

漢字構形資料庫的研發與應用

撰寫人：莊德明、鄧賢瑛
中研院資訊所文獻處理實驗室
日期：2009年7月

目錄

第一章 漢字構形資料庫的研發概要	3
1.1 字形結構分析與部件檢字	3
1.2 構字式與缺字問題	4
1.3 異體字表	5
1.4 銜接古今文字	7
1.5 增收甲骨文、金文及楚系簡帛文字	8
1.6 古漢字重文與風格碼	12
1.7 漢字構形資料庫的版本沿革	16
1.8 漢字構形資料庫的架構及特色	18
第二章 漢字構形資料庫的部件拆分	21
2.1 名詞釋義	21
2.1.1 部件	21
2.1.2 成字部件和非（成）字部件	21
2.1.3 基礎部件和合成部件	21
2.1.4 漢字結構	21
2.1.5 結構理據	21
2.1.6 部件拆分	22
2.1.7 有理據拆分和無理據拆分	22
2.1.8 部件的層級	22
2.1.9 部件在組構中的功能	23
2.2 基礎部件的規範	24
2.2.1 小篆的獨體字與楷書的基礎部件	24
2.2.2 現行兩個中文字的基礎部件標準	25
2.2.3 漢字構形資料庫的基礎部件	29
2.3 部件拆分的一些問題	30
2.3.1 部件拆分的類別與統計	31
2.3.2 非字合成部件	31
2.2.3 無理據拆分	32
2.4 五大字集部件表	32
2.4.1 五大字集基礎部件表	33
2.4.2 五大字集基礎部件組字頻率表	40
2.4.3 五大字集合成部件表	48
2.5 簡化字集部件表	72
2.5.1 簡化字基礎部件表	72
2.5.2 簡化字合成部件表	74
第三章 漢字構形資料庫的構形編碼	79

3.1 部件的組合及識別	79
3.2 構字式的制定及使用	81
3.3 構字式的處理技巧	88
3.3.1 構字式的類型	88
3.3.2 構字式的擷取	88
3.3.3 構字式在資料庫中的表達	89
3.3.4 構字式的正規化	90
3.4 風格碼的制定及使用	92
第四章 漢字構形資料庫的應用	95
4.1 系統下載及安裝	95
4.2 系統架構	95
4.3 漢字構形資料庫	96
4.4 缺字字型	97
4.5 使用界面	99
4.6 缺字增益集	103
4.7 網頁缺字	106
第五章 漢字構形資料庫的展望	110
5.1 漢字構形資料庫的應用現況	110
5.2 建立文字學入口網站	111
參考資料	113

第一章 漢字構形資料庫的研發概要

中央研究院資訊科學研究所文獻處理實驗室所研發的漢字構形資料庫，是一個記錄漢字形體知識的資料庫；這些知識包含古今文字的源流演變、字形結構及異體字表。自 1998 年 8 月推出第一個正式版本研發至今，已有超過 10 年的時間，其間為考量使用者的需求及漢字的特性，經歷了數次重大的改版更新，至 2009 年 6 月為止，版本已更新至 2.53 版，收錄古今文字 119,195 個及異體字 12,208 組。底下將就漢字構形資料庫自 1.0 至 2.5 版之間，所推出各個不同版本之各項特色，作一概括敘述：

1.1 字形結構分析與部件檢字

1998 年 8 月推出漢字構形資料庫 1.0 版，支援個人電腦作業系統 Windows 95 及 98。這是漢字構形資料庫最早的正式版本，收錄五大碼 13,051 個字形（以下簡稱五大字集），並且對五大字集進行字形結構分析。因義構形是漢字的特點，因此當對漢字進行構形分析時，可將字形依層次拆分為各級部件。以「醫」字為例，它的字形結構可如圖 1.1 所示，其中「毇」和「酉」是它的一級部件，而一級部件「毇」又可以再往下拆分為二級部件與三級部件。字形結構的一級部件可以體現全字的造字意圖，例如「醫」的一級部件「毇」表示呻吟聲，「酉」表示用以治病的酒。

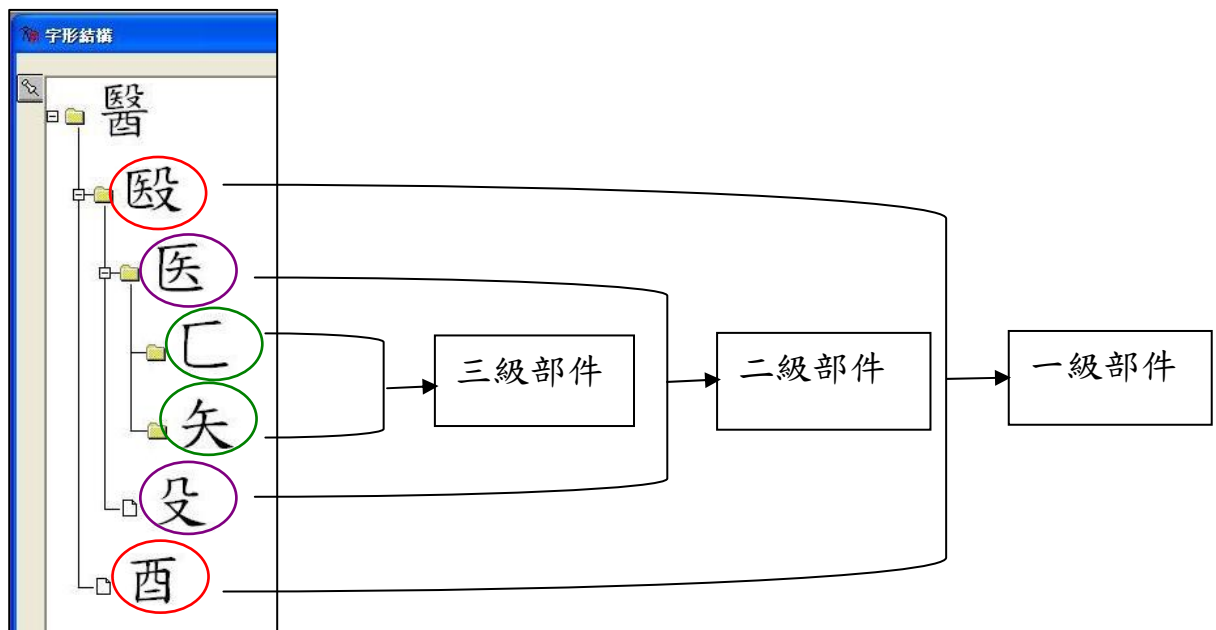


圖 1.1：「醫」字的字形結構

漢字構形資料庫的檢字技術是採用部件檢字的方式，一個字形中

的各級部件都可用來檢索字形，部件檢字的好處是不需知道每個字的部首，只要能辨認出若干部件，就可以用來檢索字形。用來檢索字形的部件可以是單一部件，也可以是一個以上的部件。假設使用者想檢索「醫」的字形，使用部件「矢」在五大字集進行部件檢字時，可以得到 82 個字形（如圖 1.2）；使用「矢」和「殳」進行檢字時，得到的字形數為 11 個（如圖 1.3）；當使用「矢」和「酉」進行檢索時，得到的字形數為 1 個（如圖 1.4）。由此例可知，使用者可以增加檢索的部件數，或使用不同的部件進行檢索，以提高檢字的精確度。

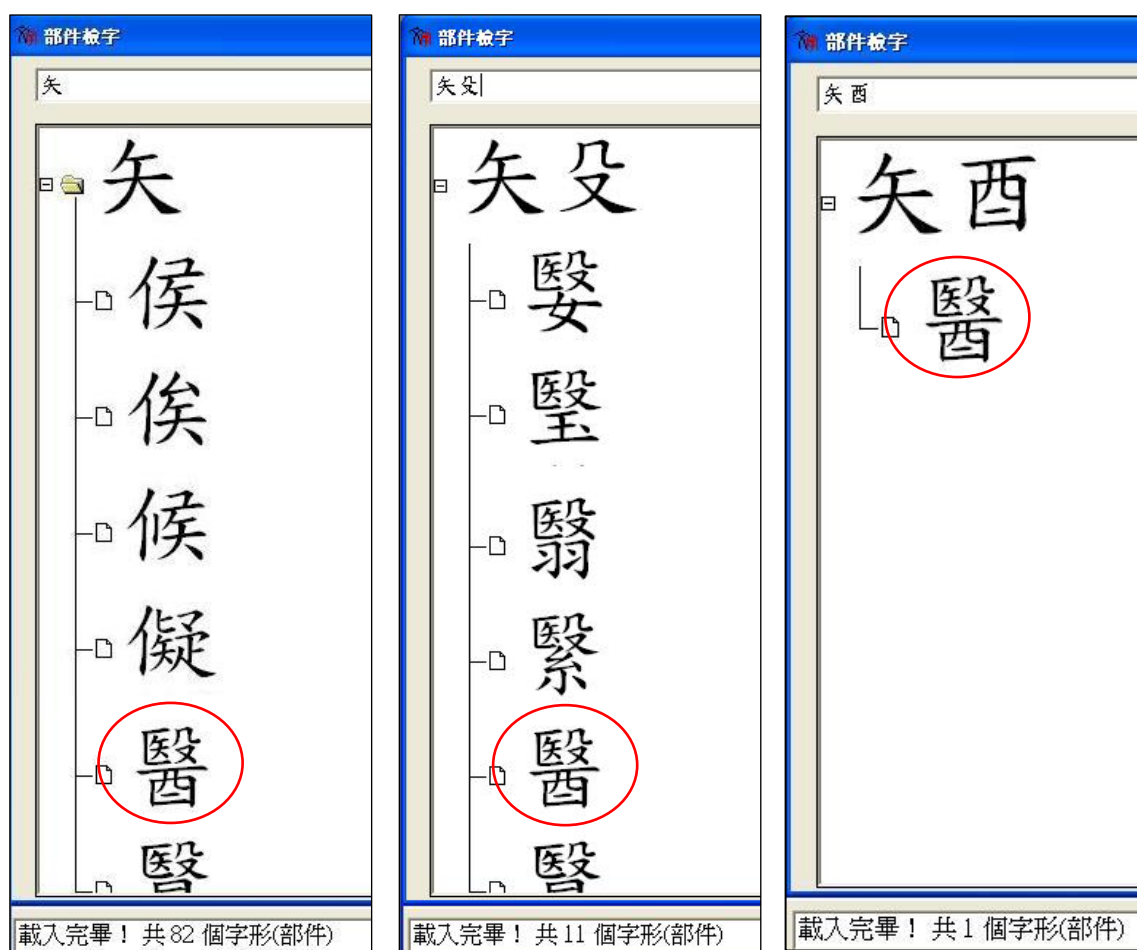


圖 1.2：以「矢」進行
部件檢字的結果

圖 1.3：以「矢」和「殳」
進行部件檢字的結果

圖 1.4：以「矢」和「酉」
進行部件檢字的結果

1.2 構字式與缺字問題

1999 年 1 月 20 日推出 1.1 版，此版本最大特色為擴充漢字構形資料庫 1.0 版的字形。除了原已收錄的五大字集以外，並採用葉健欣先生開發的漢字庫字型，收錄《漢語大字典》單字字形，使漢字構形資料庫所收錄的字數擴充至將近 5 萬字。在這 5 萬字當中，有許多是

電腦的缺字，為了在計算機中制式表達這些缺字，文獻處理實驗室將漢字部件的組合方式，簡化為橫連、直連與包含三種組合方式，分別以△、△、△三種連接符號表示。使用連接符號連接部件的字形結構表達方式稱作構字式，可利用有限的部件表示無窮的漢字。例如，「毆」的構字式可寫作「医△殳」；「醫」的構字式可寫作「毆△酉」；「医」的構字式可寫作「匚△矢」。

為了讓使用者在 Microsoft Word 文件中，能夠使用漢字構形資料庫所收錄的近 5 萬個字形，在 2000 年 10 月 18 日推出的 1.2 版中，加入了 Word 缺字預覽巨集（如圖 1.5）。使用者可以在 Word 文件中先輸入缺字字形的構字式，再透過 Word 缺字預覽，顯示文件中的缺字字形（如圖 1.6）。

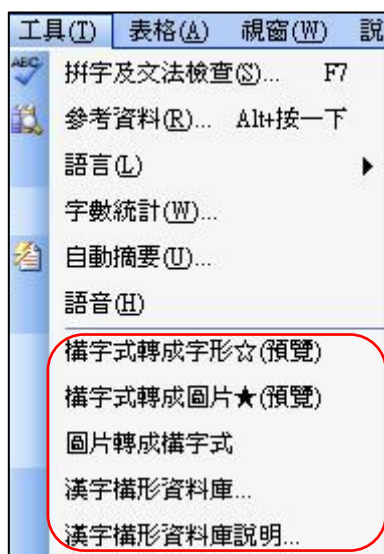
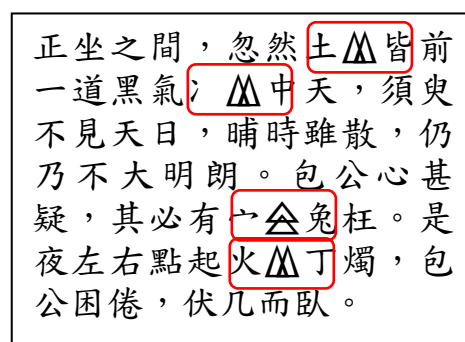


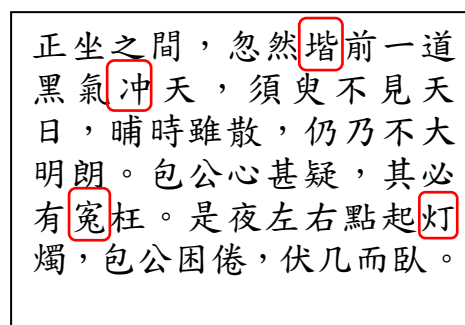
圖 1.5：Word 缺字預覽巨集



原始文件



缺字預覽



缺字預覽文件

圖 1.6：Word 缺字預覽結果

1.3 異體字表

中文缺字中有大部分是屬於異體字，這和漢字「一字多形」的特點息息相關。在考量漢字「一字多形」的特性之下，2001 年 1 月 19 日推出的漢字構形資料庫 1.3 版中，增添了異體字表的功能，收錄了

《漢語大字典》異體字表 12,208 組。

《漢語大字典》異體字表的編排是採用主體字統領異體字的方式，將同一主體字統領的簡化字、古今字、全同異體字（指音義全同而形體不同的字）、非全同異體字（指音義部分相同的異體字），集中在該主體字下編為一組。《漢語大字典》異體字表收錄異體字 12,208 組，涵蓋 36,309 個字形。

在漢字中，主體字與異體字的關係有時候是相對的，某個字可能本身是主體字，同時又是其他字形的異體字。例如當「射」字作主體字時，它的異體字有 7 個；但「射」字同時又是另一個主體字「榭」字的異體字。當使用者檢索「射」字時，漢字構形資料庫可同時顯示所有與「射」字有關的異體字表如圖 1.7 所示，這也是漢字構形資料庫的異體字表較書面字典更為便利的優點。

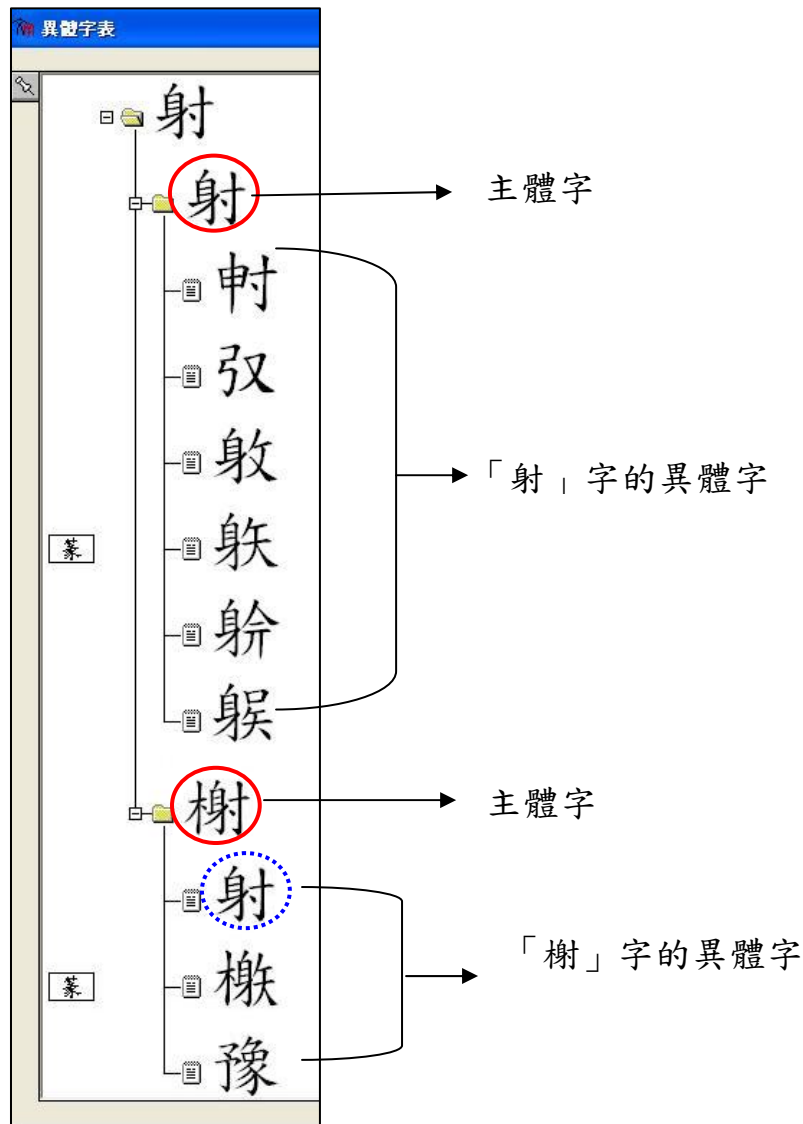


圖 1.7：「射」字的異體字表

1.4 銜接古今文字

漢字構形資料庫 2.0 版在 2002 年 7 月 2 日發佈，開始收錄古漢字，並著手收錄《說文解字詁林》中 9,831 個小篆字形。至 2003 年 3 月 17 日推出的 2.1 版時，又收錄《說文解字詁林》中 1,169 個小篆重文，這些重文包含了或體 538 個，古文 478 個，籀文 210 個，篆文 33 個，俗字 10 個以及奇字 3 個。2.1 版已完整收錄《說文解字詁林》中的小篆及重文字形一共 11,100 個。

小篆可以說是研究古文字與今文字的過渡橋樑，《說文》中的小篆不但保存古漢字演變的線索，也是現今漢字尋求字源的重要參考依據。在漢字構形資料庫中，檢索小篆字形同樣也是利用部件檢字的方式，並且依小篆的字形進行字形結構分析。假設以檢索「醫」字為例，可以先利用「醫」的部件「矢」字檢索而得到「醫」字（如圖 1.8），再點選「醫」字後，即可看見它的字形結構（如圖 1.9）：

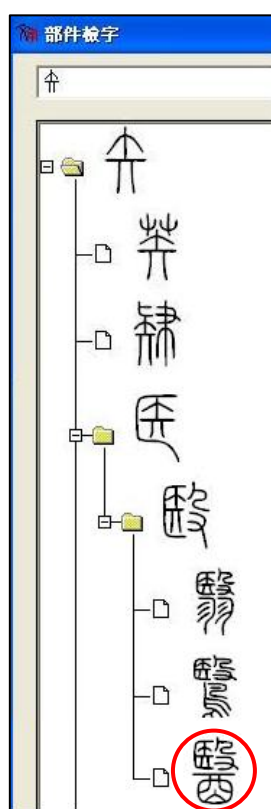


圖 1.8：以「矢」進行
小篆部件檢字的結果

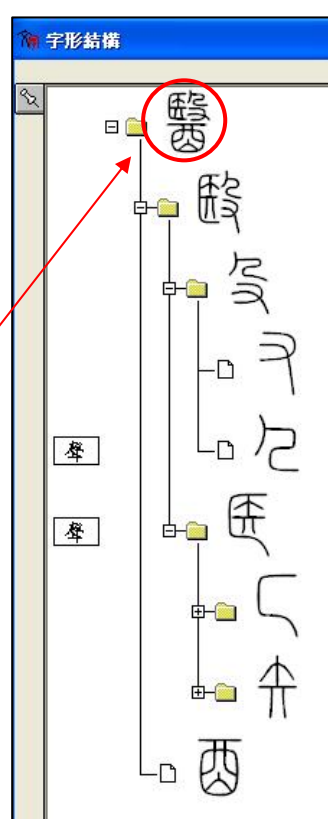


圖 1.9：「醫」字的小
篆字形結構

此外，漢字構形資料庫中的小篆字形亦根據《說文》重文而制訂小篆的異體字表，並且新增字形演變功能。以「射」字為例，現今楷書的「射」字在《說文》小篆中是「𠄎」字的或體，在漢字構形資料

庫的小篆異體字表中，即將「射」字列在「𠄎」字底（如圖 1.10），並在字形演變視窗中顯示小篆「射」字為說文或體（如圖 1.11）。



圖 1.10：「射」字的小篆異體字表



圖 1.11：「射」字的小篆字形

1.5 增收甲骨文、金文及楚系簡帛文字

在 2004 年 12 月 6 日推出的漢字構形資料庫 2.2 版，開始收錄《金文編》的金文字形；此版本新增 3,459 個金文，楷書字形也增加至 59,220 個。2005 年 8 月 3 日推出的 2.3 版，則開始收錄《楚系簡帛文字編》的楚系簡帛文字；此版本新增 2,627 個楚系簡帛文字，楷書字形也增加至 60,082 個。2006 年 8 月 2 日推出的 2.4 版，開始收錄《殷墟甲骨刻辭類纂》的甲骨文字形；此版本新增 1,970 個甲骨文，楷書字形則增加至 62,242 個。

漢字構形資料庫增收上述古文字之後，已可大致囊括古漢字字形演變的發展軌跡，例如「射」字從甲骨文到小篆的字形演變，如圖 1.12 所示：

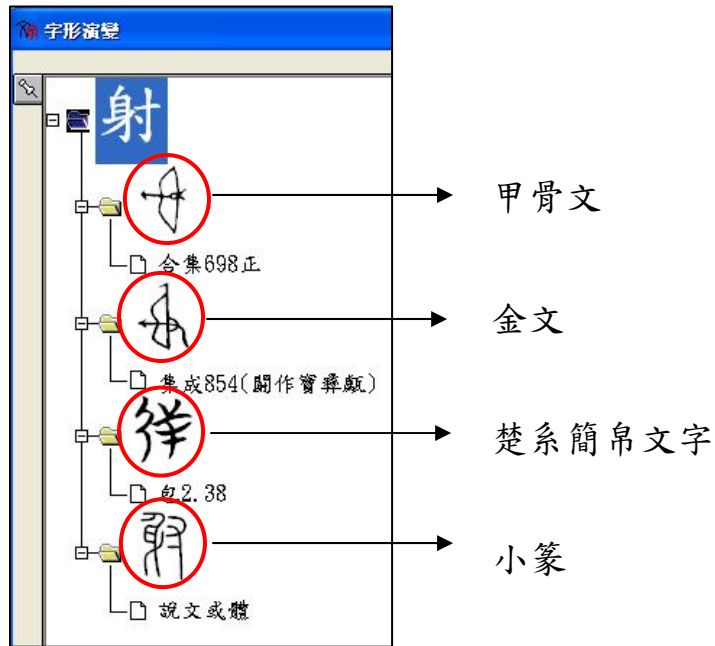


圖 1.12：「射」字的
字形演變

因為字形演變的關係，現今楷書的同一個字形，在古漢字當中可能會有不同的形體結構，從圖 1.12 可以看出「射」字原本在甲骨文中是象張弓形，在金文中又在右邊加上手執弓矢的形狀，小篆左邊從「身」即是由弓的形狀變化而來。

正因為漢字有字形演變的特點，雖然現今楷書看來是同一字，但在古漢字卻可能有不同的字形結構，所以在對古漢字進行部件檢字時，也需要依照不同的字形結構，選擇適合的部件。例如「射」字在甲骨文、金文、楚系簡帛文字及小篆中的字形結構各有不同，組成部件也有所差異，甲骨文的「射」字是由「弓」與「矢」組合而成，見圖 1.13；金文的「射」字是由「弓」、「矢」及「又」（「又」在古漢字中象手形）組合而成，見圖 1.14；楚系簡帛文字的「射」字亦是由「弓」與「矢」組合而成，見圖 1.15；而在圖 1.16 中，小篆的「射」字則是由「身」與「寸」組合而成，小篆字形左邊的「身」是由原本的弓形演變而來。

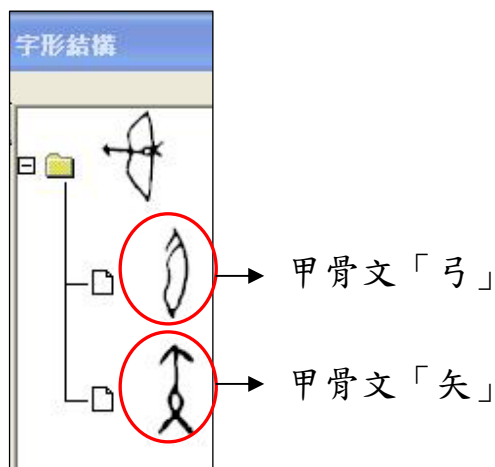


圖 1.13: 甲骨文「射」字的字形結構

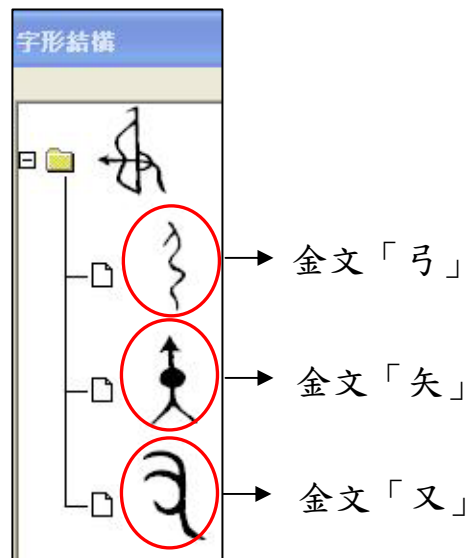


圖 1.14: 金文「射」字的字形結構

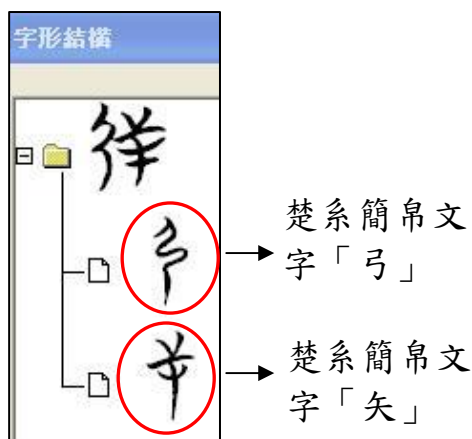


圖 1.15: 楚系簡帛文字「射」字的字形結構

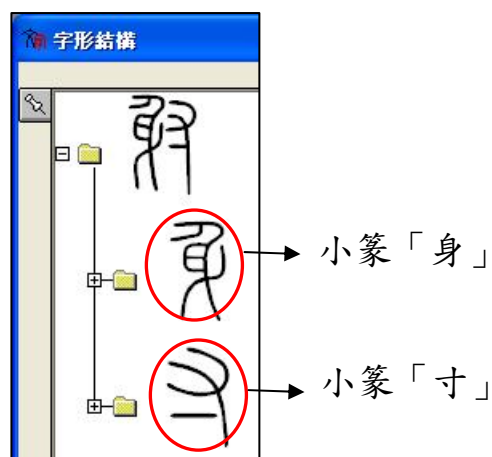


圖 1.16: 小篆「射」字的字形結構

由圖 1.13 至圖 1.16 可以得知，隨著「射」字的字形演變，也會影響它的部件組合，如果要在甲骨文、金文和楚系簡帛文字中檢索「射」字，需要以部件「弓」或「矢」進行檢索（如圖 1.17 至 1.19）；但在小篆中就要改成用部件「身」或「寸」進行檢索（如圖 1.20）。

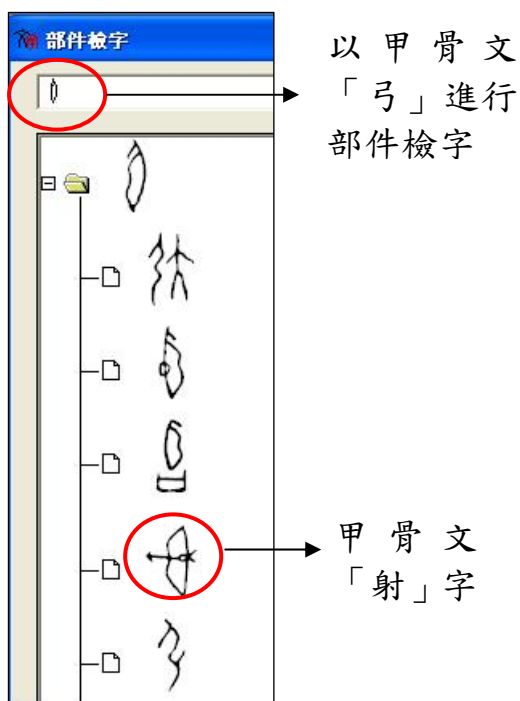


圖 1.17：以「弓」進行甲骨文部件檢字

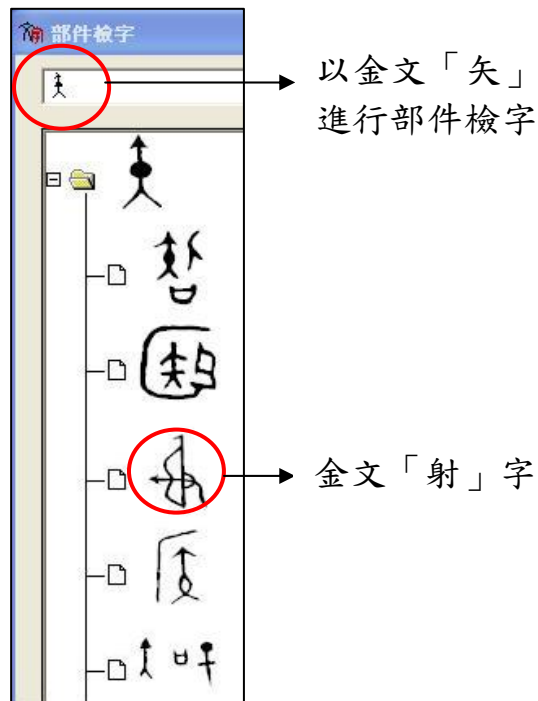


圖 1.18：以「矢」進行金文部件檢字

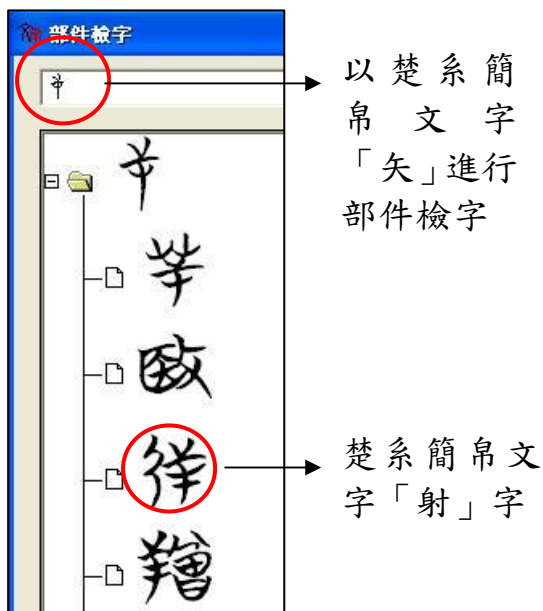


圖 1.19：以「矢」進行楚系簡帛文字部件檢字

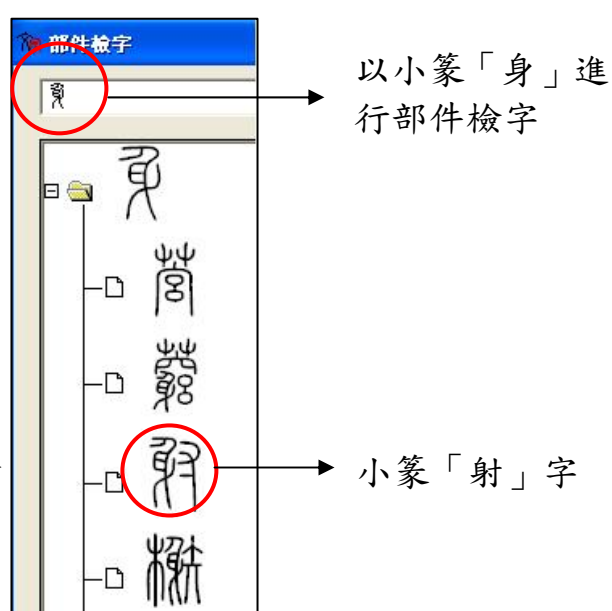


圖 1.20：以「身」進行小篆部件檢字

1.6 古漢字重文與風格碼

在 2007 年 8 月 9 日推出的漢字構形資料庫 2.5 版中，最重要的特色有三點：

1.大量增收古漢字，並且開始利用風格碼進行古漢字的編碼工作

過去在解決古漢字的顯示問題時，通常是先製作古漢字字型，再透過字型切換的方式顯示，例如在 Word 文件中，將標楷體的字型切換成「北師大說文小篆」時，結果如圖 1.21 所示：

