

【資料】

閩南語俗曲唱本「歌仔冊」全文資料庫

(<http://plaza16.mbn.or.jp/~sunliong/kua-a-chheh.htm>)

王 順隆

The Texts Database of Folk Songs in Min-Nan Dialect

ONG, Sun Liong

【序言——得諸於公，用諸於公】

本資料庫係個人於1991年撰寫學士論文時，匯集散布四方的「歌仔冊」藏書之餘所產生的一個副產品。當時網際網路尚未問世，因此建立資料庫的目的只是作為個人研究之用。近年來，多方台灣本土文化研究者得知本資料庫的存在，頻頻要求早日公開；又拜科技進步之賜，透過網路不但在世界各角落的使用者都能隨時上網連線，而且可以立時檢索查閱，因此資料庫的公開方才展現了曙光。

98年中，筆者與在漢籍全文檢索上已有輝煌成就的中央研究院接洽，提議由我提供資料，再配合中研院計算中心的技術，希望能讓所有關心台灣文化的研究者早日運用這個「歌仔冊」資料庫。幸蒙該中心林晰先生和黃佩燕小姐慨允相助，才能在年底之前就推出與國人共享。

目前本資料庫共收錄超過400目（140餘萬言）台灣及閩南地區出版的「歌仔冊」歌詞全文。版式從清道光年間的木刻版至今日的鉛印版皆有。這個縱貫百年以上的資料庫，完整的展現了閩南語口語的書寫原貌。因此相較於近年坊間充斥的台語怪字、僻字和所謂的本字，我們對原稿所用之文字儘量不加修飾與更動（如「廣」通「講」），只對異體字（如「艸」改為「草」）或是明顯的錯字（如「今日」誤寫為「令日」）稍作統一規範。希望提供給有意以台語書寫文章者另一個思考空間——「民間的傳統書寫法」。

由於早期民間的用字喜以常用字、諧音字替代本字或有音無字的字詞，故出現了極多的所謂訓音字。如「拖仔乎阮下茶科」（竹林書局《姨仔配姊夫歌》），一般人光憑字面實難理解這句話本義「抽屜給我放肥皂」，惟以閩南語照字面讀出字音後就會恍然大悟。當然，非深通閩南語漢字之文音音和白話音者，實難以自在的變換文白讀音以揣測原作者的本意。因此為顧慮非閩南籍人士的使用及檢索功能的擴充，我們計畫日後將所有的歌詞配上讀音，讓使用者也能從讀音檢索字詞，以化解諧音字氾濫的障礙。當然，此項後續作業預期將會發生很多難題，由於這是空前的創舉，我們只能從失敗中擷取經驗，期望在將來能樹立一個規範，讓往後同類型的資料庫引以為用。

另外，本資料庫的建置，截至今日都僅憑一己之力完成，在校對與設計上難免有很多不盡人意之處，希望在往後能以團隊作業的方式，縮短輸入和校對的時間，集思廣益，使其更臻完美。也歡迎各方賢能提供相關的意見或資訊，以糾正和彌補我們的缺失。

筆者手上猶有上百本的「歌仔冊」和數百部的「潮州歌冊」還未輸入資料庫，基於個人時間和經費所限，我們正向日本及台灣的相關單位申請研究經費，希望能早日將包括明清的潮州戲文在內的閩南語口語書面資料都成立資料庫。如此一來，閩南語的研究將可從往日的「點」擴充至「面」，以至「體」的層次，編織出一張具體可見的語料網，而不必再像過去一般，光憑音韻學、語音學的理論法則去推演論證早期的閩南方言。當然，其他民間文學、民俗學、戲曲小說的漢學研究者也能透過全本歌詞的查閱，取得自己所需要的台灣、廈門「歌仔冊」的全文，進行故事內容的研究。

最後，筆者當初以一個窮留學生的身分開始收集「歌仔冊」時，根本無餘力去收購已經被台灣書商炒成天價的原版「歌仔冊」，只能一點一滴的去發掘收藏「歌仔冊」的圖書館或是拜訪懇求台、日、中等民間的藏書家讓筆者複印唱本。數年來稍有所成，總算建立了一個粗具規模的收藏，雖然花費了不少旅費，卻比購買原版「歌仔冊」實惠多了。對一個研究者來說，我們需要的只是實質的內容，而不必定要實體的真本。

