

# 《楚辭》舊版造字轉碼說明

中研院資訊所文獻處理實驗室  
中央研究院語言所文獻語料小組  
2010/07/14 丁致伶

轉碼主要工作是把檔案中的舊版造字轉換成 Windows XP 能支援的 Unicode 字形，Unicode 目前共收錄漢字 70194 個字，而 XP 只能支援 20902 個字（詳如表一），不支援之字將以構字式表達。例：造字編號 936 的「鷓」字，Unicode 編碼是 4D0A，由於 XP 並不支援，仍需使用構字式「義𪗇鳥」。

表一、Unicode 的字數及編碼區段

Unicode	新增字數	新增編碼區段	總字數	WinXP
1.1 版	20902	4E00-9FFF	20902	支援
3.0 版	6582	3400-4DFF	27484	不支援
3.1 版	42710	20000-2A6D6	70194	不支援

## 一、舊版造字轉碼分析：

《楚辭》使用舊版造字 137 個，字頻 497 次，這 137 個造字中，111 個可轉成 Windows XP 能顯示的字，字頻 378 次；另外 26 個字必須轉成構字式，字頻 119 次。

轉碼完成製作轉碼分析表，請參考附件一《楚辭》轉碼分析表，欄位說明如下：

- (一) 編號：Big5 造字空間為 6217 個，編號由 1 到 6217。
- (二) 造字：舊版造字。
- (三) 頻次：舊版造字在文件的出現次數。
- (四) Big5：造字的 Big5 碼。
- (五) Unicode：造字所對應的 Unicode 碼。

- (六) WinXP：造字在 Windows XP 的對應字形。
- (七) 構字式：Windows XP 無法對應字形改用構字式。
- (八) 備註凡例：備註欄中記錄轉碼後字形及修改原因，凡例如下：
1. 異體字問題：為了使用者查詢和使用的方便，在處理異體字時最主要的方式是以標準字取代，除非是專有名詞或特殊情形，如：人名、地名等。例：造字編號 4132 的「血𠂇𠂇」字，是「眾」的異體字，以標準字「眾」取代。
  2. 錯字，以程式全部取代：與原書字形不符，並查詢教育部異體字字典確認非異體字後，皆歸類為錯字，以程式取代為正確字形。例：檔案中使用造字編號 4684 的「衤𠂇𠂇」字，而原書使用之字為「衤」，查詢教育部異體字字典確認「衤」並非「衤𠂇𠂇」之異體字，因此歸為錯字，以程式取代為正確字形「衤」。
  3. Unicode 字型呈現差異：Unicode 字型與舊版造字有些微差異，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 4091 的「臯」字，Unicode 字型呈現為「臯」，實際上仍為同一字。

