

## 基於數據庫平臺的文字編編撰和自動化發佈<sup>①</sup>

郭瑞

(華東師範大學中國文字研究與應用中心, 中國上海)

**摘要:** 文字編, 或字形表, 是在特定材料範圍內對該材料中的漢字字形進行整理研究, 並按照一定體例要求編纂彙集而成的一種字形工具書。傳統的整理手段, 形成一部文字編需要投入相當的精力, 涉及考釋、篩選、切圖、粘貼等等繁瑣的流程, 全部手工處理, 除了要求一定的專業基礎, 還需要大量的機械勞動。有些文字編, 甚至送手工摹寫。但是, 隨著現代電腦技術的飛速發展, 特別是“數據庫”這個功能強大的數字化存儲、分析平臺的出現, 給古老的文字學注入了新的活力, 推動這一學科在新的世紀與時俱進的發展。基於一個完善的數據庫平臺, 再進行文字編的工作, 效率將提高數倍。最重要的是, 繁重的純機械工作可以交由電腦來處理, 把專家學者們從繁重的體力勞動中解放出來, 更多的精力投入到研究本體工作中, 不僅對專家們智慧的尊重, 也是保證文字學科的可持續發展的有效途徑。

**關鍵字:** 數據庫; 文字編; 自動化

文字編, 或字形表, 是在特定材料範圍內對該材料中的漢字字形進行整理研究, 並按照一定體例要求編纂彙集而成的一種工具書。其在形式和內容上都不同於傳統意義上字書、辭書, 是和出土文獻有著密切關係的一種獨特的工具書類。文字編的主要特點是以直觀形式展示出該類材料中的漢字字形的變化及關係。形式上, 通常用《說文解字》小篆或通用楷體字作為字頭標目, 然後將材料內的字形羅列於相應的字頭下, 字形下給出具體出處, 有些會給出簡短語境。內容上, 或將材料中漢字窮盡性收錄, 或選擇差異性字形收錄。

但凡一種出土文獻問世, 除了文本內容的考釋工作以外, 編撰一部文字編也是一件必不可少的工作。諸如:《甲骨文編》(孫海波, 中華書局, 1965年)、《金文編》(容庚, 中華書局, 1985年)、《郭店楚簡文字編》(張守中, 文物出版社, 2000年)、《睡虎地秦簡文字編》(張守中, 文物出版社, 1994年)、《武威儀禮漢簡文字編》(徐富昌, 國家出版社, 1995年)、《馬王堆簡帛文字編》(陳松長, 文物出版社, 2001年)、《銀雀山漢簡文字編》(駢宇騫, 文物出版社, 2001年)、《張家山漢簡文字編》(張守中, 文物出版社, 2012年)、《清華大學藏戰國竹簡(壹—叁)文字編》(李學勤, 中西書局, 2014年)、《戰國文字編》(湯餘惠, 福建人民出版社, 2001年)、《漢代銅器銘文文字編》(徐正考, 作家出版社, 2007年)等等, 其徒實繁。傳統的整理手段, 形成一部文字編需要投入相當的精力, 涉及考釋、篩選、切圖、粘貼等等繁瑣的流程, 全部手工處理, 除了要求一定的專業基礎, 還需要大量的機械勞動。有些文字編, 甚至用手工摹寫。比如:較古的《隸韻》(宋·劉球)、《隸辨》(清·顧藹吉);現代的《金文編》、《睡虎地秦簡文字編》等, 即使寫者水平很高, 也難免有失真的情況。且純手工篩選整理, 不可避免的出現遺漏, 這也就是在一部文字編問世一段時間後往往還會有續編出版的原因。

然而, 我們也應該看到隨著現代電腦技術的飛速發展, 特別是“數據庫”這個功能強大的數字化存儲、分析平臺的出現, 也不斷衝擊著傳統學科, 給古老的文字學注入了新的活力, 推動這一學科在新的世紀與時俱進。基於一個完善的數據庫平臺, 再進行文字編的工作, 效率將提高數倍。最重要的是, 繁重的純機械工作可以交由電腦來處理, 把專家學者們從繁重的體力勞動中解放出來, 更多的精力投入到研究本體工作中, 不僅對專家們智慧的尊重, 也保證文字學科的可持續發展。

