

## ISO/IEC JTC1/SC2/WG2 N5189

**Doc Type:** Working Group Document  
**Title:** THX Seal glyph Correction Principles  
**Source:** Selena  
**Status:** Proposal  
**Reference** N4853, N4973, N5089, N5105, N5117  
**Date:** 2022- 06 -17

## Preface:

Since WG2 [N4688](#) (2015-10-20), it was proposed to use Tenghuaxie version Shuowen Jiezi (藤花榭本說文解字) to encode the Shuowen Seal glyphs. For this purpose, more than 20 proposals, among them, there are four papers that talk about glyph correction. The reasons for glyph correction can be summarized into six types, scattered in documents WG [N4853](#)(2017-09-01), [N4973](#)(2018-06-12), [N5089](#)(2019-06-01) and [N5117](#)(2019-09-27) files. In order to facilitate expert review and future expansion of the proposal, this time, Selena specially compiled a complete document based on the chapters in the above document that mentioned the Shuowen Seal glyphs to be modified, and named this sub-document the " Shuowen Seal Glyph Correction Principles." It can be used as a follow-up in the future expansion and review. Expert’s opinions for modification are appreciated.

The reasons for the modification of Shuowen Seal glyphs are as follows:

1. Handling of taboo glyphs(避諱字)
2. Handling of exact duplicate(重複字形)
3. Handling of the original broken glyphs (殘斷字形)
4. Handling of the inconsistent component(不一致之部件)
5. Handling of possible duplicate(近似字)
6. Handling of misprint of the original book(原書誤刻)

## Detailed descriptions and examples:



The following chapters will explain in detail the categories of the " Shuowen Seal Glyph Correction Principles ", and each category will include a case.

## 1. Handling of taboo glyphs(避諱字)

The taboos of Tenghuaxie version are presented in two ways of changing the character or omitting the strokes.

There are two ways to deal with it:

- (1)If it is confirmed “changing the character”, please refer to the other version to restore it, such

as 「正」:  → 

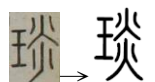
Vol 1 · 禛:

Serial No. in Tenghuaxie version	Original Glyph	TTF Font	Corresp Modern Char.	Refer to the other version
00023			禛	 

The original glyph is the character "正", which is a taboo character for "禛", because Emperor Yongzheng (雍正) was named "胤禛". Confirmed as a taboo character, so it was changed back to the original character. The real "正" character is in Volume 2, as follows:

Vol No.	Serial No.	Serial No. in Tenghuaxie version	Original Glyph	TTF Font	Corresp Modern Char.	SW Radical	SW Radical Number	Type
417	01200	01156			正	正	31	Zhengzhuang

(2) If it is confirmed 'omitting the strokes', use the original glyph as the prefix, such as 「琰」



Serial No. in Tenghuaxie version	Original Glyph	TTF Font	Corresp Modern Char.	Refer to the other version
00140			琰	 

Emperor Jiaqing's (嘉慶) name was "永琰" and later his name was changed to "顥琰", as a result, the seal glyph of the character "琰" was written in omitting strokes to express taboo.

## 2. Handling of Exact duplicate(重複字形)

WG2 N4853's Q2 clearly defines the three categories of the duplicates of the Tenghuaxie version: exact duplicates, possible duplicates, and non-duplicates. Among them, non-duplicate is treated as different character; it will be encoded separately.

Possible duplicates also will be encoded separately. In N4973: "These duplicate character sets have slight differences in line number, line type and component position, including style difference in Guwen or Zhouwen etc. Considering that Shuowen is not only a dictionary but also a comprehensive sorting and build-up of the Shuowen Seal glyph system; therefore, even if

glyphs that have only subtle differences are worth keeping, and they should be encoded individually. The Zhouwen (Chin Glyphs in the Warring States Period) and Gowen (the Six Kingdoms Glyphs in the Warring States Period) collected in the duplicates in Shuowen are valid supplements to the Shuowen Seal glyph system, they should be preserved and be encoded respectively.”




