

《醒世姻緣》(標記語料)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
2012/09/05 丁玟伶

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：

《醒世姻緣》(標記語料)使用舊版造字 10 個，字頻 16 次，這 10 個造字中，9 個可轉成 Windows XP 能顯示的字，字頻 15 次；另外 1 個字必須轉成構字式，字頻 1 次。

轉碼完成製作轉碼分析表，請參考附件一《醒世姻緣》(標記語料)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 4858 的「閱」字，是「閱」的異體字，以標準字「閱」取代。
2. 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 3098 的「揼」字，Unicode 字型呈現為「揼」，實際上仍為同一字。

附件一、《醒世姻緣》(標記語料)轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1213	睽	1	90D3	7743	睽		
1267	孺	1	914A	6B1B	孺		
1383	糠	2	91E0	7A45	糠		
1691	髻	1	93DA	4BFC		髻𠄎狄	
3029	鋤	1	9C6D	941D	鋤		
3098	揼	2	9CD4	63F8	揼		Unicode 字型呈現差異。
3968	煨	1	826A	7145	煨		
4450	狷	4	8575	524F	狷		Unicode 字型

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							呈現差異。
4858	閱	1	87F5	49A7		門△兒	閱，異體字問題。
5779	揿	2	8DE0	645A	揿		

二、手動取代表說明：

《醒世姻緣》(標記語料)檔案中需手動修改的情況分為以下幾種：

(一)《醒世姻緣》(標記語料)檔案中出現「●」字形，共計 380 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. 標記行碼：語言所標記檔裡設定的行碼。
3. 檔案原文摘錄：摘錄《醒世姻緣》(標記語料)檔案有「●」字形的文句。
4. 原書字形：「●」在原書中之字形，以此字取代檔案原文的「●」。
5. 備註：記錄修改相同詞句的次數或其他事項。

附件二、《醒世姻緣》(標記語料)「●」字手動取代表

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 6	332	●頭(Na)	幪	有 14 處
p. 12	688	●鞋(Na)	革△翁	有 4 處
p. 12	689	●帶(Na)	鞞	有 4 處

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 12	690	●(VC) 銀(Na) 順刀 (Na)	形金厂炎	
p. 12	745	●(VC)[+vrv] 定(VHC)	躡	
p. 14	747	●(Na) 鹿(Na)	麋	
p. 14	747	●(Na) 鹿(Na)	麋	
p. 15	826	●的(A) 一(Neu) 聲	嗥	
p. 18	948	三(Neu) 聲(Nf) ●	石厶充	
p. 18	966	又●子(Na)	袴	
p. 20	1103	●說蚱秀才	虫厶八	
p. 28	1480	●拉骨(A)	扌厶歪	
p. 31	1648	●臍(Na) 裏	脇	
p. 31	1662	懵懵●●	忄厶爭 忄厶爭	
p. 31	1677	●項(Na)	孛厶頁	
p. 33	1820	●條	毡	有 3 處
p. 37	2106	●●(Na) 皮	饜饜	
p. 38	2114	●●(Na) 嘗(VC) 嘗	饜饜	
p. 42	2363	●(VC) 著(Di) 我	躡	
p. 42	2393	●(VC) 頭(Na) 撒	石厶彭	
p. 42	2406	萬(Neu) 個(Nf) ●●	饜饜	
p. 42	2410	●(VH) 的(T4) 罵	扌厶歪	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 42	2413	腳(Na) ●(VC) 的	躡	
p. 48	2680	扯(VC) ●兒	擗	
p. 55	3162	一邊(Dh) 搖●	扌厶晃	
p. 58	3376	濕●●(VH) 的(漣漣	
p. 64	3767	一(Neu) 簍(Nf) 白●	養	
p. 66	3903	筍●(Na) 等(T6) 物	養	
p. 70	4191	鋪(VC) ●(Na)	毡	
p. 77	4577	窄●●	鼈鼈	有 2 處
p. 78	4662	天青(Na) ●子	縹	
p. 81	4810	賣(VD) 貓(Na) 人(Na) ●	綑	
p. 85	5105	向(P) 他(Nh) ●子	孛厶頁	
p. 85	5119	替(P) 他(Nh) ●子	孛厶頁	
p. 88	5321	膏藥(Na) ●	喻	
p. 92	5556	顧●(Na) 連裙	綉	
p. 102	6004	翰林(Na) 徐●	鞞	
p. 105	6211	隨即(Dd) ●	石厶彭	
p. 105	6211	●髻(Na)	髻厶狄	有 12 處
p. 108	6381	冷(VH) 水(Na) ●蒜	葱	
p. 114	6806	藍(VH) ●子	疋厶反	
p. 125	7536	柴●(Na) 下	土厶朶	改為朶

