

《喻世明言》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
2009/11/26 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 1520 的「𠄎」字，Unicode 編碼是 3B20，由於 XP 並不支援，仍需使用構字式「日𠄎𠄎」。

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：

《喻世明言》使用舊版造字 168 個，字頻 3109 次，這 168 個造字中，152 個可轉成 Windows XP 能顯示的字，字頻 3068 次；另外 16 個字必須轉成構字式，字頻 41 次。

轉碼完成製作轉碼分析表，請參考附件一《喻世明言》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。

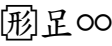
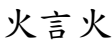

(八) 備註凡例：備註欄中記錄轉碼後字形及修改原因，凡例如下：

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 5472 的「瓦」字，而原書使用之字為「瓦」，查詢教育部異體字字典確認「瓦」並非「瓦」之異體字，因此歸為錯字，以程式取代為正確字形「瓦」。
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 3978 的「猪」字，Unicode 字型呈現為「猪」，實際上仍為同一字。

附件一、《喻世明言》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
20	藪	2	FA53			𪗇𪗇𪗇	
74	髻	2	FAAB	29B06		𪗇𪗇𪗇	
982	曾	1	8F67	80F7	曾		
988	芊	1	8F6D	7F8B	芊		
1076	高	4	8FE7	5368	高		Unicode 字型

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							呈現差異。
1154	跣	2	9076	47F6		足 厶 坐	
1170	瞳	1	90A8	7583	瞳		
1213	睽	2	90D3	7743	睽		
1278	碍	1	9155	788D	碍		
1281	儂	1	9158	349F		亻 厶 衰	
1394	啗	1	91EB	35D6		口 厶 昏	
1491	擲	8	92AF	6527	擲		
1513	糧	1	92C5	7CAE	糧		糧，異體字問題。
1520	煦	1	92CC	3B20		日 厶 煞	
1598	鞫	1	935B	4A7A		革 厶 翁	
1781	蓋	1	9475	854B	蓋		
1784	萼	2	9478	855A	萼		
1905	棍	1	9554	88E9	棍		
1926	覓	8	9569	8994	覓		
1981	讐	1	95C2	8B90	讐		讐，異體字問題。
2011	乾	3	95E0	4A50		革 厶 乞	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
2044	踈	4	9642	8E08	踈		
2046	踪	10	9644	8E2A	踪		
2053	躡	1	964B	8EA7	躡		
2147	蟾	8	96CB	87EE	蟾		
2157	酎	9	96D5	9167	酎		
2278	鎗	1	97B1	936E	鎗		
2292		1	97BF			  	躡，異體字問題。
2367	箠	1	984B	25C2D			
2408	鬧	3	9874	9599	鬧		鬧，異體字問題。
2592	贓	5	99B1	8D1C	贓		
2651	鸚	1	99EC	9E1C	鸚		
2999	嗵	2	9C4F	5523	嗵		
3789	冲	10	8154	51B2	冲		冲，異體字問題。
3791	况	2	8156	51B5	况		况，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3792	涼	1	8157	51C9	涼		涼，異體字問題。
3794	湊	18	8159	51D1	湊		湊，異體字問題。
3797	凭	3	815C	51ED	凭		
3801	剗	1	8160	5257	剗		
3802	劖	1	8161	529A	劖		
3803	鮓	11	8162	4B54		食 𩚑 追	
3804	効	8	8163	52B9	効		
3805	勅	3	8164	52C5	勅		
3814	却	546	816D	5374	却		
3815	𡗗	1	816E	537A	𡗗		
3816	厓	5	816F	5393	厓		
3818	廩	1	8171	53AB	廩		
3819	廡	56	8172	53AE	廡		廡，異體字問題。
3824	叙	1	8177	53D9	叙		
3829	咏	1	817C	548F	咏		
3830	咤	1	817D	54A4	咤		


