

方言詞典的音字系統性

以教育部國語會《台灣閩南語常用詞詞典》為例

清雲科技大學·通識教育中心

駱嘉鵬

1. 前言

方言詞典的編纂，可以分為兩大類型：一是方言標準詞典，一是方言比較詞典。前者以教學為預設目標，在各個次方言間，取得一個折衷的標準，而不羅列所有次方言的差異；後者以研究為主要目的，重在方言差異的實錄，以做為比較研究的依據。兩者的編輯旨趣既有區別，對於資料的取捨和編排的方式，自然也就有所不同，

教育部國語會的《台灣閩南語常用詞詞典》，為了照顧大多數的台灣閩南語使用者，因此在音讀方面酌收第二優勢音；在用字方面，也附註異體字。但基本上，本詞典仍屬於方言標準詞典，故其注音和用字系統的一致性，也就顯得格外重要。

然而，就實際執行效果而言，由於編輯成員母語腔調的不同、閩南語用字標準初定、以至於難以避免的人工手誤等因素，在本詞典審校結束之後，難免仍有音字系統不一的現象發生。為提升本詞典音字系統性，並降低錯誤率，乃由本人負責整理本詞典相關語料，歸納全典閩南語用字和注音詳情，以做為進一步檢視、修正的參考依據。

本文擬分享相關拙見和經驗，並藉此拋磚引玉，切磋琢磨，以做為爾後類似作業的參考。

2. 音字系統性的意義

方言詞典的音字系統性，包含兩個層面：一是注音系統，一是用字系統。所謂注音系統的一致性，是指在一部詞典中，對於任一詞素的注音，不因腔調的不同而分注不同的音讀；用字系統的一致性，則是指任一詞素的字形，不因習慣、風格……等差異而有所不同。這種一致性的原則，應該貫徹在整部詞典的所有內容，如詞目、用例和附錄等。

對於方言標準詞典而言，音字系統的一致性，攸關整體的成敗。一部注音和用字系統紊亂的詞典，將會讓使用者無所適從。果真如此，則其實用價值勢必大打折扣。

3. 音字系統性的檢校需求

本詞典的編輯過程，是台灣閩南語學界的空前大合作。由於編輯和審稿委員囊括學界菁英，南腔北調、各放異彩。本詞典編輯體例雖約定音讀以通行腔為標準，但因各委員的自然習用，即使多加留心，百密一疏，勢必難以完全倖免。

再者，由於閩南語用字向來混亂，爭議用字言人人殊；人工校稿，難以兼顧全局；加上眾人合作，更難一以貫之。教育部推薦用字公布於本詞典大致底定之後，緩不濟急。因此，本詞典的用字系統性實有必要加以全面清理。

此外，由於資料內容的龐雜，人工手誤，自然也在情理之中。凡此種種，都需要對整部詞典的音字系統一致性做一次全面性的整理。

4. 音字系統性的檢校方法

根據個人對本詞典的了解，要提高本詞典的音字系統一致性，必須先將本詞典所有閩南語用字和注音資訊加以組合，歸納出所有用字的所有注音之後，再以人工檢視問題字音，並視實況加以更正。

本詞典正文中的閩南語用字和注音，出現在詞目、詞目注音和釋義欄中的例句(詞)之中。關於詞目和詞目注音欄，因為內容單純，處理方式相對簡單許多。至於釋義欄中的的例句和注音，則在進行音字媒合之前，還需要執行一些前置處理程序，如此方能確保詞典用字的實際注音，以做為後續處理的依據。

以下依序臚列所有執行程序，並略述各程序的作業目的及其處理筆數。

順序	作業項目	作業內容	目的	處理筆數
1	前置作業	總檔新增「編號」欄，依序填入紀錄編號	做為後續處理的索引	17,243
2	標記統一	將所有「例：」、「例如：」、「例；」替換為「例：」	統一例句識別標記，做為後續擷取例句及音讀之依據	14
3	標記統一	將詞目、釋義、異體及近反義詞欄中所有「v」替換為「-v」	統一造字識別標記	98
4	標記統一	將詞目、釋義、異體及近反義詞欄中所有「--v」替換為「-v」	統一造字識別標記	27
5	釋義彙總	將各釋義欄中非空白紀錄及編號填入暫存工作表	便於後續擷取例句及音讀	19,635
6	例句篩選	將暫存表包含「例：」之紀錄轉存於「例句」工作表	便於後續擷取例句及音讀	12,885
7	例句篩選	將暫存表不含「例：」之紀錄轉存於「非例句」工作表	以備不時之需	6,750
8	欄位拆解	使用「例句」工作表，將釋義欄中「例：」標記之右方文字存入「例句」欄	清除與例句音、字和華語翻譯無關之訊息	12,885

