

《全宋詩(節選)》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
2008/4/28 黃榮順 製作

一、《全宋詩(節選)》一書(全宋詩(節選).xml)使用舊版造字 446 個，字頻 2043 次，詳如附件一。這 446 個造字中，318 個可轉成 Windows XP 可顯示的字，字頻 1465 次；另外 128 個字必須轉成構字式，字頻 578 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 681 的「𠂔」字，原書字形為「𠂔」。
- 2、重碼問題，部分採用舊漢籍造字，部分採用部件外字，人工取代：舊漢籍造字與構字式部件外字，皆是建立在 Big5 的使用者造字區中，兩者所使用的碼區是重疊的，在同時使用構字式和舊漢籍造字的新版漢籍中，這些重疊的字碼可能被用以表達兩個不同字形，我們稱之為「重碼問題」。在使用構字式的 xml 檔中，舊漢籍造字也會呈現為重碼的部件外字，因而必須將採用舊漢籍造字部分以對應的 Unicode 或構字式取代，才能呈現正確字形；至於採用部件外字者則不需轉碼。如編號 5290，字碼 8ACE，對應的舊漢籍造字為「携」，部件外字為「𠂔」，一一比對原書，確認是「携」者，再以標準字「攜」手動取代。

- 3、重碼問題，採用部件外字，不轉碼：舊漢籍造字與構字式部件外字，皆是建立在 Big5 的使用者造字區中，兩者所使用的碼區是重疊的，在同時使用構字式和舊漢籍造字的新版漢籍中，這些重疊的字碼可能被用以表達兩個不同字形，我們稱之為「重碼問題」。然而漢籍在舊轉新的過程中，若已將原舊漢籍造字的部分全數改為 Unicode 或構字式，則在 xml 出現的造字頻次皆採用部件外字，此時便不需轉碼。如編號 5801，字碼 8DF6，對應的舊漢籍造字為「𠂇 𠂇 贏」，部件外字為「𠂇」，xml 檔中的頻次皆使用部件外字「𠂇」，故不需轉碼。
- 4、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 814 的「𦵏」係「𦵏」字之異體，故以「𦵏」字取代。又如編號 5621 的「𦵏」係「淚」字之簡體，故以標準字「淚」取代。
- 5、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 1076 的「𦵏」字，Unicode 字型呈現為「𦵏」，實際上仍為同一字。
- 6、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 4792 的「𦵏 𦵏 虛」。

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 32 的「𦵏」字，Unicode 編碼為 3ABA，由於 Windows XP 並不支援，仍須改用構字式「𦵏 𦵏 斗」。

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、 txt 檔與 xml 檔頻次不同或頻次相同但出現需手動處理的情況時，造字分析表「備註」欄有備註問題情況者，詳細取代內容可參考附件二的手動取代表。附件二的手動取代表說明如下：

甲、 編號：Big5 造字空間為 6217 個，編號由 1 到 6217。

乙、 造字：舊漢籍造字。

丙、 取代：使用字形與造字欄不同時，於此欄列出。

丁、 檔案原文摘錄：摘錄檔案(全宋詩(節選).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。

戊、 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。

五、 「全宋詩(節選).xml」發現之缺字外錯字共 1 處，修正情形詳列附件三，附件三「全宋詩(節選).xml」缺字外錯字修改表體例說明如下：

甲、 檔案原文摘錄：摘錄檔案(全宋詩(節選).xml)中包含該錯字的文句段落，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。

附件一、「全宋詩(節選).xml」造字分析表

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
32	𪗇	1	1	FA5F	3ABA		𪗇 𪗇 斗	
66	𪗈	1	1	FAA3	4578		𪗈 𪗈 點	
73		1	1	FAAA			𪗉 𪗉 姜	繡，校對問題，舊版漢籍錯字，可用程式全部取代。