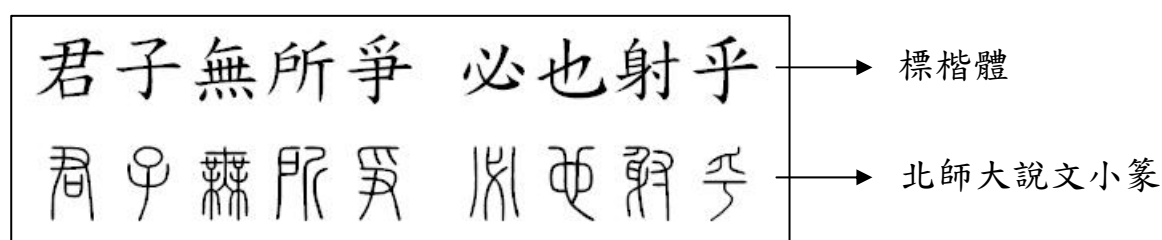


圖 1.21：利用切換字型的方式，顯示小篆字形

然而利用切換字型顯示古漢字的方式只適用於小篆，卻不適用於其他的古漢字。因為小篆已是有固定規範的文字，所以除了少部分的字形以外，大部分的小篆字形都可以與現代楷書的字形進行一對一對應。然而在更早之前的古漢字，例如甲骨文、金文或楚系簡帛文字，因為當時還沒有文字共同的規範，所以存在大量形體不固定的異體字，亦即古漢字中存在許多音義相同，但形體不同的重文。如果要運用切換字型的方式顯示這些具有大量異體字的古漢字時，就會面臨難以抉擇的困擾。例如「射」字在金文中的異體字有「𠄎」、「𠄎」、「𠄎」、「𠄎」、「𠄎」、「𠄎」……等多種寫法，在製作字型時，究竟要取哪一個字形代表金文的「射」字，答案可以說是莫衷一是。

有鑑於此，於是漢字構形資料庫 2.5 版開始採用古漢字本身的源流資訊進行編碼，與楷書運用構字式的編碼方式有所區分。例如「射」字在金文當中的異體字表如圖 1.22，但將字形前面的「+」符號展開後，即可看到同一個字形可能有許多異寫字如圖 1.23：

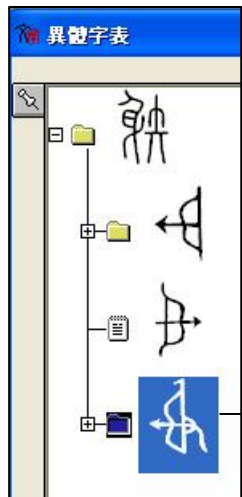
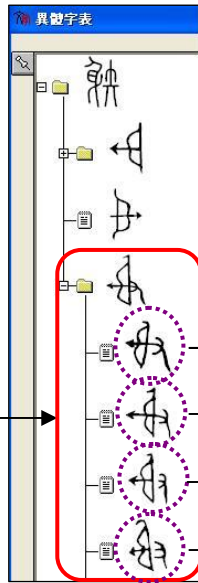


圖 1.22：金文「射」字的異體字表

點選



風格碼：

形射體集成 2803

形射體集成 4273

形射體集成 9455

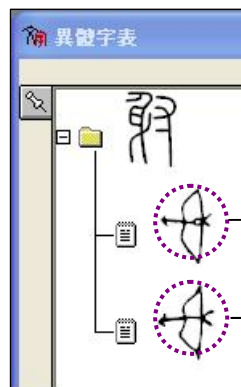
形射體集成 2784

圖 1.23：金文「射」字的異寫字

以圖 1.23 中風格碼「形射體集成 2803」為例，它的含義代表了這個字的字形出處為令鼎，著錄在《殷周金文集成》中的器號為 2803。

採用風格碼為古漢字進行編碼，可大量增收古漢字的異體字，大量擴建古漢字異體字表。在 2.5 版中一共增收了《金文編》中的金文 20,069 個、《楚系簡帛文字編》中的楚系簡帛文字 16,801 個，以及《殷墟甲骨刻辭類纂》的甲骨文 2,197 個，並且回過頭來重新為小篆字形制定風格碼。

除了前一例的金文之外，甲骨文、楚系簡帛文字、小篆字形的風格碼不同之處即在於各字形的字形出處不同，以「射」字為例，甲骨文、楚系簡帛文字及小篆的異體字表及風格碼出處說明如圖 1.24 至圖 1.26：

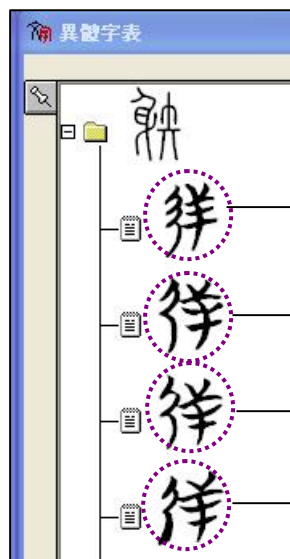


字形出處為《甲骨文合集》編號 698 正

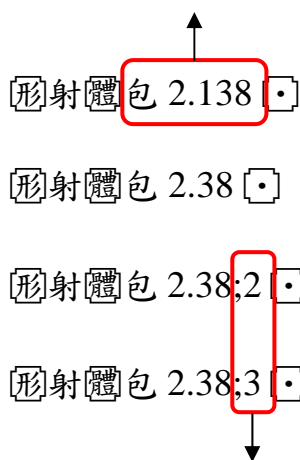
形射體合集 698 正

形射體合集 5784

圖 1.24：甲骨文「射」字的異體字表及風格碼



字形出處為《楚系簡帛文字編》，出土墓號為荊門包山 2 號墓，簡號為 138



表示「射」這個字形，在荊門包山 2 號墓，簡號 38 當中重覆的個數。

圖 1.25：楚系簡帛文字「射」字的異體字表及風格碼



表示此字為小篆「射」字的或體

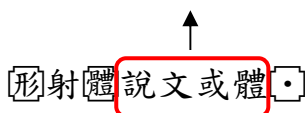


圖 1.26：小篆「射」字的異體字表及風格碼

2. 新增出處檢字功能

漢字構形資料庫 2.5 版在檢字方面，還針對古漢字的部分新增出處檢字功能，使用者能在漢字構形資料庫出處檢字的下拉選單中，選取欲檢索的字書，再利用甲骨文編號、金文器號、或楚系簡帛文字的簡號等相關資訊，檢索同一出處古漢字的字形。例如，使用者若想檢索毛公鼎（《殷周金文集成》器號 2841）上的所有字形，即可利用毛公鼎的器號在出處檢字上進行檢索，檢索畫面如圖 1.27 所示：

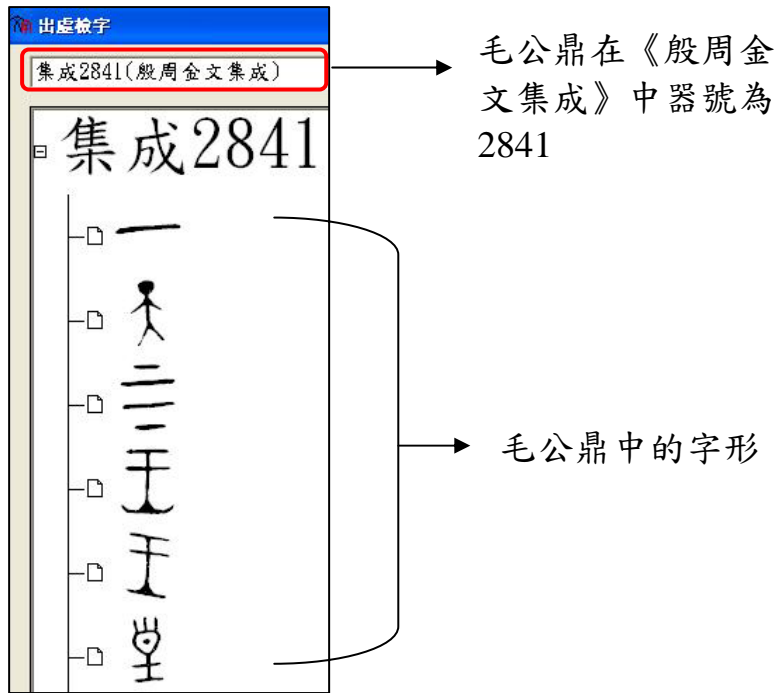


圖 1.27：利用出處檢字檢索毛公鼎上的字形

3. 新增了自動貼圖至 Microsoft Word 的功能

原本漢字構形資料庫是使用 Word 缺字預覽巨集，在 Word 文件中輸入缺字字形的構字式，再透過 Word 缺字預覽，顯示文件中的缺字字形。但這個做法會令原始文件的內容不易閱讀，增加編輯和管理文件的困難。因此在 2.5 版中，漢字構形資料庫新增了自動貼圖至 Word 的功能。使用者在處理 Word 文件中的缺字時，可先利用漢字構形資料庫檢索所需的字形，再透過快速鍵 Ctrl-W 便可立即在 Word 裡貼入缺字圖片，得到所見即所得的效果，更便利與快速地解決使用者在 Word 文件裡的缺字問題（如圖 1.28）：



圖 1.28：在 Microsoft Word 中貼入字形圖片

1.7 漢字構形資料庫的版本沿革

漢字構形資料庫的研發過程，其版本沿革可詳見表 1.1：

表 1.1：漢字構形資料庫版本沿革

版本	年/月/日	內容簡介	字數	備註
1.0	1998/8/6	收錄五大碼字形 13,051 個。	13,051	支援 Windows 95 及 98
1.1	1999/1/20	收錄五大碼字形及《漢語大字典》字頭近五萬個。	近 50,000	缺字採用葉健欣開發的漢字庫字型。
1.2	2000/10/18	收錄五大碼字形及《漢語大字典》字頭近五萬個。	近 50,000	加入 Word 缺字預覽巨集
1.3	2001/1/19	1. 收錄五大碼字形及《漢語大字典》字頭近五萬個。 2. 收錄《漢語大字典》異體字表 12,208 組。	近 50,000	《漢語大字典》異體字表共 12,208 組，包含 36,309 個字形。
1.4	2002/05/17	1. 收錄五大碼字形及《漢語大字典》字頭，合計 54,711 個字形 2. 收錄《漢語大字典》異體字表 12,208 組。	54,711	開始支援 Windows 2000
2.0	2002/07/02	1. 收錄五大碼字形及《漢語大字典》字頭，合計 54,711 個字形。 2. 收錄《漢語大字典》異體字表 12,208 組。 3. 收錄《說文解字詁林》小篆 9,831 個。	64,542	1. 開始收錄小篆字形。 2. 銜接古今漢字，並開始以建立文字學資料庫為目標。
2.1	2003/03/17	1. 收錄五大碼字形及《漢語大字典》字頭，合計 54,711 個字形。 2. 收錄《漢語大字典》異體字表 12,208 組。 3. 收錄《說文解字詁林》的小篆及重文共 11,100 個。	65,811	1. 完整收錄《說文解字詁林》的字形。 2. 完成小篆構形資料庫。 3. 缺字字型依標楷體及細明體重造。
2.2	2004/12/06	1. 收錄楷書字形 59,220 個。 2. 收錄《漢語大字典》異體字表 12,208 組。 3. 收錄《說文解字詁林》的小篆及重文共 11,100 個。 4. 收錄《金文編》中的 3,459 個金文。	73,779	1. 增收的楷書字形為金文的楷定字。 2. 開始收錄金文。 3. 本次收錄的金文以異構字為主。
2.3	2005/08/03	1. 收錄楷書字形 60,082 個。 2. 收錄《漢語大字典》異體	77,268	1. 增收的楷書字形為楚系文字的楷定字。

版本	年/月/日	內容簡介	字數	備註
		<p>字表 12,208 組。</p> <p>3. 收錄《說文解字詁林》的小篆及重文共 11,100 個。</p> <p>4. 收錄《金文編》中的 3,459 個金文。</p> <p>5. 收錄《楚系簡帛文字編》中的 2,627 個楚系簡帛文字。</p>		<p>2. 開始收錄楚系文字。</p> <p>3. 本次收錄的楚系文字以異構字為主。</p>
2.4	2006/8/2	<p>1. 收錄楷書字形 62,242 個。</p> <p>2. 收錄《漢語大字典》異體字表 12,208 組。</p> <p>3. 收錄《說文解字詁林》的小篆及重文共 11,100 個。</p> <p>4. 收錄《金文編》中的 3,781 個金文。</p> <p>5. 收錄《楚系簡帛文字編》中的 2,627 個楚系簡帛文字。</p> <p>6. 收錄《殷墟甲骨刻辭類纂》中的 1,970 個甲骨文。</p>	81,720	<p>1. 增收的楷書字形為甲骨文的楷定字。</p> <p>2. 增收的金文為圖形文字。</p> <p>3. 開始收錄甲骨文。</p> <p>4. 本次收錄的甲骨文以異構字為主。</p>
2.5	2007/8/9	<p>1. 收錄楷書字形 62,366 個。</p> <p>2. 收錄《漢語大字典》異體字表 12,208 組。</p> <p>3. 收錄《說文解字詁林》的小篆及重文共 11,100 個。</p> <p>4. 收錄《金文編》中的 20,069 個金文。</p> <p>5. 收錄《楚系簡帛文字編》中的 16,801 個楚系簡帛文字。</p> <p>6. 收錄《殷墟甲骨刻辭類纂》中的 2,197 個甲骨文。</p>	112,533	<p>1. 大量增收金文及楚系簡帛文字的重文，合計 30,462 個。</p> <p>2. 開始利用風格碼進行古漢字編碼。</p> <p>3. 除部件檢字外，另增出處檢字。</p> <p>4. 新增 Microsoft Word 自動貼圖功能。</p>
2.51	2007/12/24	<p>1. 收錄楷書字形 62,671 個。</p> <p>2. 收錄《漢語大字典》異體字表 12,208 組。</p> <p>3. 收錄《說文解字詁林》的小篆及重文共 11,100 個。</p> <p>4. 收錄《金文編》中的 20,091 個金文。</p> <p>5. 收錄《楚系簡帛文字編》中的 19,138 個楚系簡帛文字。</p> <p>6. 收錄《殷墟甲骨刻辭類</p>	115,197	<p>1. 增收楷書字形 305 個。</p> <p>2. 增收楚系簡帛文字 2,337 個。</p> <p>3. 分層顯示楚系簡帛文字異體字表中的重文。</p> <p>4. 調整說文重文的風格碼。</p>

版本	年/月/日	內容簡介	字數	備註
		纂》中的 2,197 個甲骨文。		
2.52	2008/05/23	<ol style="list-style-type: none"> 1. 收錄楷書字形 62,942 個。 2. 收錄《漢語大字典》異體字表 12,208 組。 3. 收錄《說文解字詁林》的小篆及重文共 11,100 個。 4. 收錄《金文編》中的 21,413 個金文。 5. 收錄《楚系簡帛文字編》中的 19,138 個楚系簡帛文字。 6. 收錄《殷墟甲骨刻辭類纂》中的 2,197 個甲骨文。 	116,790	<ol style="list-style-type: none"> 1. 增收楷書字形 271 個。 2. 增收金文 1,322 個。
2.53	2009/6/22	<ol style="list-style-type: none"> 1. 收錄楷書字形 64,844 個。 2. 收錄《漢語大字典》異體字表 12,208 組。 3. 收錄《說文解字詁林》的小篆及重文共 11,100 個。 4. 收錄《金文編》中的 21,413 個金文。 5. 收錄《楚系簡帛文字編》中的 19,138 個楚系簡帛文字。 6. 收錄《殷墟甲骨刻辭類纂》中的 2,700 個甲骨文。 	119,195	<ol style="list-style-type: none"> 1. 增收楷書字形 1,902 個。 2. 增收甲骨文 503 個。

1.8 漢字構形資料庫的架構及特色

漢字構形資料庫早期收錄的字形是以楷書的現代印刷字體為主，其後陸續增加小篆、金文、楚系簡帛文字及甲骨文。因此，現今的漢字構形資料庫是由甲骨文、金文、楚系簡帛文字、小篆及楷書構形資料庫組合而成，其架構圖如圖 1.29 所示：

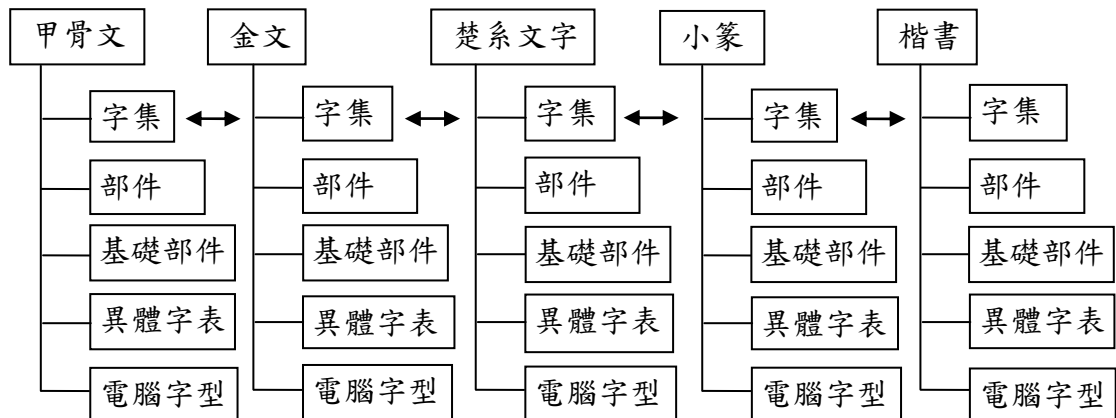


圖 1.29：漢字構形資料庫的組成

表 1.2 說明漢字構形資料庫的現況，其中列出各個構形資料庫的主要參考字書、已收字數、已分析字形的部件數及基礎部件數、異體字表組數、製作的電腦字型、提供的字書索引、合作單位等。中研院史語所負責甲骨文、金文、楚系簡帛文字的構形分析，而北京師範大學則提供了小篆字型：

表 1.2：漢字構形資料庫的現況

	甲骨文	金文	楚系文字	小篆	楷書
主要參考字書	殷墟甲骨刻辭類纂	金文編	楚系簡帛文字編	說文解字詁林	漢語大字典
字數	2,197	21,413	19,138	11,100	62,942
部件數	296	804	704	2,004	5,224
基礎部件數	228	469	464	367	982
異體字表組數	1,762	2,614	2,206	1,081	12,208
電腦字型	中研院甲骨文	中研院金文	中研院楚系簡帛文字	北師大說文小篆	標楷體及細明體外字集
字書索引	殷墟甲骨刻辭類纂、甲骨文字詁林、甲骨文字集釋	金文編、金文詁林、殷周金文集成引得	楚系簡帛文字編	說文解字詁林	漢語大字典、中文大辭典
其他屬性		殷周金文集成器號、器名	出土墓號及簡號		Unicode、Big5
合作單位	中研院史語所	中研院史語所	中研院史語所	北京師範大學、台灣師範大學	中研院史語所

綜觀漢字構形資料庫的研發過程，其主要的特色為以下四點：

- 1.銜接古今文字以反映字形源流演變。
- 2.收錄不同歷史時期的異體字表，以表達不同漢字在各個歷史層面的使用關係。
- 3.記錄不同歷史時期的文字結構，以呈現漢字因義構形的特點。
- 4.使用構字式及風格碼來解決古今漢字的編碼問題。

第二章 漢字構形資料庫的部件拆分

部件拆分是漢字構形資料庫的核心。本章先解釋部件拆分的相關名詞，然後提出基礎部件的規範，說明部件拆分所碰到的一些問題，最後列出五大字集及簡化字集的部件表。

2.1 名詞釋義

2.1.1 部件

由筆畫組成具有組配漢字功能之構字單位，當一個形體被用來構造其他的字，成為所構字的一部分時，是為所構字的部件。例如「毆」、「酉」、「医」、「殳」、「匚」、「矢」、「㇇」、「大」、「几」、「又」皆為「醫」的部件。

2.1.2 成字部件和非（成）字部件

可以獨立成字的部件稱成字部件，當它不作為其他字的部件時，本身就是一個完整的字；不能獨立成字的部件稱非字部件，非字部件不能獨立存在，必需依附於其他部件之上。例如「醫」字的部件「毆」、「酉」、「医」、「殳」、「匚」、「矢」、「大」、「几」、「又」皆可獨立成字，是為成字部件；而「㇇」無法獨立成字，是為非字部件。

2.1.3 基礎部件和合成部件

最小的不再拆分的部件稱基礎部件，由兩個以上的基礎部件組成的部件稱合成部件。例如「醫」字的基礎部件有「匚」、「㇇」、「大」、「几」、「又」、「酉」，這些部件都無法再繼續往下拆分；而「醫」字的合成部件有：「毆」、「医」、「殳」、「矢」，這些部件都是由兩個以上的基礎部件所組成，可以繼續往下拆分至基礎部件為止。

2.1.4 漢字結構

部件構成漢字的方式和規則。

2.1.5 結構理據

根據字源或參考字源，從漢字的部件組合分析出的造字意圖，稱結構理據。例如「醫」字，根據《說文》：「毆，病聲。酉，所以治病

也。」可見「醫」字的造字意圖是由「毆」與「酉」而來。根據字源解釋「醫」字造字意圖，即為「醫」字的結構理據。

2.1.6 部件拆分

將漢字拆分為部件稱部件拆分。

2.1.7 有理據拆分和無理據拆分

根據結構理據所進行部件拆分，稱有理據拆分；當無法分析理據或理據與字形發生矛盾時，依照字形所進行的部件拆分，稱無理據拆分。例如「醫」字，根據《說文》：「毆，病聲。酉，所以治病也。」因此「醫」字可以拆分為「毆」與「酉」；而「毆」根據《說文》：「擊中聲也，从殳医聲。」因此「毆」可以拆分為「殳」與「医」；又「医」根據《說文》：「盛弓弩矢器也，从匚从矢。」因此「医」可以拆分為「匚」、「矢」；但「矢」往下拆分為「𠂆」、「大」卻沒有任何字源解釋，因此在「醫」的部件當中，除「矢」的拆法為無理據拆分，其餘皆為有理據拆分。

2.1.8 部件的層級

在依層次拆分的漢字中，部件是有層級的。以「醫」字為例，「醫」字含有以下四個層級的部件：

一級部件：毆、酉

二級部件：医、殳

三級部件：匚、矢、几、又

四級部件：𠂆、大

一級部件又稱為直接部件，因為一級部件直接構成全字，體現全字的造字意圖。例如「醫」字的一級部件是「毆」與「酉」，根據《說文》：「毆，病聲。酉，所以治病也。」可見「毆」與「酉」不但直接組成「醫」字全字，而且也體現了「醫」字的造字意圖。「醫」字的各級部件如圖 2.1 所示：

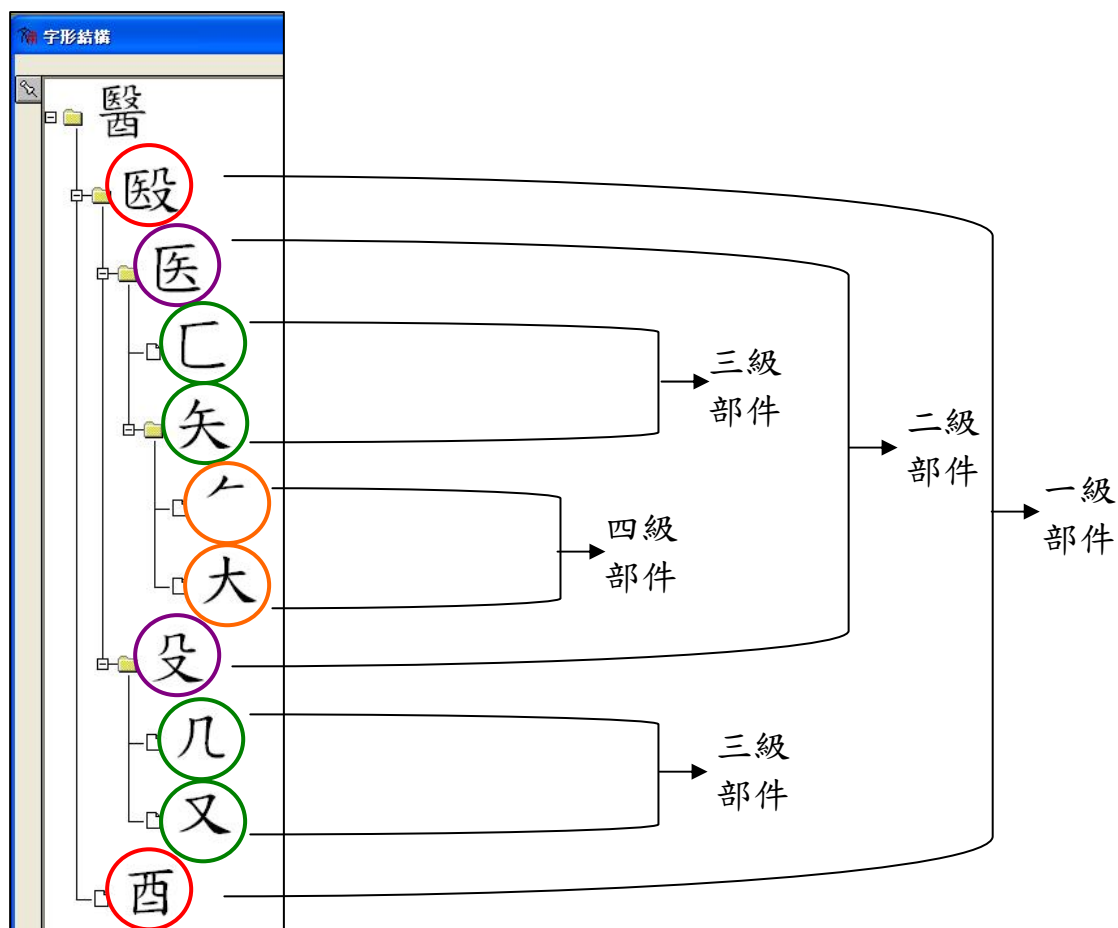


圖 2.1：「醫」字的各級部件

2.1.9 部件在組構中的功能

部件在構字時都體現一定的構意，部件所承擔的構意類別，稱為這個部件的功能。漢字部件的功能有以下五類：表形功能、表義功能、示音功能、標示功能以及替代功能。

表形功能：部件用與物象相似的形體來體現構意，例如「番」字的部件「田」表現爪印的形狀，即是承擔部件的表形功能。

表義功能：部件以它在獨用時所記錄的詞義來體現構意，即是部件的表義功能。例如「醫」字之部件「毘」表示呻吟聲；「酉」表示用來治病的酒，即是承擔部件的表義功能。

示音功能：部件在構字時體現示音的功能，部件的讀音與全字的讀音相近或相同，例如「毘」之部件「医」讀音與「毘」相同，即是承擔了示音功能。

標示功能：部件不獨立存在，而是附加在另一個部件上，作為區別和指示作用，即是具有標示功能。例如「刃」字之部件「丿」，指示出刀刃的位置，具有指示位置的作用；而「太」字之部件「丿」，

是用以和「大」字作出區別，具有區別的作用。這些部件都承擔了標示的功能。

替代功能：有些部件本身不表示構意，而是作為另一部件的替代物，稱為部件的替代功能。例如「訖」字之部件「乞」原本為「气」，後來卻寫成「乞」替代了原本的「气」；「奠」字之部件「大」原為「丌」，後來卻寫成「大」替代了原本的「丌」。此類部件本身並不表示構意，而是替代了其他部件，擔負了部件的替代功能。

2.2 基礎部件的規範

基礎部件的規範為部件拆分的首要問題。在圖 2.1「醫」字的部件拆分中，為何「矢」不是基礎部件，而要拆分成「宀」、「大」？「殳」的基礎部件「几」，由於標準字體的不同，有時也和「几」混用。以下先比較兩份基礎部件的規範，再說明漢字構形資料庫的基礎部件制定原則。

2.2.1 小篆的獨體字與楷書的基礎部件

現行中文楷書的拆分原則，可參考兩份中文字基礎部件標準：「中文字基礎部件及部件屬性」（編號 CNS 11643-2，以下簡稱 CNS 11643-2）以及 GB 13000.1 字符集「漢字部件規範」（編號 GF3001-1997，以下簡稱 GF3001）。GF3001 於 1997 年發佈，是對 GB13000.1「信息技術 通用多八位編碼字符集（UCS）第一部分：體繫結構與基本多文種平面」中的 20,902 個中文字進行拆分後得出的基礎部件表及使用原則。CNS 11643-2 則是於 2007 年發佈，是對 CNS11643「中文標準交換碼」第 1 及第 2 字面的 13,051 個中文字進行拆分後得出的基礎部件表及使用原則。這兩份基礎部件是在對現行楷書進行拆分時，重要的參考依據。

然而這兩份基礎部件標準之適用範圍在於現代楷書，並不適用於小篆文字，因為小篆字形與楷書字形在基礎部件的認定上有時並不一致。有些字在小篆中是無法拆分的基礎部件，但在楷書中卻可往下拆分為若干部件。例如「易」字在小篆中寫作「𪚩」，根據《說文》：「易，蜥易蝮蜓守宮也，象形。」可見「𪚩」整體為蜥蜴的形狀。但在楷書中的「易」字，卻可拆分成「日」與「勿」等部件；又如「癸」字在小篆中寫作「𪚩」，根據《說文》：「癸，冬時水土平可揆度也，象水從四方流入地中之形。」可見「𪚩」整體為象水從四方流入地中的樣子，但在楷書中的「癸」字可拆分成「火」與「天」等部件；再如「象」字在小篆中寫作「𪚩」，根據《說文》：「象，長鼻牙南越大獸……象耳

牙四足之形。」可見「象」整體為大象之形體，但在楷書中的「象」字可拆分成「夕」、「四」、「豕」等部件。

此外，還有些字在小篆中可以拆分成若干部件，但在楷書中卻當作基礎部件而不再往下拆分的，例如「更」字在小篆中寫作「𠄎」，根據《說文》：「更，改也，从支丙聲。」可拆分為「𠄎」和「丙」等部件，但在楷書中的「更」字是作基礎部件而不再拆分。又如「東」字在小篆中寫作「東」，根據《說文》：「官溥說：从日在木中。」可拆分為「日」與「木」等部件，但在楷書中的「東」字是當作基礎部件而不再拆分。再如「束」字在小篆中寫作「束」，根據《說文》：「縛也，从口木。」可拆分為「口」與「木」等部件，但在楷書中的「束」字則是當作基礎部件而不再拆分。

由此可知，無論是 CNS 11643-2 或是 GF3001，這兩份楷書的基礎部件均與小篆的基礎部件並不完全相符，有些字的小篆字形是無法拆分的基礎部件，但它們的楷書字形卻可往下拆分(例如「易」、「癸」、「象」等字)；而有些字則是小篆字形可往下拆分，但在楷書字形中卻是當作基礎部件，而不再拆分的(例如「更」、「東」、「束」等字)。

2.2.2 現行兩個中文字的基礎部件標準

CNS 11643-2 共有 517 個基礎部件，而 GF3001 有 560 個基礎部件，它們的基礎部件有一部分是重疊的，這些相同的部件共有 391 個，詳見如下：

一、丨、ノ、丶、丶、丶、乙、乚、丁、丁、ㄣ、ㄟ、ㄥ、一、フ、レ、レ、
二、匕、匕、匚、匚、十、厂、弓、丁、七、七、丁、与、ナ、冂、
冂、冂、卜、卜、冂、人、イ、儿、入、八、几、勺、レ、义、レ、
力、川、厂、レ、レ、マ、凡、匕、レ、レ、レ、レ、レ、レ、刀、
リ、力、尸、巳、尸、巳、ム、又、又、乃、九、了、リ、コ、ユ、
乚、レ、土、士、大、寸、尢、工、干、井、弋、廿、丈、于、才、
冂、丰、丰、牛、レ、レ、口、口、山、巾、少、夕、夕、夕、尸、
彡、彡、彡、久、又、七、千、レ、广、レ、女、子、子、小、レ、
中、中、レ、川、儿、己、己、么、又、弓、互、丑、丑、冂、巳、
巳、也、レ、五、彡、子、己、丸、戈、木、木、歹、歹、牙、犬、
彡、王、彡、弓、不、井、卅、夫、屯、廿、市、卅、无、丰、赤、
五、卅、日、日、止、レ、中、内、日、月、冂、冂、冂、冂、手、
彡、彡、彡、斤、月、彡、毛、氏、气、爪、レ、レ、父、片、牛、
生、小、衣、衣、丰、丹、及、心、小、レ、文、火、彡、レ、之、
レ、充、尸、母、母、母、水、レ、氷、月、丑、尹、巴、吊、央、

尹、尸、尹、瓦、甘、石、示、禘、世、本、未、未、夫、艸、市、
 戊、戊、田、皿、目、内、𠂔、电、电、且、冉、冊、凹、凸、史、
 央、由、甲、申、兕、呂、冊、业、瓜、白、禾、弟、手、丘、乍、
 乎、广、立、必、永、夂、月、皮、矛、弗、民、夨、𠂔、聿、业、
 而、耳、臣、西、西、西、吏、夷、束、其、卂、𠂔、𠂔、𠂔、肉、
 月、卢、虫、曲、曳、且、𠂔、𠂔、竹、竹、缶、耒、自、白、白、
 舟、年、豕、帝、缶、必、肩、米、羊、羊、羊、衣、禘、州、糸、
 糸、聿、聿、艮、艮、豕、車、酉、更、束、求、甫、臣、見、貝、
 里、串、尚、豸、身、采、我、鳥、𠂔、𠂔、言、𠂔、長、長、雨、
 事、東、走、豕、車、𠂔、亞、門、果、金、非、秉、史、承、革、
 禹、食、食、垂、禹、重、為、飛、鬥、鬼、奠、畢、亞、黑、黽、
 熏

除了相同的 391 個基礎部件外，這兩份基礎部件標準還有一部分的部件是相異的，CNS 11643-2 獨有的基礎部件有 126 個，而 GF3001 獨有的基礎部件有 169 個，它們之間的相異點可分為底下五點加以說明：

1. 是否採取認同使用

當部件因為在字中所處的部位不同而產生了筆畫變形或比例變化，例如「土」當作字形偏旁時，經常寫作提土旁「扌」，若將「扌」與「土」視為相同的部件，是為認同使用，如 GF3001 即是採取認同使用；反之，如果將這兩者視為不同的部件，即為不認同使用，如 CNS 11643-2 中將這類變化後的部件稱為附部件，視為與主部件不同的部件，是採取不認同使用。CNS 11643-2 因為不認同使用而獨有的部件共有 65 個，詳見如下，括弧內為主部件或 GF3001 採取認同使用的部件，提供對照參考：

