

# 《歸田瑣記》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  
2008/5/22 陳建安 製作

一、《歸田瑣記》一書(歸田瑣記.xml)使用舊版造字 175 個，字頻 605 次，詳如附件一。這 175 個造字中，144 個可轉成 Windows XP 能顯示的字，字頻 476 次；另外 31 個字必須轉成構字式，字頻 129 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 923 的「火产兼

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

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 3846 的「塿」字，Unicode 字型呈現為「塿」，實際上仍為同一字。

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 44 的「颶」字，Unicode 編碼為 295D7，由於 Windows XP 並不支援，仍須使用構字式「風△貝」。

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、《歸田瑣記》造字分析表

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
44	颶	1	1	FA6B	295D7		風△貝	
297	𪗇	8	8	FBED			𪗇	
298	△	6	6	FBEE			△	
396	𪗈	1	1	FCB3			𪗈	
658	疒	3	3	FE5D			疒	
759	广	3	3	FEE4			广	
821	脊	1	1	8E63	7718	脊		
841	礪	1	1	8E77	7932	礪		
857	脛	1	1	8EA9	26715		𠃉月丞巳𠃉	
893	峩	1	1	8ECD	5CE9	峩		峨，異體字問題。
923		2	2	8EEB			𠃉火产兼𠃉	爌，校對問題，舊版漢籍錯字，可用程式全部取代。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
1024	虵	2	2	8FB3	8675	虵		
1026	珣	1	1	8FB5	73E6	珣		
1133	瓌	1	1	9061	74CC	瓌		
1187	盖	4	4	90B9	76D6	盖		盖，異體字問題。
1205	眎	1	1	90CB	770E	眎		
1221	嬰	1	1	90DB	5A3F	嬰		
1226	罰	3	3	90E0	7F78	罰		
1308	蕙	1	1	9173	85FC	蕙		
1364		1	1	91CD			冗厶圭	罍，異體字問題。
1490	羸	2	2	92AE	7C5D	羸		
1591	羗	2	2	9354	7F97	羗		羗，異體字問題。
1630	郟	3	3	937B	48BF		安厶卩	
1642	脉	1	1	93A9	8109	脉		
1708	梲	3	3	93EB	3B7F		木厶昂	
1917	襍	1	1	9560	8965	襍		
1981	讐	1	1	95C2	8B90	讐		
2066	滌	1	1	9658	3D11		氵厶席	
2140	迫	2	2	96C4	5EF9	迫		迫，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
2152	蕓	2	2	96D0	979F	蕓		
2168	疇	1	1	96E0	91BB	疇		
2260	癩	1	1	977D	3FD7		疒△頰	
2502	颺	1	1	98F4	98C8	颺		
2628	莢	1	1	99D5	845C	莢		
2666	舩	1	1	99FB	4478		舟△宗	
2770	絳	1	1	9AC6	7D98	絳		絳，校對問題，舊版漢籍錯字，可用程式全部取代。
2863		1	1	9B64			艹△友	芟，異體字問題。
3003	啖	22	22	9C53	20E04		口△英	
3161	壘	1	1	9D54	58DC	壘		
3184	櫛	1	1	9D6B	6AC9	櫛		櫛，異體字問題。
3789	冲	1	1	8154	51B2	冲		冲，異體字問題。
3792	凉	1	1	8157	51C9	凉		凉，異體字問題。
3804	効	5	5	8163	52B9	効		
3805	勅	1	1	8164	52C5	勅		
3814	却	1	1	816D	5374	却		卻，異體字問題。
3815	𡗗	1	1	816E	537A	𡗗		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3816	厓	2	2	816F	5393	厓		
3817	厠	3	3	8170	53A0	厠		厠，異體字問題。