Descriptions of Exact Duplicate:

These exact duplicate sets are merely different in line shape and length, and they have no influence on glyph structure. It has been confirmed only one code point will be encoded.

There are two ways to deal with repeated glyphs of "Exact Duplicate":

(1) All are Zhengzhuan, encoded according to the leading order

For example, "右"

Vol	Serial No. in Tenghuaxie version	Original Glyph	TTF Font	Corresp Modern Char.	SW Radical	pronunciation	meaning	note
2	00939			右	口,又	you4	助也。从口从又。	encode
3	02078		-	右	又	you4	手口相助也。从又从口。	withdraw

The glyphs and meanings of the two "右" Glyphs are very similar. #00939 is in volume 2, #02078 is in volume 3, and #00939 is ranked first, so #00939 will be encoded.

However, according to the opinions of Shuowen Seal experts, different radicals may represent different meanings. Therefore, one glyph is grouped into two radicals (that is two identical Shuowen Seal glyphs with different radical 異部重文), such as #00939 has both the "口" radical and the "又" radical, which means that the character "右" is allowed to have two radicals.

(2) One is a Zhengzhuan and the other is a Chongwen/Variant, encoded in Zhengzhuan.

For example, “垂”

Vol	Serial No. in Tenghuaxie version	Original Glyph	TTF Font	Corresp Modern Char.	SW Radical	Different Radical	type	pronunciation	meaning	Note
11	07729			涇	水		Zhengz huan	tu01/ tu04	河津也。在西河西。从水丞聲。	編碼 encode
2	00880			唾	口		Zhengz huan	tu04	口液也。从口垂聲。	編碼 encode
2	00881		-	唾	口	水	Variant	tu04	唾或从水。	withdraw

#07729 is in volume 11, it's a Zhengz huan, the interpretation in Shuowen is “ferry(河津)”.

#00881 is in volume 2, it's a Chongwen/Variant of #00880, the interpretation in Shuowen is “saliva(口液)”.

Therefore, the two Glyphs only have identical Shuowen Seal glyphs but are with different meanings, however, #07729 is a Zhengz huan and will be encoded in Zhengz huan.

### 3. Handling of the original broken glyphs (殘斷字形)

After verification by experts, it was found that there are some Glyphs contains broken glyphs in the original book due to rubbing or other conditions; the finding was mainly based on the same component glyphs of the same version. Therefore, correction was processed according to the meaning of the character and the reference of other versions. Since showing broken glyphs as the representative glyphs would be confusing for the implementers, the modified representative glyphs would be used in the proposal.

See below:

No	Vol.	Vol No.	Serial No.	Serial No. in Tenghuaxie version	Original Glyph	TTF Font	Corresp Modern Char.	SW Radical	SW Radical Number	Type	Note
1	5	283	03563	03452			虞	虞	166	Zheng z huan	右上從宁，原圖不應斷筆
2	13	746	10300	x380			坊	土	480	Newly added character	依形「从土方聲」，不宜斷開
3	13	809	10363	09983			勸	力	488	Zheng z huan	勸形。原圖不應斷筆。

#### 4. Handling of the inconsistent component (不一致之部件)

After verification by experts, it was found that there are some Glyphs contain inconsistently designed glyphs in the original book; the finding was mainly based on the same component glyphs of the same version. Therefore, modification was processed to restore the original glyph according to the meaning of the character. For example, “卓” and, “遑”.

No.	Vol	Vol No.	Serial No.	Serial No. in Tenghuaxie version	Original Glyph	TTF Font	Corresp Modern Char.	SW Radical	SW Radical Number	Type
	8	290	05939	05738			卓	匕	289	Zhengzhuan
	2	563	01346	01302			遑	辵	33	Zhengzhuan
<p>說明：《說文·匕部》「高也。早、匕為卓。」            早字《說文·日部》「晨也，从日在甲上。」            甲 The seal glyph of 甲 is 「」. There is a short vertical stroke on the top of the glyph.            There are 11 characters that have glyphs relevant to the components of 「卓」, but only 「卓」 and 「遑」 have different seal glyph, so they will be modified to the original glyphs in line with consistency and Shuowen interpretation.</p>										
1	2	361	01144	01101			趙	走	26	Zhengzhuan
2	7	725	05503	05332			罩	网	279	Zhengzhuan
3	8	72	05721	05538			倬	人	287	Zhengzhuan
4	10	396	07352	07086			焯	火	382	Zhengzhuan
5	10	885	07841	07567			悼	心	408	Zhengzhuan
6	11	254	08139	07852			淖	水	410	Zhengzhuan
7	11	751	08636	08321			鮑	魚	424	Zhengzhuan

