

## 語料驅動學習融入華語課堂之教學設計\*

王炳勻

加拿大維多利亞大學亞太研究系

許展嘉

國立臺北商業大學通識教育中心

龍水水

美國賓州州立大學亞洲研究系

丁曉穎

美國印第安納大學東亞語言與文化系

### 摘要

有鑑於學界較少關注語料庫在華語課堂的應用，本文旨在探討如何將不同類型的語料驅動學習 (data-driven learning) 與各種課堂教學內容有機地結合起來，分別針對其直接和間接模式，提出引導歸納式的系統化教學流程，並藉由具體教學示例，說明語料驅動學習活動設計原則，這些示例涵蓋搭配教學、近義詞辨析教學、格式教學以及句式教學四個面向。藉由在課堂上引導學生直接檢索語料庫，或間接地分析教師從語料庫篩選出的真實語料，學習者便可透過這種發現、歸納式的方法，意識到語言使用的規律，進而促進華語的學習。最終目標為使學習者具備自主學習的能力，能夠獨立使用語料庫工具解決在華語學習過程中遇到的難題，從而提高自己的語言水平。

**關鍵詞：**索引行 發現式學習 語料庫檢索 學習者自主 歸納式學習

---

\* 本文部分成果曾於 2019 年美國中文教師學會年度會議報告，感謝與會者的評論。本期刊匿名審查委員亦提供寶貴建議，使本文更臻完善，在此一併致謝。文中如有疏漏，悉由作者負責。

## 1. 引言

語料庫語言學的興起和蓬勃發展，不僅使語言研究在方法上獲得突破性的進展，也為語言教學提供新的思路和啟發。語料庫在語言教學上的應用可大致劃分為直接與間接兩個層面 (Leech 1997; Römer 2011)。間接層面的應用相當廣泛，包括將語料庫應用於工具書的出版、教材的編寫、語言測驗的開發等，以及建置以教學為導向的語料庫（如學習者語料庫、教材語料庫等）。語料庫在語言教學上的直接應用則是教師或學習者運用語料庫（工具）輔助教學或學習，此類最具代表性的應用為「語料驅動學習」(data-driven learning)<sup>1</sup>。

至今學界對於語料庫在華語教學的應用多側重在間接層面，對其在課堂教學的應用較少關注。有鑑於此，本文旨在探討如何於華語課堂中，融入不同類型的語料驅動學習活動，將此學習模式與各種教學內容結合，引導學生藉由觀察、分析語料，從中發現語言規則，進而促進目標語的學習。

## 2. 語料驅動學習文獻綜述

### 2.1 語料驅動學習的理念與特點

語料驅動學習是由 Tim Johns (1986, 1991, 1994) 所提出的二語學習方式。Boulton 與 Cobb (2017:352) 對語料驅動學習所下的定義為「學習或使用二語時運用語料庫語言學的工具與技術」<sup>2</sup>。以下圖 1 為一典型的語料驅動學習活動示例 (Boulton 2012)，對象是以法語為母語的英語學習者，教學內容為英語的動詞結構。教師從語料庫中篩選出數個含有 stop 動詞用法的索引行 (concordance lines) 提供給學生。索引行以「語境中的關鍵詞」(keyword-in-context, 簡稱 KWIC) 方式呈現，將關鍵詞對齊置中，並加粗、加底線使之更醒目，便於學生觀察此動詞出現的語境。學生應嘗試分析語料，回答學習單上的問題，最後歸納出 stop 的動詞用法（即「stop someone (from) doing something」）。

---

<sup>1</sup> 在以中文發表的相關文獻中，「data-driven learning」大多譯為「數據驅動學習」或「資料驅動學習」。由於原文名稱中的「data」指的是「linguistic data」，本文選擇譯為「語料驅動學習」。

<sup>2</sup> 此定義的英文原文為：the use of the tools and techniques of corpus linguistics for L2 learning or use。

Verbs occur in different structures in French, e.g.:

- *demander à quelqu'un de faire quelque chose*
- *vouloir que quelqu'un fasse quelque chose*

Similar patterns exist for English too, and are a favourite in the TOEIC, with questions like:

I stopped \_\_\_\_\_ the report about the supply problem to the Korean office.

(A) David from send  
(B) David sends  
(C) David from sending  
(D) David to send

The concordances below should help you find the best answer:

---

a. was no stopping it. No matter -- nothing could have **stopped** me from doing what I did -- not prayer, not my will, not  
b. vice president have not ruled out using military force to **stop** Iran from getting a nuclear weapon. In fact, Iran might heed Israel's  
c. on his armor-toed boot, and he caught her by the hand to **stop** her falling. She thought once more of the golden retriever pup, who had  
d. . " In some places, " Ms. Raje said, " women were **stopped** from going to the film. " # Ms. Raje laments the rigidity of popular  
e. # The day Topsy died was a cold, cold day, which didn't **stop** the people from coming out to watch. Electricity was a novelty, and movies

圖 1：語料驅動學習活動示例 (Boulton 2012)

在以上的教學活動範例中可看出，學習者透過觀察、分析語料，嘗試挖掘出語言使用的規律，可稱得上是語言的研究者 (Johns 1991)。換言之，語料驅動學習與一般演繹式 (deductive) 的模式迥然不同，學生在後者通過教師對語言知識的講授，進行接受式的「規則學習」(rule-based learning)。而語料驅動學習則是發現為主 (discovery-based)、歸納式 (inductive) 的學習模式，讓學生在「規則尋覓」(pattern finding) 中進行目標語的學習 (Flowerdew 2015)。在此過程中，學習者與教師的角色也有了根本上的改變。在課堂上，學生從被動的語言知識吸收者，轉變為學習的主動參與者。教師的任務也不再是單方面傳授語言知識，而是以協助者的角色，在語料庫檢索策略以及語料分析方面，給予學生引導、建議，培養其獨立、自主學習的能力 (Chambers 2010)。

## 2.2 語料驅動學習的類型劃分

將語料驅動學習融入語言教學的具體做法可大致分為直接 (hands-on) 和間接 (hands-off) 兩大類型 (Boulton 2012)。前者是讓語言學習者親自進入語料庫，搜尋相關索引行，對其觀察、分析後，歸納出規則。在此過程中，學生可獨立操作，教師亦可介入引導 (Vyatkina 2020)。這種較直接的模式，對於某些學習者來說，尤其是目標語程度較低的，至少有以下幾點挑戰：

- 一、直接語料驅動學習對學習者的電腦資訊能力要求較高。學生需要接受訓練，熟悉語料庫介面和檢索方式，方能進行有效的語料搜索 (Boulton 2012; Warren 2016)。

- 二、學習者大多不知如何閱讀、詮釋語料庫所提供的數據和索引行，對此教師須提供指導，例如，提醒學生不需對語料中每個詞語都認識 (Sripicharn 2010)。
- 三、多數語料庫收集的語料是由母語者產出，因此，閱讀難度可能會超過學生可接受的程度 (Boulton and Cobb 2017)。
- 四、語料庫中符合檢索條件的語料筆數可能過少，不足以讓學習者發現規則，也可能過多，讓學習者感到不知所措 (Warren 2016)。
- 五、語料庫所提供的索引行可能存在過多的雜訊 (noise) (即無關的語料) (Gilquin and Granger 2010)，對缺乏目標語語感的學習者，造成分析上的困難。

為了克服以上挑戰，有些學者提出較間接的模式。一般的做法為，教師於備課階段先行進入語料庫，搜索含特定語言點的語料，根據教學目標需求、學生程度等，篩選出一些索引行，進一步分類，放入紙本的學習單中或呈現於教室的投影幕上，在課堂上提供學生觀察、分析 (Vyatkina 2020)。如此一來，學習環境較不受硬體設備限制，學生也不須接受檢索語料庫的訓練。此外，語料經由教師篩選，因此較適合學生閱讀，可避免上述提及的語料太多、太難等問題。Boulton (2012) 因而指出，此類間接的語料驅動學習模式，對程度較低的學習者較容易產生立即的成效。

進行不同類型的語料驅動學習活動時，教師介入引導的程度也可有所差異。無論是讓學生親自檢索語料庫，還是讓他們間接地分析經篩選後的語料，教師的角色都至關重要 (參 Kaltenböck and Mehlmauer-Larcher 2005)。

### 2.3 語料驅動學習的優勢與侷限性

目前學界已有不少關於語料驅動學習的實證研究，探討的議題包括學習者對於使用語料庫的態度、使用的過程與行為、學習成效等 (Boulton 2017)。然而，Vyatkina (2016) 指出此領域的研究範圍仍有待進一步拓展，大部分語料驅動學習研究以英語為目標語，參與的學習者其程度多數為中級至高級，參與的教師多為語料驅動學習研究者而非一般語言教師，而學習內容大多為詞彙與形態簡單的語法結構。Boulton 與 Cobb (2017) 進行的後設分析研究結果表明，語料庫在語言教學上通常僅做為參考工具，而非學習輔助工具。然而，該研究亦發現，語料驅動學習確實具有顯著的成效。相關文獻並指出語料驅動學習至少有以下優勢：

