

《客座贅語》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
2008/8/5 葉冠孜 製作

一、《客座贅語》一書(客座贅語.xml)使用舊版造字 295 個，字頻 1438 次，詳如附件一。這 295 個造字中，216 個可轉成 Windows XP 能顯示的字，字頻 1008 次；另外 79 個字必須轉成構字式，字頻 430 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，未確定字形，必須人工取代：在舊版漢籍電子文獻中即存在的未確定字形，持續留存在新版漢籍電子文獻中，必須人工一一比對修改，如編號 309「

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 4393 的「𡗗」字，Unicode 編碼為 39D9，由於 Windows XP 並不支援，仍須使用構字式「𠂇𠂇必」。

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、《客座贅語》一書(客座贅語.xml)中未確定字形「𠂇」共出現 4 次，詳細取代內容可參考附件二未確定字形「𠂇」手動取代表。附件二未確定字形「𠂇」手動取代表的欄位說明如下：

- 甲、檔案原文摘錄：摘錄檔案(客座贅語.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。
- 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。
- 丙、備註：記錄缺字處理情況或其他事項。

附件一、《客座贅語》造字分析表

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
14	胤	1	1	FA4D	38E7		𠂇彳么月乚口	
63	啐	1	1	FA7E	54F0	啐		
296	𡗗	58	58	FBEC			𡗗	
297	𡗗	115	115	FBED			𡗗	

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
298	△	16	16	FBEE			△	
299	形	19	19	FBEF			形	
300	□	19	19	FBF0			□	
301	∞	1	1	FBF1			∞	
304	∞	1	1	FBF4			∞	
309	?	4	4	FBF9			?	未確定字形，人工取代，參考附件二。
407	夆	1	1	FCBE			夆	
480	小	1	1	FD48			小	
489	隶	1	1	FD51			隶	
521	茈	1	1	FD71			茈	
608	夂	1	1	FDEA			夂	
616	冂	1	1	FDF2			冂	
628	繼	1	1	FDFE			繼	
629	巳	1	1	FE40			巳	
639	彳	2	2	FE4A			彳	
656	音	1	1	FE5B			音	

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
674	乚	2	2	FE6D			乚	
675	食	1	1	FE6E			食	
685	ㄛ	2	2	FE78			ㄛ	
687	ㄣ	2	2	FE7A			ㄣ	
690	ㄨ	3	3	FE7D			ㄨ	
698	ㄤ	1	1	FEA7			ㄤ	
707	ㄨㄨ	2	2	FEB0			ㄨㄨ	
710	川	1	1	FEB3			川	
717	ㄩ	4	4	FEBA			ㄩ	
730	ㄣ	1	1	FEC7			ㄣ	
746	ㄣ	1	1	FED7			ㄣ	
752	ㄣ	2	2	FEDD			ㄣ	
759	广	4	4	FEE4			广	
762	八	1	1	FEE7			八	
763	ㄩ	2	2	FEE8			ㄩ	
770	ㄣ	10	10	FEEF			ㄣ	

