

《清秘述聞》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
2008/5/15 陳建安 製作

一、《清秘述聞》一書(清秘述聞.xml)使用舊版造字 167 個，字頻 1693 次，詳如附件一。這 167 個造字中，131 個可轉成 Windows XP 能顯示的字，字頻 1514 次；另外 36 個字必須轉成構字式，字頻 179 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、構字式：Windows XP 若無對應字形，則改採用構字式

壬、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 4756 的「呂」字，原字為「臣」。
- 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 4132 的「血𠂔𠂔」係「眾」字之異體，故以「眾」字取代。又編號 1641 的「胆」字，係「膽」字之簡體字，亦以標準字的「膽」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 3935 的「槩」字，Unicode 字型呈現為「槩」，實際上仍為同一字。
- 4、待造字：Unicode 及漢字構形資料庫皆未收錄的舊漢籍造字，正在等待補造字中，所以「造字」欄空白無法看到字形。如編號 2707 的「豈𠂔支」。

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 27 的「琫」字，Unicode 編碼為 24A13，由於 Windows XP 並不支援，仍須使用構字式「王𠄎茶」。

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

附件一、《清秘述聞》造字分析表

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
27	琫	2	2	FA5A	24A13		王𠄎茶	
38		2	2	FA65			王𠄎已	玳，異體字問題。
296	𠄎	8	8	FBEC			𠄎	
297	𠄎	18	18	FBED			𠄎	
298	𠄎	2	2	FBEE			𠄎	
752	𠄎	1	1	FEDD			𠄎	
770	𠄎	1	1	FEEF			𠄎	
784	𠄎	1	1	FEFD			𠄎	
788	炣	1	1	8E42	708C	炣		
821	脊	4	4	8E63	7718	脊		
846	焜	2	2	8E7C	713B	焜		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
893	峩	17	17	8ECD	5CE9	峩		
908	爆	2	2	8EDC	71DD	爆		
938	燮	3	3	8EFA	7215	燮		
1010	珏	6	6	8FA5	73CF	珏		
1092	瑞	1	1	8FF7	747A	瑞		
1218	焯	1	1	90D8	24423		火△鼎	
1347	岌	1	1	91BC	21E05		形山弓女□	
1350	褚	3	3	91BF	7987	褚		褚，校對問題，舊版漢籍錯字，可用程式全部取代。
1381	穉	1	1	91DE	7A3A	穉		
1416	埃	4	4	9242	7AE2	埃		
1551	紘	2	2	92EB	7D8B	紘		
1584	肩	1	1	934D			肩	
1634	忬	2	2	93A1	5FEC	忬		
1641	胆	1	1	93A8	80C6	胆		膽，異體字問題。
1893	袞	3	3	9548	886E	袞		袞，異體字問題。
1959	墳	2	2	95AC	213B4		土△貢	
1966	翮	3	3	95B3	4397		喬△羽	

