

應用網路商務華語語料庫——以搭配詞為例*

王敬淳^a

陳浩然^{b1}

潘依婷^b

^a國立臺灣師範大學華語文教學研究所

^b國立臺灣師範大學國語中心暨英語學系

摘要

因應現今中外貿易交流的熱絡及外籍商務人士的倍增帶來的高度商務華語教學資源需求，及商務華語教學資源現況上的缺乏，本研究應用「將網路作為語料庫」(Web as Corpus)概念，通過網路資源於短時間內蒐集大量網路商務華語語料，並以此為基建置網路商務華語語料庫。對欲學習商務華語的外籍商務人士而言，了解商務華語語料的搭配詞資訊是流利使用商務華語的關鍵，本研究因而進一步使用智慧詞搭配搜尋引擎，配合網路商務華語語料庫之語料，提供豐富生動的網路商務華語搭配用法及大量的範例詞句供商務華語學習者、華語教師及教材編纂者使用。為突顯本網路商務華語語料庫之商務特性，本研究選用 19 個具商務性質之名詞為關鍵字透過與中央研究院平衡語料庫搭配詞檢索結果之比較，發現相對於一般綜合性之語料庫，本網路商務語料庫及智慧詞搭配搜尋引擎能檢索出更多與商務活動相關的搭配詞使用。例如：本網路商務華語語料庫檢索出的「商品」搭配詞有 19 筆，且商務特質高；而中研院平衡語料庫搭配詞檢索結果僅有 1 筆。本研究期望此網路商務華語語料庫及智慧詞搭配搜尋引擎對現今商務華語教學之迫切需求應當有所助益。

關鍵詞：商務華語 網路作為語料庫 搭配詞 語料庫應用 專業華語

* 本研究感謝教育部「邁向頂尖大學計畫」與科技部「跨國頂尖研究中心計畫」(MOST 104-2911-I-003-301),以及國立臺灣師範大學「華語文與科技研究中心」支持。

¹ 本文通訊作者

1. 前言

1.1 國外專業英語的發展概況

專業英語 (English for Specific Purposes, CSP)，是指用於某種專業領域、特定範圍和固定場合的英語，以培養學生具有某種專業領域的溝通技能為教學目標，如商務英語、外交英語、法律英語、觀光英語等。1964年，系統功能語言學創始人M. A. K. Halliday、McIntosh和Stevens合著出版了《語言科學與語言教學》(*The Linguistic Sciences and Language Teaching*) 一書，最早提出了專業英語的概念，他們首次系統性地闡明，語言教學應根據學習者的具體需要，決定教學內容和方法的原則，這標誌著應用語言學及語言教學領域裡對專業英語研究的開端。於是20世紀60年代迄今，專業英語逐步成了國外英語教學及語言學領域的一個獨立分支，不但研究成果豐碩，相關教材也不斷研發出版，許多大學（如伯明罕大學、英國阿斯頓大學）更設立了專業英語的碩士學位。同時，在英語系國家也有許多提供給英語學習者的專業英語課程，甚至還出版了許多國際性專門討論專業英文的期刊及書籍，如Elsevier出版的*English for Specific Purposes: An international Research journal*、劍橋大學出版的*Developments in English for specific purposes: A multi-disciplinary approach*及*English for Specific Purposes: A Learning-Centred Approach*、密西根大學出版的*English for Specific Purposes in Theory and Practice*、Pauline C. Robinson撰寫的*ESP (English for Specific Purposes) : The Present Position (1980)*。

此外，國外也發展了不少專業英語的語料庫，如培生出版集團設置的培生學術英文國際語料庫(Pearson International Corpus of Academic English, PICA)，從教科書、研討會、講課內容、電視廣播等蒐集了37萬英語學術口語與書面語語料，涵蓋了人文學科、社會科學、自然科學、應用科學四大學科；密西根大學設置的密西根學術英文口語語料庫(Michigan Corpus of Academic Spoken English, Micase)，從學術講座、課堂、面試等蒐羅了約一百八十五萬字的學術口語語料，呈現了許多在學術環境中所發生的自然對話內容；由瓦立克大學(University of Warwick)與雷丁大學(University of Reading)建置的英式學術口語英文(British Academic Spoken English, BASE)，收錄了160個講座和40場研討會的口語語料，共包含了一百六十萬多字，涵蓋了藝術與人文、生命與醫學科學、物理科學、社會研究和科學四大學科；以及劍橋大學建置的劍橋商務英語語料庫(Cambridge business English corpora)，從商務會議、電話會話、口頭報

