

《馬王堆漢墓帛書(參)-戰國縱橫家書》(標記) 舊版造字轉碼說明

中研院資訊所文獻處理實驗室
中央研究院語言所文獻語料小組
2009/10/20 丁致伶

轉碼主要工作是把檔案中的舊版造字轉換成 Windows XP 能支援的 Unicode 字形，Unicode 目前共收錄漢字 70194 個字，而 XP 只能支援 20902 個字（詳如表一），不支援之字將以構字式表達。例：造字編號 3853 的「𦉑」字，Unicode 編碼是 2776C，由於 XP 並不支援，仍需使用構字式「𦉑𠂇衣」。

表一、Unicode 的字數及編碼區段

| Unicode | 新增字數 | 新增編碼區段 | 總字數 | WinXP |
|---------|-------|-------------|-------|-------|
| 1.1 版 | 20902 | 4E00-9FFF | 20902 | 支援 |
| 3.0 版 | 6582 | 3400-4DFF | 27484 | 不支援 |
| 3.1 版 | 42710 | 20000-2A6D6 | 70194 | 不支援 |

一、舊版造字轉碼分析：

《馬王堆漢墓帛書(參)-戰國縱橫家書》(標記)使用舊版造字 48 個，字頻 252 次，這 48 個造字中，18 個可轉成 Windows XP 能顯示的字，字頻 66 次；另外 30 個字必須轉成構字式，字頻 186 次。

轉碼完成製作轉碼分析表，請參考附件一《馬王堆漢墓帛書(參)-戰國縱橫家書》(標記)轉碼分析表，欄位說明如下：

- (一) 編號：Big5 造字空間為 6217 個，編號由 1 到 6217。
- (二) 造字：舊版造字。
- (三) 頻次：舊版造字在文件的出現次數。
- (四) Big5：造字的 Big5 碼。

- (五) Unicode：造字所對應的 Unicode 碼。
- (六) WinXP：造字在 Windows XP 的對應字形。
- (七) 構字式：Windows XP 無法對應字形改用構字式。
- (八) 備註凡例：備註欄中記錄轉碼後字形及修改原因，凡例如下：

1. 錯字，以程式全部取代：與原書字形不符，並查詢教育部異體字字典確認非異體字後，皆歸類為錯字，以程式取代為正確字形。
例：檔案中使用造字編號 566 的「尸」字，而原書使用之字為「务」，「尸」為錯字，以程式取代為正確字形「务」。
2. 未確定部件或字形，轉碼後手動修改：造字編號 309「?」字形代表未確定的構字式部件或字形，若對應字形超過兩個以上，將在轉碼後比對原書手動修改；若該字為古漢字，則以「#圖+出處.字數」表示。例：「?」代表未確定的構字式部件或字形，在此書中對應字形超過兩個以上，其中一字「?」對應到古漢字字形，轉碼後手動修改為「#圖馬3縱橫.127.25」。手動修改部份請參考附件二《馬王堆漢墓帛書(參)-戰國縱橫家書》(標記)手動取代表。

附件一、《馬王堆漢墓帛書(參)-戰國縱橫家書》(標記)轉碼分析表

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|-----|----|----|------|---------|-------|-----|----|
| 268 | 𠂇 | 1 | FBD0 | 3561 | | 𠂇 | |
| 296 | 𠂈 | 39 | FBEC | | | 𠂈 | |
| 297 | 𠂉 | 26 | FBED | | | 𠂉 | |
| 298 | 𠂊 | 4 | FBEE | | | 𠂊 | |

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|-----|----|----|------|---------|-------|------------|-------------------|
| 299 | 𠄎 | 19 | FBEF | | | 𠄎 | |
| 300 | 𠄐 | 19 | FBF0 | | | 𠄐 | |
| 308 | ∞ | 2 | FBF8 | | | ∞ | |
| 309 | ? | 3 | FBF9 | | | ? | 未確定部件或字形，轉碼後手動修改。 |
| 397 | 帛 | 1 | FCB4 | | | 帛 | |
| 447 | 𠄎 | 1 | FCE6 | | | 𠄎 | |
| 452 | 𠄎 | 1 | FCEB | | | 𠄎 | |
| 530 | 𠄎 | 1 | FD7A | | | 𠄎 | 圭，錯字，以程式全部取代。 |
| 543 | 段 | 2 | FDA9 | | | 段 | |
| 609 | 𠄎 | 1 | FDEB | | | 𠄎冊體 說文𠄐 | 口，錯字，以程式全部取代。 |
| 637 | 𠄎 | 10 | FE48 | | | 𠄎 | |
| 679 | 𠄎 | 1 | FE72 | | | 𠄎 | |
| 687 | 𠄎 | 1 | FE7A | | | 𠄎 | |

