《新唐書》舊版造字轉碼說明

中研院資訊所文獻處理實驗室
中研院史語所漢籍電子文獻工作小組
2007/9/19 黃榮順 製作

一、 《新唐書》一書(新唐書.xml)使用舊版造字 297 個，字頻 620 次，詳如附件一。這 297 個造字中，187 個可轉成 Windows XP 可顯示的字，字頻 449 次；另外 110 個字必須轉成構字式，字頻 171 次。

二、 附件一的造字分析表說明如下：

甲、 編號：Big5 造字空間為 6217 個，編號由 1 到 6217。
乙、 造字：舊版造字
丙、 字頻(txt)：造字在「.txt」文件的出現次數
丁、 字頻(xml)：造字在「.xml」文件的出現次數
戊、 Big5：造字的 Big5 碼
己、 Unicode：造字所對應的 Unicode 碼
庚、 WinXP：造字在 Windows XP 的對應字形
辛、 構字式： Windows XP 若無對應字形，則改採用構字式
壬、 備註凡例：

1、錯字：與原書字形不符，屬於意義不同的字，則用正確的字取代原造字，如編號 4409 的「冋」字，原字應為「同」。

2、取代：如異體字，雖與原書字形不符，但在《異體字字典》（教育部）中列為異體，為廣查詢利用，盡量以正體字取代。如編號 3945 的「魚⿰⿰眾眾」，原書字形為正體「鰥」字；或發現 Unicode 另有字形者，則取代構字式。如編號 3985 的「⿱形⿱王丰⿱母⊡」，xml 檔使用「瑀」字取代。

3、txt 重複字句，致頻次和 xml 不同：手動將 xml 未修改的部份全部取代。如編號 3991 的「攜」，txt 字句重複，故頻次多 xml 一次，手動將 xml 未修改部份取代。

4、配合修改，手動取代：xml 字頻低於 txt 字頻，經比對 txt 及 xml 檔案內容後，發現一部分原舊漢籍造字之前已被人工修改，確認字形無誤後，剩餘未修改的舊漢籍造字也配合修改。如編號 4256 的「鷄」，舊漢籍造字字形為「鷄」，比對 txt 檔及 xml 檔後，發現原先在 txt 檔使

用「鷄」字者，在 xml 檔中已部份人工修改為 Unicode 的「鷄」，致頻次不同，故將未修改之部份亦手動取代為 Unicode 的「鷄」。

5、錯誤修改，手動更正及取代：xml 字頻低於 txt 字頻，經比對 txt 與 xml 檔案內容後，發現一部分原舊漢籍造字之前已被人工修改，但經比對原書，發現修改後的字形是錯誤的，必須在確認後，將錯誤修改的字形或構字式更正，並將尚未修改的舊漢籍造字以正確字形手動取代。如編號 4706 的「隥」，1 例被錯誤修改為「衛」，確認原書字形後，手動更正為「隥」，xml 未修改的部份也手動取代為「隥」。

6、Unicode 字型呈現差異：Unicode 字型與舊漢籍造字有些微差異，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 5034 的「弃」字，Unicode 字型呈現為「弃」，實際上仍為同一字。

7、待造字：舊版漢籍造字字形為電腦缺字，且漢字構形資料庫尚未收錄此字形，等待補造字中，所以「造字」欄空白無法看到字形。如編號 5079 的「忄⿰存」。

三、 Unicode 目前收錄的漢字總數為 70194，分屬於三個不同區段，詳如表一。目前 Windows XP 只支援 CJK 認同表意文字區的 20902 個字，內碼為 4E00-9FFF。所以造字編號 2223 的「骬」字，Unicode 編碼為 4BDB，由於 Windows XP 並不支援，仍須改用構字式「骨⿰宛」。

表一、Unicode 的字數及編碼區段

Unicode	字集子集合	新增字數	新增編碼區段	總字數	WinXP
1.1 版	CJK 認同表意文字區	20902	4E00-9FFF	20902	支援
3.0 版	CJK 認同表意文字擴充 A 區	6582	3400-4DFF	27484	不支援
3.1 版	CJK 認同表意文字擴充 B 區	42710	20000-2A6D6	70194	不支援