8	12	277	08941	08617			掉	手	441	Zhengzhuan
9	12	688	09352	09015			焯	女	443	Zhengzhuan

### 5. Handling of possible duplicate (近似字)

In principle, even if glyphs that have only subtle differences are worth keeping, and they would be encoded individually. (Please see the description of item 2. Handling of Exact Duplicate)

But in the N5089 Q3, two sets of very similar glyphs were adjusted. Due to Xu Shen(許慎) clearly stated that Guwen and Zhouwen should not be resembled. To improve the glyphic difference, the representative glyphs would be updated by using other versions, as explained in the following table:



王	孫	陳	藤本	毛4	毛5	華東	北師	遼甲	段	述	汪	祁
06957	06957	06957	06957	06955	06960	07906	13595	26861	06696	06714	06477	06712
06958	06958	06958	06958	06956	06961	07907	13596	37018	06687	06715	06478	06713
06959	06959	06959	06959	06957	06962	07908	13649	37009	06688	06716	06479	06714









Vol-No. 卷次-序 號	Serial No. 總序號 (含新附字)	Original Glyph 原圖	TTF Font 字型檔	Corresp Modern Char. 今字	SW Radical 正篆部 首	SW Radical Number	Type
1	06957			馬	馬	370	Zhengzhuan
2	06958			馬	馬	370	Variant
3	06959			馬	馬	370	Variant

## 6. Handling of misprint of the original book (原書誤刻)

For most of the Shuowen Seal glyphs, Shuowen gives explanations on the composition of the components. In the problematic Shuowen Seal glyph, the glyph is different from the interpretation. By comparing with other versions of Shuowen, such as Da Xu Ben, we can confirm that they are errors of Tenghuaxie version. Therefore, the incorrect glyphs will be modified.

Examples are as follows:

Vol	Vol No.	Serial No.	Serial No. in Tenghuaxie version	Original Glyph	TTF Font	Corresp Modern Char.	SW Radical	SW Radical Number	Type	Content
1	650	00650	00632			艸	艸	12	Zheng zhuan	「從艸次聲」
8	319	05968	05767			臥	卧	297	Zheng zhuan	原圖疑誤刻，「从人、臣，取其伏也。」
10	392	07348	07082			焯	火	382	Zheng zhuan	原圖疑誤刻，「从火皂聲」，字樣依陳本修正

(End of Document)



前言：

《說文》小篆編碼從 N4688(2015-10-20) 提出以藤花樹本《說文解字》進行小篆字形的編碼之後，約有 20 篇提案文件。其中有 4 個文件談及說文小篆字形的校正，今特別整理成一份完整文件，以更專家審查及爾後廣充時參考用之。

小篆字形校正的理由為：

1. 避諱字
2. 重複字形
3. 殘斷字形
4. 不一致之部件
5. 近似字
6. 原書誤刻

詳細說明如下：


### 1. 避諱字：

避諱字有兩種情形，分別以換字，或是缺筆兩種方式呈現。


























對於藤花本的避諱字，參考了其他版本的字形，還原其原形，並予以編碼。

對於避諱字，有兩種處理方式：

(1)若為「換字」，便比對他本《說文》，改回正確字形。

如「正」字：



卷一·禛：

藤本小篆序號	原圖	字型	對應字	他本比對																											
00023		禛	禛	<table border="1"><tr><td>孫</td><td>陳</td><td>N4688</td><td>TCA</td><td>毛4</td><td>華東</td><td>北師</td><td>逢甲</td><td>段</td></tr><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>00023</td><td>00023</td><td>00023</td><td>00023</td><td>00023</td><td>01079</td><td>11072</td><td>18652</td><td>00023</td></tr></table>	孫	陳	N4688	TCA	毛4	華東	北師	逢甲	段										00023	00023	00023	00023	00023	01079	11072	18652	00023
孫	陳	N4688	TCA	毛4	華東	北師	逢甲	段																							
																															