9	欄位拆解	將「例句」欄中「()」號之間的文字拆置於「華語」欄	分離出華語釋義(但一欄內存多例者仍待分析)	10,541
10	欄位拆解	校正例句紀錄，將每條例句(詞)分列一筆紀錄	校正前項之疏漏	14,885
11	欄位拆解	將閩南語例句、注音及華語釋義分欄存放	便於後續比對閩南語用字及注音	14,885
12	欄位拆解	清除「例：」之右多餘的空白	統一體例，以便音字媒合	9
13	欄位拆解	清除閩南語例句之末多餘的全形空白	統一體例，以便音字媒合	2
14	欄位拆解	清除閩南語字與音之間多餘的空白	統一體例，以便音字媒合	158
15	漢字拆解	將例句欄中各漢字代碼分別以辭典未使用之漢字代替	便於後續之音字媒合時，精確計算字數	1,672
16	漢字拆解	清除例句用字欄中之標點	便於後續之音字媒合作業	10,728
17	漢字拆解	離析所有例句之所有漢字	便於後續之音字媒合作業	95,802
18	注音拆解	將例句注音中之半形逗號改為全形逗號	統一體例，以便音字媒合	2
19	注音拆解	依空白與連字符將例句注音分音節，並計算各句音節數	便於後續之音字媒合作業	95,787
20	音字媒合	比對各例句漢字個數與注音音節數之異同	篩檢出字數與音節數不同之紀錄，以便檢視比較其差異，減少音字媒合之錯誤率	14,885
21	音字媒合	逐筆檢視例句漢字個數與注音音節數有別之紀錄，並視實況校正	校正例句中缺字、缺音及音節誤黏之紀錄	17
22	音字媒合	音字媒合	彙計各例句用字字種、讀音，及其使用次數	95,802
23	音字媒合	補足整句注音末端漏標之句號	統一體例	24
24	音字媒合	將例句注音首碼小寫者改為大寫	統一體例	33
25	音字媒合	錯誤音節檢校	篩選不合法音節，並據以校正主檔及例句注音檔，重計	12

			用字注音使用次數	
26	音字媒合	詞目及注音欄音字媒合	做為比較例句及詞目注音異同之依據	17,243
27	音讀核校	詞目音字彙總檔主音、二優音讀分析	做為核校例句及詞目注音之參考依據	1,792
28	音讀核校	例句中複音詞二優音讀篩檢	依據詞目音字彙總檔主音、二優音讀分析，篩選出例句中複音詞採用二優之紀錄	11
29	音讀核校	例句字音歸類	依據詞目音字彙總檔主音、二優音讀分析結果，區分所有例句字音的類別，以便篩選錯誤注音	95,802
30	音讀核校	例句中之非主音讀核校	檢視例句注音未採用主音讀之紀錄，並依實況予以校正	421

5.結語

方言標準詞典的音字系統性攸關整體成敗，詞典內含資訊龐雜，集體作品難免鑿柄不入。善用電腦資料庫軟體可以輔助歸納資料整體訊息，方便檢視問題資料，並適當改正疏漏，以降低詞典的錯誤率，提高其品質。

參考文獻

- 教育部國語推行委員會，2007a，《臺灣閩南語羅馬字拼音方案》，
<http://www.edu.tw/files/bulletin/M0001/151609.pdf>。
- ，2007b，《臺灣閩南語推薦用字（第1批）》，
http://www.edu.tw/files/bulletin/M0001/300iongji_960523.pdf。
- ，2008a，《臺灣閩南語推薦用字（第2批）》，
http://www.edu.tw/files/download/M0001/100iongji_970501.pdf。
- ，2008b，《臺灣客家語常用詞辭典（試用版）》，<http://hakka.dict.edu.tw/>。
- ，2008c，《臺灣閩南語常用詞辭典（試用版）》，
<http://twblg.dict.edu.tw/tw/index.htm>。
- ，2008d，《臺灣閩南語常用詞辭典編輯凡例》，
<http://twblg.dict.edu.tw/tw/index.htm>。