丿 (丿)、乚 (GF3001 視「乚」與「乚」為同一個部件)、几 (几)、
 匕 (匕)、人 (人)、儿 (儿)、入 (入)、八 (八)、几 (几)、乂 (乂)、
 又 (又)、九 (九)、土 (土)、大 (大)、工 (工)、久 (久)、乚 (乚)、
 女 (女)、小 (小)、巳 (巳)、木 (木)、木 (木)、犬 (犬)、王 (王)、
 夫 (夫)、屯 (屯)、𠂔 (𠂔)、止 (止)、牛 (牛)、文 (文)、火 (火)、
 宀 (宀)、水 (水)、本 (本)、申 (申)、且 (且)、瓜 (瓜)、生 (生)、
 禾 (禾)、禾 (禾)、丘 (丘)、立 (立)、皮 (皮)、业 (业)、耳 (耳)、
 至 (至)、束 (束)、耒 (耒)、舟 (舟)、衣 (衣)、更 (更)、束 (束)、
 求 (求)、見 (見)、里 (里)、采 (采)、雨 (雨)、金 (金)、非 (非)、
 垂 (垂)、重 (重)、熏 (熏)、丂 (丂)、与 (与)、夨 (夨)

2. 拆分上的不同

有部分字在 CNS 11643-2 中是基礎部件，但在 GF3001 中是合成部件。例如「妻」字在 CNS 11643-2 未拆，是屬於基礎部件，但 GF3001 中拆分為「𠂔」與「女」；又如「婁」字在 CNS 11643-2 未拆，是屬於基礎部件，但在 GF3001 中拆分為「𠂔」與「女」。

另外，有部分字在 GF3001 中是基礎部件，但在 CNS 11643-2 中是合成部件，例如「象」在 GF3001 未拆，是屬於基礎部件，但在 CNS 11643-2 中卻拆分為「ㄗ」、「ㄨ」、「豕」；又如「尺」在 GF3001 未拆，是屬於基礎部件，但在 CNS 11643-2 中卻拆分為「尸」、「丿」。

拆分上的不同，會影響檢索結果的呈現。例如當使用者需要檢索「妻」字時，在 CNS 11643-2 中，因為「妻」即是基礎部件，因此只能以「妻」字為部件進行檢索。但在 GF3001 中，因為「妻」字可拆分為「𠂔」與「女」二個基礎部件，所以使用「𠂔」與「女」作為部件進行檢索，都可找到「妻」字。又如使用者需檢索「象」字時，在 GF3001 中，因為「象」即是基礎部件，因此只能以「象」為部件進行檢索。但在 CNS 11643-2 中，因為「象」可拆分為「ㄗ」、「ㄨ」、「豕」三個部件，所以使用「ㄗ」、「ㄨ」、「豕」這三個部件進行檢索，都可找到「象」字。大體而言，將字形拆分得較細者，在檢索上可以運用的檢索部件選擇性也越多。

CNS 11643-2 因為拆分之不同，而獨有的部件共有 45 個，其中與 GF3001 有相關性的有 34 個。這 34 個有相關性的部件中，「𠂔」是因為 CNS 11643-2 拆分了「𠂔」而來，GF3001 則未拆「𠂔」，因而 CNS 11643-2 獨有「𠂔」，GF3001 獨有「𠂔」；另外 33 個部件則是 CNS 11643-2 未拆，但 GF3001 將之拆分的基礎部件，將留待下一段說明。

CNS 11643-2 因拆分不同而獨有的 45 個基礎部件，詳見如下，若該部件與 GF3001 之基礎部件有關，則將 GF3001 之相關部件以括弧標註其後，提供對照參考：

一、巾、毛、生、用、頁、麗、韭、首、齒、巛、勿(ㄨ)、甩(ㄨ)、至(ㄨ)、以(ㄨ)、互(ㄨ)、丙(內)、亥(夕)、羽(刁)、重(重)、董(重)、妻(𠂔)、典(曲)、雀(圭)、隹(圭)、無(𠂔)、卑(卑)、隶(𠂔)、面(月)、馬(馬)、烏(烏)、兼(兼)、鬲(𠂔)、帶(𠂔)、曹(曲)、𠂔(𠂔)、棄(棄)、婁(𠂔)、庸(庸)、壺(亞)、單(單)、鼎(鼎)、齊(月、𠂔)、龍(龍)、義(義)

GF3001 因為拆分之不同，而獨有的部件共有 68 個，其中與 CNS 11643-2 有相關性的有 33 個。這 33 個有相關性的部件中，「𠂔」的關

係已如前述，其他 32 個之關係為 CNS 11643-2 未拆，但 GF3001 將之拆分後所獨有之基礎部件，例如：「乇」為 GF3001 拆分了「甩」而來，但 CNS 11643-2 將「甩」當作基礎部件而未加以拆分；「ㄣ」為 GF3001 拆分了「勿」而來，但 CNS 11643-2 將「勿」當作基礎部件而未加以拆分，GF3001 此類的基礎部件一共有 32 個。

GF3001 因拆分不同而獨有的 68 個基礎部件，詳見如下，若該部件與 CNS 11643-2 之基礎部件有關，則將 CNS 11643-2 之相關部件以括弧標註其後，提供對照參考：

ㄥ、玉、四、丿、兀、习、丁、方、聿、亡、囧、上、東、万、自、兆、而、巨、下、予、勿、象、斥、三、月、氷、冫、丫、尺、八、乂、冫、矢、冫、ㄣ（勿）、乇（甩）、厶（至）、以（以）、夕（互）、丙（丙）、夕（亥）、习（羽）、重（重、董）、聿（聿）、典（典）、圭（佳、雀）、夨（雀）、册（無）、卑（卑）、𠂇（面）、馬（馬）、烏（烏）、兼（兼）、𠂇（鬲）、𠂇（帶）、曲（曹）、𠂇（曲）、業（棄）、聿（聿）、肅（庸）、亞（壺）、卑（單）、鼎（鼎）、月、𠂇（齊）、𠂇（龍）、𠂇（義）

3. 字形規範的差異

有部分基礎部件之差異是屬於字形規範上的差異，例如：在 CNS 11643-2 的「丰」在 GF3001 中寫作「丰」；在 CNS 11643-2 的「友」在 GF3001 中寫作「友」，這些都是屬於字形規範的差異。CNS 11643-2 因為上述原因而獨有的部件共有 15 個，詳見如下，並將採用 CNS 11643-2 與 GF3001 基礎部件標準之字形範例以括弧標註其後，提供對照參考：

匕（化化）、丰（丰丰）、友（拔拔）、夕（炙炙）、龜（龜龜）、儿（壺壺）、去（充充）、儿（恐恐）、丰（害害）、艾（艾艾）、丰（邦邦）、丰（慧慧）、𠂇（牽牽）、册（册册）、白（叟叟）

GF3001 因字形規範的差異而獨有的部件共有 12 個，詳見如下，並將採用 GF3001 與 CNS 11643-2 基礎部件標準之字形範例以括弧標註其後，提供對照參考：

月（前前）、止（延延）、𠂇（鰓鰓）、于（刊刊）、𠂇（慧慧）、戶（肩肩）、申（叟叟）、𠂇（聚聚）、𠂇（檝檝）、册（册册）、𠂇（敢敢）、𠂇（華華）

4. 字集大小不同

CNS 11643-2 主要是以CNS11643「中文標準交換碼」第1及第2字面作為字集，此字集包含大五碼 13,051 個中文字，只有繁體字，而無簡化字；而GF3001 則是以GB13000.1「信息技術 通用多八位編碼字符集（UCS）第一部分：體繫結構與基本多文種平面」為字集，該字集與國際標準化組織 ISO/IEC 10646.1:1993 一致，一共包含了20,902 個中文字，所以GF3001 的字集比CNS 11643-2 來得較大，除了繁體字外，也包含了簡化字。

GF3001 因字集較大且包含簡化字而獨有的基礎部件共有 87 個，詳見如下：

钅、纟、纟、贝、讠、鸟、车、リ、リ、弓、门、亻、夂、见、长、川、龙、丩、ツ、东、灬、东、为、韦、卩、彡、彡、乌、ス、乐、夕、飞、专、书、之、心、廿、卅、肅、彡、戸、由、吏、巾、上、卵、卵、卵、亚、亚、兕、黽、与、丹、先、甲、爲、亩、卅、円、彡、亠、市、臣、史、龙、两、市、丑、史、双、龜、フ、开、事、罍、亏、乘、市、市、罍、メ、个、夕、壺、酉

5. 其他

CNS 11643-2 遺漏的基礎部件有 2 個：弗、史。

GF3001 遺漏的基礎部件有 1 個：冊。¹

2.2.3 漢字構形資料庫的基礎部件

漢字構形資料庫在制定基礎部件時，主要是依據 CNS 11643-2，兼以 GF3001 為輔助參考，其與 CNS 11643-2 相異處如下：

1. 採取 GF3001 的認同使用原則

CNS 11643-2 因不採取認同使用的關係，產生了許多附部件，以部件「女」為例，分出位置在字形左半邊的附部件「女」，假設使用者需檢索「如」或「妒」等字形時，不能使用主部件「女」進行檢索，而必須使用它的附部件「女」進行檢索，才能檢索得到「如」和「妒」等字形。又例如部件「木」，分出位置在字形左半邊的附部件「木」和位置在字形下半邊的附部件「木」，因此當使用者需檢索「杜」字

¹ GF3001 已收錄基礎部件「弗」、「史」而CNS11643-2 遺漏之，CNS11643-2 已收錄基礎部件「冊」而GF3001 遺漏之。

時，需使用在字形左半邊的「木」為檢索部件，如果要檢索「架」字時，則需使用在字形下半邊的「禾」為檢索部件。由上述舉例可知，CNS 11643-2 之附部件在檢索上易造成混淆，也增加檢索之困難，因此漢字構形資料庫採取了 GF3001 的認同使用原則，但與 GF3001 之認同使用原則略有差異，其相異處分為底下 2 點：

(1)GF3001 認同使用，但漢字構形資料庫不認同使用者。在 GF3001 中，「乂」與「义」為同一部件，「几」與「几」也為同一部件。但漢字構形資料庫在考量字源的情況下，將「乂」、「义」以及「几」、「几」視為不同之基礎部件。

(2)GF3001 並未採取認同使用，但漢字構形資料庫認同使用者。漢字構形資料庫擴大了認同使用原則，加入認同使用的部件共有 14 個，如下所列，並在括弧中標出漢字構形資料庫所認同使用之部件：

匕 (匕)、乚 (匕)、尸 (尸)、尸 (巳)、艹 (艹)、丰 (丰)、巾 (巾)、己 (己)、巳 (巳)、丰 (丰)、丰 (丰)、丹 (丹)、白 (白)、羊 (羊)

2. 因一致性而增刪

此條原則是考慮部件拆分時的一致性，共分為底下 2 點：

(1) 因一致性而增加者。CNS 11643-2 將「水」、「氷」置於同一組，但卻拆分了「氷」。由於這三個部件有關聯性，在考量拆分一致性的情況下，漢字構形資料庫不拆分「氷」，並將之與「水」、「氷」列為同一組。

(2) 因一致性而拆分者。CNS 11643-2 將「𩚑」拆分，卻保留「𩚑」為基礎部件。在考量拆分一致性的情況下，漢字構形資料庫將「𩚑」加以拆分，不列入基礎部件表中。

3. 補上遺漏與刪除多出的五大字集部件

(1) 以五大字集為準，補上 CNS 11643-2 遺漏的五大字集中的二個部件：「弗」、「史」。

(2) 以五大字集為準，刪去 CNS 11643-2 多出的部件「𩚑」。

2.3 部件拆分的一些問題

漢字構形資料庫主要根據字形理據來進行部件拆分。當字形符合理據的，進行有理據拆分；無法分析理據或理據與字形矛盾的，依字形進行無理據拆分。另外，對多部件的漢字進行拆分時，應先依漢字組合層次做有理據拆分，直至不能進行有理據拆分而仍需拆分時，再做無理據拆分。如圖 2.1「醫」字的部件拆分。我們以《說文解字詁林》收錄的 9,820 個小篆的對應楷書為字集，依照上述原則拆分，並將所遇到的問題及處理方式整理如後。

2.3.1 部件拆分的類別與統計

根據字形理據來進行部件拆分，可分成有理據拆分及無理據拆分兩類，有理據拆分則可進一步再分成根據字源及參考字源兩類。

1. 有理據拆分—根據字源

絕大多數楷書字形拆分後的部件都和小篆相同，例如「醫」字拆成「毆」、「酉」。部分楷書部件由於隸變而稍有差異，仍然歸屬於此類。例如「江」的部件「氵」為「水」的變形；「到」的部件「刂」為「刀」的變形；「茄」的部件「艹」為「艸」的變形。9,820 個字中，部件完全相同的有 5,127 個，僅差異在變形部件的 3,338 個，以上合計 8,465 個。

2. 有理據拆分—參考字源

當楷書和小篆的部件差異不只是變形，而是由另一個部件所替代。例如小篆「奠(奠)」拆成「酋(酋)」、「兀(兀)」，楷書「奠」字拆成「酋」、「大」，「兀」為「大」所替代；小篆「香(香)」拆成「黍(黍)」、「甘(甘)」，楷書「香」字拆成「禾」、「日」，「黍」、「甘」則為「禾」、「日」所替代。9,820 個字中，有部件替代情形的字共 1,180 個。

3. 無理據拆分

當無法分析理據或理據與字形矛盾時，只好依照字形進行無理據拆分。例如小篆「兵(兵)」拆成「斤(斤)」、「冫(冫)」，楷書「兵」拆成「丘」、「八」；小篆「易(易)」為象形字，不再拆分，楷書「易」字則依 CNS11643-2 拆分成「日」、「勿」。9,820 個字中，有 175 個採用無理據拆分。

2.3.2 非字合成部件

在上述「有理據拆分—參考字源」的字形中，部分楷書的部件已由另一個部件所替代，這些替代的部件有些為非字部件。例如小篆「唐(唐)」拆成「𠩺(庚)」、「口」，楷書「唐」字中替代部件「庚」的「𠩺」，即為非字部件；小篆「制(制)」拆成「未(未)」、「刀」，楷書「制」字中替代部件「未」的「𠩺」，即為非字部件。非字部件可能是基礎部件，如「𠩺」；非字部件也可能是合成部件，如「𠩺」。

非字部件由於不是字，電腦的中文字集不會收錄。在漢字構形資料庫中，基礎部件是不可或缺的，因此非字基礎部件絕對要收錄；至於非字合成部件，由於數量較多，基於構字的需要，則可適量收錄。目前只要是具有特定構意的非字合成部件，若在《漢語大字典》有兩個（含）以上的單字用到，漢字構形資料庫即予收錄。例如「微」的非字合成部件「微」，構意為「微聲」，並可構成「微」、「微」等字，因此可收錄；至於「唐」的非字合成部件「𠩺」，構意為「庚聲」，但僅構成「唐」字，可予取消。

被取消的非字合成部件，所構之字在分析時，可直接使用該非字合成部件的下一級部件。例如「唐」的一級部件「𠩺」取消後，可直接使用二級部件「广」、「尹」構字，而將「唐」字分析成「广」、「尹」、「口」。

2.2.3 無理據拆分

對於需採無理據拆分的楷書字形，漢字構形資料庫的部件拆分原則如下：在不增加部件的情況下，使用最少的部件來拆分。例如「兵」拆分成「丘」、「八」；「易」拆分成「日」、「勿」。

2.4 五大字集部件表

1984年資策會發表了一套專為五大中文套裝軟體所設計的中文內碼，也就是俗稱的Big5中文碼或五大碼，其中含常用字5,401個，次常用字7,652個，合計13,053個字，其中有2個字重複編碼，因此實際收錄13,051個中文字²，即本文所稱之五大字集。在微軟的個人電腦作業系統中，從Windows 3.1至Windows 95/98都是採用五大碼作為繁體中文之編碼，直到Windows 2000之後的作業系統，才開始改用Unicode編碼。

依照漢字構形資料庫的部件拆分原則，五大字集拆分後的部件總數為2,297個，其中基礎部件為441個，合成部件為1,856個，詳述

²五大字集原本收錄13,053個中文字，但其中「兀」字及「設」字這兩字重複編碼，因此刪去重複兩字後為13,051個中文字。

如後。

2.4.1 五大字集基礎部件表

表 2.1 為五大字集基礎部件表，共有基礎部件 441 個，表格欄位之相關說明如下：

1.本基礎部件表包含主部件與附部件，主部件與附部件之區別在於，主部件是指部件主要字形，附部件則為主部件小幅度變化之字形。例如「心」為主部件，「小」與「忄」為其附部件。本基礎部件表共有主部件 382 個，附部件 59 個。

2.部件排序方式，以主部件筆畫數遞增式排序，亦即筆畫數少在前，多者在後。附部件跟隨於其所屬主部件之後。例如序號第 214 為該組之主部件「心」，其後跟隨附部件為序號 215 之「小」及 216 之「忄」。

3.部件若為同筆畫數者，依其筆順之筆畫序號遞增式排序。筆畫序號依 CNS11643-3「中文字筆畫分類」規定，筆畫順序為橫（一）豎（丨）撇（丿）點（丶）折（乚）。若第一筆畫是相同的，則依筆順往下比較。例如下列同為筆畫二畫之部件，當其第一筆都是橫畫時，此時就依照筆順的第二筆，依照 CNS11643-3「中文字筆畫分類」規定筆畫順序往下排列，其次序及依筆順之筆畫為：二（橫橫）、十（橫豎）、丁（橫豎）、厂（橫撇）、ナ（橫撇）、匕（橫折）。

4.所有基礎部件經排序後，序號由數字 1 起，就各個單獨基礎部件依序賦予序號，另就各組主部件依序賦予組號，附部件之組號與其所屬主部件相同，於表格中不另列出。例如序號 214 之主部件「心」，組號為 182；其附部件「小」及「忄」序號依序為 215 及 216，但組號則與主部件相同，均為 182。

5.例字編排依林樹編《中文電腦基本用字研究》（1972）之字頻排序，以字頻較高者在前，字頻較低者在後。例如，序號 214 部件「心」之例字為「想愛感寧恥」，其中「想」之字頻為 3,583 次，「愛」之字頻為 1,774 次，「感」之字頻為 721 次，「寧」之字頻為 291 次，「恥」之字頻為 192 次，因此依照字頻之高低排列後，部件「心」之例字排序為「想愛感寧恥」。

表 2.1：五大字集基礎部件表

序號	組號	部件	例	字	序號	組號	部件	例	字
1	1	一	是天同旦		41		儿	四罕	
2	2	一	刁		42	35	入	全兩糴叕	
3	3	丨	在條引叟芊		43	36	乂	艾刈	
4	4	丿	向少呂胤		44		乂	學希凶艾	
5	5	丶	主凡兔卵勺		45	37	儿	兒虎兆壺	
6	6	丶	尺		46	38	几	處凡飢冗	
7	7	乙	乾乞馱		47		几	朵	
8		乚	孔		48	39	几	夙	
9	8	丨	丁了予		49	40	勹	万別	
10	9	冂	成局司幻		50	41	勹	你	
11	10	勹	吳		51		勹	免	
12	11	勹	訊虱		52	42	勹	包旬	
13	12	勹	亡陋曷		53	43	勹	留派旅兜印	
14	13	勹	慶予壽疋		54	44	勹	叱	
15		勹	了候今		55	45	勹	六商夜京	
16	14	勹	尖		56	46	勹	班	
17	15	二	些次元仁貳互		57	47	勹	冷弱	
18	16	十	什南早古		58	48	勹	南商弟帝幸	
19	17	丁	丐		59	49	勹	於冬	
20	18	厂	原產反岸詹		60	50	勹	學軍受勞帝	
21	19	勹	有在		61	51	勹	假巨貞	
22	20	勹	能此尼旨鬯		62	52	勹	侯	
23		勹	它		63	53	勹	報命印卯	
24	21	勹	匠		64		勹	範犯卷宛厄	
25	22	勹	區		65	54	勹	叫收剋	
26	23	勹	切皂柒		66	55	勹	出匈屈凶禽	
27	24	丁	可打頂亭		67	56	勹	分切絕召賴	
28	25	勹	巧兮亟攷		68		勹	到	
29	26	勹	焉与		69	57	勹	動加勞辦務	
30	27	勹	下外赴鴻		70	58	勹	受友取隻反	
31		勹	上桌竊		71		勹	祭	
32	28	勹	同南向剛炯奐		72	59	勹	仍秀孕	
33		勹	角		73	60	勹	能公參私允	
34	29	勹	囧		74	61	勹	令甬	
35	30	勹	無傷族鹽乞		75	62	勹	究軌旭尻叕墟	
36	31	勹	介鼻丌氣		76	63	勹	乚	
37	32	勹	派遞后盾梃		77	64	勹	亨釘	
38	33	勹	以幾閃企卒		78	65	勹	鄰	
39		勹	他候		79	66	勹	幸岸幹汗刊	
40	34	勹	分六匹睿扒		80	67	勹	宇迂孟吁	

序號	組號	部件	例	字	序號	組號	部件	例	字
81	68	丰	那半奉	𠂇	123	110	弓	發張弱躬彎	粥
82	69	土	在去地社疆	𠂇	124	111	彳	急尋雪帚	
83	70	士	志毅壯王教	𠂇	125	112	五	牽	
84	71	工	江空功式左貢	𠂇	126	113	巾	出弢焮焮	
85	72	大	天因奇夾杖	𠂇	127	114	也	他拖匪池	
86	73	尢	拋尫尫	𠂇	128	115	女	要好委威瀛	
87	74	升	開算奔戒葬	𠂇	129	116	爻	建	
88	75	冫	冱	𠂇	130	117	匕	辰畏	
89	76	丈	仗	𠂇	131	118	小	你少京叔齋	
90	77	牛	降桀牽	𠂇	132		灬	當尚肖	
91	78	𠂇	虐	𠂇	133	119	子	好學孔孟困	
92	79	寸	將專付冠刊	𠂇	134		子	子	
93	80	弋	代式鳶茈雉	𠂇	135	120	彡	彙	
94	81	去	育流充棄	𠂇	136		彡	彙	
95	82	才	材閉轟	𠂇	137	121	卩	那院	
96	83	丿	北燕	𠂇	138	122	乡	鄉雍	
97	84	少	步歲賓	𠂇	139	123	子	子	
98	85	口	可和問叫各	𠂇	140	124	幺	麼後樂幼茲幻	
99	86	口	國	𠂇	141	125	川	經腦災巡牽	
100	87	山	島微岸仙峽	𠂇	142		川	順訓芥氙	
101	88	巾	市布帽飾帥匪	𠂇	143		儿	流荒侃	
102	89	巾	爾	𠂇	144	126	王	全現班弄匡閏	
103	90	乇	托	𠂇	145	127	井	耕耨井	
104	91	乇	托宅	𠂇	146	128	夫	規扶麩芙	
105	92	千	乖阡芊	𠂇	147	129	丰	責	
106	93	彳	得	𠂇	148		丰	害	
107	94	彳	形須參彥彪	𠂇	149	130	考	考	
108	95	勹	匆囟	𠂇	150	131	丐	鈔	
109	96	久	各處隆贛	𠂇	151	132	廿	度燕董	
110	97	攴	後愛致	𠂇	152	133	卩	共備展散	
111	98	久	畝疲灸羨	𠂇	153	134	木	校新條林樂查	
112	99	卯	卯	𠂇	154		木	余	
113	100	夕	多外名夢矽	𠂇	155	135	市	述癩	
114	101	丸	執芄尫	𠂇	156	136	市	沛旆芾	
115	102	凡	恐羸	𠂇	157	137	卅	卅	
116	103	广	應	𠂇	158	138	不	否杯罌礪	
117	104	宀	家	𠂇	159	139	犬	類哭默莽戾倏	
118	105	艹	前業善兼朔岡	𠂇	160		犴	狗	
119	106	尸	局屋刷殿辟	𠂇	161	140	歹	死列夙	
120	107	己	記改忌登	𠂇	162		歹	餐	
121	108	巳	起包巷熙祀	𠂇	163	141	五	吾伍	
122	109	巳	已	𠂇	164	142	屯	頓純囤窰迺	

序號	組號	部件	例	字	序號	組號	部件	例	字
165	143	旡	既蠶潛		207	176	文	產蚊斑度斐閔斌	
166	144	戈	或幾找划盞		208	177	火	勞燈滅秋灰災炎	
167	145	牙	呀穿雅迓		209		灬	然盡顯	
168	146	艹	花		210	178	宀	率函脊兆	
169	147	卝	苟		211	179	辶	這	
170	148	止	此步歸武歷企		212	180	之	芝	
171		辵	是從徙		213	181	尢	沈鳩尠髡	
172	149	日	是時間者普晶		214	182	心	想愛感寧恥	
173	150	日	書勗汨吹		215		小	恭	
174	151	冂	曼		216		忄	情	
175		月	冒胃		217	183	一	牽	
176	152	艹	衰		218	184	巾	唐	
177	153	中	忠衷冲帶		219	185	尸	倉	
178	154	囟	象		220	186	尹	爭	
179	155	冂	雋		221	187	夂	快莢	
180	156	內	納芮氤		222	188	弔	弟侖弔盃	
181	157	手	看拿掌摩		223	189	爿	將藏寢妝	
182		廴	拜		224	190	丑	紐羞	
183		扌	把		225	191	巴	把爸爬疤邑	
184	158	攴	教條變務肇煞釐		226	192	尸	聲眉	
185	159	毛	尾耗毯髦		227	193	勹	鹿	
186	160	气	氣汽芻		228	194	尹	君伊笋	
187	161	牛	解件牢犀犛		229	195	毋	毒	
188		生	先告		230		毋	貫虜	
189	162	丰	邦蚌砦傘		231		母	每姆	
190		丰	豐擘		232	196	水	冰泉尿潁盥沓	
191	163	片	牌泐		233		氵	法衍	
192	164	斤	所近質匠欣斧		234		冫	暴漆黎滕	
193	165	爪	爬抓策		235		冫	壞屬鯨	
194		心	受		236	197	夫	春	
195		扌	印褒		237	197	瓦	瓶瓷	
196	166	戶	所房妒		238	198	艹	賽襄葑	
197	167	父	交爸蚊		239	199	未	業妹釐寐	
198	168	月	有前明服朋閒		240		未	抹茱	
199		夕	望		241	200	示	款粟尉余	
200	169	氏	紙昏底		242		礻	社	
201	170	丹	坍旃彤		243	201	甘	某甜邯疳	
202	171	小	亦赤		244	202	世	葉屨貰泄泄	
203	172	衣	表		245	203	本	笨鉢栩	
204	173	艮	派旅		246	204	巾	囊	
205	174	勿	物易忽勿囫		247	205	丙	病柄曷邴	
206	175	及	級岌		248	206	石	研磨岩拓蠹磊	

序號	組號	部件	例	字	序號	組號	部件	例	字
249	207	冊	冊		291	246	聿	肅	
250	208	友	髮拔魁		292	247	夊	庚	
251	209	戊	茂戊		293	248	𠂔	假霞	
252		戊	越		294	249	弗	費佛刺氟莠	
253	210	以	似苡		295	250	民	眠啟莨	
254	211	目	看相省眼眉昊		296	251	皮	被頗疲簸髮	
255	212	且	姐助宜疽		297	252	卩	聯	
256	213	田	當界細畫福略甸		298	253	夂	發	
257	214	由	油迪笛寅甯		299	254	矛	柔茅矜淋裘	
258	215	甲	鴨押閘匣屨		300	255	冂	冂	
259	216	申	神暢氈		301	256	耳	聲取聞聶聃	
260	217	皿	盡醢齋		302	257	其	其	
261	218	𠂔	罷		303	258	臣	宦臥囂臧拒	
262	219	史	駛		304	259	𠂔	要	
263	220	央	英映盞		305		西	晒璽茜氘	
264	221	兕	免		306		𠂔	𠂔	
265	222	冉	再髯聃		307	260	吏	使	
266	223	冊	刪柵		308	261	束	策棗刺棘	
267		冊	獮		309	262	而	需耐耍沔	
268	224	冊	侖		310	263	互	恆	
269	225	业	並業		311	264	至	到屋室姪戡	
270	226	电	奄		312	265	𠂔	彘	
271		电	電		313	266	夷	姨痲莨	
272	227	内	離亂竊		314	267	虍	虎	
273	228	凹	兕		315	268	且	直具	
274	229	凸	凸		316	269	𠂔	缺	
275	230	吕	官師耜		317	270	曲	農蚰齧	
276	231	生	產星性隆甥甦眚		318	271	虫	強雖蟲蜂蛋融	
277	232	乍	作恚窄斥		319	272	曳	洩	
278	233	禾	和乘委穎穌困		320	273	业	虛	
279	234	𠂔	段		321	274	冎	骨冎	
280	235	丘	兵邱蚯		322	275	肉	腐臠馘腩	
281	236	白	的原樂怕皆皇		323		月	能育臉胡	
282	237	瓜	狐瓣馱窳蠶		324		夕	然將祭炙	
283	238	乎	呼庠		325	276	岳	寶搖缺陶鬱罄	
284	239	用	備佣甬		326	277	耒	耕誅	
285	240	甩	甩		327	278	年	哖	
286	241	𠂔	姊第越		328	279	𠂔	制	
287	242	疒	病		329	280	缶	卸	
288	243	立	位童站笠		330	281	竹	竹	
289	244	必	秘瑟謚邲闕		331		𠂔	第	
290	245	永	泳昶羨		332	282	自	息咱卽	

序號	組號	部件	例	字	序號	組號	部件	例	字
417	358	烏	烏	鳴 鄔	430	370	壺	壺	
418	359	鬼	愧	魔 魁 鬼	431	371	鼎	鼎	沸
419	360	兼	廉	歉 謙 兼	432	372	𠂔	𠂔	
420	361	雀	確	鶴 雀	433	373	單	戰	彈 闡
421	362	董	勤	謹 董	434	374	黽	蠅	鼈 鄢
422		里	董		435	375	黑	點	黨 墨 黴 嘿
423	363	莫	難	漢	436	376	熏	勳	薰 醺
424	364	帶	滯	蒂 遶	437	377	齊	濟	劑 齋 薺
425	365	曹	遭	槽	438	378	齒	齡	齧 嚙
426	366	棄	棄		439	379	龍	籠	襲 隴 龐
427	367	畢	嘩	筆 鷓	440	380	義	犧	
428	368	婁	數	屢 樓 婁	441	381	龜	鬪	
429	369	庸	傭	鄘					

2.4.2 五大字集基礎部件組字頻率表

表 2.2 為五大字集基礎部件組字頻率表，表格欄位之相關說明如下：

1.本組字頻率表之編排順序，是以部件組字頻率高者在前，依序給予編號。例如部件「口」之組字頻率為 6.29%，較部件「一」的 6.03% 為高，因此「口」之編號排在「一」字之前。

2.本組字頻率表之序號與五大字集基礎部件表中的序號一致。例如編號 1 之部件「口」序號為 98，與部件「口」在「五大字集基礎部件表」中的序號一致。

3.組字字數為該部件在五大字集的 13,051 個中文字中，所組成的字形個數，例如部件「口」在五大字集中，組成了 2,526 個字形。

4.組字頻率之計算方式為，根據林樹編《中文電腦基本用字研究》(1972)中的字頻³，可推知該部件的組字頻次。將部件的組字頻次，再除以全部 441 個部件組字之總頻次，即可算出該部件的組字頻率。以部件「束」為例，在五大字集中組字字數有 15 個，其組成字形、字頻與該字形貢獻於部件「束」之組字頻次如下所列：

刺(13, 13)、棘(5, 10)、棗(83, 166)、策(232, 232)、束(1, 1)、揀(1, 1)、棟(1, 1)、策(1, 1)、莉(1, 1)、燹(1, 2)、蔽(1, 1)、慊(1, 1)、棘(1, 2)、裨(1, 2)、鏖(1, 1)

字形後的括弧中第 1 個數字代表該字形列於《中文電腦基本用字研究》(1972)中的字頻，第 2 個數字代表「束」這個部件因為該字形而貢獻的組字頻次。例如「棘」字在《中文電腦基本用字研究》中的字頻列為 5 次，而「棘」字又是由 2 個「束」所組成，所以部件「束」因為「棘」字而貢獻的組字頻次為 10 次。將部件「束」在這 15 個字形中的組字頻次加總之後得到總和為 435 次，代表部件「束」在五大字集中的組字頻次為 435 次。將全部 441 個基礎部件的組字頻次累加，可得到 441 個基礎部件組字的總頻次為 5,233,534 次。「束」在五大字集中的組字頻次為 435 次，除以全部 441 個基礎部件組字的總頻次 5,233,534 次，即可算出「束」的組字頻次約佔全部部件組字總頻次之 0.01% 左右，此即部件「束」在五大字集中的組字頻率。

³ 若為未列在林樹編《中文電腦基本用字研究》(1972)中的字形，則給予頻次 1 次。

5. 累計頻率即為組字頻率之累加而成。由累計頻率可以看出，至部件編號第 308 個部件時，其組字的累計頻率已累計至 99.02%，代表前 308 個基礎部件已組成了五大字集中的 99.02% 的字形。

表 2.2：五大字集基礎部件組字頻率表

編號	序 號	部 件	組字字數	組字頻率(%)	累計頻率(%)
1	98	口	2526	6.29%	6.29%
2	1	一	1732	6.03%	12.32%
3	38	人	955	2.56%	14.88%
4	172	日	895	2.51%	17.39%
5	39	亻	551	2.47%	19.86%
6	5	丶	308	2.04%	21.90%
7	82	土	795	1.82%	23.72%
8	281	白	191	1.72%	25.44%
9	52	勹	292	1.64%	27.08%
10	40	八	713	1.57%	28.66%
11	18	十	572	1.48%	30.14%
12	153	木	956	1.42%	31.56%
13	373	言	364	1.24%	32.80%
14	55	宀	590	1.22%	34.02%
15	211	辶	236	1.11%	35.12%
16	99	口	262	1.08%	36.21%
17	233	讠	855	1.00%	37.21%
18	70	又	483	1.00%	38.20%
19	85	大	505	0.97%	39.17%
20	171	乚	180	0.95%	40.12%
21	22	匕	321	0.93%	41.05%
22	117	宀	450	0.92%	41.97%
23	198	月	190	0.91%	42.88%
24	21	ナ	138	0.89%	43.77%
25	45	儿	408	0.85%	44.62%
26	131	小	166	0.83%	45.45%
27	128	女	442	0.83%	46.28%
28	92	寸	227	0.82%	47.10%
29	4	丿	390	0.75%	47.85%
30	133	子	161	0.73%	48.58%
31	214	心	274	0.72%	49.30%
32	140	幺	132	0.72%	50.02%
33	127	也	34	0.71%	50.73%
34	32	冂	350	0.68%	51.41%
35	166	戈	167	0.66%	52.07%
36	106	彳	137	0.66%	52.73%
37	168	艹	873	0.65%	53.38%
38	288	立	242	0.64%	54.03%

