《清實錄》（文宗）舊版造字轉碼說明

中研院資訊所文獻處理實驗室
2007/3/14 趙苑曲

一、 《清實錄》文宗實錄部分(文宗.xml)使用舊版造字 172 個，字頻 5720 次，詳如附件一。這 172 個造字中，121 個可轉成 Windows XP 可顯示的字，字頻 3688 次；另外 51 個字必須轉成構字式，字頻 2032 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 1266 的「硋」字，原字應為「雇」。

2、取代：用另一個字取代原造字，如編號 5174 的「[形]刀兄、[⊡]」字，改用「兔」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 3778 的「併」字，Unicode 字型呈現為「併」，實際上仍為同一字。

4、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 1147 的「[形]穴矣欠[⊡]」。

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 2991 的「鹿[△]各」字，已由「麋」取代，故不另造「鹿[△]各」。

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 1306 的「㵎」字，Unicode

編碼為 3D6F，由於 Windows XP 並不支援，仍須改用構字式「氵
⿰鳳」。

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、《清實錄》「文宗.xml」造字分析表

編號	造字	字頻	Big5	Unicode	WinXP	構字式	備註
370	畺	2	FC77			畺	
379	攴	1	FCA2			攴	
400	負	53	FCB7			負	
465	邑	9	FCF8			邑	
493	胃	19	FD55	5190	胃		
559	耒	1	FDB9			耒	
675	食	1	FE6E			食	
685	衣	1	FE78			衣	
772	辶	1	FEF1			辶	
782	幺	2	FEFB			幺	

編號	造字	字頻	Big5	Unicode	WinXP	構字式	備註
853	焴	1	8EA5	7134	焴		
949	枏	1	8F46	678F	枏		
1039	稟	3	8FC2	69C0	稟		
1067	棊	1	8FDE	68CA	棊		
1111	瓃	1	904B	749D	瓃		
1147		1	906F			形穴矣欠⊡	待造字。
1184	狛	151	90B6	72D3	狛		
1186	堟	3	90B8	5845	堟		
1266	硋	1	9149	784B	硋		雇，錯字。
1306	滇	55	9171	3D6F		氵⧋鳳	
1313	砸	1	9178	25441		石⧋巴	
1379		1	91DC			言⧋恣	待造字。
1407		1	91F8			穴⧊雋	待造字。
1408	芳	1	91F9	827B	芳		
1513	粮	2	92C5	7CAE	粮		
1760	浸	1	9460	6DAD	浸		

編號	造字	字頻	Big5	Unicode	WinXP	構字式	備註
1824	墶	1	94C2	3669		土⿰奈	
1885	蠭	8	9540	882D	蠭		
1897	祕	1	954C	8890	祕		
1903	峿	4	9552	5CFC	峿		
1906	褒	11	9555	88E6	褒		
1951	味	1	95A4	20C0C		口⿰末	
1956	諜	118	95A9	8AB4	諜		
1959	墳	1	95AC	213B4		土⿰貢	
1964	謞	31	95B1	8B5E	謞		
1966	翻	73	95B3	4397		喬⿰羽	
1969	昇	1	95B6	23179		日⿱亠⿱廾	
1970	興	3	95B7	22900		興⿱心	
1993	艣	3	95CE	26A55		舟⿰虎	
2050	蹏	1	9648	8E4F	蹏		
2058	啒	3	9650	20D31		口⿰固	
2112	舾	2	96A8	4469		舟⿰古	

編號	造字	字頻	Big5	Unicode	WinXP	構字式	備註
2131	弼	34	96BB	5F3B	弼		
2230	鋥	1	975F	92E5	鋥		
2248	鑫	4	9771	9331	鑫		
2346	鐧	1	97F5	9427	鐧		間，錯字。
2592	贓	1	99B1	8D1C	贓		
2671	畐	1	9A41	7550	畐		
2701		1	9A5F			忄⿰興	待造字。
2756	款	1	9AB8	6B40	款		
2793	鍋	5	9ADD	9378	鍋		
2838	曇	1284	9B4B			晶⿱宜	
2874	饉	1	9B6F	9939	饉		
2900	彣	1	9BAB	5F63	彣		
2989		1	9C45			鳥⿰乚	待造字。
2991		1	9C47			鹿⿱各	麐，錯字。
3005	糍	2	9C55	7CCD	糍		
3159	躪	1	9D52	28206		足⿰蘭	躪，錯字。

編號	造字	字頻	Big5	Unicode	WinXP	構字式	備註
3177	罍	1	9D64	7F4E	罍		
3199	骰	27	9D7A	9ABD	骰		
3213	塔	1	9DAA	366E		土⿰答	
3216	嘁	8	9DAD	55B4	嘁		
3225	磳	2	9DB6	25562		石⿰曹	
3778	併	9	8149	5002	併		Unicode 字型呈現差異。
3793	减	1	8158	51CF	减		
3801	剗	3	8160	5257	剗		
3806	勠	54	8165	52E0	勠		
3809	匳	1	8168	5333	匳		
3816	厓	5	816F	5393	厓		
3817	厠	1	8170	53A0	厠		
3835	噍	1	81A4	564D	噍		
3840	坂	2	81A9	5742	坂		
3842	堃	73	81AB	5803	堃		
3849	奬	13	81B2	596C	奬		

