

專書導讀

課程研究

13 卷 1 期 2018 年 3 月 頁 111-119

資訊科技與華語教學的內涵 ——評介《語料庫與華語教學》

謝家浩



書名：語料庫與華語教學
作者：陳浩然、潘依婷、吳鑑城、張俊盛、謝舒凱、楊惠媚、高照明、張莉萍、
洪嘉韻、宋曜廷、謝妙玲、曾逸群、李臻儀、熊玉雯、方淑華、王瓊淑
出版社：高等教育文化事業有限公司
出版日期：2017 年 7 月

謝家浩，香港教育大學中國語言學系副教授。E-mail: kahotse@eduhk.hk
doi: 10.3966/181653382018031301006

一談到數位資訊科技與華語教學，大家馬上想到的，可能是一些很炫的演示技巧，或是一些很吸引的互動遊戲畫面。事實上，在華麗絢爛之後，有一些學者在背後為華語教學的內涵上做出貢獻，他們借助電腦在處理語言數據上愈來愈強的能力，提出把語言學習與電子資源的內涵融合，以「語料庫」為基礎，將其中的數據做分析，找出學習者的一些正誤規則，從而優化華語教學。這些學者把他們對「語料庫」的看法與實踐結集在《語料庫與華語教學》一書中。

所謂「語料庫」指的是通過大量且根據特定原則進行文本的蒐集，包括新聞、書籍、網頁、口語轉換等，以電子化方式儲存，方便使用者以字、詞、句子等關鍵資料來檢索，也可了解當中使用度的高低。語料庫一詞在語言學上意指大量的文本，經過整理，具有既定格式與標記，語料庫的英文“text corpus”就有“body of text”的意思。簡單來說，語料庫就是一個集結了大量語言使用情形的數據，並以電腦儲存提供未來在語料上的檢索或是針對語料再做進一步分析的資料庫。早在二十世紀中期，各類型語料庫的建置已經展開，隨著一日千里的資訊科技發展，不僅語料庫增多，規模亦隨之而擴大，蒐集的語料對象也愈來愈多樣化，如學習者語料庫、專業語料庫、口語語料庫等。同時語料庫也廣泛應用在於語言的教研工作上，因為語料庫不單提供了大量的語言材料供學習者參考，也方便了研究者從語料中，直接了解語言使用的真實情形，方便設計、制定華語文教學各類課文所需要的字詞，包括語義和與場景相關的資料。語料庫的發展，至今已出現了許多著名的大型語料庫，這當中自然以英語語料庫為大宗。隨著近年來正興起華語熱，學習華語的國際人士日眾，前線教師除了教學資訊科技素養外，當然希望語言專業的學者能做出可配合教學的分析，方便他們更有效地處理教學活動。在科技日新月異的環境中，國立臺灣師範大學的教研團隊與其他友校從2012年開始即對語言與科技的融合做了深入的交流，當中包括當代腦科學對於漢語學習的了解、語言科技如何運用於漢語的教學與評量上等問題。面對大量的語言材料，語料庫的建置自然順應而來。臺灣各校院的學者針對不同語料，建置各式各樣的語料庫，一方面根據語料中的語言使用情形做更深入的語言應用探究，以便在華語文教學上提出建議；另一方面也由根源著手，發展相關工具，希望為學習者提供在自學上的輔助工具。

《語料庫與華語教學》一書的作者群，都是一群關心當前華語教學趨勢的教研工作者，他們嘗試以不同方法建置各式的華語語料庫，同時蒐集華語在不同語境下的使用情形，經過研究及數據分析後，提出優化華語教學建議。作者群包括陳浩然、潘依婷、吳鑑城、張俊盛、謝舒凱、楊惠媚、高照明、張莉萍、洪嘉馥、宋

曜廷、謝妙玲、曾逸群、李臻儀、熊玉雯、方淑華、王瓊淑等16位任教於臺灣各大學的教授和研究員，本書集結了他們對「語料庫」的觀點與理念而成，核心論題環繞「語料庫」理論架構的實踐與應用。全書共收錄11篇文章，由各學者針對語料庫建置、發展及語料庫對輔助華語教學之間的關係，做出深入切題的討論，正如主編所言，本書可以分成兩大部分來看：前六章是討論語料庫與輔助華語學習的系統及工具的建置；後五章提出不同類型華語學習者語料庫的建置及分析，並提出教學建議。

