

《虛堂和尚語錄》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
2010/03/31 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 2534 的「堊」字，Unicode 編碼是 213BA，由於 XP 並不支援，仍需使用構字式「祝_厶土」。

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：

《虛堂和尚語錄》使用舊版造字 148 個，字頻 1179 次，這 148 個造字中，131 個可轉成 Windows XP 能顯示的字，字頻 1153 次；另外 17 個字必須轉成構字式，字頻 26 次。

轉碼完成製作轉碼分析表，請參考附件一《虛堂和尚語錄》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 5453 的「涅」字，是「涅」的異體字，以標準字「涅」取代。
2. 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 4512 的「口𠂔它」。
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 4262 的「麪」字，Unicode 字型呈現為「麪」，實際上仍為同一字。

附件一、《虛堂和尚語錄》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
15	畝	2	FA4E	7546	畝		
820	烟	1	8E62	70F1	烟		
881	颺	1	8EC1	98C7	颺		
895	烽	1	8ECF	71A2	烽		
896	櫛	1	8ED0	6AE9	櫛		
1067	碁	2	8FDE	68CA	碁		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1146	甕	1	906E	7506	甕		
1341	豨	1	91B6	8C7C	豨		
1354	禩	1	91C3	79A9	禩		
1394	啗	2	91EB	35D6		口 厶 音	
1400	窓	17	91F1	7A93	窓		
1409	寵	1	91FA	7AC9	寵		
1491	擲	5	92AF	6527	擲		
1536	麇	12	92DC	9E81	麇		
1591	羴	1	9354	7F97	羴		
1646	脗	1	93AD	8117	脗		
1664	腸	2	93BF	8193	腸		
1706	爍	1	93E9	242B9		火 厶 勃	火 厶 字，異體字問題。
1772	藜	2	946C	4527		艹 厶 梨	
1784	萼	1	9478	855A	萼		
1809	囡	5	94B3	56E6	囡		
1823	鬚	1	94C1	9B00	鬚		
1854	啣	2	94E0	20E95		口 厶 留	
1905	棍	10	9554	88E9	棍		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1999	贊	9	95D4	8CDB	贊		
2046	踪	2	9644	8E2A	踪		
2066	滌	1	9658	3D11		氵 厶 席	
2204	癆	1	9745	20890		疒 牙 女 历 口	
2232		2	9761			疒 厂 墨 卩 口	廊，異體字問題。
2350	慘	1	97F9	50AA	慘		
2386	瑯	12	985E	7458	瑯		
2409	闕	1	9875	95A7	闕		
2505	翻	2	98F7	98DC	翻		
2534	壑	2	9955	213BA		祝 会 土	
2602	鱈	2	99BB	9C10	鱈		
2669	獯	2	99FE	7379	獯		
2758	献	3	9ABA	732E	献		
2899		2	9BAA			亠 会 粉 会 大	亠 会 粉 会 大，異體字問題。
3004	籊	1	9C54	7BD0	籊		
3005	糍	3	9C55	7CCD	糍		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3789	冲	1	8154	51B2	冲		冲，異體字問題。
3801	剗	4	8160	5257	剗		
3803	飀	1	8162	4B54		食 厶 追	
3804	効	3	8163	52B9	効		
3805	勅	9	8164	52C5	勅		
3808	勾	2	8167	5303	勾		
3814	却	172	816D	5374	却		
3816	厓	1	816F	5393	厓		
3819	廡	8	8172	53AE	廡		廡，異體字問題。
3820	厦	1	8173	53A6	厦		
3824	叙	1	8177	53D9	叙		
3829	咏	1	817C	548F	咏		
3843	堦	2	81AC	5826	堦		
3858	寃	11	81BB	5BC3	寃		寃，異體字問題。
3864	峯	262	81C1	5CEF	峯		
3868	卮	1	81C5	5DF5	卮		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3877	迺	1	81CE	5EFC	迺		
3884	忽	1	81D5	6031	忽		
3900	挖	4	81E5	62D5	挖		
3909	擡	10	81EE	64E1	擡		
3931	椀	11	8245	6900	椀		
3932	椶	2	8246	6936	椶		
3940	欝	5	824E	6B1D	欝		
3944	氷	16	8252	6C37	氷		
3955	澗	4	825D	3D4E		汜 𠄎 間	澗，異體字問題。
3968	煨	4	826A	7145	煨		
3978	猪	5	8274	732A	猪		Unicode 字型呈現差異。
3979	猫	7	8275	732B	猫		
3984	瑠	1	827A	7460	瑠		
3990	甄	2	82A2	750E	甄		
3992	嘗	16	82A4	751E	嘗		
3994	畊	6	82A6	754A	畊		
3996	疎	50	82A8	758E	疎		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4008	癯	3	82B4	766F	癯		
4010	臯	1	82B6	7690	臯		臯，異體字問題。
4035	穽	2	82CF	7A7D	穽		
4038	竈	6	82D2	7AC8	竈		
4041	豎	21	82D5	7AEA	豎		
4042	笋	1	82D6	7B0B	笋		
4053	秃	1	82E1	20636		不会几	
4067	綉	1	82EF	7D89	綉		
4081	翱	4	82FD	7FFA	翱		
4089	脇	2	8346	8107	脇		
4090	脚	98	8347	811A	脚		
4105	菓	5	8356	83D3	菓		
4107	葱	1	8358	8471	葱		
4129	虬	1	836E	866C	虬		
4134	衛	3	8373	885E	衛		衛，異體字問題。
4139	袴	5	8378	88B4	袴		
4140	裱	1	8379	88B5	裱		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4144	霸	1	837D	8987	霸		Unicode 字型 呈現差異。
4145	羈	1	837E	898A	羈		Unicode 字型 呈現差異。
4146	覩	12	83A1	89A9	覩		
4184	迹	1	83C7	8FF9	迹		
4195	鉢	35	83D2	9262	鉢		
4203	隣	7	83DA	96A3	隣		鄰，異體字問 題。
4207	霧	1	83DE	9736	霧		
4223	顛	1	83EE	984B	顛		
4226	凜	2	83F1	98E1	凜		
4237	髡	1	83FC	9AE0	髡		
4240	髯	4	8440	9AF4	髯		
4247	鮎	1	8447	9B8E	鮎		
4251	鰕	8	844B	9C15	鰕		
4256	鷄	8	8450	9DC4	鷄		
4262	麪	1	8456	9EAA	麪		Unicode 字型 呈現差異。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4267	鼈	10	845B	9F08	鼈		
4272		4	8460			厂△龍	龐，異體字問題。
4278	亘	1	8466	4E98	亘		
4348	惚	1	84CE	506C	惚		
4447	捏	2	8572	63D1	捏		
4448		4	8573			形糸厂 墨□	纏，異體字問題。
4482	厖	1	85B7	53AA	厖		
4485	廛	1	85BA	3553		厂△墨	
4487	寢	2	85BC	41AE		穴会浸	
4504	呶	2	85CD	541A	呶		
4512		1	85D5			口△它	待造字。
4515	呪	10	85D8	546A	呪		
4566	礪	3	864C	78F5	礪		
4714	廛	2	8743	222E8		广△墨	
4735	壠	1	8758	58E0	壠		
4738	睺	1	875B	777A	睺		
4768	妬	1	8779	59AC	妬		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4778	鞞	6	87A5	97B4	鞞		
4870	壘	1	8842	97F2	壘		
4903	吋	2	8863	5C00	吋		
4957	崑	35	88BB	5D53	崑		
4969	嶋	1	88C7	5D8B	嶋		
4996	斲	1	88E2	65B5	斲		
5163	憂	2	89EC	621E	憂		
5289	慤	1	8ACD	6164	慤		
5290	携	12	8ACE	643A	携		
5300	鬪	15	8AD8	9B2A	鬪		
5315	汧	2	8AE7	3CC2		汧 𠂇	
5316	晉	2	8AE8	664B	晉		晉，異體字問題。
5360	廩	1	8B55	913D	廩		
5453	涅	19	8BD4	23D40		形 汧 日 工 冫	涅，異體字問題。
5555	迥	5	8C7B	9008	迥		
5556	隸	1	8C7C	96B7	隸		
5559	欸	19	8CA1	6B35	欸		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5590	朶	1	8CC0	6736	朶		
5591	杞	1	8CC1	233CC		木𠂇巳	杞，異體字問題。
5595	塚	2	8CC5	579C	塚		
5664	搯	1	8D4B	3A5C		扌𠂇監	
5705	廐	2	8D74	5ED0	廐		廐，異體字問題。
5721	頤	1	8DA6	9814	頤		
5742	澁	3	8DBB	6F81	澁		
5753	潛	9	8DC6	6F5C	潛		
5783	教	2	8DE4	3A8D		孛𠂇攵	