編號	序 號	部 件	組字字數	組字頻率(%)	累計頻率(%)
39	183	扌	515	0.64%	54.67%
40	144	王	330	0.64%	55.30%
41	137	冫	306	0.63%	55.94%
42	184	攴	263	0.63%	56.57%
43	73	厶	278	0.62%	57.19%
44	27	丁	149	0.59%	57.78%
45	391	門	128	0.57%	58.35%
46	353	纟	324	0.57%	58.92%
47	58	丶	398	0.56%	59.47%
48	68	丩	206	0.54%	60.02%
49	3	丨	90	0.54%	60.56%
50	83	士	188	0.54%	61.09%
51	35	亠	243	0.53%	61.63%
52	323	月	453	0.53%	62.16%
53	367	我	26	0.52%	62.68%
54	158	不	39	0.52%	63.20%
55	77	了	6	0.51%	63.71%
56	49	夕	158	0.50%	64.21%
57	60	冂	267	0.49%	64.70%
58	149	艹	83	0.48%	65.18%
59	118	艹	229	0.48%	65.67%
60	170	止	330	0.48%	66.15%
61	278	禾	264	0.48%	66.63%
62	397	隹	271	0.48%	67.11%
63	31	卜	135	0.46%	67.57%
64	256	田	529	0.46%	68.03%
65	113	夕	121	0.44%	68.47%
66	366	貝	322	0.44%	68.91%
67	116	广	264	0.44%	69.35%
68	44	乂	137	0.44%	69.79%
69	332	自	180	0.43%	70.21%
70	69	力	152	0.42%	70.63%
71	42	入	42	0.41%	71.04%
72	254	目	252	0.40%	71.44%
73	132	灬	170	0.39%	71.84%
74	84	工	230	0.39%	72.23%
75	261	四	234	0.38%	72.61%
76	17	二	153	0.37%	72.98%

編號	序 號	部 件	組字字數	組字頻率(%)	累計頻率(%)
77	209	灬	554	0.37%	73.35%
78	50	勹	121	0.34%	73.69%
79	101	巾	150	0.34%	74.02%
80	331	夂	260	0.33%	74.36%
81	10	冫	41	0.32%	74.68%
82	155	水	45	0.30%	74.98%
83	311	至	53	0.30%	75.28%
84	51	㇀	356	0.29%	75.57%
85	362	里	50	0.28%	75.85%
86	79	干	160	0.28%	76.13%
87	67	刀	195	0.28%	76.40%
88	109	久	110	0.27%	76.67%
89	192	斤	105	0.27%	76.94%
90	301	耳	125	0.27%	77.21%
91	81	丰	52	0.27%	77.48%
92	208	火	359	0.27%	77.75%
93	20	厂	245	0.27%	78.02%
94	47	几	127	0.26%	78.28%
95	348	艮	53	0.26%	78.54%
96	216	忄	282	0.26%	78.80%
97	355	車	200	0.25%	79.05%
98	344	羊	73	0.24%	79.29%
99	334	白	57	0.24%	79.53%
100	304	西	86	0.24%	79.77%
101	276	生	38	0.24%	80.01%
102	147	丰	69	0.24%	80.25%
103	365	見	69	0.24%	80.49%
104	340	衤	153	0.23%	80.72%
105	399	金	493	0.22%	80.94%
106	318	虫	448	0.22%	81.16%
107	119	尸	183	0.22%	81.37%
108	110	攴	128	0.21%	81.59%
109	181	手	38	0.21%	81.80%
110	194	心	154	0.21%	82.01%
111	177	中	45	0.21%	82.22%
112	123	弓	90	0.21%	82.42%
113	360	豕	97	0.20%	82.63%
114	196	戶	73	0.20%	82.83%
115	30	卜	25	0.20%	83.03%
116	86	尢	36	0.20%	83.23%

編號	序 號	部 件	組字字數	組字頻率(%)	累計頻率(%)
117	407	重	23	0.20%	83.42%
118	225	巴	58	0.19%	83.62%
119	253	以	5	0.19%	83.80%
120	63	尸	58	0.19%	83.99%
121	87	升	125	0.18%	84.17%
122	121	巳	61	0.18%	84.36%
123	159	犬	86	0.18%	84.54%
124	152	土	113	0.18%	84.71%
125	284	用	29	0.18%	84.89%
126	342	米	148	0.18%	85.06%
127	197	父	38	0.17%	85.24%
128	411	首	7	0.17%	85.41%
129	333	白	105	0.17%	85.58%
130	100	山	352	0.17%	85.75%
131	74	マ	89	0.17%	85.91%
132	251	戊	71	0.17%	86.08%
133	205	勿	70	0.16%	86.24%
134	66	凵	93	0.16%	86.40%
135	269	业	43	0.16%	86.56%
136	321	丹	66	0.16%	86.72%
137	415	馬	163	0.16%	86.88%
138	275	吕	51	0.15%	87.03%
139	188	牛	60	0.15%	87.17%
140	187	牛	90	0.15%	87.32%
141	146	夫	109	0.14%	87.46%
142	59	丶	50	0.14%	87.61%
143	352	糸	111	0.14%	87.75%
144	314	虍	130	0.14%	87.89%
145	41	儿	55	0.14%	88.03%
146	169	卅	77	0.14%	88.17%
147	315	且	48	0.14%	88.30%
148	203	衣	44	0.14%	88.44%
149	277	乍	31	0.13%	88.57%
150	412	為	12	0.13%	88.70%
151	105	千	30	0.13%	88.83%
152	231	母	31	0.13%	88.96%
153	107	彡	166	0.12%	89.09%
154	298	夂	44	0.12%	89.21%
155	126	中	83	0.12%	89.34%
156	163	五	27	0.12%	89.46%

編號	序 號	部 件	組字字數	組字頻率(%)	累計頻率(%)
157	151	廿	77	0.12%	89.58%
158	46	几	89	0.12%	89.70%
159	260	皿	178	0.12%	89.82%
160	339	衣	90	0.12%	89.94%
161	383	事	3	0.12%	90.06%
162	242	衤	82	0.12%	90.17%
163	384	雨	133	0.12%	90.29%
164	13	乚	89	0.11%	90.40%
165	223	爿	38	0.11%	90.52%
166	255	且	55	0.11%	90.63%
167	327	年	3	0.11%	90.74%
168	222	帛	20	0.11%	90.85%
169	302	其	40	0.11%	90.97%
170	141	ㄩ	79	0.11%	91.08%
171	241	示	90	0.11%	91.18%
172	248	石	219	0.11%	91.29%
173	173	曰	79	0.11%	91.40%
174	239	耒	43	0.11%	91.51%
175	290	永	8	0.10%	91.61%
176	377	長	15	0.10%	91.71%
177	212	之	13	0.10%	91.81%
178	232	水	47	0.10%	91.91%
179	324	夕	44	0.10%	92.01%
180	349	艮	51	0.10%	92.11%
181	357	更	18	0.10%	92.21%
182	296	皮	37	0.10%	92.30%
183	15	丿	80	0.09%	92.40%
184	404	面	12	0.09%	92.49%
185	343	羊	55	0.09%	92.59%
186	65	乚	9	0.09%	92.68%
187	94	去	30	0.09%	92.77%
188	265	冉	31	0.09%	92.86%
189	268	冊	59	0.09%	92.95%
190	120	己	18	0.09%	93.03%
191	341	亥	27	0.08%	93.12%
192	245	本	10	0.08%	93.20%
193	124	彳	55	0.08%	93.28%
194	200	氏	52	0.08%	93.37%
195	28	丂	75	0.08%	93.45%
196	309	而	67	0.08%	93.53%

編號	序 號	部 件	組字字數	組字頻率(%)	累計頻率(%)
197	295	民	20	0.08%	93.61%
198	388	果	38	0.08%	93.69%
199	14	一	79	0.08%	93.78%
200	303	臣	58	0.08%	93.85%
201	7	乙	37	0.08%	93.93%
202	33	冂	84	0.08%	94.01%
203	257	由	42	0.08%	94.09%
204	435	黑	47	0.08%	94.17%
205	160	彳	143	0.08%	94.25%
206	207	文	53	0.08%	94.32%
207	305	西	28	0.08%	94.40%
208	75	九	24	0.08%	94.48%
209	317	曲	24	0.08%	94.55%
210	96	亠	27	0.07%	94.63%
211	393	非	46	0.07%	94.70%
212	93	弋	38	0.07%	94.77%
213	369	身	17	0.07%	94.84%
214	382	東	14	0.07%	94.91%
215	361	求	20	0.07%	94.98%
216	359	酉	134	0.07%	95.05%
217	221	夂	20	0.07%	95.12%
218	238	艹	65	0.07%	95.19%
219	410	食	73	0.06%	95.25%
220	26	七	7	0.06%	95.32%
221	29	与	27	0.06%	95.38%
222	161	歹	70	0.06%	95.44%
223	394	無	18	0.06%	95.51%
224	90	牛	91	0.06%	95.57%
225	358	束	67	0.06%	95.63%
226	122	巳	1	0.06%	95.70%
227	347	聿	7	0.06%	95.76%
228	186	气	32	0.06%	95.82%
229	175	月	31	0.06%	95.88%
230	62	冫	14	0.06%	95.94%
231	335	舟	72	0.06%	96.00%
232	287	疒	157	0.06%	96.05%
233	264	兕	38	0.06%	96.11%
234	230	毌	9	0.06%	96.17%
235	114	丸	33	0.06%	96.22%
236	370	鳥	245	0.06%	96.28%

編號	序 號	部 件	組字字數	組字頻率(%)	累計頻率(%)
237	130	比	71	0.06%	96.34%
238	346	聿	31	0.05%	96.39%
239	381	虫	34	0.05%	96.44%
240	206	及	22	0.05%	96.50%
241	95	才	7	0.05%	96.55%
242	405	禺	37	0.05%	96.60%
243	154	示	36	0.05%	96.66%
244	228	尹	24	0.05%	96.71%
245	351	羽	157	0.05%	96.76%
246	108	ㄨ	14	0.05%	96.81%
247	244	世	41	0.05%	96.86%
248	423	萸	14	0.05%	96.91%
249	129	彳	49	0.05%	96.96%
250	167	牙	26	0.05%	97.02%
251	338	彡	29	0.05%	97.06%
252	289	必	33	0.05%	97.11%
253	307	吏	2	0.05%	97.16%
254	297	艸	4	0.05%	97.21%
255	433	單	39	0.05%	97.26%
256	226	尸	19	0.05%	97.30%
257	97	少	40	0.04%	97.35%
258	57	冫	34	0.04%	97.39%
259	178	凵	9	0.04%	97.44%
260	409	食	20	0.04%	97.48%
261	180	內	16	0.04%	97.52%
262	64	巳	90	0.04%	97.56%
263	25	匸	85	0.04%	97.60%
264	165	旡	25	0.04%	97.64%
265	271	电	1	0.04%	97.68%
266	236	夫	45	0.04%	97.72%
267	56	冫	2	0.04%	97.76%
268	143	儿	20	0.04%	97.80%
269	9	丿	16	0.04%	97.84%
270	356	甫	54	0.04%	97.87%
271	72	乃	20	0.04%	97.91%
272	259	申	16	0.04%	97.95%
273	125	五	45	0.03%	97.98%
274	234	水	50	0.03%	98.01%
275	36	川	22	0.03%	98.05%
276	243	甘	52	0.03%	98.08%

編號	序 號	部 件	組字字數	組字頻率(%)	累計頻率(%)
277	53	彳	34	0.03%	98.12%
278	325	缶	51	0.03%	98.15%
279	148	丰	16	0.03%	98.19%
280	210	以	46	0.03%	98.22%
281	263	央	33	0.03%	98.25%
282	371	采	54	0.03%	98.29%
283	437	齊	25	0.03%	98.32%
284	428	婁	35	0.03%	98.35%
285	220	尹	16	0.03%	98.39%
286	235	冰	26	0.03%	98.42%
287	328	帛	6	0.03%	98.45%
288	218	尹	27	0.03%	98.48%
289	61	コ	52	0.03%	98.51%
290	199	夕	1	0.03%	98.54%
291	185	毛	59	0.03%	98.56%
292	424	帶	12	0.03%	98.59%
293	280	丘	9	0.03%	98.62%
294	71	ㄨ	16	0.03%	98.65%
295	418	鬼	46	0.03%	98.68%
296	202	小	21	0.03%	98.70%
297	189	丰	51	0.03%	98.73%
298	345	州	8	0.03%	98.76%
299	413	飛	2	0.03%	98.79%
300	227	曲	73	0.03%	98.81%
301	138	乡	18	0.03%	98.84%
302	299	矛	61	0.03%	98.87%
303	135	彡	44	0.03%	98.89%
304	219	尸	22	0.03%	98.92%
305	403	革	75	0.02%	98.94%
306	380	亞	14	0.02%	98.97%
307	23	匕	19	0.02%	98.99%
308	191	片	14	0.02%	99.02%
309	294	弗	26	0.02%	99.04%
310	190	丰	18	0.02%	99.06%
311	247	丙	13	0.02%	99.09%
312	379	耳	21	0.02%	99.11%
313	8	乚	24	0.02%	99.13%
314	272	内	26	0.02%	99.16%
315	240	末	11	0.02%	99.18%
316	204	反	12	0.02%	99.20%

編號	序 號	部 件	組字字數	組字頻率(%)	累計頻率(%)
317	283	乎	11	0.02%	99.22%
318	406	垂	19	0.02%	99.24%
319	115	夂	19	0.02%	99.26%
320	376	𠂇	36	0.02%	99.28%
321	37	厂	47	0.02%	99.30%
322	24	匸	34	0.02%	99.32%
323	142	川	12	0.02%	99.34%
324	293	冫	23	0.02%	99.35%
325	111	久	7	0.02%	99.37%
326	378	𠂇	55	0.02%	99.39%
327	372	豸	35	0.02%	99.41%
328	322	肉	10	0.02%	99.42%
329	375	𠂇	25	0.02%	99.44%
330	416	鬥	6	0.02%	99.45%
331	102	巾	17	0.02%	99.47%
332	195	冫	2	0.02%	99.49%
333	34	冫	20	0.02%	99.50%
334	439	龍	35	0.02%	99.52%
335	385	丽	16	0.01%	99.53%
336	258	甲	16	0.01%	99.54%
337	421	董	19	0.01%	99.56%
338	282	瓜	31	0.01%	99.57%
339	174	冫	19	0.01%	99.59%
340	326	耒	26	0.01%	99.60%
341	414	鬲	27	0.01%	99.61%
342	400	隶	14	0.01%	99.62%
343	252	戍	9	0.01%	99.63%
344	386	冫	11	0.01%	99.65%
345	11	冫	12	0.01%	99.66%
346	215	小	11	0.01%	99.67%
347	427	畢	13	0.01%	99.68%
348	401	承	1	0.01%	99.69%
349	193	爪	5	0.01%	99.70%
350	164	屯	25	0.01%	99.71%
351	419	兼	30	0.01%	99.72%
352	250	友	24	0.01%	99.73%
353	279	冫	8	0.01%	99.74%
354	182	冫	2	0.01%	99.75%
355	89	丈	5	0.01%	99.76%
356	237	瓦	34	0.01%	99.76%

編號	序 號	部 件	組字字數	組字頻率(%)	累計頻率(%)
357	337	𠂇	13	0.01%	99.77%
358	434	𠂇	17	0.01%	99.78%
359	262	史	2	0.01%	99.79%
360	6	丿	8	0.01%	99.80%
361	145	井	6	0.01%	99.81%
362	308	束	15	0.01%	99.82%
363	392	妻	10	0.01%	99.83%
364	104	乇	20	0.01%	99.83%
365	330	竹	1	0.01%	99.84%
366	80	于	24	0.01%	99.85%
367	112	冫	17	0.01%	99.86%
368	398	卑	36	0.01%	99.86%
369	368	𠂇	21	0.01%	99.87%
370	291	聿	15	0.01%	99.88%
371	229	毋	6	0.01%	99.89%
372	201	丹	12	0.01%	99.89%
373	420	雀	7	0.01%	99.90%
374	12	冫	11	0.01%	99.91%
375	438	齒	38	0.01%	99.91%
376	417	烏	9	0.01%	99.92%
377	266	冊	7	0.01%	99.92%
378	425	曹	13	0.01%	99.93%
379	286	弟	7	0.00%	99.93%
380	88	冫	21	0.00%	99.94%
381	320	𠂇	9	0.00%	99.94%
382	224	丑	16	0.00%	99.95%
383	390	典	12	0.00%	99.95%
384	363	串	5	0.00%	99.95%
385	313	夷	16	0.00%	99.96%
386	426	棄	1	0.00%	99.96%
387	162	夕	7	0.00%	99.96%
388	374	𠂇	2	0.00%	99.97%
389	43	乂	7	0.00%	99.97%
390	441	龜	2	0.00%	99.97%
391	176	冫	6	0.00%	99.97%
392	440	義	4	0.00%	99.98%
393	270	电	20	0.00%	99.98%
394	91	冫	3	0.00%	99.98%
395	213	允	19	0.00%	99.98%
396	156	市	7	0.00%	99.98%

編號	序 號	部 件	組字字數	組字頻率(%)	累計頻率(%)
397	150	丐	2	0.00%	99.98%
398	402	韭	19	0.00%	99.99%
399	429	庸	10	0.00%	99.99%
400	408	禹	9	0.00%	99.99%
401	310	互	5	0.00%	99.99%
402	389	走	14	0.00%	99.99%
403	329	缶	7	0.00%	99.99%
404	319	曳	6	0.00%	99.99%
405	436	熏	11	0.00%	99.99%
406	387	豕	10	0.00%	99.99%
407	246	巾	9	0.00%	99.99%
408	354	臣	7	0.00%	100.00%
409	273	凹	2	0.00%	100.00%
410	48	几	3	0.00%	100.00%
411	336	肩	3	0.00%	100.00%
412	274	凸	1	0.00%	100.00%
413	396	叟	9	0.00%	100.00%
414	19	丁	7	0.00%	100.00%
415	157	卅	1	0.00%	100.00%
416	292	夊	4	0.00%	100.00%
417	431	鼎	6	0.00%	100.00%
418	179	凹	6	0.00%	100.00%
419	2	一	2	0.00%	100.00%

編號	序 號	部 件	組字字數	組字頻率(%)	累計頻率(%)
420	54	匕	1	0.00%	100.00%
421	395	秉	1	0.00%	100.00%
422	285	甩	1	0.00%	100.00%
423	300	卍	1	0.00%	100.00%
424	134	子	1	0.00%	100.00%
425	136	丩	1	0.00%	100.00%
426	350	丩	3	0.00%	100.00%
427	76	乚	1	0.00%	100.00%
428	312	戈	1	0.00%	100.00%
429	430	壺	1	0.00%	100.00%
430	432	盟	1	0.00%	100.00%
431	16	丿	1	0.00%	100.00%
432	78	ㄩ	1	0.00%	100.00%
433	103	乚	1	0.00%	100.00%
434	139	子	1	0.00%	100.00%
435	217	一	1	0.00%	100.00%
436	249	冊	1	0.00%	100.00%
437	267	冊	1	0.00%	100.00%
438	306	西	1	0.00%	100.00%
439	316	夊	1	0.00%	100.00%
440	364	弗	1	0.00%	100.00%
441	422	重	1	0.00%	100.00%

2.4.3 五大字集合成部件表

五大字集部件集中，除了 441 個基礎部件以外，還有 1,856 個合成部件。合成部件與基礎部件不同，基礎部件是無法再往下拆分的部件，而合成部件則可往下繼續拆分。

合成部件可分為成字合成部件與非字合成部件兩大類。成字合成部件是指當它不作其他字的部件時，本身就是一個完整的字，例如「冬」可當「終」、「疼」的部件，提示它們的讀音，而且「冬」本身也是一個獨立的字。非字合成部件是依附在其他部件來體現構意的部件，本身不能獨立存在，必須在字形結構的環境中才能體現它的構意。

成字合成部件與非字合成部件又根據其來源出處之不同，可細分為底下四類：

1. 為《漢語大字典》單字的成字合成部件：此類為收錄於《漢語大字典》的單字。在五大字集中，此類部件共有 1,757 個，其部件、在五大字集中的組字數及例字⁴，如表 2.3。

2 為其他字書單字的成字合成部件：此類為其他字書收錄的單字，但《漢語大字典》並未收錄。在五大字集中，此類部件共有 15 個，其部件、在五大字集中的組字數及例字，如表 2.4。

3. 包含在 GF3001 中的非字合成部件：此類為 GF3001 的基礎部件，但可依 CNS 11643-2 的拆分原則拆分，故並不將之列於基礎部件集中。在五大字集中，此類部件共有 8 個，其部件、在五大字集中的組字數及例字，如表 2.5。

4. 具有特定構意的非字合成部件：此類部件有特定的構意，且在《漢語大字典》中有兩個（含）以上的單字用到該部件，具有構形的能力，即收錄為非字合成部件，在五大字集中，此類部件有 76 個，其部件、在五大字集中的組字數及例字，如表 2.6。

⁴例字編排以林樹編《中文電腦基本用字研究》(1972)之字頻排序，以字頻較高者在前，字頻較低者在後。

表 2.3：為《漢語大字典》單字的五大字集成字合成部件

序號	部 件	組字數	例 字
1	刁	2	叨
2	三	12	仁閔
3	亍	29	行
4	亏	47	虧夸
5	下	3	恣
6	兀	112	光机虺
7	与	15	與
8	上	35	卡
9	冂	70	同
10	乞	28	吃疙
11	亼	349	會
12	义	1	紂
13	亼	39	曷
14	凡	9	帆梵
15	勺	28	的芍
16	亡	40	望忙忘芒罔
17	丫	3	丫
18	凵	8	訊迅蝨螢
19	卩	4	卵
20	又	6	汊
21	刃	19	忍韌
22	壬	14	任
23	开	54	開形研筭
24	夭	39	奏吞昊
25	元	41	完玩頑芫
26	云	24	雲魂耘
27	巾	7	師
28	支	37	枝翅度芟頰
29	冫	3	冫
30	冫	2	冫
31	太	4	汰
32	反	19	飯返
33	友	2	芟
34	尤	28	就尢疣
35	宏	15	雄宏肱
36	厄	7	扼
37	匹	26	甚鳴
38	阜	2	馮
39	比	147	皆枇鹿庇茈
40	互	5	互

序號	部 件	組字數	例 字
41	切	4	砌
42	攴	17	敲
43	少	38	省沙匙
44	午	11	許迕
45	壬	51	聽望廷
46	升	7	昇陞
47	夭	45	笑妖喬
48	仁	3	佞
49	化	11	花貨靴囫
50	今	65	念琴黔
51	兮	4	盼諡
52	介	17	界疥赧玠
53	分	49	份貧芬頒氛
54	公	32	松翁頌
55	攴	47	駁
56	从	92	雜僉
57	欠	3	傘俎
58	凶	44	匈兇酗
59	勻	16	均
60	产	38	危
61	欠	118	次芟
62	仰	5	迎仰昂
63	殳	121	設疫莛
64	勾	3	苟枸
65	卞	5	汴
66	六	16	冥
67	方	152	於房防芳
68	亢	25	抗頑匠
69	斗	27	料魁辱罈
70	冗	2	冗
71	尺	6	呎咫
72	夊	2	隼
73	引	7	蚓
74	卂	2	飢
75	虫	4	蚩
76	艮	8	報
77	叉	10	蚤
78	刃	4	梁劓
79	刃	4	盈
80	又	3	沒

序號	部 件	組 字 數	例 字
81	允	35	吮
82	予	16	序預野
83	孔	3	吼乳
84	玉	11	璧鈺
85	北	11	背乖邶
86	正	24	政整征焉症
87	卉	24	奔賁
88	疋	42	是
89	尢	45	壺
90	卬	1	筇
91	凵	1	屆
92	功	2	昂
93	去	42	法丟劫盍
94	古	93	故苦克固姑
95	打	2	叮
96	札	3	紮
97	朮	18	術怵
98	巨	26	拒苴
99	可	64	河哥
100	匚	2	筐
101	匝	5	砸
102	左	18	佐
103	丕	16	坯莖邳
104	右	19	若佑
105	布	6	佈
106	夊	8	臬
107	夂	3	冪
108	平	18	評
109	夾	25	僚
110	弍	3	貳
111	充	5	統
112	卡	3	針
113	占	39	點店乜苦
114	尙	6	睿
115	氷	1	冰
116	旦	52	但量查疸
117	号	5	號梛
118	只	8	咫枳
119	另	2	別
120	兄	9	況
121	另	2	拐
122	另	2	拐

序號	部 件	組 字 數	例 字
123	囟	1	箇
124	尸	8	岸
125	冏	101	尚炯迴
126	囚	35	泗
127	四	6	泗
128	囧	2	卣
129	矢	82	知候彘
130	失	25	跌隄
131	龟	6	拖迤
132	仕	2	莅
133	付	23	府附苻
134	代	9	袋玳
135	佞	1	侃
136	他	2	憇
137	斥	8	訴
138	令	39	冷領零囟
139	尔	2	茶
140	告	25	船兌
141	仝	2	佻
142	夆	68	參珍疹
143	戾	3	碗
144	乏	11	泛寔
145	氐	27	低底邸
146	尔	3	你
147	勹	50	夠苟劬
148	匆	2	漣
149	冬	14	終疼苓蠱
150	囟	39	曷
151	卯	12	柳昴
152	夨	31	怨宛
153	包	33	跑電庖剝魚
154	主	25	住塵
155	市	4	鬧柿
156	玄	33	率弦銜茲
157	半	21	伴判
158	羊	65	南幸
159	汜	3	范
160	宁	8	貯苧
161	穴	103	空
162	宀	1	抗
163	它	19	蛇
164	司	11	詞筍

序號	部 件	組 字 數	例 字
165	廾	9	尉
166	尼	18	呢
167	弘	14	強泓
168	疋	41	蛋疑胥
169	出	29	屈拙崇茁
170	攴	1	攷
171	圣	2	怪
172	奴	12	努叟
173	加	21	架嘉咖痂
174	召	35	招超筲邵
175	弁	5	拚筭
176	台	35	治怠苔
177	阝	2	泐
178	幼	18	黝窈
179	丞	8	丞
180	匡	8	框筐劦
181	勹	23	契
182	𠂇	3	芟
183	丟	2	銛
184	式	8	試
185	舌	10	甜恬
186	刑	6	型荊劓
187	次	22	資茨伙
188	戎	10	賊
189	圭	55	街封鞋奎恚
190	寺	23	時等痔邽
191	吉	37	臺結頡髻
192	青	15	穀慤
193	老	19	姥耆
194	考	6	烤
195	戠	16	載
196	巩	10	恐筑
197	邛	3	筇
198	共	70	供異巷鬪
199	扣	2	箝
200	朴	2	焮
201	亘	24	宣垣
202	戌	46	歲
203	在	2	茬
204	丕	1	鉶
205	百	19	陌弼
206	存	8	荐洵

序號	部 件	組 字 數	例 字
207	夸	23	跨劓
208	有	19	賄郁宥囿
209	灰	6	恢盍
210	戍	24	幾
211	死	7	屍葬
212	列	17	烈例蒨
213	成	11	城盛寔邨
214	旨	12	指嘗
215	流	15	流醢
216	尗	29	叔戚
217	此	33	些雌齷髭
218	艾	4	哎
219	光	20	耀晃觥
220	吁	2	吁
221	早	62	草
222	囗	113	哭襄
223	同	34	興洞筒
224	吊	2	屛
225	因	16	姻恩茵胤
226	光	7	微
227	屮	2	龠
228	回	51	迴嗇蛔
229	朱	22	珠茱邾
230	先	35	洗筧
231	任	8	賃荏
232	休	10	咻庥髡傑
233	伏	5	袂茯
234	伐	5	閔筏
235	伊	4	咿
236	自	32	師歸追
237	血	13	恤卹卹
238	向	7	嚮餉
239	囟	12	惱傻
240	全	18	銓痊筌
241	合	62	給拿答鴿閣
242	后	11	垢逅屋邨
243	行	28	衛珩苻
244	尗	1	隳
245	辰	8	派
246	育	1	育
247	兕	12	嶮
248	允	2	筵

序號	部 件	組 字 數	例 字
249	兆	35	跳逃窺覘
250	受	25	受
251	杂	1	染
252	朵	6	躲刹
253	夙	2	夙
254	危	19	跪
255	旬	20	詢筍郇
256	夆	7	降
257	各	54	路客閣雒
258	匈	5	胸
259	舛	39	舞桀
260	名	8	銘茗
261	多	30	夠爹夥邊
262	兒	1	兜
263	色	7	艷
264	亦	5	跡奕
265	庀	1	侂
266	囟	24	离
267	交	29	校效茭
268	叀	47	族
269	亢	4	荒
270	并	33	瓶餅屏
271	关	5	送朕
272	𠂔	27	朔逆
273	爻	1	鍍
274	汙	3	滓益
275	江	3	鴻
276	守	2	狩
277	宅	7	詫
278	窆	2	挖
279	安	19	案按
280	字	2	牴
281	扌	3	刷
282	弜	1	弼
283	收	2	菽
284	如	14	恕茹洳
285	好	2	孳
286	荔	12	協脅荔颯
287	牟	9	眸麤
288	叒	8	桑
289	亼	23	參
290	纟	55	幾樂茲幽

序號	部 件	組 字 數	例 字
291	丞	7	拯登
292	夆	1	撈
293	弄	4	哢筭
294	𠂔	1	覘
295	戒	6	械
296	走	50	起徒
297	汞	2	鯨
298	赤	16	赫
299	均	4	筠
300	坎	3	炭
301	孝	9	教酵瘳
302	志	5	誌痣
303	声	9	聲
304	𠂔	4	輓
305	克	6	兢剋
306	扶	2	菹
307	技	2	廐
308	抄	2	菽
309	折	16	浙誓逝
310	杏	2	荅
311	孛	18	勃脖荸
312	吾	25	語衙寤啟
313	豆	124	頭短豈痘
314	医	12	醫嫫
315	辰	48	農振脣
316	否	5	詬痞
317	百	132	夏
318	巫	13	靈誣覘
319	夾	25	峽頰莢
320	尨	10	彪彪
321	夾	1	陝
322	忒	2	鉞
323	皀	5	陞
324	邪	2	琊
325	矣	12	疑
326	育	10	唷
327	壘	31	經頸莖氳
328	井	5	葬莽
329	步	23	歲涉頻
330	芒	4	茫鎡
331	延	16	延
332	𠂔	1	壑

序號	部 件	組 字 數	例 字
333	卣	2	迨
334	歺	7	餐
335	肖	40	消削屑宵
336	盱	2	泚
337	旱	13	趕桿
338	呈	1	捏
339	助	5	鋤勛
340	呆	8	保
341	昊	1	撰
342	晏	14	宴
343	旱	1	旱
344	粵	7	聘
345	足	22	捉蹙
346	男	5	舅甥甥
347	困	11	網閫
348	呈	18	程郢
349	呂	14	營鋁閩
350	吟	2	吟
351	別	2	捌
352	肩	18	捐鴉胃
353	吳	12	誤虞
354	邑	21	扈挹
355	罅	28	蜀
356	岑	6	涔
357	芬	1	榜
358	罔	18	兩
359	冏	34	裔
360	囧	2	箇
361	邦	4	綁
362	告	25	造靠酷窖
363	廷	20	庭挺霆
364	利	15	梨痢痢莉
365	禿	5	頹
366	秀	5	透誘莠
367	私	2	俵
368	每	22	海霉痲
369	何	2	荷
370	攸	26	條
371	作	3	笮
372	你	2	您
373	位	2	蒞
374	皂	1	鷓

序號	部 件	組 字 數	例 字
375	兇	3	貌
376	囟	12	窗
377	余	2	賒
378	余	36	除途荼畬
379	卮	1	梔
380	含	8	頷
381	役	2	徑
382	启	1	啟
383	兇	8	殺
384	希	19	稀邴
385	兌	19	說閱
386	谷	40	容卻俗
387	坐	17	座銜判
388	孚	8	埒號
389	妥	10	綏萎
390	孚	19	乳浮葦
391	免	38	晚兔勉
392	夆	17	蜂逢
393	狂	5	逛誑
394	狃	3	獄
395	狄	5	逖荻
396	角	55	解齋桷
397	犴	3	菴犴
398	彤	3	彤
399	彘	1	頰
400	卯	4	躄
401	系	15	縣
402	亨	4	烹烹
403	序	2	埤
404	辛	60	辭宰辣
405	忘	2	惹
406	忡	2	忡
407	快	2	篋
408	弟	15	梯剃梟
409	沐	3	霖
410	沛	2	霈
411	沙	12	莎痧娑杪
412	沃	2	塗
413	沂	2	沂
414	次	3	盜羨
415	沒	2	沒
416	沈	2	沈

序號	部 件	組 字 數	例 字
417	完	22	院寇莞
418	宋	2	涿
419	宀	1	涖
420	宏	4	鎡
421	牢	2	泮
422	良	22	娘葭闍
423	亘	1	疊
424	君	18	群窘裙焄
425	那	5	哪
426	即	14	節卿聖
427	尾	5	媿
428	局	8	跼
429	忌	4	踧
430	壯	6	裝莊
431	忍	4	認
432	甬	19	通勇桶
433	矣	10	唉欸
434	夊	27	酸痠悛皴
435	奉	8	捧奉
436	珏	10	班琴
437	武	8	賦鷓
438	青	35	情靜菁園
439	表	6	錶
440	毒	4	轟礪
441	忝	7	添
442	夂	5	替
443	卦	4	掛罨
444	者	48	著都諸闍
445	壺	23	陸逵
446	菱	21	陵菱
447	幸	45	報倅圍罨
448	其	40	期基旗箕
449	取	29	最趣聚諏聚
450	昔	27	借鵲曆
451	直	10	置值矗
452	拈	2	籍
453	拔	2	菝
454	抽	2	菹
455	拍	2	啪
456	抱	2	菹
457	林	51	楚森淋彬楸
458	拉	4	啦