告、報紙等蒐集了200萬字的商務口語及書面語語料，已被用於編纂劍橋Business Advantage系列的英語教材；劍橋學術英語語料庫（Cambridge Academic English Corpus），從美國及英國的大學、研究所的講課內容、口頭報告、教科書、論文中蒐集了400萬個學術口語及書面語語料，已被用於編纂Cambridge Academic English系列的英語教材；還有劍橋法律英語語料庫（Cambridge Corpus of Legal English），從各種有關法律及司法程序的書籍、期刊及報紙中蒐集法律英語語料。

早在20世紀60年代，語言學家便察覺到發展專業英語研究的重要性，隨後在專家學者的努力下，專業英語逐步成了國外英語教學及語言學領域的一項獨立分支。迄今，專業英語無論在研究、教材，或是語料庫方面，均取得相當豐碩的成果。總體而言，國外專業英語的發展可說是已趨近成熟。

1.2 專業華語發展現況

國外專業英語已蓬勃發展，至於國內的專業華語（Chinese for Specific Purposes, CSP），也在近年來成爲海內外學習華語者漸趨重視的一個領域。根據趙悅（2007）的統計，每年有30%至40%的來華留學生以學習醫科、經濟、管理、工科、法學、教育、理科等專業爲目標，可見學習者對專業華語的需求非常大，專業華語相關的課程、教材與研究已成爲華語教學界關注的領域。

在20世紀80年代初，國外留學生開始進入中國高等院校學習相關專業，從那時起中國高等院校就先後開設了科技華語、醫用華語、理工華語等課程，並編寫了相關的教材（呂必松 1990）據不完全統計，迄今中國大陸出版的各類專業華語教材不下於150種，其中商務類華語教材占絕大多數（李泉 2011）。1982年，第一本商務華語教材《外貿洽談500句》（北京語言學院和北京外貿學院合編，華語教學出版社 1982）出版，迄今，專業華語教材的種類和數量不斷增加，到2006年，根據路志英的統計，中國大陸光是商務華語教材就至少已有45種。而臺灣方面，臺灣師範大學國語中心於1990年編寫了《實用商業會話》，並於2000年完成修訂；葉德明等人也在2011年編寫了《遠東商務漢語》。至於香港方面，以會話爲主的商務漢語教材則有莊棕嬰與張增增（2005）所主編，由中文大學所出版的《商業漢語》。

除教材外，專業華語教學相關研究亦逐步開展。從80年代後，不斷有專業華語教學研究及教材編寫研究的文獻問世，如杜厚文〈漢語科技文體的語言特

點) (1981)、石金才〈關於編寫外交人員初級漢語教材的設想〉(1990)、王硯農〈談談中醫漢語系列教材〉(1992)、張黎〈商漢語教學需求分析〉(2006)、程樂樂等〈醫學漢語教材編寫芻議〉(2008)、李曉琪〈系列商務漢語教材編寫探索〉(2010)等;臺灣方面,則有張惠雯《中醫華語教材編寫之規畫設計》(2002)、金惠淑(2005)《觀光華語導遊教材設計—針對韓國的專門大學觀光中國語系之用途》、陳淑惠《法律華語教材設計與研究》(2008)、杜昭玫〈商務漢語教材之內容分析與建議〉(2011)、李育娟(2011)〈商務華語學習者需求分析〉、陳麗宇及李欣欣〈當代初級商務華語教材之文化研究〉(2012)等等,這些研究成果對專業華語教學及教材編寫無疑有著重要的啟發意義。