本書第一章由陳浩然與潘依婷談〈語料庫與華語教學〉，概述語料庫的發展正史與各類語料庫（Randi, 2010），同時也介紹了現存華語語料庫的建置情形及在華語教學上的應用；第二章由吳鑑城、陳浩然及張俊盛三位接力談〈網路語料庫介紹與應用〉，主要簡介現存網路語料庫及說明國內網路語料庫的建置流程，同時對於編纂網路華語語料庫在臺灣的發展優勢與侷限；第三章由謝舒凱談〈中文語料與詞彙知識地圖〉的設計，提出「混搭」應用的概念帶入語言教學系統的建置中，並討論相關應用與現有的侷限，提出語言資源的設計應因應學習情境需求而調整；謝舒凱隨即在第四章〈語料庫建置工具的介紹〉概述不同類型的語料庫及建置語料庫時可做語料蒐集、標記與分析等工具的應用（Brezina, McEnery, & Wattam, 2015）；陳浩然與楊惠媚續在第五章〈語料庫與搭配詞檢索系統〉以搭配詞為主題，討論相關檢索系統的現有發展與相關檢索系統在語言學習上的應用，基於學習者在詞彙檢索上的困境，提出搭配詞與自動配詞檢索的應用（方挺，2012），藉此提升學科教材內容的針對性與涵蓋面；高照明在第六章〈詞彙知識庫與雙語平行語料庫在華語教學中的應用〉中帶出綜合中英雙語詞彙知識庫檢索系統和中英雙語對應句檢索系統在外語學習上的幫助，中英雙語平行的語料如何有效協助國際人士學習華語。

不同類型華語學習者語料庫的建置與分析及其在教學的實際應用上，張莉萍在第七章〈TOCFL學習者語料庫的偏誤標記〉首先介紹學習者偏誤類型、標記原則與標記程序，為帶有偏誤標記的TOCFL語料庫做一階段性的總結及展望（肖奚強，2011）；洪嘉馥與宋曜廷接著在第八章〈華語文寫作語料庫建置與分析〉這一章中說明寫作語料庫的建置流程，並探討各種程度的學習者在寫作上的常見偏誤類型（熊玉雯、李慧萱、宋曜廷，2014），以及如何運用語料庫找出學生在華語學習上的難點，進而優化分級語料庫；謝妙玲、曾逸群及李臻儀於第九章〈基於學習者語料庫的語法分析與教學應用——以「還」與「而」為例〉中，提取寫作語料庫內所出現的偏誤，指出學習者在學習副詞「還」和連詞「而」時的難點，並進行語法

分析與探究，認為語料庫在輔助華語教學上，應為學習者提供足夠且具代表性的語境，才能提升學習者區分與辨別的能力（梁茂成、李文中、許家金，2010）；第十章熊玉雯與宋曜廷在〈「華語學習者語音語料庫」之建置與中介語分析〉討論了學習者的語音偏誤，並分析了國立臺灣師範大學華語學習者語音語料庫的發展與前瞻，也提出相關教學建議；第十一章為〈華語中介語口語語料庫的建置與研究應用——以學習者的時間順序連貫形式為例〉，方淑華、王瓊淑與陳浩然（2013）針對學習者的時間順序連貫形式為焦點，提出分析及討論，這是國立臺灣師範大學華語中介語口語語料庫建構過程的一份報告，為讀者就現階段口語語料庫的建置、應用帶來一點新嘗試的概念。

原書輯錄11篇文章的序列，正好配應了認知教學法的教學過程，認知教學法將華語學與教的過程分為三個階段：語言理解階段、培養語言能力階段及語言運用階段。首先，認知法強調理解是語言活動的基礎，理解係指學生理解教師所教或者所學語言材料和語言規則的意義、構成和用法。在任何語言的學習活動都應建立在理解的基礎上，如句型、聽、說、讀、寫各項能力的培養。因此，理解是語言學習的第一個階段。應該注意的是：語言規則的理解並非全面依賴教師的講解，而是在教師指導下讓學生自己主動發現語言規則。建立語料庫的目的就是要讓學與教者都可以這樣的設置來了解語言。到了培養語言能力階段，認知教學法認為語言的學習不需要掌握語言的知識和結構，還要提高正確使用語言的能力。語言能力的培養要通過有意識、有組織的練習獲得。此階段施教者既要檢查學生對語言知識的理解情況，又要培養學生運用語言知識的能力。搭配詞檢索系統與平行語料庫正是基於學習者在詞彙檢索上的困境（吳佩晏、鄭綦，2011），提出搭配詞與自動配詞檢索的應用，藉此提升學習者對教材及語言應用能力的提升而設計。語言運用這個階段的學與教任務是培養學生運用語言知識、進行優化聽、說、讀、寫的能力，尤其是學生的實際交際能力。此階段將前兩個階段學得的語言知識內容與實際運用能力結合起來，目的在於使學生聽、說、讀、寫等各方面的能力都得到全面發展。

