

# 《古本竹書紀年輯證》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  
2010/08/09 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 998 的「潰」字，Unicode 編碼是 3D52，由於 XP 並不支援，仍需使用構字式「灬 𠂔 費」。

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：

《古本竹書紀年輯證》使用舊版造字 44 個，字頻 203 次，這 44 個造字中，35 個可轉成 Windows XP 能顯示的字，字頻 193 次；另外 9 個字必須轉成構字式，字頻 10 次。

轉碼完成製作轉碼分析表，請參考附件一《古本竹書紀年輯證》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. 異體字問題：為了使用者查詢和使用的方便，在處理異體字時最主要的方式是以標準字取代，除非是專有名詞或特殊情形，如：人名、地名等。例：造字編號 4134 的「衛」字，是「衛」的異體字，以標準字「衛」取代。
2. 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 2750 的「柅」字，而原書使用之字為「盼」，「柅」為錯字，以程式取代為正確字形「盼」。
3. Unicode 字型呈現差異：Unicode 字型與舊版造字有些微差異，但只是字體風格差異，實際上仍為同一個字，因此仍取 Unicode 字型。如編號 4091 的「臯」字，Unicode 字型呈現為「臯」，實際上仍為同一字。

附件一、《古本竹書紀年輯證》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
827	鮌	1	8E69	9B8C	鮌		
938	燮	1	8EFA	7215	燮		
998	潰	1	8F77	3D52		氵 ㄩ 費	
1011	垚	1	8FA6	3630		土 ㄩ 斗	
1530	紉	1	92D6	7CFA	紉		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1565	絲	2	92F9	7DDC	絲		
2610	鱣	1	99C3	9C53	鱣		
2750	梘	3	9AB2	68FF	梘		盼，錯字，以程式全部取代。
3787	冢	87	8152	51A2	冢		
3840	坂	3	81A9	5742	坂		
3921	皙	1	81FA	6673	皙		
3949	洩	1	8257	6D96	洩		
3964	威	1	8266	70D5	威		
3966	烟	2	8268	70DF	烟		
3977	狗	1	8273	72E5	狗		
3981	珉	3	8277	73C9	珉		
3987	穀	2	827D	7474	穀		
3991	攜	4	82A3	64D5	攜		
4040	竝	3	82D4	7ADD	竝		
4080	羣	5	82FC	7FA3	羣		
4091	臯	3	8348	81EF	臯		Unicode 字型呈現差異。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4093	舐	1	834A	445B		舌 厶 氏	
4104	莅	1	8355	8385	莅		
4134	衛	39	8373	885E	衛		衛，異體字問題。
4184	迹	1	83C7	8FF9	迹		
4252	鰓	1	844C			魚 厶 眾	鰓，異體字問題。
4267	鼈	1	845B	9F08	鼈		
4503		1	85CC			需 厶 兔	需 厶 兔，錯字，以程式全部取代。
4760	覲	1	8771	89B8	覲		
4961	盼	1	88BF	80B9	盼		
4996	斲	1	88E2	65B5	斲		
5026		3	8941			形 广 回 示 口	廩，錯字，以程式全部取代。
5034	弃	1	8949	5F03	弃		Unicode 字型呈現差異。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5096	紆	1	89A9	42B8		纟 厶 亏	
5116	侑	2	89BD	4F37	侑		
5221		8	8A67			𠂇 白 米 𠂇 𠂇	粵，異體字問題。
5244	嬖	1	8A7E	3738		薛 𠂇 女	
5374	盼	4	8B63	670C	盼		
5451	玳	1	8BD2	7398	玳		
5453	涅	1	8BD4	23D40		𠂇 日 工 𠂇	
5591	杞	2	8CC1	233CC		木 厶 巳	杞，異體字問題。
5619	繚	1	8CDD	470C		繚	
5625	瀆	2	8CE3	3D8F		氵 厶 隱	
5697	煑	1	8D6C	7151	煑		

## 二、手動取代表說明：

《古本竹書紀年輯證》檔案中出現「●」字形，共計9字，比對原書後以Unicode字形或構字式手動取代，並製作「●」字手動取代表，請參考附件二。欄位說明如下：

1. 頁碼：「●」字形所在位置之原書頁碼。
2. 檔案原文摘錄：摘錄《古本竹書紀年輯證》檔案有「●」字形的

文句。

3. 原書字形：「●」在原書中之字形，以此字取代檔案原文的「●」。
4. 備註：記錄修改相同詞句的次數或其他事項。

附件二、《古本竹書紀年輯證》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 16	●其名于苕華	𠄎𠄎亞 𠄎	有 2 處
p. 16	●苕華	𠄎𠄎亞 𠄎	
p. 47	●人	谷 𠄎 粵	
p. 77	洛絕於●	𠄎 𠄎 向	
p. 121	宋景●、衛公孫	善 𠄎 文	
p. 129	絳中地●	土 𠄎 席	有 2 處
p. 270	洛絕于●	𠄎 𠄎 向	