

# 語言所古漢語轉碼缺字造字記錄

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  
2009/11/19 丁致伶

**需求單位：**中央研究院語言學研究所

**連絡人：**盧秋蓉

**缺字需求說明：**「中央研究院古漢語語料庫」建構始於一九九〇年，是應漢語史研究需求而建構的語料庫。這個語料庫又依是否經過斷詞及加標詞類而分成兩類，即未加標的素語料庫以及有標注的標記語料庫。目前素語料庫所蒐集的語料已涵蓋上古漢語（先秦至西漢）、中古漢語（東漢魏晉南北朝）、近代漢語（唐五代以後）大部分的重要語料，約 403 本書；在標記語料庫方面，已完成部份上古漢語及近代漢語的語料標注工作，約 40 本書，總計 7 千多萬字，將視結果逐步提供線上檢索。

但該語料庫的中文字碼採用 Big5（收錄 13,051 字）編碼，並使用瀚典全文檢索系統造字檔（簡稱為舊版造字）以解決缺字問題，處理缺字的方式是在 Big5 的使用者造字區（FA40~C8F1）增補新字，由於 Big5 造字區只能容納 6217 字，不敷使用，且在電腦中文編碼環境整個轉換為 Unicode 之下，舊版造字於線上檢索不易，古漢語語料庫決定將原為 Big5 編碼的舊版造字轉為 Unicode 編碼，若為 Unicode 缺字則採用構字式來表達。其中需轉碼的古漢語 word 文本，漢字構形資料庫未收的缺字，將進行造字；共計需造細明體 114 字、標楷體 114 字，詳見造字清單。

## 造字清單：

### 細明體及標楷體造字清單：

細明體造字 114 字，標楷體造字 114 字，共計 228 字。

字型檔：細明體外字集九、標楷體外字集九、細明體外字集十、標楷體外字集十。

編號	細明體造字	標楷體造字	楷書編號	構字式
1.	棧	棧	70248	木△我
2.	飼	飼	70249	食△周
3.	纈	纈	70250	絞△頁
4.	遊	遊	70251	舟△遊
5.	慟	慟	70252	忄△勃
6.	塚	塚	70253	土△宗
7.	瘧	瘧	70254	疒△帝
8.	噏	噏	78578	口△爺
9.	珪	珪	78579	子△圭
10.	擲	擲	78580	扌△囂
11.	毳	毳	78581	毛△奈
12.	曜	曜	78582	形元日隹
13.	稔	稔	78583	禾△忽

編號	細明體造字	標楷體造字	楷書編號	構字式
14.	剗	剗	78584	專△刂
15.	忠	忠	78585	出△心
16.	侷	侷	78586	亻△龙
17.	嚙	嚙	78587	口△褒
18.	笨	笨	78588	竹△不
19.	囡	囡	78590	口△天
20.	灑	灑	78591	灬△厭
21.	鬣	鬣	78592	毛△亂
22.	醜	醜	78593	酉△芻
23.	襲	襲	78594	薛△衣
24.	砭	砭	78595	石△兵
25.	攢	攢	78596	扌△巖
26.	嗝	嗝	78597	形口戶乞□
27.	痢	痢	78598	疒△別
28.	藟	藟	78599	艹△榴
29.	擣	擣	78600	扌△粵
30.	巒	巒	78601	山△瞿
31.	縉	縉	78602	糸△當

編號	細明體造字	標楷體造字	楷書編號	構字式
32.	噉	噉	78603	形口广卅从口
33.	羈	羈	78604	𠂔 厶 勒
34.	葛	葛	78605	艹 厶 离
35.	熿	熿	78606	火 厶 當
36.	蝻	蝻	78607	虫 厶 甬
37.	汨	汨	78608	氵 厶 亠
38.	搨	搨	78609	扌 厶 鼠
39.	顛	顛	78610	鼻 厶 頁
40.	嚙	嚙	78611	口 厶 摩
41.	聃	聃	78612	耳 厶 丂
42.	擥	擥	78613	扌 厶 警
43.	飜	飜	78614	扁 厶 飛
44.	榬	榬	78615	舟 厶 堂
45.	𠵼	𠵼	78616	形之口衣
46.	嚙	嚙	78617	口 厶 麼
47.	揜	揜	78618	形扌矣水
48.	筵	筵	78619	形竹之幸
49.	吽	吽	78620	口 厶 牛

