

《水滸傳》(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
2012/08/14 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 1616 的「運」字，Unicode 編碼是 284E6，由於 XP 並不支援，仍需使用構字式「辶△里」。

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：

《水滸傳》(標記語料)使用舊版造字 117 個，字頻 852 次，這 117 個造字中，101 個可轉成 Windows XP 能顯示的字，字頻 761 次；另外 16 個字必須轉成構字式，字頻 91 次。

轉碼完成製作轉碼分析表，請參考附件一《水滸傳》(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 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 3235 的「𠄎門〃𠄎」字，而原書使用之字為「門△龜」，查詢教育部異體字字典確認「門△龜」並非「𠄎門〃𠄎」之異體字，因此歸為錯字，以程式取代為正確字形「門△龜」。
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 5705 的「廐」字，Unicode 字型呈現為「廐」，實際上仍為同一字。

附件一、《水滸傳》(標記語料)轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
38		42	FA65			王△巳	玨，異體字問題。
1000		31	8F79			王△巳	玨，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1170	瞳	7	90A8	7583	瞳		
1213	睽	19	90D3	7743	睽		
1321	確	1	91A2	7936	確		Unicode 字型 呈現差異。
1409	寵	2	91FA	7AC9	寵		
1491	擷	3	92AF	6527	擷		
1598	翰	1	935B	4A7A		革△翁	
1600	猷	1	935D	3E84		犛△俞	
1616	逕	36	936D	284E6		辶△里	
1660	賺	1	93BB	8181	賺		
1676	艫	1	93CB	824C	艫		
1688	魷	1	93D7	9B6B	魷		
1772	藜	1	946C	4527		艹△梨	
1808	棚	2	94B2	3BB6		木△朔	
1905	褊	7	9554	88E9	褊		
1926	覓	3	9569	8994	覓		
1995	通	1	95D0	23EE2		辶△通	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
2046	踪	1	9644	8E2A	踪		
2053	躡	2	964B	8EA7	躡		
2097	膾	32	9677	26782		月 厶 荅	
2132	癩	1	96BC	3FC0		疒 厶 班	
2154	魁	2	96D2	9B40	魁		
2157	酎	1	96D5	9167	酎		
2162	醜	2	96DA	918E	醜		
2244	蛭	1	976D			ノ 会 虫	
2345	鑽	1	97F4	9404	鑽		
2362	鑽	1	9846	93C6	鑽		
2409	闕	15	9875	95A7	闕		
2454	鍬	2	98C4	28AFC		金 厶 教	
2481	盞	1	98DF	276FE		礻 厶 盞	
2955	節	1	9BE2	83AD	節		
2999	嗶	6	9C4F	5523	嗶		
3004	籊	3	9C54	7BD0	籊		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3014	櫛	1	9C5E	3BBC		木△扇	
3235		2	9DC0			形門〃 黽	門△龜，錯字，以程式全部取代。
3789	冲	1	8154	51B2	冲		沖，異體字問題。
3797	凭	6	815C	51ED	凭		
3801	剗	2	8160	5257	剗		
3804	効	1	8163	52B9	効		
3805	勅	3	8164	52C5	勅		
3814	却	26	816D	5374	却		
3818	廩	12	8171	53AB	廩		
3819	廩	10	8172	53AE	廩		廩，異體字問題。
3824	叙	1	8177	53D9	叙		
3825	疊	1	8178	53E0	疊		
3830	咤	1	817D	54A4	咤		
3831	詰	1	817E	5586	詰		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3835	噍	2	81A4	564D	噍		
3870	帑	1	81C7	5E0B	帑		
3877	迺	4	81CE	5EFC	迺		
3886	悞	1	81D7	609E	悞		
3892	喉	4	81DD	3B0B		日 ㄩ 侯	
3900	挖	1	81E5	62D5	挖		
3909	擡	3	81EE	64E1	擡		
3931	椀	1	8245	6900	椀		
3934	桶	11	8248	69C5	桶		
3940	爵	2	824E	6B1D	爵		
3945		12	8253			彳 ㄩ 义	汉，異體字問題。
3954	戩	23	825C	622C	戩		
3963	炆	9	8265	7081	炆		
3964	威	1	8266	70D5	威		
3966	烟	48	8268	70DF	烟		
4028	稟	5	82C8	7980	稟		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4041	豎	2	82D5	7AEA	豎		
4067	綉	3	82EF	7D89	綉		
4074	繚	19	82F6	7E27	繚		
4090	脚	17	8347	811A	脚		
4105	菓	29	8356	83D3	菓		
4107	葱	2	8358	8471	葱		
4112	蔴	2	835D	8534	蔴		
4132		20	8371			血𠂇𠂇	眾，異體字問題。
4167	賫	121	83B6	8CEB	賫		
4193	鈎	1	83D0	920E	鈎		
4226	飡	13	83F1	98E1	飡		
4242	鬪	32	8442	9B2D	鬪		
4256	鷄	12	8450	9DC4	鷄		雞，異體字問題。
4267	鼈	2	845B	9F08	鼈		
4278	亘	1	8466	4E98	亘		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4287	織	2	846F	7E8E	織		
4288	鬱	4	8470	23861		形林缶 冫尉	
4307	罵	1	84A5	99E1	罵		罵，異體字問題。
4308	劫	8	84A6	5227	劫		
4316	艷	2	84AE	8276	艷		
4400	兜	1	8543	5160	兜		
4432		1	8563			形糸厂 墨	纏，異體字問題。
4450	勑	1	8575	524F	勑		Unicode 字型呈現差異。
4460	勅	2	85A1	52D1	勅		
4496	段	9	85C5	53DA	段		
4519	咲	1	85DC	54B2	咲		
4533	唸	1	85EA	552B	唸		
4604	擦	1	8672	22D32		扌	絮
4679	塏	1	86DF	5818	塏		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4702	場	3	86F6	5872	場		
4720		1	8749			土厶已	圪，異體字問題。
4764	摠	3	8775	64A7	摠		Unicode 字型呈現差異。
4779	鞣	1	87A6	97B2	鞣		
4784	鞞	1	87AB	9793	鞞		
4791	羝	1	87B2	7F93	羝		
4903	𠂔	31	8863	5C00	𠂔		
4904	尅	15	8864	5C05	尅		
5008	幪	14	88EE	5E5E	幪		
5074	𠂔	1	8971	5C2B	𠂔		
5111	旆	3	89B8	65BE	旆		
5119	博	1	89C0	613D	博		
5147	恠	1	89DC	6060	恠		
5180	拘	10	89FD	6285	拘		
5240	擗	2	8A7A	6481	擗		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5281	𪗇	1	8AC5	994A	𪗇		
5285	𪗈	1	8AC9	7742	𪗈		
5300	𪗉	1	8AD8	9B2A	𪗉		
5451	玳	1	8BD2	7398	玳		
5466	倏	1	8BE1	5010	倏		
5543	𪗊	1	8C6F	9B9D	𪗊		
5559	欸	16	8CA1	6B35	欸		
5705	廐	2	8D74	5ED0	廐		Unicode 字型 呈現差異。
5779	揿	1	8DE0	645A	揿		