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
74	髻	1	1	FAAB	29B06		髻𠂇𠂇𠂇	髻，異體字問題。
83	𧈧	1	1	FAB4	86AD	𧈧		
126		1	1	FADF				疊，校對問題，舊版漢籍錯字，可用程式全部取代。
209		4	4	FB73				杞，校對問題，舊版漢籍錯字，可用程式全部取代。
228		14	14	FBA8				椀，校對問題，舊版漢籍錯字，可用程式全部取代。
296	𠂇	72	72	FBEC			𠂇	
297	𠂇	169	169	FBED			𠂇	
298	𠂇	16	16	FBEE			𠂇	
299	𠂇	9	9	FBEF			𠂇	
300	𠂇	9	9	FBF0			𠂇	
327	𠂇	1	1	FC4C			𠂇	
336	𠂇	1	1	FC55			𠂇	
387	𠂇	1	1	FCAA			𠂇	
396	𠂇	1	1	FCB3			𠂇	
407	𠂇	4	4	FCBE			𠂇	

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
429	隹	1	1	FCD4			隹	
520	𠂔	1	1	FD70			𠂔	
658	疒	5	5	FE5D			疒	
675	食	3	3	FE6E			食	
681	𠂔	1	1	FE74			𠂔	𠂔，校對問題，舊版漢籍錯字，可用程式全部取代。
715	彡	2	2	FEB8			彡	
717	𠂔	2	2	FEBA			𠂔	
720	夕	5	5	FEBD			夕	
739	勹	1	1	FED0			勹	
752	𠂔	6	6	FEDD			𠂔	
758	𠂔	2	2	FEE3			𠂔	
766	彡	5	5	FEEB			彡	
770	扌	4	4	FEEF			扌	
775	乚	3	3	FEF4			乚	
782	幺	3	3	FEFB			幺	
797	扌	1	1	8E4B	39D6		扌 𠂔	

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
799	𪗇	1	1	8E4D	9F5A	𪗇		
804	𪗈	1	1	8E52	637E	𪗈		
814	𪗉	24	24	8E5C	87C7	𪗉		𪗉，異體字問題。
828	𪗊	1	1	8E6A	8DD4	𪗊		
832	𪗋	2	2	8E6E	6CB2	𪗋		
841	𪗌	4	4	8E77	7932	𪗌		
842	𪗍	4	4	8E78	3AB7		𪗍	
887	𪗎	13	13	8EC7	9E0E	𪗎		
899	𪗏	14	14	8ED3	99DE	𪗏		
910	𪗐	1	1	8EDE	5D78	𪗐		
912	𪗑	4	4	8EE0	9DC0	𪗑		
931	𪗒	1	1	8EF3	791A	𪗒		
934	𪗓	1	1	8EF6	7200	𪗓		
950	𪗔	1	1	8F47	4561		𪗔 𪗔	
969	𪗕	4	4	8F5A	7250	𪗕		
972	𪗖	1	1	8F5D	7287	𪗖		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
977	磨	10	10	8F62	9E8F	磨		
985	慙	2	2	8F6A	6197	慙		
1000		7	7	8F79			王𠂇己	王𠂇己，異體字問題。
1002	鳩	5	5	8F7B	9D02	鳩		
1008	玦	1	1	8FA3	739E	玦		
1024	虵	4	4	8FB3	8675	虵		
1037	鵠	6	6	8FC0	4CCF		圭𠂇鳥	
1039	稟	3	3	8FC2	69C0	稟		
1056	阤	3	3	8FD3	9641	阤		
1057	𪔐	1	1	8FD4	3687		𪔐	
1076	𪔑	1	1	8FE7	5368	𪔑		Unicode 字型差異。
1091	𪔒	8	8	8FF6	58CD	𪔒		
1133	瓌	3	3	9061	74CC	瓌		
1170	瞳	2	2	90A8	7583	瞳		
1196	濡	6	6	90C2	240E9		氵𠂇需	
1205	眎	1	1	90CB	770E	眎		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
1237	稍	1	1	90EB	77DF	稍		
1242	鬻	1	1	90F0	9B1D	鬻		
1267	櫛	1	1	914A	6B1B	櫛		
1310	𧈧	2	2	9175	45EB		虫 𧈧 麻	
1317	𠂔	1	1	917C	4197		穴 𠂔 叫	
1327	𢱤	1	1	91A8	23003		廉 𢱤 支	
1352	𢱤	1	1	91C1	29260		革 𢱤 义	𢱤，異體字問題。
1359	襪	1	1	91C8	25725		衤 𢱤 冀	
1381	穉	1	1	91DE	7A3A	穉		
1383	穰	1	1	91E0	7A45	穰		
1387	腎	4	4	91E4	7796	腎		
1391	穉	1	1	91E8	4173		禾 𢱤 芻	
1396	罕	1	1	91ED	7A82	罕		
1413	岬	4	4	91FE	37A6		山 𢱤 几	
1416	埃	1	1	9242	7AE2	埃		
1442	轍	2	2	925C	487E		車 𢱤 獻	

