

《樂府詩集》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
2009/03/09 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 959 的「揄」字，Unicode 編碼是 3A49，由於 XP 並不支援，仍需使用構字式「扌 Ⓐ 翁」。

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：

《樂府詩集》使用舊版造字 278 個，字頻 1979 次，這 278 個造字中，228 個可轉成 Windows XP 能顯示的字，字頻 1901 次；另外 50 個字必須轉成構字式，字頻 78 次。

轉碼完成製作轉碼分析表，請參考附件一《樂府詩集》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 4134 的「衛」字，是「衛」的異體字，以標準字「衛」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 2072 的「𠄎𠄎𠄎𠄎」字，而原書使用之字為「𠄎𠄎𠄎𠄎」，查詢教育部異體字字典確認「𠄎𠄎𠄎𠄎」並非「𠄎𠄎𠄎𠄎」之異體字，因此歸為錯字，以程式取代為正確字形「𠄎𠄎𠄎𠄎」。
 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 4145 的「羈」字，Unicode 字型呈現為「羈」，實際上仍為同一字。

附件一、《樂府詩集》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
820	烟	4	8E62	70F1	烟		
833	𡗗	1	8E6F	5C1E	𡗗		
842	𡗗	1	8E78	3AB7		𡗗	
869	𡗗	1	8EB5	6A5C	𡗗		
881	颺	34	8EC1	98C7	颺		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
912	鸚	1	8EE0	9DC0	鸚		
914	鸚	1	8EE2	2A23A		屬 𪗇 鳥	
943	燂	1	8F40	8975	燂		
947	燂	2	8F44	7217	燂		
959	揄	1	8F50	3A49		扌 𪗇 翁	
977	磨	2	8F62	9E8F	磨		
979	蕘	1	8F64	56CF	蕘		
1039	稟	1	8FC2	69C0	稟		
1076	高	1	8FE7	5368	高		Unicode 字型呈現差異。
1091	漶	1	8FF6	58CD	漶		
1133	瓌	3	9061	74CC	瓌		
1337	慌	1	91B2	385B		巾 𪗇 荒	
1355	袂	1	91C4			衤 𪗇 义	袂，異體字問題。
1490	籛	2	92AE	7C5D	籛		
1530	紉	1	92D6	7CFA	紉		
1547	𠂇	2	92E7	4E31	𠂇		
1600	猷	1	935D	3E84		犴 𪗇 俞	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
2278	銜	2	97B1	936E	銜		
2291	豐	1	97BE	27138		艹 厶 豐	
2298	殼	1	97C5	3C89		殼 厶 卯	
2309	棟	2	97D0	2333E		申 厶 東	
2331	槐	3	97E6	3BB0		木 厶 鬼	
2383	莘	4	985B	44A0		艹 厶 丰	
2386	瑯	1	985E	7458	瑯		
2396	醜	1	9868	28811		酉 厶 屯	
2408	鬧	2	9874	9599	鬧		
2432	鞞	1	98AE	29350		革 厶 登	
2479	鞞	1	98DD	979B	鞞		
2485	鯪	2	98E3	4C79		魚 厶 差	
2487	髀	2	98E5	29A49		骨 厶 孝	
2520	饊	1	9947	993B	饊		
2589	筴	2	99AE	7B21	筴		
2592	贓	1	99B1	8D1C	贓		
2598	鄣	1	99B7	9381	鄣		
2609	鱻	2	99C2	9C4F	鱻		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
2638	鷓	1	99DF	9D55	鷓		
2639	蓋	3	99E0	27794		衤 𠂔 蓋	
2650	鷓	2	99EB	9E18	鷓		
2651	鷓	8	99EC	9E1C	鷓		
2842	玃	1	9B4F	7FC4	玃		
3113	彡	3	9CE3	6245	彡		
3777	侏	1	8148	5008	侏		
3780	儻	1	814B	5101	儻		
3787	豕	7	8152	51A2	豕		
3788	慤	5	8153	22871		殼 忄 心	慤，異體字問題。
3789	冲	6	8154	51B2	冲		冲，異體字問題。
3791	况	3	8156	51B5	况		况，異體字問題。
3793	減	2	8158	51CF	減		減，異體字問題。
3794	湊	1	8159	51D1	湊		
3796	決	3	815B	51B3	決		決，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							題。