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
780	彳	3	3	FEF9			彳	
783	亠	1	1	FEFC			亠	
784	亻	4	4	FEFD			亻	
807	儻	1	1	8E55	50D2	儻		
825	坳	3	3	8E67	573D	坳		
827	絃	2	2	8E69	9B8C	絃		
832	淹	1	1	8E6E	6CB2	淹		
847	甌	1	1	8E7D	7F3B	甌		
899	駝	1	1	8ED3	99DE	駝		
912	鷓	1	1	8EE0	9DC0	鷓		
969	插	1	1	8F5A	7250	插		
979	蕙	1	1	8F64	56CF	蕙		
1022	瑤	1	1	8FB1	73E4	瑤		
1024	虵	1	1	8FB3	8675	虵		
1067	綦	1	1	8FDE	68CA	綦		
1092	瑞	1	1	8FF7	747A	瑞		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
1109	璫	1	1	9049	7496	璫		
1133	瓌	1	1	9061	74CC	瓌		
1149		1	1	9071			豕厶生	豨，異體字問題。
1205	眇	1	1	90CB	770E	眇		
1226	罰	1	1	90E0	7F78	罰		
1310	蠛	4	4	9175	45EB		虫厶麻	
1330		4	4	91AB			禾厶貞	待造字。
1409	寵	1	1	91FA	7AC9	寵		
1467	纂	2	2	9275	25C67		竹厶秦	
1591	羗	3	3	9354	7F97	羗		
1598	翰	2	2	935B	4A7A		革厶翁	
1642	脉	3	3	93A9	8109	脉		
1685	鵠	1	1	93D4	4CD3		老厶鳥	
1871	蟲	3	3	94F1	87C1	蟲		
1912	襍	1	1	955B	894D	襍		
1981	讐	5	5	95C2	8B90	讐		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
1999	贊	2	2	95D4	8CDB	贊		
2046	踪	1	1	9644	8E2A	踪		
2050	蹠	1	1	9648	8E4F	蹠		
2053	躡	1	1	964B	8EA7	躡		
2081	餵	1	1	9667	9926	餵		
2131	彌	1	1	96BB	5F3B	彌		
2138	缶	1	1	96C2	7F36	缶		
2159	粦	4	4	96D7	4D35		麥厶并	
2172	積	2	2	96E4	408E		矛厶贊	
2225	祴	2	2	975A	88D3	祴		
2278	鎗	1	1	97B1	936E	鎗		
2292		1	1	97BF			囟厶〇火言火囟	待造字。
2441	奢	2	2	98B7	3E19		父厶者	
2460	駮	1	1	98CA	2988D		馬厶瓜	
2485	鯪	1	1	98E3	4C79		魚厶差	
2502	颯	1	1	98F4	98C8	颯		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
2522	迨	1	1	9949	4894		辵△合	
2564	刺	1	1	9973	34E8		夾△刂	
2565	霏	1	1	9974	9714	霏		
2571		1	1	997A			𠄎; 弗𠄎	待造字。
2590	咩	2	2	99AF	54F6	咩		
2610	繹	1	1	99C3	9C53	繹		
2631	鷄	3	3	99D8	9D44	鷄		
2651	鸚	4	4	99EC	9E1C	鸚		
2742	抄	2	2	9AAA	6331	抄		
2770	絳	1	1	9AC6	7D98	絳		
2806	靱	1	1	9AEA	976D	靱		
2829	月	3	3	9B42			月	
2861		1	1	9B62			𠄎△育	待造字。
2902	憇	1	1	9BAD	6138	憇		
2906	竿	1	1	9BB1	659A	竿		
2998	棧	1	1	9C4E			木△戈	