附件一、《楚辭》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
29	𧈧	4	FA5C	2295E		忄𠂇𧈧	
827	𩺰	6	8E69	9B8C	𩺰		
832	沔	3	8E6E	6CB2	沔		
834	𧈧	6	8E70	712D	𧈧		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
850	坑	1	8EA2			土厶元	
881	颺	1	8EC1	98C7	颺		
910	崧	1	8EDE	5D78	崧		
930	鱗	1	8EF2	9C45	鱗		
931	磕	6	8EF3	791A	磕		
932	兌	1	8EF4	7683	兌		
936	鷓	2	8EF8	4D0A		義厶鳥	
961	枅	1	8F52	3B4A		木厶片	
972	犇	3	8F5D	7287	犇		
977	磨	1	8F62	9E8F	磨		
1002	鳩	1	8F7B	9D02	鳩		
1013	紆	1	8FA8	7D5A	紆		
1024	虵	1	8FB3	8675	虵		
1039	稟	1	8FC2	69C0	稟		
1044	儻	1	8FC7	511B	儻		
1056	阤	1	8FD3	9641	阤		
1067	碁	1	8FDE	68CA	碁		
1068	雷	1	8FDF	4A13		雨厶回	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1074	璿	1	8FE5	7449	璿		
1079	總	4	8FEA	7DEB	總		
1095	岫	1	8FFA	37B6		山 厶 文	
1133	瓌	1	9061	74CC	瓌		
1198	籀	1	90C4	25DDA		籀	
1416	埃	6	9242	7AE2	埃		
1530	紕	1	92D6	7CFA	紕		
1565	繇	4	92F9	7DDC	繇		
1885	蠶	2	9540	882D	蠶		
1890	忝	1	9545	44D7		艹 厶 忝	
1900	桂	1	954F	88BF	桂		
1906	裒	1	9555	88E6	裒		
1937	詢	1	9574	8A3D	詢		
2081	餵	1	9667	9926	餵		
2162	醜	1	96DA	918E	醜		
2424	藥	10	98A6	6909	藥		
2472	靄	2	98D6	9741	靄		
2502	颺	1	98F4	98C8	颺		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
2609	鱻	1	99C2	9C4F	鱻		Unicode 字型 呈現差異。
2638	鷓	2	99DF	9D55	鷓		
2863		2	9B64			𠄎 友	𠄎，異體字問題。
2950	肱	1	9BDD	8104	肱		
3172	篔	1	9D5F	7B9F	篔		
3274	嶮	4	9DE7	5D2F	嶮		
3773	鳴	1	8144	4CDF		明 鳥	
3777	俵	1	8148	5008	俵		
3796	決	2	815B	51B3	決		決，異體字問題。
3804	効	3	8163	52B9	効		
3831	喆	1	817E	5586	喆		
3834	嗶	1	81A3	5637	嗶		Unicode 字型 呈現差異。
3835	噍	2	81A4	564D	噍		
3840	坂	1	81A9	5742	坂		
3860	尠	1	81BD	5C20	尠		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3883	徧	1	81D4	5FA7	徧		
3891	慙	2	81DC	6159	慙		
3897	慙	2	81E2	615C	慙		
3899	担	1	81E4	62C5	担		
3910	擘	6	81EF	64E5	擘		
3920	皓	2	81F9	6667	皓		
3933	槩	7	8247	6998	槩		
3935	槩	2	8249	3BA3	槩		Unicode 字型 呈現差異。
3974	牀	2	8270	7240	牀		
3980	獯	4	8276	7328	獯		
4012	鼓	1	82B8	76B7	鼓		
4034	穉	1	82CE	7A49	穉		
4039	竝	2	82D3	7ADA	竝		
4040	竝	12	82D4	7ADD	竝		
4052	簾	2	82E0	7C34	簾		Unicode 字型 呈現差異。
4062	齧	1	82EA	25F7F	齧		𠄎 口 艸 及 米 𠄎，異體字問

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							題。
4063	藥	1	82EB	7CF5	藥		
4070	縹	1	82F2	7DE4	縹		
4080	羣	28	82FC	7FA3	羣		
4081	翱	7	82FD	7FFA	翱		Unicode 字型 呈現差異。
4091	皐	10	8348	81EF	皐		Unicode 字型 呈現差異。
4103	苙	9	8354	831D	苙		
4115	藁	1	8360	85C1	藁		
4129	虬	11	836E	866C	虬		
4132		52	8371			血𠄎𠄎	眾，異體字問 題。
4134	衛	3	8373	885E	衛		衛，異體字問 題。
4136	衰	1	8375	887A	衰		
4142	裛	6	837B	8943	裛		
4146	覩	5	83A1	89A9	覩		
4148	覓	33	83A3	4C1F		云𠄎鬼	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4180	臯	2	83C3	8FA0	臯		
4184	迹	6	83C7	8FF9	迹		
4186	遡	1	83C9	9061	遡		
4203	隣	2	83DA	96A3	隣		
4222	頰	1	83ED	9847	頰		
4234	羸	2	83F9	9A58	羸		
4235	馮	1	83FA	9AA9	馮		
4240	髯	2	8440	9AF4	髯		
4242	鬪	2	8442	9B2D	鬪		
4247	鮎	1	8447	9B8E	鮎		
4254	鷓	3	844E	9D7E	鷓		
4256	鷄	3	8450	9DC4	鷄		
4265	鼃	1	8459	9F03	鼃		
4266	鼃	4	845A	9F02	鼃		
4267	鼃	2	845B	9F08	鼃		
4300	臙	1	847C	267A6		月 厶 隹	
4489		1	85BE			口 厶 方	号，異體字問題。
4507	吳	11	85D0	5434	吳		吳，異體字問



