

Date: 2020/04/01

Title: The pairs in N5105 glyphs versus CCZ glyphs to be discussed

Author: suzuki toshiya

References: WG2 N5105, WG2 N5119.

As resolved in WG2 N5119, I selected the pairs of the glyphs from N5105 (based on Tonghuaxie version of Shuowen Jiezi, 藤花樹本說文解字) and CCZ (Chen Changzhi version of Shuowen Jiezi, 陳昌治本說文解字) glyphs which were circulated through Seal script mailing list, on 2019-10-08:

<http://gyvern.ipc.hiroshima-u.ac.jp/~mpsuzuki/ShuoWen/review2019/THX2CCZ-20191008-1730.pdf>

The results are summarized in the Excel file (Open Document spreadsheet file is also available), like:

seq	CCZ location	corresp. ucs	N5105 code	N5105	CCZ	tag	type	guwen ?	taboo info	connection, overshoot	struct	reprint damage	minor stroke
00116:_	陳:卷01上.葉08右.g04	璽,珠	U+4E78				A				○		
00241:_	陳:卷01上.葉14左.g05	璽,歎,璽	U+4EFA				A				○		

This document provides some additional notes for several pairs, to clarify the points to be discussed. To track the root of the glyphs in 藤花樹本, there is an excellent article by 董婧宸, “藤花樹本《說文解字》底本及校刊考” [1] which clarified the relationship among 海源閣本, 汲古閣剗改本 and 藤花樹本. For CCZ’s base 平津館本, her another excellent article “孫星衍平津館仿宋刊本《說文解字》考論” [2] is helpful to figure the relationship with 海源閣本.

1. Legends

The headers of the table are described in the table below.

column header	description of the data in the column
seq	The sequential number of the items in Chen Changzhi version of Shuowen Jiezi (CCZ, 陳昌治本說文解字).
CCZ location	The concrete location info where the item appears in CCZ. The volume, sub-volume, leaf, right-left page, the number of the glyphs in specified page.

corresp. ucs	Some corresponding UCS character(s) to identify the Shuowen Seal glyph. These characters are temporarily listed by the author, it is not the draft for the code chart (please find WG2 N5119 “Chart Format” section).
N5105 code	The temporal UCS codepoint to show the glyph, with the font used to print N5105.
CCZ	The glyph cropped CCZ printed by Zhonghua Shuju (中華書局).
tag	The tags to collect similar glyphic difference. For example, there are many glyphs including “者” as their components, and there is a group whose “者” components are different between N5105 and CCZ. In such case, they are tagged by “者”, to help to make the decision consistent.
type	The classification of the glyphic difference: A, B, C, D, G, Q and R. Please find the detailed explanation in below.
guwen	Whether the glyph is drawn by questionable stroke style. 陳 means the stroke style of CCZ glyph is questionable, 藤 means the stroke style of N5105 is questionable. Please find “the outline of type G items” below, for detail.
taboo info	If the glyph is supposed to be modified to avoid the sacred character in the name of the emperors and their families, and if it is identified whose name is avoided, the info is recorded in this column.
connection, overshoot	If the glyphic difference is regarded as the difference by connected/disconnected strokes, terminated/overshooting strokes, this column is filled by O.
struct	If 2 glyphs are made by same component, but different positions (e.g. covering vs left-right, etc), this column is filled by O.
reprint damage	If CCZ glyph is supposed to be damaged by some (re-)printing issue, this column is filled by X.
minor stroke	If the total numbers of the strokes in 2 glyphs are different, some short note is written in this column. 陳本省 means “CCZ glyph omits a few strokes in comparison with N5105 glyph”. 陳本加 means “CCZ glyph has additional strokes in comparison with N5105 glyph”.

1.1. Classification of the Glyphic Difference

The pairs are classified into 6 types:

Type “Q” (6 items): for this type, some etymological discussions are needed before making a decision by the visual presentation. Some CCZ glyphs are questionable, but there are a few questionable N5105 glyphs either.

Type “R” (28 items): for this type, it seems that the (recently available) reprint from 中華書局縮印版陳昌治本 might have some printing in the glyph. It is not recommended to use the “untouched” scanned image to produce the scalable font, because it can mislead the concept of the fine structure of the glyph.

Type “G” (28 items): for this type, there is no significant glyphic difference, but one of the pair is drawn by the needle-like strokes (for Guwen 古文 or Zhouwen 籀文), and counter part of the pair is drawn by the normal stick-like strokes (for Zhuanwen 篆文). it should be discussed which glyphs should be modified, or, both should keep

current design, before the discussion of unifications.

Type “D” (2 items): for this type, the visual shape of one glyph is quite similar to another glyph with different identity. The decision “not to encode this glyph separately, even if it has different semantics” should be included as a possible option in the discussion.






Type “C” (33 items): for this type, the glyphic difference is rather easy to notice, but it would not lead the users to different entry in Shuowen Jiezi. thus, if we don’t take Shuowen Seal as an inflating character set, the unification of the pair could be acceptable option. but there might be some objections from etymological viewpoints.








Type “B” (89 items): for this type, the glyphic difference is not so hard to notice, and the difference could be found in multiple pairs, thus, the “handling by VS” would be the recommended option.

Type “A” (632 items): for this type, the glyphic difference is (expected to be) very frequently found. In the most cases, such glyphic difference would not guide the users to different items in Shuowen Jiezi. To prevent the inflation of the codespace of Shuowen Seal, the “handling by VS” would be the strongly recommended option.