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
1530	紉	1	1	92D6	7CFA	紉		
1544	紉	3	3	92E4	7D59	紉		
1547	𠂇	7	7	92E7	4E31	𠂇		
1557	羴	2	2	92F1			羊𠂇歷	
1565	繇	8	8	92F9	7DDC	繇		
1618	𥇑	1	1	936F	2522D		目𠂇𠂇	
1736	蒺	1	1	9448	8453	蒺		
1849		1	1	94DB			冗𠂇音	𠂇𠂇音，異體字問題。
1864	𧈧	1	1	94EA	872B	𧈧		
1875	𧈧	1	1	94F5	5268	𧈧		
1885	𧈧	2	2	9540	882D	𧈧		
1900	桂	1	1	954F	88BF	桂		
1905	𧈧	2	2	9554	88E9	𧈧		
1938	𧈧	2	2	9575	8A49	𧈧		
1983	璽	1	1	95C4	24AE9		王𠂇靈	
1999	贊	2	2	95D4	8CDB	贊		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
2033	𪛗	3	3	95F6			鼠厶留	
2044	踈	5	5	9642	8E08	踈		
2076	𪛗	2	2	9662	4189		禾厶罷	
2077	轉	1	1	9663	8F4A	轉		
2078	𪛗	5	5	9664	7A0F	𪛗		
2086	𪛗	1	1	966C	4361		𪛗厶鹿	
2092	𪛗	1	1	9672	20646		𪛗厶几	
2143	火	1	1	96C7			火	火，校對問題，舊版漢籍錯字，可用程式全部取代。
2148	𪛗	1	1	96CC	8B4D	𪛗		
2168	𪛗	1	1	96E0	91BB	𪛗		
2188	𪛗	15	15	96F4	4366		𪛗厶離	
2216	𪛗	2	2	9751	25F9D		𪛗厶𪛗	
2234	𪛗	2	2	9763	38B8		𪛗厶𪛗	
2278	𪛗	1	1	97B1	936E	𪛗		
2298	𪛗	2	2	97C5	3C89		𪛗厶卵	
2339	床	1	1	97EE	221F2		广厶禾	

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
2378	颶	1	1	9856	4AFB		風△日	
2383	𦍋	1	1	985B	44A0		艹△丰	
2409	𨵿	1	1	9875	95A7	𨵿		𨵿，異體字問題。
2502	颶	10	10	98F4	98C8	颶		
2505	𪗇	1	1	98F7	98DC	𪗇		
2539	𨵿	1	1	995A	43F6		月△𨵿	
2562	𨵿	1	1	9971	385A		巾△𨵿	
2565	𨵿	1	1	9974	9714	𨵿		
2580	𪗇	1	1	99A5			艹△𪗇	
2602	鰐	3	3	99BB	9C10	鰐		
2609	鰐	1	1	99C2	9C4F	鰐		Unicode 字型差異。
2628	𪗇	2	2	99D5	845C	𪗇		
2643	𪗇	1	1	99E4	2486E		𪗇△然	
2651	鵠	11	11	99EC	9E1C	鵠		
2658	𪗇	1	1	99F3	97C2	𪗇		
2659	𪗇	1	1	99F4	201D7		𪗇△瓜	

