

# 《南史(列傳)》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  
2008/9/10 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 861 的「𪗇」字，Unicode 編碼是 4183，由於 XP 並不支援，仍需使用構字式「道𪗇禾」。

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：

《南史(列傳)》使用舊版造字 336 個，字頻 4044 次，這 336 個造字中，267 個可轉成 Windows XP 能顯示的字，字頻 3805 次；另外 69 個字必須轉成構字式，字頻 239 次。

轉碼完成製作轉碼分析表，請參考附件一《南史(列傳)》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 2399 的「金𠄎樂」字，而原書使用之字為「鑠」，查詢教育部異體字字典確認「鑠」並非「金𠄎樂」之異體字，因此歸為錯字，以程式取代為正確字形「鑠」。
  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 4878 的「宀」字，Unicode 字型呈現為「宀」，實際上仍為同一字。
  4. 待造字：Unicode 及漢字構形資料庫皆未收錄的舊漢籍造字，正在等待補造字中，所以「造字」欄空白無法看到字形。如編號 2218 的「魚𠄎賞」。

附件一、《南史(列傳)》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
38		1	FA65			王𠄎已	玳，異體字問題。
814	墓	4	8E5C	87C7	墓		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
821	脊	2	8E63	7718	脊		
826	癱	4	8E68	7655	癱		
861	導	1	8EAD	4183		道彳禾	
887	鸞	2	8EC7	9E0E	鸞		
896	櫛	1	8ED0	6AE9	櫛		
910	崧	2	8EDE	5D78	崧		
938	燮	3	8EFA	7215	燮		
940	爐	1	8EFC	7208	爐		
964	隰	2	8F55	28EF3		阝厶焉	
977	磨	1	8F62	9E8F	磨		
978	戡	1	8F63	4CD2		戠△鳥	
1039	稟	2	8FC2	69C0	稟		
1044	儻	12	8FC7	511B	儻		
1053	傲	13	8FD0	6160	傲		
1056	阨	22	8FD3	9641	阨		
1067	棊	54	8FDE	68CA	棊		
1077	稟	4	8FE8	7A3E	稟		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1091	漣	6	8FF6	58CD	漣		
1111	瓚	8	904B	749D	瓚		
1133	瓌	36	9061	74CC	瓌		
1158	咳	2	907A	7561	咳		
1205	眇	4	90CB	770E	眇		
1237	稍	36	90EB	77DF	稍		
1278	碍	3	9155	788D	碍		
1350	褚	1	91BF	7987	褚		褚，錯字，以程式全部取代。
1531	幼	2	92D7	7CFC	幼		幼，異體字問題。
1544	絇	1	92E4	7D59	絇		
1565	繇	8	92F9	7DDC	繇		
1591	羗	1	9354	7F97	羗		
1642	脉	2	93A9	8109	脉		
1690	菘	1	93D9	82EE	菘		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1784	萼	2	9478	855A	萼		
1834	蕉	2	94CC	85EE	蕉		
1885	蝨	1	9540	882D	蝨		
1893	袞	13	9548	886E	袞		
1900	桂	2	954F	88BF	桂		
1926	覓	1	9569	8994	覓		
1944	簪	20	957B	8A67	簪		
2077	鞞	2	9663	8F4A	鞞		
2193	氈	1	96F9	3CA3		叟 厶 毛	
2198	暉	22	96FE	23236		日 厶 恆	
2209	餅	3	974A	4B33		食 厶 半	
2211	朧	1	974C	2662A		月 厶 匕	
2212	襪	3	974D	277CC		衤 厶 羅	
2214	櫻	1	974F	466C		衤 厶 嬰	
2216	紉	10	9751	25F9D		纟 厶 丂	
2217	瓠	1	9752	24B1B		盧 厶 瓜	
2218		1	9753			魚 厶 賞	待造字。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
2219	慙	1	9754	22754		則ㄉ心	
2220	覓	1	9755	26CD3		艹ㄉ鳥 ㄉ几	
2221	嶷	1	9756	21F9E		山ㄉ義	
2222	菝	1	9757	44C8		艹ㄉ役	
2223	腕	2	9758	4BDB		骨ㄉ宛	
2225	穢	1	975A	88D3	穢		
2226	劓	2	975B	3505		肅ㄉ刂	
2228	腴	1	975D	818E	腴		
2399		1	986B			金ㄉ樂	鑠，錯字，以 程式全部取 代。
2478	鞏	1	98DC	978C	鞏		
2490	頷	1	98E8			名ㄉ頁	
2502	颯	4	98F4	98C8	颯		
2602	鱈	3	99BB	9C10	鱈		
2690	氈	26	9A54	23C30		爽ㄉ毛	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3769		2	8140			𠄎木西 示寸𠄎	標 𠄎 寸，錯字，以程式全部取代。
3778	併	4	8149	5002	併		併，異體字問題。
3780	儻	34	814B	5101	儻		
3787	冢	62	8152	51A2	冢		
3789	冲	32	8154	51B2	冲		
3792	涼	2	8157	51C9	涼		涼，異體字問題。
