

《華陽國志校補圖注》舊版造字轉碼說明

中研院資訊所文獻處理實驗室
中研院史語所漢籍電子文獻工作小組
2008/6/23 黃榮順 製作

一、《華陽國志校補圖注》一書(華陽國志校補圖注.xml)使用舊版造字 7 個，字頻 37 次，詳如附件一。這 7 個造字中，5 個可轉成 Windows XP 可顯示的字，字頻 34 次；另外 2 個字必須轉成構字式，字頻 3 次。

二、附件一的造字分析表說明如下：

甲、編號：Big5 造字空間為 6217 個，編號由 1 到 6217。

乙、造字：舊版造字

丙、字頻(txt)：造字在「.txt」文件的出現次數

丁、字頻(xml)：造字在「.xml」文件的出現次數

戊、Big5：造字的 Big5 碼

己、Unicode：造字所對應的 Unicode 碼

庚、WinXP：造字在 Windows XP 的對應字形

辛、構字式：Windows XP 若無對應字形，則改採用構字式

壬、備註凡例：

- 1、校對問題，舊版漢籍錯字，必須人工刪除：在舊版漢籍電子文獻中即存在的錯字，因校對時的疏漏而未更正，持續留存在新版漢籍電子文獻中，雖然頻次相同，其為原書所無之誤增字，仍須手動刪除。如編號 86 的「𠄎」（造字區未定義），原書所無，為誤增字，手動刪除。
- 2、重碼問題，採用部件外字，不轉碼：舊漢籍造字與構字式部件外字，皆是建立在 Big5 的使用者造字區中，兩者所使用的碼區是重疊的，在同時使用構字式和舊漢籍造字的新版漢籍中，這些重疊的字碼可能被用以表達兩個不同字形，我們稱之為「重碼問題」。然而漢籍在舊轉新的過程中，若已將原舊漢籍造字的部分全數改為 Unicode 或構字式，則在 xml 出現的造字頻次皆採用部件外字，此時便不需轉碼。如編號 5339，字碼為 8B40，對應的舊漢籍造字為「𠄎」，部件外字為「𠄎」，對原書發現書中皆取部件外字「𠄎」，故不需轉碼。

三、 txt 檔與 xml 檔頻次不同，或頻次雖同，但須人工處理，且造字分析表「備註」欄有備註問題情況者，詳細取代內容可參考附件二的手動取代表。附件二的手動取代表說明如下：

甲、 編號：Big5 造字空間為 6217 個，編號由 1 到 6217。

乙、 造字：舊漢籍造字。

丙、 檔案原文摘錄：摘錄檔案(華陽國志校補圖注.xml)中包含該造字的文句段落，其中紅字底線的部分，為需要手動取代的原文。

丁、 手動取代結果：變更過後的檔案內容，其中藍字底線的部分，為修改後的結果。

四、 「華陽國志校補圖注.xml」發現之未確定字形「」共 46 處，修正情形詳列附件三，附件三「華陽國志校補圖注.xml」修改表體例說明如下：

甲、 檔案原文摘錄：摘錄檔案(華陽國志校補圖注.xml)中包含「」未確定字形的文句段落，其中紅字底線的部分，為需要手動取代的原文。

乙、 手動取代結果：變更過後的檔案內容，其中藍字底線的部分，為修改後的結果。

丙、 備註：目前資料庫無法處理之字形仍維持「」，另標注原書頁數，以便日後查考與修正。

五、 「華陽國志校補圖注.xml」發現之缺字外錯字共 10 處，修正情形詳列附件四，附件四「華陽國志校補圖注.xml」缺字外錯字修改表體例說明如下：

甲、 檔案原文摘錄：摘錄檔案(華陽國志校補圖注.xml)中包含該錯字的文句段落，其中紅字底線的部分，為需要手動取代的原文。

乙、 手動取代結果：變更過後的檔案內容，其中藍字底線的部分，為修改後的結果。

丙、 備註：修改後的字形若屬於「待造字」，則於此欄標示。

附件一、「華陽國志校補圖注.xml」造字分析表

| 編號 | 造字 | 字頻 (txt) | 字頻 (xml) | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|-------------|-------------|------|---------|-------|-----|--------------------|
| 86 | | 1 | 1 | FAB7 | | | | 舊版漢籍錯字，手動刪除。 |
| 4262 | | 1 | 2 | 8456 | | | 麥△丐 | 昏，重碼問題，採用部件外字，不轉碼。 |
| 5339 | 𡗗 | 4 | 11 | 8B40 | 59E7 | 𡗗 | | 𠂇，重碼問題，採用部件外字，不轉碼。 |
| 5525 | 𧆏 | 1 | 14 | 8C5D | 8C83 | 𧆏 | | 乚，重碼問題，採用部件外字，不轉碼。 |
| 5752 | 摠 | 3 | 6 | 8DC5 | 63D4 | 摠 | | 疋，重碼問題，採用部件外字，不轉碼。 |
| 5860 | 无 | 2 | 2 | C6D3 | 65E0 | 无 | | |
| 5863 | 辵 | 1 | 1 | C6D6 | 8FB5 | 辵 | | |