「英國牛津大學東方圖書館」、「中央圖書館台灣分館」、「傅斯年圖書館」

及「日本國立亞非語言文化研究所」等公家機構都曾提供筆者館藏的「歌仔冊」影本；而私人方面則有台灣文史學家黃天橫先生、劉峰松先生、杜建坊先生、廈門的答嘴鼓曲藝家林鵬翔先生、陳勁之先生、日籍樋口靖教授、波多野太郎教授及三田裕次先生等人都不吝惠賜個人私藏的「歌仔冊」副本或相關資料給予筆者。除了家藏的數十種戰後出版的唱本之外，以上這些典藏就是本資料庫原稿的主要來源。衷心感謝這些資料所有者的支持，也懇請受惠於此資料庫的使用者不要忘了沒有他們的概允借閱，甚至提供副本，本資料庫絕無法成立；也希望還有一些坐擁「歌仔冊」的收藏家能共同參與這個後人乘涼的計畫，充實這個屬於全人類的“閩南語俗曲唱本「歌仔冊」全文資料庫”，並為台灣文化的保存及宏揚共同奮鬥。

【凡 例】

1. 本資料庫收錄之閩南語俗曲唱本，係專指以台灣閩南語、廈漳泉等閩南語次方言所記之通俗唱本（以下稱「歌仔冊」）。以其他閩南語次方言所記之俗曲唱本不包括在內。
2. 總目錄之下第一層收本資料庫序言、校勘記、凡例、使用規定、建置年表及筆者的「歌仔冊」相關論述。
3. 第二層以「歌仔冊」的出版商分類，先記出版地，再記堂號；原唱本未見堂號者，因幾為清代的木刻本，故全列入「清代刊本」一類。又，「上海開文書局」之前為清代及民國時期在大陸發行的書商；「台北黃塗活版所」至「台南博文堂」為日治時期在台灣印行的出版社；「台中文林出版社」「新竹興新出版社」「新竹竹林書局」則為戰後的「歌仔冊」出版商。
4. 第三層依歌名分筆。原冊號是依「歌仔冊」封面標題所編成，惟轉換成資料庫後，改依歌名標題分筆，因此同一冊號的資料，表示原收錄於同一本「歌仔冊」內。
5. 第三層中，同一部歌詞因篇幅過長者，取適當長度，另分割為第四層。
6. 為考慮網頁版面的視覺效果，已對非七言詩體的長短句歌詞自行斷行，

儘可能將完整的句子納於同一行內。因此本資料庫中歌詞的斷行、斷句並不一定完全同於原書。

7. 異體字、通用字、古字等爲便於輸入與檢索，一律改爲大五碼 (Big-5) 中的既有字體；民間自創的俗字、方言字，將另行提供造字檔。歌詞中摻插的數字即爲缺字的部份，惟目前缺字表和造字檔尙未完成，待完成後請自行下載中研院「瀚典全文檢索系統」的專用造字檔，以完整顯示所有的歌詞。
8. 民間用字歷來從無統一規範，雖有一部份的字詞已經廣泛地通行於民間，如「厝」、「查某」、「俗」、「趁」等幾乎少有異議，惟大多數的字詞卻隨時代、地域或作者不同而出現各式各樣的寫法，如「ka¹ ki⁷」一詞，就可發現至少有「家己」、「加忌」、「加己」三種寫法。因此檢索字詞時如僅依本字查詢，必定會有許多遺漏，造成檢索結果的錯誤。
9. 內文中的「□」，表示原稿印刷模糊或是缺葉無法判讀，而以一個「□」代替一字。
10. 書名標題後附有「◎」者，表示該歌詞已依據原稿完成校對；附「☆」者，表依其他出版社之同版本歌仔冊校畢。
11. 本資料庫目前尙屬試車階段，有許多使用上的不便，今後將隨時擴充功能或修改版面，並歡迎讀者提供意見。(sunliong@bb.mbn.or.jp)
12. 有關「歌仔冊」的相關研究，請參閱個人網站。(www.sainet.or.jp/~zhongwen/ong.htm)

【校勘記】

數個月來的校對作業，終於體會到全文資料庫的建置，最大的難關不在於輸入龐大的資料，而是在於既枯燥又似乎永無止境的校對工作。因爲，再多的文字資料都可以交由業者或專職人員進行機械式的輸入，其效率之高令人咋舌；而方言資料庫的校對卻非他人所能代工，唯有自己一字一字的確認、修正。其間所花費的精神和心力，實非經驗者不能體會。