序號	部 件	組 字 數	例 字
459	析	8	晰皙
460	松	8	鬆淞
461	尪	27	麻
462	或	27	國域惑
463	臥	2	臨
464	馭	16	緊
465	刺	2	薊
466	兩	7	輛颯
467	厓	6	捱
468	柰	4	捺
469	奔	3	湊
470	奄	20	淹菴庵
471	奇	30	寄騎
472	來	20	萊徠
473	爰	20	錢盞筴
474	到	5	倒
475	叔	17	督寂淑
476	芷	2	蒞
477	走	8	走
478	花	2	鉤
479	芬	2	芬
480	肯	4	啃
481	芳	2	芳
482	卓	15	掉罩遶
483	虎	36	號彪麓
484	犇	2	犇
485	尚	65	當倘趟敞
486	尋	2	得
487	昇	3	算
488	昊	2	溟
489	具	35	真俱颯
490	味	2	味
491	旻	1	闍
492	昆	17	混崑
493	昌	12	唱閭菖
494	昇	4	駢
495	明	4	盟萌
496	易	13	賜剔
497	胃	1	焗
498	旻	3	闍璣
499	畀	19	鼻泚
500	固	12	個痼齒

序號	部 件	組 字 數	例 字
501	响	2	响
502	岸	3	鏘
503	困	7	菌 稠
504	沓	7	踏
505	岡	9	剛 鋼 崗
506	罔	6	網 魴
507	卸	7	御
508	制	6	製 獬
509	知	7	智 蜘 痴
510	勑	5	勑
511	委	21	矮 魏 萎 痿
512	季	4	悻 瘴
513	申	17	叟
514	延	16	誕 筵 邲
515	佰	6	宿
516	版	2	蠟
517	兒	21	倪 霓 閱 郟
518	帛	7	幫 棉
519	的	2	蒟
520	阜	3	埠
521	舍	9	鋪 捨
522	所	6	質
523	欣	3	掀
524	論	32	論 崙 圖
525	陰	3	陰
526	念	11	念 斂
527	往	2	睚
528	屎	1	源
529	戾	8	淚
530	攷	2	祭
531	肩	6	捐
532	忝	1	愆
533	肴	6	淆 穀
534	戔	13	戔
535	效	25	爾 爽
536	采	13	菜 採 彩
537	呈	3	淫
538	受	4	授
539	爭	16	靜 箏
540	朋	22	崩 棚 鵬
541	服	5	蔽 鵬
542	周	28	調 週 雕

序號	部 件	組 字 數	例 字
543	昏	12	婚 闇
544	兔	22	冤 逸
545	匄	12	陶 萄
546	臼	18	陷 閭 欲 窳
547	备	1	違
548	咎	8	晷 絡
549	忽	3	惚
550	菊	13	鞠 菊 麴
551	朕	15	然
552	肥	5	肥
553	淵	6	淵
554	京	33	就 景 涼
555	享	41	敦 醇
556	尙	32	尙
557	店	4	惦
558	夜	5	液
559	府	4	腐 俯
560	底	2	麓
561	兗	1	莞
562	卒	31	碎 翠 瘁
563	庚	4	賡 鷓
564	放	45	敖
565	於	8	淤 菸 瘵
566	音	23	部 倍 菩
567	妾	9	接 霎
568	羌	4	蛻
569	卷	19	圈 倦 鬘
570	並	11	普 碰
571	炎	37	談 痰 剡
572	法	2	珞
573	河	2	荷
574	沾	3	霑
575	沮	2	菹
576	泊	2	箔
577	泓	2	霏
578	沼	2	沼
579	波	5	婆 菠
580	治	2	箔
581	宗	15	崇 粽
582	定	12	錠 楚
583	宕	2	茗
584	宜	3	誼

序號	部 件	組 字 數	例 字
585	官	15	管館筵
586	空	15	控筊
587	宛	23	碗剗苑
588	宓	6	密
589	冫	6	深
590	帚	8	婦
591	居	13	鋸
592	刷	3	涮
593	屈	10	掘窟
594	弦	4	詿
595	戕	6	臧
596	姑	2	菇
597	豸	20	綴戮
598	糸	17	綠剝氣盞
599	孟	5	猛
600	孤	3	菰
601	阿	6	啊痂
602	函	5	涵函
603	留	11	輜菑
604	契	11	楔痂
605	奏	6	湊
606	春	13	蠢椿
607	畱	6	插畝
608	咨	3	諮
609	堯	33	堯
610	封	6	幫封
611	背	5	背
612	政	2	最
613	壺	42	喜鼓
614	哉	3	澁
615	幸	9	達
616	某	6	謀
617	甚	24	堪勘甚
618	耶	4	爺椰
619	巷	3	港
620	葉	25	葉諜屨
621	故	2	做
622	胡	16	湖鬍鵠
623	南	11	喃
624	括	2	括
625	拾	2	浴
626	柰	3	隸

序號	部 件	組 字 數	例 字
627	柯	2	禽
628	相	12	想箱湘廂
629	查	6	渣
630	柱	2	崔
631	述	2	遞
632	勃	3	渤
633	軌	2	匪
634	轟	8	擊
635	軛	1	範
636	匱	13	堰郢
637	勅	1	鵠
638	刺	6	喇鬍痢
639	畱	22	福富副副
640	要	7	腰萋
641	堇	14	煙甄閨
642	東	27	練闌
643	咸	25	感鹹脣
644	威	5	歲
645	厘	3	喱
646	研	2	擘
647	頁	125	頭囂
648	奕	7	塹
649	奎	4	燻
650	昏	25	僚
651	亟	3	極
652	殄	2	饕
653	匱	2	愜篋
654	皆	14	階
655	致	2	緻
656	勁	2	勤
657	苦	3	楛
658	苛	2	澆
659	若	15	惹諾匿郡
660	苹	2	萍
661	苜	3	寬
662	苗	8	貓
663	英	11	瑛
664	苟	11	敬
665	苔	2	鎔
666	茅	2	鵠
667	貞	11	偵
668	虐	3	瘡謔

序號	部 件	組 字 數	例 字
669	省	4	省
670	削	4	箭
671	昊	8	闖
672	是	30	題提
673	眇	4	渺
674	則	8	測廁崩
675	易	38	陽瘍
676	段	23	假霞遐
677	鼎	1	鼎
678	聿	14	輯葺
679	冒	10	帽
680	星	9	醒戮
681	昨	2	詐
682	昫	2	煦
683	曷	39	喝葛歇
684	昱	2	煜
685	昭	2	照
686	畏	15	喂
687	胃	7	謂
688	虹	2	泓
689	思	23	慮鯢颶
690	馮	25	過禍窩副
691	号	11	鄂鏢号
692	品	59	區臨嶙
693	端	20	端端顛圖
694	炭	4	碳
695	春	3	騏
696	拜	2	湃
697	香	12	馨馥
698	秋	21	愁鞦菽
699	科	4	蝌蚪
700	复	15	復
701	段	8	緞
702	便	5	鞭
703	保	7	堡褒葆褓
704	皇	26	蝗凰篁
705	泉	7	線羸
706	侵	4	蓀
707	肥	1	葩
708	侯	13	猴篋儗
709	俊	2	蓀
710	盾	11	循遁

序號	部 件	組 字 數	例 字
711	奔	7	姘
712	俞	33	偷愈覩霰瘡
713	待	2	倂
714	衍	4	愆
715	律	3	律
716	扁	25	編篇翩匾
717	盆	3	溢
718	俎	2	蒞
719	卻	2	腳
720	爰	17	暖
721	冉	2	稱
722	夔	11	驥
723	匍	2	葡
724	風	34	飄瘋颶嵐
725	負	4	負
726	毘	14	毳
727	阜	1	惇
728	勉	2	勉
729	奐	7	換瘼
730	奎	3	隆
731	咎	5	糝
732	忽	1	漭
733	旬	6	鞫
734	昨	2	菲
735	胸	2	蒟
736	怨	2	慈
737	胤	2	胤
738	哀	2	依
739	亭	11	停
740	亮	3	哢
741	度	6	渡
742	彥	7	顏諺
743	席	1	蟪
744	施	8	梳施
745	游	4	遊游
746	亲	9	新
747	音	82	意暗響韻闇
748	帝	16	啼蒂啻
749	恬	2	活
750	恆	4	壻
751	美	5	鎂羹
752	酋	35	尊猶猷猷

序號	部 件	組 字 數	例 字
753	前	14	箭剪湔
754	豕	15	隊遂
755	茲	9	慈磁
756	洼	2	窪
757	洪	4	霏
758	流	3	鑿
759	洞	2	霜
760	洗	2	鑿
761	活	3	闊
762	派	3	蒹
763	洛	2	落
764	泝	2	滌
765	津	2	澗
766	宣	13	喧萱
767	突	5	淡
768	穿	2	琿
769	客	6	額喀
770	郎	15	鄉廊
771	軍	27	運揮暈黻
772	建	13	健毬
773	既	10	概暨旣
774	展	5	殿
775	屋	9	握
776	屏	4	摒
777	弭	4	麋
778	昏	4	湣
779	啟	3	愨
780	韋	44	衛偉葦韜
781	眉	10	湄鄙
782	胥	11	婿
783	皆	3	繫
784	姦	2	姦
785	盈	3	楹
786	勇	3	湧
787	癸	13	葵揆闕戮
788	發	12	發
789	柔	17	揉
790	攷	13	務駑
791	彖	20	緣篆
792	孛	6	孛
793	紅	2	葑
794	約	4	啣

序號	部 件	組 字 數	例 字
795	愨	2	愨
796	秦	3	漆
797	秦	11	臻綦
798	素	8	愨
799	菁	16	講邁斟菁
800	祿	1	蒜
801	恣	2	懿
802	貢	10	贛楨
803	穀	15	穀愨
804	袁	11	園猿袁
805	耆	7	嗜著
806	盍	21	蓋盍
807	耿	2	繫
808	菡	4	備
809	恭	2	恭
810	真	29	鎮顛闐寘
811	軌	11	幹
812	配	2	熙
813	捕	2	捕
814	桂	2	漑
815	索	4	嗦
816	軒	3	蕲
817	書	2	轡
818	專	21	博
819	哥	3	歌
820	栗	9	漂
821	敕	1	菽
822	匪	3	框筐
823	辱	11	褥蓐辱
824	差	1	陸
825	威	2	滅
826	夏	4	廈廈
827	盍	1	醢
828	原	14	願源愿
829	晉	7	緝戩
830	溼	1	溼
831	髟	47	髮
832	莧	2	萎
833	艸	4	楫
834	柴	3	噤
835	草	2	驛
836	荃	2	醛

序號	部 件	組 字 數	例 字
837	茶	3	搽
838	荅	10	搭 瘡
839	荀	2	楸
840	荒	3	慌
841	虔	3	鯪
842	犖	9	叢
843	尙	2	隙 號
844	貞	3	鎖
845	党	2	鏡
846	時	6	蒔 鮒
847	睪	19	瞿
848	晃	6	幌
849	閃	2	網
850	晏	3	鸚
851	翮	12	塌 遏 闌
852	畱	9	僵
853	𠂔	3	遣
854	𠂔	3	稷
855	畔	2	泮
856	骨	41	體 滑
857	員	14	圓 損 鄖 質
858	恩	3	摠
859	盍	2	醜
860	囷	3	溷 恩
861	豈	19	凱 愷 闐
862	敝	7	微
863	𠂔	2	堙
864	冐	11	鰓 瘰
865	盪	21	溫 瘟
866	缺	2	缺
867	氣	6	焮
868	乘	5	剩 嶮
869	租	2	菹
870	秝	15	麻
871	倩	3	箒
872	叟	17	瘦 嫂 鄆
873	修	2	鏹
874	倫	2	淪
875	隼	4	準
876	隻	3	護
877	倍	2	蓓
878	臬	8	鏹 艱 闌

序號	部 件	組 字 數	例 字
879	臭	6	嗅
880	射	4	謝 麝
881	臬	8	皦 翹
882	躬	4	窮
883	息	9	媳 憩 鄔
884	島	2	搗
885	師	7	獅 篩
886	追	6	槌
887	毘	6	篋 媿
888	倉	22	槍 創 蒼 瘡
889	徒	2	徒
890	虎	15	遞 麓
891	般	3	慇
892	扌	1	肇
893	扇	5	煽
894	瓜	3	窳
895	般	14	盤 搬 癩
896	翁	10	喻 翁
897	釜	2	滄
898	爹	2	爹
899	𠂔	7	隳
900	𠂔	11	稻
901	奚	17	雞 溪
902	鬱	4	鬱
903	桀	7	傑
904	𠂔	15	搖 遙 颯
905	脈	2	脈
906	留	19	溜 瘤 畱 鷗
907	芻	14	趨 皺 筭
908	衰	6	蓑 棧
909	高	30	敲 稿 嵩
910	席	3	蓆
911	庫	2	庫
912	庭	2	庭
913	蚤	2	缺
914	疾	5	嫉 蒺
915	效	2	傲
916	痂	2	癩
917	唐	16	糖
918	旅	2	贅
919	竟	1	競
920	均	2	蒟

序號	部 件	組 字 數	例 字
921	旁	19	磅 旁
922	畜	6	蓄 愾 鄙
923	差	19	搓 瘥
924	羔	8	糕 羔 羹
925	料	2	竊
926	迷	4	謎
927	益	14	溢 躄
928	朔	9	塑 溯
929	歛	13	厥 闕
930	逆	2	逆
931	脊	6	瘠 鵠
932	剡	4	屬
933	浦	2	蒲
934	涉	2	瀕
935	涅	2	滷
936	涅	2	涅
937	洳	1	蒞
938	海	2	晦
939	涂	3	塗
940	浪	2	蒞
941	浸	2	寢
942	寒	16	賽
943	家	6	嫁
944	害	12	割 瞎
945	容	13	溶 蓉
946	窄	4	榨
947	叢	2	濃
948	宰	3	滓
949	冢	15	蒙
950	朗	2	朧
951	冢	2	塚
952	冥	15	螟 蒙
953	冤	2	菟
954	盜	2	謚
955	退	4	腿
956	展	7	碾
957	屑	2	楮
958	屮	2	諱
959	弱	8	溺
960	蚩	4	嗤
961	挈	2	鞞
962	蚤	10	騷

序號	部 件	組 字 數	例 字
963	桑	8	噪 穎
964	脅	5	憎
965	能	12	罷 態 襪
966	孫	7	遜 蓀 獠
967	陟	2	鷹
968	除	4	滁 蔕
969	紛	2	紛
970	紡	2	驚
971	節	3	關 聯
972	邕	8	離 噤
973	烝	3	蒸
974	慧	8	慧 擘 篳
975	春	4	椿 忝
976	責	16	積 勳 簣
977	𦵏	8	釐
978	規	8	窺 闕 規 槲
979	匿	4	匿 暱
980	焉	7	媽 薦 鄔
981	赦	2	螫
982	頂	2	嶺
983	都	2	啞
984	執	12	熱
985	赦	22	傲 熬 廩
986	執	11	墊
987	殷	9	聲
988	基	2	筌
989	聊	2	踰
990	勘	3	礪
991	勒	2	鱗
992	滿	11	滿 顛
993	乾	3	墘
994	措	2	籍
995	推	3	摧
996	埜	2	溼
997	彬	3	彪
998	梵	2	漑
999	黍	3	漆
1000	麥	11	麵 麥
1001	救	2	慙
1002	斬	15	漸 暫 斬
1003	軟	2	簌
1004	連	16	蓮 鏈

序號	部 件	組 字 數	例 字
1005	專	24	傳團專專
1006	專	2	敷
1007	敕	5	整嫩邀
1008	欸	7	漱簌
1009	區	28	歐驅奩
1010	堅	5	慳
1011	粟	33	標飄癩
1012	覓	6	遷
1013	醜	2	醜
1014	毆	11	醫
1015	慝	1	嫫
1016	脣	3	滑
1017	戚	10	蹙鍼
1018	戛	2	嘎
1019	奢	4	捨
1020	奩	2	奪
1021	瓠	3	瓠
1022	爽	8	塽
1023	逐	4	筵
1024	零	7	檣鄂
1025	雪	2	鱈
1026	頃	8	傾穎廩
1027	莽	3	蟒
1028	莫	33	幕模寞鄭
1029	茶	3	琛
1030	莖	3	蓮
1031	葶	1	葶
1032	鹵	1	竊
1033	鹵	9	鹽鹹滷
1034	冢	1	敕
1035	盧	5	謔
1036	膚	22	膚
1037	滹	7	滹
1038	彪	2	滌
1039	雀	3	巖
1040	堂	10	腔
1041	常	2	嫦
1042	野	2	墅
1043	晨	2	鷗
1044	昧	2	蘇
1045	曼	19	慢蔓鄴
1046	異	15	戴翼

序號	部 件	組 字 數	例 字
1047	累	12	螺療藟
1048	國	11	愜
1049	患	3	漣
1050	唯	5	雖鷲
1051	曷	2	圖鄙
1052	眾	2	澤
1053	崔	24	摧
1054	崩	8	蹦
1055	圈	2	菌
1056	造	4	糙筵
1057	移	4	移
1058	動	2	慟
1059	笙	1	槿
1060	笠	2	鷓
1061	條	7	滌篠
1062	脩	4	滄
1063	臬	4	鄴
1064	鳥	237	鴉鷹鳶鴛
1065	參	1	穆
1066	恩	10	總蔥
1067	叟	2	傻
1068	敏	5	繁
1069	倉	3	倉
1070	貪	3	貪
1071	徙	8	屣蓰
1072	御	5	禦
1073	從	18	縱聳蓰
1074	扈	5	滄
1075	敘	2	澈
1076	殺	5	鍍
1077	悉	4	蟋窻
1078	欲	3	慾
1079	覓	2	覓
1080	匐	2	匐
1081	魚	174	魯漁鮮鯊
1082	逢	6	縫蓬
1083	猗	2	漪
1084	斛	3	榭
1085	豚	2	遯
1086	祭	15	察際療
1087	許	2	滄
1088	設	2	設

序號	部 件	組 字 數	例 字
1089	孰	3	熟
1090	郭	7	廓 擲
1091	庶	10	遮 蔗 蹠 鷓
1092	麻	25	麼 嘛
1093	庫	3	廬
1094	產	8	鏟 剷
1095	离	24	離 璃 禽 鸕
1096	康	6	慷
1097	鹿	67	麗 麈 麟 轆
1098	袞	3	滾
1099	族	7	簇 族 瘖
1100	旋	7	漩 旋
1101	章	25	障 贛 漳
1102	竟	6	鏡
1103	部	3	蔀
1104	彖	2	毅
1105	商	15	適 敵 摘
1106	商	5	熵
1107	率	5	摔
1108	牽	2	絳
1109	惟	2	罹
1110	羞	2	饑
1111	美	3	樣
1112	敝	21	幣 蔽 撇
1113	焊	2	焯
1114	凌	2	菱
1115	淑	2	菽
1116	淮	2	匯
1117	淫	2	靈
1118	淵	2	淵
1119	梁	2	樑
1120	寇	4	寇
1121	寅	10	演 贛 餞
1122	寂	2	寂
1123	宿	5	縮 蓓
1124	室	2	漶
1125	室	3	螻
1126	密	4	密
1127	焄	2	蒸
1128	尉	9	慰 蔚
1129	屠	4	廬
1130	扁	2	漏 癩

序號	部 件	組 字 數	例 字
1131	張	2	漲
1132	強	9	襁
1133	疏	2	蔬
1134	將	11	獎 蔣 鏘
1135	通	3	蓬
1136	習	13	摺 翫
1137	麥	30	膠 寥 戮 瘳
1138	翌	2	燿
1139	務	3	霧
1140	參	22	慘 毳 蓼
1141	貫	4	實 慣
1142	陳	5	陳
1143	陰	3	蔭
1144	終	2	蔘
1145	巢	11	巢 剿
1146	貳	3	膩
1147	絜	3	潔
1148	替	2	僭
1149	款	2	窾
1150	堯	32	燒 翹 堯 遶
1151	項	2	頌
1152	越	2	越
1153	敢	21	橄 愍 闕
1154	賁	19	憤 賁
1155	博	2	簿
1156	封	7	樹 廚
1157	喜	14	嘻 詬 熹
1158	彭	6	澎 髡
1159	報	2	報
1160	壹	8	懿 噎
1161	惡	2	噁
1162	斯	7	撕 廝
1163	欺	2	傲
1164	黃	34	廣 橫 簧 甞
1165	散	5	撒 霰
1166	貫	2	勸
1167	戟	3	檄
1168	朝	4	潮 廟
1169	辜	3	棒
1170	楸	5	樊
1171	惠	10	穗 蕙
1172	覃	22	潭 葷

序號	部 件	組 字 數	例 字
1173	粟	3	傑
1174	棘	4	棘 襟
1175	陌	4	爽
1176	忝	11	憂
1177	麻	14	歷
1178	寮	25	寮 遼 瞭 鷄
1179	厥	12	厥 蕨
1180	廁	3	廁
1181	欸	2	愁
1182	焱	3	颯
1183	莖	1	莖
1184	雲	7	曇 黓 澐
1185	雅	2	雅
1186	晉	11	蠶 潛 簪
1187	華	13	嘩
1188	著	2	著
1189	寬	3	寬
1190	萑	41	萑
1191	萃	2	萃
1192	虛	8	虛 覷
1193	黻	4	黻
1194	撲	15	撲
1195	敝	3	廠 斃
1196	棠	4	棠
1197	撐	3	撐
1198	撮	5	撮 蕞
1199	糧	2	糧
1200	潤	4	潤
1201	嫻	4	嫻
1202	簡	4	簡 澗
1203	閒	9	痲 儻 蘭
1204	憫	5	憫
1205	悶	3	悶
1206	唇	2	唇
1207	厥	11	厥
1208	景	9	影 璟
1209	貴	27	遺 潰 蕘
1210	蟲	18	蟲
1211	器	10	器 罍
1212	癌	2	癌
1213	選	1	選
1214	激	1	激

序號	部 件	組 字 數	例 字
1215	賈	23	賈
1216	罾	3	罾
1217	圍	2	圍
1218	撫	15	撫 廡 甌
1219	智	2	智
1220	毳	5	毳 竈
1221	煖	21	贊
1222	稍	3	稍
1223	黍	7	黎 黏
1224	黎	1	藜
1225	稭	2	稭
1226	喬	28	橋 蕎 屨 敵
1227	筑	3	築
1228	策	3	錄
1229	答	4	劄
1230	箎	2	箎
1231	筍	2	筍
1232	筆	2	筆
1233	備	2	備
1234	牌	2	牌
1235	貸	2	贖
1236	集	4	礁
1237	焦	21	瞧 蕉 鷓
1238	進	3	遲
1239	烏	1	鴉
1240	烏	5	寫
1241	臬	1	臬
1242	皓	2	皓
1243	翕	4	歛 滄
1244	復	5	覆 履
1245	須	4	鬚 嬰
1246	雇	3	顧 僱
1247	鈞	2	鈞
1248	欽	4	欽
1249	番	35	翻 播 蕃
1250	舜	3	瞬
1251	鬻	4	亂
1252	象	9	像
1253	猶	2	猶
1254	然	4	燃
1255	飧	2	餐
1256	賈	2	鄧

序號	部 件	組 字 數	例 字
1257	就	7	蹴驚
1258	敦	13	燉慙
1259	斌	2	贇
1260	疾	2	瘞
1261	袞	2	袞
1262	童	20	鐘量
1263	惰	2	惰
1264	善	8	膳鄩
1265	黍	2	黍
1266	普	5	譜
1267	彘	25	憐鄰遴
1268	尊	16	遵蹲
1269	奠	4	鄭
1270	馮	2	憑
1271	曾	20	增層甞
1272	焱	2	焱
1273	煒	1	檣
1274	勞	12	撈癆
1275	湛	2	霑
1276	渠	4	螺葉
1277	湯	6	盪蕩
1278	渝	2	滄
1279	寧	14	寧
1280	蕊	4	蕊榮
1281	尋	13	潯鄩
1282	畫	5	劃燼
1283	逮	2	逮
1284	肅	15	繡蕭
1285	犀	6	遲墀
1286	屨	6	滌
1287	費	4	費
1288	巽	11	選撰筭
1289	𤝵	3	𤝵
1290	絮	2	絮
1291	登	19	燈凳鄧登
1292	發	11	廢潑
1293	喬	19	橘鷓適
1294	隋	7	墮橢
1295	陽	2	錫
1296	隆	3	窿癩
1297	隊	3	墜
1298	隊	1	墜

序號	部 件	組 字 數	例 字
1299	鄉	8	響鄉
1300	結	2	結
1301	絕	2	絕
1302	絲	24	驚響
1303	幾	13	機
1304	瑟	4	瑟
1305	資	4	資
1306	戴	2	鐵
1307	鼓	12	鞞鞞
1308	達	9	鞞聞
1309	聖	3	聖
1310	勤	2	勤
1311	幹	3	幹
1312	構	2	構
1313	摺	1	摺
1314	禁	5	襟
1315	楚	5	礎
1316	楊	2	鶯
1317	楸	2	愁
1318	毆	8	擊
1319	薑	9	僵薑
1320	剝	2	剝
1321	賈	3	價
1322	遠	3	隨
1323	感	5	憾
1324	嗇	10	牆蓄
1325	殄	1	薤
1326	殮	1	養
1327	雷	7	鐺
1328	零	3	零
1329	頓	2	頓
1330	肆	2	肆
1331	葉	4	蝶
1332	歲	12	穢歲
1333	葬	2	駢
1334	背	2	嘴
1335	莠	1	蕩
1336	萬	19	厲邁蕘萬
1337	葛	8	藹藹
1338	董	3	懂
1339	敬	10	驚傲
1340	葱	1	葱

序號	部 件	組 字 數	例 字
1341	蔦	2	偽
1342	祭	4	燦
1343	虐	2	戲
1344	虞	15	據劇
1345	虞	4	嘆
1346	虜	4	擄
1347	業	4	鄴
1348	當	14	擋簷
1349	睢	2	淮
1350	賊	2	蟹
1351	遐	2	遼
1352	戢	5	蔽戢
1353	歇	2	蠍
1354	號	2	饗
1355	路	7	露鷺潞
1356	豐	6	體
1357	農	14	濃
1358	過	8	搨邁
1359	臬	23	操
1360	署	4	薯曙
1361	罽	21	譯圍敦
1362	叢	21	環園鬢翹
1363	罍	1	嫖
1364	罪	3	擇
1365	蜀	28	獨屬歌
1366	嵬	2	巍
1367	雉	2	薙
1368	稜	2	稜
1369	愁	2	憊
1370	筮	4	噬
1371	節	6	櫛癩
1372	鼠	26	竄鼯
1373	雋	6	鑄雋
1374	毀	4	燬
1375	裊	2	嫵
1376	敦	21	激邀窳
1377	奧	11	澳畷
1378	僉	28	臉劍簽
1379	會	22	繪郇蒼
1380	愈	2	癒
1381	薇	5	薇
1382	禽	4	擒

序號	部 件	組 字 數	例 字
1383	愛	10	瓊菱
1384	豸	2	墜
1385	亂	2	亂
1386	詹	19	擔簷
1387	劍	1	薊
1388	解	11	蟹懈癖薛
1389	煞	2	煞
1390	亶	23	檀氈遄
1391	凜	6	凜凜
1392	雁	3	應
1393	廉	8	簾鎌
1394	薦	4	薦
1395	新	2	薪
1396	戢	11	識
1397	歆	2	嗽
1398	意	13	憶蕙
1399	羸	9	羸
1400	雍	10	擁壅
1401	義	11	議
1402	羨	2	羨
1403	道	3	導
1404	遂	12	隧
1405	慈	3	滋
1406	煩	2	煩
1407	溥	5	簿
1408	溫	2	溘
1409	溪	2	瀉
1410	漉	2	瀉
1411	塞	2	塞
1412	運	2	璉
1413	殿	5	澱腎癩
1414	辟	29	壁避闢擗薛
1415	贖	1	贖
1416	預	3	瀕預
1417	遠	2	遠
1418	臺	6	臺檯
1419	赫	2	嚇
1420	壽	20	籌壽燾翹
1421	綦	3	綦
1422	聚	4	驟鄴
1423	幹	2	澣
1424	憲	3	愼

序號	部 件	組 字 數	例 字
1425	輕	2	鑿
1426	匱	2	櫃
1427	監	18	鹽 艦 檻
1428	甄	3	甄
1429	嚴	9	嚴
1430	厭	10	壓 蹇
1431	奩	2	奩
1432	爾	17	彌 邇 璽 爾
1433	臧	5	藏 臧
1434	豨	3	豨 豨
1435	需	24	儒
1436	疑	11	礙 疑
1437	蓋	5	溢
1438	尊	2	磚
1439	蓐	2	蓐
1440	幕	3	冪
1441	夢	3	儻
1442	萑	15	護
1443	蒿	2	噶
1444	蒙	14	矇
1445	睿	4	濬 叡
1446	壑	1	壑
1447	虐	1	虧
1448	對	6	懟
1449	嘗	3	鱮
1450	睪	16	嬰
1451	睽	2	葵
1452	閣	2	擱
1453	焜	7	顯 濕
1454	遣	3	譴
1455	團	2	糰
1456	還	2	嚮
1457	罰	2	蒟
1458	舞	2	躡
1459	藉	4	籍
1460	稭	1	藹
1461	熙	1	熙
1462	箸	2	櫛
1463	箕	2	簸
1464	算	2	匱
1465	與	14	舉 嶼 歟 蕝
1466	僕	3	濮

序號	部 件	組 字 數	例 字
1467	鼻	15	鼾 鼻
1468	魁	2	櫛
1469	銛	2	慙
1470	嶺	2	嶺
1471	歛	2	撇
1472	銀	2	椶
1473	慙	6	穩
1474	狸	3	靈
1475	貌	3	藐 邈
1476	獄	3	嶽 鷺
1477	詰	1	讒
1478	歛	2	歛
1479	豪	8	壕
1480	麼	2	嫵
1481	旖	2	旖
1482	辨	5	辦
1483	榮	5	嶸
1484	榮	2	滌
1485	熒	2	熒
1486	滿	3	懣
1487	漸	2	灑
1488	漂	2	藻
1489	滯	2	懣
1490	拳	2	拳
1491	賓	17	濱 鬢
1492	窳	2	嗜
1493	察	3	擦
1494	寧	13	惺
1495	盡	10	儘 蓋
1496	頗	2	櫛
1497	澀	2	澀
1498	熊	2	熊
1499	翟	18	耀 戮 籊
1500	翠	4	澤
1501	瞿	2	霏
1502	維	12	羅 灘
1503	綿	2	睇
1504	齏	3	斷 繼
1505	黜	2	黜
1506	慧	2	慧
1507	爽	2	襖
1508	輦	3	攆

序號	部 件	組 字 數	例 字
1509	頤	4	頤
1510	賣	20	讀竇覲
1511	穀	2	澌
1512	樞	2	菴
1513	樊	5	攀
1514	敷	2	璫
1515	遯	1	遯
1516	賢	2	磳
1517	遷	4	韃
1518	厲	10	勵蠣
1519	磊	2	荔
1520	憂	11	優鄴
1521	鴈	2	屨
1522	豬	4	潞
1523	蓺	5	藝
1524	菴	2	樞
1525	蔑	9	襪
1526	蓬	2	鏗
1527	蔡	3	擦
1528	慮	7	濾
1529	戩	2	蠶巖
1530	賞	2	償
1531	暴	11	爆
1532	賤	2	濺
1533	賜	2	飭
1534	閭	2	欄
1535	數	4	藪擻
1536	畠	22	壘儷畧
1537	罽	1	獸
1538	罷	6	擺
1539	墨	4	嚳
1540	耦	2	藕
1541	黎	3	藜
1542	箴	2	鱗
1543	鼻	4	邊
1544	樂	14	藥爍
1545	質	6	蹟
1546	徵	4	懲癥
1547	衛	3	衛
1548	盤	2	盪
1549	魯	6	魯
1550	夔	6	瓊

序號	部 件	組 字 數	例 字
1551	滕	3	藤
1552	劉	5	瀏
1553	諸	5	儲
1554	廚	4	櫛
1555	廣	14	礦
1556	摩	2	摩
1557	廛	4	纏
1558	痲	2	菴
1559	鹿	10	鑣
1560	慶	2	癯
1561	廢	2	廢
1562	毅	2	藪
1563	適	4	適
1564	養	5	養養
1565	頰	4	頰
1566	鄭	3	擲
1567	翦	2	翹
1568	瑩	2	澄
1569	潭	2	潭
1570	潘	3	藩
1571	寮	2	屎
1572	寬	3	腕
1573	寫	2	瀉
1574	審	5	審
1575	窮	2	窮
1576	慰	2	慰
1577	隕	3	隕
1578	緣	2	緣
1579	鼠	10	獵選鬣
1580	冀	3	驥
1581	穀	2	澌
1582	燕	10	嚳
1583	翰	2	瀚
1584	擇	3	擇
1585	賴	8	懶癩籟
1586	噩	5	鱷
1587	融	2	融
1588	歷	11	壑廔
1589	霍	9	藿臃
1590	頻	4	頻顰頻
1591	萑	2	鄣
1592	菅	2	管

序號	部 件	組 字 數	例 字
1593	蕭	4	瀟
1594	壑	2	漑
1595	廌	8	獻
1596	盧	21	爐 蘆 廬 顱
1597	虬	2	贊
1598	懸	3	懸 轟
1599	閻	3	蘭
1600	閻	4	爛
1601	遺	3	墮
1602	異	1	嫖
1603	園	2	摑
1604	積	2	癢
1605	頽	2	瀕
1606	錐	7	雙 錐
1607	錢	2	錢
1608	錄	2	錄
1609	禦	2	藥
1610	衡	2	衡
1611	穌	4	蘇
1612	謁	2	靄
1613	裏	7	壞
1614	磨	2	磨
1615	親	7	襯
1616	羸	3	羸
1617	齶	1	齶
1618	滌	3	滌
1619	澡	2	藻
1620	褰	2	攘
1621	憲	4	愷
1622	疆	2	疆
1623	鬻	1	鬻
1624	辭	3	繫
1625	隨	2	濇
1626	緼	2	蘊
1627	戴	3	襪
1628	盪	2	盪
1629	聯	2	聯
1630	葑	2	葑
1631	鞠	3	鞠
1632	翳	3	蠟
1633	羣	3	霸
1634	霜	5	霜

