

《馬王堆漢墓帛書(壹)》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
2009/10/14 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 959 的「揄」字，Unicode 編碼是 3A49，由於 XP 並不支援，仍需使用構字式「扌𠂔翁」。

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(壹)》使用舊版造字 104 個，字頻 1092 次，這 104 個造字中，34 個可轉成 Windows XP 能顯示的字，字頻 467 次；另外 70 個字必須轉成構字式，字頻 625 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(壹)》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 4132 的「血𠂇𠂇」字，是 Unicode 字形「眾」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「眾」取代。
 2. **部份使用不同字形，手動修改**：檔案中為同一舊版造字，校對時卻發現原書使用兩個以上不同的字形，可能部份舊版造字為錯字或原書使用兩種異體字形，導致同一舊版造字轉碼時對應到兩個以上的字形；將一部份手動修改，剩餘的頻次則用程式取代成對應的 Unicode 或構字式。例：造字編號 3976 的「𠂇」字頻次為 6，5 字為錯字，依原書以程式取代為正確字形「无」，剩餘 1 字手動修改為「无」。手動修改部份請參考附件二《馬王堆漢墓帛書(壹)》手動取代表。
 3. **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 416 的「𠂇」字，而原書使用之字為「百」，查詢教育部異體字字典確認「𠂇」並非「百」之異體字，因此歸為錯字，以程式取代為正確字形「百」。

附件一、《馬王堆漢墓帛書(壹)》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
----	----	----	------	---------	-------	-----	----

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
265	𠄎	1	FBCD			𠄎	
280	欸	1	FBDC			欸	
296	𠄎	65	FBEC			𠄎	
297	𠄎	175	FBED			𠄎	
298	𠄎	20	FBEE			𠄎	
299	𠄎	47	FBEF			𠄎	
300	𠄎	47	FBF0			𠄎	
301	∞	1	FBF1			∞	
308	∞	2	FBF8			∞	
321	臣	6	FC46			臣	
327	穀	1	FC4C			穀	
330	朝	5	FC4F			朝	
336	亩	1	FC55			亩	由，錯字，以程式全部取代。
359	維	2	FC6C			維	
371	𠄎	1	FC78			𠄎	
374	𠄎	1	FC7B			𠄎	
379	支	4	FCA2			支	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
386	崇	4	FCA9			崇	
411	取	41	FCC2			取	
416	亘	1	FCC7			亘	百，錯字，以程式全部取代。
435	朮	1	FCDA			朮	
436	吉	4	FCDB			吉	
437	吋	1	FCDC			吋	
474	敷	2	FD42			敷	
481	屯	3	FD49			屯	
486	毆	14	FD4E			毆	
510	葬	1	FD66			葬	
528	县	11	FD78			县	
543	段	1	FDA9			段	
552	豕	1	FDB2			豕	
556	商	7	FDB6	5547	商		
562	彘	1	FDBC			彘	
566	尸	1	FDC0			尸	务，錯字，以程式全部取

