

# 《清實錄》(世宗)舊版造字轉碼說明

中研院資訊所文獻處理實驗室  
2007/1/25 趙苑曲

一、《清實錄》世宗實錄部分(世宗.xml)使用舊版造字 106 個，字頻 470 次，詳如附件一。這 106 個造字中，71 個可轉成 Windows XP 可顯示的字，字頻 398 次；另外 35 個字必須轉成構字式，字頻 72 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：造字在文件中的出現次數

丁、Big5：造字的 Big5 碼

戊、Unicode：造字所對應的 Unicode 碼

己、WinXP：造字在 Windows XP 的對應字形

庚、構字式：Windows XP 若無對應字形，則改採用構字式。

辛、備註凡例：

- 1、錯字：與原書字形不符，用正確的字取代原造字，如編號 5189 的「宄」字，原字應為「穴」。
- 2、取代：用另一個字取代原造字，如編號 4720 的「土已」字，改用「圮」字取代。
- 3、重碼造字，手動取代：同時使用舊漢籍造字及構字式，導致一個字碼代表兩種字形的重碼問題，必須一一比對原書字形，以判斷應取何者之字形，取舊漢籍造字字形者再以手動取代。如編號 5797，舊漢籍造字字形為「灑」，構字式符號為「」，一一比對原書字形為「灑」者再以手動取代。
- 4、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 3975 的「牕」字，Unicode 字型呈現為「牕」，實際上仍為同一字。
- 5、待造字：舊版漢籍造字字形為電腦缺字，且漢字構形資料庫尚未收錄此字形，等待補造字中，所以「造字」欄空白無法看到字形。如編號 43 的「口六」。

部分「造字」欄空白者，因已使用其他字形取代，故不另外補造字。如編號 4720 的「土~~厶~~已」字，已由「圪」取代，故不另造「土~~厶~~已」。

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。目前 Windows XP 只支援 CJK 認同表意文字區的 20902 個字，內碼為 4E00-9FFF。所以造字編號 52 的「繚」字，Unicode 編碼為 4307，由於 Windows XP 並不支援，仍須改用構字式「纛~~厶~~索」。

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、 造字分析表「備註」欄標示「手動取代」者，詳細取代內容可參考附件二的手動取代表。附件二的手動取代表說明如下：

- 甲、 編號：Big5 造字空間為 6217 個，編號由 1 到 6217。此處編號同附件一的「造字分析表」。
- 乙、 造字：舊版造字字形
- 丙、 部件：構字式字形，通常當部件使用。
- 丁、 檔案原文摘錄：摘錄檔案(世宗.xml)中包含該造字的文句段落，其中紅字底線的部分，為需要手動取代的原文。
- 戊、 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、《清實錄》「世宗.xml」造字分析表

編號	字頻	造字	Big5	Unicode	WinXP	構字式	備註
43	1		FA6A			口 厶 六	待造字。
52	1	縲	FA73	4307		糸 厶 索	
543	1	段	FDA9			段	
556	1	商	FDB6	5547	商		
624	1	曷	FDF A			曷	
986	1	猥	8F6B	7302	猥		
1000	1		8F79			王 厶 巳	待造字。
1067	1	碁	8FDE	68CA	碁		
1073	1		8FE4			算 厶 么	待造字。
1175	1		90AD			土 厶 郎	待造字。
1255	3	緯	90FD	7DC8	緯		
1355	1		91C4			衤 厶 义	待造字。
1530	1	紉	92D6	7CFA	紉		
1551	1	紘	92EB	7D8B	紘		
1606	1		9363			汜 厶 丙	待造字。

編號	字類	造字	Big5	Unicode	WinXP	構字式	備註
1622	1	艘	9373	8267	艘		
1663	1		93BE			氵 厶 戎	待造字。
1677	1	邕	93CC	5C9C	邕		
1701	1	衢	93E4			雨 厶 衢	
2021	2	趋	95EA	8D82	趋		
2244	1		976D			ノ 厶 虫	待造字。
2415	2	鬪	987B	95D8	鬪		
2572	7	銜	997B	8858	銜		
2618	1	缶	99CB	26222		缶	
2955	1	節	9BE2	83AD	節		
2994	2	揀	9C4A	6352	揀		
3225	1	磴	9DB6	25562		石 厶 曹	
3793	9	減	8158	51CF	減		
3811	1	貳	816A	8CAE	貳		
3816	2	厓	816F	5393	厓		
3817	1	廁	8170	53A0	廁		