00023	00023	00023	00023	00023	01079	11072	18652	00023																							

原圖為「正」字，這是「禛」的避諱字，因為雍正皇帝名「胤禛」。確認為避諱字，所以改回原字。真正的「正」字則在卷二，如下：



Vol No.	Serial No. 序號	Serial No. in Tenghuaxie version 藤本序號	Original Glyph 原圖	TTF Font 字形	Corresp Modern Char. 今字	SW Radical	SW Radical Number 部首序	Type 類型
417	01200	01156			正	正	31	Zhengzhuān 正篆

(2)如果經確認是「缺筆」的避諱字，將補上缺筆，恢復原形。如「琰」→.

藤本小篆序號	原圖	字型	對應字	他本比對
00140			琰	

因嘉慶皇帝名「永琰」，後來改名「顥琰」，所以「琰」字篆形就以缺筆來表示避諱。

## 2. 重複字形

WG2 [N4853](#) 的 Q2，明確定義了藤花樹本「重複字形」的三種分類：基本同形(exact duplicate)、字形微別(possible duplicate)以及非同形(non-duplicate)。其中，「非同形」視為不同的字，會分開編碼。「字形微別」也會分開編碼，說明參見 [N4973](#)。「基本同形」說明如下：

「基本同形」的重複字形有兩種處理方式：

(1) 皆為正篆(Zhengzhuān)，以排序在前的進行編碼

如「右」字

Vol	Serial No. in Tenghuaxie version	Original Glyph	TTF Font	Corresp Modern Char.	SW Radical	pronunciation 音讀	meaning 說文釋義	note 備註
2	00939			右	口,又	you4	助也。从口从又。	編碼 encode
3	02078		-	右	又	you4	手口相助也。从又从口。	移除

兩個「右」字的字形、字義都非常相近，#00939 在卷二，#02078 在卷三，#00939 排序在

前，所以將#00939 進行編碼。但是，因為不同部首可能代表不同的意義，所以，將一個字形，歸到兩個部首(即異部重文)，如#00939 同時是「口」部、「又」部，表示「右」字可以有兩個部首。

(2)一為正篆(Zhengzhuan)，一為重文(Chongwen/Variant)，以正篆字形編碼如「涇」字

Vol	Serial No. in Tenghuaxie version	Original Glyph	TTF Font	Corresp Modern Char.	SW Radical 正篆部首	Diffrent Radical 異部部首	type 類型	pronunciation 音讀	meaning 說文釋義	Note 備註
11	07729			涇	水		Zhengzhuan	tu01/tuo4	河津也。在河西。从水丞聲。	編碼 encode
2	00880			唾	口		Zhengzhuan	tu04	口液也。从口垂聲。	編碼 encode
2	00881		-	唾	口	水	Variant	tu04	唾或从水。	移 withdraw

#07729在卷十一，是正篆(Zhengzhuan)，《說文》釋義為「河津」。#00881在卷二，是#00880的重文(Chongwen/Variant)，《說文》釋義為「口液」，因此這兩字僅有小篆字形相同，但意義不同，而#07729為正篆，以正篆進行編碼。

### 3. 殘斷字形


經專家驗證，發現原書中有某些字形由於拓印或其他原因而有殘斷字形。因此，根據字形的意義和比對其他版本而進行更正。由於將殘斷字形會對使用者造成混淆，因此在提案中將使用更正後的字形。

