

《朝野僉載》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

2008/4/11 陳建安 製作

一、《朝野僉載》一書(朝野僉載.xml)使用舊版造字 146 個，字頻 606 次，詳如附件一。這 146 個造字中，122 個可轉成 Windows XP 能顯示的字，字頻 483 次；另外 24 個字必須轉成構字式，字頻 123 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 3800 的「杀𠂔」係「剝」字之異體，故以「剝」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 3978 的「猪」字，Unicode 字型呈現為「猪」，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 2432 的「𨮑」字，Unicode 編碼為 29350，由於 Windows XP 並不支援，仍須使用構字式「革𠂔登」。

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、《朝野僉載》造字分析表

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
296	𡗗	5	5	FBEC			𡗗	
297	𡗘	18	18	FBED			𡗘	
299	𡗙	1	1	FBEF			𡗙	
300	𡗚	1	1	FBF0			𡗚	
521	𡗛	2	2	FD71			𡗛	
634	𡗜	1	1	FE45			𡗜	
766	𡗝	4	4	FEEB			𡗝	
768	𡗞	2	2	FEED			𡗞	
799	𡗟	1	1	8E4D	9F5A	𡗟		
814	𡗠	3	3	8E5C	87C7	𡗠		
821	𡗡	1	1	8E63	7718	𡗡		
841	𡗢	1	1	8E77	7932	𡗢		
842	𡗣	3	3	8E78	3AB7		𡗣	
899	𡗤	6	6	8ED3	99DE	𡗤		
1024	𡗥	1	1	8FB3	8675	𡗥		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
1067	碁	3	3	8FDE	68CA	碁		碁，異體字問題。
1091	塹	1	1	8FF6	58CD	塹		
1109	璫	1	1	9049	7496	璫		
1133	瓊	10	10	9061	74CC	瓊		
1187	蓋	1	1	90B9	76D6	蓋		蓋，異體字問題。
1226	罰	5	5	90E0	7F78	罰		
1237	稍	1	1	90EB	77DF	稍		
1267	櫛	2	2	914A	6B1B	櫛		
1278	碍	1	1	9155	788D	碍		
1530	紂	2	2	92D6	7CFA	紂		
1565	繇	2	2	92F9	7DDC	繇		
1871	蟲	1	1	94F1	87C1	蟲		
1893	袞	1	1	9548	886E	袞		
1981	讐	1	1	95C2	8B90	讐		
2021	趲	1	1	95EA	8D82	趲		
2046	踪	2	2	9644	8E2A	踪		
2147	蟥	2	2	96CB	87EE	蟥		
2157	酎	2	2	96D5	9167	酎		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
2197	鉄	1	1	96FD	9244	鉄		鐵，異體字問題。
2232		1	1	9761			𠂔厂墨卩𠂔	鄺，異體字問題。
2402	勳	3	3	986E	52EE	勳		
2432	鞞	1	1	98AE	29350		革𠂔登	
2454	鍬	3	3	98C4	28AFC		金𠂔教	
2494	麻需	3	3	98EC	2A3B1		麻𠂔需	
2520	饒	2	2	9947	993B	饒		
2578	勛	2	2	99A3	20892		𠂔角比力𠂔	
3035	賔	2	2	9C73	8CD4	賔		賓，異體字問題。
3160	坟	1	1	9D53	575F	坟		
3199	駁	1	1	9D7A	9ABD	駁		
3789	冲	2	2	8154	51B2	冲		沖，異體字問題。
3794	湊	1	1	8159	51D1	湊		湊，異體字問題。
3800		2	2	815F			杀𠂔𠂔	刹，異體字問題。
3801	剗	1	1	8160	5257	剗		
3802	劓	1	1	8161	529A	劓		
3804	効	4	4	8163	52B9	効		
3805	勅	15	15	8164	52C5	勅		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
3814	却	33	33	816D	5374	却		
3817	廁	6	6	8170	53A0	廁		廁，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		
3840	坂	4	4	81A9	5742	坂		
3843	堦	3	3	81AC	5826	堦		
3846	墻	7	7	81AF	58FB	墻		Unicode 呈現差異。
3849	獎	4	4	81B2	596C	獎		獎，異體字問題。
3850	姪	1	1	81B3	59D9	姪		
3858	寃	4	4	81BB	5BC3	寃		冤，異體字問題。
3864	峯	2	2	81C1	5CEF	峯		峰，異體字問題。
3876	廻	7	7	81CD	5EFB	廻		迴，異體字問題。
3883	徧	2	2	81D4	5FA7	徧		
3921	暫	6	6	81FA	6673	暫		
3925	寻	6	6	81FE	3775		寻	
3931	椀	5	5	8245	6900	椀		
3935	槩	6	6	8249	69E9	槩		概，異體字問題。
3942	毡	1	1	8250	6BE1	毡		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
3949	涖	1	1	8257	6D96	涖		
3966	烟	6	6	8268	70DF	烟		煙，異體字問題。
3973		2	2	826F			𠂇 𠂇扁	編，異體字問題。
3974	牀	4	4	8270	7240	牀		
3976	犁	5	5	8272	7282	犁		
3978	猪	18	18	8274	732A	猪		Unicode 呈現差異。
3979	猫	15	15	8275	732B	猫		
3984	瑠	1	1	827A	7460	瑠		
3989	璿	1	1	82A1	74C8	璿		
3996	疎	2	2	82A8	758E	疎		
4013	盃	2	2	82B9	76CC	盃		
4016	着	36	36	82BC	7740	着		著，異體字問題。
4025	礪	1	1	82C5	7934	礪		
4028	稟	3	3	82C8	7980	稟		稟，異體字問題。
4041	豎	1	1	82D5	7AEA	豎		
4044	筋	2	2	82D8	7B6F	筋		
4069	綑	1	1	82F1	7DB3	綑		
4080	羣	20	20	82FC	7FA3	羣		群，異體字問題。