3801	剗	1	8160	5257	剗		
3804	効	14	8163	52B9	効		
3805	勅	1	8164	52C5	勅		
3814	却	119	816D	5374	却		
3817	廁	6	8170	53A0	廁		廁，異體字問題。
3819	廝	1	8172	53AE	廝		廝，異體字問題。
3824	叙	3	8177	53D9	叙		
3826	叶	11	8179	53F6	叶		
3829	咏	1	817C	548F	咏		
3830	咤	3	817D	54A4	咤		
3833	嘑	1	81A2	5611	嘑		
3835	噍	1	81A4	564D	噍		
3836	喻	2	81A5	564F	喻		
3840	坂	42	81A9	5742	坂		
3846	壻	25	81AF	58FB	壻		
3851	姤	4	81B4	59F9	姤		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3856	轆	1	81B9	283B9		車厶稟	
3858	寃	4	81BB	5BC3	寃		冤，異體字問題。
3863	岷	1	81C0	5C85	岷		
3866	崑	1	81C3	5D10	崑		
3868	卮	9	81C5	5DF5	卮		
3872	鵠	4	81C9	4CF5		勅厶鳥	
3874		2	81CB			忄厶稟	慄，異體字問題。
3877	迺	1	81CE	5EFC	迺		
3895	淑	2	81E0	6E86	淑		
3900	挖	2	81E5	62D5	挖		
3909	擡	1	81EE	64E1	擡		
3910	擘	3	81EF	64E5	擘		
3915	斲	1	8459	6585	斲		
3920	皓	1	81F9	6667	皓		
3921	皙	7	81FA	6673	皙		
3930	桴	1	8244	6901	桴		
3931	椀	8	8245	6900	椀		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3932	椶	3	8246	6936	椶		
3936	棹	2	824A	6A70	棹		
3937	檝	7	824B	6A9D	檝		
3946	汹	1	8254	6C79	汹		
3957	凜	1	825F	6F9F	凜		
3966	烟	246	8268	70DF	烟		
3974	牀	107	8270	7240	牀		
3979	猫	1	8275	732B	猫		
3980	獾	6	8276	7328	獾		
3981	珉	2	8277	73C9	珉		
3984	瑠	8	827A	7460	瑠		
4002	瘡	1	82AE	7607	瘡		
4011	皞	1	82B7	76A1	皞		皞，異體字問題。
4013	盃	6	82B9	76CC	盃		
4015	眦	2	82BB	7726	眦		
4016	着	18	82BC	7740	着		
4025	礪	1	82C5	7934	礪		
4028	稟	2	82C8	7980	稟		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4035	窳	4	82CF	7A7D	窳		
4038	竈	8	82D2	7AC8	竈		
4039	竝	6	82D3	7ADA	竝		
4041	豎	2	82D5	7AEA	豎		
4042	笋	2	82D6	7B0B	笋		
4043	筭	1	82D7	7B53	筭		
4044	筓	14	82D8	7B6F	筓		
4045	筴	4	82D9	7B92	筴		
4052	簾	3	82E0	7C34	簾		Unicode 字型 呈現差異。
4055	粃	1	82E3	7C83	粃		
4061	糲	2	82E9	7CE6	糲		
4063	糲	1	82EB	7CF5	糲		
4066	綺	2	82EE	7D5D	綺		
4067	綉	1	82EF	7D89	綉		
4069	綑	1	82F1	7DB3	綑		
4070	縹	2	82F2	7DE4	縹		
4080	羣	252	82FC	7FA3	羣		
4085	耻	4	8342	803B	耻		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4087	鮠	3	8344	4C47		魚𩺰旦	
4090	脚	14	8347	811A	脚		
4097	芑	1	834E			艹厶巳	
4100	菰	1	8351	82FD	菰		
4104	莅	3	8355	8385	莅		
4107	葱	20	8358	8471	葱		
4109	蒟	2	835A	84AD	蒟		
4111	蒨	1	835C	8493	蒨		
4114	藜	3	835F	8616	藜		
4115	藁	4	8360	85C1	藁		
4129	虬	12	836E	866C	虬		
4134	衛	93	8373	885E	衛		衛，異體字問題。
4137	袜	5	8376	889C	袜		
4139	袴	11	8378	88B4	袴		
4140	裱	3	8379	88B5	裱		
4145	羈	2	837E	898A	羈		Unicode 字型呈現差異。
4146	覩	28	83A1	89A9	覩		睹，異體字問