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
2742	抄	1	1	9AAA	6331	抄		
2820	摩	3	3	9AF8	64EA	摩		Unicode 字型差異。
2890	𪛗	1	1	9BA1	24CE6		𪛗𪛗大	𪛗，校對問題，舊版漢籍錯字，可用程式全部取代。
2906	𪛖	2	2	9BB1	659A	𪛖		𪛖，異體字問題。
2938	簪	2	2	9BD1	7C2E	簪		簪，異體字問題。
2956	荅	2	2	9BE3	839F	荅		
3019	𪛙	1	1	9C63	817C	𪛙		
3020	𪛚	4	4	9C64			𪛚	
3113	𪛛	2	2	9CE3	6245	𪛛		
3124	𪛜	2	2	9CEE	543A	𪛜		
3223	𪛝	2	2	9DB4	6EE3	𪛝		
3770	个	2	2	8141	4E2A	个		
3779	𪛞	2	2	814A	290F1		雨𪛞沱	
3780	𪛟	7	7	814B	5101	𪛟		
3795	𪛠	1	1	815A	51D3	𪛠		
3797	凭	5	5	815C	51ED	凭		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
3801	剗	8	8	8160	5257	剗		
3803	餽	1	1	8162	4B54		食 𩚑 追	
3806	勳	2	2	8165	52E0	勳		
3809	匱	5	5	8168	5333	匱		
3810	匱	1	1	8169	5335	匱		
3815	𪗇	1	1	816E	537A	𪗇		
3816	厓	2	2	816F	5393	厓		
3823	擦	1	1	8176	6456	擦		
3830	㗎	10	10	817D	54A4	㗎		
3831	𪗇	2	2	817E	5586	𪗇		
3833	嘑	2	2	81A2	5611	嘑		
3834	嘑	2	2	81A3	5637	嘑		
3835	噍	6	6	81A4	564D	噍		
3840	坂	37	37	81A9	5742	坂		
3843	堦	18	18	81AC	5826	堦		
3844	城	1	1	81AD	5884	城		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
3851	姹	2	2	81B4	59F9	姹		
3852	嫻	1	1	81B5	5A63	嫻		
3856	輾	2	2	81B9	283B9		車𨋖稟	
3868	卮	46	46	81C5	5DF5	卮		
3880	獲	1	1	81D1	5F60	獲		
3883	徧	3	3	81D4	5FA7	徧		
3886	悞	11	11	81D7	609E	悞		
3891	慙	77	77	81DC	6159	慙		慚，異體字問題。
3900	挖	1	1	81E5	62D5	挖		
3910	擘	5	5	81EF	64E5	擘		
3918	旻	3	3	81F7	65DC	旻		
3921	皙	2	2	81FA	6673	皙		
3925	寻	1	1	81FE	3775		寻	
3930	憵	4	4	8244	6901	憵		
3931	椀	25	25	8245	6900	椀		
3932	椶	9	9	8246	6936	椶		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
3933	渠	2	2	8247	6998	渠		
3935	槩	6	6	8249	69E9	槩		概，異體字問題。
3937	檝	7	7	824B	6A9D	檝		
3940	欝	3	3	824E	6B1D	欝		
3946	泓	1	1	8254	6C79	泓		
3949	洳	1	1	8257	6D96	洳		
3956	渚	1	1	825E	6F74	渚		Unicode 字型差異。
3974	牀	5	5	8270	7240	牀		
3976	犁	2	2	8272	7282	犁		
3977	狗	2	2	8273	72E5	狗		
3980	獾	3	3	8276	7328	獾		
3981	珉	12	12	8277	73C9	珉		
3982	珅	1	1	8278	73F7	珅		
3984	瑠	7	7	827A	7460	瑠		
3987	穀	1	1	827D	7474	穀		
3989	璦	9	9	82A1	74C8	璦		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
3996	疎	19	19	82A8	758E	疎		
3997	臚	1	1	82A9	9AD7	臚		
4000	疴	1	1	82AC	75B4	疴		
4004	癭	2	2	82B0	763B	癭		
4006	癢	1	1	82B2	764F	癢		
4008	癰	20	20	82B4	766F	癰		
4012	鼓	1	1	82B8	76B7	鼓		鼓，異體字問題。
4013	盃	1	1	82B9	76CC	盃		
4015	眦	2	2	82BB	7726	眦		
4018	矚	1	1	82BE	77AF	矚		
4019	矪	1	1	82BF	77F4	矪		
4023	礚	1	1	82C3	78D3	礚		
4024	礫	1	1	82C4	792E	礫		
4025	礪	3	3	82C5	7934	礪		
4028	稟	1	1	82C8	7980	稟		稟，異體字問題。
4031	杭	8	8	82CB	79D4	杭		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
4033	稽	1	1	82CD	7A2D	稽		
4034	穉	8	8	82CE	7A49	穉		
4035	穽	11	11	82CF	7A7D	穽		
4038	竈	12	12	82D2	7AC8	竈		
4039	竝	5	5	82D3	7ADA	竝		
4040	竝	3	3	82D4	7ADD	竝		
4041	竝	3	3	82D5	7AEA	竝		
4042	筍	2	2	82D6	7B0B	筍		
4044	筍	5	5	82D8	7B6F	筍		
4045	簾	3	3	82D9	7B92	簾		
4052		3	3	82E0			形竹虛八	簾，異體字問題。
4055	粃	9	9	82E3	7C83	粃		
4058	糴	3	3	82E6	7CC9	糴		
4059	糴	2	2	82E7	7CC7	糴		
4063	縻	5	5	82EB	7CF5	縻		
4070	縻	1	1	82F2	7DE4	縻		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
4077	罇	10	10	82F9	7F47	罇		
4082	耇	1	1	82FE	8008	耇		
4089	脇	2	2	8346	8107	脇		
4091	皐	5	5	8348	81EF	皐		
4103	茝	2	2	8354	831D	茝		
4104	莅	4	4	8355	8385	莅		蒞，異體字問題。
4105	菓	5	5	8356	83D3	菓		
4107	葱	2	2	8358	8471	葱		
4109	蒟	2	2	835A	84AD	蒟		
4111	蒔	18	18	835C	8493	蒔		
4114	藥	2	2	835F	8616	藥		
4115	藁	14	14	8360	85C1	藁		
4118	藥	9	9	8363	8602	藥		
4124	蚰	2	2	8369	8771	蚰		
4129	虬	7	7	836E	866C	虬		
4131	𪗇	2	2	8370	8842	𪗇		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
4139	袴	17	17	8378	88B4	袴		
4140	衽	4	4	8379	88B5	衽		
4142	褰	16	16	837B	8943	褰		
4145	羈	1	1	837E	898A	羈		Unicode 字型差異。
4147	莛	1	1	83A2	8480	莛		
4158	譌	1	1	83AD	8B4C	譌		Unicode 字型差異。
4160	謫	2	2	83AF	8B81	謫		
4165	獮	4	4	83B4	8C9B	獮		
4167	賁	5	5	83B6	8CEB	賁		
4172	蹙	2	2	83BB	8E77	蹙		
4173	𪗇	2	2	83BC	8EAD	𪗇		
4176	輒	6	6	83BF	8F19	輒		輒，異體字問題。
4178	輓	1	1	83C1	8F2D	輓		
4184	迹	1	1	83C7	8FF9	迹		跡，異體字問題。
4186	邈	4	4	83C9	9061	邈		
4188	邨	1	1	83CB	90A8	邨		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
4189	港	3	3	83CC	6DC3	港		
4190	邨	1	1	83CD	90C4	邨		
4193	鈎	5	5	83D0	920E	鈎		
4195	鉢	33	33	83D2	9262	鉢		
4198	鑛	2	2	83D5	945B	鑛		
4201	陲	2	2	83D8	967B	陲		
4203	隣	1	1	83DA	96A3	隣		鄰，異體字問題。
4207	霧	13	13	83DE	9736	霧		
4208	磧	1	1	83DF	9767	磧		
4212	韞	1	1	83E3	97B8	韞		
4213	韞	4	4	83E4	97BE	韞		
4216	韞	2	2	83E7	97C8	韞		
4217	韞	1	1	83E8	97E0	韞		
4218	韭	1	1	83E9	97EE	韭		
4219	頰	4	4	83EA	982E	頰		
4221	頰	7	7	83EC	9833	頰		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
4222	頤	7	7	83ED	9847	頤		
4223	顙	7	7	83EE	984B	顙		
4225	顛	9	9	83F0	9866	顛		
4226	飡	2	2	83F1	98E1	飡		
4236	瞞	1	1	83FB	77D2	瞞		
4237	髡	5	5	83FC	9AE0	髡		髡，異體字問題。
4242	鬪	43	43	8442	9B2D	鬪		
4243	鬪	1	1	8443	9B34	鬪		
4249	鰱	7	7	8449	9BFF	鰱		
4250	鰲	6	6	844A	9C0C	鰲		
4251	鰕	2	2	844B	9C15	鰕		
4254	鴟	7	7	844E	9D7E	鴟		
4259	麕	2	2	8453	9E95	麕		Unicode 字型差異。
4260	麋	8	8	8454	9E9E	麋		
4263	麴	1	1	8457	9EAF	麴		
4267	鼈	7	7	845B	9F08	鼈		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
4269	鼯	1	1	845D			鼠𪛗晏	
4270	鼯	2	2	845E	9F45	鼯		
4271	齧	1	1	845F	9F69	齧		
4278	亘	10	10	8466	4E98	亘		
4342	倮	1	1	84C8	502E	倮		
4348	倮	1	1	84CE	506C	倮		
4356	𪛗	3	3	84D6	976B	𪛗		
4359	棧	2	2	84D9	6827	棧		
4365	𪛗	1	1	84DF	4C65		制𪛗魚	
4387	𪛗	2	2	84F5	20E57		口𪛗哲	
4399	壺	2	2	8542	213D6		土𪛗壺	
4406	黠	1	1	8549	2A46E		黑𪛗南	
4493	𪛗	2	2	85C2	3E40		𪛗𪛗字	𪛗，異體字問題。
4495	𪛗	3	3	85C4	9EA8	𪛗		
4505	碯	1	1	85CE	25500		石𪛗𪛗	
4515	𪛗	4	4	85D8	546A	𪛗		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
4533	唵	2	2	85EA	552B	唵		
4546	鴟	8	8	85F7	9D76	鴟		
4554	唬	1	1	8640	55C1	唬		
4566	礪	5	5	864C	78F5	礪		
4575	鑑	1	1	8655	7E7F	鑑		
4579	隽	5	5	8659	347A		亻 厶 隹	
4605	橙	2	2	8673	6195	橙		
4607	狔	5	5	8675	3E60		犭 厶 屯	
4617	襍	1	1	86A1	4671		衤 厶 屬	
4625	敼	1	1	86A9	3A7B		危 厶 支	
4632	欒	1	1	86B0			薛 厶 木	
4704	犝	1	1	86F8	7271	犝		
4712	灋	2	2	8741	7059	灋		
4716	塢	2	2	8745	57FF	塢		
4735	壠	30	30	8758	58E0	壠		
4739	韞	1	1	875C	97DB	韞		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
4743	𢀇	1	1	8760	22C72		扌𠂔𠂔	
4750	𠂔	19	19	8767			𠂔	𠂔，校對問題，舊版漢籍錯字，可用程式全部取代。
4763	𠂔	1	1	8774	6586	𠂔		
4768	𠂔	2	2	8779	59AC	𠂔		𠂔，異體字問題。
4772	𠂔	2	2	877D	9BF6	𠂔		
4778	𠂔	4	4	87A5	97B4	𠂔		
4779	𠂔	6	6	87A6	97B2	𠂔		
4782	𠂔	1	1	87A9	97BA	𠂔		
4784	𠂔	9	9	87AB	9793	𠂔		
4785	𠂔	1	1	87AC	8D0B	𠂔		𠂔，異體字問題。
4791	𠂔	3	3	87B2	7F93	𠂔		
4792		1	1	87B3			𠂔𠂔𠂔	待造字。
4797	𠂔	2	2	87B8	57A1	𠂔		
4801	𠂔	1	1	87BC	5A2C	𠂔		
4804	𠂔	1	1	87BF	5E42	𠂔		
4806	𠂔	2	2	87C1	56B1	𠂔		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
4816	𠂇	2	2	87CB			𠂇	
4822	𠂇	1	1	87D1	5C29	𠂇		𠂇，異體字問題。
4836	𠂇	3	3	87DF	20D32		口 𠂇 𠂇	
4839	𠂇	3	3	87E2	6637	𠂇		
4856	𠂇	3	3	87F3	3CB2		𠂇 𠂇 毛	
4871	𠂇	1	1	8843	5B4F	𠂇		
4893	𠂇	1	1	8859	263EA		羊 𠂇 𠂇	
4904	𠂇	3	3	8864	5C05	𠂇		
4907	𠂇	5	5	8867	69D6	𠂇		
4908	𠂇	1	1	8868	5C4A	𠂇		𠂇，異體字問題。
4934	𠂇	2	2	88A4	9B09	𠂇		
4948	𠂇	2	2	88B2	2308F		𠂇 𠂇 斗	
4953	𠂇	2	2	88B7	22835		𠂇 𠂇 𠂇	
4990	𠂇	5	5	88DC	5DBB	𠂇		
5002	𠂇	3	3	88E8	4A6D		𠂇	𠂇，重碼問題，採用部件外字，不轉碼。
5008	𠂇	2	2	88EE	5E5E	𠂇		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
5013	汙	8	8	88F3			汙	
5031	聞	5	5	8946	95B4	聞		聞，異體字問題。
5034	弃	11	11	8949	5F03	弃		棄，異體字問題。
5041	獾	2	2	8950	26D67		犴 𤝵 狼	
5043	彊	3	3	8952	5F4D	彊		
5046	彝	1	1	8955	5F5C	彝		
5057	豕	1	1	8960	39FB		豕 𪚩 豕	
5066	瑣	1	1	8969	24A0F		王 𪚩 賈	瑣，異體字問題。
5072	扃	1	1	896F	6244	扃		扃，異體字問題。
5074	𪚩	1	1	8971	5C2B	𪚩		𪚩，異體字問題。
5075	豕	1	1	8972			豕	
5091	鄒	5	5	89A4	92E3	鄒		
5099	惣	4	4	89AC	60E3	惣		
5103	臆	1	1	89B0	43F0		月 𪚩 忍	
5111	旆	19	19	89B8	65BE	旆		旆，異體字問題。
5115	颶	2	2	89BC	4B03		風 𪚩 立	