這種全面發展聽、說、讀、寫四項技能所追求的教學目標是培養學生實際而全面地運用華語，且希望部分學生可經由進行正誤分析後加以自我糾正。此過程中學生出現錯誤在所難免，學與教者都要對錯誤進行分析，了解產生錯誤的原因，進行針對性的糾正，逐步培養學生正確運用語言的能力。語言是受規則支配的系統，學習外語就是要學習這個系統規則，且語言活動是一種受規則支配的創造性活動，只有在理解語言規則的基礎上有意義地學習，才能有效地處理學與教的語言規則。

華語的學習既要學習口語，又要學習書面語，同時重視聽、說、讀、寫四項技能的培養。對於學生所犯的錯誤，當然要理解，也要進行分析，然後針對地進行糾正，不是有錯必糾。這個理解便是語料庫所能提供的一大功能，因為它從學習者的角度出發，整理了學生的學習，讓學與教都可從主動的理解語言教學活動和有意義地學習語言知識入手，促進學生學習語言知識的活力，從而引發學生語言能力的全面發展。

隨著認知語言學與語料庫研究的深入與發展，愈來愈多的教育實踐者把認知語言學與語料庫的研究成果運用於教學實踐中，這就是輔以語料庫的新認知教學法（*new cognitive approach assisted by corpora*）。之所以把該教學法命名為輔以語料庫的新認知教學法，主要理由有：一、教學的很多語料來自真實的語料庫，不像原認知法和其他教學法的語料很多是由內省法獲得的；二、該教學法的主要語言學理論——認知語言學有別於原認知教學法的語言學理論——轉換生成語言學（從廣義上說，該語言學理論也是從人的認知方面來解釋語言的，但其認為語言是人類先天所具有的能力，是人生下來大腦中就固有的能力，即主張語言學習「天賦觀念論」）。Garside、Leech與McEnery（1997）首先提出了用語料庫來教學的想法，一旦學生掌握了必備的語料庫知識和技巧，語言學習活動的中心將是學習者。「用語料庫來教學」意味著採用基於語料庫的方法來教授與語言或語言學學科相關的課程，使學習者從所謂的「資料驅動學習」（*data-driven learning, DDL*）中獲益。就基於語料庫的語言教學而言，傳統的3P教學模式（即“*presentation*”、“*practice*”、“*production*”）並不太合適，而具有探索性的3I模式（即“*illustration*”、“*interaction*”、“*induction*”）較為適用。“*illustration*”指學習者觀察真實語料；“*interaction*”指學習者討論並分享在語料中的發現；“*induction*”則指學習者針對某個語言點歸納出自己的規則，而這樣得出的規則在隨後更多語料的基礎上逐步加以完善。為適應不同層面的學習需求，DDL可以由教師指導，也可由學生主導，但總的來說，DDL以學習者為中心。這種自主式學習過程「讓學生作為一個研究者，期待在語言學習中有獨到的發現，從而做出個人應有的貢獻」。

現時教育發展的大趨勢，學習與電子裝備的關係日見緊密，華語學習在全球掀起的熱潮，以及電腦在語言數據上處理能力的增加，學與教者也愈來愈習慣語言學習與電子資源的融合，兩者之間的互動也必定更加密切。從學習工具的層面而言，因語料庫工具的逐漸發展成熟，將可直接被學習者運用，成為自學上的有力輔助工具，如學習者可直接利用語料庫檢索詞彙，並取得用法、搭配詞、例句等相關

資訊。從教學層面來看，各種語料的蒐集及各式語料庫的建置將提供華語教材編寫的參考及建議，藉由軟體輔助所得出的語料數據，教材編輯者可以了解語言的真實使用情形，包含常用字詞與應用頻率等，對於教材、教法都有具體而實質的建議。而在華語教學的研究上，研究者更可利用語料庫的大量語料及數據分析，依循著脈絡歸類出華語語用的規則或是找出學習者常見或規律性的偏誤，也可順帶提供華語教師在教學上的設計或建議等。如國立臺灣師範大學曾金金提過的一個問題：「一石二鳥、一箭雙雕、一舉兩得，這三道成語，你會選擇哪一個作為初階華語的材料？」你會怎麼選？憑印象？還是依賴什麼？要是有了語料庫，便可由當中找到具體的理據。同樣在教材設計上，字詞用語在初階與進階的課文中自有差別，遑論段、篇與當中的修辭手法。

語料庫除了仰賴前賢的創設外，不妨考量試用Concise¹來建立一個你個人專屬的語料庫，把華語教學、課程與教材設計、教學技能、跨文化教學、評估的信度、效度等理論與實務的經驗，通過語料庫有系統地與人分享，提供華語文數位學習多面向的省思。除了書中一些專研的語料庫外，以下也提供一些在華語教研上有參考價值的語料庫，供讀者瀏覽（依筆畫排序）：