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3002	鞞	1	1	9C52	97B1	鞞		
3004	籊	1	1	9C54	7BD0	籊		
3005	糍	2	2	9C55	7CCD	糍		
3019	腩	1	1	9C63	817C	腩		
3028	禘	1	1	9C6C			禘	
3048	叅	1	1	9CA2	53C5	叅		
3096	歲	1	1	9CD2	5D57	歲		
3161	壘	2	2	9D54	58DC	壘		
3168	瘠	1	1	9D5B	24EA5		疒△答	
3177	蠶	6	6	9D64	7F4E	蠶		
3255	善	106	106	9DD4	5584	善		
3770	个	1	1	8141	4E2A	个		
3789	冲	3	3	8154	51B2	冲		
3792	凉	3	3	8157	51C9	凉		
3794	凑	1	1	8159	51D1	凑		
3799	刦	4	4	815E	5226	刦		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3801	剗	3	3	8160	5257	剗		
3804	効	1	1	8163	52B9	効		
3813	邛	1	1	816C	536D	邛		
3814	却	1	1	816D	5374	却		
3816	厓	5	5	816F	5393	厓		
3817	廁	1	1	8170	53A0	廁		
3824	叙	4	4	8177	53D9	叙		
3826	叶	1	1	8179	53F6	叶		
3829	咏	16	16	817C	548F	咏		
3830	咤	2	2	817D	54A4	咤		
3833	嘑	2	2	81A2	5611	嘑		
3835	噍	1	1	81A4	564D	噍		
3836	喻	1	1	81A5	564F	喻		
3843	堦	4	4	81AC	5826	堦		
3846	壻	8	8	81AF	58FB	壻		Unicode 呈現差異。
3858	寃	1	1	81BB	5BC3	寃		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3864	峯	25	25	81C1	5CEF	峯		
3869	币	1	1	81C6	5E00	币		
3876	廻	1	1	81CD	5EFB	廻		
3883	徧	4	4	81D4	5FA7	徧		
3886	悞	1	1	81D7	609E	悞		
3889	慤	1	1	81DA	6164	慤		
3891	慙	2	2	81DC	6159	慙		
3892	睺	1	1	81DD	3B0B		日 厶 侯	
3894	筋	1	1	81DF	4225		竹 厶 勛	
3921	皙	5	5	81FA	6673	皙		
3931	椀	1	1	8245	6900	椀		
3932	椶	1	1	8246	6936	椶		
3934	楬	1	1	8248	69C5	楬		
3935	槩	3	3	8249	69E9	槩		Unicode 呈現差異。
3937	檝	2	2	824B	6A9D	檝		
3949	涖	1	1	8257	6D96	涖		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3966	烟	28	28	8268	70DF	烟		
3974	牀	12	12	8270	7240	牀		
3977	狗	6	6	8273	72E5	狗		
3978	猪	2	2	8274	732A	猪		Unicode 呈現差異。
3979	猫	5	5	8275	732B	猫		
3984	璫	2	2	827A	7460	璫		
3985	璫	3	3	827B	7447	璫		Unicode 呈現差異。
3994	畊	1	1	82A6	754A	畊		
3996	疎	8	8	82A8	758E	疎		
4008	癯	1	1	82B4	766F	癯		
4013	盃	1	1	82B9	76CC	盃		
4015	眦	1	1	82BB	7726	眦		
4016	着	9	9	82BC	7740	着		Unicode 呈現差異。
4027	裨	1	1	82C7	7986	裨		
4028	稟	6	6	82C8	7980	稟		
4031	杭	2	2	82CB	79D4	杭		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4036	窰	6	6	82D0	7AB0	窰		
4038	竈	13	13	82D2	7AC8	竈		
4042	笋	3	3	82D6	7B0B	笋		
4043	筴	2	2	82D7	7B53	筴		Unicode 呈現差異。
4044	筴	3	3	82D8	7B6F	筴		
4052	簾	1	1	82E0	7C34	簾		
4056	莽	2	2	82E4	7CA6	莽		
4060	穀	1	1	82E8	7CD3	穀		
4063	藜	1	1	82EB	7CF5	藜		
4065	紮	1	1	82ED	7D25	紮		
4068	綫	3	3	82F0	7DAB	綫		
4074	繚	2	2	82F6	7E27	繚		
4078		1	1	82FA			巳𠄎升	待造字。
4080	羣	36	36	82FC	7FA3	羣		
4081	翱	1	1	82FD	7FFA	翱		
4089	脇	3	3	8346	8107	脇		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4090	脚	15	15	8347	811A	脚		
4091	臯	2	2	8348	81EF	臯		
4105	菓	5	5	8356	83D3	菓		
4107	葱	5	5	8358	8471	葱		
4109	葛	1	1	835A	84AD	葛		
4112	蔴	2	2	835D	8534	蔴		
4114	藥	1	1	835F	8616	藥		
4124	蟲	1	1	8369	8771	蟲		
4129	虬	2	2	836E	866C	虬		
4132		61	61	8371			血𠄎𠄎	眾，異體字問題。
4134	衛	103	103	8373	885E	衛		
4139	袴	1	1	8378	88B4	袴		
4146	覩	3	3	83A1	89A9	覩		
4158	譌	1	1	83AD	8B4C	譌		Unicode 呈現差異。
4178	輓	5	5	83C1	8F2D	輓		
4184	迹	2	2	83C7	8FF9	迹		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4186	遡	2	2	83C9	9061	遡		
