the variations separately if their origins are clearly different and non-unifiable. Ken Whistler supported this suggestion by the explanation with 3D model of glyph shapes, phonetics and semantics; the character identity is corresponding to a continuous region in 3D space, too significant glyph shape difference could be a rationale to encode multiple characters sharing same phonetics and meaning. This suggestion may have some influence on the purpose of the

standardization of the character set, so China delegates postponed the official response after the discussion with other experts.

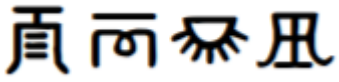
Also Ken pointed that there is a difference between encoding the script and designing an orthography. The removal of the variants looks like as if the purpose were an introduction of new orthography of Shuishu into the industrial standard. China delegates responded their background study (to be published in future) is meant to be inclusive of all major variants, instead of being specific to regional materials or based on a particular calligraphy practitioner's dictionary. But the experts find the difficulty to evaluate the quality of the proposed character set without the publishing of the background studies.

b) Repertoire

In WG2 N4956 section 3.2, Andrew and Eiso tested the cross sections of the character or glyph collections in the past researches and PDAM2.2 character set, and 273 characters in PDAM2.2 are not found in the past researches. Andrew asked about the requirement of the 273 characters. China delegates responded the 273 characters have been checked by Shuishu masters and they are really required. But, some experts were concerned about its stability, even if there is no procedural error; if the requirement is supported by the selected source collection, the characters taken from other published studies should be dealt with in the same way, so there is no reason to reject characters in the past researches. Andrew and suzuki suggested to give the priority to the stable characters which are found in all of PDAM2.2 and past researches (including evidence in published dictionaries), and have the clear identities such as calendar-related characters, numbers, directions. The additional characters should be discussed and proposed later. However, Michael Everson pointed the number of the "stable" characters is unclear yet, so it is questionable whether this idea is practical solution for the users. China delegates decided to bring the comments in this meeting, discuss with other experts, and respond later.

c) Font design

In WG2 N4956 section 3.6, Andrew and Eiso commented the quality of Shuishu font used in PDAM2.2 needs improvement (and recommended Wéi Zōnglín (韦宗林)'s work: Shuizu Shuǐzú wénzì yánjiū (水族文字研究)). China delegates wondered which points of PDAM2.2 fonts were problematic, because they were designed by the authorized Shuishu masters & calligraphers, and regarded as sufficiently beautiful. Andrew and other experts explained that the preferred designs may differ between the calligraphic work and the computer fonts; especially for the code chart in the character encoding standard, the representative glyphs are expected to be designed consistently, and the calligraphic decoration should be minimized. A good model would be the normalized shapes from

Wéi Shìfāng's 2007 dictionary, such as (p. 46 of N4956): 

China delegates understood which points are regarded as problematic, but the official response is postponed after the discussion with other experts.

In summary, no consensus to make some concrete action items, and the experts decided to wait for the responses from the experts in China. The experts recommended an early stage circulation of the analysis of the existing Shuishu documents and the rationale of the character set to be proposed, to the Script Ad Hoc Group (which has monthly meetings) and UTC (which has quarterly meetings). (Script Ad Hoc submissions should be sent to dwanders@berkeley.edu.)

(end of document)