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							題。
4154	譽	7	83A9	8AD0	譽		
4158	譌	1	83AD	8B4C	譌		
4172	屨	1	83BB	8E77	屨		
4173	躄	2	83BC	8EAD	躄		
4184	迹	34	83C7	8FF9	迹		
4185	邊	2	83C8	9037	邊		
4186	遡	2	83C9	9061	遡		
4193	鈎	38	83D0	920E	鈎		鈎，異體字問題。
4195	鉢	2	83D2	9262	鉢		
4209	鞞	1	83E0	9789	鞞		
4221	頰	2	83EC	9833	頰		
4223	頤	2	83EE	984B	頤		
4237	髡	4	83FC	9AE0	髡		
4242	鬪	61	8442	9B2D	鬪		
4246	鑛	1	8446	28B97		金 𠩺 虞	
4251	緞	6	844B	9C15	緞		
4254	鷓	11	844E	9D7E	鷓		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4256	鷄	1	8450	9DC4	鷄		雞，異體字問題。
4260	麀	6	8454	9E9E	麀		
4267	鼈	3	845B	9F08	鼈		
4268	叢	4	845C	9F17	叢		
4269	鯤	1	845D	9F39	鯤		Unicode 字型呈現差異。
4278	亘	15	8466	4E98	亘		
4284		1	846C			形土厂 墨	土厶廛，異體字問題。
4296	堰	2	8478	5878	堰		
4307	罵	1	84A5	99E1	罵		罵，異體字問題。
4343	驥	2	84C9	4BC0		馬厶聶	
4344	鷓	1	84CA	9DE0	鷓		
4350	侶	2	84D0	5058	侶		
4356	鞞	1	84D6	976B	鞞		
4359	棧	2	84D9	6827	棧		
4384	鞞	1	84F2	97A7	鞞		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4388	縻	1	84F6	4328		纟 厶 蔡	
4394	苈	1	84FC	82BF	苈		
4428	躡	27	855F	8E70	躡		
4437	廩	1	8568	5EEA	廩		
4460	勅	7	85A1	52D1	勅		
4485	廛	1	85BA	3553		厂 厶 墨	廛，異體字問題。
4507	吳	60	85D0	5434	吳		吳，異體字問題。
4542	塿	1	85F3	21484		土 厶 稟	
4544	醅	3	85F5	288E7		酉 厶 齊	
4558	廐	5	8644	53A9	廐		廐，異體字問題。
4566	礪	1	864C	78F5	礪		
4637	吋	1	86B5	65F6	吋		
4680	臚	1	86E0	2680A		月 厶 品 止	
4681	𦉳	1	86E1	25859		奚 厶 禾	
4684		1	86E4			衤 厶 受	衤，錯字，以

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							程式全部取代。
4713	灑	1	8742	2410A		灑 灑 墨 灑	
4714	廛	1	8743	222E8		廛 廛 墨	廛，異體字問題。
4735	壠	3	8758	58E0	壠		
4744	絺	2	8761	25FC9		絺 絺 末	
4767	懷	1	8778	61B9	懷		
4768	妬	33	8779	59AC	妬		妒，異體字問題。
4778	鞞	1	87A5	97B4	鞞		
4779	鞞	1	87A6	97B2	鞞		
4806	噉	1	87C1	56B1	噉		
4807	媮	1	87C2	5A6E	媮		
4815	廂	1	87CA	53A2	廂		
4832	愴	2	87DB	61C0	愴		
4876	羴	2	8848	437D		羊 羴 歷	
4884		3	8850			躔 躔 厂	躔，異體字問題

