

# 建構詞彙語意知識庫與語言教學策略的多元文化思考：

## 以「網路展書讀」網站語言文化教學設計為例

羅鳳珠

砂岡和子

台灣元智大學

日本早稻田大學

### 壹、前言

隨著全球化的進展，大量的地區文化以及歷史語言正在加速消失，應用資訊科技和網路作為搶救、典藏和推廣文化遺產的工具，可以滿足日益迫切的語言教學策略的多元文化要求。筆者（羅）自 1992 年以來持續建置的「網路展書讀」文學網站，將文學作品及相關文物數位化，並從中建立文本、詞彙、語意、語音、文物資料庫，以 GIS 顯示其地理位置分佈狀態，開發多種教學功能，是一個兼具不同時代，不同文體、文本、媒介、功能的中國文學網站。本文以描寫台灣人民生活、情感、思想、山川文物及風土民情的台灣文學為範圍，探討如何以數位化方式典藏先民留下的文獻，透過網路無遠弗屆的傳播功能，向全球介紹台灣文學；藉助日新月異的資訊科技工具，吸引年輕人親炙文學；建構整合文字、影音媒體的教學模式，在網路上虛擬語言學習實境，落實以多語言教學，多元文化兼容並蓄的方式，達到保存語言、文化的教學目標。

詞彙語意知識庫是電腦能否實現文本內容理解的關鍵因素，詞彙語意的計算和理解，是語意分析中最為關鍵的一步。本研究的主要目標是以歷時性與共時性語言、多族群語言、多文體對映的語料庫作為基礎，建構符合人類知識體系概念的詞彙語意知識庫（LSKB）。透過詞彙語意的計算和理解、語意標記與語意概念的分類、語法資訊的建構，建立包含詞彙之語意、語法、語用、語音等訊息的歷代語言知識庫，讓電腦具備理解語意，識別語法、分辨語用，自動標註語音的能力，進而用以作為分析、判斷、歸納篇章主題特徵的基礎，使電腦具備提取篇章主題特徵與主題分類的能力，提供語言、文學、資訊科學研究及漢語教學之用。

### 貳、研究價值

歷代語言知識庫建置的成果，將可以成為目前語料來源最多元、意義功能最豐富的中國歷代語言知識庫。

經過長期的努力與發展，各種本體知識、概念詞典或詞彙語意知識庫的建立，已取得很多成果。目前國際上具有代表性的工作包括普林斯頓大學的 WordNet、加州柏克萊大學的 FrameNet、微軟公司的 MindNet 和中科院的 HowNet 等。以上知識庫，除了 MindNet 採自動建構，其餘都是以人工處理，而且多數以當代自然語言為主。本研究所建構的語言知識庫，優於前人之處包含：1. 語料來源多元；2. 語料包含不同時代、文體、語言；3. 語料信息包含詞彙語意、語法、語用知識以及詞彙語音資料；4. 語意概念分類可以據以建立不同語料詞彙的語意對映機制，使計算機對文本的計算與理解能力大幅度提高；5. 研究模式具有跨文體，跨

時代(歷時語料與共時語料),融合詞彙之語法、語意、語音知識及建立語意對映機制的特性,除了可以用來建立歷代語言知識庫,還可以提供做為其他語言知識庫建置時參照使用。

### 參、歷代語言知識庫建置方法

本研究結合文本內容、語言學、計算語言學、圖書資訊學及資訊科學等不同學科領域的學者組成研究團隊,在各自既有的基礎上,以專長互補,資源共享的方式進行合作,也邀請數位學習領域學者以此資料庫進行數位教學設計與實證。

本研究總計結合了下列資源:1.元智大學之文學文本資料庫及詞彙語意、語音資料庫;2.北京大學之現代漢語資料庫、人民日報加工語料庫及語法信息資料庫、概念詞資料庫;3.香港理工大學之自然語言中文信息提取技術、句法、語法研究;4.日本早稻田大學之歷史語言知識庫、漢語教學語言知識庫;5.韓國首爾市立大學之漢語語法、漢語教學、語料庫語言學等,本研究在五校現有的資源基礎上建構「中國歷代語言知識庫」。