<sup>①</sup> 中國國家社科基金重大項目“秦漢六朝字形全譜(13&ZD131)”成果。

筆者所在的中國教育部人文社會科學重點研究基地華東師範大學中國文字研究與應用中心，經過數年的科研實踐探索，已經逐漸形成了一套比較成熟的文字編類成果研製的程序，積累成功的經驗。最新推出的一批文字編類成果：《新金文編》（董蓮池，作家出版社，2010年）、《魏晉南北朝隋唐五代字形表》（臧克和，南方日報出版社，2008年）、《中國文字異體字大系·篆書編》（劉志基，上海書畫出版社，2007年）、《中國文字異體字大系·隸書編》（臧克和，上海書畫出版社，2010年）、《中國文字異體字大系·楷書編》（王平，上海書畫出版社，2008年）、《日藏唐代漢字抄本字形表》（臧克和，華東師範大學出版社，2015年）等。

筆者結合曾經參與的一些具體工作，針對新技術手段下文字編類成果實現的步驟、關鍵問題等內容做一下匯總，與諸位專家學者分享，以期對大家的研究工作能夠起到一定的幫助。

## 一、 計算機技術支持下的文字編類成果的編撰步驟

第一步：準確的釋文和最好的圖片資料。

釋文是文字編的基礎，釋文的準確與否直接關係到文字編字形歸屬的正確與否。而高品質的圖片資料，保證文字編字形的清晰準確，提升使用價值。因此，通常情況下要在所整理資料的範圍內做到全面、真實，並隨時關注相關材料的最新考釋信息，及時對材料進行補充更新，盡可能保證釋文的準確性，並能反映出最新的研究成果。同時，在窮盡性搜集材料的基礎上，校重查漏，“同”中選優，把最好的圖片資料甄選出來。

第二步：數據庫平臺建設。

基於釋文和圖片資料，根據文字編編撰的實際要求，建設一個內容準確、功能完備的數據庫，是實現文字編自動化發佈的重要平臺。這一步也是減輕機械勞動負擔的關鍵所在。

我們主要採用微軟辦公軟件自帶的數據庫 Microsoft Office Access 2003 和 SQL Server 2005 數據庫相結合的研發平臺。Access 數據庫屬於小型數據庫，具有安裝便捷、容量較小、操作直觀的特點，適用於不同時期不同類型的文字材料的單獨處理。待各個獨立的數據庫完成後，再通過統一的標準整合到 SQL Server 數據庫平臺，該數據庫也是美國微軟公司推出的一種關係型數據庫系統，屬於大型數據庫系統。是一個可擴展的、高性能的、為分散式客戶機/伺服器計算所設計的數據庫管理系統。並通過 Web 服務系統向用戶端開放訪問。

通常某種材料的數據庫平臺應該包括四張表：文獻表、字形表、單字表、部首表。

(1)文獻表，主要儲存相關資料的文獻屬性信息，不同種類的資料所需要設置的文獻屬性字段是有差別的，但是有幾個字段是共性的，包括：時代、著錄信息、釋文、對應圖片編號、資料類型。

(2)字形表，以資料中的每一個漢字為一條記錄，並記錄該字的各類屬性信息。字形表是基於文獻表的釋文，運用電腦編程逐字分解出來的。字形表字段要包括：通用字、漢字位置信息（該漢字在原始材料中的具體位置，包括篇號、列號、字序）、字形編號（以“篇號-列號-自序”這樣的形式記錄）、字形圖片的位置信息（該字形在原始圖片上的座標位置）。這些是共性字段，不同的材料還會有些特定的字段信息，比如：字體標記、異體字、假借字、語境片段、字形頻率、出處等等。

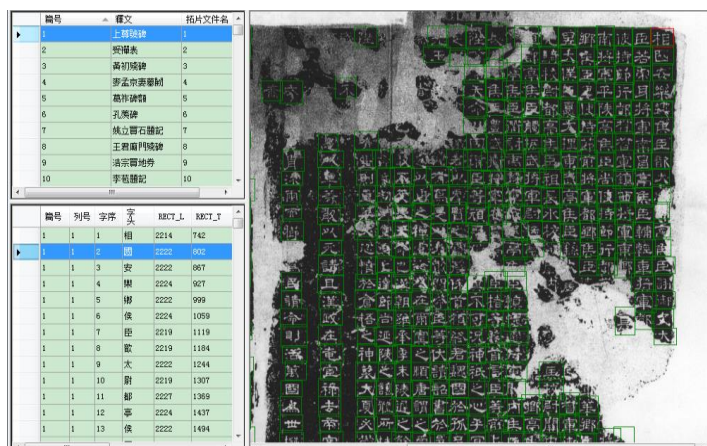
(3)單字表，是字形表“通用字”字段去重後得到一個字表，反應的是該材料中實際漢字字量，也是文字編字頭標目和索引編制的來源。因此，單字表必須包括：通用字頭（或《說文》楷字）、部首、筆劃數；根據需要有些還要有：說文小篆、讀音、前三筆筆劃類型、異體字等等。