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
						墨 \square	題。
4888	荔	1	8854	8318	荔		
4904	尅	5	8864	5C05	尅		
4925	輻	1	8879	8F65	輻		
4927	岝	1	887B	5C9E	岝		
4934	髮	2	88A4	9B09	髮		
4950	嶮	1	88B4	5D18	嶮		
4990	截	1	88DC	5DBB	截		
4996	斲	5	88E2	65B5	斲		
5064	帆	1	8967	5FDB	帆		
5065	忼	2	8968	5FFC	忼		
5066	璵	2	8969	24A0F		王 \triangle 貫	
5067	閨	1	896A	28D7D		門 \triangle 昏	
5072	扃	1	896F	6244	扃		
5105	埴	1	89B2	5860	埴		
5111	旆	13	89B8	65BE	旆		
5115	颶	1	89BC	4B03		風 \triangle 立	
5133	憇	1	89CE	6199	憇		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5144	構	1	89D9	8920	構		
5154	璿	2	89E3	7468	璿		
5162	嶷	1	89EB	37EA		山厶畏	
5163	憂	3	89EC	621E	憂		
5176		35	89F9			冫厶冗	沉，異體字問題。
5199		1	8A51			九厶虎	虢，異體字問題。
5211	駟	1	8A5D	99F5	駟		
5216	鞫	1	8A62	97CA	鞫		
5225	邨	2	8A6B	460F		血厶卩	
5263		1	8AB3			彳厶力	勞，異體字問題。
5290	携	3	8ACE	643A	携		
5298	昨	2	8AD6	65FF	昨		
5310	烏	1	8AE2	8204	烏		
5319	晰	10	8AEB	6663	晰		
5327	嶂	2	8AF3	5D2A	嶂		
5347	憾	2	8B48	617D	憾		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5355	𡗗	1	8B50	20AE4		𡗗	
5360	𡗘	2	8B55	913D	𡗘		
5437	𡗙	1	8BC4	5E77	𡗙		
5442	𡗚	8	8BC9	67BF	𡗚		
5452	𡗛	3	8BD3	8F5C	𡗛		
5459	𡗜	1	8BDA	21647		垂厶夂	
5466	𡗝	5	8BE1	5010	𡗝		
5541	𡗞	2	8C6D	8827	𡗞		
5542	𡗟	2	8C6E	93C1	𡗟		
5555	𡗠	1	8C7B	9008	𡗠		
5564	𡗡	1	8CA6	6ABF	𡗡		
5583	𡗢	1	8CB9	98F1	𡗢		
5590	𡗣	2	8CC0	6736	𡗣		
5643	𡗥	4	8CF5	6D8F	𡗥		
5653	𡗧	1	8D40	3B09		日厶𠂇	
5654	𡗨	2	8D41	7C4B	𡗨		
5690	𡗪	1	8D65	2260D		弗厶心	
5700	𡗬	1	8D6F	447D		舟厶𠂇	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5702	厨	4	8D71	53A8	厨		
5703	𪗇	1	8D72	9F53	𪗇		
5704	𪗈	5	8D73	6C0A	𪗈		𪗈，異體字問題。
5706	𪗉	33	8D75	654D	𪗉		𪗉，異體字問題。
5713	𪗊	2	8D7C	481E		足𪗊戚	
5719	麋	1	8DA4	4D25		鹿𪗋加	
5749	牖	2	8DC2	7257	牖		
5797	灑	2	8DF2	7054	灑		
5808	賀	1	8DFD	27D4D		𪗌𪗍貝	

二、手動取代表說明：

《樂府詩集》檔案中出現「●」字形，共計 99 字，比對原書後以 Unicode 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《樂府詩集》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、《樂府詩集》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 31	不黎不●	𠂔△酉	
p. 38	乾●詔百靈	𠂔	
p. 43	儀●下基	𠂔	
p. 57	或●或蔗	蓑	
p. 97	蔗●惟興	蓑	
p. 119	拂漢涵●	𠂔△膏	待造字
p. 200	左●回衿帶	衽	
p. 209	契●化	宀△日△六	改為冥
p. 228	蒲葦●●	宀△日△六 宀△日△六	改為冥冥
p. 243	中山浮羽●	卮	
p. 263	菱芒乍●絲	冈△冂	改為冑
p. 294	平●望烟合	冈△且	改為置
p. 304	屈●猛	𠂔虎武貝□	
p. 305	苞枿●矣	黑△對	待造字
p. 313	同前 車 ●	𠂔△文	有 2 處
p. 340	日起罌●外	冈△思	改為思
p. 341	百尺●金埒	日△敢	
p. 352	●多塵亂飛	足△甫	
p. 355	梁·車 ●	𠂔△文	
p. 355	意欲●●走	走△參	