在語料庫部分,所使用的文本分為「中國歷代韻文」、「中國歷代語體文」、「現代漢語」三大類,以這三大類文本進行三種加工步驟:(一)文本詞彙切分;(二)將所獲得的詞彙進行:1.詞彙語意標記;2.語意概念分類;3.詞形結構標記;4.詞性標記;5.句法結構標記;6.語篇結構標記;7.詞彙語音標記。(三)經過標記的詞彙語料即可據以建立歷代語言知識庫。(四)詞彙語音標記可以作為文本語音自動標記使用。經過這四項步驟,即可建立包含「詞彙語意」、「詞彙語法」、「篇章主題分類(特徵)」的「中國歷代韻文」、「中國歷代語體文」、「現代漢語」等三類文體的語言知識庫。

### 肆、應用於教學之設計

本研究利用筆者(羅)已建置的文學網站之「韻文詞彙語意知識庫」及「詞彙語音資料庫」,以優勢互補為前提,結合海內外不同學術領域的學者,將之應用於地區文化研究和語言教學。本文以語言、文字、音韻、方言的教學設計為例,論述如何引用資訊科技工具,建立含有多元文化的語言文化教學網站,以GIS呈現語言及文化分佈狀態,期能落實多元文化共現,多語言共存的教學目標。網站所使用的媒體包括:文字、影音、教學軟體、GIS(Geographic information system);所使用的資料庫包括文學文本資料庫、漢語及閩、客語詞彙資料庫、詞彙語意知識庫、語音資料庫。於教學上的應用設計包括:文本詞彙教學、語音教學、GIS輔助語言文化教學,詳述如下:

#### 一、文本詞彙教學:

對閱聽者而言,閱讀文本時最常遇到的問題是對文本內容之生難字詞、典故、方言無法理解而產生閱讀障礙,影響閱讀意願。為了降低閱讀障礙,提高理解力,建立符合人類知識體系架構的語意概念分類知識庫,提供文本內容之語意概念顯示的功能,可以提高學習效率。

(一)「詞彙語意知識庫」在教學上的設計與應用包括:1.典故詞彙顯示功能:在網頁上點選「顯示典故」標記,含有典故之詞彙便會顯示超連結,使用者點選後可以連結到〔詩詞曲典故網站〕<sup>1</sup>,顯示該典故之出處內容等相關訊息,可以幫助理解該典故之語意與寓意。2.擴增詞彙知識學習:詞彙語意知識庫能將相同語意概念的詞彙聚集,以提供擴充詞彙知識學習的功能。3.一詞泛讀功能:「一詞泛讀」的概念,意指「針對一詞

<sup>1</sup> 羅鳳珠建置,〔詩詞曲典故網站〕,網址:<http://cls.hs.yzu.edu.tw/orig/>,2000年教育部卓越計畫經費補助建置,收有詩詞曲典故詞條二萬餘筆。

廣泛閱讀」，這個概念首先由鄭錦全院士提出，並以此概念建置〔全球華語文數位教與學資源中心〕網站<sup>2</sup>。「一詞泛讀」可以幫助學習者在最短時間內集中閱讀該詞彙在原著文本之篇章句子的用法，除了藉此充分了解詞義，還可以學習該詞彙在該文本之語用及語境。