二、手動取代表說明：

《水滸傳》(標記語料)檔案中需手動修改的情況分為以下幾種：

- (一)《水滸傳》(標記語料)檔案中出現「●」字形，共計 264 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《水滸傳》(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、《水滸傳》(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	290	●●兒(Na)	餽蝕	有 2 處
p. 7	412	猥●(VH)	獮	有 2 處
p. 64	3809	●揲(VC)[+vrv] 起 來	扌△周	
p. 67	4070	如何(Dj) ●(VC) 得(T3)	嚏	
p. 86	5221	騎(VC) 著(Di) ● (A) 馬	擡	
p. 152	9466	那里(Nc) ●(VC) 幾	門△坐	
p. 156	9750	先(Dd) ● (VC)[+vrv] 倒	攢	
p. 156	9757	又(D1) ● (VC)[+vrv] 倒	扌△戳	待造字
p. 188	11558	馬鞍●	鞦	有 3 處
p. 203	12389	貪(VK) ●(VC) 這 (Nh) 口(Nf) 黃湯	嚏	
p. 242	14916	開剝推●(U)	尋△司	
p. 264	16469	●哨	唵	有 7 處

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 273	16876	酷吏●官都殺盡	𦘔	
p. 282	17621	●(U) 的(T3) 火雜 雜	𦘔	
p. 291	18016	一(Neu) 片(Nf) 青 (VH) ●皮	豸𦘔瓜	
p. 294	18252	自(Dh) 有(V_2) ● 劃	百𦘔𠂔	
p. 349	21614	牽(VC) 一(Neu) ● (Nf) 羊	羸	
p. 406	25134	●骨池(Nc) 內(Ng)	澈	
p. 406	25140	●(VC) 在(P) 池子 (Nc) 裏	澈	
p. 408	25244	挽(VC) 了(Di) 個 (Nf) ●兒	𦘔𦘔角	
p. 416	26030	臉(Na) 上(Ng) 便 (D1) ●(VA)	扌𦘔閉	
p. 416	26030	兩(Neu) ●(VA)	扌𦘔閉	
p. 441	27462	許多(Neqa) 蒸(VC) ●兒	食𦘔卷	
p. 441	27463	雞(Na) 來(VCL) ● (VC) 了	扌𦘔系	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 453	28323	與(VD) 他(Nh) ● (VC)	𠵼	
p. 468	29320	左(Ng) 手(Na) ● (U) 來	扌 𠵼系	
p. 476	29558	●(VC) 得(T3) 這 等(𠵼	
p. 478	29785	出(VC) 得(T3) 鬆 (VHC) ●	月 𠵼桑	
p. 551	34198	又(D1) 唸●(VH)	噫	
p. 583	36092	吹風胡哨(Dh) ● (VA)	是 𠵼干	
p. 606	37583	一(Neu) 點(Nf) ● 兒	𠵼 𠵼角	有 2 處
p. 606	37648	紅(Na) 俏(VH) ● 兒	𠵼 𠵼角	
p. 626	38898	得(VK) 出去(VCL) ●(VC)	𠵼	
p. 644	39916	兩(Neu) 箇(Nf) ● 扎	扌 𠵼匾	
p. 674	41597	●行(VA)[+vrv] 到	之 𠵼贊	改為之 𠵼贊
p. 674	41647	今日(Nd) 之(DE)	𠵼 𠵼 𠵼 𠵼	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
		●		
p. 693	42774	抓●兒(A) 頭巾	髟△角	
p. 695	42899	髻髻(Na) ●(Na) 邊	髻	
p. 695	42989	李達(Nb) 匹●(Dh)	髟△角	
p. 698	43239	前面(Ng) 石(Na) ●(Nc) 上	石△皆	
p. 702	43527	●(VC)[+vrv] 住 (VHC)[+vrr] 李達	欸	
p. 703	43699	肉(Na) 裏(Ng) 多 (Dab) ●(VC) 些	米△叅	改為糝
p. 719	44635	張保(Nb) ●●(VH) 不(Dc) 是	允△兼 允△介	改為慙尪
p. 720	44709	●命三	扌△弃	有 9 處
p. 741	45983	露(VC)[+vrv] 出 (VCL)[+vrr] 光●	立△戈	
p. 745	46295	●項(Na)	宀△頁	有 6 處
p. 762	46815	自(Dh) 把(P) ●兒	髟△角	
p. 765	47058	鼓上●	虫△造	改為蚤 有 6 處
p. 781	47831	口(Na) 邊(Ng) ●	月△尔	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
		腥		
p. 840	51389	一(Neu) 條(Nf) 皂 (VH) ●	繚	
p. 878	53500	戴宗(Nb) ●(VC)	睭	有 2 處
p. 888	54344	酒肉(Na) ●(VC)	噉	
p. 901	54503	一(Neu) 頓(Nf) ● 子拳	孛△頁	
p. 904	54741	馬鞍(Na) 前(A) ● (Na)	鞦	
p. 906	54870	●州(Nc)	完△女	有 2 處
p. 909	55219	量(VE) 此(Nh) 草 ●	完△女	
p. 973	58335	手(Na) ●(VC)	搯	有 24 處
p. 1004	59911	天地(Nc) 相(Dh) ●	拗	
p. 1007	60144	違●	拗	有 4 處
p. 1022	60954	鈴●(Na)	金△午	有 2 處
p. 1031	61577	●(VC) 著(Di) 手 (Na) 中	搯	
p. 1053	62889	●(VC)[+vrv] 住 (VHC)[+vrr] 盧俊義	𠂔才 北白儿 𠂔	改為才 △兜

