

《雞肋編》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
2008/7/9 陳維君 製作

一、《雞肋編》一書(雞肋編.xml)使用舊版造字 143 個，字頻 473 次，詳如附件一。這 143 個造字中，101 個可轉成 Windows XP 能顯示的字，字頻 318 次；另外 42 個字必須轉成構字式，字頻 155 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 4132 的「血𠄎𠄎」係「眾」字之異體，故以「眾」字取代。又編號 1513 的「糧」字，係「糧」字之簡體字，亦以標準字的「糧」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 4450 的「翔」字，Unicode 字型呈現為「翔」，實際上仍為同一字。
- 3、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 2244 的「ノ𠄎虫」。

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 1277 的「璣」字，Unicode 編碼為 24AC9，由於 Windows XP 並不支援，仍須使用構字式「王𠄎燮」。

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、《雞肋編》造字分析表

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
38		2	2	FA65			王𠂇己	王𠂇己，異體字問題。
296	𠂇	7	7	FBEC			𠂇	
297	𠂇	32	32	FBED			𠂇	
299	𠂇	1	1	FBEF			𠂇	
300	𠂇	1	1	FBF0			𠂇	
314	𠂇	1	1	FBFE			𠂇	
374	𠂇	5	5	FC7B			𠂇	
549	𠂇	5	5	FDAF			𠂇	
658	𠂇	1	1	FE5D			𠂇	
675	𠂇	4	4	FE6E			𠂇	
721	𠂇	1	1	FEBE			𠂇	
752	𠂇	1	1	FEDD			𠂇	
755	𠂇	1	1	FEE0			𠂇	
763	𠂇	2	2	FEE8			𠂇	

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
770	扌	2	2	FEEF			扌	
792	帛	1	1	8E46	5E2C	帛		
1067	碁	4	4	8FDE	68CA	碁		棋，異體字問題。
1146	甕	1	1	906E	7506	甕		
1261	庸	3	3	9144	5EAF	庸		
1277	璽	4	4	9154	24AC9		王 璽	
1310	蠲	4	4	9175	45EB		虫 麻	
1381	穉	2	2	91DE	7A3A	穉		
1416	埃	2	2	9242	7AE2	埃		
1433	窆	3	3	9253	53DC	窆		
1513	糧	1	1	92C5	7CAE	糧		糧，異體字問題。
1544	絇	1	1	92E4	7D59	絇		
1565	繇	9	9	92F9	7DDC	繇		
1642	脉	4	4	93A9	8109	脉		
1841	塿	3	3	94D3	5896	塿		
1893	袞	2	2	9548	886E	袞		袞，異體字問題。
2046	踪	1	1	9644	8E2A	踪		
2168	疇	2	2	96E0	91BB	疇		