- (二) 閩客語詞彙知識庫：民間文學以口耳傳唱，蘊含豐富的生活語言及歷史語言，真實呈現先民的生活方式、信仰習俗以及各地的風土民情，成為後人學習閩、客語等生活語言及民俗文化的最好教材。〔台灣民間文學館網站〕<sup>3</sup>便根據文本內容及從中抽取建置的「閩客語詞彙庫」設計了幾項教學功能：1.詞語學習：提供學習者以閩、客語詞彙練習造句以及教師線上批改的功能。如果對該詞彙的語意與用法不熟悉，可以從「閩客語詞彙庫」查詢該詞彙的讀音及解釋，也可以從台灣民間文學館網站的文本，查詢到所有使用該詞彙的篇章文句，並進一步閱讀全文。與「閩客語詞彙資料庫」及〔台灣民間文學館網站〕連結的「詞語學習」平台之設計，同時兼具「一詞泛讀」的功能。2.兒童劇場：將民間文學原著改編為閩南話、客家話、普通話三種語言的兒童劇本，增加學習的趣味性，降低閱讀時語文能力的門檻。3.生難詞彙之語意隨選隨示功能：使用者遇到不了解語意之詞彙，只需要將滑鼠移到該詞彙位置，系統便會自動顯示該詞彙之讀音及語意，降低閱讀時因語言隔閡所帶來的閱讀障礙，提升閱讀之方便性。

## 二、語音教學：

- (一) 閩客語語音教學：使用者閱讀文本時，即可點選「線上聆聽」按鈕，同步聆聽口述者之語音，學習該語言之讀書音與說話音。
- (二) 讀音自動標誌：讀音自動標示功能可以自動標註不同方言音標，保留方言讀音，作為語言教學及方便閱讀之用。
- (三) 詩詞曲文多語吟唱教學：〔但傷知音稀---詩詞曲文三語吟唱讀教學〕<sup>4</sup>網站收錄了二千餘篇(首)詩詞曲賦文的吟唱聲音資料，所使用的語言包含普通話、閩南語、客語，以提高教學與自學效率。

## 三、GIS 輔助語言文化教學：

建立「台灣民間文學地理資訊系統」，以 GIS 顯示民間文學及語言文化分佈，彰顯民間文學具有濃厚地域色彩的特色，以及語言隨著居民遷移而互相滲透，內容也隨著新地區的地域特徵而有所變動的情形。使用者透過電子地圖，正確得知每一篇民間歌謠的分佈地區、所使用的語言以及語言的分佈狀況。

## 伍、應用於漢語教學之驗證

日本人長期以來親炙中國文化，可惜現在年輕一輩對中國文學的素養已不如以前深厚，但中國文學仍然在日本脈脈相傳。日本的主要高等教育單位每年培養一定數量的專業研究隊員，對

<sup>2</sup>. 〔全球華語文數位教與學資源中心〕網站，由鄭錦全院士主持，黃居仁、羅鳳珠、蔡美智共同主持，【兼具教學與研究功能的全球華語文數位教與學資源中心】，國科會數位學習國家型科技計畫，2004年5月1日至2005年12月31日。網址：<http://elearning.ling.sinica.edu.tw/>。

<sup>3</sup>. 羅鳳珠、胡萬川，〔台灣民間文學館網站〕專題網站計畫，文化建設委員會，執行期限：2007年3月1日至2008年11月30日，分三期建置。

<sup>4</sup>. 吳明德教授、朱則剛教授主持，數位典藏融入九年一貫教學數位學習網站發展計畫，語文領域羅鳳珠負責，【不惜歌者苦，但傷知音稀—詩詞曲文三語吟唱讀教學計畫】，網址：[http://cls.hs.yzu.edu.tw/CL\\_POEM/](http://cls.hs.yzu.edu.tw/CL_POEM/)，2005年7月1日至2005年12月31日。

這些外籍中國通，建構歷代語言知識庫的貢獻是深遠的。他們的語言和欣賞文學能力並不亞於母語人，從知識庫中能獲得不同概念，從而能夠豐富自己的母文化，特別是【網路展書讀】<sup>5</sup>網站上的語音資料對研究中國文學的外籍專家受益非凡。對很少有機會接觸真實文藝的研究生來說，網路上影音媒體的語言實境會提供他們親身感受中國文學的教學環境。

