《清實錄》（聖祖）舊版造字轉碼說明

中研院資訊所文獻處理實驗室
2007/1/22 趙苑曲

一、 《清實錄》聖祖實錄部分(聖祖.xml)使用舊版造字 134 個，字頻 531 次，詳如附件一。這 134 個造字中，82 個可轉成 Windows XP 可顯示的字，字頻 401 次；另外 52 個字必須轉成構字式，字頻 130 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 2989 的「鳥⿰⿱人人乚」字，原字應為「鳦」。

2、取代：用另一個字取代原造字，如編號 4275 的「义」字，改用「義」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 1122 的「禨」字，Unicode 字型呈現為「禨」，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 1000 的「王⿰⿱人人巳」。

> 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 2989 的「鳥⿰⿱人人乚」字，已由「鳦」取代，故不另造「鳥⿰⿱人人乚」。

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「瑹」字，Unicode 編碼為 24A13，由於 Windows XP 並不支援，仍須改用構字式「王

⿱⿰丷丷茶」。

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、《清實錄》「聖祖.xml」造字分析表

編號	字碼	字頻	Big5	Unicode	WinXP	構字式	備註
27	琛	4	FA5A	24A13		王⿱⿰丷丷茶	
543	叚	1	FDA9			叚	
624	啚	2	FDFA			啚	
687	犭	1	FE7A			犭	
742	勹	6	FED3			勹	
797	抳	1	8E4B	39D6		扌⿱⿰丷丷尼	
1000		2	8F79			王⿱⿰丷丷巳	待造字。
1052	琌	1	8FCF	741C	琌		
1122	禐	1	9056	798C	禐		Unicode 字型呈現差異。
1138	珺	2	9066	743E	珺		
1147		1	906F			⿴形穴矣欠⿴囗丶	待造字。

編號	字碼	字頻	Big5	Unicode	WinXP	構字式	備註
1218	熚	3	90D8	24423		火⿰鼎	
1325		1	91A6			礻⿰今	待造字。
1381	穉	1	91DE	7A3A	穉		
1477	唏	1	92A1	35ED		口⿰悉	
1478		1	92A2			竹⿱欺	待造字。
1530	糺	1	92D6	7CFA	糺		
1551	綋	1	92EB	7D8B	綋		
1606		1	9363			氵⿰丙	待造字。
1736	蓱	1	9448	8453	蓱		
1885	蠭	1	9540	882D	蠭		
1978	讚	1	95BF	8B83	讚		
2112	舫	2	96A8	4469		舟⿰古	
2254	嗤	2	9777	20DE3		口⿰虫	
2505	飜	1	98F7	98DC	飜		
2572	銜	1	997B	8858	銜		
2601	岏	1	99BA	21D7B		山⿰亢	

編號	字碼	字頻	Big5	Unicode	WinXP	構字式	備註
2609	鰧	2	99C2	9C4F	鰧		Unicode 字型呈現差異。
2642	汊	2	99E3	3CBC		氵⿰乂	
2704		3	9A62			髟⿱今了	待造字。
2719	函	1	9A71	51FE	函		
2838	曇	1	9B4B			晶⿱今宜	
2842	翍	2	9B4F	7FC4	翍		
2955	茚	1	9BE2	83AD	茚		
2989		1	9C45			鳥⿰乂乚	鳦，錯字。
3029	鏺	6	9C6D	941D	鏺		
3098	揸	2	9CD4	63F8	揸		
3788	慤	4	8153	22871		殼⿱今心	
3793	减	4	8158	51CF	减		
3815	卺	2	816E	537A	卺		
3817	厠	3	8170	53A0	厠		
3835	噍	1	81A4	564D	噍		
3838	鋒	4	81A7	2899F		金⿰乂夆	

編號	字碼	字頻	Big5	Unicode	WinXP	構字式	備註
3840	坂	3	81A9	5742	坂		
3843	堦	1	81AC	5826	堦		
3844	堿	1	81AD	5884	堿		
3866	崐	1	81C3	5D10	崐		
3875		4	81CC			豈⿰丶丶支	待造字。
3882	靂	1	81D3	5F72	靂		
3889	愨	2	81DA	6164	愨		
3895	淑	2	81E0	6E86	淑		
3916	敭	3	81F5	656D	敭		
3922	暙	1	81FB	23293		執⿱亼日	執⿱亼目，錯字。
3924	朞	2	81FD	671E	朞		
3938	橩	1	824C	6AC8	橩		
3945		7	8253			氵⿰丶丶义	氵⿰丶丶乂，錯字。
3950	浜	2	8258	6D5C	浜		
3959	潛	1	8261	6FF3	潛		
3981	珉	9	8277	73C9	珉		

