

# 《禮記》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  
2010/04/23 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 4299 的「焯」字，Unicode 編碼是 242F2，由於 XP 並不支援，仍需使用構字式「火𠂇芮」。

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：

《禮記》使用舊版造字 76 個，字頻 458 次，這 76 個造字中，70 個可轉成 Windows XP 能顯示的字，字頻 448 次；另外 6 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《禮記》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：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 5497 的「裕」字，而原書使用之字為「裏」，「裕」為錯字，以程式取代為正確字形「裏」。
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 3935 的「槩」字，Unicode 字型呈現為「槩」，實際上仍為同一字。

附件一、《禮記》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
982	曾	2	8F67	80F7	曾		
1039	稟	3	8FC2	69C0	稟		
1062		4	8FD9			馬𠂔亞	駟，異體字問題。
1077	稟	1	8FE8	7A3E	稟		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1079	總	10	8FEA	7DEB	總		
1344		9	91B9			衤 𠂔 易	裼，錯字，以程式全部取代。
1431	筓	1	9251	7AFE	筓		
1893	袞	1	9548	886E	袞		
1982	調	3	95C3	8B87	調		
1999	贊	2	95D4	8CDB	贊		
2761	碇	3	9ABD	7888	碇		
2860	箴	1	9B61	7C1A	箴		
2947	聰	2	9BDA	8066	聰		
2950	肱	1	9BDD	8104	肱		
3770	个	11	8141	4E2A	个		
3787	豕	13	8152	51A2	豕		
3791	况	1	8156	51B5	况		况，異體字問題。
3829	咏	4	817C	548F	咏		
3835	噍	3	81A4	564D	噍		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3846	壻	17	81AF	58FB	壻		
3868	卮	1	81C5	5DF5	卮		
3883	徧	8	81D4	5FA7	徧		
3889	慤	7	81DA	6164	慤		
3891	慙	1	81DC	6159	慙		
3929	黎	1	8243	68C3	黎		
3930	憵	21	8244	6901	憵		
3935	槩	1	8249	3BA3	槩		Unicode 字型 呈現差異。
3949	洩	8	8257	6D96	洩		
3965	裁	2	8267	70D6	裁		
3974	牀	14	8270	7240	牀		
4011	皞	6	82B7	76A1	皞		Unicode 字型 呈現差異。
4012	鼓	5	82B8	76B7	鼓		鼓，異體字問 題。
4038	竈	6	82D2	7AC8	竈		
4063	藜	1	82EB	7CF5	藜		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4070	縹	1	82F2	7DE4	縹		
4080	羣	36	82FC	7FA3	羣		
4091	皐	8	8348	81EF	皐		Unicode 字型 呈現差異。
4097	芑	1	834E			艹 厶 巳	
4100	菘	2	8351	82FD	菘		
4107	葱	7	8358	8471	葱		
4114	藜	1	835F	8616	藜		
4130	衅	1	836F	8845	衅		
4132		90	8371			血 厶 丞	眾，異體字問 題。
4134	衛	27	8373	885E	衛		衛，異體字問 題。
4138	裘	1	8377	88A0	裘		
4139	袴	1	8378	88B4	袴		
4140	裱	1	8379	88B5	裱		
4142	裛	4	837B	8943	裛		
4208	躓	2	83DF	9767	躓		
4209	鞞	1	83E0	9789	鞞		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4212	鞞	4	83E3	97B8	鞞		
4217	鞞	2	83E8	97E0	鞞		
4238	髭	1	83FD	9AE2	髭		
4242	鬪	6	8442	9B2D	鬪		
4259	麇	6	8453	9E95	麇		
4267	鼈	8	845B	9F08	鼈		
4268	鼈	1	845C	9F17	鼈		
4299	焯	1	847B	242F2		火 ㄩ 苒	
4342	倮	1	84C8	502E	倮		
4371	倜	2	84E5	50F4	倜		
4455	劔	5	857A			僉 ㄩ 刃	
4478	淙	1	85B3	3CFF		灬 ㄩ 奈	
4512		1	85D5			口 ㄩ 它	咤，錯字，以程式全部取代。
4987	躡	1	88D9	4831		足 ㄩ 屬	
5034	弃	13	8949	5F03	弃		
5173	漱	1	89F6	6F44	漱		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5254	紂	1	8AAA	42B7		纟 厶才	
5289	慤	1	8ACD	6164	慤		
5302	昏	25	8ADA	662C	昏		
5374	盼	1	8B63	670C	盼		
5447	躡	1	8BCE	8E83	躡		
5497	𡗗	2	8C41	8873	𡗗		裏，錯字，以程式全部取代。
5591	杞	6	8CC1	233CC		木 厶 巳	杞，異體字問題。
5705	廐	2	8D74	5ED0	廐		Unicode 字型呈現差異。
5739		1	8DB8			冗 厶 古	𠄎，異體字問題。
5752	摠	6	8DC5	63D4	摠		

## 二、手動取代表說明：

《禮記》檔案中出現「●」字形，共計10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《禮記》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、《禮記》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 20	孺子●之喪	享△黃	
p. 47	聘禮以脯●	月△盍	
p. 60	祭為己●卑	亻△奈	
p. 75	●雜子女	形亻申向	
p. 89	君大夫●爪	髟△脊	
p. 102	●畝之中	亩△犬	
p. 115	不●于儀	促△言	
p. 115	章善●惡	疒△亘	
p. 116	下民卒●	疒△亘	
p. 131	蓋●有存	厂△勤	改為广△勤