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 126	7597	●●包(Na)	饹饹	
p. 139	8438	●●皮子(Na)	饹饹	
p. 156	9506	有(V_2) ●頭(Na)	鉢	
p. 180	10796	●(VC) 在(P) 臀	綳	
p. 193	11622	麥(Na) 的(DE) ●●	饹饹	
p. 204	12243	炮●(Na) 一般	火△章	
p. 216	12893	一(Neu) ●(Nf) 酒	盃	
p. 218	13037	肫●(Na)	月△荅	
p. 218	13051	沒(VJ) 些(Neqa) ●倣	係	
p. 232	13871	金(Nb) 公(Na) 是 (SHI) ●	躡	
p. 242	14474	肫●紬(Na)	月△荅	
p. 247	14832	合(P) 他(Nh) ●別	拗	
p. 252	15124	蒸(VC) ●●	饹饹	有 6 處
p. 252	15125	切(VC) ●子	食△卷	
p. 252	15130	(COMMACATEGORY) ●● (Na)	饹饹	
p. 252	15130	●子(Na) 成(VHC) 幾	食△卷	
p. 253	15160	大(VH) ●●(Na)	饹饹	
p. 253	15192	拿(VC) 著(Di) ●●	饹饹	
p. 253	15203	兩(Neu) 個(Nf) ●●	饹饹	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 261	15743	四(Neu) 個(Nf) ●魚	鯨	
p. 266	16033	促●(VH) 短命	搭	
p. 276	16790	●●拉拉(Dh)	扌厶歪 扌厶歪	
p. 279	16972	●拉(VH) 潑(VH) 婦	扌厶歪	
p. 284	17379	做(VC) ●●(Na)	饜饜	
p. 288	17623	老(VH) ●拉(Na)	扌厶歪	有 3 處
p. 288	17637	●頭的(Na)	石厶斤	有 5 處
p. 289	17718	●的(Dh) 頭(Na)	石厶彭	
p. 291	17798	●(VC) 了(Di) 頭	石厶斤	
p. 291	17807	這(Nh) 夥子(Nf) ●	石厶斤	
p. 292	17844	吃(VC) 個(Nf) ●●	饜饜	
p. 292	17859	那(Nh) 夥(Nf) ●(VC) 頭	石厶斤	
p. 296	18161	端(VC) ●●(Na)	饜饜	
p. 297	18209	●舌地獄	犁	
p. 316	19598	臉(Na) ●稠(A)	食厶強	
p. 316	19610	平(VH) 旁人(Na) ●	躡	
p. 322	19859	●秭(Na)	禾厶蜀	有 7 處
p. 344	21372	儒●	縹	有 2 處
p. 345	21413	供獻(VD) ●●	饜饜	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 346	21528	●(Na) 襪(Na)	毡	
p. 350	21789	攙(VC) 麵(Na) ●	躡	
p. 350	21789	了(Di) 酒●(Na)	麪	
p. 359	22441	老(VH) ●(VC) 頭	石△斤	
p. 372	23388	囉●	咍	有 5 處
p. 397	25215	麻花●子	餛	
p. 401	25435	持●守錢	鉢	
p. 410	26075	大大(VH) 的(DE) ●●	饅饅	
p. 411	26100	●●(Na) 蒜麵	饅饅	
p. 417	26484	●穀(Na)	糝	
p. 418	26553	雞巴(Na) ●(VC)	□	
p. 418	26560	你(Nh) 待(Dd) ●	□	
p. 418	26570	那(Nh) 人●(Na)	□	
p. 418	26600	我(Nh) ●(VC) 他	□	
p. 419	26637	哈兒●兒(VA)	口△每	
p. 421	26779	●子(Na)	糝	有 9 處
p. 423	26931	誰(Nh) ●烤(VF)	焗	
p. 425	27010	程學究●中	棍	
p. 427	27152	斷間(Na) 板●(Na)	榻	
p. 430	27405	捍(VC) 炮●(Na)	火△長	
p. 451	29090	休(Dc) 想(VE) ●	躡	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 458	29452	眼睛(Na) ●(VC)	睇	
p. 461	29647	陰●(Na)	隲	有 2 處
p. 466	29972	千(Neu) 淫(VH) 萬 (Neu) ●	𠂇 𠂇 歪	
p. 467	30058	武●(Nb)	嬰	
p. 484	31392	兩(Neu) 個(Nf) ●	鬪	改為鬥
p. 486	31493	王●正可存	迹	
p. 493	31978	王●(U)	迹	
p. 494	32111	靴●(Na)	絛	