四、 造字分析表「備註」欄標示「手動取代」者，詳細取代內容可參考附件二的手動取代表。附件二的手動取代表說明如下：

甲、 編號：Big5 造字空間為 6217 個，編號由 1 到 6217。
乙、 造字：舊漢籍造字字形
丙、 取代：實際使用字形與造字欄不同時，於此欄列出。

丁、 檔案原文摘錄：摘錄檔案(新唐書.xml)中包含該造字的文句段落，其中紅字底線的部分，為需要手動取代的原文。

戊、 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。

五、 「新唐書.xml」缺字外發現之錯字，修正情形詳列附件三，附件三體例同上文第四點丁、戊兩項，編號則為流水號。

附件一、「新唐書.xml」造字分析表

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
8	賦	1	1	FA47	27E10		貝⿰貳	
9	鉚	1	1	FA48	491D		金⿰卯	
799	齗	1	1	8E4D	9F5A	齗		
860	褱	1	1	8EAC	890F	褱		
866	奐	4	4	8EB2	247ED		㐫⿱犬	㐫⿱大，取代。
961	枅	1	1	8F52	3B4A		木⿰片	
968	瑩	2	2	8F59	4741		烝⿱豆	
1002	鳩	1	1	8F7B	9D02	鳩		
1013	絙	1	1	8FA8	7D5A	絙		
1015	嶴	4	3	8FAA	5D4F	嶴		嶴，錯誤修改，手動取代。
1062		1	1	8FD9			馬⿰弔	騮，錯字。

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
1067	棊	1	1	8FDE	68CA	棊		
1109	璂	2	1	9049	7496	璂		璂，錯誤修改，手動取代。
1146	甍	1	1	906E	7506	甍		
1170	疃	2	2	90A8	7583	疃		
1258	袖	1	1	9141	4102		礻⿰由	
1383	穠	1	1	91E0	7A45	穠		
1412	鄼	1	1	91FD	913C	鄼		
1430		1	1	9250			竹⿱毛	待造字。
1513	粮	1	1	92C5	7CAE	粮		
1605	翎	1	1	9362	7FD6	翎		
1641	胆	1	1	93A8	80C6	胆		
1653	脿	1	1	93B4	8142	脿		
1813	瞡	1	1	94B7	25251		目⿰雀	
1897	祕	1	1	954C	8890	祕		祕，取代。
1908		1	1	9557			礻⿰苒	待造字。
1937	詢	2	2	9574	8A3D	詢		

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
1982	讟	1	1	95C3	8B87	讟		
2168	醻	1	1	96E0	91BB	醻		
2223	骬	2	2	9758	4BDB		骨⿰宛	
2225	祴	1	1	975A	88D3	祴		
2228	腤	1	1	975D	818E	腤		
2256	鹹	3	3	9779	4D1A		鹵⿰亢	
2272	鼂	1	1	97AB	2A4DF		句⿱黽	
2281	髓	1	1	97B4	4BDA		⿸形骨幺月⿶	
2334	絃	1	1	97E9	25FED		糸⿰丟	
2346	鐧	1	1	97F5	9427	鐧		
2352	綢	1	1	97FB			糸⿰間	
2367	箋	2	2	984B	25C2D		竹⿱差	
2369	艅	1	1	984D	26A34		舟⿰同	
2371	玤	5	5	984F	3EB9		王⿰弁	
2400	忿	1	1	986C	5E09	忿		
2409	閧	3	3	9875	95A7	閧		

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
2427		4	4	98A9			山⿰郭	待造字。
2430	踖	1	1	98AC	473A		豆⿰昔	
2450	盧	1	1	98C0	354E		厂⿱盍	
2451		1	1	98C1			麥⿱果	麥⿰果，取代。
2454	鐓	1	1	98C4	28AFC		金⿰敖	
2456	鞣	2	2	98C6	29308		革⿰索	
2457	鞨	2	2	98C7	2935C		革⿰罩	
2460	駈	1	1	98CA	2988D		馬⿰瓜	
2473	譠	4	3	98D7	27B24		言⿰童	言⿰童，錯誤修改，手動取代。
2476	鶝	1	1	98DA	7588	鶝		
2477	騔	2	2	98DB	4BB7		馬⿰蜀	
2482	鉈	1	1	98E0	4929		金⿰宅	
2484	猘	1	1	98E2	3E6B		犭⿰夷	
2486	鯪	2	2	98E4	9BAC	鯪		
2488	骴	1	1	98E6	29A38		骨⿰出	
2489		1	1	98E7			⿰形氵⿱余甘⿴囗⿰	涂⿱甘，取代。