編號	字碼	字頻	Big5	Unicode	WinXP	構字式	備註
3988	瑉	3	827E	7482	瑉		
3990	甄	4	82A2	750E	甄		
3995	晦	1	82A7	756E	晦		
4003	塚	3	82AF	585C	塚		
4015	眦	1	82BB	7726	眦		
4027	裨	85	82C7	7986	裨		
4031	杭	1	82CB	79D4	杭		
4036	窰	12	82D0	7AB0	窰		
4039	竚	1	82D3	7ADA	竚		
4124	蝨	1	8369	8771	蝨		
4136	衰	1	8375	887A	衰		
4142	褒	4	837B	8943	褒		
4190	郄	3	83CD	90C4	郄		
4198	鑛	2	83D5	945B	鑛		
4216	韈	2	83E7	97C8	韈		
4251	鰕	2	844B	9C15	鰕		

編號	字碼	字頻	Big5	Unicode	WinXP	構字式	備註
4254	鶡	1	844E	9D7E	鶡		
4271	齩	1	845F	9F69	齩		
4275	义	14	8463	4E49	义		義，取代。
4308	刼	7	84A6	5227	刼		
4348	偬	1	84CE	506C	偬		
4350	品	3	84D0	5058	品		
4378	隲	12	84EC	28F65		陟△馬	
4392	儁	2	84FA	5136	儁		倠△冏，錯字。
4400	兟	1	8543	5160	兟		
4558	厩	1	8644	53A9	厩		Unicode 字型呈現差異。
4563		3	8649			山△王	待造字。
4627	礙	99	86AB	791F	礙		
4630		3	86AE			形匕示欠⊙	形上示欠⊙，錯字。
4668	器	7	86D4	5650	器		
4684		1	86E4			礻△受	待造字。
4720		2	8749			土△已	圮，取代。

編號	字碼	字頻	Big5	Unicode	WinXP	構字式	備註
4746	鼇	3	8763	28932		形牙攵厘⊡	
4754	辰	1	876B	200A2		辰	
4762	旃	1	8773	3ACB		㫃△冉	
4769	恩	1	877A	60E5	恩		
4812	羌	2	87C7	26351		羌△、	
4827	煢	2	87D6	3DC0		○○火△孔	
4845		3	87E8			女△間	待造字。
4910		1	886A			氵△畜	待造字。
4933	耑	7	88A3	5CDD	耑		
4961	肣	1	88BF	80B9	肣		
4981	鳧	1	88D3	9CEC	鳧		
4997	褒	2	88E3	2768C		衣△壯	
5031	閒	1	8946	95B4	閒		
5095	来	1	89A8	6765	来		
5108	惛	1	89B5	60FD	惛		
5131	窻	1	89CC	7ABB	窻		Unicode 字型呈現差異。

編號	字碼	字頻	Big5	Unicode	WinXP	構字式	備註
5173	漱	1	89F6	6F44	漱		
5200		1	8A52			形ノ冂米大⊡	待造字。
5223		3	8A69			形扌穴人夕⊡	待造字。
5230		9	8A70			形土亩且⊡	待造字。
5253		1	8AA9			⺊⌂冈⌂同	⺊⌂咼，錯字。
5292	㝵	1	8AD0	65C9	㝵		
5305	昞	18	8ADD	661E	昞		
5306	簡	1	8ADE			竹⌂閒	
5312		3	8AE4			形女ㄗ冈比⊡	形女⺈比⊡，錯字。
5335	䖟	1	8AFB	459F		亡⌂虫	
5342	曆	3	8B43			形厂林日⊡	
5395	規	1	8B78	4093		矢⌂見	
5418	邞	1	8BB1	28697		冉⌂阝	
5423	栐	1	8BB6	6810	栐		
5549	藁	2	8C75	85F3	藁		
5552	弆	1	8C78	5F06	弆		

編號	字碼	字頻	Big5	Unicode	WinXP	構字式	備註
5567	嶼	4	8CA9	3817		山⿰奥	
5583	飧	1	8CB9	98F1	飧		
5707	桄	1	8D76	6830	桄		
5710	稆	1	8D79	7A06	稆		Unicode 字型呈現差異。
5727		1	8DAC			霖⿰彡	待造字。
5729	漈	1	8DAE	6F34	漈		
5739		3	8DB8			囚⿱古	待造字。
5742	澁	2	8DBB	6F81	澁		
5753	潛	25	8DC6	6F5C	潛		
5764	籉	4	8DD1	7C5E	籉		
5770	汧	5	8DD7			氵⿰开	