p. 496	32259	拍(VC) 著(Di) ●骨	𠂇月火亏𠂇	
p. 499	32504	六(Neu) ●(Nf) 酒	卮	
p. 528	34549	礮●子(Na)	石 𠂇 察	
p. 536	35146	火●●(Dh)	綳綳	
p. 537	35280	●包(Na) 盒子	毡	
p. 551	36244	囚姬蟬脫●	壳	
p. 551	36275	●(VC) 狗尾	躡	
p. 552	36335	●(Na) 弔拷	木 𠂇 匣	
p. 554	36500	大(VH) 盒(Na) ●●	饒饒	
p. 554	36501	樣(Na) 的(DE) ●子	菓	
p. 556	36654	●(Na) 上(Ng)	木 𠂇 匣	有 2 處
p. 556	36659	在(P) ●(Na) 裏	木 𠂇 匣	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 563	37182	●燂(VC) 喜果	炆	
p. 563	37203	儒(Na) ●	繚	
p. 570	37726	地下(Nc) 舖(VC) ●	毡	
p. 575	38004	酒醉●防	踈	改為疏
p. 582	38655	一(Neu) 床(Nf) ●	毡	
p. 590	39307	●包(Na) 端(VC) 著	毡	
p. 592	39466	就(D1) ●(VC)[+vrv] 折	扌 ㄩ 歪	
p. 593	39488	該(VL) ●(VC)	木 ㄩ 𣎵	
p. 593	39488	一(Neu) ●子(Nf)	木 ㄩ 𣎵	
p. 610	40686	腿(Na) ●(VH) 跨	扌 ㄩ 歪	
p. 614	41062	●拉(VC)	扌 ㄩ 歪	
p. 616	41145	秭●	稽	有 3 處
p. 624	41843	炭火(Na) ●煉	煨	
p. 638	42863	天目山(Nc) ●●	笋 𩚑	
p. 638	42878	萬●(Na)	笋	
p. 644	43319	盅●兒(Na)	筴	
p. 644	43354	●魚(Na)	𩚑	
p. 645	43464	在(P) ●子	𠃉 月 𠃉 𠃉	
p. 661	44768	潑(Na)[+spo] 的(DE) ●拉	扌 ㄩ 歪	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 662	44858	鋪(VC) ●(Na) 結(VC)	毡	
p. 662	44870	拜●(Na)	毡	
p. 662	44880	癩癩●●	扌△歪 扌△歪	
p. 663	44949	●拉(VC)[+vrv] 他	駿	有 2 處
p. 674	45785	白(VH) ●(Na) 套襪	毡	
p. 678	46087	綾(Na) 襪(Na) ●鞋	毡	
p. 678	46140	蒸(VC) ●(Na)	饊	
p. 680	46245	秣●	禾△蜀	
p. 684	46575	●黏(Na)	食△強	
p. 685	46636	●性(Na) 歪慙	扌	
p. 691	47080	●辣(VH) 麟	扌△歪	
p. 698	47691	番子手(Na) ●訪	躡	
p. 700	47763	吃(VC) 的(T3) ●飽	响	
p. 702	47907	老(VH) ●辣	扌△歪	有 3 處
p. 704	48091	爺(Na) ●拉	駿	
p. 705	48142	●枝捲(Na)	餓	
p. 705	48143	煎(VC) ●●	饊饊	
p. 710	48513	●肫	〇〇火△宗	
p. 710	48513	●(Na) 堵(VC)	月△荅	
p. 710	48543	狄希陳(Nb) ●趨	趨	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 717	49001	白(VH) ●帽	毡	
p. 721	49271	●(U) 得(T3) 惡心	藁	
p. 726	49545	那(Nh) ●兒(Na)	壳	
p. 726	49574	炮●(Na)	火△長	有 9 處
p. 726	49613	●的(A) 一(Neu) 聲	石△彭	
p. 726	49619	蹬●	扌△歪	有 2 處
p. 728	49735	麵(Na) 的(DE) ●●	饅饅	
p. 728	49745	婆娘(Na) ●(U)	扌△歪	
p. 728	49745	呀(T) ●(U) 的	扌△歪	
p. 740	50681	●拉(VC)[+vrv]	駿	
p. 766	52671	●●屑屑(Dh)	乚乚	
p. 769	52803	●鱔(Na) 見	魚△曲	
p. 770	52850	●聲(Na)	礮	
p. 773	53119	做(VC) 的(T3) ●●	饅饅	
p. 785	53948	那(Nh) ●(Na) 上	橈	
p. 785	53955	急(VH) 得(T3) ●(VC) 頭	石△彭	
p. 792	54487	大(VH) 籃(Na) ●●	饅饅	
p. 792	54546	秃(VH) ●(VH)	扌△歪	
p. 796	54841	●●嚶嚶	啞啞	