編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
2244		1	1	976D			ノ厶虫	待造字。
2278	銜	1	1	97B1	936E	銜		
2298	殼	1	1	97C5	3C89		殼△卵	
2331	槐	1	1	97E6	3BB0		木△毘	
2381	礪	2	2	9859	255EC		石△壘	
2408	鬧	1	1	9874	9599	鬧		鬧，異體字問題。
2421	傷	4	4	98A3	3465		亻△易	
2520	饊	1	1	9947	993B	饊		
2592	贓	3	3	99B1	8D1C	贓		
2596	辮	1	1	99B5	3869		辮△巾	
2622		1	1	99CF			王△恭	待造字。
3021	卜	1	1	9C65			卜	
3770	个	1	1	8141	4E2A	个		個，異體字問題。
3789	冲	1	1	8154	51B2	冲		沖，異體字問題。
3791	况	4	4	8156	51B5	况		況，異體字問題。
3792	凉	1	1	8157	51C9	凉		涼，異體字問題。
3801	剗	2	2	8160	5257	剗		
3814	却	4	4	816D	5374	却		卻，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
3817	廁	1	1	8170	53A0	廁		廁，異體字問題。
3819	廡	1	1	8172	53AE	廡		廡，異體字問題。
3844	城	2	2	81AD	5884	城		
3846	墻	8	8	81AF	58FB	墻		Unicode 呈現差異。
3850	姪	1	1	81B3	59D9	姪		
3858	寃	2	2	81BB	5BC3	寃		寃，異體字問題。
3864	峯	2	2	81C1	5CEF	峯		峰，異體字問題。
3883	徧	9	9	81D4	5FA7	徧		
3909	擡	1	1	81EE	64E1	擡		
3923		1	1	81FC			晶△宜	疊，異體字問題。
3927	柏	1	1	8241	6855	柏		
3935	槩	1	1	8249	69E9	槩		概，異體字問題。
3968	煨	1	1	826A	7145	煨		
3970	煊	1	1	826C	714F	煊		
3974	牀	7	7	8270	7240	牀		
3987	穀	1	1	827D	7474	穀		穀△土，異體字問題。
3990	甄	7	7	82A2	750E	甄		
4007	癩	1	1	82B3	7667	癩		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
4008	癯	3	3	82B4	766F	癯		
4013	盃	3	3	82B9	76CC	盃		
4016	着	2	2	82BC	7740	着		著，異體字問題。
4021	礪	2	2	82C1	78AF	礪		
4024	礪	1	1	82C4	792E	礪		
4033	稽	1	1	82CD	7A2D	稽		
4038	竈	1	1	82D2	7AC8	竈		
4041	豎	1	1	82D5	7AEA	豎		
4042	笋	1	1	82D6	7B0B	笋		
4044	筍	1	1	82D8	7B6F	筍		
4068	綫	6	6	82F0	7DAB	綫		線，異體字問題。
4073	纒	1	1	82F5	7E6E	纒		
4080	羣	16	16	82FC	7FA3	羣		群，異體字問題。
4081	翱	1	1	82FD	7FFA	翱		
4085	耻	1	1	8342	803B	耻		
4090	脚	5	5	8347	811A	脚		腳，異體字問題。
4105	菓	1	1	8356	83D3	菓		
4114	藥	1	1	835F	8616	藥		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
4118	藜	2	2	8363	8602	藜		
4124	蝨	1	1	8369	8771	蝨		
4132		48	48	8371			血𠂇𠂇	眾，異體字問題。
4134	衛	11	11	8373	885E	衛		衛，異體字問題。
4139	袴	7	7	8378	88B4	袴		
4142	褰	9	9	837B	8943	褰		
4149	覩	1	1	83A4	89BC	覩		
4167	賈	1	1	83B6	8CEB	賈		
4178	輓	3	3	83C1	8F2D	輓		
4184	迹	8	8	83C7	8FF9	迹		
4202	隲	2	2	83D9	9696	隲		
4213	鞞	1	1	83E4	97BE	鞞		
4232	隲	1	1	83F7	96B2	隲		隲，異體字問題。
4237	髡	2	2	83FC	9AE0	髡		髡，異體字問題。
4242	鬪	7	7	8442	9B2D	鬪		Unicode 呈現差異。
4251	鰕	1	1	844B	9C15	鰕		
4260	麋	3	3	8454	9E9E	麋		
4262	麪	13	13	8456	9EAA	麪		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
4267	鼈	1	1	845B	9F08	鼈		
4268	叢	1	1	845C	9F17	叢		
4278	亘	1	1	8466	4E98	亘		
4307	罵	1	1	84A5	99E1	罵		罵，異體字問題。
4383	櫛	1	1	84F1	3C1A		木△離	
4450	勑	5	5	8575	524F	勑		Unicode 呈現差異。
4461	勑	2	2	85A2	52D4	勑		
4507	吳	39	39	85D0	5434	吳		吳，異體字問題。
4568	衿	1	1	864E	4633		衤△金	
4618	鏤	1	1	86A2	4956		金△奠	
4649	賈	1	1	86C1	8552	賈		
4702	場	1	1	86F6	5872	場		
4750	卅	2	2	8767			卅	卅，異體字問題。
4762	旃	1	1	8773	3ACB		舛△冉	
4790	賺	2	2	87B1	81C1	賺		
4836	嗜	1	1	87DF	20D32		口△沓	
4878	冗	1	1	884A	5B82	冗		冗，異體字問題。
4902	兗	1	1	8862	5156	兗		兗，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
4921	曬	1	1	8875	77D6	曬		
4940	韻	10	10	88AA	97F5	韻		
5025		1	1	8940			千△干	𠂇，異體字問題。
5176		2	2	89F9			氵△冗	沉，異體字問題。
5198	甌	1	1	8A50	3F3E		鹿△瓦	
5218		1	1	8A64			𠂇△覓、 	寬，異體字問題。
5281	餓	2	2	8AC5	994A	餓		
5290	携	4	4	8ACE	643A	携		
5338	暎	2	2	8AFE	668E	暎		
5442	柈	1	1	8BC9	67BF	柈		
5476	椿	1	1	8BEB	3B7C		木△沓	
5503		1	1	8C47			木△茶	待造字。
5559	欸	1	1	8CA1	6B35	欸		
5582	鞣	1	1	8CB8	97B5	鞣		
5585	寓	3	3	8CBB	3762		宀△禹	
5590	朶	3	3	8CC0	6736	朶		朶，異體字問題。
5591	杞	1	1	8CC1	233CC		木△巳	杞，異體字問題。
5645	軼	2	2	8CF7	4836		軼	

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXp	構字式	備註
5706	敘	9	9	8D75	654D	敘		敘，異體字問題。
5729	滌	1	1	8DAE	6F34	滌		
5854	么	1	1	C6CD	5E7A	么		