- 一、此模式可將真實性 (authenticity) 帶入課堂 (Gilquin and Granger 2010)。學習者在語料驅動學習活動中所接觸到的語言，自然且未經簡化 (Flowerdew 2015)，與傳統教材所呈現的有所差異。
- 二、進行語料驅動學習活動時，學習者密集接觸大量含特定語言點的用例，較容易注意到語言規律 (Gilquin and Granger 2010)。
- 三、此模式的發現式特點可讓學習變得更有興趣，增進學習者的動機 (Gilquin and Granger 2010)。語料驅動學習增強學生的積極性和主動性，為自身學習增能賦權 (empowered) (Mair 2002)，自信隨之提升。
- 四、語料驅動學習活動不但可增加學習者的心理活動，亦有助於發展心理認知能力，如觀察、分析、推理、假設、猜測、反思等 (O'Sullivan 2007)。

語料驅動學習雖有其優勢，亦有些侷限，其中讓教師和學習者感受最明顯的一點為過程費時。由於語料庫操作相對繁瑣，熟悉其介面與使用方式、搜尋及篩選語料等步驟，皆需要時間，因此教師備課時投入的時間成本較高 (Boulton 2010)。對學生而言，觀察語料、閱讀索引行、完成任務更是耗時 (Gilquin and Granger 2010)。有鑑於這項侷限性，教師在設計語料驅動學習任務時，須斟酌活動進行的頻率以及每次活動所涵蓋的語言點和語料多寡。

## 2.4 語料驅動學習與華語教學相關研究

在華語教學領域目前僅有零星的文獻提及語料驅動學習 (Xu 2019)，一類是針對語料庫工具的研發及其於華語教學的應用，另一類是基於語料庫的華語課堂教學。

第一類文獻中最常提及的語料庫工具為 Sketch Engine。Kilgariff 等 (2014) 提及此工具已廣泛運用於英語教學，在華語、日語和阿拉伯語等教學亦有所應用。多項文獻 (如 Smith, Chen and Kilgariff 2008; Kilgariff, Keng and Smith 2015; Smith 2011) 舉例說明 Sketch Engine 的主要功能及其可提供的訊息，並建議可使用該工具，設計面向華語學習者的語料驅動學習任務，例如，讓學習者使用「詞彙特性速描」(Word Sketch) 這項功能，找出常與「吃」搭配的賓語，並嘗試將這些賓語分類 (Smith 2011)。Smith 等 (2008) 則指出，參與其研究的華語學習者對於 Sketch Engine 多給予正面的評價，例如：容易上手、有趣、功能強大等。

Chen、Wu、Yang 與 Pan (2016) 開發了 Intelligent Collocation Engine (ICE)，該工具內建具有詞類標記的新聞語料庫。使用者限定查詢詞的詞性和搭配類型，即可得到符合檢索條件的搭配用法、其出現頻率和相關索引行。此外，作者以「動名」搭配類型為例，透過翻譯任務測試發現，華語學習者能夠成功地使用 ICE 找到正確的搭配用法，且對此工具表達正面的態度。另外，亦有在職教師於使用 ICE 之後提供回饋意見。這些教師指出，ICE 便於他們找出合適的搭配用法，且他們有意願在備課時使用此工具。

第二類文獻以 Yeh 與 Zhang (2018) 的實證研究作為代表。該研究從自行建立的學習者口語語料庫發現，學習者在敘事時，鮮少使用標記兩個連續事件的「就」(例如：「昨天我們吃完飯，就去看電影」)。作者針對這項弱點，設計了基於語料庫的課堂教學，其中以下幾個做法與語料驅動學習特別相關：

- 一、教師引導學生多次比較由母語者和學習者產出的敘事語料有何不同，促使學生對「就」在故事敘事中的用法，有更強的意識。
- 二、教師提供一段由母語者產出的完整敘事語料，其中出現多次「就」。學生在引導問題的指引下（例如：「就」於句子的位置為何？在故事敘事中，「就」有何連接功能？），針對該虛詞的功能及使用規律進行分析，並提出假設。其後教師提供更多類似語料，供學生驗證假設，在學生報告初步分析結果後，進行確認和進一步解釋。
- 三、教師設計多種產出任務，讓學生練習使用該虛詞，例如從語料庫中選取一些敘事的語料，將其中所有的「就」移除，讓學生在適當處將「就」補回。

此研究透過實驗組與控制組、前測與後測的設計，檢視此教學法的成效。實驗組接受上述基於語料庫的教學，而對照組並無接受其他針對「就」的教學。後測結果顯示，實驗組對「就」語篇連接功能的掌握優於對照組。

從本小節的文獻回顧可知，尚未有研究者使用現有較具規模的華語語料庫工具，結合教學目標設計語料驅動學習活動，由教師在課堂上引導學生進行。本文旨在填補以上研究缺口，依作者群在北美不同大學的實際操作經驗，嘗試將融入語料驅動學習的教學流程系統化，並提供具體教學示例和建議，以供對於使用語料庫進行教學感興趣的華語教師作為參考。

### 3. 融入直接語料驅動學習的教學設計

儘管目前已有多種華語語料庫可供大眾使用，絕大多數的華語學習者對相關資源仍相當陌生。為了培養學生日後能有效將語料庫作為自主學習工具，教師需先在課堂上給予充分的指導，以豐富具體的實例帶領學生實際操作。本節闡述如何在課堂上融入直接語料驅動學習活動，透過教師引導學生親自檢索語料庫，並對相關語料加以觀察、分析，來進行有針對性的華語教學活動。

#### 3.1 融入直接語料驅動學習的教學活動流程

Kennedy 與 Miceli (2001) 以中級義大利語的學習者為研究對象，讓學生以修改作文為目的，獨立進行語料庫檢索，探討此過程中學生可能遇到的挑戰。該文獻將語料庫檢索的過程區分成圖 2 的四個基本操作步驟。本文即在此基礎上，從教師引導的角度，提出融入直接語料驅動學習的教學流程與建議。

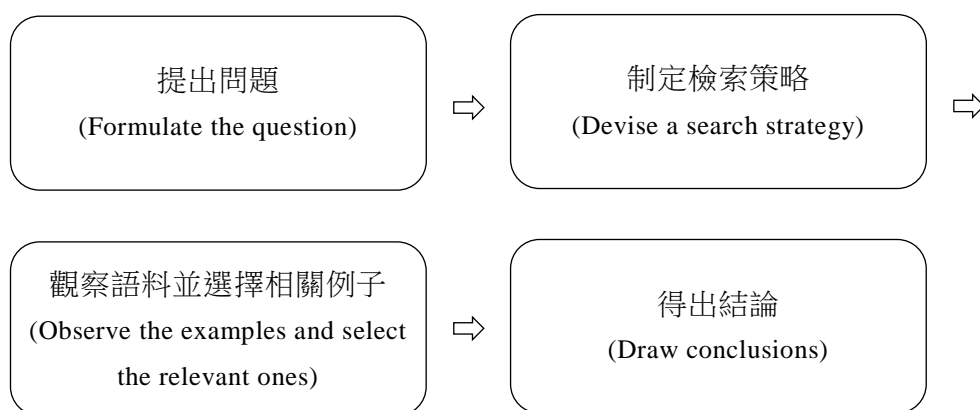


圖 2：融入直接語料驅動學習的教學活動流程

在課堂上進行融入直接語料驅動學習的教學活動時，應先由教師針對目標語言點提出特定問題。為了設計出明確的問題，教師須先行檢索語料庫，對相關語料進行分析，找出最適合學生學習的面向。若檢索結果中含有過多雜訊，則不適合讓學生觀察分析。提問方式盡量提供線索，提示學生接下來可如何檢索語料庫，找到問題的答案。教師還可對學生進行檢索的語料加以限定，以節省課堂時間。為了讓學生能夠制定出合適的檢索策略，教師有必要先以淺顯易懂的方式介紹語料庫（包含其主要功能與檢索方式等方面），培

養學習者的「語料庫素養」(corpus literacy) (Mukherjee 2006:14)。在觀察語料過程中，教師可提醒學生盡量避免對語料中不認識的詞彙逐一查詢。學生於分析語料後，可與同儕先行討論，初步歸納結果，再由教師確認結論的正確性。在整個操作流程中，教師可依情況適時介入引導，給予提示。

以下 3.2 小節首先說明教師如何給學生進行基本的語料庫教學，使其具備與語料庫「互動」的能力。其後三小節分別說明如何依圖 2 引導歸納式的教學流程，將直接語料驅動學習融入中、高年級華語課堂中的搭配教學、近義詞辨析教學、以及格式教學三大面向。

### 3.2 面向學習者的語料庫教學

語料庫教學的課前準備包括提醒學生攜帶筆記型電腦，或是螢幕較大的平板電腦、智慧型手機、電子閱讀器等可上網的裝置。教師也可選擇在電腦教室進行教學。另外，若選用的語料庫要求使用者登入後才能使用，教師可事先提供該語料庫的網址，要求學生先註冊、取得帳號，以節省課堂上寶貴的時間。