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
5121		1	1	89C2			𠂇 丰 方 攴	𠂇 𠂇 敎，校對問題，舊版漢籍錯字，可用程式全部取代。
5126	𦵒	1	1	89C7	8B29	𦵒		
5133	𦵑	3	3	89CE	6199	𦵑		
5134	𦵒	1	1	89CF	619E	𦵒		
5144	𦵒	1	1	89D9	8920	𦵒		
5173	漱	1	1	89F6	6F44	漱		漱，異體字問題。
5176		1	1	89F9			氵 𠂇 冗	沉，異體字問題。
5241	鵞	1	1	8A7B	9D5E	鵞		鵞，異體字問題。
5251	𠂇	1	1	8AA7	53B7	𠂇		
5253		1	1	8AA9			𠂇 𠂇 𠂇 𠂇 𠂇 𠂇	𠂇，校對問題，舊版漢籍錯字，可用程式全部取代。
5264	獮	1	1	8AB4	7371	獮		
5269	擗	1	1	8AB9	6511	擗		
5288	淮	1	1	8ACC	23D82		氵 𠂇 往	
5289	𦵑	3	3	8ACD	6164	𦵑		𦵑，異體字問題。
5290	携	6	6	8ACE	643A	携		重碼問題。部分採用舊漢籍造字「携」，以「攜」字取代；部分採用部件外字「𠂇」。