(4)部首表，所有單字所涉及的部首類型。目前的文字編主要採用兩種部首分類，一種是以《說文》540部為本；一種是以現代工具書編撰的部首為依據，通常採用《漢語大字典》（第二版）的200部。該表也主要是為了索引編排的需要。

第三步：字形圖片切割和儲存。

這一步也是相當重要的步驟，因為文字編是字形圖片彙編而成。因此，從完整的原始圖片中分

割出單個的字形圖片，就成為一個繁重而枯燥的機械操作，以往傳統方法就是耗時耗力，還不能避免缺漏、錯誤。針對這一難題，我們經過多年的摸索、改進，研製出了一種可以隨時釋文、圖片相互對照修改的切圖軟件，如圖 1：左上為釋文信息、左下為字形表信息，右側為對應拓片。只需要在右側圖片對應的字形上拉一個方框，該字形就會自動保存在電腦中指定的位置，字形圖片編號為：“篇號-列號-自序”的組合形式，簡牘類材料“列號”實際為“簡號”，如圖 2；字形圖片的位置信息自動保存到字形表相應字段中；同時字形表的游標跳至下一條記錄，繼續在圖片上拉方框選字即可。發現錯誤可以及時修改訂正。此外，利用該軟件的報表功能，可以對某字下的所有字形進行查看、篩選，如圖 3。



(圖 1：“字形切割軟件”，左上為釋文信息、左下為字形表信息，右側為對應拓片。)



(圖 2：所切下的字形，字形圖片的命名以篇號、列號、字序組成，便於查檢和調取。)



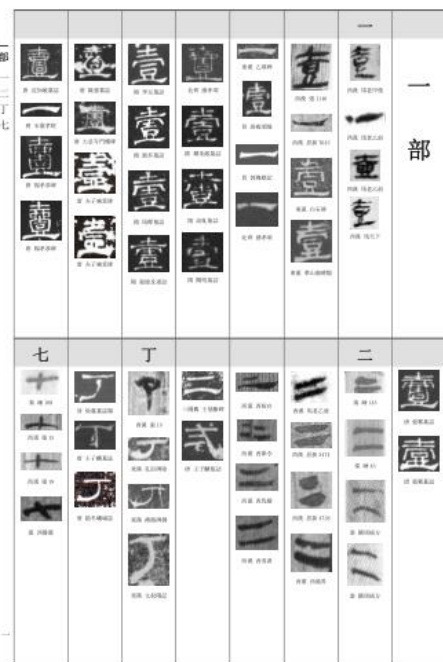
(圖3：“群”字下所列字形，“群”字右側數字為統計字頻。)

第四步：成果版式的確定。

關於版式，目前通常的形式為單欄雙行表格式，第一行為單字字頭，或選用通用字，或使用說文小篆的對應楷字；第二行列出該字頭下的所用字形，並附有該字形的出處信息，如《漢魏六朝隋唐五代字形表》，圖4。有時也採用上下雙欄雙行式，比如：《中國異體字大系》，圖5。



漢魏六朝隋唐五代字形表  
一四四



(圖4：《漢魏六朝隋唐五代字形表》) (圖5：《中國異體字大系·篆書編》)

有些特殊需要的會靈活增減，比如《日藏唐代漢字抄本字形表》，由於加注了日本讀音，所以變為了三行的編排形式，見圖6。

(圖 6:《日藏唐代漢字抄本字形表》)

第五步：相關程序的編寫。

利用電腦編程語言編寫發佈程序，實現文字編內容的自動化發佈，以及索引的自動生成，不僅僅大大提高了編寫效率，也提升了成果的準確率。

不同材料種類的文字編所需要編寫的程序有些差別，但是有更多內容是共性的。以《漢魏六朝隋唐五代字形表》發佈程序為例，做簡單介紹說明，由於篇幅限制下列程序有關屬性限定的程序作了省略。

