

《元刊雜劇三十種新校(下冊)》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
2009/12/10 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 1691 的「鬚」字，Unicode 編碼是 4BFC，由於 XP 並不支援，仍需使用構字式「鬚𠄎狄」。

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：

《元刊雜劇三十種新校(下冊)》使用舊版造字 31 個，字頻 58 次，這 31 個造字中，19 個可轉成 Windows XP 能顯示的字，字頻 31 次；另外 12 個字必須轉成構字式，字頻 27 次。

轉碼完成製作轉碼分析表，請參考附件一《元刊雜劇三十種新校(下冊)》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. 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 4067 的「綉」字，而原書使用之字為「氈」，以程式取代為正確字形「氈」。

附件一、《元刊雜劇三十種新校(下冊)》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
296	厶	5	FBEC			厶	
297	厶	10	FBED			厶	
298	厶	1	FBEE			厶	
658	疒	1	FE5D	7592		疒	
746	忄	2	FED7	5FC4		忄	
770	扌	1	FEEF	624C		扌	
1310	麻	1	9175	45EB		虫厶麻	
1491	擷	1	92AF	6527	擷		
1691	髻	2	93DA	4BFC		髻厶狄	
2022	漚	1	95EB	703D	漚		
2206	祆	7	9747	8884	祆		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
2302	忉	1	97C9	38FF		忉 厶 勺	
2346	鐳	1	97F5	9427	鐳		
2432	鞞	1	98AE	29350		革 厶 登	
3777	俵	2	8148	5008	俵		
3807	驗	3	8166	9A10	驗		
3830	咤	2	817D	54A4	咤		
4039	竝	1	82D3	7ADA	竝		
4042	笋	1	82D6	7B0B	笋		
4067	綉	1	82EF	7D89	綉		氈，錯字，以程式全部取代。
4074	縑	1	82F6	7E27	縑		
4111	蕤	1	835C	8493	蕤		
4195	鉢	1	83D2	9262	鉢		
4226	飡	1	83F1	98E1	飡		
4254	鷓	1	844E	9D7E	鷓		
4460	勅	1	85A1	52D1	勅		
4816	彳	1	87CB			彳	
4953	愨	1	88B7	22835		忉 厶 敞	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5008	幪	3	88EE	5E5E	幪		
5559	欸	1	8CA1	6B35	欸		
5570	繇	1	8CAC	6ABE	繇		

二、手動取代表說明：

《元刊雜劇三十種新校(下冊)》檔案中需手動修改的情況分為以下幾種：

(一)《元刊雜劇三十種新校(下冊)》檔案中出現「？」字形，共計33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《元刊雜劇三十種新校(下冊)》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、《元刊雜劇三十種新校(下冊)》「？」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 3	聽的??地叫	吂吂	
p. 5	斜了油?髻	髻	
p. 5	天?地	吂	
p. 6	俺叔叔癩?跛臂	賺	

頁碼	檔案原文摘錄	原書字形	備註
p. 6	鄰莊近?都	瞳	
p. 8	做?倒了	擲	
p. 35	開宴?肥鹿	瞳	
p. 35	錦?蕩蝶翅蜂鬚	鞞	改為鞞
p. 35	金貂繡?	袄	
p. 36	?袖垂肩仕女	驪	
p. 48	解了你的?	繚	
p. 52	住環?禮拜他	繚	
p. 52	麼?這個是你妻	?	
p. 54	癩?踐塵埃	賺	
p. 55	抹了?盂	鉢	
p. 69	韞?而藏諸	匱	
p. 85	肚皮?包著氣	衡	
p. 87	玉帝?, 東岳	勅	
p. 88	不能夠?賜封	勅	
p. 90	夫人?愁悶	衡	
p. 90	無百△?劃	司	
p. 104	輕擺著馬?環	金△前	
p. 112	制誥?怎生	勅	
p. 140	想?是禍患	衡	
p. 140	不知?是利害	衡	

頁碼	檔案原文摘錄	原書字形	備註
p. 140	?是禍害	衡	
p. 186	石?桑枯	灿	改為爍
p. 200	?地，★將軍	吖	
p. 205	你叫??賭賽	吖吖	
p. 206	?著，折末甚物	拞	
p. 256	錦?香，薦微臣	鞞	改為鞞

(二)《元刊雜劇三十種新校(下冊)》檔案中，對於舊版造字以外發現的錯字，以手動修改，並製作錯字修改表，請參考附件三。

欄位說明如下：

1. 檔案原文摘錄：摘錄檔案《元刊雜劇三十種新校(下冊)》中的文句段落，其中紅字底線的部分，為需要手動取代的原文。文句後方標示（全部取代）者，表示執行全部取代的動作，不一一列舉所有原文。
2. 手動取代結果：變更過後的檔案內容，其中藍字底線的部分，為修改後的結果。文句後方標示（全部取代）者，表示執行全部取代的動作，不一一列舉所有原文。
3. 備註：記錄修改相同詞句的次數或其他事項。

附件三、《元刊雜劇三十種新校(下冊)》缺字外錯字修改

檔案原文摘錄	手動取代結果	備註
<u>么</u> （全部取代）	<u>么</u> （全部取代）	有 32 處