當中研院計算中心於1998年底將“閩南語俗曲唱本「歌仔冊」全文資

料庫”加入該院的「漢籍電子文獻資料庫」，並先於院內上線公開之後，校對工作也同時正式地展開了。由於當初將原稿寄往大陸請業者輸入，完工後欲透過郵局寄回日本時，磁片須經中國的保密局檢查，方可寄出國外。1994年底業者將磁片和原稿一併送檢，卻沒料到保密局以歌仔冊中有黃色文字，又以竹林書局早期發行的歌仔冊封面印有「反攻大陸」、「反共抗俄」、「解救同胞」等反動文字為由將原稿扣押了。即便日後親自上門催討，對方依然大打官腔，相應不理，許多原稿就因此失落了。如此荒唐可笑之事，竟然發生在20世紀即將結束之際，真是「秀才遇到兵，有理說不清」。

校對作業，首先交由一位學生將原稿和檔案逐字比對，並製作造字表。本想如此就可達到一定的標準，完成校對的工作。不料，該生校完一部份的原稿後，才發現並不是只要認識漢字的人就可以勝任校對。不解閩南語，不解歌仔冊用字習慣，就如同讓一位非國文系的學生去校對「十三經」一般，他只能做到原稿一字，檔案一字地機械式校對。數十萬字的校對工作本既單調又費眼力，如此機械式的比對更容易發生失誤；即便是一個完美主義者，他所能達到的也只是原稿的百分之百重現，而無法自行判讀老舊木刻版，石印版歌仔冊中模糊難辨的字跡，或是修正原稿中明顯的錯誤。至此，筆者終於覺悟到所有的原稿必須親自校對，無法偷工省事。因此在學生做完初校後，所有的檔案又親自做了第二校。

在校對的過程中，我們秉持了如下幾個原則：

1. 歌仔冊作者的時代、出身籍貫不同，語言腔調各異，還有許多互不通用的方言詞彙，因此有很多因素造成我們無法判別是原稿本身的訛字，還是作者用了一個我們未知的字詞。為忠實展現原歌仔冊的書寫習慣，我們原則上不更動原稿中的任何文字，只對很明顯的訛字，如「令日」為「今日」之誤，「天鳥」係「天烏」之誤等加以修正。
2. 民間刊刻的通俗讀物為了省時、省工，在刻板時充斥了簡體字、俗字、異體字或古字，為了便於電腦輸入和檢索上的統一，我們將這些俗字、簡體字等都改為台灣大五碼中的通用字形。對於借音字、訓讀字、別字等則不予更動。

3. 尚未校對的歌詞中往往穿插了許多數字，這是當初在輸入時凡是電腦無法鍵入的漢字都暫時按編號列入一張缺字表。如今因無原稿校勘補字，故只能暫時擱置。已校完的部份，已重新建了一張造字表，並在原稿的數字前加上「*」號，以資區別。只待造字檔完成，即可上線轉換。
4. 早期木刻版歌仔冊的版式紊亂，有少數是以七言三句為一行者，為求統一格式，我們在輸入時將之改為四句一行。另有些長短句的小曲，原版中未加斷句或是斷句有誤者，為便利讀者閱讀，自行加以斷句。若有誤斷，尚請指正。
5. 早期的木刻版和石印版歌仔冊，保存年代較久，紙質較差，因蟲蝕風化的結果致使原稿模糊難辨或殘缺之處，依前後文盡量予以復原。無法復原之處，依應有字數加記「□」。
6. 為便於閱覽、查詢或日後研究者引用，我們在所有歌詞的行首加上了行號。又，為縮短網路傳送資料的時間，我們將超過70行的歌詞予以切割分段，以每次傳送歌詞不超過70行為原則。

如先前所述，手上的原稿失去了大半，故此次校對作業未能一鼓作氣，以竟全功，至為遺憾。希望在日後重新取得影本之後，能早日開始第二期的校勘作業。本次第一期作業中校畢的原稿，我們在書名標題後加上了「◎」號，而標題後附有「☆」號者（如166冊、177冊等），表示已依其他出版社發行的歌仔冊校對，待日後若能取得原版，再予重新修正。未附記號者，就表示尚未校稿，難免有些錯誤，請讀者在引用時多加留意。根據完工後的初步統計，未校的部份其錯誤率約在千分之五左右；二校完成後的歌詞，相信其錯誤率應已在萬分之一以下。如果讀者對已校歌詞有不解其義，或是懷疑有錯誤的字詞，應是原稿本身的問題，而非輸入的失誤。