p. 803	55369	使(VC) ●包(Na)	毡	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 806	55569	翻(VAC) ●(Na)	甄	
p. 811	55975	●痕(Na) 在	泪	改為淚
p. 823	56762	●蝮瘡(Na)	瘦	
p. 823	56812	●●斜斜(Dh)	七七	
p. 825	56944	黃●(Na)	芪	
p. 828	57174	●亮骨(Na)	賺	
p. 828	57188	●瘡(Na)	賺	
p. 835	57720	●(Na) 帽(Na)	毡	
p. 836	57760	個(Nf) ●(Na) 來	毡	
p. 836	57798	一似(VG) ●(VC)	鼈	
p. 836	57801	叫(VL) 他(Nh) ●	鼈	
p. 841	58137	●(VC) 他(Nh)	嚏	
p. 842	58178	●趲(VH) 的	趲	
p. 849	58771	火●●(VH)	石△彭 石△彭	
p. 853	59052	爍(VC) 的(DE) ●	餓	
p. 855	59226	●條(Na)	毡	有 2 處
p. 856	59305	麵(Na) ●●	饜饜	
p. 861	59653	肫●	月△荅	有 2 處
p. 867	60089	椅子(Na) ●(VC)	躡	
p. 871	60503	●(VC) 的(DE) 地	躡	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 882	61362	捆(VC) ●包	毡	
p. 884	61453	造生●薛素姐	坟	改為墳
p. 890	61901	他(Nh) ●看	躡	
p. 894	62243	●●洩洩	乚乚	
p. 898	62505	盆(Na) 鐵(Na) ●	箒	
p. 900	62670	奴才(Na) 的(DE) ●	𠂇△𠂇	
p. 901	62780	賊(VH) ●辣	扌△歪	有 2 處
p. 905	63080	擺(VC) ●菜(Na)	菓	
p. 908	63297	屋(Nc) 裏(Ng) ●	睇	
p. 908	63374	天●瘡(Na)	疤	
p. 910	63561	魚●(Na) 蝦蟹	鼈	有 3 處
p. 928	64985	●(U) 著(Di) 過	目△農	
p. 928	64998	大紅(Na) ●包	毡	
p. 931	65084	放(VC) 炮●(Na)	火△章	
p. 951	66604	絲線(Na) 鸞●	絛	
p. 966	67729	大(VH) ●牛(Na)	牛△孛	
p. 976	68570	他(Nh) ●躡	乚	
p. 983	69095	淫婦(Na) ●拉	扌△歪	
p. 990	69605	大(VH) ●(Nf) 銀子	𠂇	
p. 994	69947	好(Dbb) ●(VC) ● (VC)	揼揼	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 1007	70982	●辣骨	扌 厶 歪	
p. 1015	71539	瀾●	涑 厶 鳥	
p. 1022	72154	屁股(Na) ●(VC)	疔	
p. 1024	72269	●(VC) 腦	石 厶 彭	
p. 1026	72413	束(Nf) 牙●	筋	
p. 1028	72585	的(DE) ●子	身 厶 兆	
p. 1032	72919	個(Nf) ●子	櫛	
p. 1044	73752	爺(Na) 們(T6) ●	扌 厶 歪	
p. 1046	73900	做(VC) ●●	饘 饘	
p. 1055	74599	二(Neu) 的(DE) ●魚	鮎	
p. 1057	74768	●青(Na) 坐馬	但	
p. 1058	74893	沒有(VJ) 床●	欖	
p. 1062	75150	羊羔(Na) 酒●	响	
p. 1070	75782	使手●(U)	扌 厶 歪	
p. 1080	76616	葫●子(U)	櫛	
p. 1084	76871	糖蜜(Na) 芝●	蔴	
p. 1095	77751	沒俶沒●	係	
p. 1099	78090	連(VJ) ●(Na) 帶	稽	
p. 1100	78195	糠●(Na) 草子	粃	有 2 處
p. 1104	78439	包裹(VC) 的(DE) ●書	勅	改為敕
p. 1120	79721	成(VHC) 了(Di) 鼻●	劬	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 1139	81182	帽●(Na)	盞	
p. 1157	82652	●(U) 了個	石△彭	
p. 1160	82926	是(SHI) 個(Nf) ●頭	才△粵	
p. 1173	83928	使(VC) ●包	毡	
p. 1183	84765	溫●粧燕	鰲	
p. 1188	85178	了(Di) 鼻●	屮	
p. 1203	86353	取(VC) 拜●	毡	
p. 1212	87052	狄希陳(Nb) ●(VC)	石△彭	
p. 1219	87527	後面(Ng) ●(VC) 頭	石△彭	