1.正文內容的發佈程序：

(1) 定義變數，包括字頭、字形變數和版式變數

(略)

(2)發佈部首

```
Set rstbushou = db.OpenRecordset("select * from 汉语大字典部首表 where 部首编号=" & bshbh)
```

```
Do Until rstbushou.EOF
```

```
ipage = 1
```

```
ihangtou = 1
```

```
ihangxing = 2
```

```
ilie = 0
```

```
i = 1
```

```
With Word.ActiveDocument.Tables(1).Cell(2, 8 - ilie).Range
```

```
.InsertAfter Chr(13) & Chr(10) & rstbushou!bushou & Chr(13) & Chr(10) & Chr(13) & Chr(10) & "部"
```

(略)

```
.Collapse (wdCollapseEnd)
```

```
End With
```

```
ilie = ilie + 1
```

(3)发布字头

```
Set rstzitou = db.OpenRecordset("select * from danzi where buid=" & ibushou & " order by bihua")
```

```
Do Until rstzitou.EOF
```

```
    If ihangxing = 2 Then
```

```
        With Word.ActiveDocument.Tables(ipage).Cell(ihangtou, 8 - ilie).Range
```

```
            .InsertAfter rstzitou!zitou
```

```
            If rstzitou!ziku = "fz" Then
```

```
                .Font.Name = "宋体-方正超大字符集"
```

```
            Else
```

```
                .Font.Name = "宋体"
```

```
            End If
```

```
            .Font.Size = 24
```

```
            .Font.Bold = True
```

```
            .Collapse (wdCollapseEnd)
```

```
        End With
```

```
    ElseIf ihangxing = 4 Then
```

```
        With Word.ActiveDocument.Tables(ipage).Cell(ihangtou, 16 - ilie).Range
```

```
            .InsertAfter rstzitou!zitou
```

```
            If rstzitou!ziti = "fz" Then
```

```
                .Font.Name = "宋体-方正超大字符集"
```

```
            Else
```

```
                .Font.Name = "宋体"
```

```
            End If
```

```
            .Font.Size = 24
```

```
            .Font.Bold = True
```

```
            .Collapse (wdCollapseEnd)
```

```
        End With
```

```
    End If
```

(4)發佈字形位置信息

```
Set rstzixing = db.OpenRecordset("select * from chx_zixing where unicode="" & rstzitou!unicode & "" order by ziti,jianchen,id")
```

```
Do Until rstzixing.EOF
```

```
    szixing = Replace(rstzixing!dizhi, "zipu", "D:\字表图片")
```

```
    schuchu = rstzixing!shidai & " " & rstzixing!jianchen
```

```
    n1 = CInt(rstzixing!ziti)
```

```
    n = rstzixing!ziti2
```

```
    If i hangxing = 2 Then
```

```
        If n = "z" And n1 < 17 Then
```

```
With Word.ActiveDocument.Tables(ipage).Cell(ihangxing, 8 - ilie).Range
```

```
    .InsertAfter "【篆書】" & Chr(13) & Chr(10)
```

```
    .Collapse (wdCollapseEnd)
```

```
End With
```

```
ElseIf n = "l" And n1 > 16 And n1 < 30 Then
With Word.ActiveDocument.Tables(ipage).Cell(ihangxing, 8 - ilie).Range
.InsertAfter "【隸書】" & Chr(13) & Chr(10)
.Collapse (wdCollapseEnd)
End With
ElseIf n = "k" And n1 > 29 Then
With Word.ActiveDocument.Tables(ipage).Cell(ihangxing, 8 - ilie).Range
.InsertAfter "【楷書】" & Chr(13) & Chr(10)
.Collapse (wdCollapseEnd)
End With
End If
With Word.ActiveDocument.Tables(ipage).Cell(ihangxing, 8 - ilie).Range
.InsertAfter "▲" & szixing & "▲" '嵌入圖片用
End With
With Word.ActiveDocument.Tables(ipage).Cell(ihangxing, 8 - ilie).Range
.InsertAfter Chr(13) & Chr(10) & schuchu & Chr(13) & Chr(10)
.Collapse (wdCollapseEnd)
End With
End If
If i < 10 Then
i = i + 1
Else
ilie = ilie + 1
i = 1
End If
If ilie = 8 Then
ilie = 0
ipage = ipage + 1 '換頁
ihangtou = 1
ihangxing = 2
End If
```