以下說明如何對中、高級程度以上的學習者進行循序漸進的語料庫教學。首先，教師應先解釋何謂「語料庫」，並簡介與語料庫相關的概念，如「索引行」、「搭配詞」、「頻率」等。這些概念的內涵不難理解，教師僅需給予清楚的定義，稍加解釋、適時舉例說明即可。此處不需過度延伸說明，例如，不必詳列語料庫的類型，也不必詳述一般語料庫建置的過程。學習者此時對於語料庫的了解一般已足以進行接下來的教學流程。

學生有了基本概念之後，教師即可引入計畫讓學生檢索的華語語料庫，簡單介紹其主要功能，為後面的教學活動做準備<sup>3</sup>。對於初次接觸語料庫的學習者而言，面對各種選項、數值設定等，難免望之卻步。因此，教學時宜避免一一說明各項設定，而應讓學習者體認到，大量語料在介面上一次呈現，語言之中的規律可一目瞭然。

華語語料庫根據不同面向可分為以下幾大類：書面語語料庫、口語語料庫、學習者語料庫、專業語料庫等（參陳浩然、潘依婷 2017）。教師可依課

---

<sup>3</sup> 教師若擔心立即引入華語語料庫，其全華語的介面可能讓學習者降低使用語料庫的意願，可依學生母語背景，先導入合適的非華語語料庫。例如，針對以英語為母語或英語能力較高的學習者，教師可先引入英語語料庫（如 *Corpus of Contemporary American English*），有助於提高學生對語料庫操作的熟稔度。



程教學目標以及學生漢語水平，選擇合適的語料庫。例如，若針對程度有限的學生，可選擇學習者語料庫，而非母語者語料庫。在專業華語課程，則可選用專業語料庫，例如，可使用商務漢語語料庫輔助商務漢語的教學。

在設計直接語料驅動學習的教學活動時，本文以一般語言課程的教學目標為導向，選用了兩個華語語料庫。第一個為 **Sketch Engine** 平台中較新的網路語料庫：**Chinese Simplified Web 2017 sample**<sup>4</sup>。使用者可透過該平台檢索許多語料庫，因此，教師宜先限定學生使用某個語料庫。操作介面以英語呈現，適合英語程度較高或是以英語為母語的學習者。接著，教師向學生介紹如何使用以下 **Sketch Engine** 的主要功能來檢索語料。

- 一、語彙索引 (Concordance)：使用者輸入某個詞語後，即可得到語料庫中含有該詞語的索引行，以 **KWIC** 方式呈現。教師還可視教學目標需要，示範如何進一步處理這些索引行。例如，可透過 **sort** 功能，將所有索引行依上下文所出現的詞語排序；亦可透過 **filter** 功能，篩選出符合需求的語料。
- 二、詞彙特性速描 (Word Sketch)：此項功能可以提供某個詞語的搭配詞，且這些搭配詞是根據與關鍵詞的語法關係加以分類。教師可先利用如「學生」這類學習者較熟悉的詞彙作為關鍵詞，介紹此功能的檢索結果中出現的重要語法關係：例如，「**Subject\_of**」（如：「學生＋學習」）、「**Object\_of**」（如：「引導＋學生」）、「**Modifies**」（如：「學生＋宿舍」）、「**A\_Modifier**」（如：「畢業的學生」）、「**Possession**」（如：「學生的能力」）、「**Measure**」（如：「位＋學生」）等。若點選某個搭配詞，即可看到大量含有該搭配用法的索引行，以 **KWIC** 方式呈現。
- 三、詞彙對照速描 (Word Sketch Difference)：使用者輸入兩個欲比較的詞語後，檢索結果中，綠底的詞語較常與第一個詞語搭配，而紅底的詞語較常與第二個詞語搭配；顏色深淺代表搭配關係的密切程度。  
「詞彙對照速描」與近義詞辨析教學特別相關，其優勢在於可將近

---

<sup>4</sup> 使用者可由以下網頁進入 **Sketch Engine** 的平台（可免費試用 30 天）：<https://app.sketchengine.eu/>。若學習者具備閱讀正體字語料的能力，可在以下網頁免費註冊帳號，使用中央研究院「中文詞彙特性素描系統」(Chinese Word Sketch Engine)：<https://wordsketch.ling.sinica.edu.tw/>。

義詞的搭配偏好透過顏色、圖像呈現，便於學生在短時間內更直觀地觀察到近義詞所搭配詞語的相似與相異之處，非常適用於語料驅動學習的初學者與詞彙量相對較少的華語學習者。以「漂亮」、「美麗」為例，檢索結果中 **Modifies** 欄位內的搭配詞語（圖 3）可進一步經由右上方「Show Visualization」功能加以視覺化（圖 4）。靠左側的詞語較常以「美麗」修飾，多與景色有關；而靠右側的詞語較常以「漂亮」修飾，多與女性有關。

除上述三大功能外，**Sketch Engine**也提供「詞彙表列」(Wordlist)、「同義辭典」(Thesaurus) 等其他功能，教師可視教學目標與需要，斟酌是否介紹給學生。



	漂亮 11,996x	美丽 11,548x
裙子	33	0
李慧珍	23	0
女生	38	0
女孩子	36	6
脸蛋	35	8
女人	90	69
图画	7	49
景色	12	80
花朵	5	46
风景	8	161
风光	0	73
画卷	0	74

圖 3：「詞彙對照速描」檢索結果示例



圖 4：「詞彙對照速描」檢索結果視覺化圖表示例

本文所使用的第二個華語語料庫為「北京語言大學語料庫中心（BLCU Corpus Center，簡稱BCC）（荀恩東、饒高琦、肖曉悅、臧嬌嬌 2016）<sup>5</sup>。BCC 提供多種語體的漢語語料，包括文學、報刊、微博、科技等，總字數約150億字，為目前最具規模的華語語料庫之一。在BCC線上搜尋語料時，所輸入的檢索式具有線性（linear）的特點，不僅能輸入單詞，亦可輸入詞串（如「說的也是」）、格式（如「v 來 v 去」）或是搭配組合（如「強烈的 n」）。BCC 還可統計檢索式在語料庫中的分布情況，所有實例可依頻率高至低排列呈現。教師可根據網站上的使用說明，簡單說明搜尋語料的方式。檢索式中變項成分的詞性以英文字母表示，例如，以 n 代表名詞、以 v 代表動詞、以 a 代表形容詞等。教師可同樣以「學生」為例，針對某一語體的語料，示範如何制定檢索式，所舉的例子可涵蓋不同詞性（如「學生的 n」、「v 學生」、「a 學生」等），取得相關統計結果（圖5為微博語料中「學生的 n」的部分統計結果），教師還可從統計結果點選其一（如「學生的感覺」），展示如何查看含特定高頻搭配組合的索引行。

<sup>5</sup> 使用者可由以下網頁免費線上檢索 BCC 的語料：<http://bcc.blcu.edu.cn/>（使用時不必註冊或登錄）。該網頁還提供功能介紹（包括檢索、篩選、統計等功能）、檢索式示例、檢索式說明等相關訊息。



共 1212 个结果

下载      首页   上页   下页   末页

学生的作文	165	学生的时候	138
学生的感觉	134	学生的文章	94
学生的老师	79	学生的家长	62
学生的心理	59	学生的素质	56

圖 5：BCC 微博語料中「學生的 n」的部分統計結果<sup>6</sup>

教師可透過各種檢索任務，引導學生比較不同語料庫所提供的訊息。此一做法主要的優點為，學生未來面對各式各樣的語言問題時，較能選用合適的語料庫來找出答案。

### 3.3 搭配教學示例

詞語之間的搭配關係在二語學習過程中無疑為一大難點。本小節提供兩種搭配關係（「動詞＋名詞」和「量詞＋名詞」）的教學示例，說明如何善用語料庫工具所提供的訊息，提高學習者對詞語搭配的意識，加強其詞彙知識的深度。

#### 3.3.1 動詞與名詞的搭配

教師可從學生正在學習的單元中，挑選數個生詞，教授動詞與名詞（賓語）的搭配。本小節以動詞「推出」為例，依圖 2 中的四個步驟，說明如何將直接語料驅動學習活動用於教授此動詞所搭配的名詞。

（一）提出問題：教師可從動詞出發，針對與其搭配的名詞（賓語）提問；反之亦可，即從名詞出發，針對其作為賓語時所搭配的動詞提問。具體

<sup>6</sup> 在圖 5 的統計結果中，學習者可能會特別注意到，「學生的時候」此詞串之中，「學生」與「時候」之間的關係，不同於「學生」與該畫面中其他名詞之間的關係。教師可藉此向學生說明，BCC 所提供的檢索結果為所有符合檢索式的用例，並未對其進行語法關係的分析，這點與 Sketch Engine 不同。

的範例如下：

請使用 Sketch Engine 中的網路語料庫 (Chinese Simplified Web 2017 sample) 來回答以下問題<sup>7</sup>：哪些名詞最常作為「推出」這個動詞的賓語？這些名詞在意思上有什麼類似的地方？

(二) 制定檢索策略：透過教師先前對 Sketch Engine 的介紹 (3.2 小節)，學生應可判斷出此處該使用「詞彙特性速描」(Word Sketch) 這項功能，以得到常與「推出」搭配的賓語。