編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3840	坂	1	81A9	5742	坂		
3843	堦	16	81AC	5826	堦		
3846	堦	45	81AF	58FB	堦		
3858	寃	83	81BB	5BC3	寃		
3864	峯	23	81C1	5CEF	峯		
3868	卮	1	81C5	5DF5	卮		
3876	廻	1	81CD	5EFB	廻		
3883	徧	2	81D4	5FA7	徧		
3886	悞	39	81D7	609E	悞		
3891	慙	3	81DC	6159	慙		
3909	擡	68	81EE	64E1	擡		
3931	椀	4	8245	6900	椀		
3934	桶	4	8248	69C5	桶		
3938	橈	9	824C	6AC8	橈		
3942	毡	2	8250	6BE1	毡		
3944	冰	1	8252	6C37	冰		
3966	烟	14	8268	70DF	烟		
3971	熇	8	826D	71BA	熇		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3974	牀	96	8270	7240	牀		
3975	牕	20	8271	7255	牕		
3976	犁	3	8272	7282	犁		
3978	猪	14	8274	732A	猪		Unicode 字型 呈現差異。
3979	猫	10	8275	732B	猫		
3991	攜	8	82A3	64D5	攜		
3996	疎	12	82A8	758E	疎		
4001	瘧	1	82AD	7602	瘧		
4012	鼓	1	82B8	76B7	鼓		鼓，異體字問 題。
4016	着	832	82BC	7740	着		
4028	稟	36	82C8	7980	稟		
4035	穽	1	82CF	7A7D	穽		
4036	窰	3	82D0	7AB0	窰		
4038	竈	15	82D2	7AC8	竈		
4039	竚	3	82D3	7ADA	竚		
4042	笋	1	82D6	7B0B	笋		
4043	筴	1	82D7	7B53	筴		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4044	筋	6	82D8	7B6F	筋		
4045	箒	4	82D9	7B92	箒		
4057	莖	19	82E5	585F	莖		
4068	綫	34	82F0	7DAB	綫		
4069	綑	1	82F1	7DB3	綑		
4074	縑	5	82F6	7E27	縑		
4078		2	82FA			已合併	异，異體字問題。
4080	羣	25	82FC	7FA3	羣		
4090	脚	144	8347	811A	脚		
4091	臯	9	8348	81EF	臯		
4094	館	4	834B	8218	館		
4105	菓	9	8356	83D3	菓		
4107	葱	1	8358	8471	葱		
4109	蒟	1	835A	84AD	蒟		
4112	蔴	8	835D	8534	蔴		
4115	藁	4	8360	85C1	藁		
4128	虿	1	836D	87DA	虿		
4129	虬	1	836E	866C	虬		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4132		140	8371			血𠂇𠂇	眾，異體字問題。
4134	衛	12	8373	885E	衛		衛，異體字問題。
4138	裘	7	8377	88A0	裘		
4139	袴	4	8378	88B4	袴		
4144	霸	30	837D	8987	霸		Unicode 字型呈現差異。
4145	羈	3	837E	898A	羈		Unicode 字型呈現差異。
4146	覩	6	83A1	89A9	覩		
4167	賚	4	83B6	8CEB	賚		
4172	屨	5	83BB	8E77	屨		
4173	𨔵	5	83BC	8EAD	𨔵		
4184	迹	11	83C7	8FF9	迹		
4193	鈎	6	83D0	920E	鈎		
4203	隣	54	83DA	96A3	隣		
4221	頰	1	83EC	9833	頰		
4231	駢	1	83F6	9A0C	駢		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4232	隲	5	83F7	96B2	隲		
4237	髡	2	83FC	9AE0	髡		
4242	鬪	27	8442	9B2D	鬪		
4256	鷄	9	8450	9DC4	鷄		
4260	麐	1	8454	9E9E	麐		
4278	亘	2	8466	4E98	亘		
4287	織	5	846F	7E8E	織		
4307	罵	12	84A5	99E1	罵		罵，異體字問題。
4309	刼	24	84A7	523C	刼		
4316	艷	1	84AE	8276	艷		
4432		39	8563			 	纏，異體字問題。
4447	掙	4	8572	63D1	掙		
4448		2	8573			 	  , 異體字問題。
4460	勅	3	85A1	52D1	勅		
4461	勅	1	85A2	52D4	勅		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4507	吳	2	85D0	5434	吳		吳，異體字問題。
4515	呪	12	85D8	546A	呪		
4518	响	2	85DB	54CD	响		
4522	瞞	6	85DF	3B13		日厶票	
4531	攢	1	85E8	6505	攢		
4542	塿	1	85F3	21484		土厶稟	
4762	旃	1	8773	3ACB		舛厶冉	
4768	妬	15	8779	59AC	妬		
4870	壘	1	8842	97F2	壘		
4872	褶	1	8844	276D5		衤厶冒	
4903	吋	3	8863	5C00	吋		
4904	尅	8	8864	5C05	尅		
5008	僕	2	88EE	5E5E	僕		
5111	旃	5	89B8	65BE	旃		
5147	恠	1	89DC	6060	恠		
5173	漱	2	89F6	6F44	漱		
5174		1	89F7			𠄎刀兂 、 	兔，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5290	携	6	8ACE	643A	携		
5437	并	65	8BC4	5E77	并		
5439	栢	3	8BC6	6822	栢		
5472	瓦	20	8BE7		瓦		瓦，錯字，以程式全部取代。
5551	費	8	8C77	8CF7	費		
5556	隸	5	8C7C	96B7	隸		
5590	朶	20	8CC0	6736	朶		
5591	杞	3	8CC1	233CC		木𠃉巳	杞，異體字問題。
5595	朶	1	8CC5	579C	朶		
5621	泪	3	8CDF	6CEA	泪		
5622	邳	5	8CE0	48B5		云𠃉卩	
5694	槩	3	8D69	3BA3	槩		Unicode 字型呈現差異。
5697	煮	3	8D6C	7151	煮		
5702	厨	7	8D71	53A8	厨		
5705	廐	1	8D74	5ED0	廐		Unicode 字型