(二) 《醒世姻緣》素語料中之「●」字，為日後線上檢索方便，在標記語料多修正為古今通用字，與文獻處理實驗室較以原書為準修改其轉碼後素語料檔案的「●」字，兩者用字略有不同，其中多為異體字差異，特製作用字差異記錄表，以供未來參考。請參考附件三《醒世姻緣》素語料與標記語料「●」字差異記錄表。欄位說明如下：

1. 頁碼：修改字形所在位置之原書頁碼。
2. 素語料原文摘錄：摘錄《醒世姻緣》素語料檔案的文句段落，其中紅字部分，是經由文獻處理實驗室轉碼修改後的「●」字。
3. 標記語料原文摘錄：摘錄《醒世姻緣》標記語料檔案的文句段落，其中藍字的部分，為標記語料修正素語料中之「●」字形。文句後方標示（全部取代）者，表示標記語料將素語料使用之字形全部取代，不一一列舉所有原文。
4. 備註：記錄修改相同辭句的次數或其他事項。

附件三、《醒世姻緣》素語料與標記語料

「●」字差異記錄表

頁碼	素語料原文摘錄	標記語料原文摘錄	備註
p. 2	菓	果 (全部取代)	
p. 17	牙筋	牙箸	
p. 31	牀	床 (全部取代)	
p. 40	睺了	瞅(VC) 了(Di)	
p. 41	鬪	鬥 (全部取代)	
p. 60	耽	耽 (全部取代)	
p. 60	勅	敕 (全部取代)	
p. 97	慙	慚 (全部取代)	
p. 110	妬	妒 (全部取代)	
p. 123	孛△頁	脖	有 18 處
p. 156	劫	劫 (全部取代)	
p. 169	不响	不(Dc)[+vrd] 響	
p. 179	迴	迴 (全部取代)	
p. 210	墳	墳 (全部取代)	
p. 213	叙	敘 (全部取代)	
p. 214	徧	遍 (全部取代)	
p. 218	疎	疏 (全部取代)	
p. 225	猪	豬 (全部取代)	
p. 259	綉花汗巾	繡花(A) 汗巾	

頁碼	素語料原文摘錄	標記語料原文摘錄	備註
p. 261	綳兒	繃兒	
p. 267	紅毡	紅(VH) 氈	
p. 267	毡上	氈(Na) 上	
p. 268	莖	莽 (全部取代)	
p. 285	芝麻鹽	芝麻鹽	
p. 292	冷竈	冷(VH) 灶(Na)	
p. 298	擡在後園	壘(VC) 在(P) 後(A) 園	
p. 308	挑葱	挑(VC) 蔥(Na)	
p. 310	下了筯	下(VCL)[+vrr] 了(Di) 箸	
p. 311	鍾筯	鍾箸(Na)	
p. 312	大筯	大(VH) 箸(Nf)	
p. 361	罈	罈 (全部取代)	
p. 362	鐵火筯	鐵火箸(Na)	
p. 382	一雙筯	一(Neu) 雙(Nf) 箸	有 2 處
p. 391	死鼈一般	死(VH) 鱉(Na) 一般	
p. 430	豎	豎 (全部取代)	
p. 437	龜鼈	龜鱉(Na)	
p. 439	虬髯	虬髯(Na)	
p. 470	奉酌	奉酬(VD)	