編號	細明體造字	標楷體造字	楷書編號	構字式
50.	譚	譚	78621	言△摹
51.	撈	撈	78622	扌△劣
52.	塵	塵	78623	形广秝土□
53.	簣	簣	78624	从△宀△艹△衣
54.	蹊	蹊	78625	足△奕
55.	揀	揀	78626	形扌亡艹□
56.	鯁	鯁	78627	魚△韋
57.	鞮	鞮	78628	革△曼
58.	獬	獬	78629	犴△角△羊
59.	榷	榷	78630	形木乚幸□
60.	徃	徃	78631	彳△在
61.	嚙	嚙	78632	口△鶯
62.	逕	逕	78633	辵△架
63.	勑	勑	78634	耒△力
64.	狹	狹	78635	彳△夾
65.	𩚑	𩚑	78636	厭△肉
66.	𦑔	𦑔	78637	月△察
67.	魁	魁	78638	鬼△王

編號	細明體造字	標楷體造字	楷書編號	構字式
68.	禘	禘	78639	禘 $\triangle$ 亞
69.	窳	窳	78641	穴 $\triangle$ 侵
70.	鞣	鞣	78642	革 $\triangle$ 扁
71.	胎	胎	78643	月 $\triangle$ 多
72.	煥	煥	78644	光 $\triangle$ 奐
73.	劬	劬	78645	勹 $\triangle$ 乃
74.	粢	粢	78646	夂 $\triangle$ 米
75.	盍	盍	78647	彳 $\triangle$ 蓋
76.	瑀	瑀	78648	王 $\triangle$ 禺
77.	儻	儻	78649	亻 $\triangle$ 量
78.	燹	燹	78650	〇來 $\triangle$ 人
79.	織	織	78651	纟 $\triangle$ 戠
80.	璫	璫	78652	王 $\triangle$ 衮
81.	睦	睦	78653	日 $\triangle$ 壺
82.	眇	眇	78654	日 $\triangle$ 少
83.	胎	胎	78655	月 $\triangle$ 右
84.	捩	捩	78656	扌 $\triangle$ 徙
85.	顛	顛	78657	長 $\triangle$ 頁

編號	細明體造字	標楷體造字	楷書編號	構字式
86.	蝨	蝨	78658	凡△虫
87.	糶	糶	78659	米△離
88.	僉	僉	78660	亻△奐
89.	秃	秃	78661	禾△丸
90.	葩	葩	78662	艹△白△匕△
91.	刃	刃	78663	石△刀
92.	霰	霰	78664	雨△浸
93.	凌	凌	78665	辶△麥
94.	忭	忭	78666	忄△成
95.	蒸	蒸	78667	𠃉△羔
96.	頽	頽	78668	无△頁
97.	蒙	蒙	78669	艹△豪
98.	輶	輶	78670	車△票
99.	𠂔	𠂔	78671	申△丁
100.	爨	爨	78672	火△爨
101.	羝	羝	78673	羊△丘
102.	驥	驥	78674	馬△棗
103.	膏	膏	78675	灬△膏

編號	細明體造字	標楷體造字	楷書編號	構字式
104.	榎	榎	78676	形木名攵
105.	蝨	蝨	78677	虫厶厨
106.	甄	甄	78678	雲厶允
107.	苳	苳	78679	艹厶阡
108.	箴	箴	78680	竹厶戠
109.	巒	巒	78681	山厶麗
110.	萌	萌	78682	艹厶朋
111.	寘	寘	78683	穴厶具
112.	爨	爨	78684	火厶農
113.	唸	唸	78685	口厶企
114.	睨	睨	78709	目厶思