編號	造字	字頻	Big5	Unicode	WinXP	構字式	備註
3866	崐	38	81C3	5D10	崐		
3869	帀	31	81C6	5E00	帀		
3874		15	81CB			忄⿰禀	待造字。
3897	慾	1	81E2	615C	慾		
3914	攸	3	81F3	6542	攸		
3916	敭	23	81F5	656D	敭		
3945		9	8253			氵⿰义	待造字。
3955	澗	5	825D	3D4E		氵⿰閒	
3959	潛	134	8261	6FF3	潛		
3990	甄	27	82A2	750E	甄		
3995	晦	1	82A7	756E	晦		
4017	蠏	1	82BD	880F	蠏		
4027	裨	3	82C7	7986	裨		
4031	秔	1	82CB	79D4	秔		
4035	穽	1	82CF	7A7D	穽		
4036	窰	25	82D0	7AB0	窰		

編號	造字	字頻	Big5	Unicode	WinXP	構字式	備註
4044	筋	21	82D8	7B6F	筋		
4049	篽	89	82DD	7BFA	篽		
4051	簬	9	82DF	7C35	簬		
4061	糦	2	82E9	7CE6	糦		
4073	纘	1	82F5	7E6E	纘		
4111	莼	4	835C	8493	莼		
4135	衚	1	8374	885C	衚		
4136	衰	1	8375	887A	衰		
4142	褒	12	837B	8943	褒		
4149	覷	1	83A4	89BC	覷		
4195	鉢	1	83D2	9262	鉢		
4198	鑛	4	83D5	945B	鑛		
4202	隖	1	83D9	9696	隖		
4234	贏	61	83F9	9A58	贏		
4243	鬴	1	8443	9B34	鬴		
4254	鶡	10	844E	9D7E	鶡		

編號	造字	字頻	Big5	Unicode	WinXP	構字式	備註
4259	麕	87	8453	9E95	麕		
4266	鼉	1	845A	9F02	鼉		
4271	齩	2	845F	9F69	齩		
4274		4	8462			扌⿰爻	待造字。
4308	刼	34	84A6	5227	刼		
4348	偬	2	84CE	506C	偬		
4370	傯	251	84E4	50E1	傯		
4378	隲	2	84EC	28F65		陟⿺馬	
4400	兟	1	8543	5160	兟		
4465	磬	1	85A6	3C88		殼⿱召	
4483	嵇	1	85B8	5D70	嵇		
4533	唫	47	85EA	552B	唫		
4547		2	85F8			口⿰美	待造字。
4554	嘧	2	8640	55C1	嘧		
4566	磵	1	864C	78F5	磵		
4616	壳	2	867E	58F3	壳		

編號	造字	字頻	Big5	Unicode	WinXP	構字式	備註
4622	該	1	86A6	8A43	該		
4657	埈	2	86C9	57C8	埈		
4712	灠	7	8741	7059	灠		
4766	捡	1336	8777	6366	捡		
4790	臁	2	87B1	81C1	臁		
4800	嚫	2	87BB	5654	嚫		
4816	亻	1	87CB			亻	
4845		1	87E8			女⿰間	待造字。
4888	荔	3	8854	8318	荔		
4932	嵗	1	88A2	21DEB		山⿱成	
4933	峝	7	88A3	5CDD	峝		
4955	龜	3	88B9	9F9C		龜	
5003		1	88E9			⿸形女穴⿰人又⿴⿱一口	待造字。
5043	彏	1	8952	5F4D	彏		
5071	岔	1	896E	5FE9	岔		
5083	瑈	8	897A	7419	瑈		

編號	造字	字頻	Big5	Unicode	WinXP	構字式	備註
5086		1	897D			冗⿰卓	冈⿰卓，取代。
5159	鵠	4	89E8	9D60	鵠		
5170	戺	2	89F3	623A	戺		
5173	漱	4	89F6	6F44	漱		
5174		8	89F7			⿰刀兄丶⊡	兔，取代。
5189	穴	7	8A47	5B82	穴		
5195		1	8A4D			⿰木亩且⊡	待造字。
5220	毌	1	8A66	6BCC	毌		毋，錯字。
5223		384	8A69			⿰扌穴人夕⊡	待造字。
5229	撐	1	8A6F	6490	撐		
5230		6	8A70			⿰土亩且⊡	待造字。
5305	昞	1	8ADD	661E	昞		
5306		69	8ADE			竹⿰閒	待造字。
5311	晄	1	8AE3	6644	晄		
5344	敫	2	8B45	55F8	敫		
5368	蠱	1	8B5D	66E7	蠱		

編號	造字	字頻	Big5	Unicode	WinXP	構字式	備註
5389	笮	1	8B72	7B07	笮		
5398	湹	614	8B7B	6E7B	湹		
5410	枏	3	8BA9	67AC	枏		
5445	梘	3	8BCC	6898	梘		
5463	峇	6	8BDE	65F9	峇		
5491	橫	2	8BFA	6967	橫		
5516		3	8C54			○○束	待造字。
5522	嶪	1	8C5A	5DAB	嶪		
5549	藁	85	8C75	85F3	藁		
5551	賣	1	8C77	8CF7	賣		
5562	稑	14	8CA4	7A09	稑		
5582	鞵	2	8CB8	97B5	鞵		
5654	籥	3	8D41	7C4B	籥		
5718	澤	1	8DA3	3D62		氵⧋雷	
5739		5	8DB8			冈⧊古	待造字。
5753	潛	5	8DC6	6F5C	潛		

編號	造字	字頻	Big5	Unicode	WinXP	構字式	備註
5776	澛	1	8DDD	7002	澛		
5854	幺	4	C6CD	5E7A	幺		