然李泉(2011)認為,總體來看目前專業華語教學、教學理論研究、教材編寫及研究仍相當薄弱,對此,他歸納出七點理由:(1)現有華語教學學科理論體系研究的重要論著及相關文獻,基本上都還未提及專業華語這個領域,表明人們對這一領域重視不夠。(2)專業華語教學自身的基礎理論研究還未廣泛地開展,迄今還沒有形成一套成熟的專業華語教學理論和多種完善的教學模式。(3)各類專業華語教學還缺乏教學大綱、教學標準和教學規範。(4)現有專業華語教學領域還不夠寬泛,發展也不夠平衡。(5)教材種類還不夠齊全,缺乏品牌教材,許多領域尚無正式出版的教材。(6)現有教材研究成果主要集中在早期的科技華語和近年的商務華語教材的編寫研究,而且大都是教材編寫構想或所編教材的理論和實踐總結,僅有此還不夠,迄今有關專業華語教學及教材研究的文獻不足百篇,專著更是鳳毛麟角。(7)目前專業華語教學師資少之又少。

由以上七點可知,專業華語教學雖已取得了一些教學經驗和研究成果,但總體上並沒有受到應有的重視,亦沒有獲得較成熟的發展,故無法與通用華語教學並駕齊驅。因此,為滿足各類專業華語學習者的需求,開展專業華語教學理論研究和教材實勢在必行。

1.3 商務華語語料庫建置

目前專業華語教學、教學理論研究、教材編寫及研究雖已取得了一些成果,但在專業華語語料庫方面,相較於國外各種專業英語語料庫的建設成果,專業華語語料庫的建置仍顯得不足。而語料庫不僅是語言研究和自然語言處理研究的資料來源,其提供的大量實際語料,對語言教材的編纂及課堂教學也具有相當重要的價值,故各類專業華語語料庫的建置實有其必要性和重要性。

21 世紀以來由於中外經貿交流的擴大，經貿華語特別是商務華語教學研究及教材編寫呈現一枝獨秀的趨勢。據李曉琪（2010）的統計，目前中國大陸已出版的商務漢語教材已逾 80 種，在各類專業華語教材中占了大多數。此外，為評估第一語言非漢語者從事商務活動所應具有的漢語水準，中國國家漢辦亦在 2006 年研製了《商務漢語考試大綱》，「商務漢語考試」(Business Chinese Test, BCT) 已在新加坡、韓國、美國、日本、泰國、法國等 20 多個國家舉行，以上種種顯示商務華語在各種專業華語教學研究中是最為熱門的。

綜合以上所述，相較於其他各種專業華語語料庫（如法律華語、中醫華語等），商務華語語料庫的建設顯得迫切許多。有鑒於此，本研究擬發展一個商務華語語料庫，並嘗試提出將商務華語語料運用在教材編寫及課堂教學的方式，也希望本研究建置語料庫技術及方法能夠拋磚引玉，為國內更多相關研究提供參考。

2. 文獻探討

2.1 商務語料庫現況

目前國際上的商務語料庫仍不多見，較知名的當屬劍橋商務英語語料庫 (Cambridge business English corpora) 以及香港理工大學商務語料庫 (The Hong Kong Polytechnic University Business Corpora)。

2.1.1 劍橋商務英語語料庫 (Cambridge business English corpora)

劍橋大學建置的劍橋商務英語語料庫 (Cambridge business English corpora) 收錄了總字數達 200 萬字的商務口語及書面語語料，其中更收錄了來自劍橋大學與諾丁漢大學共同開發的劍橋-諾丁漢商務英語語料庫 (Cambridge and Nottingham Business English Corpus, CANBEC) 的口語語料。語料來源為商務會議、電話會話、口頭報告、報紙等，已被用於編纂劍橋 Business Advantage 系列的英語教材，但目前不提供對外使用。

2.1.2 香港理工大學商務語料庫 (The Hong Kong Polytechnic University Business Corpora)

香港理工大學英語系開發了一個總字數一千兩百多萬，包含了口語及書面語語料的「語言庫」(Language Bank)，其中又根據語料性質歸類為一般 (general)