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|----|------|---------|-------|-----|-----------------|
| 690 | 𠄎 | 2 | FE7D | | | 𠄎 | |
| 748 | 𠄎 | 6 | FED9 | | | 𠄎 | |
| 750 | 𠄎 | 1 | FEDB | | | 𠄎 | |
| 752 | 𠄎 | 2 | FEDD | | | 𠄎 | |
| 759 | 𠄎 | 2 | FEE4 | | | 𠄎 | |
| 763 | 𠄎 | 2 | FEE8 | | | 𠄎 | |
| 766 | 𠄎 | 1 | FEEB | | | 𠄎 | |
| 768 | 𠄎 | 2 | FEED | | | 𠄎 | |
| 770 | 𠄎 | 6 | FEEF | | | 𠄎 | |
| 772 | 𠄎 | 1 | FEF1 | | | 𠄎 | |
| 780 | 𠄎 | 4 | FEF9 | | | 𠄎 | |
| 784 | 𠄎 | 9 | FEFD | | | 𠄎 | |
| 1958 | 諧 | 1 | 95AB | 8ACE | 諧 | | |
| 2143 | 火 | 1 | 96C7 | | | 火 | 火，錯字，以程式全部取代。 |
| 2745 | | 16 | 9AAD | | | 日𠄎寅 | 日𠄎寅，錯字，以程式全部取代。 |

| 編號 | 造字 | 頻次 | Big5 | Unicode | WinXP | 構字式 | 備註 |
|------|----|----|------|---------|-------|----------------|---------------|
| 2770 | 絳 | 1 | 9AC6 | 7D98 | 絳 | | |
| 2776 | 薈 | 1 | 9ACC | 842E | 薈 | | |
| 3814 | 却 | 1 | 816D | 5374 | 却 | | |
| 3853 | 弊 | 1 | 81B6 | 2776C | | 敝 _衣 | |
| 3946 | 汹 | 2 | 8254 | 6C79 | 汹 | | |
| 4085 | 耻 | 2 | 8342 | 803B | 耻 | | |
| 4190 | 郄 | 1 | 83CD | 90C4 | 郄 | | |
| 4469 | 勻 | 3 | 85AA | 5304 | 勻 | | |
| 4533 | 唸 | 1 | 85EA | 552B | 唸 | | |
| 4750 | 卅 | 6 | 8767 | 5344 | 卅 | | 卅，錯字，以程式全部取代。 |
| 4768 | 妬 | 1 | 8779 | 59AC | 妬 | | |
| 5169 | 戇 | 1 | 89F2 | 6239 | 戇 | | |
| 5353 | 役 | 1 | 8B4E | 4F07 | 役 | | |
| 5355 | 羴 | 1 | 8B50 | 20AE4 | | 羴 | |
| 5617 | 迺 | 2 | 8CDB | 900E | 迺 | | |
| 5860 | 无 | 39 | C6D3 | 65E0 | 无 | | |

二、手動取代表說明：

《馬王堆漢墓帛書(參)-戰國縱橫家書》(標記)轉碼後檔案中出現「❓」字形，比對原書後以對應之字形手動修改；若該字為古漢字，則以「#圖+出處.行數.字數」表示，並製作「❓」字手動取代表，請參考附件二。欄位說明如下：

1. 標記行碼：語言所標記檔裡設定的行碼。
2. 檔案原文摘錄：摘錄檔案《馬王堆漢墓帛書(參)-戰國縱橫家書》(標記)中的文句段落，其中紅字底線的部分，為轉碼後需要手動取代的「❓」字形。
3. 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。
4. 備註：記錄修改相同辭句的次數或其他事項。

附件二、《馬王堆漢墓帛書(參)-戰國縱橫家書》(標記) 「❓」字手動取代表

| 標記行碼 | 檔案原文摘錄 | 手動取代結果 | 備註 |
|------|------------------------------------|--|----|
| 176 | <𠄎 ⁺⁺ 才 ❓(VC1)>{%} 拜 | <𠄎 ⁺⁺ 才 𠄎(VC1)>{%} 拜 | |
| 746 | 𠄎; 口 ❓(NB1) | 𠄎; 口 𠄎(NB1) | |
| 825 | 至(VA) ❓(b) 而(C) | 至(VA) <u>#圖馬3縱 橫.127.25</u> (b) 而(C) | |