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							呈現差異。
5706	敘	45	8D75	654D	敘		敘，異體字問題。
5797	灑	1	8DF2	7054	灑		

二、手動取代表說明：

《喻世明言》檔案中出現「●」字形，共計 88 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《喻世明言》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、《喻世明言》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 5	整日的不徹不●	係	
p. 20	赤條條的●	扌厶雙	
p. 47	却在那里●酒	疒	
p. 47	皮匠家●底	厶厶厶	
p. 47	你●得與他	扌厶弃	
p. 51	●一口氣	疒厶敝	

頁碼	檔案原文摘錄	原書字形	備註
p. 66	●●不得	門△爭 門△坐	
p. 69	容得這等塵●	齒△曹	
p. 80	便去偷情●院	闕	
p. 82	門外●喝道	么	
p. 96	便去●喫	噏	
p. 97	在門生家●酒	噏	
p. 98	沒半個人●他	係	
p. 136	皇甫●判度	王△專	待造字
p. 167	●單、去吉	〇△公甲	有 2 處
p. 167	吳矣、●台	呂	有 2 處 改為呂
p. 192	●身便跳	扌△雙	
p. 195	●喇一聲	唎	
p. 197	●身望下便跳	扌△雙	
p. 205	山下●草	乍△刂	
p. 219	賣●糜王公處	米△羨	待造字
p. 219	推折了門●	形木戶冘□	
p. 234	碧油●擁	扌△童	
p. 260	百姓中禿髮●髻	髻△賴	
p. 263	那里●他	係	

頁碼	檔案原文摘錄	原書字形	備註
p. 300	身材矯●	𠂔彳卜ヨ匕𠂔	待造字
p. 303	小●船來到	擗	
p. 303	●船上人應道	擗	
p. 304	●哨一聲	唵	
p. 371	官脚●行遲	𠂔厶廣	
p. 392	要●身子不起	門厶爭	
p. 407	若不●不賭	闕	
p. 469	名為「螿●」	蚰	
p. 478	●談便	咲	
p. 492	大伯伸着●項	孛厶頁	
p. 495	劍靶●在手裏	拈	
p. 515	賣鵲鶉●●	餽蝕	
p. 515	●●兒	餽蝕	有 4 處
p. 515	條篋黃穿那●	餽	
p. 515	●兒，捏些鹽	蝕	
p. 517	門來關上，●來	扌厶辰	
p. 517	●了，謊得僧兒	扌厶辰	
p. 518	扯出賣●●	餽蝕	
p. 518	并賣●●的僧兒	餽蝕	
p. 522	教賣●●的	餽蝕	
p. 523	●●性命	門厶爭	

頁碼	檔案原文摘錄	原書字形	備註
		門△坐	
p. 529	關●子	椵	有 3 處
p. 534	買一百錢●肉	火△麤	
p. 534	買了●肉	火△麤	
p. 534	買●肉共蒸餅	火△麤	
p. 535	留這●肉	火△麤	
p. 535	裹了那●肉	火△麤	
p. 535	解開●肉裹兒	火△麤	
p. 535	肥底●肉	火△麤	
p. 535	買的●肉裏面	火△麤	
p. 536	提一裹●肉	火△麤	
p. 542	房門上打一●	扌△寨	
p. 543	觀察一半衫●	衤△奚	
p. 552	小火赤●蛇	虫△連	
p. 552	●着高高的	門△爭	
p. 575	●塊磚兒	丿△羊	待造字
p. 585	●索長釘	擲	
p. 600	連放三箇大●	石△充	
p. 602	我●去送官	擲	
p. 603	乘箇小●船	擲	
p. 603	依舊●轉	擲	

頁碼	檔案原文摘錄	原書字形	備註
p. 609	董四背剪●起	擲	
p. 618	將他●頭	替△卩	
p. 631	推的推，●的	扌△雙	
p. 631	●，霎時間	扌△雙	
p. 631	飛●出城	透	