3793	減	1	8158	51CF	減		減，異體字問題。
3794	湊	1	8159	51D1	湊		湊，異體字問題。
3800		4	815F			杀 𠄎 刂	刹，異體字問題。
3801	剗	2	8160	5257	剗		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3804	効	51	8163	52B9	効		
3806	勳	7	8165	52E0	勳		
3812	𠂇	1	816B	20BCC		𠂇	
3814	却	28	816D	5374	却		
3817	廁	5	8170	53A0	廁		廁，異體字問題。
3819	廡	5	8172	53AE	廡		廡，異體字問題。
3820	厦	3	8173	53A6	厦		
3821	梛	1	8174	23485		木𠂇	
3823	擦	1	8176	6456	擦		𠂇扌𠂇示，錯字，以程式全部取代。
3826	叶	2	8179	53F6	叶		
3829	咏	1	817C	548F	咏		
3835	噍	1	81A4	564D	噍		
3836	喻	1	81A5	564F	喻		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3839	舩	1	81A8	4480		舟 𦨭 符	
3840	坂	6	81A9	5742	坂		
3843	堦	1	81AC	5826	堦		
3846	壻	13	81AF	58FB	壻		
3849	獎	25	81B2	596C	獎		獎，異體字問題。
3853	褻	1	81B6	2776C		敝 衤 衣	
3858	冤	14	81BB	5BC3	冤		冤，異體字問題。
3860	尠	2	81BD	5C20	尠		
3864	峯	7	81C1	5CEF	峯		
3865	瀟	36	81C2	24162		灬 𦨭 蕭	
3866	崐	2	81C3	5D10	崐		
3868	卮	3	81C5	5DF5	卮		
3869	币	1	81C6	5E00	币		
3874		1	81CB			忄 𦨭 稟	慄，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3877	迺	4	81CE	5EFC	迺		
3879	警	1	81D0	27AF7		軟ㄟ言	軟ㄟ言，錯字，以程式全部取代。
3883	徧	49	81D4	5FA7	徧		
3884	忽	8	81D5	6031	忽		
3889	慤	17	81DA	6164	慤		
3892	喉	3	81DD	3B0B		日ㄩ侯	
3903	筆	1	81E8	25BE4		ㄨㄥㄟ韋	
3904	墟	1	81E9	588C	墟		
3910	擘	1	81EF	64E5	擘		
3915	斃	39	8459	6585	斃		
3919	𧈧	1	81F8	23C0B		𧈧ㄩ毛	𧈧ㄩ毛，異體字問題。
3920	皓	4	81F9	6667	皓		
3921	皙	10	81FA	6673	皙		
3924	朞	41	81FD	671E	朞		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3929	梨	1	8243	68C3	梨		
3930	捰	4	8244	6901	捰		
3931	椀	4	8245	6900	椀		
3935	槩	3	8249	3BA3	槩		Unicode 字型 呈現差異。
3937	檝	6	824B	6A9D	檝		
3944	氷	1	8252	6C37	氷		
3949	涖	1	8257	6D96	涖		
3966	烟	4	8268	70DF	烟		
3974	牀	122	8270	7240	牀		
3975	牕	1	8271	7255	牕		
3977	狗	1	8273	72E5	狗		
3978	猪	13	8274	732A	猪		Unicode 字型 呈現差異。
3980	獾	4	8276	7328	獾		
3981	珉	11	8277	73C9	珉		
3983	琕	26	8279	7415	琕		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3984	瑠	5	827A	7460	瑠		
3985	璫	4	827B	7447	璫		
3987	穀	7	827D	7474	穀		
3990	甄	1	82A2	750E	甄		
3991	攜	3	82A3	64D5	攜		
3996	疎	60	82A8	758E	疎		
4001	瘧	1	82AD	7602	瘧		
4005		1	82B1			鬼△天	待造字。
4013	盃	2	82B9	76CC	盃		
4015	眦	2	82BB	7726	眦		
4016	着	6	82BC	7740	着		
4022	甯	3	82C2	3CB0		寧△毛	
4023	碓	1	82C3	78D3	碓		
4028	稟	17	82C8	7980	稟		
4029	稊	1	82C9	799D	稊		
4032	秣	1	82CC	2578A		禾△未	
4035	穽	3	82CF	7A7D	穽		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4038	竈	5	82D2	7AC8	竈		
4041	豎	9	82D5	7AEA	豎		
4042	笋	2	82D6	7B0B	笋		
4044	筋	8	82D8	7B6F	筋		
4045	箒	1	82D9	7B92	箒		
4052	簾	1	82E0	7C34	簾		竹 厶 虞，異體字問題。
4055	粃	1	82E3	7C83	粃		
4063	蘖	1	82EB	7CF5	蘖		
4064		1	82EC			艹 厶 主 厶 母	艹 厶 毒，異體字問題。
4077	罇	2	82F9	7F47	罇		
4078		28	82FA			巳 厶 升	异，異體字問題。
4080	羣	244	82FC	7FA3	羣		
4081	翱	1	82FD	7FFA	翱		
4085	耻	7	8342	803B	耻		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4087	鮎	4	8344	4C47		魚𩺰旦	
4090	脚	51	8347	811A	脚		
4091	臯	13	8348	81EF	臯		
