

# 《孔子家語》舊版造字轉碼說明

中研院資訊所文獻處理實驗室  
中研院史語所漢籍電子文獻工作小組  
2008/3/10 陳建安 製作

一、《孔子家語》一書(孔子家語.xml)使用舊版造字 74 個，字頻 289 次，詳如附件一。這 74 個造字中，65 個可轉成 Windows XP 能顯示的字，字頻 271 次；另外 9 個字必須轉成構字式，字頻 18 次。

二、附件一的造字分析表說明如下：

甲、編號：Big5 造字空間為 6217 個，編號由 1 到 6217。

乙、造字：舊版造字

丙、字頻(txt)：造字在「.txt」文件的出現次數

丁、字頻(xml)：造字在「.xml」文件的出現次數

戊、Big5：造字的 Big5 碼

己、Unicode：造字所對應的 Unicode 碼

庚、WinXP：造字在 Windows XP 的對應字形

辛、備註凡例：

- 1、校對問題，舊版漢籍錯字，可用程式全部取代：在舊版漢籍電子文獻中即存在的錯字，因校對時的疏漏而未更正，持續留存在新版漢籍電子文獻中；若該造字的所有頻次，皆屬於錯誤使用的錯字情形，可以用程式全部取代為正確字形。如編號 5336 的「睹」字，原字為「睹」。
- 2、異體字問題：新版漢籍考量到使用者檢索及使用時的便利性，將用字原則改為除專詞等特殊情形之外，一律改用標準字呈現。如編號 4134 的「衛」係「衛」字之異體，故以「衛」字取代。
- 3、Unicode 字型呈現差異：Unicode 字型與舊漢籍造字有些微差異，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 4242 的「鬪」字，Unicode 字型呈現為「鬪」，實際上仍為同一字。

三、Unicode 目前收錄的漢字總數為 70194，分屬於三個不同區段，詳如表一。Windows XP 只支援 CJK 認同表意文字區的 20902 個字，內碼為 4E00-9FFF。所以如造字編號 1032 的「憊」字，Unicode 碼為 2039D，Windows XP 並不支援，故仍須使用構字式「𠄎1 𠄎心□」。

表一、Unicode 的字數及編碼區段

Unicode	字集子集合	新增字數	新增編碼區段	總字數	WinXP
1.1 版	CJK 認同表意文字區	20902	4E00-9FFF	20902	支援
3.0 版	CJK 認同表意文字擴充 A 區	6582	3400-4DFF	27484	不支援
3.1 版	CJK 認同表意文字擴充 B 區	42710	20000-2A6D6	70194	不支援

四、《孔子家語》中未確定字形「●」共 42 處，詳細處理內容可參考附件二的未確定字形「●」手動取代表。附件二《孔子家語》未確定字形「●」手動取代表欄位說明如下：

- 甲、檔案原文摘錄：摘錄檔案(孔子家語.xml)中包含未確定字形「●」的文句段落，其中紅字底線的部分，為需要手動取代的字形。
- 乙、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。

附件一、《孔子家語》造字分析表

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
834	煢	2	2	8E70	712D	煢		煢，異體字問題。
977	麿	1	1	8F62		麿		
980	犛	2	2	8F65	72AB	犛		
1032	憇	5	5	8FBB	2039D		𠃉 𠃉 𠃉 𠃉 𠃉	
1044	儻	1	1	8FC7	511B	儻		
1530	紉	3	3	92D6	7CFA	紉		
1885	蠶	1	1	9540	882D	蠶		
1893	袞	3	3	9548	886E	袞		
2042	䟽	1	1	9640	47FD		足 𠃉 充	䟽，異體字問題。