第一期的校勘作業，經統計約50萬字，約佔全文的三分之一強。儘管如此，這也花費了近三個月的時間。衡量個人的集中力和眼力，充其量一天只能校對一、兩萬字，若是要求完美，進行二校，則時間就必須加倍。因此每當思及剩餘未校的部份，除了再次前往各地複印原稿之外，還必須再花上

兩倍的時間校稿，不由得雙眼發昏。各方諸賢若有這方面的經驗，還請多多惠賜良策。

【使用規程】

1. 本資料庫對使用者和查閱筆數不予設限，任何人皆可自由檢索閱覽。
2. 透過網際網路使用本資料庫不收取任何費用。
3. 檢索的結果可任意為研究報告或教學所引用，惟請加註資料出處。（閩南語俗曲唱本「歌仔冊」全文資料庫 “The Texts Database of Folk Songs in Min-Nan Dialect”）
4. 因使用本資料庫而發生的任何損失，概由使用者自行負責。
5. 本資料庫的著作權屬王順隆所有。自本資料庫擷取的所有資料、文獻、檔案等不得以任何形式從事任何以營利為目的之行爲。本資料在庫不影響內容的前提下，已略加暗記，以茲辨認。對惡意違犯者，將于網頁上公佈其惡行。
6. 本資料庫之所有檔案未經許可不得下載，使用者請以列印或複製方式保存使用。
7. 對本資料庫有任何意見、疑問、指正，請以電子郵件直接聯繫。
(sunliong@bb.mbn. or. jp)

【建置年表】

時間	事項
1991/ 3	首次接觸「歌仔冊」，開始收集文獻及訪書
1991/12	提出學士論文〈閩南語「歌仔冊」所使用的台語漢字及詞彙〉
1993/ 9	於《台灣風物》發表〈談台閩「歌仔冊」的出版概況〉
1994/ 6	與北京業者洽商輸入作業事宜
1994/ 8	第一批原稿輸入開始
1994/ 9	第一批原稿輸入完畢，第二批原稿輸入開始
1994/ 9	於《台灣文獻》發表〈閩台「歌仔冊」書目・曲目〉
1994/11	第二批原稿輸入完畢，同時原稿遭中國保密局以反動文字及黃色文字為由扣押
1995/ 8	親訪中國保密局索討原稿，未果
1996/ 3	於《台灣文獻》發表〈「歌仔冊」書目補遺〉
1996/ 9	以「短期共同研究員」身分，調查日本亞非語言文化研究所藏「歌仔冊」
1996/10	提出研究成果〈閩南語通俗唱本「歌仔冊」による方言語彙索引の試作〉
1997/ 2	於「第5屆國際閩方言研討會」發表〈閩南語「歌仔冊」的詞彙研究〉
1997/ 6	於《台灣文獻》發表〈從七種全本《孟姜女歌》的語詞・文體看「歌仔冊」的進化過程〉
1997/10	於《文教大學文學部紀要》發表〈閩南語「歌仔冊」的進化過程——從七種全本《孟姜女歌》的語詞・文體來看〉
1998/ 2	於《文教大學文學部紀要》發表〈論台灣「歌仔戲」的語源與台灣俗曲「歌仔」的關係〉
1998/10	於《文教大學文學部紀要》發表〈「潮州歌冊」研究中的幾個問題〉
1998/11	與中央研究院計算中心第一次洽商資料庫合作事宜
1998/11	撰寫閩南語「歌仔冊」全文資料庫【序言】
1998/11	閩南語「歌仔冊」全文資料庫檢索系統雛形完成
1998/12	與中央研究院計算中心第二次洽商資料庫改修事宜
1998/12	閩南語「歌仔冊」全文資料庫第一版完成，開始試用
1998/12	原稿校對作業開始
1998/12	於日本亞非所發表〈閩南語通俗歌謠集「歌仔冊」の進化と變容〉
1998/12	撰寫閩南語「歌仔冊」全文資料庫【凡例】、【使用規程】、【建置年表】
1999/ 4	原稿校勘作業完成（第一期）
1999/ 4	撰寫閩南語「歌仔冊」全文資料庫【校勘記】
1999/ 9	資料庫正式版完全公開

【使用範例】

為介紹本資料庫的檢索方法，特以「日頭」一詞為例，實際進行檢索。
列出結果於下，並簡介資料庫的構造：

瀚典·閩南語「歌仔冊」(通俗唱本)