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
								待造字。
2491	鼬	1	1	98E9	2A545		鼠⿺乂屯	
2493	狩	1	1	98EB	24560		爻⿺乂守	
2494	麻需	1	1	98EC	2A3B1		麻⿺乂需	
2502	颱	1	1	98F4	98C8	颱		
2506	餪	1	1	98F8	991D	餪		
2508	鯣	1	1	98FA	29E44		魚⿺乂曷	
2510	鮟	1	1	98FC	9B8D	鮟		
2511	簊	1	1	98FD	25DE8		竹⿱今聶	
2519	翊	1	1	9946	26480		豈⿺乂羽	
2524	饍	1	1	994B	994D	饍		
2539	脞	1	1	995A	43F6		月⿺乂坐	
2540	諱	1	1	995B	46ED		言⿺乂幸	
2541	誡	4	4	995C	8A97	誡		
2546	略	2	2	9961	8849	略		
2547	彥	3	3	9962	3881		广⿱多	

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
2558	鵡	1	1	996D	2A026		式⿰鳥	
2562	幘	2	2	9971	385A		巾⿰冓	
2563	攝	1	1	9972	6502	攝		
2565	霪	1	1	9974	9714	霪		
2567	稡	4	4	9976	7A21	稡		
2568	沪	1	1	9977			氵⿰厄	
2570	鼽	1	1	9979	21C08		臭⿰兀	
2571		1	1	997A			⿸氵弗⿱鬲⿰	沸⿱鬲，取代。
2572	衒	1	1	997B	8858	衒		
2573		1	1	997C			王⿰弃	待造字。
2574	勵	1	1	997D	3899		广⿱勤	
2577	鋟	1	1	99A2	28967		金⿰乏	
2578	勖	1	1	99A3	20892		⿸角比力⿰	
2581	帋	1	1	99A6	383F		此⿱巾	
2582	癉	1	1	99A7	7612	癉		
2584	苜	1	1	99A9	44A4		艹⿱日	

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
2587	脜	2	2	99AC	80D2	脜		
2588	齾	1	1	99AD	9F7D	齾		
2591		1	1	99B0			口⿰龍	待造字。
2592	贓	1	1	99B1	8D1C	贓		
2593		1	1	99B2			言⿰光	待造字。
2595	髟	3	3	99B4	29AF9		髟⿱巾	
2613	醷	1	1	99C6	28898		酉⿰益	
2614	綢	1	1	99C7	368B		多⿰周	
2615	縈	1	1	99C8	4318		敝⿱糸	
2652	縻	1	1	99ED			戚⿱糸	
2653	攝	1	1	99EE	22E04		扌⿰翦	
2654	帾	1	1	99EF	3841		巾⿰夸	
2693		1	1	9A57			⿸形竹行缶示⿰	籲，取代。
2694	餇	1	1	9A58	296DE		飠⿰孚	
2695		1	1	9A59			兌⿰支	待造字。
3770	个	6	6	8141	4E2A	个		

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
3774	觿	1	1	8145	89FD	觿		
3775	岊	2	2	8146	21D52		巳⿱山山	
3778	併	1	1	8149	5002	併		
3796	决	2	1	815B	51B3	决		决，錯誤修改，手動取代。
3798	凷	1	1	815D			凷	
3801	剗	10	9	8160	5257	剗		剗，錯誤修改，手動取代。
3805	勅	1	1	8164	52C5	勅		
3810	匱	8	8	8169	5335	匱		
3816	厓	15	15	816F	5393	厓		
3819	厮	3	3	8172	53AE	厮		
3824	叙	1	1	8177	53D9	叙		
3829	咏	1	1	817C	548F	咏		
3830	咤	9	9	817D	54A4	咤		
3834	嗥	3	3	81A3	5637	嗥		Unicode 字型差異。
3838	鋒	1	1	81A7	2899F		金⿰夂夂夆	
3843	堦	1	1	81AC	5826	堦		