編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							題。
4530	咬	5	85E7	6648	咬		
4533	唸	2	85EA	552B	唸		
4562	逕	1	8648	28563		形之口 午	
4684		1	86E4			衤 彡 受	衤，錯字，以程式全部取代。
4727	壘	12	8750	58C4	壘		
4756	呂	51	876D	382F		呂	
4794	藁	1	87B5	8448	藁		
4858	閱	1	87F5	49A7		門 兒	
4970	繒	1	88C8	430C		纟 習	
5031	閨	1	8946	95B4	閨		
5034	弃	7	8949	5F03	弃		棄，異體字問題。
5048	婚	1	8957	36F0		女 昏	
5065	忼	2	8968	5FFC	忼		
5082	詠	1	8979	279BD		言 尔	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5108	惛	2	89B5	60FD	惛		
5271	嶮	1	8ABB	3801		山厶參	
5339	奸	1	8B40	59E7	奸		
5343	蠱	1	8B44	27492		夆厶虫	
5351	肫	1	8B4C	80AC	肫		
5353	役	1	8B4E	4F07	役		
5406	尻	1	8BA5	51E5	尻		
5434	湟	1	8BC1	6E7C	湟		
5459	缺	1	8BDA	21647		垂厶夂	
5463	岢	4	8BDE	65F9	岢		
5559	欸	2	8CA1	6B35	欸		
5578	弛	2	8CB4	38AE		弓厶屯	
5645	𨾏	7	8CF7	4836		𨾏	
5689	𨾏	1	8D64	20087		𨾏	
5706	敍	6	8D75	654D	敍		敍，異體字問題。
5717	陘	3	8DA2	967F	陘		
5724	稽	1	8DA9	7A38	稽		
5728	飢	1	8DAD	98E4	飢		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5759	寢	1	8DCC	21ACF		形宀 變口	
5808	賀	1	8DFD	27D4D		夂會貝	

## 二、手動取代表說明：

《楚辭》檔案中出現「●」字形，共計54字，比對原書後以Unicod字形或構字式手動取代，並製作「●」字手動取代表，請參考附件二。欄位說明如下：

1. 頁碼：「●」字形所在位置之原書頁碼。
2. 檔案原文摘錄：摘錄《楚辭》檔案有「●」字形的文句。
3. 原書字形：「●」在原書中之字形，以此字取代檔案原文的「●」。
4. 備註：記錄修改相同詞句的次數或其他事項。

### 附件二、《楚辭》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 17	●芙蓉以為裳	夬	
p. 35	●九州之博大	息	
p. 67	●芳椒兮	丩 采	
p. 75	鳴●兮吹竽	龠 虎	
p. 92	南北順●	隋 木	
p. 125	羹者而吹●兮	軟 韭	
p. 152	●母媯而	莫 女	

頁碼	檔案原文摘錄	原書字形	備註
p. 171	皆暖●其	暵	
p. 183	氣清，●●兮	宀△未 艸△翠	
p. 183	蟬●漠	宀△未	
p. 190	心尚●其弗濟	夭△并	
p. 192	四時●來	通	
p. 192	悅而日●兮	夭△并	
p. 195	恂●而自苦	愁	
p. 197	山澤，云△鬼●	白△鬼	
p. 198	●離散	白△鬼	
p. 198	對曰：「掌●	癆	
p. 200	●而得脫	季	
p. 201	秦●齊縷	竹△世△冉	
p. 216	憂●煩亂	息	
p. 216	青春受●	言△𠂔	
p. 217	湯谷●只	宀△未	
p. 220	不●嗑只	𠂔	
p. 220	清馨凍●	形入音欠□	
p. 237	上葳●而防露	形竹麥玉□	
p. 244	●母勃屑而日	莫△女	
p. 251	神●兩而無舍	冈△亡	

頁碼	檔案原文摘錄	原書字形	備註
p. 259	心紆●而增傷	車△尔	
p. 263	氣涸●其若	沸△鬲	
p. 266	生天●之若	形自象土□	
p. 270	天門兮●戶	形自象土□	
p. 271	寶遷兮份●	礪	
p. 278	過萬首兮●●	山△疑	
p. 282	楚懷之●連	形女○○△早□	
p. 283	淫暄而道●	靡△土	
p. 288	●湘流而南極	形汜△△△	
p. 292	航而●湘兮	汜△△	
p. 296	流俗。盪漚●	淩	
p. 299	從容於山●	慶	
p. 311	●揚舟於會稽	汜△△	
p. 315	忌●專兮郢	豈△否	
p. 315	歎兮氣●結	形食△勺吉□	
p. 317	蠱●兮號西	截△蝮	
p. 319	●靡兮成俗	飢	
p. 319	●縈兮青	△△△△	
p. 323	鹿蹊兮●●	足△斷 足△斷	
p. 323	鶉●兮甄甄	禽△鳥	

頁碼	檔案原文摘錄	原書字形	備註
p. 325	●蚩兮噍噍	虫厶尹	
p. 326	神光兮●●	𠃉上火頁𠃉 𠃉上火頁𠃉	改為頰頰
p. 326	●蜴兮進龜	蛛厶片	