(三) 觀察語料並選擇相關例子：「詞彙特性速描」功能所提供的搭配詞，是根據與「推出」之間的語法關係分類。要找出其常搭配的賓語，學生應選擇觀察 Object 欄位之下的搭配詞，頻率最高的包括：「車型」、「產品」、「新品」、「系列」、「版本」等。

(四) 得出結論：「推出」這個動詞最常搭配的名詞通常是某項產品、活動或服務，而與產品相關的名詞更具多樣化，包括其型號、系列、版本等。

以上動名搭配教學示例也適用於其他搭配關係（如形容詞與名詞）的教學，下一小節則聚焦漢語中較為獨特的量詞與名詞搭配關係。

### 3.3.2 量詞與名詞的搭配

量詞是漢語的特點之一，也是華語學習者較容易出現偏誤的詞類之一。本小節依圖 2 中的四個步驟，提供兩個教學示例，說明如何透過直接語料驅動學習活動，引導學生觀察、分析量詞與名詞的搭配關係。

在進行詞彙教學時，若學習的目標詞語為名詞，有時須同時讓學生知道常與該名詞搭配的量詞。此時，便可進行類似如下的教學活動：

(一) 提出問題：教師可從名詞出發，針對與此名詞搭配的量詞提問。具體的範例如下：

哪些量詞最常與「新聞」這個名詞搭配？請使用 Sketch Engine 中的網路語料庫 (Chinese Simplified Web 2017 sample) 以及 BCC 中的報刊語料來回答以上問題，並比較兩個語料庫提供的答案有哪些相同和不同的地方。

<sup>7</sup> 此處的提問範例並未要求學生使用 BCC。若教師備課時先行使用 BCC 以「推出 n」來進行檢索，便會發現檢索結果中含大量雜訊，不利學生在觀察後得出預期的結論。

(二) 制定檢索策略：學生進入 **Sketch Engine**，選擇了指定的網路語料庫後，再使用「詞彙特性速描」這項功能，即可得到「新聞」的搭配詞。教師可提示學生，在 **BCC** 輸入檢索式時以 **q** 代表量詞。學生起初想出的檢索式可能較為簡單，例如「**q 新聞**」或「**一 q 新聞**」，教師亦可視情況引導學生制定出較為複雜的檢索式，以搜索出較完整的語料<sup>8</sup>。此處宜使用「**[m r]q 新聞**」，該檢索式表示量詞前可出現數量詞 (**m**) 或指示代詞 (**r**)，方括弧代表其內所列出的任一出現即可 (即「或」的關係)。

(三) 觀察語料並選擇相關例子：學生應從「詞彙特性速描」統計結果中觀察 **Measure** 欄位之下的搭配詞，包括「則」、「日」、「條」、「篇」、「周」等。在 **BCC** 中，若使用「**[m r]q 新聞**」這個檢索式，可以得知在報刊語料裡，較常與「新聞」搭配的量詞包括「項」、「條」、「個」、「則」、「周」等。兩個語料庫的檢索結果皆出現了一些雜訊。因此，教師應引導學生，點選每個量詞，瀏覽其索引行，才能決定該量詞是否確實為與「新聞」搭配。以 **Sketch Engine** 的網路語料庫為例，出現了「日」、「周」等雜訊，學生觀察其索引行後可發現，這些雜訊大多源自固定的詞串。例如，「每日經濟新聞」似乎是某個新聞網站的專欄名稱，在語料庫中出現多次，經斷詞後，語法分析器 (**parser**) 將「日」誤析為「新聞」的量詞。**BCC** 中亦有類似的雜訊來源。例如，學生在統計功能所產生的頻率表中點選「一項新聞」，則可發現該詞串通常出現於「一項新聞公告」之中，顯然「新聞」用於修飾「公告」，而「項」是搭配「公告」的量詞，而非搭配「新聞」。若學生未查看索引行，則很可能將誤析的用例與正確的混為一談。在教學活動中，教師除了提問、檢視學生搜尋的過程以及歸納出的結論外，亦應引導學生對語料進行批判性的思考。

(四) 得出結論：兩個語料庫的檢索結果皆顯示「新聞」常與「則」、「條」這兩個量詞搭配，而檢索結果的相異之處則多為雜訊。教師可以此鼓勵學生探索不同的語料庫，方能透過比對、檢核，獲得可靠的結果。得出結論後，教師可進一步引導學生觀察索引行，分析「新聞」與不同量詞搭配時的差異。

---

<sup>8</sup> 儘管學生想出的檢索式有些無法取得所有包含目標詞語搭配的用例，然而，在實際課堂教學時，不須拘泥於此目標。只要檢索結果足以讓學習者觀察、分析，並得出大致的結論、回答提問，即已達成學習目標。教師先行搜索語料庫時，可先從學生容易上手的檢索式出發，然而，若檢索式過於簡略而導致檢索結果雜訊過多，仍應對檢索式加以限定。

當學習的目標詞語為量詞，而學生不知該量詞除了與課本所列舉的例子搭配以外，尚可與其他哪些名詞搭配，便可從量詞出發，針對與此量詞搭配的名詞進行教學。具體的操作流程如下：

（一）提出問題：除了引導學生搜尋常與此量詞搭配的名詞，還可比較不同語體的語料。具體的範例如下：

哪些名詞最常與「片」這個量詞搭配？哪些是具體名詞？哪些是抽象名詞？請使用 Sketch Engine 中的網路語料庫 (Chinese Simplified Web 2017 sample) 以及 BCC 中的文學和報刊語料來回答以上問題，並比較檢索結果有哪些相同和不同的地方。

（二）制定檢索策略：學生使用 Sketch Engine 檢索指定的網路語料庫時，在「詞彙特性速描」這項功能之下，輸入「片」。然而，檢索結果中並沒有任何欄位直截了當地列出與該量詞搭配的名詞。使用 BCC 時，大多數學生可能會制定出如「片 n」或「一片 n」這類較為簡單的檢索式，有些學生可能會考量到，量詞前除了數量詞外還可能與指示代詞連用，制定出較為複雜的檢索式，例如「[m r]片 n」，因而取得更為完整的語料。

（三）觀察語料並選擇相關例子：在 BCC 中輸入「[m r]片 n」此檢索式，透過其統計功能可得到如圖 6（文學語料）和圖 7（報刊語料）的結果。學生可清楚地觀察到，在兩個不同語體中「片」與名詞（以及數量詞／指示代詞）的搭配偏好，有何相似與相異之處。此外，教師還可引導學生瀏覽索引行，觀察各個搭配用法常見的語境。例如，「一片空白」之前通常出現與「腦」相關的名詞（如「大腦」、「腦子」、「腦袋」等）。透過觀察某個搭配組合的常見用法，有助於學習者歸納出該搭配組合的典型語境。如此一來，學生不但能學到目標詞語的各個搭配詞，亦能學到各個搭配組合的用法，兼具廣度與深度。

共 5433 个结果

下载 首页 上页 下页 末页

这片土地	2021	一片废墟	677
一片空白	456	一片爱心	439
这片热土	430	一片欢呼声	428
一片掌声	414	一片深情	402
一片绿色	376	一片火海	340

圖 6：BCC 文學語料中「[m r]片 n」的部分統計結果

共 2910 个结果

下载 首页 上页 下页 末页

这片土地	308	一片空白	280
一片好心	128	一片树叶	113
一片树林	100	两片嘴唇	96
一片叶子	92	一片废墟	88
一片空地	82	一片面包	74

圖 7：BCC 報刊語料中「[m r]片 n」的部分統計結果

(四)得出結論：透過比對兩個子語料庫的檢索結果，學生應可歸納出，在文學語體中，常與「片」搭配的名詞較為具體，而在報刊語體中的則較為抽象。教師還可進一步鼓勵學生嘗試將出現在檢索結果中的抽象名詞分類(許多與聲音、情感有關)，引導學生歸納出「片」的引申用法，深化對該量詞的認識。



以上示例表明，就量詞與名詞的搭配而言，**Sketch Engine** 所提供的訊息對學習者的幫助較為有限。相比之下，**BCC** 的檢索方式較符合直覺，檢索結果亦相當直觀，且教師可善用其中多個子語料庫，引導學生比較量詞與名詞在不同語體之中的搭配用法。

綜合本節所有的搭配教學示例可知，不同語料庫工具應用於教授各類詞語搭配關係時，皆有其優勢與侷限。教師須提醒學生不應被動地全盤接受語料庫所提供的檢索結果，而應對搜尋出的語料進行批判性思考，方能避免將雜訊誤認為學習要點。此歷程在直接語料驅動學習的初期階段格外重要，有助於學習者建立正確的心態與做法，提高課外使用語料庫獨立學習的能力。

### 3.4 近義詞辨析教學示例

隨著二語學習者整體程度提高，詞彙量增加，如何掌握目標語中近義詞的細微差異也益趨重要。洪煒（2013）的實證研究結果顯示，在課堂上就近義詞辨析進行顯性教學（**explicit instruction**），對習得近義詞的差異有顯著的效果。該研究亦表明，與傳統的接受式教學模式相比，發現式教學模式在促進近義詞習得與保持教學效果方面，皆有較大的優勢。**Liu 與 Zhong**（2016）則建議教師於教學中盡可能運用語料庫分析，引導學生注意最常與目標近義詞搭配的詞語和其他相關語境訊息。本小節以「尊敬」與「尊重」為例，依圖 2 的四個步驟，說明如何將直接語料驅動學習融入於課堂上近義詞辨析教學之中。