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
2044	踈	1	1	9642	8E08	踈		
2232		2	2	9761			𠄎厂墨卩𠄎	𠄎，異體字問題。
2828	蝸	1	1	9B41	8744	蝸		
3787	豕	9	9	8152	51A2	豕		
3789	冲	2	2	8154	51B2	冲		冲，異體字問題。
3791	况	1	1	8156	51B5	况		况，異體字問題。
3804	効	3	3	8163	52B9	効		
3814	却	2	2	816D	5374	却		
3831	喆	2	2	817E	5586	喆		
3834	嗶	1	1	81A3	5637	嗶		
3846	墻	1	1	81AF	58FB	墻		Unicode 呈現差異。
3873		1	1	81CA			广△段	廩，異體字問題。
3883	徧	3	3	81D4	5FA7	徧		
3921	皙	4	4	81FA	6673	皙		
3924	菴	5	5	81FD	671E	菴		
3930	憵	2	2	8244	6901	憵		
3974	牀	1	1	8270	7240	牀		
3981	珉	5	5	8277	73C9	珉		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
3996	疎	14	14	82A8	758E	疎		
4011	皞	6	6	82B7	76A1	皞		Unicode 呈現差異。
4035	穽	1	1	82CF	7A7D	穽		
4038	竈	5	5	82D2	7AC8	竈		
4055	粃	3	3	82E3	7C83	粃		
4063	藜	1	1	82EB	7CF5	藜		
4077	罇	3	3	82F9	7F47	罇		
4080	羣	12	12	82FC	7FA3	羣		群，異體字問題。
4091	皐	2	2	8348	81EF	皐		Unicode 呈現差異。
4105	菓	3	3	8356	83D3	菓		
4109	葛	1	1	835A	84AD	葛		
4114	藜	1	1	835F	8616	藜		
4115	藁	2	2	8360	85C1	藁		
4134	衛	49	49	8373	885E	衛		衛，異體字問題。
4140	衽	1	1	8379	88B5	衽		
4146	覩	8	8	83A1	89A9	覩		
4176	輒	2	2	83BF	8F19	輒		
4184	迹	3	3	83C7	8FF9	迹		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
4201	陲	2	2	83D8	967B	陲		
4203	隣	1	1	83DA	96A3	隣		鄰，異體字問題。
4223	顛	2	2	83EE	984B	顛		
4226	凜	2	2	83F1	98E1	凜		
4242	鬪	7	7	8442	9B2D	鬪		Unicode 呈現差異。
4250	鱸	4	4	844A	9C0C	鱸		
4267	鼈	1	1	845B	9F08	鼈		
4342	倮	1	1	84C8	502E	倮		
4460	勅	2	2	85A1	52D1	勅		
4528	嶼	1	1	85E5	5D60	嶼		
4643	徃	42	42	86BB	5F83	徃		
4762	旃	1	1	8773	3ACB		𠂇△冉	旃，異體字問題。
4768	妬	2	2	8779	59AC	妬		妒，異體字問題。
4792		3	3	87B3			𠂇△虛	虞，異體字問題。
4904	尅	5	5	8864	5C05	尅		
5147	恠	1	1	89DC	6060	恠		
5192	隙	1	1	8A4A	969F	隙		
5225	邨	1	1	8A6B	460F		血△卩	

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
5249	蝻	1	1	8AA5	873D	蝻		
5289	慤	10	10	8ACD	6164	慤		
5336	睹	1	1	8AFC	668F	睹		睹，校對問題，舊版漢籍錯字，可用程式全部取代。
5389	竿	3	3	8B72	7B07	竿		
5578	弛	2	2	8CB4	38AE		弓𠂔𠂔	弛，異體字問題。
5588	恠	2	2	8CBE	6061	恠		
5703	𪗇	3	3	8D72	9F53	𪗇		
5705	廐	1	1	8D74	5ED0	廐		廐，異體字問題。
5706	敍	1	1	8D75	654D	敍		敍，異體字問題。
5709	蟀	2	2	8D78	872F	蟀		
5853		2	2	C6CC				𠂔，校對問題，舊版漢籍錯字，可用程式全部取代。

附件二、《孔子家語》未確定字形「●」手動取代表

檔案原文摘錄	手動取代結果
以兵鼓●劫定公	以兵鼓諺劫定公
雷鼓曰●	雷鼓曰諺
寡人且愚●	寡人且愚冥
言恚愚●暗也	言恚愚冥暗也
古人以●	古人以尤合旨
醴●齊也	醴盍齊也
三曰●齊	三曰盍齊
至于中●	至于中冥

檔案原文摘錄	手動取代結果
中日中●映中	中日中冥映中
聞諸惜其腐●	聞諸惜其腐食△念
彼有●	彼有竇
●折者不終	罍△寸折者不終
●則折矣	罍△寸則折矣
●紕於此	禾△忽紕於此
引●也	引弓△几也
則民引●	則民引弓△几
●其志而無不容	弓△几其志而無不容
齟●者九竅	齟齒△爵者九竅
●疾雀反	齒△爵疾雀反
●土之人醜	土△毛土之人醜
●耗字也	土△毛耗字也
●土麤疎	土△毛土麤疎
降脫●	降脫履
而鳥不●	而鳥不羽△氏
而獸不●	而獸不彳△越
●●飛走之貌也	羽△氏彳△越飛走之貌也
●況必反	羽△氏況必反
不郊●以●	不郊冥以冥
若復郊則●	若復郊則冥
郊●者也	郊冥者也
魚之大者名為●	魚之大者名為魚△壽
●宜為鱣	魚△壽宜為鱣
徑五寸而●組綬	徑五寸而糸△其組綬
●雜色組綬	糸△其雜色組綬
●巨箕反	糸△其巨箕反
馬●，封	馬髟△葛，封
●力葉反	髟△葛力葉反
●明欲毀鄉校	鬲△夔明欲毀鄉校
●明然明	鬲△夔明然明
毀宗而●行也	毀宗而足△葛行也