序號	部 件	組 字 數	例 字
1635	靈	8	靈 橋 鄗
1636	薛	4	孽
1637	薦	3	韜
1638	薄	3	樽
1639	戲	2	戲
1640	遽	3	遽
1641	嬰	12	櫻 鸚 瘦
1642	闌	15	蘭 爛
1643	灑	3	灑
1644	舉	2	擲
1645	繁	3	潔 繁
1646	鍾	2	鐘
1647	龠	11	鑰 龠 籥
1648	斂	3	澈 斂
1649	籤	13	織 籤
1650	爵	7	嚼
1651	龜	14	纒 劓
1652	鮮	4	癬 鮮
1653	變	2	躑
1654	襄	28	讓 襄 勳
1655	糜	2	醪
1656	麋	3	麋
1657	糞	2	糞
1658	營	3	嶸
1659	蹇	3	捷
1660	榮	2	藥
1661	彌	4	彌
1662	牆	2	牆
1663	翼	2	翼
1664	隱	4	癮
1665	豐	8	艷 豐
1666	贅	2	贅
1667	燾	2	燾
1668	聶	9	攝
1669	職	2	職
1670	藪	1	藪
1671	轉	2	轉
1672	鬻	2	鬻
1673	藍	2	濫
1674	藏	3	臟
1675	萑	24	歡 權
1676	藉	2	躑

序號	部 件	組 字 數	例 字
1677	叢	2	叢
1678	瞿	19	懼 衢 矍 毘
1679	壘	4	壘
1680	蟲	3	蠱
1681	嵩	11	攜 鄴
1682	魏	3	魏
1683	雙	3	艘
1684	歸	3	巋
1685	雜	2	噉
1686	犇	1	羸
1687	竄	3	擯
1688	斷	2	斷
1689	離	4	癱 灘
1690	壞	2	壞
1691	難	7	灘 癱 難
1692	顛	3	巔 癩
1693	攀	2	攀
1694	繫	2	繫
1695	醮	2	蘸
1696	麗	16	曬 麗 邈
1697	嬰	2	嬰
1698	蝨	1	蝨
1699	藝	3	藝
1700	蠱	2	蠱
1701	繭	2	繭
1702	壇	2	壇
1703	羅	10	羅 鑼 邏
1704	贊	20	鑽 贊 鄴
1705	儻	2	儻
1706	邊	2	邊
1707	繚	20	變
1708	靡	6	靡 靡
1709	離	5	籬
1710	壑	1	壑
1711	懷	2	懷
1712	類	3	類
1713	瀝	2	瀝
1714	躡	2	躡
1715	斡	2	斡
1716	蘇	2	蘇
1717	孽	2	孽

序號	部 件	組 字 數	例 字
1718	獻	6	獻
1719	穀	1	鑿
1720	黨	10	儻
1721	矍	9	攫
1722	闕	2	闕
1723	闡	2	瀾
1724	瞿	1	瞿
1725	嚴	9	儼
1726	纂	2	攥
1727	豐	3	臺
1728	覺	3	攪
1729	羸	2	羸
1730	糴	2	糴
1731	澱	2	澱
1732	寶	2	贖
1733	齧	2	齧
1734	覽	6	攬
1735	霸	3	壩
1736	蘭	4	欄
1737	鼻	2	屨
1738	顛	2	灑
1739	纍	4	纍 標
1740	嚙	1	嚙
1741	夔	3	夔
1742	屬	9	囑
1743	灌	1	灌
1744	蠢	4	蠢
1745	聽	2	廳
1746	囊	5	囊
1747	彎	2	灣
1748	顯	2	灑
1749	籟	2	籟
1750	樂	3	灤
1751	靈	3	檯
1752	贛	3	慈
1753	鼈	1	鼈
1754	蒙	2	蒙
1755	癢	2	癢
1756	豔	2	豔
1757	鬱	3	鬱

表 2.4：為其他字書單字的五大字集成字合成部件

序號	部件	組字數	例字
1758	傘	9	舍
1759	𠂔	42	卷
1760	高	55	高
1761	声	73	慶
1762	汧	1	梁
1763	宀	6	寢
1764	廾	1	虞
1765	朕	11	勝

序號	部件	組字數	例字
1766	絲	7	濕
1767	畧	1	摺
1768	惠	3	聽
1769	害	4	憲
1770	穀	1	澌
1771	陸	2	陸
1772	斲	2	斲

表 2.5：在 GF3001 中的五大字集非字合成部件

序號	部件	組字數	例字
1773	丂	1	寡
1774	𠂔	2	齋 齏
1775	勿	8	黎
1776	夂	5	夜

序號	部件	組字數	例字
1777	𠂔	5	那
1778	𠂔	26	留 劉
1779	𠂔	11	盡
1780	𠂔	42	會

表 2.6：具有特定構意的五大字集非字合成部件

序號	部件	組字數	例字
1781	司	11	司
1782	才	10	在
1783	冂	20	蘭
1784	冂	19	向 粵
1785	凡	40	風
1786	𠂔	25	索
1787	𠂔	25	希
1788	冂	1	沉
1789	冂	8	局
1790	𠂔	12	吳
1791	𠂔	1	壺
1792	𠂔	30	貴
1793	冂	15	岡
1794	冂	28	周
1795	冂	8	弱
1796	冂	2	疏
1797	𠂔	3	畚
1798	舌	21	話 刮 颯
1799	冂	15	復

序號	部件	組字數	例字
1800	𠂔	2	旅
1801	冂	1	梅
1802	冂	7	眾
1803	𠂔	3	佩
1804	𠂔	13	奐
1805	𠂔	15	產
1806	𠂔	19	郎
1807	𠂔	5	屨
1808	𠂔	9	囊
1809	𠂔	1	鍔
1810	𠂔	2	𠂔
1811	𠂔	1	𠂔
1812	𠂔	184	路
1813	𠂔	2	琰
1814	𠂔	20	度
1815	𠂔	21	差
1816	𠂔	14	侵 寢
1817	𠂔	2	款
1818	𠂔	14	華

序號	部 件	組字數	例 字
1819	青	5	散
1820	卓	22	乾
1821	具	4	篡
1822	豈	1	激
1823	泉	15	原
1824	疋	18	從
1825	商	5	商
1826	尸	15	屬
1827	尪	11	將
1828	陟	7	隋
1829	殸	1	澌
1830	渠	4	渠
1831	育	8	情
1832	男	2	愣
1833	尢	2	稽
1834	皇	5	毀
1835	囟	3	腦
1836	與	16	與
1837	齊	2	儕

序號	部 件	組字數	例 字
1838	勞	45	勞
1839	幾	1	畿
1840	橐	2	橐
1841	撤	4	撤
1842	莖	9	夢
1843	莖	1	薨
1844	對	6	對
1845	傷	6	傷
1846	徵	7	徵
1847	燕	11	燕
1848	潛	1	潛
1849	鼉	2	鼉
1850	齊	1	齋
1851	學	14	學
1852	覽	7	覽
1853	篡	4	篡
1854	囊	5	囊
1855	釁	4	釁
1856	壘	4	壘

2.5 簡化字集部件表

簡化字集指的是《簡化字總表》收錄的 2,235 個簡化字。依照漢字構形資料庫的部件拆分原則，簡化字集拆分後的部件總數為 1,122 個，其中基礎部件為 367 個，合成部件為 755 個，詳述如後。

2.5.1 簡化字集基礎部件表

簡化字集基礎部件共有 367 個，扣除和五大字集基礎部件重複的 326 個外，共有 41 個，詳見表 2.7。表 2.7 的排序是依照部件的筆畫數遞增式排序。若為同筆畫之部件，則依照筆順，按筆畫順序橫（一）豎（丨）撇（丿）點（丶）折（乚）進行排序。

表 2.7：簡化字集基礎部件表

序號	部件	例 字
1	丨	竖
2	丿	师
3	川	啸
4	讠	计狱
5	灬	泼
6	丷	劲
7	弓	马
8	纟	丝
9	火	烧
10	亻	壮寝
11	饣	饭
12	乌	乌
13	门	们问
14	讠	学应
15	灬	农
16	ヨ	当寻归
17	勹	场痲筋
18	飞	飞
19	习	习
20	纟	级辨辮
21	寿	寿

序號	部件	例 字
22	韦	伟围苇韬
23	专	传
24	巾	带
25	车	军库轮阵
26	小	变
27	内	锅纳
28	贝	贪则坝赢
29	见	现觅魁
30	长	张
31	鸟	岛
32	为	伪
33	书	书
34	戈	栈笈盖划
35	龙	拢宠垄庞
36	东	冻崇鸫
37	东	练
38	钅	衔针
39	乐	烁
40	甲	单
41	两	满俩魍

2.5.2 簡化字集合成部件表

簡化字集合成部件共有 755 個，扣除和五大字集合成部件重複的 530 個外，共有 225 個。這 225 個合成部件，亦分為成字合成部件與非字合成部件二大類，又根據其來源出處之不同，可細分為底下四類：

1. 為《漢語大字典》單字的成字合成部件：此類為收錄於《漢語大字典》的單字。在簡化字集中，此類部件共有 209 個，其部件及例字，如表 2.8。

2. 為其他字書單字的成字合成部件：此類為其他字書收錄的單字，但《漢語大字典》並未收錄。在簡化字集中，此類部件共有 2 個，其部件及例字，如表 2.9。

3. 包含在 GF3001 中的非字合成部件：此類為 GF3001 中的基礎部件，但可依 CNS 11643-2 的拆分原則拆分，故並不列於基礎部件集中。在簡化字集中，此類部件共有 1 個，其部件及例字，如表 2.10。

4. 具有特定構意的非字合成部件：此類部件有特定的構意，且在《漢語大字典》有兩個（含）以上的單字用到該部件，具有構形的能力，即收錄為非字合成部件。在簡化字集中，此類部件有 13 個，其部件及例字，如表 2.11。

表 2.8：為《漢語大字典》單字的簡化字集成字合成部件

序 號	部 件	例 字
1	万	厉蚤
2	马	吗騰笃驰
3	无	吮芜庠
4	艺	屹
5	历	圻痲雳
6	区	佞欧奘
7	冈	刚岗
8	仑	伦囡瘰
9	仓	伦创疮苍
10	风	岚枫疯颶
11	鸟	鸣邬
12	办	协苏
13	双	轰
14	击	陆
15	巧	窍
16	节	栉
17	厉	励砺
18	卢	垆颅
19	归	岾
20	务	雾
21	刍	绉邹趋
22	鸟	鳩鶯鸵鷹
23	兰	栏
24	头	买
25	汇	扌
26	写	泻
27	对	树忒
28	发	废泼
29	队	坠
30	丝	咄鷺
31	动	恻
32	亚	严垭壺氫
33	执	衰势
34	仄	厝悞
35	页	顷囂
36	夹	峡莢頰愜
37	轨	甌
38	毕	哔篳
39	尧	晓茺翹

序 號	部 件	例 字
40	贞	侦
41	师	狮筛
42	壮	装
43	当	档
44	岁	秽刈
45	岂	凯皑闾
46	则	测厠
47	乔	桥荞
48	印	鲫
49	华	哗
50	会	绘荟刽
51	杀	铄
52	负	赖
53	争	诤阐
54	庄	脏
55	刘	浏
56	齐	济霁剂齏
57	产	铲颜
58	兴	举
59	汤	烫荡
60	军	浑晕鞞
61	许	浒
62	农	侏
63	边	筵
64	寻	浔荨
65	尽	焮苾
66	买	卖
67	纟	葑
68	孙	獠赫遜
69	红	荳
70	约	哟药
71	寿	俦焘筹
72	麦	唛麸
73	贡	喷贛
74	壳	恧
75	覓	寬
76	严	伊
77	劳	涝痲
78	丽	俪邈鸱

序 號	部 件	例	字
79	来	徠萊賚	
80	鹵	鹵	
81	坚	鯉	
82	时	埶詩	
83	吴	誤	
84	县	懸	
85	围	圍	
86	冎	剛媧窩	
87	员	勛損圓	
88	兵	賓	
89	金	儉劍筌	
90	龟	闖	
91	条	洿	
92	犹	莠	
93	鸟	搗	
94	库	褲	
95	闰	潤	
96	闲	嫻痲鸛	
97	间	澗簡	
98	闷	悶	
99	闷	炯	
100	穷	勞	
101	进	璉	
102	连	璉蓮	
103	过	捫	
104	达	咤闕	
105	迁	跼	
106	灵	稭	
107	张	漲	
108	纯	莼	
109	阴	蔭	
110	责	債簣	
111	规	窺	
112	择	擇	
113	画	媧	
114	卖	腴窶覲	
115	转	轉	
116	斩	塹嶄慚	
117	顷	傾頤穎	
118	齿	啮齡	
119	虏	擄	
120	国	帼	

序 號	部 件	例	字
121	龟	繩	
122	罗	萝啰邏	
123	质	躋	
124	征	懲	
125	鱼	漁穌魯鯿騰	
126	单	彈箏鞞闡	
127	审	嬋	
128	肃	嘯簫	
129	录	綠	
130	弥	猕	
131	参	渗疹	
132	贯	慣	
133	贲	噴	
134	贲	勦	
135	带	滯	
136	荣	嵘	
137	貳	膩	
138	览	纜	
139	将	蔣鏘	
140	尝	償	
141	显	濕	
142	显	蘊	
143	贵	饋匱簣	
144	虽	強	
145	贱	濺	
146	钦	嵌揄	
147	急	穩	
148	弯	灣	
149	彦	諺顏	
150	冏	桐	
151	阁	摺	
152	郑	擲	
153	娄	樓屢數窶	
154	总	聰	
155	举	桴	
156	觉	攪	
157	费	鑽	
158	艳	滌	
159	聂	攝顛	
160	晋	縉	
161	贾	檟	
162	荐	鞞	

序號	部件	例	字
163	慮	濾	
164	監	檻 藍	
165	罷	摆 罨	
166	筆	滄	
167	隹	錨	
168	拿	鎗	
169	愛	暖	
170	朶	灑	
171	廔	纏	
172	寬	腕	
173	賓	賓 鬢	
174	諸	儲	
175	難	攤	
176	預	瀕	
177	黃	糞	
178	蕭	瀟	
179	嗇	墻	
180	虛	覷	
181	嬰	櫻 癭 鸚	
182	斂	潑	
183	斷	齧	
184	剪	灑	
185	渴	藹	
186	維	淮	

序號	部件	例	字
187	輦	撵	
188	頤	擷	
189	雁	贖	
190	魯	魯 穩	
191	闌	爛	
192	窳	擻	
193	屬	囑	
194	強	強	
195	緼	蘊	
196	隱	癮	
197	緣	椽	
198	猷	猷	
199	賴	懶 癩 籟	
200	頻	瀕 颯	
201	數	擻 藪	
202	滿	懣	
203	藺	躡	
204	鮮	癬 藓	
205	闕	咽	
206	顛	巔 癲	
207	贊	瓚 趲	
208	顛	灑	
209	贛	慇	

表 2.9：為其他字書單字的簡化字集成字合成部件

序號	部件	例	字
210	友	钹	
211	呂	營	

表 2.10：在 GF3001 中的簡化字集非字合成部件

序號	部件	例	字
212	宀	臨	

表 2.11：具有特定構意的簡化字集非字合成部件

序號	部件	例	字
213	収	堅	
214	宀	貯	
215	艹	勞	

序號	部件	例	字
216	叺	監	
217	宀	學	
218	彳	擇	

序 號	部 件	例 字
219	至	劲径瘞荃
220	𠂆	奖
221	亦	弯
222	𠂆	断继

序 號	部 件	例 字
223	兔	挽
224	苒	满颞
225	岳	谣鷓

第三章 漢字構形資料庫的構形編碼

漢字可拆分成有限的部件，而這些有限的部件可組成數以萬計的漢字。除了不同的部件可組成不同的漢字外，相同的部件也可利用相對位置或部件的個數來組成不同的漢字。部件及其相對位置可用來定義漢字的字形結構式，即構字式，而構字式可作字形的識別碼。本章首先說明部件的組合及識別，再說明構字式的制定、使用及處理技巧，最後再根據構字式延伸出用來處理古漢字的風格碼。

3.1 部件的組合及識別

漢字是由有限的部件所組成，除了不同的部件可組成不同的漢字外，相同的部件也可利用相對位置或部件的個數來組成不同的漢字。以五大字集的 13,501 個字形為例，其中有 12,817 個字使用的部件均不相同，部件相同但相對位置不同的字共有 115 組，234 個字。例如「架」、「枷」、「拐」三個字均由部件「力」、「口」、「木」所組成，不同的是部件的相對位置。這 115 組字如下：

加（另、叻）、旭（見、沓）、架（枷、拐）、翌（翊、玃）、可（叮）、召（叨）、只（叭）、妃（改）、吝（叻）、否（呸）、呆（杏）、含（吟）、岑（嶺）、岌（岌）、旱（旰）、李（籽）、防（邗）、阮（邗）、阮（邗）、盆（坳）、周（垌）、坪（幸）、委（姝）、帕（帛）、招（拐）、易（吻）、杏（杲）、盲（眈）、邸（阨）、眈（旻）、炅（耿）、邗（阨）、保（咻）、垢（垠）、某（柑）、查（担）、毗（毘）、泵（砵）、珀（皇）、省（眇）、崑（崑）、衍（沂）、郁（隋）、呷（聃）、些（吡）、响（响）、崑（崑）、峇（峇）、峇（峇）、昶（咏）、邗（阨）、哲（晰）、員（員）、宴（晏）、响（响）、案（按）、烘（糞）、杳（杳）、邗（阨）、啦（培）、唯（售）、堊（堊）、娶（媿）、脣（脣）、部（陪）、都（陪）、陴（郟）、陴（郟）、唛（培）、媿（媿）、勛（賀）、啻（啼）、景（晾）、晷（晷）、棘（棗）、椎（集）、渺（涓）、郵（陸）、岳（岳）、椶（聚）、棊（棊）、焚（楨）、澇（澇）、郇（隄）、暉（暉）、蜃（蜃）、訾（訾）、賁（賁）、隕（郟）、楸（楸）、輶（輶）、塵（塵）、墓（墓）、幕（幕）、摧（摧）、窪（窪）、蜥（蜥）、障（鄣）、唯（唯）、蜓（蜃）、暮（暮）、楸（楸）、楸（楸）、籐（籐）、謀（謀）、蕩（蕩）、寫（寫）、麋（麋）、塵（塵）、樞（樞）、樞（樞）、藹（藹）、礪（礪）、鳩（鳩）、蠟（蠟）

另外，五大字集還有 67 個字僅由單一部件重複組合而成：

比、爻、圭、多、朋、林、炎、玨、品、姦、哥、弱、茲、晶、棘、棗、兢、赫、磊、聶、蟲、競、轟、轟、荔、姁、孖、屾、艸、爰、田、林、牝、垚、焱、姓、砬、秝、煖、憇、犇、淼、焱、焱、焱、晒、睥、銜、躄、覓、豕、賈、晶、晶、焱、虤、雥、聶、囍、鑫、龔、轟、轟、轟、轟、轟、轟、轟

綜合上述，絕大多數的漢字皆可由其組成的部件來識別，少數字形則須再描述部件的相對位置。部件的組合及識別原則進一步說明如下：

1. 漢字係由一或多個部件依層次逐級組合而成，絕大多數字形可透過各級部件的組合來識別。例如「謝」字的各級部件組合「言射」、「言身寸」都可用來識別「謝」字；「寶」字的一級部件組合「宀王缶貝」可用來識別「寶」字。

2. 少數漢字的差異僅在於部件的相對位置不同，要識別此類字形，除部件本身外，還須描述部件的相對位置。此類字形通常僅有兩個一級部件，而部件的相對位置可簡化成左右、上下及內外三類。例如「暉」、「暈」二字的差別僅在部件「日」、「軍」的相對位置為左右或上下；「見」、「旭」二字的差別僅在部件「日」、「九」的相對位置為上下或內外。

3. 少數漢字是由單一部件重複組合而成，此類部件的組合方式通常為由左至右，由上至下，或呈三角狀、四角狀排列。例如「林」字由單一部件「木」由左至右排列，「棗」字由單一部件「束」由上至下排列，「轟」字由單一部件「車」呈三角狀排列，「焱」字由單一部件「火」呈四角狀排列。

4. 大多數漢字部件的相對位置，可由前後部件之性質判斷得知。例如部件「言」與其他部件組合字形時，部件的相對位置通常為左右，而「言」在左方，因此「言」、「射」組合成「謝」；部件「雨」與其他部件組合字形時，部件的相對位置通常為上下，而「雨」在上方，因此「雨」、「相」組合成「霜」；部件「口」和其他部件組合字形時，部件的相對位置通常為內外，而「口」在外，因此「口」、「員」組合成「圓」。

5. 基礎部件為構形識別的最小單位，不得再拆分成其他基礎部件來識別字形。例如部件「貝」為 CNS 11643-2 規定之基礎部件，不得再拆分成「目」、「八」，因此「目八」無法用以識別「貝」字，「宀王







缶目八」也無法用以識別「寶」字。

3.2 構字式的制定及使用







構字式即字形結構表達式，可用來識別漢字。依據部件組合及識別原則，構字式的定義及使用方式如下：











1. 構字式可用於識別字形，它係由部件及構字符號組成，詳見表 3.1。














2. 表 3.1 的構字符號分為形標及定位符號兩種，定位符號可用於描述部件的相對位置，形標則不涉及部件的相對位置。

3. 形標為字形的標示，由構字符號「」及「」構成，其中「」為字形標示的起點，「」表示結束，用以識別字形的部件則包夾其中。例如構字式「宀王缶貝」可用以識別「寶」字。

4. 定位符號分成連接符號及重疊符號兩類，連接符號主要用以描述不同部件的組合情形，重疊符號僅用以描述單一部件之重複組合。

5. 連接符號有三個，「」為橫連，用於描述部件的組合順序為由左至右；「」為直連，用於描述部件的組合順序為由上至下；「」為包含，用於描述部件的組合順序為由外至內。在構字式中，連接符號應置於部件之間。例如構字式「言射」可用以識別「謝」字；「雨相」可用以識別「霜」字；「口員」可用以識別「圓」字。

6. 在構字式中，不同的連接符號不可併用，而相同的連接符號可連續使用。例如「寶」字的構字式不可寫成「宀王缶貝」，應寫成「宀王缶貝」；「謝」字的構字式可寫成「言射」或「言身寸」；「晷」字的構字式可寫成「日共言」。

7. 採用形標的構字式，雖然形標內的部件順序基本上不影響字形的識別，但仍應制定部件排序之原則。目前採用之排序原則為部件間的連接順序：先假設不同的連接符號可併用，再抽離連接符號，留下的部件順序即是。例如先假設「寶」字的構字式可寫成「宀王缶貝」，抽離連接符號後，構字式為「宀王缶貝」；假設「虜」字的構字式可寫成「虍毋力」，抽離連接符號後，構字式為「虍毋力」；假設「獸」字的構字式可寫成「豶酉欠」，抽離連接符號後，構字式為「豶酉欠」。

8.重疊符號共八個，常用的有「∞」、「∞」、「∞」、「∞」四個，其中「∞」用以描述兩個相同部件的組合方式為由左至右，「∞」用以描述兩個相同部件的組合方式為由上至下，「∞」用以描述三個相同部件的組合方式為三角狀排列，「∞」用以描述四個相同部件的組合方式為四角狀排列；其餘四個不常用的重疊符號為「∞」、「∞」、「∞」、「∞」，其中「∞」用以描述三個相同部件的組合方式為由左至右，「∞」用以描述三個相同部件的組合方式為由上至下，「∞」用以描述四個相同部件的組合方式為由左至右，「∞」用以描述四個相同部件的組合方式為由上至下。在構字式中，重疊符號應置於部件之前。例如構字式「∞木」可用以識別「林」字，「∞束」可用以識別「棗」字，「∞車」可用以識別「轟」字，「∞火」可用以識別「燚」字。

9.在構字式中，形標和連接符號不可併用。例如「謝」字的構字式可寫成「言△射」、「言△身△寸」或「形言射□」、「形言身寸□」，但不可寫成「形言△射□」或「形言△身△寸□」。

10.在構字式中，重疊符號可和形標或連接符號併用。例如「堯」字的構字式可寫成「堯△兀」，也可寫成「∞土△兀」；「歎」字的構字式可寫成「形歎西欠□」，也可寫成「形又又又又西欠□」或「形∞又西欠□」。

11.由於一個字可透過各級部件的組合來識別，所以同一個字也可能有數種不同的構字式，但僅有一級部件所組合的構字式最貼近造字意圖。例如「謝」字的構字式可寫成「言△射」或「言△身△寸」，其中「言」、「射」為一級部件，而「身」、「寸」為二級部件，因此「言△射」比「言△身△寸」更貼近「謝」的造字意圖；「廬」字的構字式可寫成「雨△歷」、「形雨麻止□」、「形雨厂秝止□」、或「形雨厂禾禾止□」，其中「雨」、「歷」為一級部件，「麻」、「止」為二級部件，「厂」、「秝」為三級部件，而「禾」為四級部件，因此「雨△歷」最貼近「廬」的造字意圖。

表 3.2 為五大字集最後五百個次常用字的構字式，每個字的構字式都儘可能採用一級部件來組合，以貼近該字的造字意圖。

表 3.1：構字符號及構字式

構字符號		說明	構字式	
形標		◻為字形標示的起點，◻表示結束，部件則包夾於其中。	寶 = 宀王缶貝	
連接符號		橫連—當部件的組合順序為由左至右	謝 = 言射	
		直連—當部件的組合順序為由上至下	霜 = 雨相	
		包含—當部件的組合順序為由外至內	圓 = 口員	
定位符號	重疊符號	○○	兩個相同部件的組合方式為由左至右	林 = ○○木
		⊗	兩個相同部件的組合方式為由上至下	棗 = ⊗束
		∞	三個相同部件的組合方式為由左至右	孖 = ∞子
		⊗	三個相同部件的組合方式為由上至下	𡗗 = ⊗小
		⊙	三個相同部件的組合方式為三角狀排列	轟 = ⊙車
		∞∞	四個相同部件的組合方式為由左至右	𠄎 = ∞∞
		⊗	四個相同部件的組合方式為由上至下	𠄎 = ⊗一
⊙	四個相同部件的組合方式為四角狀排列	燚 = ⊙火		

表 3.2：五大字集的構字式(次常用字的後五百個)

Big5	字	構字式	Big5	字	構字式	Big5	字	構字式	Big5	字	構字式
F6B9	鞫	韋△鳥	F6E1	攆	扌△靡	F74A	蕪	艹△歸	F772	鞫	韋△蜀
F6BA	鶻	栗△鳥	F6E2	攏	扌△麗	F74B	蕪	艹△壘	F773	顛	歲△頁
F6BB	鸚	晏△鳥	F6E3	攢	扌△贊	F74C	蕪	艹△魏	F774	顛	僉△頁
F6BC	鶻	骨△鳥	F6E4	權	木△瞿	F74D	蠃	虫△龍	F775	颯	風△瑟
F6BD	鶻	芻△鳥	F6E5	檣	木△嵩	F74E	蠃	虫△蕭	F776	饕	雍△食
F6BE	鶻	泉△鳥	F6E6	檣	木△叢	F74F	蠃	虫△盧	F777	饕	食△蒙
F6BF	鶻	留△鳥	F6E7	瞿	瞿△毛	F750	蠃	龍△虫	F778	驕	馬△葬
F6C0	鶻	脊△鳥	F6E8	灑	灑△離	F751	蠃	辭△虫	F779	驕	馬△曾
F6C1	鶻	倉△鳥	F6E9	灑	灑△靡	F752	龍	龍△龍	F77A	驕	馬△單
F6C2	鶻	鳥△虎	F6EA	灑	灑△壇	F753	龍	龍△見	F77B	驕	馬△肅
F6C3	鶻	突△鳥	F6EB	灑	灑△贊	F754	龍	龍△見	F77C	驕	馬△屏
F6C4	鶻	紡△鳥	F6EC	蠃	火△蟲	F755	龍	龍△樂	F77D	驕	馬△喬
F6C5	鶻	鳥△馬	F6ED	蠃	火△霍	F756	龍	龍△翦	F77E	驕	馬△華
F6C6	鶻	鳥△師	F6EE	蠃	魏△牛	F757	龍	龍△晶	F7A1	驕	馬△虛
F6C7	鶻	泉△鳥	F6EF	蠃	牙△嬰	F758	龍	龍△夙	F7A2	驕	馬△單
F6C8	鶻	翁△鳥	F6F0	璽	王△霍	F759	龍	龍△言	F7A3	驕	馬△敦
F6C9	鶻	鹵△差	F6F1	璽	彌△玉	F75A	龍	龍△審	F7A4	驕	骨△堯
F6CA	麋	鹿△栗	F6F2	璽	王△燾	F75B	龍	龍△慧	F7A5	驕	髟△曾
F6CB	黠	黑△聖	F6F3	璽	王△嵩	F75C	龍	龍△賣	F7A6	鬪	鬥△敢
F6CC	黠	黑△甚	F6F4	瘦	疒△嬰	F75D	龍	龍△廚	F7A7	鬪	鬪△米
F6CD	黠	黑△弁	F6F5	皦	白△爵	F75E	龍	龍△塵	F7A8	鬪	鬼△虛
F6CE	馨	鼓△谷	F6F6	礪	石△霜	F75F	龍	龍△遷	F7A9	鬪	鬼△幾
F6CF	蕭	鼓△崩	F6F7	禱	礻△龠	F760	龍	龍△樂	F7AA	鱒	魚△章
F6D0	襲	鼓△長	F6F8	穰	禾△襄	F761	龍	龍△巖	F7AB	鱒	魚△雪
F6D1	鼯	鼠△青	F6F9	穰	禾△爵	F762	龍	龍△蠟	F7AC	鱒	魚△責
F6D2	齋	齊△貝	F6FA	籊	竹△霍	F763	龍	龍△足	F7AD	鱒	魚△專
F6D3	齋	齒△曳	F6FB	籊	竹△擇	F764	龍	龍△晶	F7AE	鱒	魚△堅
F6D4	齋	犬△齒	F6FC	籊	竹△錄	F765	龍	龍△樂	F7AF	鱒	魚△勒
F6D5	龔	龍△示	F6FD	籊	竹△錢	F766	龍	龍△卞	F7B0	鱒	魚△逐
F6D6	齋	一△豐	F6FE	籊	竹△盧	F767	龍	龍△賓	F7B1	鱒	魚△習
F6D7	囁	口△蠆	F740	耀	耀△入	F768	龍	龍△需	F7B2	鱒	魚△條
F6D8	鞫	單△展	F741	鞫	辭△米	F769	龍	龍△萑	F7B3	徽	微△魚
F6D9	噴	口△贊	F742	縷	纟△盧	F76A	龍	龍△金	F7B4	鰲	敖△魚
F6DA	變	絲△大	F743	縷	岳△盧	F76B	龍	龍△寧	F7B5	鰲	魚△曹
F6DB	嬈	女△麗	F744	鞫	鞫△革奇	F76C	龍	龍△齊	F7B6	鰲	魚△祭
F6DC	變	絲△女	F745	瞿	月△瞿	F76D	龍	龍△榮	F7B7	鸚	庸△鳥
F6DD	嶸	山△孽	F746	艦	舟△盧	F76E	龍	龍△爾	F7B8	鸚	專△鳥
F6DE	噴	山△贊	F747	豐	豐△豐	F76F	龍	龍△蓋	F7B9	鸚	爽△鳥
F6DF	麗	广△麗	F748	職	職△職	F770	龍	龍△對	F7BA	鸚	麥△鳥
F6E0	攏	扌△離	F749	藎	藎△藎	F771	龍	龍△督	F7BB	鸚	茶△鳥

Big5	字	構字式
F7BC	鷓	晨△鳥
F7BD	鷓	婁△鳥
F7BE	鷓	笠△鳥
F7BF	鷓	族△鳥
F7C0	鷓	敝△鳥
F7C1	鷓	執△鳥
F7C2	鷓	敕△鳥
F7C3	鷓	毆△鳥
F7C4	鷓	屠△鳥
F7C5	鷓	唯△鳥
F7C6	鷓	畢△鳥
F7C7	鷓	麥△离
F7C8	鷓	黑△真
F7C9	鷓	鼠△突
F7CA	鷓	鼠△昊
F7CB	鷓	鼠△軍
F7CC	鷓	鼻△隶
F7CD	鷓	齒△困
F7CE	鷓	龕△龍
F7CF	鷓	龕△禾
F7D0	鷓	亻△粟
F7D1	鷓	蟲△彳
F7D2	鷓	罍△土
F7D3	鷓	土△嚴
F7D4	鷓	奢△單
F7D5	鷓	女△嚴
F7D6	鷓	山△獻
F7D7	鷓	廡△虫
F7D8	鷓	弓△雙
F7D9	鷓	難△心
F7DA	鷓	忄△黨
F7DB	鷓	忄△雙
F7DC	鷓	扌△黨
F7DD	鷓	扌△纂
F7DE	鷓	文△豐
F7DF	鷓	緜△日
F7E0	鷓	攢△贊
F7E1	鷓	樂△木
F7E2	鷓	擲△羅
F7E3	鷓	殼△喬
F7E4	鷓	灬△闡
F7E5	鷓	灬△覺

Big5	字	構字式
F7E6	麤	靡△彡
F7E7	獠	犛△薪
F7E8	獠	犛△嚴
F7E9	獠	犛△雙
F7EA	癰	疒△雖
F7EB	瞞	目△瞞
F7EC	籬	竹△遽
F7ED	鐘	钅△鍾
F7EE	纓	纟△襄
F7EF	艘	舟△冕
F7F0	離	艹△離
F7F1	壑	艹△壑
F7F2	懷	艹△懷
F7F3	靡	艹△靡
F7F4	類	艹△類
F7F5	繫	艹△繫
F7F6	壞	艹△壞
F7F7	蟻	虫△襄
F7F8	蠲	益△蜀
F7F9	蠲	虫△翳
F7FA	嬰	虫△嬰
F7FB	襪	衤△戴
F7FC	襪	衤△闌
F7FD	襪	衤△鐵
F7FE	鱗	角△燕
F840	燕	言△燕
F841	讎	讎△言
F842	讎	龍△言
F843	讎	言△歷
F844	讎	谷△龍
F845	贊	虢△貝
F846	躡	足△龍
F847	輻	車△盧
F848	輻	車△歷
F849	醜	酉△燕
F84A	鑄	金△慮
F84B	鑄	金△質
F84C	鑄	金△蓬
F84D	鑄	金△黎
F84E	鑄	金△鼠
F84F	鞣	革△萑
F850	鞣	革△焮

