

# 《洛陽伽藍記校注》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  
2009/02/05 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 2235 的「屨」字，Unicode 編碼是 21849，由於 XP 並不支援，仍需使用構字式「麻△女」。

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：

《洛陽伽藍記校注》使用舊版造字 92 個，字頻 227 次，這 92 個造字中，85 個可轉成 Windows XP 能顯示的字，字頻 215 次；另外 7 個字必須轉成構字式，字頻 12 次。

轉碼完成製作轉碼分析表，請參考附件一《洛陽伽藍記校注》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. 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 3978 的「猪」字，Unicode 字型呈現為「猪」，實際上仍為同一字。
3. 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 5309 的「日𠄎旬」。

附件一、《洛陽伽藍記校注》轉碼分析表

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
910	嵒	1	8EDE	5D78	嵒		
1530	紉	1	92D6	7CFA	紉		
2235	屨	2	9764	21849		麻△女	
2305	鑽	1	97CC	93BB	鑽		
2677	坵	1	9A47	24B2A		土△瓦	
3099	栳	2	9CD5	68B9	栳		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3787	冢	3	8152	51A2	冢		
3789	冲	1	8154	51B2	冲		沖，異體字問題。
3791	况	4	8156	51B5	况		況，異體字問題。
3796	决	2	815B	51B3	决		決，異體字問題。
3805	勅	2	8164	52C5	勅		
3814	却	4	816D	5374	却		
3825	叠	1	8178	53E0	叠		
3826	叶	1	8179	53F6	叶		
3840	坂	3	81A9	5742	坂		
3843	堦	4	81AC	5826	堦		
3846	堦	3	81AF	58FB	堦		
3868	卮	2	81C5	5DF5	卮		
3883	徧	1	81D4	5FA7	徧		
3891	慙	2	81DC	6159	慙		
3906	摧	1	81EB	6409	摧		
3929	黎	2	8243	68C3	黎		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
3944	冰	1	8252	6C37	冰		
3974	牀	5	8270	7240	牀		
3975	牕	3	8271	7255	牕		
3976	犁	1	8272	7282	犁		
3977	狗	1	8273	72E5	狗		
3978	猪	3	8274	732A	猪		Unicode 字型 呈現差異。
3990	甄	7	82A2	750E	甄		
3996	踈	7	82A8	758E	踈		
4016	着	2	82BC	7740	着		
4017	蟬	1	82BD	880F	蟬		
4039	疇	1	82D3	7ADA	疇		
4055	粃	1	82E3	7C83	粃		
4067	綉	1	82EF	7D89	綉		
4080	羣	7	82FC	7FA3	羣		
4081	翱	1	82FD	7FFA	翱		
4090	脚	3	8347	811A	脚		
4091	臯	3	8348	81EF	臯		
4094	館	1	834B	8218	館		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4107	葱	9	8358	8471	葱		
4118	藥	1	8363	8602	藥		
4132		19	8371			血会爪	眾，異體字問題。
4134	衛	6	8373	885E	衛		衛，異體字問題。
4139	袴	1	8378	88B4	袴		
4142	裊	1	837B	8943	裊		
4146	覩	4	83A1	89A9	覩		睹，異體字問題。
4176	輒	2	83BF	8F19	輒		
4184	迹	1	83C7	8FF9	迹		
4186	遡	1	83C9	9061	遡		
4190	邳	1	83CD	90C4	邳		
4195	鉢	4	83D2	9262	鉢		
4203	隣	3	83DA	96A3	隣		鄰，異體字問題。
4220	穎	1	83EB	9834	穎		
4226	飡	1	83F1	98E1	飡		

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4242	鬪	3	8442	9B2D	鬪		
4252	鰓	1	844C			魚厶眾	鰓，異體字問題。
4256	鷄	2	8450	9DC4	鷄		雞，異體字問題。
4267	鼈	2	845B	9F08	鼈		
4275	义	12	8463	4E49	义		
4278	亘	4	8466	4E98	亘		
4460	勅	4	85A1	52D1	勅		
4496	段	2	85C5	53DA	段		
4515	呪	7	85D8	546A	呪		
4518	响	1	85DB	54CD	响		
4519	咲	1	85DC	54B2	咲		
4574	勛	1	8654	228E2		思厶勳	
4723	壇	1	874C	58C3	壇		
4735	壠	1	8758	58E0	壠		
4738	睺	2	875B	777A	睺		
4813		1	87C8			形厂 墨	灑，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
4826	塚	2	87D5		塚		
4844		1	87E7			𠄎𠄎凡	𠄎，異體字問題。
4858	閱	1	87F5	49A7		門△兒	
4981	鳧	1	88D3	9CEC	鳧		
5188	捫	1	8A46	62EA	捫		
5309		1	8AE1	3AEC		日△旬	待造字。
5339	𦉳	1	8B40	59E7	𦉳		
5453	涅	5	8BD4	23D40		形日 工	
5478	飴	1	8BED	4D77	飴		
5542	鑠	4	8C6E	93C1	鑠		
5551	賈	2	8C77	8CF7	賈		
5558	龕	2	8C7E	9E84	龕		
5559	欸	1	8CA1	6B35	欸		
5583	殮	1	8CB9	98F1	殮		
5588	恠	1	8CBE	6061	恠		
5591	杞	1	8CC1	233CC		木△巳	杞，異體字問題。

編號	造字	頻次	Big5	Unicode	WinXP	構字式	備註
5706	敘	2	8D75	654D	敘		敘，異體字問題。
5749	牖	2	8DC2	7257	牖		
5752	摠	1	8DC5	63D4	摠		
5769	郛	2	8DD6	90CD	郛		
5799	罰	1	8DF4	262FA		𠄎 𠄎 剗	

## 二、手動取代表說明：

《洛陽伽藍記校注》檔案中需手動修改的情況分為以下幾種：

(一) 《洛陽伽藍記校注》檔案中出現「●」字形，共計16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《洛陽伽藍記校注》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、《洛陽伽藍記校注》「●」字手動取代表

頁碼	檔案原文摘錄	原書字形	備註
p. 100	●(鶴)生於異縣	需 𠄎 鳥	
p. 132	●磴舂簞	石 𠄎 輦	



頁碼	檔案原文摘錄	原書字形	備註
p. 133	文宗學府，●班	𠂔𠂔𠂔	
p. 160	久閻阿那●來朝	舟𠂔𠂔	待造字
p. 160	可班那●蕃王	舟𠂔𠂔	待造字
p. 237	光、●、榮	日𠂔𠂔	
p. 288	●噠國	口𠂔𠂔	有 3 處
p. 288	北盡勅●	勒𠂔𠂔	待造字
p. 300	遮嫚●之處	女𠂔𠂔	
p. 317	為●噠所滅	口𠂔𠂔	
p. 318	與●(𠂔)賓爭境	𠂔	
p. 318	●噠、烏場王	口𠂔𠂔	
p. 341	云此浮圖●(陷)	𠂔𠂔𠂔	
p. 350	●(栖)禪寺	𠂔𠂔𠂔	待造字

(二)《洛陽伽藍記校注》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，請參考附件三。欄位說明如下：

1. 檔案原文摘錄：摘錄檔案《洛陽伽藍記校注》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。

附件三、《洛陽伽藍記校注》缺字外錯字修改

檔案原文摘錄	手動取代結果	備註
<u>今</u> 玉及水精珍異	<u>金</u> 玉及水精珍異	