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
3844	堿	1	1	81AD	5884	堿		
3858	寃	10	10	81BB	5BC3	寃		
3866	崐	1	1	81C3	5D10	崐		
3869	帀	1	1	81C6	5E00	帀		
3892	睺	2	2	81DD	3B0B		日⿰侯	
3900	拕	1	1	81E5	62D5	拕		
3912	拵	1	1	81F1	62C3	拵		
3915	斅	1	1	81F4	6585	斅		Unicode 字型差異。
3930	椁	6	6	8244	6901	椁		
3932	樓	2	1	8246	6936	樓		樓，錯誤修改，手動取代。
3945		1	1	8253			氵⿰又	汊，取代。
3963	炁	2	2	8265	7081	炁		
3966	烟	2	2	8268	70DF	烟		
3980	猨	2	2	8276	7328	猨		
3985		5	5	827B			⿸形王丰母⿷一	瑀，取代。
3991	攜	2	1	82A3	64D5	攜		攜，手動取代。 txt 重複字句，致頻次多 xml

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
								一次。
3993	畀	1	1	82A5	2233F		畀	
3995	晦	2	2	82A7	756E	晦		
4012	皷	2	2	82B8	76B7	皷		
4015	眦	3	3	82BB	7726	眦		
4016	着	1	1	82BC	7740	着		
4027	禆	4	4	82C7	7986	禆		
4030	秆	1	1	82CA	79C6	秆		
4032	秣	2	2	82CC	2578A		禾⿰未未	秣，錯字。
4034	穉	2	2	82CE	7A49	穉		
4036	窰	3	3	82D0	7AB0	窰		
4037	寮	1	1	82D1	7AC2	寮		寮，錯字。
4045	箒	1	1	82D9	7B92	箒		
4055	粃	1	1	82E3	7C83	粃		
4059	糇	1	1	82E7	7CC7	糇		
4063	蘖	3	3	82EB	7CF5	蘖		
4070	緤	1	1	82F2	7DE4	緤		

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
4085	耻	1	1	8342	803B	耻		
4094	舘	3	3	834B	8218	舘		
4106	蕾	1	1	8357	8458	蕾		
4114	蘖	1	1	835F	8616	蘖		
4115	藁	24	21	8360	85C1	藁		藁，錯誤修改，手動取代。
4118	藥	1	1	8363	8602	藥		
4119		1	1	8364			形艹正頁巳夂⊡	夔，取代。
4124	蝨	2	2	8369	8771	蝨		
4131	衂	5	5	8370	8842	衂		
4136	衰	3	1	8375	887A	衰		衰，配合修改，手動取代。
4137	袜	3	3	8376	889C	袜		
4149	覷	1	1	83A4	89BC	覷		
4157	謝	1	2	83AC	8B3F	謝		并，部份錯字，手動取代。
4167	賁	1	1	83B6	8CEB	賁		
4172	蹙	2	2	83BB	8E77	蹙		
4178	輿	2	2	83C1	8F2D	輿		

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
4193	鈎	4	4	83D0	920E	鈎		
4198	鑛	3	2	83D5	945B	鑛		鑛，配合修改，手動取代。
4208	疇	1	1	83DF	9767	疇		
4209	靴	1	1	83E0	9789	靴		
4214	鞽	3	2	83E5	97C0	鞽		鞽，錯誤修改，手動取代。
4215	鞘	1	1	83E6	97D2	鞘		
4217	韠	2	2	83E8	97E0	韠		
4226	飡	6	6	83F1	98E1	飡		
4227	飣	2	1	83F2	98E6	飣		飣，錯誤修改，手動取代。
4234	贏	2	2	83F9	9A58	贏		
4247	鮎	1	1	8447	9B8E	鮎		
4250	鰌	2	2	844A	9C0C	鰌		
4252	鰥	16	13	844C			魚⿰丷眾	鰥，錯誤修改，手動取代。
4256	鷄	35	33	8450	9DC4	鷄		鷄，配合修改，手動取代。
4259	麕	4	4	8453	9E95	麕		
4267	鼈	6	6	845B	9F08	鼈		