(5)結束部分（略）

## 2.字形圖片處理宏代碼

### (1)嵌入字形圖片

```
For i = 1 To 10000
Selection.Find.ClearFormatting
With Selection.Find
.Text = "▲*▲"
.Forward = True
.Wrap = wdFindContinue
(略)
End With
```

(略)

```
Selection.Find.Execute
```

```
srep = Mid(Selection.Text, 2, Len(Selection.Text) - 2)
```

```
With Selection
```

```
    .Text = ""
```

```
    .InlineShapes.AddPicture FileName:=srep, LinkToFile:=True, SaveWithDocument:=False
```

```
End With
```

```
Next i
```

(2)調整圖片大小

' 同寬

```
For i = 0 To Word.ActiveDocument.InlineShapes.Count
```

```
    With Word.ActiveDocument.InlineShapes(i)
```

```
        iheight = .Height
```

```
        iwidth = .Width
```

```
        .LockAspectRatio = msoTrue
```

```
        If .Width > 45 Then
```

```
            .Width = 45
```

```
            .Height = iheight * (45 / iwidth)
```

```
        ElseIf .Width < 30 Then
```

```
            .Width = 30
```

```
            .Height = iheight * (30 / iwidth)
```

```
        End If
```

```
    End With
```

```
Next i
```

' 同高

```
For i = 0 To Word.ActiveDocument.InlineShapes.Count
```

```
    With Word.ActiveDocument.InlineShapes(i)
```

```
        iheight = .Height
```

```
        iwidth = .Width
```

```
        .LockAspectRatio = msoTrue
```

```
        If .Height > 45 Then
```

```
            .Height = 48
```

```
            .Width = 48 * iwidth / iheight
```

```
        End If
```

```
    End With
```

```
Next i
```

3.索引內容的發佈程序:

(1)部首索引

定義變量(略)

```
Set rst = db.OpenRecordset("select distinct bushoubihua from 汉语大字典部首表")
```



Do Until rst.EOF

    With Word.ActiveDocument.ActiveWindow.Selection

        .Text = Chr(13) & Chr(10)  
        .Font.Size = 10.5  
        .Collapse (wdCollapseEnd)  
        .Text = cc(rst!bushoubihua) & "畫" & Chr(13) & Chr(10)  
        .Font.Size = 14  
        .Font.Name = "黑體"  
        .Collapse (wdCollapseEnd)

    End With

Set rstbihua = db.OpenRecordset("select \* from 汉语大字典部首表 where bushoubihua=" & rst!bushoubihua & " order by bushouyema")

    Do Until rstbihua.EOF

        With Word.ActiveDocument.ActiveWindow.Selection

sbushou = rstbihua!bushou & "部"

        .Text = sbushou  
        .Font.Size = 11.5  
        .Font.Name = "宋體"  
        .Collapse (wdCollapseEnd)

        Select Case Len(rstbihua!bushouyema)

            Case 1

                .Text = "    " & "    " & cc(rstbihua!bushouyema) & Chr(13) & Chr(10)

            Case 2

                .Text = "    " & "    " & cc(rstbihua!bushouyema) & Chr(13) & Chr(10)

            Case 3

                .Text = "    " & cc(rstbihua!bushouyema) & Chr(13) & Chr(10)

            Case 4

                .Text = "    " & cc(rstbihua!bushouyema) & Chr(13) & Chr(10)

        End Select

        .Font.Spacing = -0.4  
        .Font.Name = "華文仿宋"  
        .Font.Size = 9  
        .Collapse (wdCollapseEnd)

