

# 《黃龍慧南禪師語錄》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  
2010/03/18 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 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：

《黃龍慧南禪師語錄》使用舊版造字 30 個，字頻 101 次，這 30 個造字中，27 個可轉成 Windows XP 能顯示的字，字頻 98 次；另外 3 個字必須轉成構字式，字頻 3 次。

轉碼完成製作轉碼分析表，請參考附件一《黃龍慧南禪師語錄》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. 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 895 的「燵」字，而原書使用之字為「火 𤇀 蓬」，查詢教育部異體字字典確認「火 𤇀 蓬」並非「燵」之異體字，因此歸為錯字，以程式取代為正確字形「火 𤇀 蓬」。
  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、《黃龍慧南禪師語錄》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
895	燵	1	8ECF	71A2	燵		火 𤇀 蓬，錯字，以程式全部取代。
1205	眇	1	90CB	770E	眇		
1536	麪	2	92DC	9E81	麪		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
1706	焮	1	93E9	242B9		火△焮	火△焮字，異體字問題。
1905	焜	1	9554	88E9	焜		
2278	鎗	1	97B1	936E	鎗		
2598	鄒	1	99B7	9381	鄒		
3805	勅	1	8164	52C5	勅		
3814	却	20	816D	5374	却		
3864	峯	18	81C1	5CEF	峯		
3900	挖	1	81E5	62D5	挖		
3909	擡	1	81EE	64E1	擡		
3955	澗	2	825D	3D4E		氵△澗	澗，異體字問題。
3992	嘗	1	82A4	751E	嘗		
3996	踈	5	82A8	758E	踈		
4090	脚	19	8347	811A	脚		
4121	懼	1	8366	39AC		忄△羅	
4139	袴	1	8378	88B4	袴		
4146	覩	1	83A1	89A9	覩		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4184	迹	3	83C7	8FF9	迹		
4195	鉢	5	83D2	9262	鉢		
4251	鰈	1	844B	9C15	鰈		
4262	麪	1	8456	9EAA	麪		Unicode 字型呈現差異。
4267	鼈	1	845B	9F08	鼈		
4546	鷓	2	85F7	9D76	鷓		
4768	妬	2	8779	59AC	妬		
5008	幞	1	88EE	5E5E	幞		
5074	𠂔	1	8971	5C2B	𠂔		
5300	鬪	1	8AD8	9B2A	鬪		
5453	涅	4	8BD4	23D40		 	涅，異體字問題。

## 二、手動取代表說明：

《黃龍慧南禪師語錄》檔案中出現「●」字形，共計9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《黃龍慧南禪師語錄》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、《黃龍慧南禪師語錄》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 630	不勝●	𠂇 𠂇 麼	
p. 630	●斧子	鉤	有 3 處
p. 633	中路●泥沙	馬 𠂇 展	
p. 634	扇子●跳	𠂇 𠂇 字	
p. 636	子●跳	𠂇 𠂇 字	
p. 636	●利人難得	𠂇	
p. 638	良久云·●	𠂇	