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
1981	讐	1	1	95C2	8B90	讐		
2046	踪	1	1	9644	8E2A	踪		
2087	霽	8	8	966D	291A9		雨△澍	
2127	漣	4	4	96B7	3D8C		氵△聯	
2213	鉷	1	1	974E	9277	鉷		
2274	鎋	1	1	97AD	9348	鎋		
2345	鑽	6	6	97F4	9404	鑽		
2359	鑄	7	7	9843	9408	鑄		
2361	鎚	1	1	9845	93F3	鎚		
2363	翮	6	6	9847	4396		曾△羽	
2472	靄	1	1	98D6	9741	靄		
2565	靄	4	4	9974	9714	靄		
2664	焯	2	2	99F9	3DC6		火△阜	
2707		1	1	9A65			豈△支	待造字。
2741		1	1	9AA9			厯△心	歷，異體字問題。
2793	鋤	1	1	9ADD	9378	鋤		
2811		1	1	9AEF			亞△田	留，異體字問題。
2906	竽	1	1	9BB1	659A	竽		竽，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3096	歲	3	3	9CD2	5D57	歲		歲，異體字問題。
3242	霽	1	1	9DC7	9731	霽		
3770	个	2	2	8141	4E2A	个		
3780	儁	4	4	814B	5101	儁		
3789	冲	8	8	8154	51B2	冲		冲，異體字問題。
3791	况	3	3	8156	51B5	况		况，異體字問題。
3813	邛	1	1	816C	536D	邛		
3814	却	2	2	816D	5374	却		卻，異體字問題。
3816	厓	13	13	816F	5393	厓		
3820	廈	3	3	8173	53A6	廈		廈，異體字問題。
3824	叙	1	1	8177	53D9	叙		敘，異體字問題。
3826	叶	6	6	8179	53F6	叶		
3829	咏	1	1	817C	548F	咏		
3831	喆	8	8	817E	5586	喆		
3842	堃	12	12	81AB	5803	堃		
3843	堦	1	1	81AC	5826	堦		
3860	尠	1	1	81BD	5C20	尠		
3864	峯	36	36	81C1	5CEF	峯		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3866	崐	1	1	81C3	5D10	崐		
3876	廻	1	1	81CD	5EFB	廻		迴，異體字問題。
3883	徧	2	2	81D4	5FA7	徧		
3888	憇	6	6	81D9	60B3	憇		
3891	慙	1	1	81DC	6159	慙		慚，異體字問題。
3895	淑	1	1	81E0	6E86	淑		淑，異體字問題。
3897	慙	1	1	81E2	615C	慙		
3911	攡	1	1	81F0	651F	攡		
3915	斨	2	2	8459	6585	斨		
3916	斨	2	2	81F5	656D	斨		
3921	皙	4	4	81FA	6673	皙		皙，異體字問題。
3923		1	1	81FC			晶△宜	疊，異體字問題。
3931	椀	1	1	8245	6900	椀		
3933	桀	3	3	8247	6998	桀		
3935	槩	1	1	8249	69E9	槩		Unicode 呈現差異。
3948	沝	2	2	8256	6CAD	沝		
3949	涖	1	1	8257	6D96	涖		
3966	烟	5	5	8268	70DF	烟		煙，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3969	煊	2	2	826B	714A	煊		
3975	牕	1	1	8271	7255	牕		Unicode 呈現差異。
3987	穀	2	2	827D	7474	穀		
3988	璦	7	7	827E	7482	璦		
3991	攜	1	1	82A3	64D5	攜		攜，異體字問題。
3994	畊	8	8	82A6	754A	畊		
3996	疎	6	6	82A8	758E	疎		
4008	癯	2	2	82B4	766F	癯		
4010	臯	1	1	82B6	7690	臯		
4034	穉	5	5	82CE	7A49	穉		
4039	竚	1	1	82D3	7ADA	竚		
4068	綫	1	1	82F0	7DAB	綫		線，異體字問題。
4080	羣	41	41	82FC	7FA3	羣		
4085	耻	2	2	8342	803B	耻		
4097	芑	10	10	834E			艹 厶 巳	芑，異體字問題。
4104	莅	2	2	8355	8385	莅		
4107	葱	1	1	8358	8471	葱		
4108	莽	9	9	8359	844A	莽		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4111	蕤	4	4	835C	8493	蕤		
4118	藜	1	1	8363	8602	藜		
4129	虬	1	1	836E	866C	虬		
4132		34	34	8371			血𠄎𠄎	眾，異體字問題。
4134	衛	49	49	8373	885E	衛		衛，異體字問題。
4138	裘	1	1	8377	88A0	裘		
4147	蓋	1	1	83A2	8480	蓋		
4176	輒	1	1	83BF	8F19	輒		輒，異體字問題。
4184	迹	4	4	83C7	8FF9	迹		
4186	遡	1	1	83C9	9061	遡		
4188	邨	2	2	83CB	90A8	邨		
4194	鈇	4	4	83D1	9241	鈇		
4195	鉢	1	1	83D2	9262	鉢		
4202	鴟	7	7	83D9	9696	鴟		
4220	穎	13	13	83EB	9834	穎		穎，異體字問題。
4237	髡	1	1	83FC	9AE0	髡		髡，異體字問題。
4254	鷓	5	5	844E	9D7E	鷓		
4256	鷄	1	1	8450	9DC4	鷄		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4258	麇	7	7	8452	9E90	麇		
4264	麇	3	3	8458	9EC1	麇		
4267	鼈	1	1	845B	9F08	鼈		
4369	晫	3	3	84E3	666B	晫		
4409	冏	1	1	854C	518B	冏		
4480	岫	1	1	85B5	21D84		山⊕包	
4507	吳	657	657	85D0	5434	吳		吳，異體字問題。
4623	玩	1	1	86A7	5743	玩		
4629	埶	1	1	86AD	577F	埶		
4634		1	1	86B2			弓⊕女	癸，校對問題，舊版漢籍錯字，可用程式全部取代。
4750	卅	3	3	8767			卅	卅，校對問題，舊版漢籍錯字，可用程式全部取代。
4756	呂	1	1	876D	382F		呂	臣，校對問題，舊版漢籍錯字，可用程式全部取代。
4760	覲	1	1	8771	89B8	覲		
4902	兗	2	2	8862	5156	兗		兗，異體字問題。
4940	韵	2	2	88AA	97F5	韵		
4950	嶮	1	1	88B4	5D18	嶮		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4957	崑	2	2	88BB	5D53	崑		
5013	汜	1	1	88F3			汜	
5019	畧	3	3	88F9			畧	略，異體字問題。
5126	暮	3	3	89C7	8B29	暮		
5154	璿	1	1	89E3	7468	璿		
5164	肇	3	3	89ED	8088	肇		
5173	漱	12	12	89F6	6F44	漱		漱，異體字問題。
5218		33	33	8A64			𠄎→覓、𠄎	寬，異體字問題。
5290	携	2	2	8ACE	643A	携		攜，異體字問題。
5292	粵	2	2	8AD0	65C9	粵		
5305	晒	3	3	8ADD	661E	晒		
5318	蔭	2	2	8AEA	4516		艹 厶 陵	
5338	暎	1	1	8AFE	668E	暎		
5410	柎	1	1	8BA9	67AC	柎		
5418	邨	1	1	8BB1	28697		冉 厶 卩	
5439	栢	4	4	8BC6	6822	栢		
5451	玳	1	1	8BD2	7398	玳		
5470	棟	3	3	8BE5	68C5	棟		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5484	楛	3	3	8BF3	6973	楛		
5493	之	2	2	8BFC			之	
5503		1	1	8C47			木△茶	待造字。
5549	藁	2	2	8C75	85F3	藁		
5553	檉	4	4	8C79	6A52	檉		
5556	隸	343	343	8C7C	96B7	隸		隸，異體字問題。
5560	槁	4	4	8CA2	6A99	槁		
5571	榦	1	1	8CAD	6A45	榦		
5585	寓	2	2	8CBB	3762		宀△禹	
5586	獻	1	1	8CBC	6B57	獻		
5591	杞	17	17	8CC1	233CC		木△巳	杞，異體字問題。
5629	淦	2	2	8CE7	6D24	淦		
5706	敘	15	15	8D75	654D	敘		敘，異體字問題。
5757	滢	12	12	8DCA	6F0B	滢		
5773	滢	2	2	8DDA	6FDA	滢		