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
4085	耻	4	4	8342	803B	耻		
4089	脇	3	3	8346	8107	脇		
4090	脚	15	15	8347	811A	脚		
4129	虬	1	1	836E	866C	虬		
4132		38	38	8371			血𩇛𩇛	眾，異體字問題。
4134	衛	3	3	8373	885E	衛		衛，異體字問題。
4139	袴	1	1	8378	88B4	袴		
4158	譌	11	11	83AD	8B4C	譌		Unicode 呈現差異。
4160	適	1	1	83AF	8B81	適		
4167	賁	1	1	83B6	8CEB	賁		
4169	趲	1	1	83B8	8DA6	趲		
4193	鈎	5	5	83D0	920E	鈎		
4195	鉢	4	4	83D2	9262	鉢		
4197	盞	2	2	83D4	936B	盞		
4213	鞞	10	10	83E4	97BE	鞞		
4221	頰	1	1	83EC	9833	頰		
4222	頰	1	1	83ED	9847	頰		
4225	顚	1	1	83F0	9866	顚		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
4226	𪗇	1	1	83F1	98E1	𪗇		
4232	𪗈	1	1	83F7	96B2	𪗈		𪗈，異體字問題。
4256	鷄	9	9	8450	9DC4	鷄		
4260	𪗊	9	9	8454	9E9E	𪗊		
4267	𪗋	2	2	845B	9F08	𪗋		
4271	𪗌	3	3	845F	9F69	𪗌		
4278	𪗍	1	1	8466	4E98	𪗍		
4307	罵	11	11	84A5	99E1	罵		罵，異體字問題。
4309	𪗏	1	1	84A7	523C	𪗏		𪗏，異體字問題。
4428	𪗐	2	2	855F	8E70	𪗐		𪗐，異體字問題。
4493	𪗑	4	4	85C2	3E40		𪗑 𪗑字	𪗑，異體字問題。
4515	𪗒	12	12	85D8	546A	𪗒		
4546	𪗓	2	2	85F7	9D76	𪗓		
4575	𪗔	1	1	8655	7E7F	𪗔		
4607	𪗕	1	1	8675	3E60		𪗕 𪗕屯	
4768	𪗖	13	13	8779	59AC	𪗖		𪗖，異體字問題。
4778	𪗗	2	2	87A5	97B4	𪗗		
4904	𪗘	1	1	8864	5C05	𪗘		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
4996	斲	5	5	88E2	65B5	斲		
5008	幪	1	1	88EE	5E5E	幪		
5094	翯	4	4	89A7	26407		翯	
5153	瓚	1	1	89E2	74C6	瓚		
5176		10	10	89F9			氵 𠂔 冗	沉，異體字問題。
5218		8	8	8A64			𠂔 𠂔 覓、𠂔	寬，異體字問題。
5222	亶	4	4	8A68	4EB6	亶		
5239	擎	2	2	8A79	6486	擎		
5241	鵞	7	7	8A7B	9D5E	鵞		鵞，異體字問題。
5434	湮	1	1	8BC1	6E7C	湮		湮，異體字問題。
5439	栢	6	6	8BC6	6822	栢		
5466	倏	1	1	8BE1	5010	倏		Unicode 呈現差異。
5558	龕	4	4	8C7E	9E84	龕		
5559	欸	2	2	8CA1	6B35	欸		
5575	𡵓	2	2	8CB1	8A14	𡵓		
5583	殮	2	2	8CB9	98F1	殮		
5585	寓	1	1	8CBB	3762		宀 𠂔 禹	
5587	𡵓	1	1	8CBD	6B58	𡵓		

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	構字式	備註
5702	厨	1	1	8D71	53A8	厨		廚，異體字問題。
5704	𣎵	2	2	8D73	6C0A	𣎵		
5705	廐	1	1	8D74	5ED0	廐		廢，異體字問題。
5706	敘	2	2	8D75	654D	敘		敘，異體字問題。
5707	棧	2	2	8D76	6830	棧		