4094	館	1	834B	8218	館		
4103	苙	1	8354	831D	苙		
4104	菘	30	8355	8385	菘		
4107	葱	2	8358	8471	葱		
4114	藜	2	835F	8616	藜		
4115	藁	3	8360	85C1	藁		
4129	虬	32	836E	866C	虬		
4131	𠂔	4	8370	8842	𠂔		
4132		726	8371			血𠂔𠂔	眾，異體字問題。
4134	衛	491	8373	885E	衛		衛，異體字問題。
4137	袜	1	8376	889C	袜		
4138	裘	2	8377	88A0	裘		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4139	袴	29	8378	88B4	袴		
4140	袵	1	8379	88B5	袵		
4142	褰	38	837B	8943	褰		
4145	羈	1	837E	898A	羈		
4146	覩	19	83A1	89A9	覩		睹，異體字問題。
4157	譎	3	83AC	8B3F	譎		
4160	適	8	83AF	8B81	適		
4172	屨	1	83BB	8E77	屨		
4173	駝	1	83BC	8EAD	駝		
4177	駟	1	83C0	8F27	駟		
4184	迹	89	83C7	8FF9	迹		
4185	邊	10	83C8	9037	邊		
4193	鈎	3	83D0	920E	鈎		鈎，異體字問題。
4195	鉢	8	83D2	9262	鉢		
4203	隣	42	83DA	96A3	隣		鄰，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							題。
4210	愉	1	83E1	2468E		牛 厶 俞	
4213	鞞	1	83E4	97BE	鞞		
4218	菲	4	83E9	97EE	菲		
4222	頰	3	83ED	9847	頰		
4225	顛	1	83F0	9866	顛		
4226	凜	5	83F1	98E1	凜		
4240	髯	9	8440	9AF4	髯		
4241	萁	1	8441	84B5	萁		
4242	鬪	28	8442	9B2D	鬪		
4249	鯪	1	8449	9BFF	鯪		
4250	鱈	4	844A	9C0C	鱈		
4254	鷓	10	844E	9D7E	鷓		
4256	鷄	3	8450	9DC4	鷄		雞，異體字問題。
4259	麇	2	8453	9E95	麇		
4260	麇	7	8454	9E9E	麇		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4267	鼈	2	845B	9F08	鼈		
4275	义	1	8463	4E49	义		
4278	亘	7	8466	4E98	亘		
4282	黠	2	846A	2A4A0		黑厶敢	
4296	堰	1	8478	5878	堰		
4300	臙	1	847C	267A6		月厶雀	
4304	舩	1	84A2	26A35		舟厶行	
4308	刼	2	84A6	5227	刼		
4310		1	84A8			革厶健	待造字。
4314	滄	1	84AC	6D5B	滄		
4340	撮	1	84C6	3A67		扌厶暴	
4342	倮	8	84C8	502E	倮		
4348	惣	1	84CE	506C	惣		
4358	髻	1	84D8	4BF0		髟厶介	
4371	倜	2	84E5	50F4	倜		
4428	躡	1	855F	8E70	躡		
4436	劦	2	8567	40AE		勹厶石	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4437	廩	1	8568	5EEA	廩		
4439		1	856A			形足巳 心	蹠，異體字問題。
4442	潛	1	856D	6F98	潛		
4448		3	8573			形糸厂 墨	纏，錯字，以程式全部取代。
4458	勑	1	857D	52CC	勑		
4460	勑	1	85A1	52D1	勑		
4461	勑	69	85A2	52D4	勑		
4477		1	85B2			形艹勺 一	芍，錯字，以程式全部取代。
4507	吳	1	85D0	5434	吳		吳，異體字問題。
4511	味	1	85D4	548A	味		
4515	呪	16	85D8	546A	呪		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4535	纘	1	85EC	7E89	纘		
4536	𧈧	1	85ED	3F10		扁𠂔瓜	𧈧，錯字，以程式全部取代。
4537	絹	10	85EE	7DED	絹		
4542	塿	2	85F3	21484		土𠂔稟	
4546	鷓	1	85F7	9D76	鷓		
4558	厩	1	8644	53A9	厩		厩，異體字問題。
4565	嶠	1	864B	5D74	嶠		
4575	縊	2	8655	7E7F	縊		
4603	憇	1	8671	6187	憇		
4605	橙	15	8673	6195	橙		
4607	狫	2	8675	3E60		豸𠂔屯	
4608	𦉳	1	8676	7581	𦉳		
4617	禡	1	86A1	4671		礻𠂔屬	
4622	詒	1	86A6	8A43	詒		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4633	嚼	10	86B1	3B2D		日 𠂔 爵	
4671	鷓	1	86D7	4D0F		燕 𠂔 鳥	
4699	暄	11	86F3	6685	暄		
4711	隧	2	8740	2145E		土 𠂔 遂	
4735	壠	2	8758	58E0	壠		
4753		3	876A			𠂔 𠂔 比 𠂔 大	待造字。
4768	妬	15	8779	59AC	妬		妒，異體字問題。
4779	鞣	1	87A6	97B2	鞣		
4785	贗	1	87AC	8D0B	贗		
4808	健	1	87C3	3493		亻 𠂔 達	食 𠂔 半，錯字，以程式全部取代。
4822	𠂔	2	87D1	5C29	𠂔		
4836	嗜	3	87DF	20D32		口 𠂔 沓	
4845		1	87E8			女 𠂔 間	嫺，異體字問