No	Vol.	Vol No.	Serial No. in Tenghuaxie version 藤本序號	Serial No. in Tenghuaxie version 藤本序號	Original Glyph 原圖	TTF Font 字形	Corresp Modern Char. 今字	SW Radical 正篆部首	SW Radical Number 部首序	Type 類型	Note 備註
1	5	283	03563	03452			盧	盧	166	Zhengzhuan	右上從宁，原圖不應斷筆
2	13	746	10300	x380			坊	土	480	Newly added character	依形「从土方聲」，不宜斷開
3	13	809	10363	09983			勸	力	488	Zhengzhuan	虎形。原圖不應斷筆。

#### 4. 不一致之部件

經過專家驗證，發現原書中有不一致的字形；該發現主要基於相同版本的相同部件的字形。因此，根據字理恢復原形。

如「卓」與「連」：

No.	卷次	Vol No.	Serial No. 序號	Serial No. in Tenghuaxie version 藤本序號	Original Glyph 原圖	TTF Font 字形	Corresp Modern Char. 今字	SW Radical 正篆部首	SW Radical Number 部首序	Type 類型
	8	290	05939	05738			卓	匕	289	Zhengzhuan
	2	563	01346	01302			連	辵	33	Zhengzhuan
<p>說明：《說文·匕部》「高也。早、匕為卓。」            早字《說文·日部》「晨也，从日在甲上。」            甲字篆形為「」。上方有一短豎筆。            「卓」字部件相關字形有 11 字，僅「卓」「連」篆形不同，依字理與一致性恢復原形。</p>										
1	2	361	01144	01101			趙	走	26	Zhengzhuan
2	7	725	05503	05332			罩	网	279	Zhengzhuan
3	8	72	05721	05538			倬	人	287	Zhengzhuan
4	10	396	07352	07086			焯	火	382	Zhengzhuan
5	10	885	07841	07567			悼	心	408	Zhengzhuan
6	11	254	08139	07852			淖	水	410	Zhengzhuan
7	11	751	08636	08321			鮑	魚	424	Zhengzhuan
8	12	277	08941	08617			掉	手	441	Zhengzhuan
9	12	688	09352	09015			嫖	女	443	Zhengzhuan

## 5. 近似字

原則上，即使只有細微差別的字形也值得保留，並且它們將被單獨編碼。參見 [WG2 N4973](#)。

但是在 N5089 Q3 中，提出兩組非常相似的字形。由於許慎明確指出，古文和籀文不應該相似。所以，為了增加字形差異，將使用其他版本來代換字形，如下表所示：









王	孫	陳	藤本	毛4	毛5	華東	北師	逢甲	殷	述	汪	祁
06957	06957	06957	06957	06955	06960	07906	13595	26361	06696	06714	06477	06712
06958	06958	06958	06958	06956	06961	07907	13596	37018	06687	06715	06478	06713
06959	06959	06959	06959	06957	06952	07908	13640	37009	06688	06716	06479	06714

Vol-No. 卷次-序 號	Serial No. 總序號 (含新附字)	Original Glyph 原圖	TTF Font 字型檔	Corresp Modern Char. 今字	SW Radical 正篆部 首	SW Radical Number	Type
1	06957			馬	馬	370	Zhengzhuan
2	06958			馬	馬	370	Variant
3	06959			馬	馬	370	Variant

## 6. 原書誤刻

大多數小篆字形，《說文》釋義都有說明部件組成。在有問題的小篆字形中，字形與釋義不同。通過與其他《說文》版本的比對，如大徐本，可以確認它們是藤本的錯誤。以字理、比對他本而修正。舉例如下：

卷次	Vol No.	Serial No. 序號	Serial No. in Tenghuaxie version 藤本序號	Original Glyph 原圖	TTF Font 字形	Corresp Modern Char. 今字	SW Radical 正篆部首	SW Radical Number 部首序	Type 類型	Content 說明
一	650	00650	00632			苳	艸	12	Zheng zhuan	「從艸次聲」
八	319	05968	05767			臥	卧	297	Zheng zhuan	原圖疑誤刻，「从人、臣，取其伏也。」
十	392	07348	07082			焯	火	382	Zheng zhuan	原圖疑誤刻，「从火皂聲」，字樣依陳本修正

(End of Document)