3820	廈	11	11	8173	53A6	廈		廈，異體字問題。
3824	叙	3	3	8177	53D9	叙		叙，異體字問題。
3825	疊	6	6	8178	53E0	疊		疊，異體字問題。
3826	叶	4	4	8179	53F6	叶		
3829	咏	2	2	817C	548F	咏		
3830	咤	1	1	817D	54A4	咤		
3846	壻	4	4	81AF	58FB	壻		Unicode 呈現差異。
3852	媼	1	1	81B5	5A63	媼		
3858	寃	5	5	81BB	5BC3	寃		寃，異體字問題。
3864	峯	22	22	81C1	5CEF	峯		峰，異體字問題。
3868	卮	2	2	81C5	5DF5	卮		卮，異體字問題。
3883	徧	16	16	81D4	5FA7	徧		
3886	悞	2	2	81D7	609E	悞		
3891	慙	1	1	81DC	6159	慙		慙，異體字問題。
3892	睺	1	1	81DD	3B0B		日 𠄎 侯	
3895	淑	1	1	81E0	6E86	淑		淑，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3897	慇	1	1	81E2	615C	慇		
3909	擡	1	1	81EE	64E1	擡		
3916	敷	1	1	81F5	656D	敷		
3929	黎	1	1	8243	68C3	黎		
3931	椀	1	1	8245	6900	椀		
3934	桶	1	1	8248	69C5	桶		
3937	檝	1	1	824B	6A9D	檝		
3949	泣	4	4	8257	6D96	泣		
3961	灾	1	1	8263	707E	灾		災，異體字問題。
3966	烟	4	4	8268	70DF	烟		煙，異體字問題。
3974	牀	3	3	8270	7240	牀		
3977	狗	6	6	8273	72E5	狗		
3978	猪	4	4	8274	732A	猪		Unicode 呈現差異。
3979	猫	1	1	8275	732B	猫		
3991	攜	1	1	82A3	64D5	攜		
3996	疎	3	3	82A8	758E	疎		
4000	疴	6	6	82AC	75B4	疴		
4011	皤	1	1	82B7	76A1	皤		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4013	盃	1	1	82B9	76CC	盃		
4014	蓋	3	3	82BA	8462	蓋		蓋，異體字問題。
4016	着	9	9	82BC	7740	着		著，異體字問題。
4024	礮	3	3	82C4	792E	礮		
4028	稟	3	3	82C8	7980	稟		稟，異體字問題。
4042	笋	1	1	82D6	7B0B	笋		
4068	綫	1	1	82F0	7DAB	綫		線，異體字問題。
4080	羣	18	18	82FC	7FA3	羣		群，異體字問題。
4081	翱	3	3	82FD	7FFA	翱		
4085	耻	1	1	8342	803B	耻		
4090	脚	5	5	8347	811A	脚		腳，異體字問題。
4097	芑	5	5	834E			𠄎 𠄎 巳	芑，異體字問題。
4098	芪	1	1	834F	82AA	芪		
4103	茝	8	8	8354	831D	茝		
4104	莅	9	9	8355	8385	莅		
4111	蕓	1	1	835C	8493	蕓		
4112	蔴	1	1	835D	8534	蔴		
4115	藁	1	1	8360	85C1	藁		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4129	虬	1	1	836E	866C	虬		
4132		48	48	8371			血𠂇𠂇	眾，異體字問題。
4134	衛	9	9	8373	885E	衛		衛，異體字問題。
4138	表	2	2	8377	88A0	表		
4139	袴	1	1	8378	88B4	袴		
4146	覩	2	2	83A1	89A9	覩		
4149	覩	4	4	83A4	89BC	覩		
4167	賈	1	1	83B6	8CEB	賈		
4173	耽	2	2	83BC	8EAD	耽		
4176	輒	7	7	83BF	8F19	輒		輒，異體字問題。
4184	迹	5	5	83C7	8FF9	迹		
4193	鈎	1	1	83D0	920E	鈎		鈎，異體字問題。
4195	鉢	3	3	83D2	9262	鉢		
4198	鑛	3	3	83D5	945B	鑛		
4203	隣	5	5	83DA	96A3	隣		鄰，異體字問題。
4208	磧	1	1	83DF	9767	磧		
4220	穎	5	5	83EB	9834	穎		穎，異體字問題。
4242	鬪	1	1	8442	9B2D	鬪		Unicode 呈現差異。