附件二、「華陽國志校補圖注.xml」手動取代表

| 編號 | 造字 | 檔案原文摘錄 | 手動取代結果 |
|----|----|---------------------|---------------|
| 86 | | 人士，十一人女 <u> </u> | 人士，十一人女（手動刪除） |

附件三、「華陽國志校補圖注.xml」 「㊦」修改表

| 檔案原文摘錄 | 手動取代結果 | 備註 |
|---|---|-------|
| ㊦盛㊦中 | ㊦盛㊦中 | p.56 |
| ㊦，盛器也 | ㊦，盛器也 | p.56 |
| ㊦，鹽屑也 | ㊦，鹽屑也 | p.56 |
| 作※，不作㊦ | 作※，不作㊦ | p.56 |
| 作正十字加㊦ | 作正十字加㊦ | p.56 |
| 西字，小篆作㊦ | 西字，小篆作  西  說文  | |
| 作㊦與㊦相似 | 作㊦與㊦相似 | p.56 |
| 必字，小篆作㊦ | 必字，小篆作㊦ | p.56 |
| 最原始的寫法作㊦ | 最原始的寫法作㊦ | p.57 |
| 其先或只作㊦ | 其先或只作㊦ | p.57 |
| 甲文作㊦ | 甲文作㊦ | p.57 |
| 金文作㊦ | 金文作  監  說文  | |
| 古文，有作㊦的 | 古文，有作㊦的 | p.58 |
| 篆文給我看，作㊦ | 篆文給我看，作㊦ | p.58 |
| 蜀，其字作㊦ | 蜀，其字作㊦ | p.220 |
| 其㊦字，象蠶形 | 其㊦字，象蠶形 | p.220 |
| 又㊦字，象二 | 又㊦字，象二 | p.220 |
| 可知此㊦字乃蜀人 | 可知此㊦字乃蜀人 | p.220 |
| 用地方字，作㊦ | 用地方字，作㊦ | p.220 |
| ㊦(蠶)表複數 | ㊦(蠶)表複數 | p.220 |
| 讀㊦為蠶字音也 | 讀㊦為蠶字音也 | p.220 |
| 其字從㊦從人 | 其字從  羊  說文  從人 | |
|     (表示的一般羌人) |     (表示的一般羌人) | p.223 |
|     (表示賣石器 |     (表示賣石器 | p.223 |
|     (表示賣絲的 |     (表示賣絲的 | p.223 |
| 內無㊦附 | 內無  附 | |
| 廖本作㊦附 | 廖本作  附 | |
| ㊦附見《詩》 |  附見《詩》 | |
| 錢、劉、李、《函》作㊦ | 錢、劉、李、《函》作  附 | |
| 宋、明舊刻皆作㊦ | 宋、明舊刻皆作  附 | |
| “㊦附”二字 | “  附”二字 | |
| 墨批“㊦附見 | 墨批“  附見 | |

| 檔案原文摘錄 | 手動取代結果 | 備註 |
|----------|----------|----|
| 𠄎即今胥字 | 胥即今胥字 | |
| 作“𠄎附” | 作“胥附” | |
| 為𠄎為骨。張、吳 | 為胥為骨。張、吳 | |

附件四、「華陽國志校補圖注.xml」缺字外錯字取代表

| 檔案原文摘錄 | 手動取代結果 | 備註 |
|---------------|--------------|-----|
| 樹為簇(cu) | 樹為簇(cù) | |
| 語音作叢(cong) | 語音作叢(còng) | |
| 元豐本作𠄎尸𠄎又𠄎 | 元豐本作𠄎尸𠄎久𠄎 | 待造字 |
| 建寧爨【𠄎𠄎𠄎一】 | 建寧爨【𠄎】 | |
| 廖本作𠄎𠄎𠄎一 | 廖本作𠄎 | |
| 量作𠄎𠄎𠄎一 | 量作𠄎 | |
| 【𠄎𠄎𠄎采𠄎】(蜀)穀二石 | 【𠄎𠄎𠄎𠄎】(蜀)穀二石 | 待造字 |
| 𠄎𠄎𠄎采𠄎字譌 | 𠄎𠄎𠄎𠄎字譌 | |
| 二本作𠄎𠄎𠄎穀 | 二本作𠄎𠄎𠄎𠄎穀 | |
| 嫂，古作火𠄎𠄎 | 嫂，古作女𠄎𠄎 | |