

# 《明覺禪師語錄》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  
2010/03/24 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 4121 的「懼」字，Unicode 編碼是 39AC，由於 XP 並不支援，仍需使用構字式「忄 厶羅」。

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：

《明覺禪師語錄》使用舊版造字 76 個，字頻 654 次，這 76 個造字中，68 個可轉成 Windows XP 能顯示的字，字頻 638 次；另外 8 個字必須轉成構字式，字頻 16 次。

轉碼完成製作轉碼分析表，請參考附件一《明覺禪師語錄》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. 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 1925 的「靄」字，Unicode 字型呈現為「靄」，實際上仍為同一字。

附件一、《明覺禪師語錄》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
64	颺	2	FAA1	98B7	颺		
77	哪	1	FAAE	35BF		口 𠂇 耶	
1387	腎	1	91E4	7796	腎		
1400	窓	5	91F1	7A93	窓		
1536	麤	12	92DC	9E81	麤		
1706	焯	1	93E9	242B9		火 𠂇 勃	火 𠂇 孛，異體字問題。
1772	藜	1	946C	4527		艹 𠂇 梨	

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1854	啣	1	94E0	20E95		口𠂔留	
1925	羈	5	9568	8989	羈		Unicode 字型 呈現差異。
1978	讚	1	95BF	8B83	讚		讚，異體字問 題。
2157	酌	2	96D5	9167	酌		
2502	颺	2	98F4	98C8	颺		
3789	冲	7	8154	51B2	冲		冲，異體字問 題。
3797	凭	3	815C	51ED	凭		
3804	効	2	8163	52B9	効		
3805	勅	2	8164	52C5	勅		
3814	却	115	816D	5374	却		
3824	叙	1	8177	53D9	叙		
3864	峯	194	81C1	5CEF	峯		
3884	忽	4	81D5	6031	忽		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3909	擡	5	81EE	64E1	擡		
3931	椀	3	8245	6900	椀		
3937	檝	1	824B	6A9D	檝		
3940	藟	4	824E	6B1D	藟		
3944	氷	19	8252	6C37	氷		
3955	澗	6	825D	3D4E		汜 𠂔 間	澗，異體字問題。
3966	烟	2	8268	70DF	烟		
3979	猫	1	8275	732B	猫		
3981	珉	2	8277	73C9	珉		
3984	瑠	1	827A	7460	瑠		
3992	嘗	1	82A4	751E	嘗		
3996	疎	8	82A8	758E	疎		
4001	瘞	3	82AD	7602	瘞		
4028	稟	1	82C8	7980	稟		
4036	窰	1	82D0	7AB0	窰		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4041	豎	14	82D5	7AEA	豎		
4042	笋	2	82D6	7B0B	笋		
4045	箒	2	82D9	7B92	箒		
4052	簾	1	82E0	7C34	簾		Unicode 字型 呈現差異。
4090	脚	23	8347	811A	脚		
4121	懼	1	8366	39AC		忄 ㄇ 羅	
4134	衛	1	8373	885E	衛		衛，異體字問 題。
4146	覩	9	83A1	89A9	覩		
4174	躑	2	83BD	8EB6	躑		
4184	迹	3	83C7	8FF9	迹		
4195	鉢	14	83D2	9262	鉢		
4203	隣	12	83DA	96A3	隣		鄰，異體字問 題。
4207	霧	1	83DE	9736	霧		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4256	鷄	1	8450	9DC4	鷄		
4267	鼈	7	845B	9F08	鼈		
4278	亘	2	8466	4E98	亘		
4381	爾	54	84EF	511E	爾		
4447	捏	5	8572	63D1	捏		
4448		1	8573			 	纏，異體字問題。
4450	狃	1	8575	524F	狃		Unicode 字型呈現差異。
4739	鞞	1	875C	97DB	鞞		
4778	鞞	1	87A5	97B4	鞞		
4806	噉	1	87C1	56B1	噉		
4893	羴	9	8859	263EA			羊 𠂔 靈，異體字問題。
4948	𠂔	1	88B2	2308F			
5008	幞	2	88EE	5E5E	幞		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5111	旆	2	89B8	65BE	旆		
5289	慤	1	8ACD	6164	慤		
5290	携	11	8ACE	643A	携		
5300	鬪	9	8AD8	9B2A	鬪		
5442	枿	3	8BC9	67BF	枿		
5453	湴	10	8BD4	23D40		形: 日 工: 工	湴，異體字問題。
5479	蠲	1	8BEE	9F03		圭: 圭 電: 電	蠲，異體字問題。
5555	迥	2	8C7B	9008	迥		
5559	欸	7	8CA1	6B35	欸		
5590	朶	5	8CC0	6736	朶		
5595	朶	5	8CC5	579C	朶		
5683	攆	1	8D5E	6508	攆		攆，異體字問題。
5742	澁	1	8DBB	6F81	澁		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5753	潜	4	8DC6	6F5C	潜		
5783	教	1	8DE4	3A8D		𠄎𠄎𠄎	

## 二、手動取代表說明：

《明覺禪師語錄》檔案中出現「●」字形，共計55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《明覺禪師語錄》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、《明覺禪師語錄》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 670	摸●	扌𠄎索	有 6 處
p. 671	●斧子	鉤	有 2 處
p. 671	黎莫●速	言𠄎斯	
p. 672	●漢	𠄎𠄎 粟心 𠄎	有 2 處
p. 672	便乃性●平	𠄎𠄎 粟心 𠄎	
p. 675	●訛	譎	有 3 處
p. 676	曬●皮草	日𠄎良	

頁碼	檔案原文摘錄	原書字形	備註
p. 676	●老觀	鷓	
p. 676	一隊●酒	噶	
p. 677	●玄沙云	口𠂔爾	待造字
p. 677	打野●漢	裡	
p. 681	●跳	足𠂔字	有 3 處
p. 682	●麼	溜	有 7 處
p. 682	更是●	卜𠂔麼	
p. 684	生是●子	穉	
p. 684	生是●杖	穉	
p. 688	不是●公	叢	
p. 688	把定●老	叢	
p. 693	闕●子	椈	
p. 698	●●吹	風△弗 風△弗	
p. 700	春風●●	風△弗 風△弗	有 4 處
p. 700	酬行●長老	叢	有 2 處
p. 700	一何●速	言𠂔斯	
p. 708	●●兮步	風△弗 風△弗	
p. 710	藤●輕屨	足𠂔舜	

頁碼	檔案原文摘錄	原書字形	備註
p. 710	樓以●之	形 <sup>火</sup> 虎共	
p. 711	中列梵●	广△罩	
p. 713	斤●俊	月△吏	