二、手動取代表說明：

《虛堂和尚語錄》檔案中出現「●」字形，共計69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. 檔案原文摘錄：摘錄《虛堂和尚語錄》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、《虛堂和尚語錄》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 986	第三●	灑	
p. 986	團●	團	有 3 處
p. 988	其由 · ●●●●底	米△吳 米△吳 士△吳 士△吳	
p. 988	闍梨莫●速	斯△言	
p. 989	●訛	誦	有 2 處
p. 991	息畊●	灾△文	
p. 993	頸云 · ●	口△力	
p. 994	力●唏	口△力	
p. 995	●酒糟	噶	有 4 處
p. 997	當路坐●	灑	
p. 998	主丈 · 月●	灑	
p. 998	冬來事●	灑	
p. 998	無闍梨●	灑	
p. 998	●徇	彳△京	有 3 處
p. 999	● · 師云	灑	
p. 999	摸●	扌△索	
p. 1000	向汝道●	灑	
p. 1002	磨便出●	灑	

頁碼	檔案原文摘錄	原書字形	備註
p. 1002	一塊泥●	土△字	
p. 1005	若不是●	彘	
p. 1005	句歸何處●	擻	
p. 1005	有甚●●	饌饌	
p. 1006	眼●●地	目△虫 目△虫	
p. 1012	隈隈●●	毛△崔 毛△崔	
p. 1012	親承記●根	荊	
p. 1012	道·●·爾	口△力	
p. 1013	白髮●●箇	毛△參 毛△參	
p. 1016	得全●	擻	
p. 1016	祖云·也好●	擻	
p. 1022	被●上座	彘	
p. 1023	優波●多	穉	
p. 1023	湍●	汜△盾	待造字
p. 1024	後底●	擻	
p. 1025	諸方●徒	匸△于	
p. 1028	●跳	足△字	有 2 處
p. 1028	處云·牛●	擻	

頁碼	檔案原文摘錄	原書字形	備註
p. 1029	●上堂參	叢	
p. 1029	●云	叢	有 4 處
p. 1029	●無語	叢	
p. 1029	●上座到百丈	叢	
p. 1029	何須諮●	言 厶 室	
p. 1029	●擘開胸云	叢	
p. . 1030	風靜●衣輕	白 厶 旬	
p. 1042	逼戶寒●	灑	
p. 1045	再三撈●始	橈	
p. 1045	老手舊●膊	骨 厶 乞	
p. 1049	咩●吒喇	口 厶 發	
p. 1055	更加搯●	扌 厶 奎	
p. 1060	餓狗絳●	糸 厶 虎	
p. 1061	●驀然眼皮	口 厶 力	