Big5	字	構字式
F851	護	音△萑
F852	馱	馬△戴
F853	馱	馬△宣
F854	鬢	彡△農
F855	鬢	彡△農
F856	鬢	彡△會
F857	鱒	魚△尊
F858	鱒	魚△尋
F859	鱒	魚△肅
F85A	鱒	魚△喬
F85B	鱒	魚△發
F85C	鱒	魚△虛
F85D	鱒	魚△番
F85E	鱒	魚△堯
F85F	鱒	魚△象
F860	鱒	魚△喬
F861	鵠	鳥△敦
F862	鵠	尊△鳥
F863	鵠	奈△鳥
F864	鵠	覃△鳥
F865	鵠	肅△鳥
F866	鵠	喬△鳥
F867	鵠	單△鳥
F868	鵠	買△鳥
F869	鵠	無△鳥
F86A	鵠	喬△鳥
F86B	鵠	焦△鳥
F86C	鵠	就△鳥
F86D	鵠	燕△鳥
F86E	鵠	厥△鳥
F86F	鵠	黃△鳥
F870	鵠	閑△鳥
F871	鵠	閑△鳥
F872	鵠	華△鳥
F873	鵠	番△鳥
F874	鵠	麻△賁
F875	鵠	黍△离
F876	鵠	黑△參
F877	鵠	毆△黑
F878	鵠	冥△黽
F879	鵠	鼓△蚤
F87A	鵠	鼠△兼

Big5	字	構字式
F87B	鼯	鼠△奚
F87C	鼯	鼠△虎
F87D	鼯	鼻△曷
F87E	齋	齊△非
F8A1	齋	齒△取
F8A2	齋	齒△昔
F8A3	齋	齒△奇
F8A4	齋	齒△兒
F8A5	齋	口△齧
F8A6	齋	∞喜
F8A7	孀	女△屬
F8A8	廩	尸△聶
F8A9	蠶	扌△蠶
F8AA	曠	日△黨
F8AB	曠	日△嚴
F8AC	曠	木△黨
F8AD	灑	灬△屬
F8AE	灑	灬△蘭
F8AF	灑	灬△穎
F8B0	灑	灬△覽
F8B1	燻	火△黨
F8B2	燻	王△獻
F8B3	蠶	蟲△瓜
F8B4	管	緜△目
F8B5	磧	石△贊
F8B6	襪	衤△類
F8B7	襪	衤△贊
F8B8	斷	斤△斷
F8B9	縞	纟△嵩
F8BA	羅	四△緜
F8BB	艘	舟△雙
F8BC	灑	灬△灑
F8BD	蠶	虫△萑
F8BE	蠶	虫△瞿
F8BF	蠶	虫△嵩
F8C0	蠶	聿△爾△血
F8C1	讎	言△隱
F8C2	讎	言△闌
F8C3	躡	足△燮
F8C4	躡	足△襄
F8C5	躡	薛△足
F8C6	躡	足△闌

Big5	字	構字式
F8C7	醜	酉△麋
F8C8	醜	酉△需
F8C9	醜	酉△爵
F8CA	鑫	品△金
F8CB	鎧	金△龍
F8CC	錘	金△璽
F8CD	隼	品△隹
F8CE	隼	雲△逮
F8CF	霍	雨△雒
F8D0	寵	雨△龍
F8D1	贛	革△賣
F8D2	贛	韋△賣
F8D3	贛	馬△寶
F8D4	贛	骨△寶
F8D5	灑	漸△鬼
F8D6	鱣	魚△亘
F8D7	鱣	魚△豐
F8D8	鯢	魚△睪
F8D9	鯢	魚△栗
F8DA	鯢	魚△叢
F8DB	鯢	魚△會
F8DC	鵠	溪△鳥
F8DD	鵠	意△鳥
F8DE	鵠	亘△鳥
F8DF	鷺	鳥△義
F8E0	鷺	虞△鳥
F8E1	鷺	畢△鳥
F8E2	鷺	蜀△鳥
F8E3	羸	羸△鳥
F8E4	鷺	楊△鳥
F8E5	鷺	辟△鳥
F8E6	鷺	與△鳥
F8E7	鷺	敦△鳥
F8E8	廬	鹿△量
F8E9	鑿	鼓△堂
F8EA	鸚	鼻△邕
F8EB	齶	齒△彥
F8EC	齶	齒△禺
F8ED	齶	齒△号
F8EE	嚙	口△囊
F8EF	攢	扌△囊
F8F0	屬	屬△斤

Big5	字	構字式
F8F1	屬	木△屬
F8F2	標	木△景
F8F3	欄	木△蘭
F8F4	檣	木△蟲
F8F5	囊	汜△囊
F8F6	覽	火△覽
F8F7	變	牛△變
F8F8	黨	目△黨
F8F9	闕	目△闕
F8FA	嚴	石△嚴
F8FB	邊	竹△邊
F8FC	贊	竹△贊
F8FD	耀	耀△出
F8FE	麗	彡△麗
F940	贊	彡△贊
F941	縣	毒△縣
F942	羅	彡△羅
F943	肉	豸△肉
F944	難	難△肉
F945	景	彡△景
F946	彡	彡△彡
F947	啣	彡△啣
F948	麗	彡△麗
F949	繭	彡△繭
F94A	藝	彡△藝
F94B	攀	彡△攀
F94C	犛	角△犛
F94D	聶	言△聶
F94E	萑	言△萑
F94F	竄	足△竄
F950	藉	足△藉
F951	瞿	足△瞿
F952	薄	金△薄
F953	闌	金△闌
F954	鐵	金△鐵
F955	兔	金△兔
F956	寒	金△寒
F957	愛	雲△愛
F958	頁	齶△頁
F959	襄	食△襄
F95A	嘗	魚△嘗
F95B	與	魚△與

Big5	字	構字式
F95C	齊	魚△齊
F95D	鳥	寧△鳥
F95E	鳥	爾△鳥
F95F	鳥	翟△鳥
F960	鳥	蒙△鳥
F961	鳥	與△鳥
F962	鳥	獄△鳥
F963	齊	鹿△齊
F964	詹	黑△詹
F965	毘	單△毘
F966	虛	鼻△虛
F967	益	齒△益
F968	真	齒△真
F969	芻	齒△芻
F96A	齒	差△齒
F96B	樂	口△樂
F96C	顯	汜△顯
F96D	贏	竹△贏
F96E	雙	虫△雙
F96F	贊	走△贊
F970	贊	足△贊
F971	麗	西△麗
F972	嵩	金△嵩
F973	壘	金△壘
F974	藏	金△藏
F975	萑	金△萑
F976	燕	馬△燕
F977	蔑	魚△蔑
F978	樂	魚△樂
F979	厲	魚△厲
F97A	箴	魚△箴
F97B	暴	鳥△暴
F97C	鳥	晶△鳥
F97D	黑	厭△黑
F97E	毘	辟△毘
F9A1	皆	龠△皆
F9A2	贛	汜△贛
F9A3	泉	品△泉
F9A4	蘭	米△蘭
F9A5	儵	虎△儵
F9A6	屬	虫△屬
F9A7	蝮	蝮△蝮

Big5	字	構字式
F9A8	蝮	虫△蝮
F9A9	獻	言△獻
F9AA	雙	豸△雙
F9AB	雙	足△雙
F9AC	寶	身△寶
F9AD	瀝	雨△瀝
F9AE	頁	聶△頁
F9AF	頁	萑△頁
F9B0	風	萑△風
F9B1	贊	食△贊
F9B2	香	品△香
F9B3	襄	馬△襄
F9B4	霜	馬△霜
F9B5	鞠	馬△鞠
F9B6	襄	彡△襄
F9B7	鳥	盧△鳥
F9B8	鳥	龍△鳥
F9B9	農	鼻△農
F9BA	心	贛△心
F9BB	靈	木△靈
F9BC	靈	火△靈
F9BD	龍	彡△龍
F9BE	變	足△變
F9BF	黨	金△黨
F9C0	獻	金△獻
F9C1	雙	金△雙
F9C2	萑	馬△萑
F9C3	嵩	馬△嵩
F9C4	龜	鬥△龜
F9C5	鳥	龠△鳥
F9C6	鬱	火△鬱
F9C7	譽	彡△譽
F9C8	賣	言△賣
F9C9	屬	金△屬
F9CA	萑	魚△萑
F9CB	豐	麥△豐
F9CC	齶	疒△齶
F9CD	馬	品△馬
F9CE	麗	魚△麗
F9CF	鳥	麗△鳥
F9D0	豔	汜△豔
F9D1	鬱	汜△鬱

Big5	字	構字式
F9D2	麤	品鹿

Big5	字	構字式
F9D3	齧	獻𠂇齒

Big5	字	構字式
F9D4	龕	鼻𠂇囊

Big5	字	構字式
F9D5	龕	品龍

3.3 構字式的處理技巧

本節說明構字式的處理技巧。首先分析構字式的類型，並針對不同類型的構字式，說明擷取時如何判斷起始及終結字元，以及構字式如何在資料庫中表達；最後，針對同一個字的不同構字式，說明如何進行構字式正規化。

3.3.1 構字式的類型

構字式可依構字符號分成五個類型，詳如表 3.3。前三個類型是由連接符號決定，型四採用形標，型五只包含重疊符號。

表 3.3：構字式的類型

類 型	說 明	字 例	構 字 式	備 註
1	由部件、橫連符號或重疊符號構成	頰	桑△頁	只要出現橫連符號即屬於此型
		掰	手△分△手	
		斂	貝△女	
2	由部件、直連符號或重疊符號構成	霏	雨△郭	只要出現直連符號即屬於此型
		簞	竹△甫△皿	
		鬻	〇无△鬲	
3	由部件、包含符號或重疊符號構成	麕	鹿△兒	只要出現包含符號即屬於此型
		龠	龠△〇〇口	
4	由部件、形標或重疊符號構成	遼	形△遼△彖△	只要出現形標即屬於此型
		豐	形△山〇△丰△豆△	
5	只由重疊符號構成	虬	〇〇虎	只有出現重疊符號
		戔	〇戈	
		聶	〇〇貝	
		焱	〇〇火	

3.3.2 構字式的擷取

構字式可夾雜在字串中使用，由字串擷取構字式，首先得先找出起始和終結字元。例如「大桑△頁虬髯骨相奇」中構字式的起始和終結字元分別為「桑」、「頁」。構字式的擷取類型可依起始字元來判斷，並依起始字元是否為連接符號、形標或重疊符號分成三個類型，這三個類型的終結字元判斷詳如表 3.4。

表 3.4：構字式的起始和終結字元

類型	符號	字 例	構 字 式	起始	終結	備 註
1	△	頰	桑△頁	桑	頁	當第一個構字符號為連接符號時，則起始字元為△、△、△的前一個；終結字元為連續兩個非構字符號的第一個，例如「大桑△頁虬髯骨相奇」之「頁」。
		掰	手△分△手	手	手	
	△	霽	雨△郭	雨	郭	
		簞	竹△甫△皿	竹	皿	
	△	麕	鹿△兒	鹿	兒	
		侖	侖△∞口	侖	口	
2	形	邊	形之备糸□	形	□	當第一個構字符號為形時，則起始字元為形，終結字元為□。
		豐	形山∞丰豆□	形	□	
3	∞	斂	∞貝△女	∞	女	當第一個構字符號為重疊符號時，則起始字元為重疊符號，終結字元為連續兩個非構字符號的第一個。
	∞	虤	∞虎	∞	虎	
	品	聶	品貝	品	貝	
	∞	焱	∞火	∞	火	

3.3.3 構字式在資料庫中的表達

構字式可採用關聯式資料表儲存，並以「構字符號」及「部件序」兩個欄位來表達；「構字符號」用來表達構字式的類型，「部件序」儲存部件。不再拆分的基礎部件可視為型○，「構字符號」為0，「部件序」為基礎部件本身。例如表 3.5 中「頰」字的構字式「桑△頁」為型一，「構字符號」為1，「部件序」為「桑頁」；「桑」字的構字式「叕△木」為型二，「構字符號」為2，「部件序」為「叕木」；「叕」字的構字式「品又」為型五，「構字符號」為5，「部件序」為「品又」；「又」、「木」為基礎部件，「構字符號」為0，「部件序」為「又」、「木」。

構字式另有兩個相關的欄位「基礎部件序」、「基礎部件組」，基礎部件序可由部件的基礎部件序組合而得，而基礎部件組則由基礎部件重新依筆畫、筆順排序而得。例如表 3.5 「頰」字部件「桑」、「頁」的基礎部件序分別為「又又又木」、「一自八」，因此「頰」字的基礎部件序為「又又又木一自八」；基礎部件序「又又又木一自八」中的每個基礎部件依筆畫、筆順重新排序成「一八又又又木自」，即為基礎部件組。表 3.6 為表 3.3 各字例在資料庫中的表達。

表 3.5：「穎」字及其部件在資料庫中的表達

字	構字式	構字符號	部件序	基礎部件序	基礎部件組	
穎	桑△頁	△	1	桑頁	又又又木一自八	一八又又又木自
桑	叒△木	△	2	叒木	又又又木	又又又木
叒	品又	品	5	品又	又又又	又又又
木	木		0	木	木	木
又	又		0	又	又	又
頁	百△八	△	2	百八	一自八	一八自
百	一△自	△	2	一自	一自	一自
自	自		0	自	自	自
八	八		0	八	八	八

表 3.6：構字式在資料庫中的表達

類 型	字 例	構 字 符 號	部 件 序	基 礎 部 件 序	基 礎 部 件 組	
1	穎	△	1	桑頁	又又又木一自八	一八又又又木自
	掰	△	1	手分手	手八刀手	八刀手手
	散	△	1	貝貝女	貝貝女	女貝貝
2	霽	△	2	雨郭	雨宀口子卩	宀口子卩雨
	簞	△	2	竹甫皿	竹甫皿	皿竹甫
	鬻	△	2	〇无鬲	无无鬲	无无鬲
3	麋	△	3	鹿兒	广曲匕匕白儿	儿匕匕广曲白
	龠	△	3	龠〇口	人一冊口口口	一人口口口冊
4	遼	形口	4	辵备条	辵夕田互水	夕互辵水田
	豐	形口	4	山〇丰豆	山丰丰一口艹	一口山艹丰丰
5	虤	〇〇	5	〇〇虎	虤儿虤儿	儿儿虤虤
	戔	〇	5	〇戈	戈戈	戈戈
	鬲	品	5	品貝	貝貝貝	貝貝貝
	燚	〇〇	5	〇〇火	火火火火	火火火火

3.3.4 構字式的正規化

在實際應用上，使用者表達的構字式可能和儲存在資料庫的不同，此時可利用基礎部件序或基礎部件組進行構字式的正規化。例如「戔」字的構字式在資料庫記錄為「〇戈」，使用者可能表達成「戈△戈」，但是這兩個構字式的基礎部件序皆同為「戈戈」，因此可利用基礎部件序進行構字式的正規化。

表 3.7：記錄在資料庫的構字式及使用者表達的構字式

類型	字 例	資料庫的構字式	使 用 者 的 構 字 式
1	數	貝貝女	貝貝女
2	驚	无无鬲	无无鬲
3	龠	人一口口口	人一口口口
4	遼	辵久田彡	辵久田彡
	豐	山丰丰豆	山丰丰豆
5	虤	虎虎	虎虎
	戔	戈戈	戈戈
	聶	貝貝貝	貝貝貝
	燂	火火火火	火火火火

表 3.8：構字式的正規化

字 例	構 字 式	基 礎 部 件 序	基 礎 部 件 組	正 規 化 欄 位
數	貝貝女	貝貝女	女貝貝	
	貝貝女	貝貝女	女貝貝	基礎部件序
驚	无无鬲	无无鬲	无无鬲	
	无无鬲	无无鬲	无无鬲	基礎部件序
龠	人一口口口	人一口口口	一口口口	
	人一口口口	人一口口口	一口口口	基礎部件組
遼	辵久田彡	辵久田彡	久彡辵彡田	
	辵久田彡	辵久田彡	久彡辵彡田	基礎部件序
豐	山丰丰豆	山丰丰一口	一口山丰丰	
	山丰丰豆	山丰丰一口	一口山丰丰	基礎部件序
虤	虎虎	虎虎	虎虎	
	虎虎	虎虎	虎虎	基礎部件序
戔	戈戈	戈戈	戈戈	
	戈戈	戈戈	戈戈	基礎部件序
聶	貝貝貝	貝貝貝	貝貝貝	
	貝貝貝	貝貝貝	貝貝貝	基礎部件序
	貝貝貝	貝貝貝	貝貝貝	基礎部件序
燂	火火火火	火火火火	火火火火	
	火火火火	火火火火	火火火火	基礎部件序
	火火火火	火火火火	火火火火	基礎部件序
	火火火火	火火火火	火火火火	基礎部件序

表 3.7 針對表 3.6 的部分字例，列出記錄在資料庫的構字式及使用者可能表達的構字式，這些不同的構字式通常是因為挑選了不同的部件。例如「遽」字的一級部件為「辵备彖」，因此資料庫記錄為「 \square 辵备彖 \square 」；使用者若將一級部件誤認為「辵久田彖」，構字式就會表達成「 \square 辵久田彖 \square 」。表 3.8 則針對表 3.7 的構字式，計算出基礎部件序及基礎部件組，大多數的構字式皆可透過基礎部件序進行正規化，少數的構字式則需透過基礎部件組。例如「遽」字的構字式「 \square 辵久田彖 \square 」可透過基礎部件序「辵久田彖」正規化成「 \square 辵久田彖 \square 」；「侖」字的構字式「亼厶 ∞ 口厶冊」可透過基礎部件組「一人口口口冊」正規化成「侖 \triangle ∞ 口」。

3.4 風格碼的制定及使用

風格碼是構字式的延伸，構字式是利用字形結構來區分字形的字形碼，而風格碼則是利用出處來區分同一個字形而風格迥異的字體碼。構字式適用於楷書，風格碼則適用於金文、甲骨文等古漢字，制定風格碼的目的在於解決古漢字重文的編碼問題。

以往總利用字型的切換來顯示古漢字，但是這種方式只適用於小篆，卻不適用於其他的古漢字。因為小篆是規範的文字，而甲骨文、金文或戰國文字由於尚未規範，仍存在著大量的重文。圖 3.1 為《金文編》字頭下的字形數，《金文編》收錄的 24,261 個字形歸在 2,420 個字頭下，其中有 927 個字頭只有一個字形，字形超過 10 個的字頭有 504 個，而字形最多的字頭是「寶」，有 273 個。

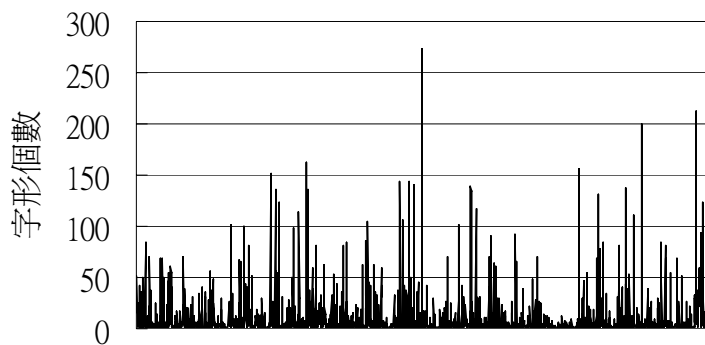


圖 3.1：《金文編》字頭下的字形數

風格碼的形式為「 \square 字形 \square 出處 \square 」或「 \square 字形 \square 出處;n \square 」，其中新增的構字符號「 \square 」用來表示同一個字的不同形體，夾在「 \square 」、「 \square 」中的為出處，若同一字在同一出處有兩個（含）以上的字形，則將字形依序編號，n 即表該字重複出現時的編號。例如金文「 𠄎 」的風格碼為「 \square 員 \square 員父尊 \square 」；說文古文「 𠄎 」、「 𠄎 」的風格碼為「 \square

中體說文古文□」，「形中體說文籀文□」

金文的引用常以器名標註出處，然而器名並不是唯一的。例如《集成》器號 23 到 30 的青銅器，器名都是「中義鐘」。另外，由於金文楷定的不同，青銅器的命名，也常和《集成》有所出入。例如器號 38 的青銅器，《集成》的器名為「挀簠鐘」，然而「挀」的後起字形為「荊」，「簠」為「曆」，所以「挀簠鐘」也可稱作「荊曆鐘」。因此，金文的風格碼為避免器名混淆，改以器號標記。例如「員父尊」的集成器號為 5861，「員」的風格碼改寫成「形員體集成 5861 □」。表 3.9 列出了漢字構形資料庫已採用的甲骨文、金文、戰國楚系簡帛文字及小篆的出處。

表 3.9：古漢字出處

古漢字	出處
甲骨文	合集(甲骨文合集)、屯(小屯南地甲骨)、英(英國所藏甲骨集)、懷(懷特氏等所藏甲骨集)。 ⁵
金文	集成器號(約 12,000 件) ⁶
楚系簡帛文字	牌 406、仰 25、常 2、望 1、望 2、天卜、天策、兩 21、馬 1、磚 370、秦 1、秦 13、秦 99、范 27、滕 1、包 2、信 1、信 2、曾、帛甲、帛乙等出土墓號及簡號。 ⁷
小篆	說文、說文或體、說文古文、說文籀文、說文篆文、說文俗字、說文奇字

表 3.10 列出古漢字「員」的風格碼，由風格碼的出處可判斷古漢字的字體。例如「員」的風格碼「形員體合集 10978 □」，出處為「合集」，因此可判斷「員」為甲骨文；「員」的風格碼為「形員體集成 9534 □」，出處為「集成」，因此可判斷「員」為金文；「員」的風格碼為「形員體秦 1.2 □」，出處為「秦 1」，因此可判斷「員」為楚系簡帛文字；「員」的風格碼為「形員體說文 □」，出處為「說文」，因此可判斷「員」為小篆。

當古漢字的楷定字需以構字式表達時，由於古漢字字集較小，部件已足夠識別該楷定字，因此構字式原有的定位符號可省略，只須將部件序置於「形、體」之間即可。例如表 3.11 甲骨文「員」的出處為「合集 35249」，其楷定字「員」的構字式為「宀員貝」，則「員」的風格碼可作「形宀員體合集 35249 □」；楚系簡帛文字「員」的出處為「天

⁵ 見《殷墟甲骨刻辭類纂》所引甲骨著錄書目

⁶ 見《殷周金文集成》目錄

⁷ 見《楚系簡帛文字編》所引文字的出土墓號簡稱表






卜」，其楷定字「𦉰」的構字式為「 水示」，則「𦉰」的風格碼可作「 水示 卜」。

表 3.10：古漢字「員」的風格碼

















































古漢字	風 格 碼	古漢字	風 格 碼
	 員  合集 10978 		 員  集成 6432 
	 員  英 1782 		 員  集成 6432;2 
	 員  集成 5861 		 員  集成 2789 
	 員  集成 5966 		 員  秦 1.2 
	 員  集成 9367 		 員  說文 
	 員  集成 9534 		 員  說文籀文 

表 3.11：古漢字的風格碼

古漢字	楷定	風 格 碼	古漢字	楷定	風 格 碼
	員	 員  合集 35249 		𦉰	 彳 水示  天 卜 
	𦉰	 臣 戈  合集 6404 反 		𦉰	 𦉰  包 2.257 
	𦉰	 金 須  集成 4388 		𦉰	 土 則  說文 
	𦉰	 木 吹  集成 3687 		𦉰	 大 巨  說文 

第四章 漢字構形資料庫的應用

漢字構形資料庫目前主要用來解決缺字問題。缺字問題由來已久，是因為目前電腦中文字碼字形不足而引起的，在處理古文獻時尤其嚴重。例如中央研究院漢籍電子文獻資料庫發展到 1998 年時，累積的五大字集缺字已超過 9,600 個。現行中文字碼的最大問題，是將漢字視同拉丁字母，採用相同的編碼架構，而完全忽略漢字是表意文字，是由有限的基礎部件所組成的開放字集。現行的中文字碼最基本的元素是「字」，而構字式最基本的元素則是「部件」；相較於現行的中文字碼，構字式更適合用來表達缺字。

漢字構形資料庫自 2002 年起，即提供免費下載。2008 年，基於技術推廣及公益考量，漢字構形資料庫、部件檢字程式及相關字型整合成「中研院漢字部件檢字系統」，依照公眾授權模式提供程式原始碼及其運作所需相關資料授權他人使用。以下說明中研院漢字部件檢字系統的下載、功能簡介及使用方式。

4.1 系統下載及安裝

中研院漢字部件檢字系統是以漢字構形資料庫為核心，下載網址為 <http://cdp.sinica.edu.tw/cdphanzi/>。漢字構形資料庫 2.53 版的下載檔案 cdphanzi253.exe 大小為 86M，cdphanzi253.exe 可自動解壓縮成 cdphanzi.cab、setup.exe、setup.lst 三個檔案，然後再執行 setup.exe 進行安裝。系統預設的安裝目錄為 C:\Program Files\Cdphanzi，安裝成功後，在「開始」功能表中可找到「缺字公用程式」之群組。詳細的安裝程序可參考 <http://cdp.sinica.edu.tw/cdphanzi/setup.htm>。

4.2 系統架構

中研院漢字部件檢字系統的架構如圖 4.1，漢字構形資料庫是整個系統的核心，目前可藉由構字式來解決 Microsoft Office 文件及網頁的缺字，其使用流程如下：

1. 使用者可透過漢字構形資料庫的操作界面，查詢所需缺字之構字式。
2. 透過漢字構形資料庫的構字式及字型對應表，擷取缺字；構字式亦可透過字形產生器直接產生缺字。
3. 使用者可在 Microsoft Office 文件中嵌入缺字之構字式，再透過缺字處理增益集在 Microsoft Office 文件中顯示缺字。

4. 網頁管理者亦可在網頁中嵌入構字式與 JavaScript 程式碼，在使用者開啟並讀取網頁時，即連至主機的缺字處理 API，並由主機傳出缺字的字形圖片至網頁，並顯示使用者所讀取的網頁中之缺字。

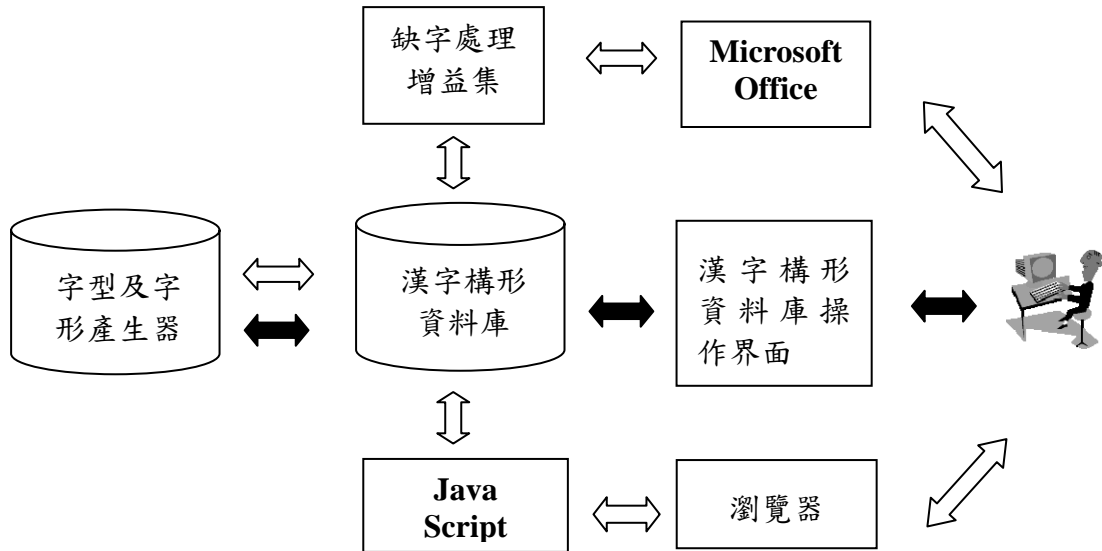


圖 4.1：應用漢字構形資料庫來解決缺字問題

4.3 漢字構形資料庫

漢字構形資料庫 2.53 版是由甲骨文、金文、楚系簡帛文字、小篆及楷書構形資料庫組合而成，其架構圖如第一章的圖 1.29。表 4.1 列出在安裝目錄中這五個資料庫的對應檔案(Microsoft Access 格式)。

表 4.1：漢字構形資料庫的對應檔名

資料庫	甲骨文	金文	楚系文字	小篆	楷書
檔案	cdpjiagu.mdb	cdpbronz.mdb	cdpchubs.mdb	cdpseal.mdb	cdphanzi.mdb

每個資料庫都至少包含「檢字表」、「異體字表」兩個資料表。檢字表儲存構字資訊，包含第三章表 3.5 的「字形」、「連接符號」、「部件序」、「基礎部件序」、「基礎部件組」等欄位，異體字表則儲存異體字資訊。古今文字的銜接資訊也是儲存在檢字表中，楷書檢字表儲存古漢字的銜接資訊，古漢字的檢字表則只儲存楷書的銜接資訊，因此古漢字間的銜接仍須透過楷書。圖 4.2 表達「員」的古今文字銜接，甲骨文「𠄎」可透過楷書「員」銜接到小篆「員」。缺字字型與構字式的對應同樣是儲存在檢字表中，包含「字體」、「字碼」兩個欄位，這

部分待 4.4 節再來說明。

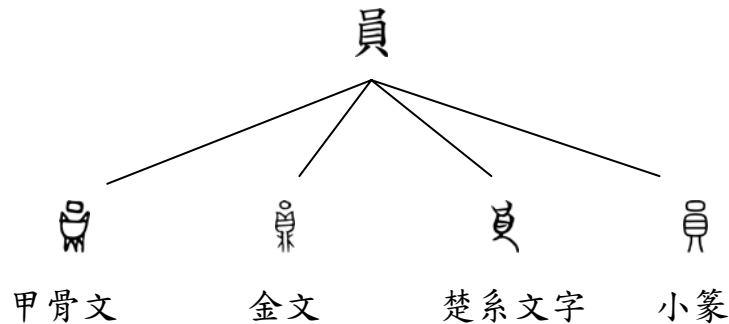


圖 4.2：「員」的古今文字銜接

4.4 缺字字型

缺字字型檔包含了楷書的細明體與標楷體，也包含甲骨文、金文、楚系簡帛文字及小篆等古漢字字型。安裝中研院漢字部件檢字系統後，在 Microsoft Office 中即可見到多出的字型如圖 4.3。



圖 4.3：古漢字字型及缺字字型

每套缺字字型只用到五大碼的外字區，最多只能編到 6,217 個字。表 4.2 中的標楷體及細明體部件外字，即是一般所謂的造字檔，其中細明體部件外字為系統預設外字集，標楷體部件外字目前只對應到標楷體。表 4.3 為古漢字字型，甲骨文取自《類纂》，金文取自《金文編》，楚系簡帛文字則取自《楚系簡帛文字編》。受到編碼空間的限制，楷書缺字目前依標楷體及細明體各分成九套，金文及楚系簡帛文字重文字型則各分成四套，每個缺字所對應的字型及字碼，則儲存在檢字表的「字體」、「字碼」兩個欄位。這些字型無法直接使用，必須以構字式或風格碼透過 Microsoft Office 中的缺字增益集進行轉換，

詳見 4.6 節。

表 4.2：楷書缺字字型檔

標楷體缺字字型	檔案名稱	細明體缺字字型	檔案名稱
標楷體部件外字	cdpeudck.tte	細明體部件外字	cdpeudc.tte
標楷體外字集一	hzcdp01k.ttf	細明體外字集一	hzcdp01m.ttf
標楷體外字集二	hzcdp02k.ttf	細明體外字集二	hzcdp02m.ttf
標楷體外字集三	hzcdp03k.ttf	細明體外字集三	hzcdp03m.ttf
標楷體外字集四	hzcdp04k.ttf	細明體外字集四	hzcdp04m.ttf
標楷體外字集五	hzcdp05k.ttf	細明體外字集五	hzcdp05m.ttf
標楷體外字集六	hzcdp06k.ttf	細明體外字集六	hzcdp06m.ttf
標楷體外字集七	hzcdp07k.ttf	細明體外字集七	hzcdp07m.ttf
標楷體外字集八	hzcdp08k.ttf	細明體外字集八	hzcdp08m.ttf
標楷體外字集九	hzcdp09k.ttf	細明體外字集九	hzcdp09m.ttf

表 4.3：古漢字字型檔

古漢字	電腦字型	字數	檔案名稱	備註
甲骨文	中研院甲骨文	1,970	cdpjiagu.ttf	收錄《類纂》可楷定的甲骨文。
	中研院甲骨文重文	4,351	hzcdp01j.ttf	收錄《類纂》的甲骨文。
金文	中研院金文	3,781	cdpbronz.ttf	收錄《金文編》的異構字。
	中研院金文重文一	6,217	hzcdp01b.ttf	收錄《金文編》24,261 個金文，由於編碼的限制，所以分成四個字型。
	中研院金文重文二	6,217	hzcdp02b.ttf	
	中研院金文重文三	6,217	hzcdp03b.ttf	
	中研院金文重文四	5,610	hzcdp04b.ttf	
楚系簡帛文字	中研院楚系簡帛文字	2,627	cdpchubs.ttf	收錄《楚系簡帛文字編》的異構字
	中研院楚系簡帛文字重文一	6,217	hzcdp01cbs.ttf	收錄《楚系簡帛文字編》19,250 個楚文字，由於編碼的限制，所以分成四個字型。
	中研院楚系簡帛文字重文二	6,217	hzcdp02cbs.ttf	
	中研院楚系簡帛文字重文三	6,217	hzcdp03cbs.ttf	
	中研院楚系簡帛文字重文四	599	hzcdp04cbs.ttf	
小篆	北師大說文小篆	11,100	bnucdp.ttf	北京師範大學研發，中研院資訊所重新編碼。