編號	字頻	造字	Big5	Unicode	WinXP	構字式	備註
3835	1	噍	81A4	564D	噍		
3838	20	鋒	81A7	2899F		金 ㄩ 夂	
3840	13	坂	81A9	5742	坂		
3843	1	堦	81AC	5826	堦		
3844	1	城	81AD	5884	城		
3916	1	敷	81F5	656D	敷		
3924	1	朥	81FD	671E	朥		
3927	1	柏	8241	6855	柏		
3932	1	櫻	8246	6936	櫻		
3942	1	毡	8250	6BE1	毡		
3945	4		8253			氵 ㄩ 义	待造字。
3946	1	汹	8254	6C79	汹		
3959	1	潛	8261	6FF3	潛		
3967	1	炳	8269	712B	炳		
3975	3	牕	8271	7255	牕		Unicode 字型呈現差異。
3981	6	珉	8277	73C9	珉		

編號	字頻	造字	Big5	Unicode	WinXP	構字式	備註
3988	23	璦	827E	7482	璦		
3989	4	璦	82A1	74C8	璦		
3990	1	甄	82A2	750E	甄		
4003	2	塚	82AF	585C	塚		
4027	24	裨	82C7	7986	裨		
4031	1	杭	82CB	79D4	杭		
4036	8	窰	82D0	7AB0	窰		
4044	1	筋	82D8	7B6F	筋		
4078	1		82FA			巳厶升	待造字。
4096	1	蒨	834D	84E2	蒨		
4097	1	芭	834E			艹厶巳	
4118	1	藜	8363	8602	藜		
4136	1	衰	8375	887A	衰		
4144	2	霸	837D	8987	霸		Unicode 字型呈現差異。
4198	1	鑛	83D5	945B	鑛		
4216	2	鞮	83E7	97C8	鞮		

編號	字頻	造字	Big5	Unicode	WinXP	構字式	備註
4234	5	羸	83F9	9A58	羸		
4259	1	麇	8453	9E95	麇		
4308	9	劫	84A6	5227	劫		
4400	3	堯	8543	5160	堯		
4442	2	潛	856D	6F98	潛		
4630	1		86AE			形匕示欠	待造字。
4657	1	竣	86C9	57C8	竣		
4679	2	塏	86DF	5818	塏		
4720	4		8749			土厶已	圮，取代。
4746	5	齧	8763	28932		形牙女厘	
4762	1	旃	8773	3ACB		舛厶冉	
4816	1	亻	87CB			亻	
4845	1		87E8			女厶間	待造字。
4908	1	届	8868	5C4A	届		
4933	19	崗	88A3	5CDD	崗		
4969	1	嶋	88C7	5D8B	嶋		

編號	字頻	造字	Big5	Unicode	WinXP	構字式	備註
4997	3	裛	88E3	2768C		衣△壯	
4999	1	巖	88E5	5DD7	巖		
5005	1	幫	88EB	5E2E	幫		
5043	1	彙	8952	5F4D	彙		
5150	1	伏	89DF	3432		亻△大	
5165	1		89EE			晶△宜	待造字。
5189	1	宀	8A47	5B82	宀		宀，錯字。
5298	1	昨	8AD6	65FF	昨		
5395	7	規	8B78	4093		矢△見	
5451	2	玨	8BD2	7398	玨		
5493	1	辵	8BFC			辵	
5516	1		8C54			〇〇束	待造字。
5549	2	藁	8C75	85F3	藁		
5562	2	稷	8CA4	7A09	稷		
5567	1	嶼	8CA9	3817		山△奧	
5583	2	殮	8CB9	98F1	殮		

編號	字頻	造字	Big5	Unicode	WinXP	構字式	備註
5632	1	替	8CEA			𠄎△日	
5667	1	燻	8D4E			𠄎火日欠丨欠𠄎	
5707	1	棧	8D76	6830	棧		
5729	2	滌	8DAE	6F34	滌		
5737	1		8DB6			𠄎灬 𠄎 𠄎	待造字。
5741	13	籽	8DBA	79C4	籽		
5742	1	澁	8DBB	6F81	澁		
5749	1		8DC2			片△庸	待造字。
5753	4	潛	8DC6	6F5C	潛		
5764	1	籓	8DD1	7C5E	籓		
5797	177	灑	8DF2	7054	灑		重碼造字，手動取代。

附件二、《清實錄》「世宗.xml」手動取代表

編號	造字	部件	檔案原文摘錄	手動取代結果
5797	?		△灑堆等	<a href="#">灑</a> 灑堆等