- LIVAC共時語料庫簡介 <http://www.livac.org/intro.php?lang=tc>
- LIVAC漢語共時語料庫 <http://www.livac.org/index.php?lang=sc>
- 中文句結構樹資料庫 <http://turing.iis.sinica.edu.tw/treesearch/>
- 中文詞彙網路 <http://cwn.ling.sinica.edu.tw/>
- 中文語言資源聯盟 <http://www.chineseldc.org/>
- 中文斷詞系統 <http://ckipsvr.iis.sinica.edu.tw/>
- 中央研究院上古漢語標記語料庫 <http://lingcorpus.iis.sinica.edu.tw/ancient/>
- 中英雙語知識本體詞網 <http://bow.ling.sinica.edu.tw/>
- 中央研究院漢籍電子文獻 <http://hanji.sinica.edu.tw/>
- 文國尋寶記 <http://wen.ling.sinica.edu.tw/>
- 北京大學CCL語料庫 http://ccl.pku.edu.cn:8080/ccl_corpus
- 古代漢語語料庫檢索 <http://www.aihanyu.org/cncorpus/ACindex.aspx>
- 近代漢語標記語料庫 <http://lingcorpus.iis.sinica.edu.tw/early/>

¹ 參見Concise語料庫工具網址：<http://concise.sustudio.org/p/download.html>。

- 現代漢語平衡語料庫 <http://asbc.iis.sinica.edu.tw/>
- 現代漢語語料庫檢索 <http://www.aihanyu.org/cncorpus/CnCindex.aspx>
- 搜文解字 <http://words.sinica.edu.tw/>
- 媒體語言語料庫 <http://ling.cuc.edu.cn/RawPub/>
- 傳媒語言語料庫線上分詞標註系統 <http://ling.cuc.edu.cn/cucseg/>

參考文獻

- 方挺（2012）。語料庫為本分析以中文為母語之英學習者增強配詞使用及母語影響。應用外語學報，18，35-55。[Fang, T. (2012). A corpus-based study of amplifier collocations by Chinese learners of English and L1 Influence. *Journal of Applied Foreign Languages*, 18, 35-55.]
- 方淑華、王瓊淑、陳浩然（2013）。中介語語料庫之建置及應用語料庫於華語教學之實證研究。載於崔希亮、張寶林（主編），第二屆漢語中介語語料庫建設與應用國際學術討論會論文選集（頁107-118）。北京市：北京語言大學出版社。[Fang, S.-H., Wang, Q.-S., & Chen, H.-R. (2013). An empirical study of the construction of interlanguage corpus and the application of corpus in Chinese language teaching. In X.-L. Cui & B.-L. Zhang (Eds.), *Proceedings of the 2nd International Conference of Construction and Application of Chinese Interlanguage Corpu* (pp. 107-118). Beijing, China: Beijing Language and Culture University Press.]
- 肖奚強（2011）。漢語中介語研究論略。語言文字應用，2，109-115。[Xiao, X.-Q. (2011). On the study of Chinese interlanguage. *Applied Linguistics*, 2, 109-115.]
- 吳佩晏、鄭縈（2011）。從「依」、「據」與「依據」談華語近義詞教學。臺灣華語教研究，2（1），23-46。[Wu, P.-Y., & Cheng, Y. (2011). Near synonyms in teaching Chinese as second language: A case study to “yī”, “jù”, and “yījù”. *Taiwan Journal of Chinese as a Second Language*, 2(1), 23-46.]
- 梁茂成、李文中、許家金（2010）。語料庫應用教程。北京市：外語教學與研究。[Liang, M.-C., Li, W.-Z., & Xu, J.-J. (2010). *Using corpora: A practical coursebook*. Beijing, China: Foreign Language Teaching and Research.]
- 熊玉雯、李慧萱、宋曜廷（2014）。基於ACTFL之華語文寫作評分規準。華語文教學研究，11（4），111-140。[Hsiung, Y.-W., Lee, H.-H., & Sung, Y.-T. (2014). Examining the ACTFL writing assessment rating scale for L2 Chinese learners. *Journal of Chinese Language Teaching*, 11(4), 111-140.]
- Brezina, V., McEnery, T., & Wattam, S. (2015). Collocations in context: A new perspective on collocation networks. *International Journal of Corpus Linguistics*, 20(2), 139-173.
- Garside, R., Leech, G., & McEnery, A. (Eds.). (1997). *Corpus annotation: Linguistic information*

from computer text corpora. London, UK: Longman.

Randi, R. (2010). *Using corpora in the language classroom*. Cambridge, UK: Cambridge University Press.

(本篇已授權收納於高等教育知識庫，<http://www.ericdata.com>)