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							代。
583	兂	6	FDD1			兂	无，錯字，以程式全部取代。 部份使用不同字形「无」，手動修改。
629	巳	6	FE40			巳	巳，錯字，以程式全部取代。
634	殷	1	FE45			殷	
637	戔	2	FE48			戔	
653	冐	6	FE58			冐	
655	以	5	FE5A			以	乚，錯字，以程式全部取代。 部份使用不同字形「衣」，手動修改。
658	疒	1	FE5D			疒	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
679	叢	2	FE72			叢	
682	惠	14	FE75			惠	
685	衤	2	FE78			衤	
691	牛	1	FE7E			牛	
711	衤	4	FEB4			衤	
717	足	3	FEBA			足	
720	夕	1	FEBD			夕	
738	心	18	FECF			心	
746	忄	2	FED7			忄	
750	冫	3	FEDB			冫	
752	𠂇	8	FEDD			𠂇	
759	广	5	FEE4			广	
763	冫	2	FEE8			冫	
766	糸	6	FEEB			糸	
768	艹	10	FEED			艹	
769	亼	1	FEEE			亼	
770	扌	8	FEEF			扌	
771	攴	6	FEF0			攴	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
772	之	2	FEF1			之	
777	宀	5	FEF6			宀	
780	彳	10	FEF9			彳	
784	亻	10	FEFD			亻	
826	癱	1	8E68	7655	癱		
959	揄	1	8F50	3A49		扌 厶 翁	
980	攣	3	8F65	72AB	攣		
1059	取	2	8FD6	51A3	取		
1532	婚	2	92D8	636A	婚		
1565	繇	2	92F9	7DDC	繇		
2125	羸	2	96B5	4369		羊 厶 文	
2297	筭	1	97C4	25BA5		𠂇 厶 析	
2349	悞	2	97F8	2263D		忄 厶 曳	
2721	吞	5	9A73	53F4	吞		
3021	忄	2	9C65			忄	
3028	忄	2	9C6C			忄	
3095	鈞	1	9CD1	91F0	鈞		
3134	囁	5	9CF8	5405	囁		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3791	况	1	8156	51B5	况		
3814	却	3	816D	5374	却		
3819	廝	1	8172	53AE	廝		
3853	弊	2	81B6	2776C		敝 _衣	
3860	尠	1	81BD	5C20	尠		
3887	息	6	81D8	60A4	息		
3900	挖	1	81E5	62D5	挖		
3930	捰	1	8244	6901	捰		
3964	威	12	8266	70D5	威		
3996	疎	1	82A8	758E	疎		
4062	齧	1	82EA	25F7F		齧	
4066	綺	1	82EE	7D5D	綺		
4080	羣	6	82FC	7FA3	羣		
4104	莅	4	8355	8385	莅		
4132		30	8371			血 _血 𠂇	眾，異體字問題。
4184	迹	3	83C7	8FF9	迹		
4193	鈎	1	83D0	920E	鈎		
4242	鬪	4	8442	9B2D	鬪		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4256	鷄	3	8450	9DC4	鷄		
4417		1	8554			𠂇 △ ^{oo} 十	𠂇，錯字，以程式全部取代。
4481	垓	3	85B6	57C9	垓		
4604	擦	1	8672	22D32		扌 △ 絜	
4693	慵	1	86ED	22931		忄 △ 龍	
4750	卅	5	8767	5344	卅		卅，錯字，以程式全部取代。
5119	博	1	89C0	613D	博		
5215		1	8A61			毘 △ 免	毘，異體字問題。
5654	籥	2	8D41	7C4B	籥		
5860	无	340	C6D3	65E0	无		

二、手動取代表說明：

《馬王堆漢墓帛書(壹)》檔案中需手動修改的情況分為以下幾種：

- (一) 《馬王堆漢墓帛書(壹)》檔案中之舊版造字，因字形部份不同而無法全部以程式取代，皆手動修改成Windows XP能顯示的字

或構字式，並製作手動取代表，請參考附件二。欄位說明如下：

1. 編號：Big5 造字空間為 6217 個，編號由 1 到 6217。
2. 造字：舊版造字字形。
3. 取代：實際使用字形與舊版造字不同時，於此欄列出。
4. 檔案原文摘錄：摘錄檔案《馬王堆漢墓帛書(壹)》中的文句段落，其中紅字底線的部分，為需要手動取代的原文。
5. 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。
6. 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(壹)》手動取代表

編號	造字	取代	檔案原文摘錄	手動取代結果	備註
583	𠄎	无	<u> </u> ，倉甚虛	<u>无</u> ，倉甚虛	
655	𠄎	衣	因民以為師。弗因𠄎 <u> </u> 也	因民以為師。弗因 𠄎 <u>衣</u> 也	

(二)《馬王堆漢墓帛書(壹)》檔案中，對於舊版造字以外發現的錯字，以手動修改，並製作錯字修改表，請參考附件三。欄位說明如下：

1. 檔案原文摘錄：摘錄檔案《馬王堆漢墓帛書(壹)》中的文句段落，其中紅字底線的部分，為需要手動取代的原文。
2. 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。
3. 備註：記錄修改相同詞句的次數或其他事項。

附件三、《馬王堆漢墓帛書(壹)》缺字外錯字修改

檔案原文摘錄	手動取代結果	備註
--------	--------	----

檔案原文摘錄	手動取代結果	備註
忖（仇）雒 会 心（目 △ 瞿）	忖（仇）雒 会 心（ <u>雙</u> ）	