    End With

rstbihua.MoveNext

Loop

rstbihua.Close

rst.MoveNext

Loop

結束部分（略）

(2)筆劃索引

定義變量 (略)

```
Set rstbihua = db.OpenRecordset("select distinct bihua from 笔画索引")
Do Until rstbihua.EOF
    With Word.ActiveDocument.ActiveWindow.Selection
        .Text = Chr(13) & Chr(10)
        .Font.Size = 10.5
        .Collapse (wdCollapseEnd)
        .Text = cc(rstbihua!bihua) & "畫" & Chr(13) & Chr(10)
        .Font.Size = 12
        .Font.Name = "黑体"

        .Collapse (wdCollapseEnd)
    End With
    Set rstzi = db.OpenRecordset("select * from 笔画索引 where bihua=" & rstbihua!bihua & "
order by bihua,yibi,erbi,sanbi,qubie,id")
    Do Until rstzi.EOF
        With Word.ActiveDocument.ActiveWindow.Selection
            .Text = Replace(rstzi!zitou, Chr(13), "")
            If rstzi!ziku <> "" Then ' "kaishu" Then
                .Font.Name = "lishuzi"
                .Font.Size = 10
            Else
                .Font.Name = "宋体"
                .Font.Size = 10
            End If
            .Collapse (wdCollapseEnd)
            Select Case Len(rstzi!zipage)
                Case 1
                    .Text = " " & " " & cc(rstzi!zipage)
                Case 2
                    .Text = " " & " " & cc(rstzi!zipage)
                Case 3
                    .Text = " " & cc(rstzi!zipage)
                Case 4
                    .Text = " " & cc(rstzi!zipage)
            End Select
            (略)
        End With
        rstzi.MoveNext
    Loop
    rstzi.Close
    rstbihua.MoveNext
Loop
```

結束部分(略)

## 二、編撰過程中的兩個關鍵問題

上述部分只是就文字編類成果自動化編排實現的具體操作程序作了羅列和分析。以上程序最終的目的，就是要把科研人員從這些機械性的勞動中解放出來，更專注於學術研究層面的工作。事實上，要保證這些程序執行結果的準確性，則依靠研究人員的研究工作，而這些工作真正是研究人員主要花費精力所在，也是體現研究人員科研水準和智慧的部分。具體來說就是原始資料釋讀問題和電子錄入的標準化問題。

### 1. 文本釋讀的準確性

這個問題的重要性，在上文的第一步中已經做了解釋。這裏，有必要交代一下我們對原始圖片掃描標準，經過一段時間的實驗，我們統一採用精益 A3 幅面高速掃描器，掃描規格為灰度、600dpi 精度、.jpg 格式儲存。這種條件下處理的原始圖片，既保證圖片處理的速度，便於電腦上的閱讀操作，又能使切割下的字形具有符合印刷要求的清晰度。

### 2. 文字處理的標準化

文字編往往面臨數量較大的出土文獻資料的加工，而且要依託數據庫平臺，要保證自動化生成的文字編內容準確，標準化工作尤為重要。主要涉及：數據庫建設的標準化、文本錄入的標準化、文字編字頭和字形選擇的標準。

#### (1) 數據庫建設的標準化

中國和相關行業關於數據庫建設發佈了一系列標準規範文件，可以作為數據庫建設、屬性標注體系的規範化、精細化的科學依據。這些文件法規主要包括：《GB 18030-2005 信息技術中文編碼字符集》《TR-REC-014\_資料集核心元資料標準》《TR-REC-032\_元資料訪問服務介面規範》《TR-REC-033\_資料跨域互操作技術規範》《GB/T20273-2006 信息安全技術 數據庫管理系統安全技術要求》《GB/T 13725—1992 建立術語數據庫的一般原則與方法》《GF 3001—1997 信息處理用 GB 13000.1 字符集漢字部件規範》《GB 18030—2005 信息技術 中文編碼字符集》等等。這些標準化文件雖然都由中國相關部門發佈的，但是很多技術層面的標準與國際接軌，這樣可以保證我們所創建的數據庫具有開放性、包容性的特點，不僅僅用於文字編的編撰使用，而且也可以作為其他相關學科研究的工作平臺。

例如我們所創建的《金文語料庫》《中古石刻文獻語料庫》等，既可作為發佈文字編的平臺，也可以在互聯網上共用使用。文字編只不是是數據庫平臺的一個最基本的功能而已。

#### (2) 文本錄入的標準化

文本錄入嚴格按照《中華人民共和國國家通用語言文字法》，參照中國語言文字工作委員會頒佈的系列法規文件，包括：《簡化字總表》(1986年6月24日)《通用規範漢字表》(2013年6月5日)《GF 1001—2001 第一批異形詞整理表(試行)》(2001年12月19日)《GF 2001—2001 GB13000.1 字符集漢字折筆規範》(2001年12月19日)等。當然，我們不能僵化的使用這些文件規定，由於所處理材料的不同，比如一些出土古文字文獻，既要盡可能的符合國家法規，也要顧及古文字學界約定俗成的一些習慣，比如對於古文字隸古定的字形處理。這些就要具體問題具體分析了，只要保證同類型文字材料處理標準的保持一致。