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 1079	64172	兩(Neu) 眉(Na) 入 (VCL) ●	鬢	
p. 1090	64525	●箕掌(Na)	𠂇△皮	有 3 處
p. 1090	64526	拷●圈(Na)	扌△老	有 4 處
p. 1121	66021	便(D1) 要(VK) 催 ●	辶△贊	
p. 1126	66459	硫黃(Na) ●硝	火△𠂇	改為焰 有 4 處
p. 1129	66721	匹●兒(Dh) 揪	𠂇△角	
p. 1137	66979	趙●(Nb)	日△拆	有 4 處
p. 1142	67477	成(VHC)[+vrr] ● (VC) 餅	畫	改為畫
p. 1163	68813	捨(VC) 身(Na) ● (VC)[+spv] 命	扌△弃	
p. 1170	68983	二(Neu) 十(Neu) ●棍	木△𠂇	
p. 1171	69124	彈(VC) 打(VC) ● 鳩	班△鳥	
p. 1171	69130	膽(Na) 包(VC) 身 ●	躰	
p. 1172	69222	頭髻(Na) ●鬆	鬢	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 1186	70170	禿●(Na) 腦袋	駱	改為驢
p. 1194	70484	如(VG) 拷●(Na) 之	扌△老	
p. 1195	70501	以(Cbb) 充(VG) ● 資	貝△親	
p. 1205	71031	積●(VC)[+vrv] 得	辶△贊	
p. 1205	71082	●邊不可無黃菊	髻	
p. 1213	71299	那里(Dj) 執●	扌	
p. 1220	71902	戎裝(Na) ●帶	巾△貫	
p. 1227	72269	腳步(Na) ●行	辶△贊	
p. 1234	72936	便(D1) 叫(VF) ●	煮	改為煮
p. 1236	73106	雲(Na) ●(Na) 花 (Na) 顏	髻	
p. 1236	73128	房(Nc) 中(Ng) 積 ●	辶△贊	
p. 1236	73151	紅(VH) 日(Na) ● (VC) 山	𠃉口行缶𠃉	
p. 1237	73177	開(VHC) 絳(A) ●	蓋	
p. 1244	73540	出(VCL) 草●亭	叅	
p. 1244	73541	●拜(VC) 了(Di) 四(Neu) 拜	叅	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 1245	73639	●(VC) 神(Na) 已 (Dd) 罷	叅	
p. 1245	73662	●膊(Na)	月△荅	
p. 1245	73665	一(Neu) 聲(Nf) ● (VC)	叅	
p. 1248	73995	都(Dab) 來(VCL) ●見	叅	
p. 1263	74872	差(VF) 人(Na) ● 運	之△贊	
p. 1270	75354	長壯(VH) 軍士(Na) ●	扌△晃	
p. 1272	75461	鬣(Na) 邊(Ng) 都 (Dab) 插	鬣	
p. 1286	76193	不(Dc) 知(VK) ● 倒	亻△顛	
p. 1286	76197	馬(Na) 後(A) 蹄 (Na) ●	之△坐	
p. 1297	76894	兩(Neu) 手(Na) ●	搯	有 2 處
p. 1318	78261	●謀(Na)	叅	有 11 處
p. 1319	78357	帥府(Nc) ●見	叅	
p. 1320	78453	要(VK) ●(U) 完備	之△贊	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 1321	78505	那里(Nc) 掙●	之△贊	
p. 1322	78610	太尉(Na) 催●(VF)	之△贊	
p. 1335	79396	你(Nh) ●倒(Dh)	イ△顛	
p. 1340	79890	罪惡難分●倒	イ△顛	
p. 1342	80099	●謀大事(U)	叅	
p. 1371	81569	宋江(Nb) 催●	之△贊	
p. 1375	81899	住(VHC)[+vrr] 鞍● (Na)	鞦	