頁碼	素語料原文摘錄	標記語料原文摘錄	備註
p. 503	徼彳△幸	徼(VC)[+spv] 倖	有 2 處
p. 521	烏木筴	烏木(Na) 箸	有 2 處
p. 521	兩雙筴	兩(Neu) 雙(Nf) 箸	有 2 處
p. 527	泪	淚 (全部取代)	
p. 536	覩物	睹(VE) 物(Na)	
p. 542	形迹	形跡(Na)	
p. 563	銀鑲筴	銀鑲(A) 箸	
p. 570	用筴挑	用(VC) 箸(Na) 挑	
p. 579	兩盃麵	兩(Neu) 碗(Nf) 麵	
p. 580	慇懃	慇懃(VF)	
p. 582	瞇目△逢	瞇縫(VC)	
p. 638	羊尾筴	羊尾(Na) 筴	
p. 645	足△另的跌	拐(VC) 的(T3) 跌	
p. 681	漏卮	漏(VA) 卮(Na)	有 3 處
p. 684	主人拗別	主人(Na) 拗別	
p. 684	韭	韭 (全部取代)	
p. 685	剋落	剋落(VC)	
p. 699	拗東	拗東	
p. 756	用蔴繩	用(VC) 麻(Na) 繩	
p. 760	懶惰悞	懶惰(VH) 誤	
p. 764	髻髻	髻佛(VG)	

頁碼	素語料原文摘錄	標記語料原文摘錄	備註
p. 769	魚鱉	魚鱉(Na)	
p. 775	嗅	嗅(全部取代)	
p. 778	棒椎一响	棒椎(Na) 一(Dd) 響	
p. 781	呪	呪(全部取代)	
p. 807	碇碌	谷碌	有 2 處
p. 830	寧喂狼	寧(VK) 喂(VC) 狼	
p. 848	一顛	一(Neu) 顛(VA)	有 2 處
p. 850	足△比蹬	趾蹬(VA)	
p. 894	虛壳	虛(VH) 殼	
p. 900	蹲在	蹲(VA) 在(P)	
p. 942	唾沫	唾沫(Na)	
p. 972	熱火筴	熱(VHC) 火筴	
p. 974	軟軟的	軟軟(VH) 的	
p. 984	還木△歹	還(Dd) 搵(VC)[+vrv]	
p. 998	木△歹一木△歹	搵(VC) 一(Neu) 搵	
p. 1004	捱一木△歹	捱(VJ) 一(Neu) 搵	
p. 1008	月△合肢	肱肢(Na)	
p. 1017	佻佻不係	佻佻(Dh) 不(Dc) 睬	
p. 1030	使筴夾	使(VC) 筴(Na) 夾	
p. 1055	攞食	攞(VC) 食(Na)	
p. 1065	靚著臉	靚(VC) 著(Di) 臉	

頁碼	素語料原文摘錄	標記語料原文摘錄	備註
p. 1084	紅麴	紅麴(Na)	
p. 1092	覷著	覷(VC) 著(Di)	
p. 1117	筴來	筴(Na) 來	
p. 1125	椀	碗(全部取代)	
p. 1145	弄筴	弄(VC) 筴(Na)	
p. 1165	西ㄩㄣㄉㄨㄥ你處	酬(VC) 你(Nh) 處	
p. 1177	口ㄩㄣㄊㄨㄥ屁股	舔(VC) 屁股	
p. 1182	造了詞詞	造(VC) 了(Di) 歌詞	
p. 1222	軀壳	軀殼(Na)	

(三)《醒世姻緣》(標記語料)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，請參考附件四。欄位說明如下：

1. 檔案原文摘錄：摘錄檔案《醒世姻緣》(標記語料)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。
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。
3. 備註：記錄修改相同詞句的次數或其他事項。

附件四、《醒世姻緣》(標記語料)缺字外錯字修改

檔案原文摘錄	手動取代結果	備註
不(Dc) 瞅(VC) 不(Dc) <u>瞅</u>	不(Dc) 瞅(VC) 不(Dc) <u>睬</u>	