（一）提出問題：教師針對近義詞辨析提問時，應力求明確，避免過於寬泛的問題（如「這兩個詞語意思有何不同？」）。為此教師應自行先檢索語料庫，對統計結果和索引行進行分析，在設計提問時，聚焦於能夠有效協助學生辨析目標近義詞的搭配關係。提問的面向不宜過多，以兩到三個為宜，以免學生覺得過於複雜。以下示例中的提問並非以目標語呈現，而是使用學習者較熟悉的語言（假設學生的母語為英語），以避免學生因不熟悉中文的語法術語、不了解提問，而無法進行活動。此外，對於某些學習者來說，「尊敬」與「尊重」此組近義詞的難點在於其英文釋義非常接近，以英文提問則可有效引導學生於辨析時與母語之中的對應詞比較。以下為具體提問範例：

Consult the corpus—Chinese Simplified Web 2017 sample—via Sketch Engine to answer the following questions:

1. The verb “to respect” in English can take human objects (as in *to respect your parents*) or abstract ones (as in *to respect privacy*)? Which one of the Chinese words is more commonly used with human and abstract objects, respectively?
2. What do the nouns that are frequently modified by 尊敬 have in common? Analyze the relevant concordance lines to find out the typical context in which the Chinese word is used as a modifier.

(二)制定檢索策略：透過之前教師對 Sketch Engine 的簡介(3.2 小節)，學生應能判定與近義詞相關的問題應使用其中「詞彙對照速描」(Word Sketch Difference) 功能，並輸入兩個比較詞語（即「尊敬」與「尊重」）來進行檢索。

(三)觀察語料並選擇相關例子：學生須依據提問的線索從檢索結果所顯示的諸多語法關係之中，選擇最相關的部分來回答問題。針對第一個問題，學生應選擇觀察 Object 欄位。如圖 8 所示，常作「尊敬」賓語的詞語包括「長者」、「祖先」、「師長」等，而常作「尊重」賓語的則包括「權利」、「選擇」、「人格」等。教師還可引導學生點擊「Show Visualization」以得到較直觀、視覺化的結果，也可鼓勵學生點擊向下箭頭符號以得到更多檢索結果。第二個引導問題針對的是「尊敬」所修飾的對象，因此學生應觀察 Modifies 欄位中綠底的詞語，包括「賓客」、「客人」、「客戶」等名詞（圖 8）。學生還須點擊所列詞語旁的「…」圖標，進一步瀏覽相關的索引行，例如，點選「尊敬 + 客戶」即可觀察到「尊敬的客戶」後常接「您好」或「感謝」等禮貌用語（圖 9）。

尊敬 3,137×				尊重 10,278×			
Object				Modifies			
长者	4	0	...	宾客	30	0	...
祖先	4	0	...	客人	34	0	...
师长	82	10	...	网友	75	0	...
长辈	29	23	...	纳税人	17	0	...
老人	38	16	...	客户	134	0	...
老师	49	49	...	用户	135	0	...
权利	0	106	...	态度	4	22	...
选择	0	141	...	职业	0	8	...
人格	0	71	...	基础	0	24	...
劳动	0	102	...	同时	0	6	...
规律	0	234	...	需要	0	10	...
意愿	0	216	...	前提	0	7	...

圖 8：「尊敬」與「尊重」詞彙對照速描部分檢索結果

1	  yhj9.com	哎，红包也不容易拿啊。</s><s>要红酒客服：尊敬 的客户，您好！</s><s>感谢您的惠顾与支持哦！
2	  yhj9.com	光临，要红酒客服Fairy祝您生活愉快！</s><s>尊敬的客户，您好！</s><s>一般17:00前下的订单可
3	  ysslc.com	<s><s>关于预约兑换2017年贺岁普通纪念币的公告 尊敬的客户：</s><s>中国人民银行将于2017年1月4E
4	  icxo.com	”的短信声，这是银行发来的第二条短信：“尊敬的客户，由于利率调整，你下个月应还金额为3
5	  weldinfo.net	<s>联系人：李小姐，联系电话：028-83</s><s>尊敬的客户：</s><s>您好！</s><s>如果中国焊接资讯
6	  072550.cn	家，苦心经营，做良心产品，用心做事购物须知 尊敬的客户：感谢您选购香港鸿博办公家具产品，我
7	  jd.com	感怎么样是纯法国红酒吗？</s><s>年份是哪年 尊敬的客户，您好！</s><s>您咨询这款酒是2014年
8	  lawze.com	：lawyer分类：委托必读标签：关于我们</s><s>尊敬的客户：</s><s>为了节省您宝贵的时间，避免笔
9	  cxfund.com.cn	寸宝充值，什么时候才能添加新的银行卡</s><s>尊敬的客户您好，感谢您对长信基金的关注与支持，
10	  letfind.com.cn	描述</s><s>房源描述房源描述房源描述房源描述 尊敬的客户您好非常感谢您能在百忙之中抽空关顾我！

圖 9：「尊敬+客戶」部分索引行

（四）得出結論：就第一個問題而言，學生觀察語料後應可總結出此對目標近義詞作動詞用時，與之共現的賓語類型有所不同：與人相關的名詞常作「尊敬」的賓語，而抽象名詞僅能作「尊重」的賓語。教師還可進一步引導學生歸納出「尊敬」的賓語多為年齡更長者。就第二個問題而言，學生瀏覽索引行後應可歸納出，「尊敬」與其所修飾的名詞出現的典型語境為商業書信或公告開頭的稱呼，而「尊重」則沒有此種用法。

以上示例闡釋了如何在教師的提問引導下，讓學生直接檢索語料庫、觀察語料統計結果、閱讀分析索引行，最後總結出近義詞實際用法的差別。此發現式的過程可讓學生透過較多用例，更具體地了解如何使用目標詞語。

### 3.5 格式教學示例

格式為結構較為固定的詞組，一般含有兩部分，即常項詞語和變項詞語，兩者搭配表達特定的格式義。然而，格式的意義不一定可完全從其組成成分的語意得知，且格式中詞語之間的搭配關係也有所限制。二語學習者若能精確掌握格式的用法，便能有效提高目標語的交際能力。本節以「既…又…」為例，依圖 2 中的四個步驟，說明如何將直接語料驅動學習融入在此格式的教學中。

（一）提出問題：教師提問時必須限定目標格式之中變項成分的詞類。此外，若使用的語料庫涵蓋不同語體，教師可進一步限定學生在某一子語料庫中進行檢索。具體的範例如下：

在 BCC 的科技語料中，哪些形容詞最常一起出現在「既…又…」這個格式？這些形容詞是單音節的多還是雙音節的多？一起出現在這個格式的形容詞在形式或意義上有什麼特點？

（二）制定檢索策略：根據問題的條件，學生應在 BCC 中選擇科技子語料庫，並以「既 a 又 a」（a 代表形容詞）進行檢索，透過「統計」功能便可得到如圖 10 的檢索結果。教師也可讓學生嘗試在 Sketch Engine 平台上檢索，使其認知到在該系統中無法使用格式作為檢索詞。

共 1656 个结果

下载 首页 上页 下页 末页

既费时又费力	64	既古老又年轻	38
既省时又省力	29	既简单又复杂	24
既美观又实用	21	既合理又合法	17
既简单又实用	17	既抽象又具体	17
既简单又方便	16	既省时又省钱	14

圖 10：BCC 科技語料中「既 a 又 a」的部分統計結果

（三）觀察語料並選擇相關例子：具有中級水平以上的學習者應至少可看懂圖 10 中「既古老又年輕」、「既省時又省力」、「既簡單又方便」等相關例子，即使不認識統計結果中的某些形容詞，也能得出初步結論。

（四）得出結論：學生透過語料庫的檢索不難發現，進入「既…又…」的形容詞大多為雙音節成分。此外，從統計結果中，學生亦可歸納出，在科技語料中該格式的高頻用例大多帶有正面的色彩。此外，在這些用例中共現的形容詞許多具有共同語素（如「既費時又費力」）或語意相反（如「既抽象又具體」）的特點。對學習者而言，這些發現很難僅從課本所提供的幾個例句中得到。若學生提出其他合理的觀察點，教師也應給予鼓勵。回答問題之後，教師還可進一步引導學生閱讀相關索引行，使其對此格式的使用語境有更完整的認識（如「既古老又年輕」大多用於描述城市、行業或學科等）。

教師可視其需要進行延伸活動，例如，學生在初級課堂已學過「又…又…」，此格式與「既…又…」用法類似，但較為口語化。教師可針對該格式設計提問，引導學生透過觀察、分析語料，比較兩者使用情況的異同，深化學習者對已學語言點的理解與應用。最後，教師也可鼓勵學生比較同一格式在不同語體中的分布和使用情況。