本文前述由第一作者所敘述的教學設計，主要對象是以華語作母語或語言文化水平相近的人。以下介紹將以華語為第二學習語言的非華人為對象，將此網站應用於公共文化課程的教學實況。在早稻田大學政治經濟學部 2010 年春期的「東亞地域文化研究」課程講義，使用了【網路展書讀】的網站內容。71 名受講學生包含 56 名早稻田大學政經學部的日本學生，他們都對東亞政治、文化、社會表現出濃厚興趣；12 名來自大陸清華大學及北京大學的雙學位本科生；3 名來自韓國。雖然受講學生都已接受或正在接受公共漢語課程的教學，但僅為初級水平，因此上課都採用日語進行教學。15 名外國學生的日語能力，已能聽懂老師以日語教學，但並沒有達到以日語寫作的水平，因此學期最後要求學生提交報告時，所有學生允許使用日、中、英三種語文寫作。在早稻田大學，華人留學生人數年年增加，他們跟同胞和日本學生之間雖然可以用日語或國語進行一般交流，但很難有機會涉及到更深層次的文化內涵交流。

在課程第 12 回台灣香港語言生活的講義中，使用到了【網路展書讀】網站中的【唐詩三百首】、【斯文脈脈永留香：台灣文學數位文物館】、【腹有詩書氣自華：唐宋詩詞資料庫】等內容。還利用【但傷知音稀---詩詞曲文三語吟唱讀教學】網站的吟唱聲音素材，比如聲律啓蒙之東韻（雲對雨，雪對風），閩南話的陶淵明〈飲酒〉、王昌齡〈出塞〉詩，普通話的李煜〈玉樓春〉等等。對於大陸南方出身的學生而言，唐宋詩詞並不是新的學習內容，因此他們在聽了閩南話的詩歌朗誦之後，特別感興趣，在課後提出了各種感想和問題。

在早稻田大學這樣一個國際性的教學環境之中，這些文學素材能引發華籍學生內在心境和情感的感動，能為課堂提供師生之間互動的機會。另外，通過對這類網站內容的介紹，大家都可以瞭解到，彼此有著多元的方言體系及與之相關聯的地域文化。課堂上介紹此網站後，在 2010 年 7 月早稻田大學期末報告中，要求學生以自己的母語進行調查時，出乎意料的是，無論是留學生還是日本學生，都對各自使用的方言及地域文化有著較深的體會，因此特別將教學結果在本文提出，以作為驗證。

然而，對於只在高中學到一點古日語朗讀漢詩的日本學生而言，他們幾乎沒有別的機會接觸中國的文藝和文學作品，連基本的國語發音都很難聽明白，對閩南話的朗讀更不能理解。對於接觸中國古典文化機會越來越少的大多數現代日本學生而言，對該類網站所提供的內容進行理解時，有較大的局限性，也並不一定能產生共鳴。針對這一點，筆者（砂岡）在講義中進行了相關的問卷調查，調查的結果是：在被要求寫出自己的座右銘時，九成的日本學生選擇用母語的日語語句，並且在被要求將自己的座右銘翻譯成外國語時，八成的學生使用了英語。如果是在明治時代之前，多數日本知識份子可能會使用漢文典籍中的語句。此外，當被問及「現代在東亞地區有沒有必要使用共同語言？如果有需要，你會選擇哪種語言？」的問題時，除了包括留學生在內的極少數學生選擇漢語以外，90%的學生認為沒有必要另訂共同語言，理由是英語已經成為世界共通語言，沒有必要在東亞範圍內提出新的共通語言。在當代日本社會普遍對中國文化產生疏離感的環境中，若希望這類網站在對外文化傳播及漢語教學能產生積極的作用，應當需要在教學思想和措施方面投入相對的努力。