ASCC
中央研究院計算中心
ACADEMIA Sinica
Computing Centre

【根目錄】

檢索條件

目錄深度

- 嘉義 玉珍書局
 - 245冊 戒世驚某歌
 - 246冊 最新手巾歌
 - 247冊 最新百果歌
 - 248冊 火燒紅蓮寺歌 第二集
 - 248冊 火燒紅蓮寺歌 第三集
 - 248冊 火燒紅蓮寺第五集
 - 248冊 火燒紅蓮寺歌第六集
 - 248冊 火燒紅蓮寺歌 第七集
 - 250冊 金快運河記新歌
 - 251冊 新編勤改修身歌
 - 252冊 最新勸世人心知足歌
 - 253冊 最新開天闢地經紀歌
 - 254冊 菜瓜花鸞英為夫守節歌
 - 255冊 萬能機關真正談天說地歌 上本
 - 256冊 新編流行英臺回家想思歌
 - 257冊 新編流行蔡端浩洛陽橋歌
 - 258冊 自嘆煙花修身歌 下本
 - 259冊 最新戒煙歌
 - 259冊 三孀祝壽新歌
 - 259冊 最新昌門賢母的歌
 - 259冊 最新懺悔的歌
 - 260冊 十二門人新歌
 - 261冊 天作良緣新歌
 - 262冊 談天說地新歌
 - 263冊 最新愛情與黃金歌 上集
 - 264冊 菜瓜花鸞英為夫守節歌 上本
 - 265冊 最新流行青冥三會新歌
 - 265冊 新編煙花勸世歌
 - 266冊 蕊翠新歌
 - 266冊 絨口歌
 - 266冊 知足歌
 - 266冊 醒世歌
 - 266冊 烏貓歌

進入第一層，可見以出版社分類的目錄。選擇其中之一，即可進入第二層：

第二層

瀚典: 【閩南語俗曲唱本「歌仔冊」全文資料庫】

ASCC

中央研究院計算中心
ACADEMIA Sinica
Computing Centre



檢索條件

日頭

目錄深度

• 閩南語俗曲唱本「歌仔冊」全文資料庫

1. 清代刊本
2. 廈門 林國清
3. 廈門 文德堂書局
4. 廈門 會文堂書局
5. 廈門 博文齋書局
6. 上海 開文書局
7. 台北 黃塗活版所
8. 台北 周協隆書局
9. 嘉義 捷發出版社
10. 嘉義 玉珍書局
11. 台中 瑞成書局
12. 台北 光明社
13. 台北 禮樂活版所
14. 台南 博文堂
15. 台中 文林書局
16. 新竹 興新書局
17. 新竹 竹林書局



第二層依同一出版社所發行的「歌仔冊」書目分項。再選擇其一即可進入閱覽歌詞全文。若該唱本歌詞篇幅較長，則須進入第三層。

在第一至第三的任何一層裡，於「檢索條件」欄中鍵入字詞，再於選擇方格中指定檢索對象，然後按「執行」即開始檢索，並自動顯示檢索結果如下：

瀚典: /閩南語「歌仔冊」(通俗唱本)/廈門.../159冊 最新張繡英林無宜相褒 第 1 頁, 共 1 頁

ASCC
中央研究院計算中心
ACADEMIA SINICA
Computing Centre

【資料庫】 【根目錄】 【回上層】 【檢索條列】 【檢索報表】 【全段顯示】

檢索條件: 日頭
找到 87 段, 此為第 26 段

(略 ...)

49. 有時會到娘家庄	君身一時卜返到	日頭斜西都卜落	娘看日頭亦久未
50. 今日十外也有月	日頭那暗即點火	照寔言語共君說	君卜返厝著恰早
51. 又驚警察來巡查	路頭走慌著恰早	日頭未暗好行踏	娘頭看日未過午
52. 行到君厝未半埔	卜是離遠久日路	不敢君赶你歸埔	又卜來挑下午茶

(略 ...)

55. 共娘門挽永有時	恁翁看見有怪意	敢會致害娘仔你	勸君寬心免挂疑
56. 真情說出郎君知	阮厝丈夫真古意	免驚害娘個事志	日頭未暗著回返
57. 敢著回返君家門	想卜起行又快斷	目尾相對心頭酸	娘看日頭未半天
58. 親哥緊返是卜咩	一脚共娘門挽擲	乎娘加趁十外錢	娘你錢銀教即愛

(略 ...)

65. 行到崙頭學目看	咱娘聲音載心肝	看見日頭有即晏	緊緊來去娘茶山
66. 舉頭看見咱君到	娘身歡喜在心頭	恰恨分開君回返	思娘快得到天光

(略 ...)

此時，若選擇 [全段顯示]，亦可閱覽歌詞全文。