第三節闡述了如何將直接語料驅動學習融入於華語課堂中的搭配教學、近義詞辨析教學和格式教學。對於教授不同類型的語言點，各華語語料庫之

適用程度有所差別。教師讓學生親自檢索語料庫前，須先培養學習者的「語料庫素養」。指派給學生的檢索任務務必明確，針對每個目標詞語或格式的教學面向不宜過多。為了減少學習者因訊息量過大而引發的焦慮感，教師可引導學生先對相關統計結果進行觀察分析，再進一步瀏覽含高頻或特殊用法的索引行，如此較容易讓學習者獲得成就感。

#### 4. 融入間接語料驅動學習的教學設計

先前文獻已指出，在實際課堂中進行直接語料驅動學習的活動，侷限較多。相比之下，基於間接語料驅動學習的教學活動較為靈活，不必在電腦教室中進行或要求學生自備電腦，學習者也不須接受語料庫檢索的訓練。教師在設計間接語料驅動學習活動時，應考慮以下三個方面：

- 一、語料庫的選用：Whistle (1999) 強調，索引行對教學的價值和用處大多取決於語料庫本身。在進行間接語料驅動學習活動的過程中，學習者不須親自檢索語料庫。因此，在選擇語料庫時，較為重要的標準為，是否能從其語料庫中篩選出具有代表性和可讀性的索引行，供學習者觀察、分析。
- 二、索引行的篩選：教師於課前須對語料庫中含目標語言點的語料進行篩選。為了避免徒增學生進行間接語料驅動學習活動時的工作量，教師須根據學習者的語言程度，篩選出學生可理解範圍內的索引行作為例句，避免選用過長、生詞過多、或結構過於複雜的語料。最後，索引行數量可視目標語言點的使用複雜程度以及課堂時間而定，建議每個活動不超過十條。
- 三、索引行的排版：在篩選出適合學生分析的索引行之後，下一個考量點便是如何將這些語料呈現給學生。為了讓學習者迅速注意到目標語言點的語境，建議教師排版時採用 KWIC 的方式，也就是關鍵詞在各行語料中整齊地呈現於正中央，且使用粗體或是套色等醒目提示。如此排版方式可幫助學習者有意識地注意到「輸入」(input) 中的特定語言形式特徵，進而轉化為「吸收」(intake) (參 Schmidt 1990)，有利於目標語言點的習得。

本節闡述如何在課堂上融入間接語料驅動學習活動，透過引導學生對教師從語料庫篩選出的索引行進行觀察、分析，並歸納出語言規則，從中調動

學習者在學習過程中的主動性和積極性。

#### 4.1 融入間接語料驅動學習的教學活動流程

Carter 與 McCarthy (1995) 指出，由於口語的許多特色未於現有的語法書和教科書中完全涵蓋，在教授口語語法時，傳統的 3P 教學法（即 Presentation-Practice-Production）並不完全合適。對此，他們提出 3I 教學策略，即 Illustration-Interaction-Induction。第一步「觀察分析」(Illustration) 是讓學生觀察、分析真實（口語）語料；第二步「分享討論」(Interaction) 是讓學生分享、討論自己對語料的初步觀察；第三步「總結歸納」(Induction) 是讓學生歸納出自己的（或小組的）語言規則，這規則可再藉由分析更多語料來慢慢修正。Flowerdew (2009) 則在此 3I 的基礎上增加第四個步驟：「介入引導」(Intervention)，由教師視情況在「總結歸納」之前或之後執行，給學生提供提示性訊息。

在以上文獻基礎上，本文提出間接語料驅動學習的教學流程（圖 11）。此流程包含六個步驟。除上述四個步驟外，前後各加了一個步驟：「暖身導入」(Introduction) 和「應用表達」(Implementation)。下一小節將說明如何將此引導歸納式的教學流程應用於課堂上的句式教學。

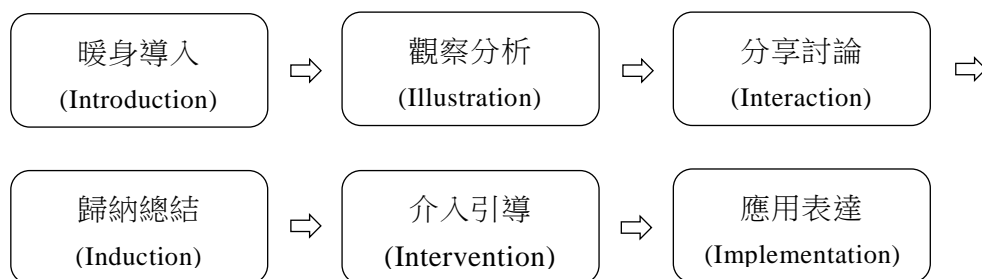


圖 11：融入間接語料驅動學習的教學活動流程

#### 4.2 句式教學示例

在傳統的華語課堂中進行句式教學時，大多由教師先行對句式規則進行講解，大量操練相關形式，從而讓學生習得該語法點。此傳統演繹式的教學方法與融入語料驅動學習的教學活動，本質截然不同。本小節以「無論…都…」此語法點為例，闡述如何將間接語料驅動學習融入中級華語課堂中的句式教學，引導學生觀察、分析教師從語料庫中篩選出的索引行，進而歸納出

句式規則，並產出正確的句子。

一般而言，課本將「無論…都…」列為中級語法點。以《中文聽說讀寫》(Integrated Chinese)為例，該句式首次出現於第三冊。此套教材將「無論」譯為「regardless of...; whether it be」，對「無論…都…」這個語法點提出以下英文說明，並提供(A)到(D)四個例句 (Liu, Yao, Bi, Shi and Ge 2018:109-110)：

無論, meaning “regardless,” signifies that a result will remain the same under any condition or circumstance. It must be used with either a question pronoun, as in (A) and (B), or an alternative construction, as in (C) and (D).

(A)明天無論誰請客我都不去。

(B)Q：你想去城裏的哪個購物中心？

A：城裏的購物中心我都没去過，所以無論去哪個都可以。

(C)我們已經說好明天去買東西，你無論願意不願意都得跟我們去。

(D)他無論在家裡還是在學校，都在學習，很少看見他玩兒。

「無論…都…」此句式的生成性很高，且形式眾多，但從上可知，一般教材限於篇幅，不太可能在舉例時涵蓋所有使用情況。為了讓學生對「無論…都…」的使用情況有更完整的認識，我們依據語料驅動學習的理念並參考文獻所指出的常見偏誤 (Bai 2009)，訂定了以下教學目標：

透過對語料庫索引行的觀察與分析，學習者能夠：

一、歸納出「無論…都…」的數種常見使用形式。

二、主動發現：(a)「都」的存在，以及該副詞在句中的位置；(b)

「無論…都…」出現在選擇疑問句時，該問句中的連詞應為「還是」，而非「或者」。

三、進行自我糾錯，正確使用該句式。

考慮到「無論…都…」的教學對象為具有中級漢語水平的學習者，在本教學活動中選用北京語言大學 HSK 動態作文語料庫 (2.0 版)<sup>9</sup>。該語料庫為最具代表性的華語學習者語料庫之一，語料總數達到 11,569 篇作文，共計 424

<sup>9</sup> 讀者可從以下網址登入此語料庫並獲得更多相關訊息：<http://yuyan ziyuan.blcu.edu.cn/info/1043/1501.htm>。



萬詞。語料經斷句處理，錯誤亦經過類型標記。教師在備課時可透過語料中出現的錯誤，提前了解學生可能出現的偏誤，從而設計更具針對性的教學活動。本文從該語料庫一共選了十條索引行，涵蓋「無論…都…」的多種基本使用形式。索引行節選自修改後的作文，皆以完整句子呈現。排版時以 KWIC 形式將「無論」置中，並使用粗體、加上底線，可讓學生立即注意到此關鍵詞（圖 13）。

以下便依圖 11 的六個步驟逐一說明操作原則，完成針對「無論…都…」的教學流程需要約一小時的課時。

（一）暖身導入（10 分鐘）：導入目標語言點時，教師可設計圍繞同一主題的暖身活動，結合「形式聚焦」(focus on form) (Long 1991) 教學法，即在以溝通和意義為中心的課堂中，引導學生關注語言形式。此處以「語言誓約」(Language Pledge) 為主題，引入「無論…都…」這個句式（圖 12）。學生須根據教師提供的引導提示，嘗試使用目標句式在此主題下造句（如「無論在哪兒，都不可以說英文」）。與傳統教學不同的是，此步驟為學習者的試錯階段，教師暫且不對學生的造句提供回饋意見或糾錯。

In a number of Chinese immersion programs, students are asked to sign a language pledge promising that they will only speak Chinese for the duration of the program, both in and outside of the classroom. Imagine you are going to one of these programs for Chinese learning:

☞ 在這兒學習的時候，可以說英文嗎？

(Please create sentences using “無論” ‘regardless of; no matter’ based on the hints provided.)