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
4268	鼖	4	4	845C	9F17	鼖		
4271	齩	1	1	845F	9F69	齩		
4309	刼	1	1	84A7	523C	刼		
4313	敺	1	1	84AB	22FBA		匪⿰文	
4314	浛	3	3	84AC	6D5B	浛		
4342	倮	2	2	84C8	502E	倮		
4348	偬	1	1	84CE	506C	偬		
4380	驉	2	2	84EE	9A3C	驉		
4389	瑥	1	1	84F7	746B	瑥		
4392	儰	1	1	84FA	5136	儰		
4409	同	2	2	854C	518B	同		同，錯字。
4428	蹢	1	1	855F	8E70	蹢		
4436	碧	1	1	8567	40AE		刱⿱石	
4437	廩	1	1	8568	5EEA	廩		
4458	勸	1	1	857D	52CC	勸		Unicode 字型差異。
4465	磬	1	1	85A6	3C88		殼⿱召	

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
4495	麩	1	1	85C4	9EA8	麩		
4496	叚	1	1	85C5	53DA	叚		
4515	呪	8	8	85D8	546A	呪		
4536	瓞	1	1	85ED	3F10		扁⿰瓜	
4546	鵛	5	5	85F7	9D76	鵛		
4549	暎	1	1	85FA			日⿰奐	
4554	唬	1	1	8640	55C1	唬		
4592	猬	1	1	8666	247E4		犭⿰軍	
4618	鏌	3	1	86A2	4956		金⿰莫	金⿰莫，錯誤修改，手動取代。
4633	曶	1	1	86B1	3B2D		日⿰爵	
4649	蕢	1	1	86C1	8552	蕢		
4655	垻	2	2	86C7	57BB	垻		
4695	槲	1	1	86EF	6A8A	槲		
4704	牁	8	8	86F8	7271	牁		
4706	隥	2	1	86FA	96A5	隥		隥，錯誤修改，手動取代。
4707	悛	1	1	86FB	3944		忄⿰夋	

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
4715	噘	2	2	8744	5688	噘		
4784	鞋	1	1	87AB	9793	鞋		
4792	簾	1	1	87B3			竹⿱虛	
4822	彽	2	2	87D1	5C29	彽		
4890		3	3	8856			宀⿱幸	宰，錯字。
4907	槖	1	1	8867	69D6	槖		
4910		1	1	886A			氵⿰鹿	待造字。
4919		1	1	8873			亻⿰鹿	待造字。
4925	轡	1	1	8879	8F65	轡		
4934	鬘	1	1	88A4	9B09	鬘		
4948	奭	2	2	88B2	2308F		奭⿰斗	
4957	嵒	1	1	88BB	5D53	嵒		
4990	巖	1	1	88DC	5DBB	巖		
5029		2	2	8944			⿸形⿱口⿱門丁⿶⿰⿱⿰⿱	鬲，取代。
5034	弃	2	2	8949	5F03	弃		Unicode 字型差異。
5046	彝	1	1	8955	5F5C	彝		

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
5065	忼	3	3	8968	5FFC	忼		
5074	厖	1	1	8971	5C2B	厖		
5079		1	1	8976			忄⿰存	待造字。
5106	怷	3	3	89B3	5FF2	怷		
5108	愖	2	2	89B5	60FD	愖		
5111	旆	1	1	89B8	65BE	旆		
5123	愰	1	1	89C4	6130	愰		
5133	憙	3	3	89CE	6199	憙		
5144	禱	5	5	89D9	8920	禱		
5156	禾	1	1	89E5	353B		禾	
5189	宂	5	5	8A47	5B82	宂		
5219	妃	1	1	8A65			女⿰巳	
5220	毋	1	1	8A66	6BCC	毋		毋，取代。
5221		3	3	8A67			⿸形白米丐⿴囗·	粵，取代。
5239	擎	1	1	8A79	6486	擎		
5250		1	1	8AA6			㮇⿱手手	攀，錯字。

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
5292	專	1	1	8AD0	65C9	專		
5302	昏	3	3	8ADA	662C	昏		
5311	晄	1	1	8AE3	6644	晄		
5312		1	1	8AE4			形女 冫 冈 比⊙	形女 勹 比⊙，取代。
5324	裘	1	1	8AF0	88E0	裘		
5331	翕	2	10	8AF7	5963	翕		辶，用作部件，不轉碼。
5342	曆	2	6	8B43			形厂 林 日⊙	重碼。 4 例作「 阝 」，不轉碼； 2 例將「形厂 林 日⊙」依原書手動取代為「曆」。
5363	曄	1	1	8B58	66CE	曄		
5364	殦	1	1	8B59	239D1		歹 ⿰ 凶	
5402	术	1	1	8BA1	672F	术		
5407	簊	4	4	8BA6	7C3A	簊		
5408	紝	2	2	8BA7	7D4D	紝		
5434	湟	2	2	8BC1			形氵 白 工⊙	
5442	枅	1	1	8BC9	67BF	枅		
5450	飡	1	1	8BD1	6E4C	飡		