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
5310	鳥	2	2	8AE2	8204	鳥		
5328		1	1	8AF4			目𠂔𠂔	眇，異體字問題。
5335	𧈧	2	2	8AFB	459F		亡𠂔虫	虻，異體字問題。
5336	睹	1	1	8AFC	668F	睹		Unicode 字型差異。
5338	暎	1	1	8AFE	668E	暎		
5343	蠱	5	5	8B44	27492		𧈧𠂔虫	
5344	磬	3	3	8B45	55F8	磬		
5360	鄺	1	1	8B55	913D	鄺		
5392		5	5	8B75			𠂔𠂔白𠂔心𠂔	忽，異體字問題。
5421		2	2	8BB4			𠂔𠂔𠂔𠂔𠂔說文𠂔	𠂔，重碼問題，採用部件外字，不轉碼。
5434		1	1	8BC1			𠂔𠂔𠂔白𠂔工𠂔	涅，異體字問題。
5439	栢	11	11	8BC6	6822	栢		
5442	栢	26	26	8BC9	67BF	栢		
5443	𧈧	2	2	8BCA	27FF6		足𠂔𠂔	
5454	𧈧	1	1	8BD5	68B8	𧈧		
5456		1	1	8BD7			韋𠂔𠂔未	𧈧，重碼問題，採用部件外