1. 學生、老師
2. 在宿舍、在教室、在圖書館、在咖啡館 ...
3. 吃飯的時候、上課的時候、運動的時候 ...
4. 跟同學、跟朋友、跟老師 ...
5. 累、不累

圖 12：「無論…都…」暖身導入活動示例

（二）觀察分析（15 分鐘）：在此步驟，學生須快速閱讀教師提供的語料，並進一步歸類，分析「無論…都…」的形式。教師須在此提醒學生，閱讀語料時能理解句子大概意思即可，毋須知道每個詞的意思。由圖 13 可看出，例句 1、2、3、4、6、8 皆含有疑問詞短語，屬於同一類；例句 7 和 9 含有選

擇疑問詞短語，屬於同一類；例句 5、10 皆含有正反疑問短語，屬於同一類。

1.	<u>无论</u> 有多吵我都睡得着了。
2.	<u>无论</u> 你年齡多大都有父母。
3.	我 <u>无论</u> 是什么事都告诉我妈妈。
4.	因为 <u>无论</u> 是谁，都喜欢好听的音乐。
5.	他爱喝酒， <u>无论</u> 身体好不好，都找借口喝。
6.	我非常喜欢流行歌， <u>无论</u> 哪个国家的流行歌都喜欢。
7.	<u>无论</u> 是老人还是小孩儿，我们都听音乐。
8.	<u>无论</u> 在什么时间，什么地方，都可以看到人。
9.	<u>无论</u> 玩电脑、学习还是走路，我都一边听音乐一边唱歌。
10.	<u>无论</u> 自己抽不抽，如果自己的旁边的人吸烟，都会损害健康。

圖 13：「無論…都…」索引行示例

（三）分享討論（5 分鐘）：在完成步驟二後，教師可將學生分組，討論各自的觀察，實現生生互動。

（四）歸納總結（10 分鐘）：在此步驟，學生以小組形式，對語料中出現的「無論…都…」形式進行歸納、總結。初次使用語料驅動學習活動進行句式教學時，教師可用學生已熟悉的語法點為例（如「一…就…」），說明如何描述句式規則。

（五）介入引導（10 分鐘）：教師於此時確認各小組初步歸納出的語言規則。如有需要，可進一步提示或修正。例如，若學生針對每條索引行各寫一條規則，教師則可提醒學生注意索引行之間類似之處。

（六）應用表達（10 分鐘）：為了強化學習者對該語法點的了解與掌握，教師須設計不同的表達練習，鞏固學習成效。此時可先讓學生重新檢視在暖身導入活動時所造的句子，應用所學進行修改與自我糾錯（如將遺漏的「都」補上）。教師還可根據學生對目標句式的掌握程度，進行不同形式的操練活動。

上述程序可依實際教學情況、需求，彈性調整。例如，若課堂時間有限，流程中的暖身導入和觀察分析活動可安排為預習作業，部分應用表達練習也

可安排為複習作業，如此一來，學生可在不受時間壓力之下完成任務。教師還可提供額外的語料，針對目標句式的非典型用法，進行延伸教學。例如，語料中可出現「無論」與「都」以外的副詞（如「也」、「皆」等）共現的用例，讓已掌握基本用法的學習者深化對該語法點的認識。

本小節說明如何將間接語料驅動學習活動融入至課堂上的句式教學。在此過程中，學生不再被動地接受教師提供的語法規則，而是基於教師篩選的語料，主動地對語言形式進行分析與歸納。此學習模式還可訓練學生提高對生詞的容忍度，採用選擇性、掃描的策略閱讀文本，有別於以往線性、逐字逐句的方式。

## 5. 結語與建議

隨著語料庫工具日益普及，相關資源逐漸被視為語言教學的利器，語料驅動學習模式應運而生。本文旨在探討如何將不同類型的語料驅動學習與各種華語課堂教學內容有機地結合起來，分別針對其直接和間接模式，提出引導歸納式的系統化教學流程，並藉由具體教學示例，說明語料驅動學習活動設計原則，這些示例涵蓋搭配教學、近義詞辨析教學、格式教學以及句式教學四個面向。在課堂上引導學生直接檢索語料庫、或間接地分析教師從語料庫篩選出的真實語料，學習者便可透過這種發現、歸納式的方法，意識到語言使用的規律，進而促進華語的習得。最終目標為使學習者具備自主學習的能力，能夠獨立使用語料庫工具解決在華語學習過程中遇到的難題，從而提高自己的語言水平。

本文根據文獻的研究結果，結合作者的實踐經驗，針對以下四個面向提出幾點操作建議：

- 一、語言點的選擇：某些語言點，例如複雜的詞語搭配關係、差別細微的近義詞、生成性較高的格式或句式、仰賴語境的虛詞等，一般教材僅透過有限的例句，不易簡明扼要地將其用法解釋清楚。而透過語料驅動學習對此類語言點進行教學，學生可藉由閱讀較多實例，接觸到更多相關形式與用法，進一步培養對目標語的語感。在初始階段，教師亦可考慮先以已學過的語言點來進行其延伸用法的教學，學習者可在較短的時間內完成語料驅動學習任務，累積信心與成就感。

- 二、教師方面的任務：教師的課前準備工作應儘量達到細緻、充分。教師須事先檢索慎選的語料庫，結合學生的語言程度及教學目標，設計具有針對性的活動，課堂提問力求具體、明確，以利學生按照教師要求、快速進行有目的性的檢索。在進行語料驅動學習的課堂教學活動時，教師應循序漸進，起初以較封閉式的問題引導，慢慢過渡到較開放式的檢索，以免學生在語料中迷失，感到挫折，且浪費課堂時間。隨著學生對此學習方法的逐漸適應及語言程度的提升，教師可適量減少直接幫助，逐步增加間接引導，最終實現學習者的自主學習。
- 三、學習者方面的考量：對於習慣於傳統教學模式的學生而言，初次體驗語料驅動學習時，難免會產生牴觸的心理。教師應事先與學生溝通，使其了解進行語料驅動學習活動的目的及好處，緩解學生的焦慮感和排斥感。相關文獻指出，語料驅動學習活動對中、高級學習者的效果更為顯著 (Granath 2009)。對於詞彙量較豐富、漢字閱讀水平也較高的學習者而言，從事語料驅動學習活動有助於深化所學知識，對此模式的接受度也較高。然而，語料驅動學習其相對開放的特質以及對分析推理能力的要求，對某些學習者來說會更具挑戰。因此，教師需要針對學習者的特點在教學活動中進行適當調整。
- 四、活動設計的多樣性：教師備課時，須考量語料驅動學習活動設計的多樣性。多元的活動不僅可以避免學生產生倦怠感，同時可以滿足學習者各自不同的學習偏好 (Chambers 2007)。教師在設計語料驅動學習活動時，可善用語料庫的不同功能和數據呈現方式，增添活動多樣性。

綜上所述，在具體實踐中，教師應綜合教學目標及學習者的特點，以靈活多樣的 forms 將語料驅動學習融入課堂教學活動之中。儘管此模式在諸多方面具有優勢及益處，然而並不能全面取代傳統的課堂教學。教師應思考如何將語料驅動學習任務作為常規教學的有效補充，藉以豐富課堂活動的種類，刺激學習者使用不同的學習策略，提升其主動性和自主性，以達到更佳的教学效果。本文希望能起到拋磚引玉的作用，期待引起更多華語教師對語料驅動學習的關注和探討，並為後續探討此模式學習成效的實證研究，提供啟發和參考。

## 引用文獻

- Bai, Jianhua. 2009. *Chinese Grammar Made Easy: A Practical and Effective Guide for Teachers*. New Haven and London: Yale University Press.
- Boulton, Alex. 2010. Data-driven learning: Taking the computer out of the equation. *Language Learning* 60.3: 534-572.
- Boulton, Alex. 2012. Hands-on/hands-off: Alternative approaches to data-driven learning. *Input, Process and Product: Developments in Teaching and Language Corpora*, eds. by James Thomas, and Alex Boulton, 152-168. Brno: Masaryk University Press.
- Boulton, Alex. 2017. Corpora in language teaching and learning. *Language Teaching* 50.4: 483-506.
- Boulton, Alex, and Tom Cobb. 2017. Corpus use in language learning: A meta-analysis. *Language Learning* 67.2: 348-393.
- Carter, Ronald, and Michael McCarthy. 1995. Grammar and the spoken language. *Applied Linguistics* 16.2: 141-158.
- Chambers, Angela. 2007. Popularising corpus consultation by language learners and teachers. *Corpora in the Foreign Language Classroom*, eds. by Encarnación Hidalgo, Luis Quereda, and Juan Santana, 3-16. Amsterdam: Rodopi.
- Chambers, Angela. 2010. What is data-driven learning? *The Routledge Handbook of Corpus Linguistics*, eds. by Anne O'Keefe, and Michael McCarthy, 345-358. London and New York: Routledge.
- Chen, Howard Hao-jan, Jian-cheng Wu, Christine Ting-yu Yang, and Iting Pan. 2016. Developing and evaluating a Chinese collocation retrieval tool for CFL students and teachers. *Computer Assisted Language Learning* 29.1: 21-39.
- Flowerdew, Lynne. 2009. Applying corpus linguistics to pedagogy: A critical evaluation. *International Journal of Corpus Linguistics* 14.3: 393-417.
- Flowerdew, Lynne. 2015. Data-driven learning and language learning theories: Whither the twain shall meet. *Multiple Affordances of Language Corpora for Data-driven Learning*, eds. by Agnieszka Leńko-Szymańska, and Alex Boulton, 15-36. Amsterdam and Philadelphia: John Benjamins.