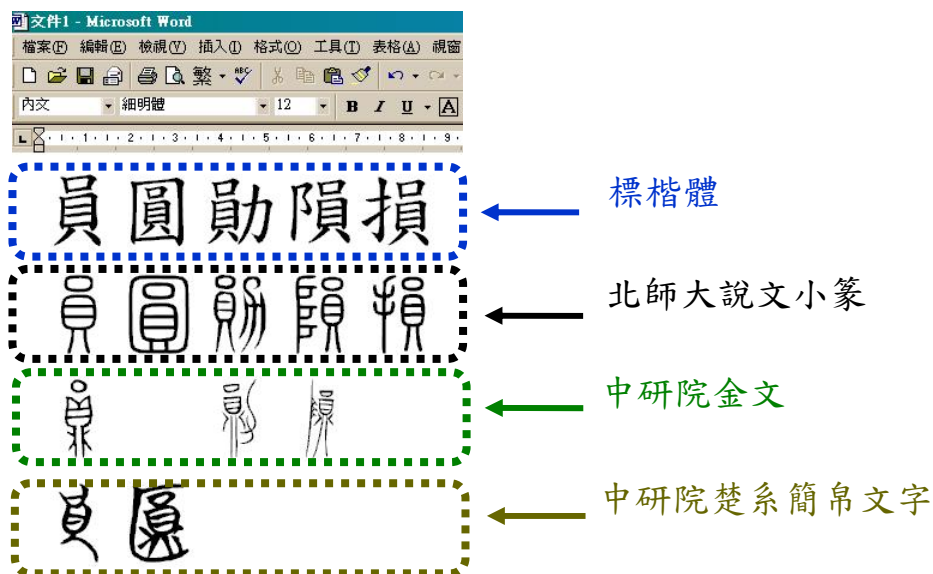


圖 4.4：使用古漢字字型

古漢字字型中的「北師大說文小篆」、「中研院楚系簡帛文字」、「中研院金文」、「中研院甲骨文」，可在 Microsoft Word 以替換字型的方式直接使用。例如圖 4.4，第一排為標楷體，第二排為將字型替換至北師大說文小篆，第三排替換至中研院金文，第四排替換至中研院楚系簡帛文字。從圖 4.4 中可以看出當將楷書的字型切換至古漢字字型時，有些字在古漢字中不一定有相對應的字形，因此會顯示為空白。

4.5 使用界面

系統安裝成功之後，在「開始」功能表中可找到「缺字公用程式」，並由此啟動「漢字構形資料庫」，如圖 4.5，此即為漢字構形資料庫使用界面，對應的檔案為 cdphanzi.exe。使用界面是採用 Visual Basic 6.0 開發。

漢字構形資料庫使用界面的主功能表分成「字集」、「字形」、「部件」及「編輯」等次功能表。

字集次功能表如圖 4.6，在進行部件檢字之前，應挑選適當的字集。例如挑選「常用字」來檢索楷書的常用字，或是挑選「說文解字」來檢索小篆。

字形次功能表如圖 4.7，包含部件檢字、出處檢字、字形結構、字形演變、字形索引及異體字表，這些是漢字構形資料庫使用界面的最主要功能，其中以部件檢字及異體字表最為重要。

楷書部件檢字的使用範例可參考第一章的圖 1.2、圖 1.3、圖 1.4，古漢字部件檢字的範例可參考圖 1.8、圖 1.17、圖 1.18、圖 1.19 及圖

1.20。

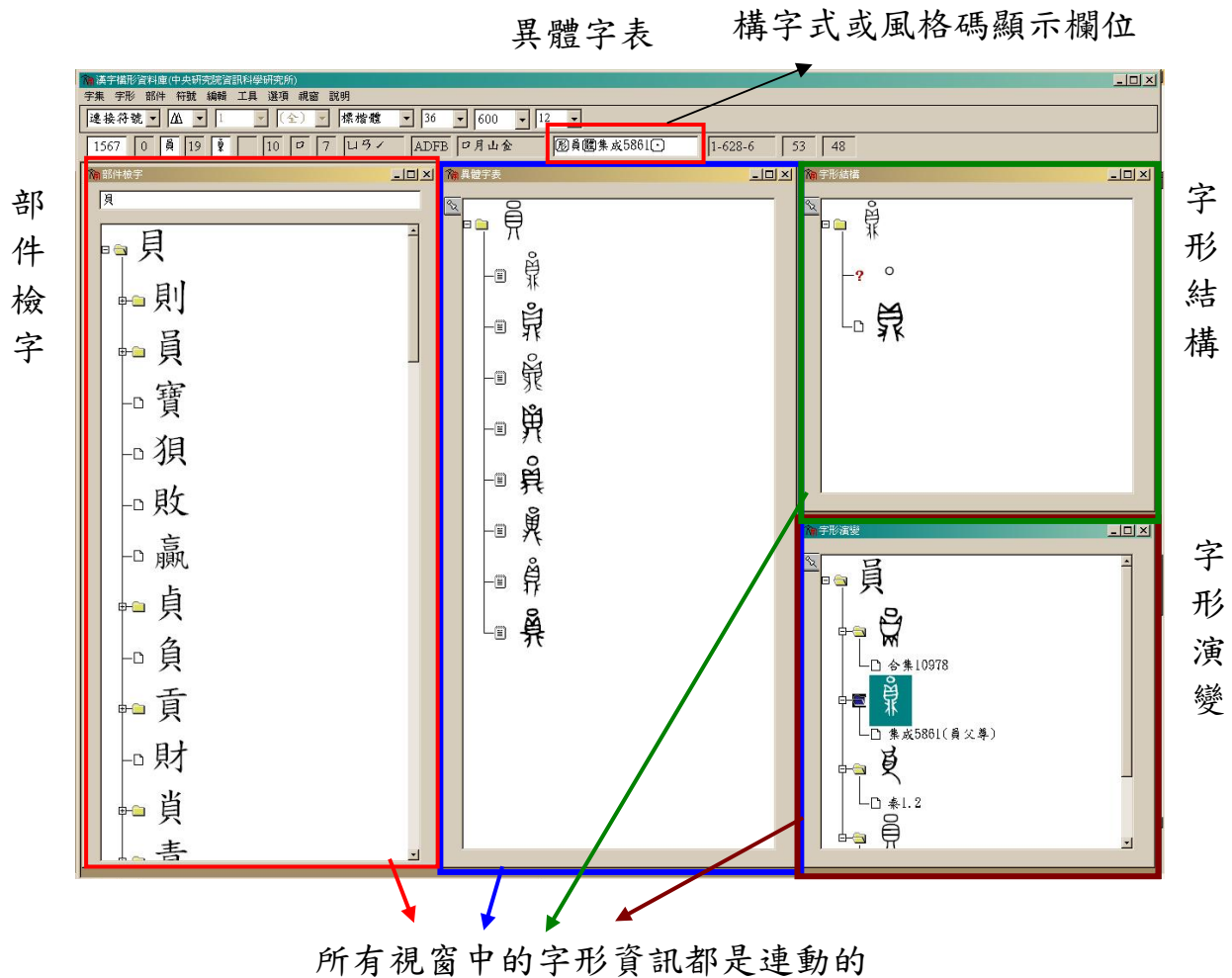


圖 4.5：漢字構形資料庫的使用界面

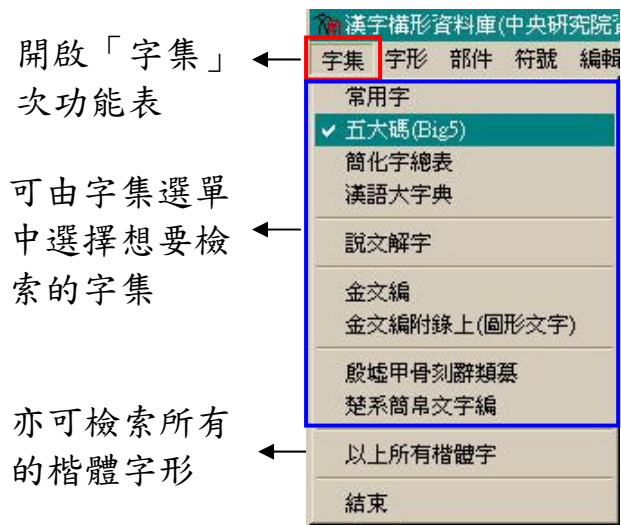


圖 4.6：漢字構形資料庫字集次功能表

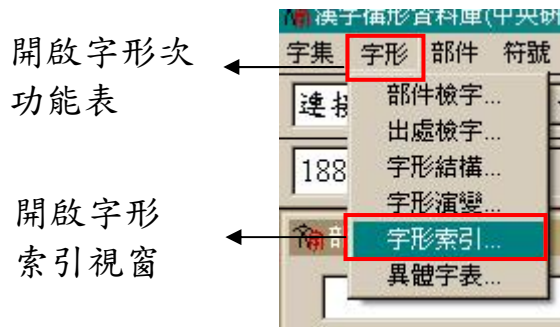


圖 4.7：漢字構形資料庫字形次功能表

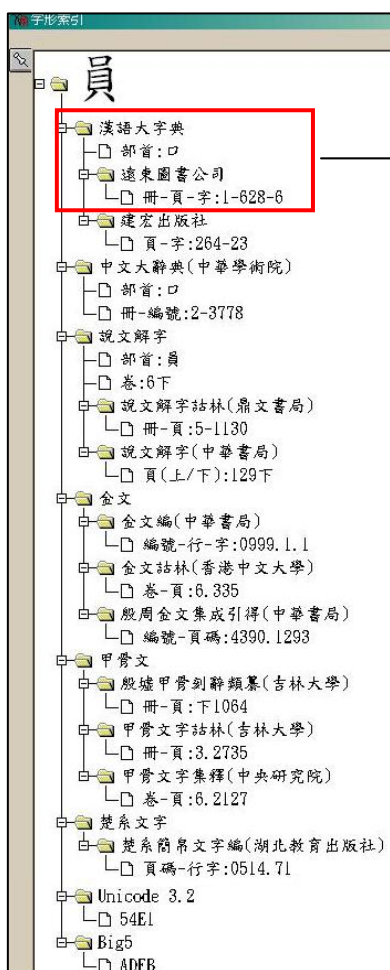
楷書字形結構的範例可參考第一章的圖 1.1，古漢字字形結構範例可參考圖 1.9、圖 1.13、圖 1.14、圖 1.15 及圖 1.16。

楷書異體字表的範例可參考第一章的圖 1.7，古漢字異體字表的範例可參考圖 1.10、圖 1.22、圖 1.23、圖 1.24、圖 1.25 及圖 1.26。

字形演變的範例可參考第一章的圖 1.11、圖 1.12。

出處檢字的範例可參考第一章的圖 1.27。

字形索引的內容詳見第一章的表 1.2 的字書索引及其他屬性，範例如圖 4.8。



「員」字在遠東圖書公司出版的《漢語大字典》中的索引資料




圖 4.8：漢字構形資料庫的字形索引

編輯	工具	選項	視窗	說明
複製				Ctrl+C
複製圖片				
複製到Microsoft Word				Ctrl+W
貼上				Ctrl+V
設定顯示字型...				
設定複製圖片...				

圖 4.11：漢字構形資料庫編輯次功能表

選項	視窗	說明
部件檢字	▶	▼ 24 ▼
字形結構	▶	△
字形索引	▶	0
字形演變	▶	字形結構
複製到Microsoft Word	▶	複製成字形或構字式
其他	▶	複製成圖片
		✓ 智慧型複製
智慧型複製可將尚未編碼的古漢字複製成圖片		

圖 4.12：漢字構形資料庫選項次功能表

圖 4.5 的漢字構形資料庫界面的每個視窗中都有不同的漢字資訊，且彼此之間是連動的。當使用者在「部件檢字」輸入某一部件進行檢字，並在其檢字結果中點選某一字形，此時其他視窗內的字形資訊也會同時進行變動。例如當使用者使用「貝」進行部件檢字，並在出現的楷書字形結果中，點選「員」字，此時即會在其他視窗出現「員」的楷書字形資訊。若使用者此時再點選「字形演變」視窗中的「員」字金文字形，在「異體字表」與「字形結構」視窗中即會同步顯示「員」字金文字形的相關資訊。在整個漢字構形資料庫界面的上端還有一欄是構字式及風格碼的顯示欄位，從該欄位中可看到我們點選的「員」字金文字形之風格碼為「員集成 5861 」。

4.6 缺字增益集

在 Microsoft Windows 下，Microsoft Office 不但是使用最廣泛的應用軟體，同時還可充當應用軟體的開發平台，安裝漢字構形資料庫之後，即會在 Microsoft Word 建立缺字增益集，這個增益集包含「構字符號」工具列及缺字處理次功能表。

「構字符號」工具列如圖 4.13，可協助使用者輸入構字符號。缺

字處理功能表如圖 4.14，它是位於「工具」功能表之下，由「構字式轉成字形」、「構字式轉成圖片」、「圖片轉成構字式」、「開啟漢字構形資料庫」及「漢字構形資料庫說明」五個項目組成。



圖 4.13：構字符號工具列

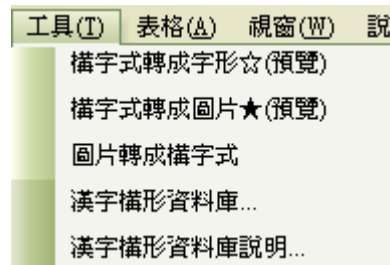


圖 4.14：缺字處理功能表

如上一節所述，在漢字構形資料庫查詢的字形都有相應的構字式或風格碼。在 Word 中缺字的輸入，可以構字式或風格碼來表達。使用者輸入構字式或風格碼的文件，稱為「原始文件」。執行「構字式轉成字形」或「構字式轉成圖片」，可將構字式或風格碼轉成字形或字形圖片，並存為新增的文件，這個新增的文件稱為「缺字預覽文件」。如圖 4.15 中，左邊的文件為原始文件，缺字使用構字式表達；右邊的文件則為經過缺字增益集將構字式轉換成字形，而新增的缺字預覽文件。

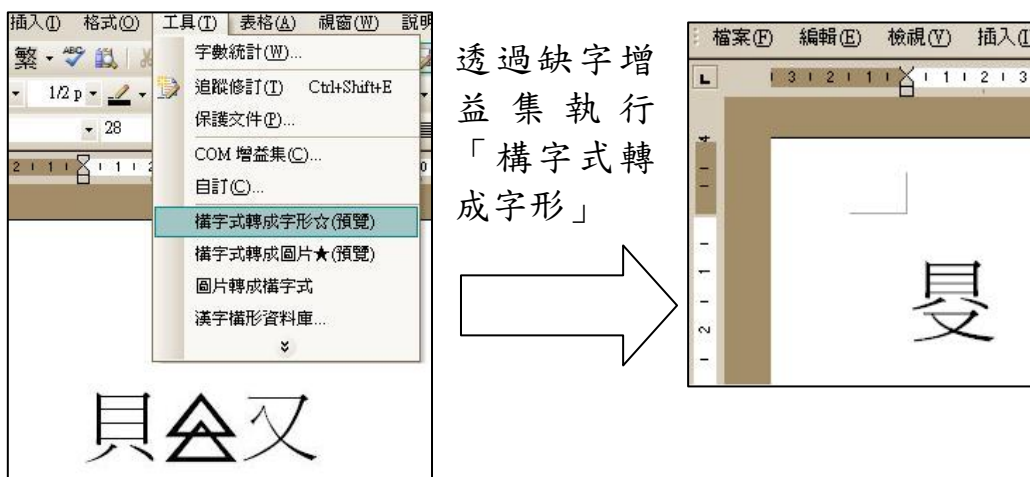


圖 4.15：透過缺字增益集顯示缺字

在缺字增益集當中，「將構字式轉成字形」與「將構字式轉成圖片」這二個預覽功能，都可幫助將嵌有構字式之原始文件轉成適合閱讀和列印的缺字預覽文件，差別只在於「構字式轉成字形」是將構字式表達的缺字以字型檔的方式顯示，而「構字式轉成圖片」是將構字式表達的缺字以圖片的方式顯示。因此以「構字式轉成圖片」方式產生的缺字預覽文件，不但可以提供閱讀和列印，同時也適用於檔案的

交換。即使另一方的電腦中沒有安裝漢字構形資料庫的字型檔，同樣也可顯示缺字字形的圖片，同時每個字形圖片皆已內嵌相對應的構字式資訊，如圖 4.16 所示。

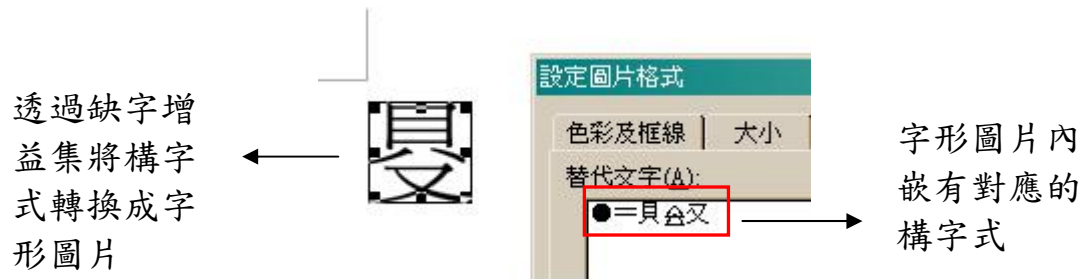


圖 4.16：透過缺字增益集顯示缺字圖片

由於構字式可當交換碼，所以原始文件適合交換及修改。缺字預覽文件只供閱讀及列印，修改時必須回到原始文件進行。為了讓使用者更便利編輯與管理文件，在 2007 年 8 月釋出的漢字構形資料庫 2.5 版中，新增了自動貼圖至 Microsoft Word 的功能，使用者在漢字構形資料庫中看到的任何字形，都可以利用快速鍵 Ctrl-W，將字形圖片自動貼至 Microsoft Word 中，字形圖片之大小則與欲編輯之 Microsoft Word 文件中的字級大小一致，如圖 4.17。

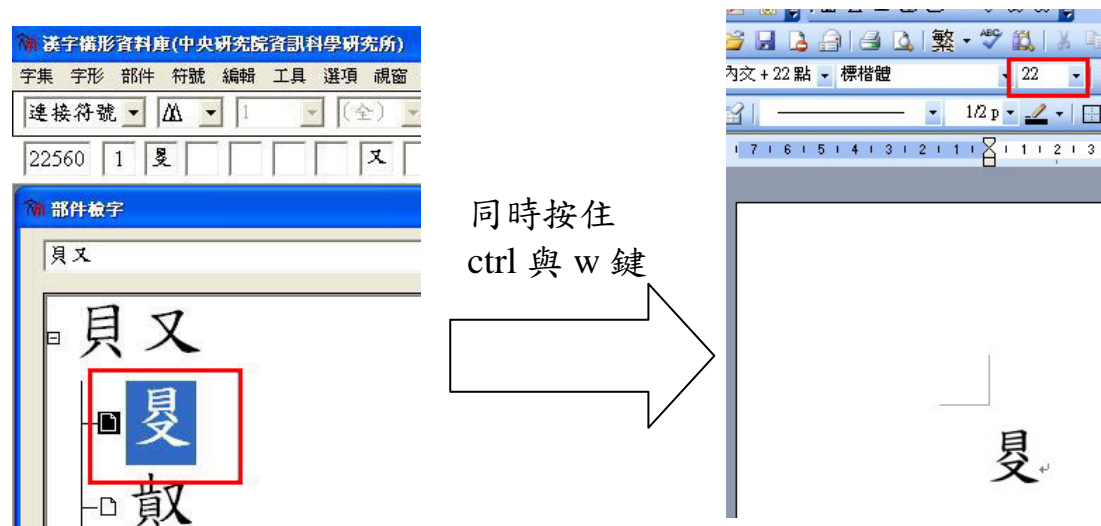


圖 4.17：利用快速鍵自動貼圖至 Microsoft

無論是由缺字增益集中的將「構字式轉成圖片」，或是在漢字構形資料庫的使用界面上，使用快速鍵貼圖至 Microsoft Word 文件，這些貼至 Microsoft Word 的字形圖片都嵌有對應的構字式資訊，所以使用者可以再利用缺字處理次功能表中的「圖片轉成構字式」將字形圖片轉回構字式，並可在嵌有構字式的原始文件中，以構字式或部件進

行查詢。

4.7 網頁缺字

網頁缺字技術目前由數位典藏技術發展組支援，詳細的說明可參考數位典藏技術發展組設計的缺字系統網站

(<http://char.ndap.org.tw/index.htm>)。

當使用者讀取含有構字式及技術發展組所提供之 JavaScript 程式碼的網頁時，其缺字顯示流程如圖 4.18 所示。

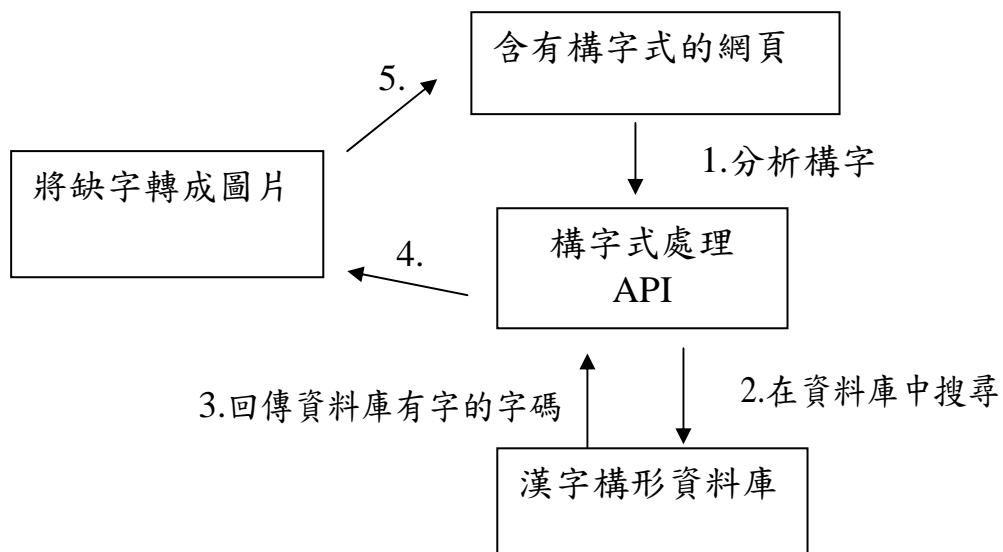


圖 4.18：網頁缺字呈現技術

網頁中的構字式經由放置於主機的構字式處理 API 之分析，並在資料庫中搜尋之後，如果有相對應的資料，即會從資料庫中傳出字碼，經由構字式處理 API 將字碼轉成圖片，並傳到 Client 端的瀏覽器。其具體的建置網頁缺字方式如下：

1. 網頁建置者如果想要在網頁中運用漢字構形資料庫所提供之缺字圖片時，在輸入資料時，必須先以構字式表達電腦缺字，如圖 4.19，此畫面即是含有構字式之網頁資料在未上傳之前的形式。

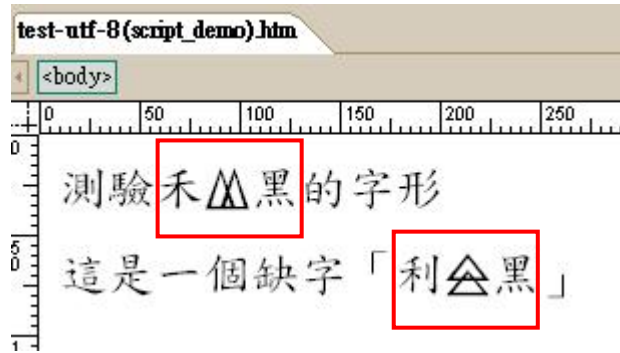


圖 4.19：以構字式表達網頁上的缺字字形

2.接著在網頁的程式碼中加入 JavaScript 程式碼。首先在<head>與</head>之間加入以下的程式碼，如圖 4.20。

```
<SCRIPT src="http://char.ndap.org.tw/API/ics.js"
language="javascript"></SCRIPT>
```

```

1 <html>
2
3 <head>
4
5 <meta http-equiv="Content-Type" content="text/html; charset=utf-8">
6
7
8 <SCRIPT src="http://char.ndap.org.tw/API/ics.js" language="javascript"></SCRIPT>
9
10 <title>test</title>
11 </head>
```

圖 4.20：網頁缺字需加入之程式碼

3.之後在</html>之後加入呼叫轉換整個頁面構字式的函式，如圖

```

67 </body>
68 </html>
69
70
71
72 <SCRIPT LANGUAGE="JavaScript">
73 <!--
74
75 processPage('red','20','','','DFKai-sb');
76 //-->
77 </SCRIPT>
```

圖 4.21：轉換整個頁面的構字式

4.21。

```
<SCRIPT LANGUAGE="JavaScript">
<!--
processPage('Red','20','','','DFKai-sb');
//-->
```

</SCRIPT>

此一函式可供使用者變化整個頁面上欲顯示缺字的顏色、字形大小及字體（目前可支援細明體 Mingliu 與標楷體 DFKai-sb）。例如圖 4.21 中函式 `processPage('red','20','','','DFKai-sb')`，括弧中的第一個參數代表欲顯示的缺字字形顏色為紅色；第二個參數代表字形大小為 20 點；第五個參數代表字體為標楷體。

如果使用者在網頁上有許多區塊，而各個區塊的缺字字形都有不同的外觀設定時，亦可呼叫轉換區塊缺字的函式，其置於</html>之後的程式碼代換如下：

```
<SCRIPT LANGUAGE="JavaScript">
<!--
processObject(document.getElementById('t4'),'black','18','','','DFKai-sb');
//-->
</SCRIPT>
```

此一函式可供使用者變化網頁中不同區塊欲顯示缺字的顏色、字形大小及字體，其參數設定除了需加上顯示缺字的區塊名稱之外，其餘設定與轉換整個頁面的函式設定相同。例如圖 4.22 中，代表網頁中名為 t4 的區塊裡的缺字字形，字形顏色為黑色；字體大小為 18 點；字體為標楷體。

```
67 </body>
68 </html>
69
70
71
72 <SCRIPT LANGUAGE="JavaScript">
73 <!--
74 processObject(document.getElementById('t4'),'black','18','','','DFKai-sb');
75 processPage('red','20','','','DFKai-sb');
76 //-->
77 </SCRIPT>
78
```

圖 4.22：轉換區塊的構字式

網頁資料輸入完成後，即可上傳至網路，當瀏覽網頁者開啟網頁後，網頁中的構字式在經過缺字處理 API 處理之後，其缺字字形會以圖片的形式傳至讀取網頁的瀏覽器，其顯示結果如圖 4.23。

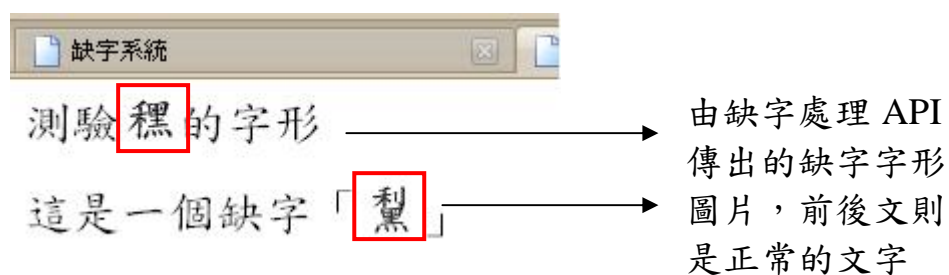


圖 4.23：網頁缺字的顯示結果

第五章 漢字構形資料庫的展望

漢字構形資料庫研發至今，已有超過 10 年的時間，收錄的古今漢字也將近十二萬個，並成功地解決中央研究院漢籍電子文獻資料庫及數位典藏與數位學習國家型科技計畫的缺字問題。然而，漢字構形資料庫不論在資料庫本身或推廣方面，仍有很大的改善空間，短期內研發漢字構形資料庫的網路版勢在必行；然而無論在推廣或內容的擴充方面，建立文字學入口網站才是長期努力的目標。

5.1 漢字構形資料庫的應用現況

利用漢字構形資料庫開發的漢字部件檢字系統，可用來解決缺字問題。經過這些年來的驗證，這個缺字問題的解決方案，基本上是可行的。目前數位典藏與數位學習國家型科技計畫至少已有 23 個資料庫採用漢字構形資料庫來解決缺字問題，其中規模最大的為中央研究院漢籍電子文獻資料庫，漢籍電子文獻已上線古籍達三億五千八百萬字。在解決缺字問題方面，除了持續擴充漢字構形資料庫外，我們還進行下列工作：

1. 以公眾授權模式提供「中研院漢字部件檢字系統」的原始程式碼及相關資料，釋出給大眾使用。⁸
2. 推動漢字構字標準，撰寫「中文字構形識別序列」標準草案，並函送標檢局審議。

雖然歷經十餘年的開發，漢字構形資料庫不論在資料庫本身或推廣方面，仍有很大的改善空間：

1. 漢字有形、音、義三要素，漢字構形資料庫目前只著重在字形知識的整理，尚未涉及字音與字義；建立一個形、音、義俱備的漢字知識庫，仍是我們長遠的目標。
2. 漢字構形資料庫的缺字解決方案，採用單位仍侷限於中央研究院及數位典藏與數位學習國家型科技計畫，仍須再推廣。
3. 部件檢字雖然比部首檢字便利，比筆畫檢字有效率，但仍應再增加部首、筆畫等其他檢字方式，以求完備。
4. 漢字構形資料庫雖可透過網際網路免費下載⁹，然而由於檔案太大（約 84M），而且只能在繁體中文版下使用，以致下載人次增加有限。目前每天的下載人次約 10 次，總下載人

⁸請參考 <http://cdp.sinica.edu.tw/cdphanzi/declare.htm>

⁹漢字構形資料庫下載網址：<http://cdp.sinica.edu.tw/cdphanzi/>

次為 15,788 次（截至 2009 年 7 月 27 日止）。若要增加使用人次，漢字構形資料庫應再開發中文簡體、英文、日文、韓文等各國語言版，或是網路應用版。

上述問題的解決關鍵，其實就在網際網路。短期內可先研發漢字構形資料庫的網路版以解決缺字、檢字及下載安裝的問題，然後再將網路版擴充成文字學入口網站，以建立形、音、義俱備的漢字知識庫，同時也能更有效地推廣漢字構形資料庫。

5.2 建立文字學入口網站

文字學入口網站主要是利用漢字的形、音、義知識來檢字，並提供漢字字典網站的連結。這樣一個入口網站不同於 Google、Yahoo 等用關鍵詞來找網頁的入口網站，而這樣的需求也非 Google、Yahoo 等入口網站所能達成。規畫中的文字學入口網站應具有下列特色：

1. 多語的使用界面。作為一個入口網站，多語的界面是必須的，除了繁體、簡體中文、英文外，至少也須再兼顧其他漢字區的使用者，因此日文及韓文界面也須列入考量。
2. 多樣的檢字方式。漢字的形、音、義、字碼等屬性都可用來檢字。字形檢字包括部首、筆畫、部件等，字音檢字包括注音、拼音、聲韻等，字義檢字包括英文、日文、韓文等，字碼檢字包括 Big5、Unicode、CCII、CNS11643 等。
3. 可檢索不同字集。入口網站應可讓使用者檢索不同的字集，這些字集包括常用字、簡化字，甚至甲骨文、金文、小篆等古漢字。
4. 可解決缺字問題。缺字指的是中文字碼裡沒有收錄的字，例如 Big5 收錄了 13,051 個字，13,051 之外的字即是缺字；現在常用的 Unicode 隨著不同版本字數雖然有所不同，少者 20,902 個字，多者可到 70,194 字，但是缺字問題仍然存在。無論使用者電腦中文字碼收錄字數的多寡，即使未安裝漢字字型，入口網站都應能正常顯示古今漢字。
5. 提供字典網站連結。目前字典網站彼此間的連結不夠，若要多瀏覽幾個網站，常需重複進行檢字，這對使用者而言並不方便。入口網站應在使用者找到所要的字後，提供現有的字典網站連結，以節省重複檢字的時間，並提供使用者更多的漢字參考資訊。

具有上述特色的文字學入口網站，其效益如下：

1. 推動字書上網。入口網站提供的檢字及缺字技術可降低字典上網的門檻，而入口網站提供的連結又可增加字典的能見度，如此不但可增加現有字典上網的意願，同時也有助於線上字典的編纂。
2. 建立漢字知識庫。古今字書的數位化工作，實非個別單位能獨立完成，如今正可透過網際網路集合眾人之力共同推動並予以整合，以建立完善的漢字知識庫。
3. 便利漢字知識的擷取。目前的入口網站主要透過關鍵字及關鍵詞來搜尋網頁，以漢字為主的網頁不易突顯；相形之下，文字學入口網站則只提供這些網頁的連結，因此更有利於使用者擷取到這些漢字知識。
4. 闡揚漢字文化。入口網站所提供的連結，除了字典網站外，也包括其他和漢字相關的文化、藝術網站，這些連結及網站可用來闡述及發揚漢字文化。

參考資料

參考文獻

1. 中國社科院考古所，《殷周金文集成》，中華書局，1984-1994年
2. 王寧，《漢字漢語基礎》，北京科學出版社，1996年7月
3. 周法高等，《金文詁林》，香港中文大學，1974年
4. 姚孝遂，《殷墟甲骨刻辭類纂》，中華書局·北京，1989年
5. 容庚，《金文編》，中華書局，北京，1985年7月
6. 徐中舒，《遠東·漢語大字典》，遠東圖書公司，1991年9月
7. 張亞初，《殷周金文集成引得》，中華書局，北京，2001年7月
8. 莊德明、許婉蓉，《漢字構形資料庫使用手冊》，中研院資訊所，台北，2002年7月
9. 莊德明、許永成、謝清俊，如何使用電腦處理古今文字的銜接——以小篆為例，第十四屆中國文字學全國學術研討會，高雄，2003年3月
10. 莊德明，《漢字構形資料庫的建置與應用》，漢字與全球化國際學術研討會，台北，2005年1月
11. 莊德明，〈如何利用風格碼解決古漢字的編碼問題〉，第三屆中國文字學國際學術研討會，河南開封，2006年8月
12. 莊德明，〈漢字數位化的困境及因應：談如何建立漢字構形資料庫〉，第五屆兩岸三院資訊交流與數位資源共享研討會，台北，2007年10月
13. 莊德明、鄧賢瑛，〈文字學入口網站的規畫〉，第四屆中國文字學國際學術研討會，山東煙台，2008年8月
14. 楊家駱，《說文解字詁林》，鼎文書局，1994年3月
15. 滕壬生，《楚系簡帛文字編》，湖北教育出版社，1995年7月
16. 謝清俊，〈電子古籍中的缺字問題〉，第一屆中國文字學會學術討論會，天津，1996年8月
17. 謝清俊，〈談中國文字在電腦中的表達〉，「中國文字的未來」學術研討會，台北，1991年6月

參考網站

1. 漢字構形資料庫下載網址 <http://cdp.sinica.edu.tw/cdphanzi/>
2. 缺字系統首頁 <http://char.ndap.org.tw/>