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
								字，不轉碼。
5482	璽	8	8	8BF1	8812	璽		
5483	眡	1	1	8BF2	7721	眡		
5484	楫	2	2	8BF3	6973	楫		
5488	霽	6	6	8BF7	4A34		雨𠂇對	
5509	駢	3	3	8C4D	9A00	駢		
5510	駢	2	2	8C4E	99CA	駢		
5512	𪗇	3	3	8C50	8C3A	𪗇		
5541	蠹	3	3	8C6D	8827	蠹		蠹，異體字問題。
5542	鑠	20	20	8C6E	93C1	鑠		
5543	𩺰	1	1	8C6F	9B9D	𩺰		
5549	藁	2	2	8C75	85F3	藁		
5555	迥	3	3	8C7B	9008	迥		
5558	𪗇	1	1	8C7E	9E84	𪗇		
5559	欸	12	12	8CA1	6B35	欸		
5564	𪗇	1	1	8CA6	6ABF	𪗇		
5570	𪗇	2	2	8CAC	6ABE	𪗇		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
5571	撫	3	3	8CAD	6A45	撫		
5572	埠	3	3	8CAE	57BE	埠		
5582	鞣	1	1	8CB8	97B5	鞣		
5584	慍	1	1	8CBA	611E	慍		
5587	歛	3	3	8CBD	6B58	歛		
5591	杞	2	2	8CC1	233CC		木𠩺巳	杞，異體字問題。
5621	汨	6	6	8CDF	6CEA	汨		淚，異體字問題。
5625	瀝	1	1	8CE3	3D8F		氵𠩺隱	
5643	涎	2	2	8CF5	6D8F	涎		
5654	籊	2	2	8D41	7C4B	籊		
5669	蛭	1	1	8D50	2722B		工𠩺虫	臨，重碼問題，採用部件外字，不轉碼。
5682	蠹	1	1	8D5D	881A	蠹		
5704	𧯛	12	12	8D73	6C0A	𧯛		
5707	棧	1	1	8D76	6830	棧		
5710	𥻗	2	2	8D79	7A06	𥻗		Unicode 字型差異。
5717	陜	3	3	8DA2	967F	陜		

