

Title: Comments on WG2 N4973

Source: suzuki toshiya

Date: 2018-06-17

Type: individual contribution

Although it was a pity that there was no traffics from China and TCA experts in the Seal adhoc mailing list, the continuous efforts and the submission WG2 N4973 is appreciated. The members of Seal adhoc and the subscribers of the list are encouraged to review that, to form a consensus before the official report from the adhoc to WG2. Besides of the discussion in the adhoc, WG2 N4973 raised an interesting point which WG2 experts comments are expected.

1. There are 3 Different Scripts in Shuowen?

In the duplicate list category B, “possible duplicate” cases, there is a mention about the pairs of Zhuanwen (篆文) versus Zhouwen (籀文) or Guwen (古文), why they should be coded separately.

3	07909	水	沿		yan2	緣水而下也。从水占聲。《春秋傳》曰：“王沿夏。”
	07647	水	沿		yun3	水。出河東東垣王屋山，東爲涕。从水允聲。古文。
4	03714	章	章		guo1	度也，民所度居也。从回，象城口之重，兩亭相對也。或但从口。（音章。）凡口之屬皆从口。
	09841	土	章		yong1	城垣也。从土庸聲。古文墉
5	06088	欠	歛		xiao4	吟也。从欠肅聲。《詩》曰：“其歛也訶。”
	00932	口	歛		yu4	吹聲也。从口肅聲。籀文嘯从欠
6	06019	見	尋		de2	取也。从見从寸。寸，度之，亦手也。
	01350	彳	尋		de2	行有所得也。从彳尋聲。古文省彳。

Figure 1: Zhuanwen versus Guwen pairs in the possible duplicate category B list of WG2 N4973

The position in WG2 N4973 could be summarized as:

- Their glyph differences are not structural, but the line type or very subtle position difference.
- The existence of Zhouwen (籀文) and Guwen (古文) in Shuowen is an important record of the history about the differentiation of Old Hanzi in Warring States Period, between Qin and 6 Kingdoms¹.
- So they are worthful to be coded separately in the small seal glyph system.

2. Some Extra Cases for Category B?

It seems that this position requests the separated encoding, by the recognition; the Zhuanwen (篆文), Zhouwen (籀文) and Guwen (古文) are different scripts (at least, diverged scripts), and the difference of the scripts should be considered as more significant than the difference of the glyphs in the same script. It is supposed from the fact that “馬” variants are not listed in the category B list.

孫	陳	N4688	TCA	毛4	華東	北師	逢甲	段
								
06957	06957	06956	06957	06955	07906	13585	26861	06687
								
06958	06958	06957	06958	06956	07907		37018	06688
								
06959	06959	06958	06959	06957	07908		37009	06689

Figure 2: Zhuanwen (篆文), Guwen (古文) and Zhouwen (籀文) of “Horse” in various editions of Shuowen

It is difficult to design Guwen “Horse” and Zhouwen “Horse” distinctively. Somebody may think the proposed glyph (in TCA column) for Guwen has a curve to the right but that for Zhouwen has a curve to the left, but it is not stable in other editions. For example, the glyphs in Songben did not have such contrast.



Figure 3: Guwen and Zhouwen of “Horse” in Two Songben of Shuowen

¹ There is no reference, but I guess this statement refers “戰國時秦用籀文六國用古文說” by 王國維.

3. Expected Impact by Dealing Zhuanwen, Guwen and Zhouwen as Different Scripts

If the distinction of these 3 scripts is regarded as stably presented by the typeface design, is it possible to encode Guwen 魏三体石經, when somebody proposes? Also, the materials like 汗簡 or 說文古籀補 give the characters whose abstract structures are almost same with Zhuanwen, but the typeface design is quite similar to Guwen.



Figure 4: Example of 魏三体石經 (Chinese 3 Script Stone, Wei dynasty), in each column, from upper to lower, Guwen (古文), Zhuanwen (篆文) and Lishu (隸書) for “same characters” are inscribed.

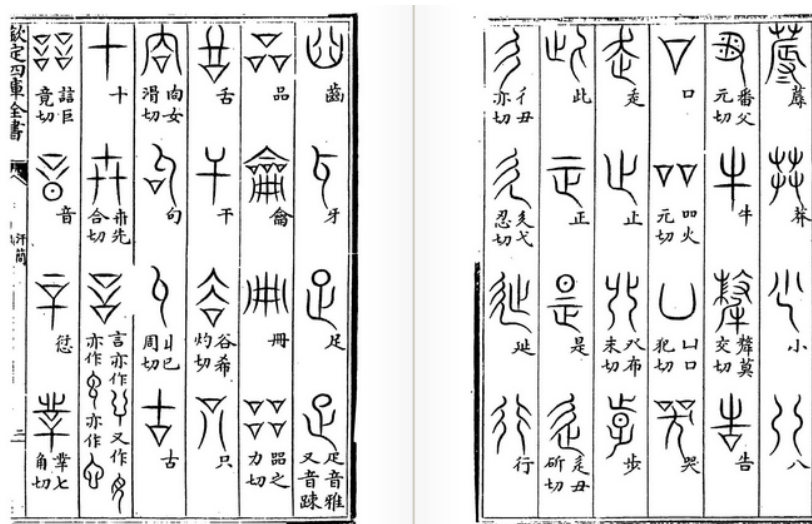


Figure 5: Example of Guwen listed in 汗簡 (HanJian, 四庫全書本).

WG2 N4973 proposes the separated encoding of the Zhuanwen, Guwen and Zhouwen, by relating the distinctions to the archaeological ideas. But current standardization of Seal script is focused to the single reference, Shuowen Jiezi, it is not good idea to justify the proposed encoding architecture by post-Shuowen archaeological ideas. Because if we do that, maybe the number of the characters we should consider would be far greater than 11000. Also there would be the suggestions to separate Zhuanwen, Guwen and Zhouwen into different blocks, because the Guwen or Zhouwen characters which Shuowen does not give (e.g. Shuowen do not give the Guwen for 口, but Hanjian gives) would be proposed in future, by referring non-Shuowen materials.

4. Proposal

I understand, the original motivation to encode Shuowen Seal, is the limiting the scope of the materials to compact. But at the same time, the characters in Shuowen are hoped to be coded separately, regardless with a point whether they are really usable distinctively. Considering such requests, my proposal how to handle Zhuanwen, Guwen and Zhouwen is following.

- The separation of Guwen and Zhouwen are exceptionally defined for the characters clearly given in DaXu version of Shuowen. Regardless with the shape similarity (like Guwen Horse and Zhouwen Horse).
 - Representative glyph shapes should be designed to clarify whether the character is Zhuanwen, or Guwen, or Zhouwen, as much as possible.
 - Seal adhoc group should discuss whether they should be distinguished in NFKD/NFKC normalization.
- If Guwen-like or Zhouwen-like characters are proposed by referring to non-DaXu versions of Shuowen (like 唐寫本木部殘卷) or non-Shuowen materials (like 汗簡, 說文古籀補), they should be tested whether they have the non-unifiable shape except of the typeface design. If the difference is only the typeface design, they should be unified with existing Zhuanwen.
 - Of course, the introduction of the variation sequence should be discussed.

I hope to receive some feedbacks to my proposal from China, TCA and other WG2 experts.

(end of document)