頁碼	檔案原文摘錄	原書字形	備註
		走△覃	
p. 357	連蹠復連●	足△甫	
p. 362	利●女	羊△丘	
p. 363	●鉞鷓尾條	金△互	
p. 363	齊著鐵●鉞	金△互	
p. 370	●跋黃塵下	足△必	
p. 372	何處●觴來	舌△業	
p. 379	白馬●●赤	走△參 走△覃	
p. 383	確●砉砉	石△嵬	
p. 388	臨醉欲●驕	扌△弃	
p. 411	風瑟瑟，木●●	形木彡文□ 形木彡文□	待造字
p. 462	青●未垂孤鳳餓	竹△復	
p. 465	一作●★入居	隹	
p. 472	鬱金香●特香衣	白△為	
p. 479	秦涇含藥☆（●）	火△鳥	
p. 484	弓曉月逾明。●●	凜凜	
p. 493	谷望河陽縣，☆（●）	木△度	
p. 504	當學織女嫁●牛	牽	改為牽
p. 572	●土裏浴	土△孛	

頁碼	檔案原文摘錄	原書字形	備註
p. 576	●●踏沙人似鬼	趑趑	
p. 579	覽取為吾●	𠄎	
p. 600	●如☆如字下	齒 𠄎 玄	
p. 607	大人●岫當安之	山 𠄎 兒	
p. 618	坎●元淑賦	土 𠄎 廩	
p. 671	觀騎立蚺●	𠄎 虫 厂 豆 寸 𠄎	
p. 672	道逢播●郎	拏	
p. 695	船絲作●	糸 𠄎 乍	
p. 707	三山●行舟	隱	改為隱
p. 711	●機倚不織	隱	改為隱
p. 728	珠佩媠●戲金闕	女 𠄎 厄	
p. 804	夜天●●不見	燿燿	
p. 827	●娟好且修	女 𠄎 便	
p. 895	戚里●曾闕	遡	
p. 902	主人起舞●盤	沙 𠄎 皿	
p. 920	金樓旦●嶮	山 𠄎 蹇	
p. 933	香●雙珠環	衤 𠄎 登	
p. 956	珠落璀璨白●袍	𠄎 𠄎 剌	
p. 970	玄雲暎●合	雲 𠄎 尢	待造字
p. 974	膽●臟腑	撐	
p. 976	一作為★●☆	勺 𠄎 盍	

頁碼	檔案原文摘錄	原書字形	備註
p. 976	一作●★葉垂	勾	
p. 986	梁·車 ●	𨮒△文	
p. 991	一作●★娟	女△便	
p. 1012	突兀●	山△孝	
p. 1025	鯉魚鰭●赤	鱻	
p. 1063	生蔑☆ (●)	艹△阡	待造字
p. 1077	援綠☆ (●)	竹△戢	待造字
p. 1081	坐對鷓●嬌不語	鷓	
p. 1088	毳●●五木香	毛△荅 毛△登	
p. 1088	艾●及都梁	蒹	
p. 1096	●●胡馬蹄	趨起	
p. 1099	玉溜☆一作●	霈△凶	改為霈△田
p. 1179	據鞍長歎●	息	
p. 1185	銅鳴周國☆ (●)	旒	
p. 1209	●●沈湘語	口△祭 口△祭	
p. 1213	敕勒金☆ (●)	墳	
p. 1249	天公誅●汝	謫	
p. 1290	牛馬●腥草	齒△來	
p. 1312	牛●●	口△毛	

頁碼	檔案原文摘錄	原書字形	備註
		口𠂔毛	
p. 1315	變●撩機	糸𠂔聶	
p. 1320	嘈●琵琶青幕	噴	
p. 1322	大陵高岸失峇●	山𠂔堯	改為嶢
p. 1322	陟巒丘之●陀	山𠂔麗	待造字
p. 1358	蒺藜後●碗	危𠂔臬	
p. 1359	挑紋變●力	糸𠂔聶	
p. 1359	赤者為●騮	馬𠂔棗	待造字
p. 1382	藕絲蛛☆ (●)	蝸	