編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4243	𩇛	1	1	8443	9B34	𩇛		
4249	𩇛	2	2	8449	9BFF	𩇛		
4267	𩇛	1	1	845B	9F08	𩇛		
4278	𩇛	1	1	8466	4E98	𩇛		
4297	躔	2	2	8479	281E0		𠃉足厂墨	躔，異體字問題。
4307	罵	2	2	84A5	99E1	罵		罵，異體字問題。
4308	劫	1	1	84A6	5227	劫		劫，異體字問題。
4309	劫	2	2	84A7	523C	劫		劫，異體字問題。
4348	惚	1	1	84CE	506C	惚		
4447	捏	1	1	8572	63D1	捏		捏，異體字問題。
4448		2	2	8573			𠃉糸厂墨	纏，異體字問題。
4450	𠃉	3	3	8575	524F	𠃉		𠃉，異體字問題。
4455	劍	1	1	857A			僉 刀	劍，異體字問題。
4507	吳	69	69	85D0	5434	吳		吳，異體字問題。
4553		1	1	85FE			𠃉 卩 久 生	蔭，校對問題，舊版漢籍錯字，可用程式全部取代。
4603	憇	1	1	8671	6187	憇		
4629	坴	1	1	86AD	577F	坴		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4727	壘	1	1	8750	58C4	壘		
4746		1	1	8763	28932		𠄎牙女厘𠄎	釐，異體字問題。
4762	旃	1	1	8773	3ACB		从△冉	
4763	斨	2	2	8774	6586	斨		
4785	贗	1	1	87AC	8D0B	贗		贗，異體字問題。
4894		1	1	885A			宀△心△冉	寧，異體字問題。
4902	兗	1	1	8862	5156	兗		兗，異體字問題。
4904	尅	3	3	8864	5C05	尅		
4908	届	1	1	8868	5C4A	届		届，異體字問題。
4996	斲	1	1	88E2	65B5	斲		
5074	尫	1	1	8971	5C2B	尫		尫，異體字問題。
5085	苜	2	2	897C	26B12		艹△艹	苜，異體字問題。
5131	窻	1	1	89CC	7ABB	窻		Unicode 呈現差異。
5173	漱	1	1	89F6	6F44	漱		漱，異體字問題。
5186	尗	1	1	8A44	5C17	尗		
5218		1	1	8A64			𠄎宀寬、𠄎	寬，異體字問題。
5225	卹	2	2	8A6B	460F		血△卩	卹，異體字問題。
5241	鷺	7	7	8A7B	9D5E	鷺		鷺，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5275	掇	1	1	8ABF	3A16		扌 ㄩ致	
5289	慤	3	3	8ACD	6164	慤		慤，異體字問題。
5290	携	1	1	8ACE	643A	携		
5300	鬪	1	1	8AD8	9B2A	鬪		
5310	烏	1	1	8AE2	8204	烏		
5405	藪	1	1	8BA4	6A37	藪		Unicode 呈現差異。
5434	湟	2	2	8BC1	6E7C	湟		湟，異體字問題。
5466	倏	2	2	8BE1	5010	倏		Unicode 呈現差異。
5555	迥	2	2	8C7B	9008	迥		迥，異體字問題。
5556	隸	1	1	8C7C	96B7	隸		隸，異體字問題。
5558	麓	1	1	8C7E	9E84	麓		
5559	欸	14	14	8CA1	6B35	欸		欸，異體字問題。
5582	鞣	1	1	8CB8	97B5	鞣		
5590	朶	3	3	8CC0	6736	朶		朶，異體字問題。
5591	杞	2	2	8CC1	233CC		木 ㄩ已	杞，異體字問題。
5678	量	1	1	8D59	757A	量		
5697	煑	3	3	8D6C	7151	煑		煮，異體字問題。
5702	厨	16	16	8D71	53A8	厨		廚，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5706	敘	17	17	8D75	654D	敘		敘，異體字問題。
5720	垧	4	4	8DA5	578D	垧		
5753	潛	1	1	8DC6	6F5C	潛		潛，異體字問題。
5772	澆	1	1	8DD9	6FD9	澆		