

《涑水記聞》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
2008/5/7 陳維君 製作

一、《涑水記聞》一書(涑水記聞.xml)使用舊版造字 124 個，字頻 847 次，詳如附件一。這 124 個造字中，103 個可轉成 Windows XP 能顯示的字，字頻 603 次；另外 21 個字必須轉成構字式，字頻 244 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、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 893 的「峩」係「峨」字之異體，故以「峨」字取代。又編號 3966 的「烟」字，係「煙」字之簡體字，亦以標準字的「煙」字取代。

2、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 4242 的「鬪」字，Unicode 字型呈現為「鬪」，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 1946 的「嗚」字，Unicode 編碼為 3B19，由於 Windows XP 並不支援，仍須使用構字式「日 𠂔 為」。

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、《涑水記聞》造字分析表

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXp	構字式	備註
296	𠄎	5	5	FBEC			𠄎	
297	𠄎	6	6	FBED			𠄎	
500	𠄎	4	4	FD5C			𠄎	
572	𠄎	1	1	FDC6			𠄎	凶，異體字問題。
717	𠄎	1	1	FEBA			𠄎	
777	𠄎	1	1	FEF6			𠄎	
842	𠄎	3	3	8E78	3AB7		𠄎	
893	𠄎	1	1	8ECD	5CE9	𠄎		峨，異體字問題。
899	𠄎	1	1	8ED3	99DE	𠄎		
938	𠄎	1	1	8EFA	7215	𠄎		燮，異體字問題。
982	𠄎	2	2	8F67	80F7	𠄎		胸，異體字問題。
1076	𠄎	4	4	8FE7	5368	𠄎		Unicode 呈現差異。
1077	𠄎	2	2	8FE8	7A3E	𠄎		
1133	𠄎	16	16	9061	74CC	𠄎		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXp	構字式	備註
1146	甕	1	1	906E	7506	甕		
1271		4	4	914E			者△多	待造字。
1382	醬	1	1	91DF	3F66		灺△田	
1417	凌	22	22	9243	35AB		口△菱	
1544	絙	3	3	92E4	7D59	絙		
1565	絲	2	2	92F9	7DDC	絲		
1642	脉	1	1	93A9	8109	脉		脈，異體字問題。
1668	膈	1	1	93C3	81C8	膈		
1710	爍	1	1	93ED	3E01		火△業	
1867	菥	1	1	94ED			形△禾九△	
1885	蝨	1	1	9540	882D	蝨		
1945	詢	6	6	957C	8A7E	詢		
1946	暘	8	8	957D	3B19		日△為	
1981	讐	2	2	95C2	8B90	讐		讎，異體字問題。
2046	踪	1	1	9644	8E2A	踪		
2157	酌	1	1	96D5	9167	酌		
2339	床	7	7	97EE	221F2		广△禾	
2505	翻	2	2	98F7	98DC	翻		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXp	構字式	備註
2750	梘	2	2	9AB2	68FF	梘		
2834	攵	2	2	9B47	6535	攵		
2977	却	1	1	9BF8			去厶卩	卻，異體字問題。
3789	冲	37	37	8154	51B2	冲		沖，異體字問題。
3791	况	4	4	8156	51B5	况		況，異體字問題。
3797	凭	1	1	815C	51ED	凭		憑，異體字問題。
3799	刂	1	1	815E	5226	刂		
3805	勅	40	40	8164	52C5	勅		
3814	却	46	46	816D	5374	却		卻，異體字問題。
3817	廁	1	1	8170	53A0	廁		廁，異體字問題。
3819	廡	9	9	8172	53AE	廡		廡，異體字問題。
3824	叙	1	1	8177	53D9	叙		敘，異體字問題。
3826	叶	3	3	8179	53F6	叶		
3843	堦	2	2	81AC	5826	堦		
3846	堦	3	3	81AF	58FB	堦		Unicode 呈現差異。
3858	寃	1	1	81BB	5BC3	寃		冤，異體字問題。
3883	徧	6	6	81D4	5FA7	徧		
3891	慙	10	10	81DC	6159	慙		慚，異體字問題。