編號	造字	字頻 (xml)	字頻 (txt)	Big5	Unicode	WinXP	構字式	備註
5719	麇	3	3	8DA4	4D25		鹿厶加	
5720	埴	1	1	8DA5	578D	埴		
5723	牂	1	1	8DA8	4367		牛厶羊	
5742	澁	9	9	8DBB	6F81	澁		
5744	萑	1	1	8DBD			萑厶又	
5746		13	13	8DBF			每厶系	繁，異體字問題。
5764	籩	10	10	8DD1	7C5E	籩		
5783	敦	1	1	8DE4	3A8D		𠂔厶攵	
5797	灑	5	5	8DF2	7054	灑		
5800	𨔵	6	6	8DF5	908C	𨔵		〇〇，重碼問題，採用部件外字，不轉碼。
5801		1	1	8DF6			𠂔厶羸	〇〇，重碼問題，採用部件外字，不轉碼。
5807	睥	2	2	8DFC	4079		目厶辟	𠂔，重碼問題，採用部件外字，不轉碼。
5808	賀	2	2	8DFD	27D4D		𠂔厶貝	𠂔，重碼問題，採用部件外字，不轉碼。

附件二、「全宋詩(節選).xml」手動取代表

編號	造字	取代	檔案原文摘錄	手動取代結果
5290	携	攜	<u>𢶏</u> 壺相勸醉方休	<u>攜</u> 壺相勸醉方休

附件三、「全宋詩(節選).xml」缺字外錯字取代表

檔案原文摘錄	手動取代結果
攀條 <u>臨</u> 收柔萼	攀條 <u>攀</u> 柔萼