

# 《元宮詞百章箋注》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  
2008/5/20 葉冠孜 製作

一、《元宮詞百章箋注》一書(元宮詞百章箋注.xml)使用舊版造字 97 個,字頻 426 次,詳如附件一。這 97 個造字中,79 個可轉成 Windows XP 能顯示的字,字頻 374 次;另外 18 個字必須轉成構字式,字頻 52 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 3276 的「女𠃉耍」字,原字為「女𠃉栗」。
- 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 3864 的「峯」係「峰」字之異體,故以「峰」字取代。
- 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 4259 的「麇」字,Unicode 字型呈現為「麇」,實際上仍為同一字。
- 4、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 1729 的「火𠃉動」。

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 2660 的「𪛗」字，Unicode 編碼為 3ECB，由於 Windows XP 並不支援，仍須使用構字式「王𪛗束」。

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、《元宮詞百章箋注》造字分析表

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
60	碩	2	2	FA7B	78BD	碩		
296	𪛗	1	1	FBEC			𪛗	
297	𪛗	5	5	FBED			𪛗	
771	𪛗	1	1	FEF0			𪛗	
911	鶯	3	3	8EDF	9D1C	鶯		
979	蕘	1	1	8F64	56CF	蕘		
1240		1	1	90EE			𪛗革卅𪛗伐𪛗	襪，異體字問題。
1467	纂	2	2	9275	25C67		𪛗𪛗秦	
1513	糧	1	1	92C5	7CAE	糧		糧，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
1545	媿	1	1	92E5	372E		女厶監	
1641	胆	2	2	93A8	80C6	胆		膽，異體字問題。
1685	鵠	4	4	93D4	4CD3		老厶鳥	
1729		4	4	9441			火厶動	待造字。
1735	樁	3	3	9447	3C02		木厶審	
1797		1	1	94A7			彳厶遷	待造字。
2021	趁	1	1	95EA	8D82	趁		
2278	鋤	1	1	97B1	936E	鋤		
2408	鬧	2	2	9874	9599	鬧		鬧，異體字問題。
2415	鬪	1	1	987B	95D8	鬪		鬥，異體字問題。
2660	球	1	1	99F5	3ECB		王厶東	
2663	巉	20	20	99F8	5DCE	巉		
3264	梦	1	1	9DDD	68A6	梦		夢，異體字問題。
3276		6	6	9DE9			女厶耍	女厶耍，舊漢籍錯字，可用程式全部取代。
3780	儻	1	1	814B	5101	儻		
3789	冲	1	1	8154	51B2	冲		沖，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3791	况	1	1	8156	51B5	况		况，異體字問題。
3792	凉	26	26	8157	51C9	凉		凉，異體字問題。
3801	剗	1	1	8160	5257	剗		
3805	勅	11	11	8164	52C5	勅		敕，異體字問題。
3814	却	11	11	816D	5374	却		卻，異體字問題。
3819	廡	2	2	8172	53AE	廡		廡，異體字問題。
3824	叙	4	4	8177	53D9	叙		敘，異體字問題。
3825	疊	3	3	8178	53E0	疊		疊，異體字問題。
3829	咏	2	2	817C	548F	咏		
3843	堦	4	4	81AC	5826	堦		
3859	尔	5	5	81BC	5C14	尔		
3864	峯	14	14	81C1	5CEF	峯		峰，異體字問題。
3892	睭	2	2	81DD	3B0B		日 厶 侯	
3899	担	2	2	81E4	62C5	担		擔，異體字問題。
3909	擡	1	1	81EE	64E1	擡		
3932	櫻	8	8	8246	6936	櫻		
3966	烟	5	5	8268	70DF	烟		煙，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
3971	熇	2	2	826D	71BA	熇		
3974	牀	4	4	8270	7240	牀		
3978	猪	1	1	8274	732A	猪		豬，異體字問題。
4014	蓋	1	1	82BA	8462	蓋		蓋，異體字問題。
4016	着	17	17	82BC	7740	着		著，異體字問題。
4041	豎	1	1	82D5	7AEA	豎		豎，異體字問題。
4067	綉	13	13	82EF	7D89	綉		繡，異體字問題。
4068	綫	1	1	82F0	7DAB	綫		線，異體字問題。
4080	羣	15	15	82FC	7FA3	羣		群，異體字問題。
4090	脚	5	5	8347	811A	脚		腳，異體字問題。
4094	館	1	1	834B	8218	館		館，異體字問題。
4107	葱	3	3	8358	8471	葱		
4109	葛	2	2	835A	84AD	葛		
4129	虬	3	3	836E	866C	虬		
4132		16	16	8371			血𠄎𠄎	眾，異體字問題。
4134	衛	8	8	8373	885E	衛		衛，異體字問題。
4193	鈎	2	2	83D0	920E	鈎		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4195	鉢	45	45	83D2	9262	鉢		
4202	鴟	1	1	83D9	9696	鴟		
4216		1	1	83E7	97C8	鞮		襪，異體字問題。
4218	菲	3	3	83E9	97EE	菲		
4242	鬪	3	3	8442	9B2D	鬪		Unicode 呈現差異。
4254	鷓	1	1	844E	9D7E	鷓		
4256	鷄	11	11	8450	9DC4	鷄		雞，異體字問題。
4259	麇	2	2	8453	9E95	麇		Unicode 呈現差異。
4262		1	1	8456	9EAA	麪		麵，異體字問題。
4265	鼃	3	3	8459	9F03	鼃		
4277		1	1	8465			金𠩺义	釵，異體字問題。
4278	亘	3	3	8466	4E98	亘		
4294	仔	2	2	8476	4F03	仔		
4307	罵	1	1	84A5	99E1	罵		罵，異體字問題。
4381	爾	1	1	84EF	511E	爾		
4507	吳	9	9	85D0	5434	吳		吳，異體字問題。
4558	厩	2	2	8644	53A9	厩		廄，異體字問題。

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
4686	埤	1	1	86E6	3654		土厶韋	
4702	場	1	1	86F6	5872	場		
4768	妬	1	1	8779	59AC	妬		妒，異體字問題。
4779	鞣	1	1	87A6	97B2	鞣		
4805	瀆	1	1	87C0	51DF	瀆		瀆，異體字問題。
4815	廂	1	1	87CA	53A2	廂		廂，異體字問題。
4866	鼃	2	2	87FD	9F04	鼃		
4870	壑	1	1	8842	97F2	壑		
4996	斲	2	2	88E2	65B5	斲		
5019	畧	1	1	88F9			畧	略，異體字問題。
5032	焄	1	1	8947	3B0E		焄	
5046	彝	1	1	8955	5F5C	彝		
5095	来	1	1	89A8	6765	来		來，異體字問題。
5290	携	2	2	8ACE	643A	携		攜，異體字問題。
5300	鬪	8	8	8AD8	9B2A	鬪		鬥，異體字問題。
5402	朮	4	4	8BA1	672F	朮		
5559	欸	2	2	8CA1	6B35	欸		

編號	造字	字頻 (txt)	字頻 (xml)	Big5	Unicode	WinXP	構字式	備註
5582	鞣	2	2	8CB8	97B5	鞣		
5590	朶	29	29	8CC0	6736	朶		朶，異體字問題。
5702	廚	9	9	8D71	53A8	廚		廚，異體字問題。
5742	澁	1	1	8DBB	6F81	澁		澁，異體字問題。