2. Detail of Type Q Items


A brief explanation what should be discussed are written in below. I’m not asking for further glyph correction at all (I respect all experts working to complete WG2 N5105, so I have no doubt about the requirement. however, I want to reconfirm some stability).


paired images (N5105 vs CCZ)	Explanation
	<p>陳昌治本 might have mistakenly used “水” instead of “泉”, although 平津館本, the source of 陳昌治本, have used “泉” correctly. The requirement of the separated encoding of 陳昌治本 glyph is really questionable, but too significant to unify them. I want to ask 中華書局 experts whether they would “correct” this glyph in future reprint of 縮印版陳昌治本, or, they would keep the current glyph as it is.</p>
	<p>陳昌治本 glyph includes a “口” component on the above the “鼎” component, which is not included in 藤花樹本 glyph. I could not find other versions of Shuowen Jiezi without the “口” component. Before the discussion of the unification, I want to reconfirm the stability the current glyph, by the experts.</p>
	<p>There is no recognizable glyphic difference between N5105 and 陳昌治本. But, if we collect the glyphs including “畏”, you would be able to notice this pair is exceptionally using  (under-part of 辰), instead of . I want to reconfirm the stability of the current glyph, by the experts. It seems that this glyph shape has been since 宋刊小字本, it is an option to use current glyph as “we follow Sung-ben glyph, even if it is etymologically incorrect” (段注本 and 祁騫藻本 seem to have corrected this glyph).</p>

	<p>The glyphic difference is 口 or 白. N5105 glyph follows 藤花樹本原本 , which is similar to 海源閣本  or 王昶旧藏本 . But, this glyph shape does not match with its description, saying as “从火、毀聲” (毀 is drawn as , I could not find other versions using 毀 instead of 毀 (石印版藤花樹本 by 商務印書館 changed this character to use 白). I want to reconfirm the stability of current glyph, whether it is preferred shape, and would be kept as it is now.</p>
	<p>There is no recognizable glyphic difference between N5105 and 陳昌治本. But, when we try to find 尋 from Shuowen, we can find 尋, but no 尋 as self-standing item. If the description of the characters using 尋 say as “尋 is simplified form of 尋, it would be reasonable explanation, but often it is described as “尋聲”, no mention about the simplification. It has been already known by 段玉裁, so he “corrected” this inconsistency by replacing all “尋” by “尋”. Current glyph is preferred and already stabilized, even if there might be some etymological claims? Maybe a few more characters including 尋 should be checked.</p>
	<p>The unification of 田 and 囧 is questionable (although there is no existing glyph , and looking like a mistakenly designed component in 陳昌治本 (平津館本 uses 田). Before going to the unification discussion, I want to ask 中華書局 experts whether they would “correct” this glyph in future reprint of 縮印版陳昌治本, or, they would keep the current glyph as it is. “not to encode 陳昌治本 glyph” might be another option.</p>




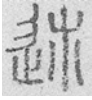
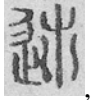




















3. Outline of Type G Items




In most cases, the found Guwen (古文) / Zhouwen (籀文) glyphs are drawn by normal stick-like strokes. But the last one, #08422, “泉” is a contra case. “泉” is the representative of the radical “泉” (部首筆頭字), and its description does not contain Guwen or Zhouwen keyword, and but it is drawn by needle-like strokes, as if it were Guwen or

Zhouwen. 藤花樹本原本 shows the glyph with stick-like strokes like , but the font in N5105 is drawn by the

needle-like strokes, like .

4. Detail of Type D Items

paired images (N5105 vs CCZ)	explanation
	<p>The contrast of #01224 vs #01336 in 藤花樹本(原本) is:  versus . But, the contrast in 海源閣本 is:  versus , quite subtle (maybe 2 side strokes are snaking or not).</p> <p>That in 陳昌治本 is almost same:  versus .</p> <p>If we decide  and  are not unifiable and decide to encode them separately, there would be 2 similar glyphs encoded separately,  and .</p> <p>The points to be discussed are two.</p> <p>① The glyphic difference in 海源閣本 & 陳昌治本 are sufficient to encode 2 glyphs separately? From the viewpoint of the semantics, they are different characters.</p> <p>② If their glyphic difference is insufficient to assign 2 codepoints, is it acceptable to encode  as a (VS-handled) variant of  with same semantics? Or, it should be encoded as a (VS-handled) variant of  with different semantics?</p>
	<p>The contrast of #06922 vs #06924 in 藤花樹本(原本) is:  versus . But, the contrast in 海源閣本 is:  versus , quite subtle.</p> <p>That in 陳昌治本 is almost same:  versus .</p> <p>If we decide  and  are not unifiable and decide to encode them separately, there would be 2 similar glyphs encoded separately,  and .</p> <p>The points to be discussed are two.</p>

	<p>① The glyphic difference in 海源閣本 & 陳昌治本 are sufficient to encode 2 glyphs separately? From the viewpoint of the semantics, they are different characters.</p> <p>② If their glyphic difference is insufficient to assign 2 codepoints, is it acceptable to encode  as a (VS-handled) variant of  with same semantics? Or, it should be encoded as a (VS-handled) variant of  with different semantics?</p>
--	--

References

- [1] 董婧宸：藤花榭本《說文解字》底本及校刊考，《文献》，2019年06期，p. 47-61, http://www.360doc.com/content/19/1207/19/90557_878119552.shtml
- [2] 董婧宸：孫星衍平津館仿宋刊本《說文解字》考論，《勵耘語言學刊》，2018年01期，p. 219-234, https://www.sohu.com/a/273624871_717218

Acknowledgements

Here I express many thanks for helping the listing and drafting works: Masaki Tsukada, Yosuke Nakayama, Tsukasa Oi, Taichi Kawabata and Hajime Okuma. This work was supported by JSPS Grant-in-Aid 16K00460, 19K12716.

(end of document)