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXp	構字式	備註
3910	擘	1	1	81EF	64E5	擘		攬，異體字問題。
3916	敷	1	1	81F5	656D	敷		
3924	朞	1	1	81FD	671E	朞		期，異體字問題。
3933	槩	1	1	8247	6998	槩		
3935	槩	16	16	8249	69E9	槩		概，異體字問題。
3966	烟	1	1	8268	70DF	烟		煙，異體字問題。
3967	炳	1	1	8269	712B	炳		
3974	牀	1	1	8270	7240	牀		
3977	狗	1	1	8273	72E5	狗		
3996	疎	12	12	82A8	758E	疎		
4008	癯	2	2	82B4	766F	癯		
4013	盃	1	1	82B9	76CC	盃		
4015	眦	3	3	82BB	7726	眦		眦，異體字問題。
4016	着	8	8	82BC	7740	着		著，異體字問題。
4019	矻	1	1	82BF	77F4	矻		
4024	礮	1	1	82C4	792E	礮		
4034	穉	1	1	82CE	7A49	穉		穉，異體字問題。
4038	竈	2	2	82D2	7AC8	竈		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXp	構字式	備註
4040	竝	1	1	82D4	7ADD	竝		
4041	豎	2	2	82D5	7AEA	豎		豎，異體字問題。
4044	筋	1	1	82D8	7B6F	筋		
4067	綉	1	1	82EF	7D89	綉		繡，異體字問題。
4080	羣	73	73	82FC	7FA3	羣		群，異體字問題。
4081	翱	1	1	82FD	7FFA	翱		翱，異體字問題。
4085	耻	3	3	8342	803B	耻		
4089	脇	1	1	8346	8107	脇		脅，異體字問題。
4090	脚	7	7	8347	811A	脚		腳，異體字問題。
4109	葛	2	2	835A	84AD	葛		
4115	藁	4	4	8360	85C1	藁		
4132		162	162	8371			血𠂇𠂇	眾，異體字問題。
4134	衛	38	38	8373	885E	衛		衛，異體字問題。
4146	覩	6	6	83A1	89A9	覩		
4167	賚	1	1	83B6	8CEB	賚		
4174	躒	1	1	83BD	8EB6	躒		
4184	迹	11	11	83C7	8FF9	迹		
4213	鞞	1	1	83E4	97BE	鞞		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXp	構字式	備註
4237	髡	4	4	83FC	9AE0	髡		髡，異體字問題。
4240	髡	1	1	8440	9AF4	髡		
4242	鬪	23	23	8442	9B2D	鬪		Unicode 呈現差異。
4258	麿	1	1	8452	9E90	麿		
4262	麪	1	1	8456	9EAA	麪		
4268	叢	2	2	845C	9F17	叢		
4437	廩	1	1	8568	5EEA	廩		廩，異體字問題。
4460	勅	1	1	85A1	52D1	勅		
4469	勺	5	5	85AA	5304	勺		
4487		1	1	85BC			穴 宀 浸	寢，異體字問題。
4507	吳	48	48	85D0	5434	吳		吳，異體字問題。
4607	狫	1	1	8675	3E60		犭 厶 屯	
4725	敷	1	1	874E			𠂔 出 方 支 𠂔	
4768	妬	2	2	8779	59AC	妬		妒，異體字問題。
4805	瀆	1	1	87C0	51DF	瀆		瀆，異體字問題。
4822	尫	1	1	87D1	5C29	尫		尫，異體字問題。
4839	昱	10	10	87E2	6637	昱		
4878	宄	3	3	884A	5B82	宄		宄，異體字問題。

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXp	構字式	備註
4902	兗	10	10	8862	5156	兗		兗，異體字問題。
4904	尅	1	1	8864	5C05	尅		
5008	幞	4	4	88EE	5E5E	幞		
5028	昇	16	16	8943	662A	昇		
5193	號	1	1	8A4B	8662	號		
5290	携	3	3	8ACE	643A	携		
5300	鬪	1	1	8AD8	9B2A	鬪		
5437	并	1	1	8BC4	5E77	并		并，異體字問題。
5439	栢	1	1	8BC6	6822	栢		
5478	飴	3	3	8BED	4D77	飴		
5537	离	2	2	8C69	79BC	离		
5542	鑠	2	2	8C6E	93C1	鑠		
5543	養	1	1	8C6F	9B9D	養		養，異體字問題。
5549	藁	1	1	8C75	85F3	藁		
5552	弃	3	3	8C78	5F06	弃		
5568	开	1	1	8CAA	5F00	开		
5591	杞	12	12	8CC1	233CC		木叒巳	杞，異體字問題。
5702	厨	1	1	8D71	53A8	厨		廚，異體字問題。

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXp	構字式	備註
5704	𧯛	21	21	8D73	6C0A	𧯛		
5859	支	1	1	C6D2			支	