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
5452	輭	4	4	8BD3	8F5C	輭		
5463	峀	1	1	8BDE	65F9	峀		
5541	蠹	9	9	8C6D	8827	蠹		
5542	鑠	4	3	8C6E	93C1	鑠		鑠，配合修改，手動取代。
5549	藁	4	1	8C75	85F3	藁		藁，錯誤修改，手動取代。
5551	賣	3	2	8C77	8CF7	賣		賣，錯誤修改，手動取代。
5559	欸	1	1	8CA1	6B35	欸		
5562	稑	1	1	8CA4	7A09	稑		
5580	离	1	1	8CB6			离	
5584	愞	1	1	8CBA	611E	愞		
5595	垛	1	1	8CC5	579C	垛		
5605	阭	1	1	8CCF	28E30		阝⿰尸尼	阭，錯字。
5625	濦	1	1	8CE3	3D8F		氵⿰阝隱	
5626		3	2	8CE4			氵⿰阝卡	汴，錯字。 汴，錯誤修改，手動取代。
5666	汚	5	5	8D4D	6C5A	汚		

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
5673	郊	1	1	8D54	28678		支⿰卩	
5693	鰜	3	3	8D68	9E7B	鰜		
5697	煑	9	6	8D6C	7151	煑		煑，錯誤修改，手動取代。
5702	厨	1	1	8D71	53A8	厨		
5703	齔	1	1	8D72	9F53	齔		
5705	廐	1	1	8D74	5ED0	廐		Unicode 字型差異。
5707	桟	1	1	8D76	6830	桟		
5709	蜂	1	1	8D78	872F	蜂		
5710	稆	2	2	8D79	7A06	稆		Unicode 字型差異。
5723	牂	1	1	8DA8	4367		牛⿰羊	
5730	阨	3	2	8DAF	9638	阨		阨，錯誤修改，手動取代。
5742	澁	1	1	8DBB	6F81	澁		
5761	慮	1	1	8DCE			廣⿱心	
5769	郝	1	1	8DD6	90CD	郝		
5770	汧	5	4	8DD7			氵⿰开	氵⿰开，錯誤修改，手動取代。
5781		1	1	8DE2			⿸形氵⿱艹兩⿰⊙	滿，取代。

編號	造字	字頻(txt)	字頻(xml)	Big5	Unicode	WinXP	構字式	備註
5790	憪	2	2	8DEB	22880		忄 ⿰門間	

附件二、「新唐書.xml」手動取代表

編號	造字	取代	檔案原文摘錄	手動取代結果
1015	嶴		虜徙營九衆之陰 （錯誤修改）	虜徙營九嶴之陰 （手動取代）
			獵于九__	獵于九嶴
			營九__山為陵	營九嶴山為陵
			帝為築觀九__山	帝為築觀九嶴山
1109	璪		間以衛貝珊瑚 （錯誤修改）	間以璪貝珊瑚 （手動取代）
			飾以金珠__玉	飾以金珠璪玉
2473	譿		李周幷，乃引去 （錯誤修改）	李周⿰言門童，乃引去 （手動取代）
			內樞密使李周__	內樞密使李周⿰言門童
			樞密使李周__	樞密使李周⿰言門童
			斬重遂、周__	斬重遂、周⿰言門童