4195	鉢	3	3	83D2	9262	鉢		
4213	鞞	2	2	83E4	97BE	鞞		
4220	穎	2	2	83EB	9834	穎		
4226	凜	1	1	83F1	98E1	凜		
4238	髭	1	1	83FD	9AE2	髭		
4240	髯	3	3	8440	9AF4	髯		
4242	鬪	10	10	8442	9B2D	鬪		Unicode 呈現差異。
4247	鮎	1	1	8447	9B8E	鮎		
4249	編	1	1	8449	9BFF	編		
4251	鰕	3	3	844B	9C15	鰕		
4256	鷄	6	6	8450	9DC4	鷄		
4260	麋	1	1	8454	9E9E	麋		
4262	麪	8	8	8456	9EAA	麪		
4263	麴	2	2	8457	9EAF	麴		
4267	鼈	2	2	845B	9F08	鼈		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4270	𪗇	1	1	845E	9F45	𪗇		
4271	𪗈	1	1	845F	9F69	𪗈		
4275	𪗊	2	2	8463	4E49	𪗊		
4278	𪗋	2	2	8466	4E98	𪗋		
4307	𪗏	2	2	84A5	99E1	𪗏		
4322	備	1	1	84B4	4FCC	備		
4342	𪗒	4	4	84C8	502E	𪗒		
4356	𪗖	1	1	84D6	976B	𪗖		
4393	𪗙	1	1	84FB	39D9		扌 𪗙 必	
4412	𪗛	1	1	854F	2726E		虫 𪗛 龟	
4437	廩	1	1	8568	5EEA	廩		
4447	捏	1	1	8572	63D1	捏		
4450	翔	4	4	8575	524F	翔		Unicode 呈現差異。
4460	勅	35	35	85A1	52D1	勅		
4495	𪗞	1	1	85C4	9EA8	𪗞		
4496	段	11	11	85C5	53DA	段		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4507	吳	150	150	85D0	5434	吳		
4528	嶼	1	1	85E5	5D60	嶼		
4529	嘜	1	1	85E6	54F4	嘜		
4566	礶	1	1	864C	78F5	礶		
4607	狫	3	3	8675	3E60		犭 厶 屯	
4618	鑛	1	1	86A2	4956		金 厶 奠	
4629	坩	1	1	86AD	577F	坩		
4649	蕒	1	1	86C1	8552	蕒		
4659	垲	1	1	86CB	57DB	垲		
4718	屯	1	1	8747	340C		屯	
4735	壠	4	4	8758	58E0	壠		
4750	卅	5	5	8767			卅	
4767	懷	1	1	8778	61B9	懷		
4768	妬	1	1	8779	59AC	妬		
4772	鱣	3	3	877D	9BF6	鱣		
4774	鯪	2	2	87A1	9BA7	鯪		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4778	鞞	1	1	87A5	97B4	鞞		
4789	蝮	1	1	87B0	878B	蝮		
4805	瀆	4	4	87C0	51DF	瀆		
4814	健	6	6	87C9	5FA4	健		
4866	鼃	7	7	87FD	9F04	鼃		
4874	邈	1	1	8846	6530	邈		
4878	宄	1	1	884A	5B82	宄		Unicode 呈現差異。
4880	輶	2	2	884C	8F28	輶		
4902	兗	1	1	8862	5156	兗		
4904	尅	2	2	8864	5C05	尅		
4909	弥	1	1	8869	5F25	弥		Unicode 呈現差異。
5019	畧	16	16	88F9			畧	
5049	鬚	1	1	8958	29B6A		鬚𠂇屈	
5074	𠂇	2	2	8971	5C2B	𠂇		
5131	窻	3	3	89CC	7ABB	窻		Unicode 呈現差異。
5163	憂	2	2	89EC	621E	憂		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5169	危	1	1	89F2	6239	危		
5173	漱	2	2	89F6	6F44	漱		
5241	鷺	7	7	8A7B	9D5E	鷺		
5242	飢	1	1	8A7C	2064F		飢 飢	
5281	餓	1	1	8AC5	994A	餓		
5290	携	2	2	8ACE	643A	携		
5292	粵	2	2	8AD0	65C9	粵		
5315	公	1	1	8AE7	3CC2		公 公	
5327	嶂	1	1	8AF3	5D2A	嶂		
5344	替	2	2	8B45	55F8	替		
5402	术	1	1	8BA1	672F	术		
5405	鞞	1	1	8BA4	6A37	鞞		Unicode 呈現差異。
5418	卹	1	1	8BB1	28697		卹 卹	
5422	梯	2	2	8BB5	67F9	梯		
5439	栢	3	3	8BC6	6822	栢		
5464	燠	1	1	8BDF	7157	燠		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5466	倏	4	4	8BE1	5010	倏		Unicode 呈現差異。
5515	軌	2	2	8C53			車厶丸	
5542	鑠	2	2	8C6E	93C1	鑠		
5558	龕	2	2	8C7E	9E84	龕		
5579	宐	1	1	8CB5	5D08	宐		
5583	殮	2	2	8CB9	98F1	殮		
5587	歛	1	1	8CBD	6B58	歛		
5588	恡	1	1	8CBE	6061	恡		
5591	杞	1	1	8CC1	233CC		木厶巳	
5687	貜	1	1	8D62	4AC9		貜	
5697	煑	6	6	8D6C	7151	煑		
5702	厨	6	6	8D71	53A8	厨		
5704	羶	5	5	8D73	6C0A	羶		
5706	敍	11	11	8D75	654D	敍		
5739		1	1	8DB8			冗厶古	待造字。
5768	濶	9	9	8DD5	6FF6	濶		

附件二、《客座贅語》未確定字形「𠄎」手動取代表

檔案原文摘錄	手動取代結果	備註
<p>押字形為𠄎𠄎兀<qe>押字形為 𠄎𠄎兀 「𠄎𠄎兀」形原為「𠄎 𠄎兀」，據萬曆本臨摹。</qe></p>	<p>押字形為𠄎𠄎兀<qe>押字形 為𠄎𠄎兀 「𠄎𠄎兀」形原 為「𠄎𠄎兀」，據萬曆本臨 摹。</qe></p>	<p>此處「𠄎」字形 為符號，暫不修 改。</p>