p. 1376	81952	●戰(VA) 廝(Dh) 殺	拏	
p. 1376	81999	帳(Na) 前(Ng) ● 見	叅	
p. 1429	85509	熒●火星	或△火	
p. 1436	85910	女(Nb) 、 (PAUSECATEGORY) ●	虛	
p. 1437	85981	西●(Na) 七炁星君	八△允	待造字
p. 1437	85994	各(Dab) ●(VC) 伏 手	拏	
p. 1439	86129	紫絲韁牢絆山●	鞦	
p. 1471	88195	戎裝(Na) ●帶	巾△貫	
p. 1493	89889	不(Dc) 要(VK) ●	係	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
		(VJ)		
p. 1501	90562	●(VC) 兩(Neu) 斧	拞	有 2 處
p. 1502	90644	●(VC) 板斧(Na)	拞	有 2 處
p. 1503	90754	項●(Nb)	一△允	依前文改為 充
p. 1518	91977	只(Daa) ● (VC)[+spv] 一(Neu) 死	扌△弃	
p. 1545	94328	正在(Dd) ●劃	百△司	
p. 1566	96005	不(Dc) 去(VCL) ● (VJ) 他	係	
p. 1568	96175	●勦大業(U)	贊△力	待造字
p. 1574	96594	動●(VA)[+vrv] 不 得	驛	有 3 處
p. 1576	96838	火(Na) 般(T4) 熾 ●	火△音	改為焰
p. 1579	97073	不(Dc) 來(VCL) ● (VJ) 著	係	
p. 1579	97112	●●喝喝	口△么 口△么	有 2 處
p. 1585	97500	果(Dbb) 是(SHI)	丑	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
		雪●		
p. 1588	97774	耳根(Na) 連(VJ) ●子	亭△頁	
p. 1592	98101	村(Nc) 姑(Na) 農 ●	嬪	
p. 1593	98141	●(VC) 著	睇	有 2 處
p. 1595	98293	輸敗(VH) 脛●(VH)	月△贊	改為贖
p. 1595	98359	他(Nh) 炙●(VA) 殺	火△專	
p. 1600	98783	兄弟(Na) ●助(VC)	幫	改為幫
p. 1601	98871	今日(Nd) ●(VH) 著(Di)	疒△敝	
p. 1606	99342	堆(Na) 距(VC) ● (Na)	足△聖	待造字
p. 1617	100173	馬●(Nb)	𩇛	有 7 處
p. 1618	100269	瓊英(Nb) 向(P) 鞍 ●	鞦	
p. 1625	100775	馬(Na) 後(A) 蹄 (Na) ●	足△虛	
p. 1625	100829	一(Neu) 頭(Nf) ● 鬆	蓬	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 1627	100960	動●(VA)[+vrv]	驛	有 2 處
p. 1659	103714	●源洞(Nc) 中	幫	
p. 1660	103764	●(VC) 了(T) 上	拈	
p. 1668	104121	裝(VC) 了(Di) ●	𠃉木宀肖	待造字
p. 1669	104204	就此(Dd) ●見	叅	
p. 1672	104477	上來(VCL) ●見	叅	
p. 1673	104486	過去(VCL)[+vrr] ● 拜	叅	
p. 1682	105060	●戰(VA)	拈	有 7 處
p. 1699	106228	一(Neu) ●(Nf) 羊 (Na)	羸	
107135	107135	王仁(Nb) 卻(D1) ●	拈	
p. 1716	107419	穿心紅(Na) 的(DE) ●兒	𠃉厶角	
p. 1732	108452	火雜雜(VH) 地(T4) ●	拈	
p. 1735	108629	口(Na) 裡(Ng) ● (VC) 著	𠃉口行缶	
p. 1768	110924	●見(VC) 盧(Nb) 先鋒	叅	