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							題。
4846	愷	1	87E9	398E		忄 厶 畫	
4870	壘	1	8842	97F2	壘		
4878	宄	4	884A	5B82	宄		Unicode 字型 呈現差異。
4888	荔	1	8854	8318	荔		
4889	毳	1	8855	3CAA		登 厶 毛	
4907	橐	1	8867	69D6	橐		
4950	嶮	1	88B4	5D18	嶮		
4951	礪	1	88B5	7930	礪		
4966	痲	1	88C4	764E	痲		
4980	藝	1	88D2			藝 厶 火	
5008	幪	1	88EE	5E5E	幪		
5026		1	8941			形 广 回 示 口	廩，錯字，以 程式全部取 代。
5031	閨	1	8946	95B4	閨		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5054	獾	1	895D	3E90		犴 ㄉㄨㄢ	
5106	忤	1	89B3	5FF2	忤		
5111	旆	5	89B8	65BE	旆		
5116	侑	4	89BD	4F37	侑		
5119	博	1	89C0	613D	博		
5121		2	89C2			形忤 主 方文 ㄉㄨㄢ	傲，錯字，以 程式全部取 代。
5126	暮	1	89C7	8B29	暮		
5132	愜	5	89CD	6198	愜		
5133	愜	2	89CE	6199	愜		
5136	嶰	1	89D1	5D83	嶰		
5169	危	1	89F2	6239	危		
5176		5	89F9			汜 ㄉㄨㄢ	沉，異體字問 題。
5192	隙	1	8A4A	969F	隙		
5195		6	8A4D			形木 ㄉㄨㄢ	檀，異體字問