通常我們採用了“通用字(原始字形)”的錄入方式，比如：北魏武泰元年(528)《徐天寶造像記》中“夫至貌(𠄎)希(𠄎)微，非世賞之能(𠄎)模(摸)；真(𠄎)姿妙(讀為渺)滅，豈情近情之髣(𠄎)響。”圓括號內呈現的是原始字形信息，當然在數據庫中，這些異體字形並非以圖片形式出現，而是通過造字來實現的，即仍然是可以檢索到的電腦字符形式；另外具體的用字情況也可以在圓括號內顯示，例如“妙滅”的“妙”為“渺”字的通假字，所以在其後括弧內標注“(讀為渺)”。

便於讀者理解文義。

### (3)文字編字頭和字形選擇的標準

字頭標準的選擇，目前主要是兩種形式，一種是以《說文解字》小篆為字頭選擇標準，如：《新金文編》、《新甲骨文編》(劉釗，福建人民出版社，2009年)、《齊魯文字編》(張振謙，學苑出版社，2014年)、《三晉文字編》(湯志彪，作家出版社，2014年)等；一種以通用漢字(規範的繁體字)為字頭標準，如：《中國異體字大系》(篆書編、隸書編、楷書編)、《魏晉南北朝隋唐五代字形表》等。第一種常見於古文字類的文字編，屬於傳統做法，這樣古文字和小篆的形體對照更加直接一些，對於熟悉《說文》的專家查檢便利，但並非所有學者都能夠很清楚的記住《說文》部首、漢字的編排次序，這種情況下查找起來就比較費時了，有些文字編為了查檢便利，還附錄了楷字的筆劃檢字表，等於是對小篆、古文字進行了隸定，然後在按照筆劃編排檢字表；第二種常見於漢字隸變以後的文字編，直接按照通用漢字的部首、筆劃數進行編排，符合現代工具書編撰習慣，查檢便捷。為了清晰呈現漢字演變脈絡，還會在楷書字頭旁標出對應的《說文》小篆。

字形選擇的標準，這個問題在筆者寫過的一篇文章《中國中古石刻文獻文字數位化處理標準芻議》<sup>①</sup>一文中，以石刻文獻為例做了討論。主要涉及：現行字符集集外異體字處理標準、字際關係梳理標準等內容。在此不再贅述。

### [參考文獻]

- [1] 孫海波：《甲骨文編》，中華書局，1965年。
- [2] 容庚：《金文編》，中華書局，1985年。
- [3] 張守中：《郭店楚簡文字編》，文物出版社，2000年。
- [4] 張守中：《睡虎地秦簡文字編》，文物出版社，1994年。
- [5] 徐富昌：《武威儀禮漢簡文字編》，國家出版社，1995年。
- [6] 陳松長：《馬王堆簡帛文字編》，文物出版社，2001年。
- [7] 駢宇騫：《銀雀山漢簡文字編》，文物出版社，2001年。
- [8] 張守中：《張家山漢簡文字編》，文物出版社，2012年。
- [9] 李學勤：《清華大學藏戰國竹簡(壹—叁)文字編》，中西書局，2014年。
- [10] 湯餘惠：《戰國文字編》，福建人民出版社，2001年。
- [11] 徐正考：《漢代銅器銘文文字編》，作家出版社，2007年。
- [12] 劉釗：《新甲骨文編》，福建人民出版社，2009年。
- [13] 張振謙：《齊魯文字編》，學苑出版社，2014年。
- [14] 湯志彪：《三晉文字編》，作家出版社，2014年。
- [15] 臧克和：《漢魏六朝隋唐五代字形表》，南方日報出版社，2008年。
- [16] 劉志基等：《中國異體字大系》，上海書畫出版社，2007年-2010年。

【作者簡介】郭瑞(1980-)，華東師範大學副教授，研究方向：漢字數位化處理，中古漢字發展史。

<sup>①</sup> 郭瑞，中國中古石刻文獻文字數位化處理標準芻議[J]，漢字研究，2014，第10輯：261-272。