

# 《銀雀山漢墓竹簡(壹)-孫子&孫臏》(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  
2009/10/28 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 5534 的「塢」字，Unicode 編碼是 4CA8，由於 XP 並不支援，仍需使用構字式「工𠂔鳥」。

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：

《銀雀山漢墓竹簡(壹)-孫子&孫臏》(標記)使用舊版造字 51 個，字頻 336 次，這 51 個造字中，19 個可轉成 Windows XP 能顯示的字，字頻 94 次；另外 32 個字必須轉成構字式，字頻 242 次。

轉碼完成製作轉碼分析表，請參考附件一《銀雀山漢墓竹簡(壹)-孫子&孫臏》(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. **異體字問題**：檔案內的舊版造字與 Unicode 或漢字構形資料庫所收之字，僅僅只有部件位置或異寫的差異，並無改變該字字義，以 Unicode 字形或漢字構形資料庫所收字形的構字式取代。例：造字編號 712 的「开」字，是 Unicode 字形「𠂔」的異體字，僅僅只有部件位置或異寫的差異，並無改變該字字義，因此以 Unicode 字形「𠂔」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 5746 的「每𠂔系」字，而原書使用之字為「𧯛」，查詢教育部異體字字典確認「每𠂔系」並非「𧯛」之異體字，因此歸為錯字，以程式取代為正確字形「𧯛」。
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 1059 的「𠂔」字，Unicode 字型呈現為「𠂔」，實際上仍為同一字。

附件一、《銀雀山漢墓竹簡(壹)-孫子&孫臏》(標記)轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
296	𠂔	26	FBEC			𠂔	
297	𠂔	64	FBED			𠂔	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
298	△	5	FBEE			△	
299	𠄎	10	FBEF			𠄎	
300	◻	10	FBF0			◻	
350	𠄎	1	FC63			𠄎	
379	支	1	FCA2			支	
387	𠄎	3	FCAA			𠄎	
405	正	2	FCBC			正	
438	竟	1	FCDD	7ADE	竟		
447	𠄎	14	FCE6			𠄎	
463	𠄎	1	FCF6			𠄎	
546	券	1	FDAC	52B5	券		
551	𠄎	2	FDB1			𠄎	
556	商	10	FDB6	5547	商		
637	𠄎	5	FE48			𠄎	
658	𠄎	3	FE5D			𠄎	
679	𠄎	4	FE72			𠄎	
685	𠄎	1	FE78			𠄎	
712	开	1	FEB5			开	𠄎，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
720	夕	1	FEBD			夕	
750	𠂇	1	FEDB			𠂇	
752	𠂈	2	FEDD			𠂈	
756	𠂉	1	FEE1			𠂉	
759	𠂊	1	FEE4			𠂊	
763	𠂋	2	FEE8			𠂋	
766	𠂌	3	FEEB			𠂌	
768	𠂍	6	FEED			𠂍	
770	𠂎	3	FEED			𠂎	
771	𠂏	1	FEF0			𠂏	
777	𠂐	2	FEF6			𠂐	
780	𠂑	7	FEF9			𠂑	
784	𠂒	9	FEFD			𠂒	
915	遲	1	8EE3	905F	遲		
1059	𠂓	1	8FD6	51A3	𠂓		Unicode 字型 呈現差異。
2757	𠂔	1	9AB9	6B71	𠂔		
3791	况	1	8156	51B5	况		
3817	厠	1	8170	53A0	厠		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3833	嘍	1	81A2	5611	嘍		
3869	币	1	81C6	5E00	币		
3964	威	6	8266	70D5	威		
4029	褱	1	82C9	799D	褱		
4080	羣	1	82FC	7FA3	羣		
4184	迹	1	83C7	8FF9	迹		
4190	邳	1	83CD	90C4	邳		
4242	鬪	9	8442	9B2D	鬪		
5317	巒	1	8AE9	26248		巒𠃉𠃉	巒𠃉𠃉，異體字問題。
5534	塢	1	8C66	4CA8		工𠃉鳥	
5746		1	8DBF			每𠃉系	鯨，錯字，以程式全部取代。
5788	𠃉	49	8DE9	23AA0		𠃉	
5860	无	54	C6D3	65E0	无		