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
						且 	題。
5209	摑	2	8A5B	6386	摑		
5227		3	8A6D			 戶戈 木 	屎  戈，異體字問題。
5230		6	8A70			 土亩 且 	壇，異體字問題。
5236	鯪	1	8A76	4C52		魚  邑	
5289	慤	29	8ACD	6164	慤		
5290	携	3	8ACE	643A	携		
5298	昨	4	8AD6	65FF	昨		
5315	汧	1	8AE7	3CC2		汧  公	
5316	晉	1	8AE8	664B	晉		晉，異體字問題。
5338	暎	9	8AFE	668E	暎		
5360	鄺	1	8B55	913D	鄺		
5401	齋	1	8B7E	2A5CB		齊  衣	
5434	湟	10	8BC1	6E7C	湟		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5437	并	146	8BC4	5E77	并		
5439	栢	28	8BC6	6822	栢		
5442	枿	2	8BC9	67BF	枿		
5452	輻	9	8BD3	8F5C	輻		
5456		1	8BD7			韋厶未	韋，異體字問題。
5459	缺	1	8BDA	21647		垂厶夂	
5466	倏	1	8BE1	5010	倏		
5478	飴	1	8BED	4D77	飴		
5483	眡	1	8BF2	7721	眡		
5539	醜	2	8C6B	29C5B		鬲厶變	醜，異體字問題。
5547	廢	4	8C73	4560		艹厶廢	
5549	藁	2	8C75	85F3	藁		
5551	賈	3	8C77	8CF7	賈		
5559	欸	6	8CA1	6B35	欸		
5585	寓	39	8CBB	3762		宀厶禹	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5589	廚	1	8CBF	2228D		𠩺广豆 寸	廚，異體字問題。
5591	杞	6	8CC1	233CC		木	杞，異體字問題。
5594	异	97	8CC4	5F02	异		
5609	沟	4	8CD3	6C6E	沟		
5653	曷	1	8D40	3B09		日	
5666	污	4	8D4D	6C5A	污		
5695	勝	1	8D6A	3B3A		朕	
5697	煮	6	8D6C	7151	煮		
5700	崩	4	8D6F	447D		舟	
5702	厨	27	8D71	53A8	厨		
5704	氈	4	8D73	6C0A	氈		氈，異體字問題。
5705	廐	6	8D74	5ED0	廐		廐，異體字問題。
5706	敘	34	8D75	654D	敘		敘，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
							題。
5707	棧	1	8D76	6830	棧		
5708		2	8D77			麥△井	麥△并，錯字，以程式全部取代。
5709	蟀	1	8D78	872F	蟀		
5729	滌	2	8DAE	6F34	滌		
5730	隄	1	8DAF	9638	隄		
5753	潛	1	8DC6	6F5C	潛		
5775	湯	1	8DDC	7003	湯		

## 二、手動取代表說明：

《南史(列傳)》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，請參考附件二欄位說明如下：

1. 檔案原文摘錄：摘錄檔案《南史(列傳)》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。

### 附件二、《南史(列傳)》缺字外錯字修改

檔案原文摘錄	手動取代結果	備註
與儀同黃 <u>爽</u> △毛法	與儀同黃 <u>法</u> 爽△毛	