頁碼	標記行碼	檔案原文摘錄	原書字形	備註
p. 1768	111004	天氣(Na) 昏●(VH)	登𠄎𠄎	
p. 1787	112308	來(VCL)[+vrr] ● (VC) 了	擲	
p. 1787	112364	諸(Nes) 將(Na) ● 拜	叅	

(二) 《水滸傳》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《水滸傳》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《水滸傳》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《水滸傳》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、《水滸傳》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. 16	遶	奔（全部取代）	

頁碼	素語料原文摘錄	標記語料原文摘錄	備註
p. 29	欸紐狼腰	欸紐(VC) 狼腰(Na)	
p. 29	欸欸揪住	欸欸(VH) 揪 (VC)[+vrv] 住(VHC)	
p. 30	惧	懼 (全部取代)	
p. 33	宰了一鞞大	宰(VC) 了(Di) 一 (Neu) 腔(Nf) 大(VH)	
p. 70	草蒂兒	草(Na) 蒂兒(Na)	有 2 處
p. 73	必△剝△剝 剝剝	必剝剝(VH)	
p. 84	捏着兩把	捏(VC) 著(Di) 兩 (Neu) 把	
p. 103	浼	浼 (全部取代)	
p. 104	叅	參	
p. 114	瞅	瞅	
p. 151	允△兼允△介	尷尬	有 2 處
p. 155	央浼	央浼(VF)	
p. 164	顧盼	顧盼(VC)	
p. 165	獐皮窄革△幻	獐皮(Na) 窄(VH) 鞞	
p. 177	闢了	鬥(VC) 了	
p. 183	行△岳環背鞍	銜環背鞍	
p. 189	勃△刺△刺	勃喇喇(Dh)	

頁碼	素語料原文摘錄	標記語料原文摘錄	備註
p. 213	撻	划 (全部取代)	
p. 217	肯帶才 ㄟ 挈	肯(VL) 帶挈(VC)	
p. 311	𠵼 ㄟ 顛倒	顛倒(Dh)	
p. 315	鼻 ㄟ 勾	𠵼 (全部取代)	
p. 348	𠵼 ㄟ 忽𠵼 ㄟ 聿心	忽律心(Na)	
p. 960	恪	吝 (全部取代)	
p. 1137	口 ㄟ 毫	𠵼	
p. 1619	冒面具	冒(VC) 面具(Na)	