- Gilquin, Gaëtanelle, and Sylviane Granger. 2010. How can data-driven learning be used in language teaching? *The Routledge Handbook of Corpus Linguistics*, eds. by Anne O'Keefee, and Michael McCarthy, 359-370. London and New York: Routledge.
- Granath, Solveig. 2009. Who benefits from learning how to use corpora? *Corpora and Language Teaching*, ed. by Karin Aijmer, 47-65. Amsterdam: John Benjamins.
- Johns, Tim. 1986. Microconcord: A language-learner's research tool. *System* 14.2: 151-162.
- Johns, Tim. 1991. Should you be persuaded: Two examples of data-driven learning. *Classroom Concordancing*, eds. by Tim Johns, and Philip King, 1-16. *Birmingham English Language Research Journal* 4.
- Johns, Tim. 1994. From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *Perspectives on Pedagogical Grammar*, ed. by Terence Odlin, 293-313. Cambridge: Cambridge University Press.
- Kaltenböck, Gunther, and Barbara Mehlmauer-Larcher. 2005. Computer corpora and the language classroom: On the potential and limitations of computer corpora in language teaching. *ReCALL* 17.1: 65-84.
- Kennedy, Claire, and Tiziana Miceli. 2001. An evaluation of intermediate students' approaches to corpus investigation. *Language Learning & Technology* 5.3: 77-90.
- Kilgarriff, Adam, Nicole Keng, and Simon Smith. 2015. Learning Chinese with the Sketch Engine. *Corpus Linguistics in Chinese Contexts*, eds. by Bin Zou, Simon Smith, and Michael Hoey, 63-73. London: Palgrave and Macmillan.
- Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíče, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, and Vít Suchomel. 2014. The Sketch Engine: Ten years on. *Lexicography* 1.1: 7-36.
- Leech, Geoffrey. 1997. Teaching and language corpora: A convergence. *Teaching and Language Corpora*, eds. by Anne Wichmann, Steven Fligelstone, Tony McEnery, and Gerry Knowles, 1-23. London: Longman.
- Liu, Dilin, and Shouman Zhong. 2016. L2 vs. L1 use of synonymy: An empirical study of synonym use/acquisition. *Applied Linguistics* 37.2: 239-261.

- Liu, Yuehua, Tao-chung Yao, Nyan-ping Bi, Yaohua Shi, and Liangyan Ge. 2018. *Integrated Chinese* [中文聽說讀寫] (4<sup>th</sup> edition), Vol. 3. Boston: Cheng & Tsui Company Inc.
- Long, Mike. 1991. Focus on form: A design feature in language teaching methodology. *Foreign Language Research in Cross-Cultural Perspective*, eds. by Kees de Bot, Ralph B. Ginsberg, and Claire Kramsch, 39-52. Amsterdam: John Benjamins.
- Mair, Christian. 2002. Empowering non-native speakers: The hidden surplus value of corpora in continental English departments. *Teaching and Learning by Doing Corpus Analysis*, eds. by Bernhard Kettemann, and Georg Marko, 119-130. Amsterdam: Rodopi.
- Mukherjee, Joybrato. 2006. Corpus linguistics and language pedagogy: The state of the art—and beyond. *Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods*, eds. by Sabine Braun, Kurt Kohn, and Joybrato Mukherjee, 5-24. Frankfurt am Main: Peter Lang.
- O’Sullivan, Íde. 2007. Enhancing a process-oriented approach to literacy and language learning: The role of corpus consultation literacy. *ReCALL* 19.3: 269-286.
- Römer, Ute. 2011. Corpus research applications in second language teaching. *Annual Review of Applied Linguistics* 31: 205-225.
- Schmidt, Richard W. 1990. The role of consciousness in second language learning. *Applied Linguistics* 11.2: 129-158.
- Smith, Simon. 2011. Corpus-based tasks for learning Chinese: A data-driven approach. *The Asian Conference on Technology in the Classroom Official Conference Proceedings 2011*, 48-59.
- Smith, Simon, Alice Chen, and Adam Kilgariff. 2008. A corpus query tool for SLA: Learning Mandarin with the help of Sketch Engine. *Corpus Linguistics, Computer Tools, and Applications—State of the Art*, ed. by Barbara Lewandowska-Tomaszczyk, 673-686. Frankfurt am Main: Peter Lang.
- Sripicharn, Passapong. 2010. How can we prepare learners for using language corpora? *The Routledge Handbook of Corpus Linguistics*, eds. by Anne O’Keefee, and Michael McCarthy, 371-384. London and New York:

Routledge.

- Vyatkina, Nina. 2016. Data-driven learning of collocations: Learner performance, proficiency, and perceptions. *Language Learning & Technology* 20.3: 159-179.
- Vyatkina, Nina. 2020. Corpora as open educational resources for language teaching. *Foreign Language Annals* 53.2: 359-370.
- Warren, Martin. 2016. Introduction to data-driven learning. *The Routledge Handbook of Language Learning and Technology*, eds. by Fiona Farr, and Liam Murray, 337-347. London and New York: Routledge.
- Whistle, Jeremy. 1999. Concordancing with students using an “off-the-Web” corpus. *ReCall* 11.2: 74-80.
- Xu, Jiajin. 2019. The corpus approach to the teaching and learning of Chinese as an L1 and an L2 in retrospect. *Computational and Corpus Approaches to Chinese Language Learning*, eds. by Xiaofei Lu, and Berlin Chen, 33-53. Singapore: Springer.
- Yeh, Meng, and Xian Zhang. 2018. Corpus-based instruction: Teaching discourse-linking *jiu* (就) in storytelling. *Chinese as a Second Language* 53.1: 1-23.
- 洪煒。2013。〈漢語作為第二語言的近義詞教學實證研究〉，《世界漢語教學》，第 27 卷第 3 期，424-432。[Hong, Wei. 2013. An empirical study on the teaching of near-synonyms to learners of Chinese as a second language. *Chinese Teaching in the World* 27.3: 424-432.]
- 荀恩東、饒高琦、尚曉悅、臧嬌嬌。2016。〈大數據背景下 BCC 語料庫的研製〉，《語料庫語言學》，第 3 卷第 1 期，93-118。[Xun, Endong, Gaoqi Rao, Xiaoyue Xiao, and Jiaojiao Zang. 2003. The construction of the BCC Corpus in the age of Big Data. *Corpus Linguistics* 3.1: 93-118.]
- 陳浩然、潘依婷。2017。〈語料庫與華語教學〉，《語料庫與華語教學》，1-47。臺北：高等教育出版。[Chen, Hao-jan Howard, and I-ting Pan. 2017. Corpora and teaching Chinese as a second language. *Corpus and Teaching Chinese as a Second Language*, ed. by Hao-jan Howard Chen, 1-47. Taipei: Higher Education.]



[審查：2020.6.30 修改：2020.8.6 接受：2020.9.9]

王炳勻

Ben Pin-Yun WANG

Department of Pacific and Asian Studies

University of Victoria

Clearihue Building C205, Victoria, BC V8P5C2, Canada

benpywang@gmail.com

許展嘉

Chan-Chia HSU

10051 臺北市濟南路一段 321 號 國立臺北商業大學通識教育中心

Center for General Education

National Taipei University of Business

No.321, Sec. 1, Jinan Rd., Taipei City 10051, Taiwan

chanchiah@gmail.com

龍水水

Shuishui LONG

Department of Asian Studies

The Pennsylvania State University at University Park

102 Old Botany Building, University Park, PA 16802, U.S.A.

kellylong523@gmail.com

丁曉穎

Xiaoying LILES

Department of East Asian Languages and Cultures

Indiana University at Bloomington

355 North Jordan Avenue, Bloomington, IN 47405-1105, U.S.A.

dingxiao@iu.edu

## **Designing Data-Driven Learning Activities for the Chinese as a Second Language Classroom**

**Ben Pin-Yun WANG**  
**Department of Pacific and Asian Studies**  
**University of Victoria**

**Chan-Chia HSU**  
**Center for General Education**  
**National Taipei University of Business**

**Shuishui LONG**  
**Department of Asian Studies**  
**The Pennsylvania State University at University Park**

**Xiaoying LILES**  
**Department of East Asian Languages and Cultures**  
**Indiana University at Bloomington**

### **Abstract**

Little attention has been paid to the use of corpus tools in the Chinese as a second language classroom. To fill this gap, this article explores how to incorporate distinct approaches to data-driven learning (DDL) in teaching different aspects of Chinese lexico-grammar in the classroom. We have proposed systematic instructional procedures that involve hands-on and hands-off DDL, respectively. Such guided-inductive procedures are further illustrated by model DDL instructional activities, which cover the teaching of collocation, near-synonyms, as well as phrasal and sentential patterns. The inductive DDL tasks designed in this project require learners to directly consult Chinese corpora, or indirectly analyze the concordance lines prepared beforehand, all under the guidance of the instructor. The short-term goal is to facilitate the learning of Chinese by helping students discover language patterns based on corpus data.

The ultimate goal is to develop learner autonomy, empowering students with the ability to use corpora as a learning aid.

**Keywords:** concordance lines, corpus consultation, discovery learning, inductive learning, learner autonomy