---

<sup>5</sup> 羅鳳珠建置，【網路展書讀】網站，網址：<http://cls.hs.yzu.edu.tw/>（建置時間：1993 年迄今）

## 陸、結語

經過早稻田大學教學實驗的驗證，【網絡展書讀】網站內容發揮了促進國際交流的效果，教學設計也對研究中國語言和文化的專業人才提供了極大的方便。隨著全球化的進展，在大量的地區文化加速消失以及歷史語言的學習逐漸減少的現實中，該類網站必將能夠讓學生重新認識到本地文化和多元化文化的重要性及優越性。我們從事人文學科的教研工作者，要堅信地區歷代文學確實能令人深刻瞭解自身文化的淵源，人文學科的教研工作者在教學上努力提倡，才能有助於認同自己的文化資源。

如何有效應用於中國籍之外的日、韓籍學生的教學，需要了解多元語言文化的教學目標，以及第二外語學習的理論和實踐。

一般母語的學習是處在該語言的實際環境，自小耳濡目染，所以學習的障礙比較小，學習效率也比較高。然而，學習第二語言時，由於多數人是在中學或成人以後才開始學習，學習時間被大幅度壓縮，而且已經錯過了以循序漸進方式長期累積閱讀數量、累積詞彙數量、體會詞彙語境的機會，特別是成人的語音理解能力遠不如兒童那樣容易記憶、模仿與接受。

因此推廣以中國的文藝和文學作品為範圍的語言政策多元文化教學，首先需要提高日本人基本的漢語聽、說、讀、寫能力，以促使培養出一批高層專業人才。除了日常交流，在書面漢語方面，至少應當達到通過交流，能夠引起對本地區文化產生共鳴的水平。

其次，此類網站今後應該考慮為外國使用者提供更基礎的學習輔助機能，比如，提供讓外國使用者也能一目了然的網站地圖（site map）。因為網站提供了龐大的內容，檢索閱覽頗為費時，特別對外國使用者來講，在使用時會遇到目標不明確，顯示速度慢等障礙。

第三，由於該類網站專業性較強，對於異域文化教學而言，如果沒有相應的基礎知識，理解較為困難。例如「現代亞洲地區華語使用人口分佈」以及「方言和少數民族語言使用歷史演變」、「閩南話的傳播路徑（推測）」、「中、港、台等地區的基本語言政策」等等。因此教師必須另外編制相關的資料作為補充教材，要介紹中國的文藝及語言的簡史，首先需要讓學生瞭解一定的歷史背景。如果能夠提供與網站內容有關的基礎知識連接，將會更有效果。比如在日本的教學過程中，如能提供相應的日語維基百科的連接，或是日本三省堂【言語學大辭典】<sup>6</sup>等開放式學術網站的連接，將會更便於製作教材或學習應用。

第四，外國人漢語學習者的特殊需要，應當連接於更為廣泛的漢語基礎知識的教學網站。比如，中國北京外研社與香港鳳凰衛視歐洲臺合作的漢語電視教學節目「空中孔子課堂」的網站<sup>7</sup>進行連接，以便為外國使用者學習現代華語提供更基礎的詞彙語法等輔助教材。

謝詞：本文教學設計之內容由羅鳳珠執筆，教學驗證內容由砂岡和子執筆，結論由二人共同執筆。本研究得到：1. 羅鳳珠等主持，「歷代語言知識庫建置計畫」，九十八年度蔣經國國際學術交流基金會（RG013-D-09）。2. 2010—2013 年度日本文部科學省科學研究費補助金〔基盤（C）課題番号：22520445、課題”東アジア複言語コミュニケーション・データベースの構築”：研究代表者砂岡和子〕。3. 國科會「斯文脈脈永留香：台灣古典漢詩傳唱教學地理資訊系統」（NSC 99-2631-H-155-003）及文建會「台灣民間文學館計畫」（皆由羅鳳珠、范毅軍共同主持）的資助，特此表示感謝。也感謝陸儉明教授、鄭錦全院士、朱則剛教授指導。

<sup>6</sup> 日本言語學大辭典データベース編纂委員會（代表・町田和彦），網址：  
<http://www.joao-roiz.jp/SLE/>

<sup>7</sup> 外研漢語網站（Learn Chinese with FLTRP）空中孔子課堂，網址：<http://www.chineseplus.cn>