編號	造字	取代	檔案原文摘錄	手動取代結果
3796	决		欲引衛 （錯誤修改）	欲引决 （手動取代）
			未聞雪冤庠事	未聞雪冤决事
3801	剡		衛多難 （錯誤修改）	剡多難 （手動取代）
			埶（全部取代）	剡（全部取代）
3932	樓		為衛櫚賦自託 （錯誤修改）	為樓櫚賦自託 （手動取代）
			豹席穷屩	豹樓穷屩
3991	携		范陽有闰	范陽有携
4115	藁	藁	巡縛衛為人千餘	巡縛藁為人千餘
			久，乃衛 （以上錯誤修改）	久，乃藁 （以上手動取代）
			凡租及艹厽稾秸	凡租及藁秸（Unicode 字型取代）
			樊（全部取代）	藁（全部取代）
4136	衰		蓑人穢夫爭候	衰人穢夫爭候
4157	謝		脅奚眾弍附突厥 （錯誤修改）	脅奚眾并附突厥 （手動取代）

編號	造字	取代	檔案原文摘錄	手動取代結果
			詔學士弍之	詔學士謝之
4198	鑛		金在茨何足貴	金在鑛何足貴
			丹彩紫茨	丹彩紫鑛
4214	韉		以蒲衛乘牛 （錯誤修改）	以蒲韉乘牛 （手動取代）
			百獸毛為回面	百獸毛為韉面
			五百騎釘轡藻回	五百騎釘轡藻韉
4227	飧		衛鬻不給 （錯誤修改）	飧鬻不給 （手動取代）
			瓦鼎為菫	瓦鼎為飧
4252	鰥	鰥	頛（全部取代）	鰥（全部取代）
			恤幷寡 （以下錯誤修改）	恤鰥寡 （以下手動取代）
			以幷寡少者進考	以鰥寡少者進考
			免幷寡惸獨不濟者	免鰥寡惸獨不濟者
4256	鷄		奎（全部取代）	鷄（全部取代）
4618	鏌		宇文華㠯一具	宇文華金⋏奠一具
			褫冠衛待罪	褫冠金⋏奠待罪

編號	造字	取代	檔案原文摘錄	手動取代結果
			（以下錯誤修改）	（以下手動取代）
			復冠衜乃入	復冠金从奠乃入
4706	隥		衜彴橫邪 （錯誤修改）	隥彴橫邪 （手動取代）
			丮道曲狹	隥道曲狹
5342	曆	曆	大卩、貞元間	大曆、貞元間
			大卩初	大曆初
5542	鏁		鉤鍵關启	鉤鍵關鏁
			以鐵启維車	以鐵鏁維車
			鐵启頸	鐵鏁頸
5549	藁		其令有司設券席	其令有司設藁席
			畋漁、芻幷 （以下錯誤修改）	畋漁、芻藁 （以下手動取代）
			艮賤、芻幷	艮賤、芻藁
			芻幷、運漕	芻藁、運漕
5551	賫		故敕有司裝衜視 （錯誤修改）	故敕有司裝賫視 （手動取代）
			又遣使段金帛	又遣使賫金帛

編號	造字	取代	檔案原文摘錄	手動取代結果
			聿叚五城	聿賫五城
5626	氵△△ 卡	汴	尚衡擊衞州賊 （錯誤修改）	尚衡擊汴州賊 （手動取代）
			李宀△介亂扌州	李宀△介亂汴州
			與扌人戰琵琶溝	與汴人戰琵琶溝
5697	煑		帍（全部取代）	煑（全部取代）
			衞紙而食 （以下錯誤修改）	煑紙而食 （以下手動取代）
			衞鎧弩以食	煑鎧弩以食
			山東衞海之饒	山東煑海之饒
5730	陒		或以困衞歸者 （錯誤修改）	或以困陒歸者 （手動取代）
			成而玄堂埄	成而玄堂陒
			且地狹埄	且地狹陒
5770	汧		銳兵三千夜入衆陽 （錯誤修改）	銳兵三千夜入氵△△幵陽 （手動取代）
			李晟敗之于冗陽	李晟敗之于氵△△幵陽
			吐蕃寇冗陽	吐蕃寇氵△△幵陽

編號	造字	取代	檔案原文摘錄	手動取代結果
			趨隴、冗陽間	趨隴、氵 ⿰幺幺开陽間
			剽冗陽、華亭	剽氵 ⿰幺幺开陽、華亭

附件三、「新唐書.xml」缺字外錯字修改表

編號	檔案原文摘錄	手動取代結果
1	奇禽怪羽為 園	奇禽怪羽為園
2	修城郡公言⿰幺幺辶⿱⿰幺幺已	修城郡公言⿰幺幺眾