語料庫、學術 (academic) 語料庫、商務 (business) 語料庫、新聞 (journalistic) 語料庫、法律/公文 (legal and documentary) 語料庫以及文學 (literature) 語料庫。其中，由 Li Lan 博士和 Grahame Bilbow 教授於 1999 至 2000 年間開發的商務語料庫，從商務報告、書信、備忘錄、新聞、報紙、傳單等收錄了大量的商務語料，由總字數分別約為 116 萬字、123 萬字、132 萬字的英語、中文、日語三大語料庫組成，並提供線上檢索的功能。

2.1.3 其他商務語料庫

除了以上二商務語料庫外，由深圳職業技術學院應用外國語學院亦於2006年建立了一個商務英語語料庫，從商務企業、商業信函、電子郵件、年度報告、商務書籍與教材等擷取了600多萬字的商務語料；關西大學染谷泰正 (SOMEYA Yasumasa) 教授也開發了一個商務信件語料庫檢索系統 (Online Business Letter Corpus KWIC Concordancer)，從商務信件中蒐集商務語料，並提供線上檢索功能。曼徹斯特大學 Mike Nelson 也建置了約100萬字的商務英語語料庫 (The Business English Corpus, BEC)，從商務相關書籍、電子郵件、傳真、報紙、報告等材料中蒐集口語與書面語商務語料，目前不對外公開使用。

2.2 商務語料庫的教學應用

Leech (1997) 將語料庫在語言教學的應用劃分為直接應用和間接應用。前者指學生和教師在語言課堂上使用語料庫輔助教學，從而影響教與學的方式，即「數據驅動學習」(Data-Driven Learning)；後者指將語料庫運用於工具書的出版、教材開發和語言測試等，幫助擬定教學內容。

語料透過中心詞檢索 (concordancing)、關鍵字分析 (keyword analysis)、詞塊分析 (cluster analysis) 等方法，可從詞彙、詞頻、搭配詞、句型、語義及語用等方面進行分析，這些分析結果可直接或間接地應用在商務語言教學課程大綱設計、測驗、教材編寫等方面，周玉林 (2010) 即在〈試論「商務英語語料庫」的建設及其應用〉一文中，對商務語料的教學應用作了進一步的說明：

(1) 運用於教學大綱設計與教材編寫

教學大綱的編寫應當以實證為依據，而不是以直覺為基礎 (Mindt 2001)。語料庫中所反映的語言使用情形可為教學大綱設計提供實證依據，有助於安排教學內容先後順序。商務語料庫即可用來計算商務詞彙的詞頻和分佈情況，從而制定商務詞彙表，使之成為商務語言教學大綱設計、教材編寫的重要依據。

(2) 直接運用於商務語言教學

商務語料庫亦可直接應用在商務語言教學中。教師可先向學生說明語料庫及語料分析工具等的使用方式，引導學生使用語料庫進行資料分析。學生一旦掌握了語料庫的檢索方式，就可通過語料庫進行商務詞彙與通用詞彙、句法等方面的比較、進行詞語搭配、主題分析、詞彙語義和話語分析等，從而培養自主學習的能力。

此外，也能讓學生運用語料庫觀察真實語料，從大量語料中觀察到某一語言現象，然後討論和分享在語料中的發現，形成與語料庫的互動。接著再歸納該語言現象的規則，並在教師的指導下，通過觀察更多語料，逐步修正歸納出的規則，即「數據驅動學習」(Data-Driven Learning) 的學習模式。

(3) 運用於商務語言測驗

語料庫可透過中心詞索引、搭配詞索引及詞頻統計工具，展示商務語料庫中，詞彙在不同商務語境中的詞頻、用例和詞與詞間的細微差別，並可對一些疑難用法、慣用搭配和語篇結構特徵進行歸類及排序。此分析結果可為編制試題提供真實素材，確保試題的代表性和真實性，提升試題的效度。當前，語料庫已廣泛用於國際大型英語測試中，如美國教育考試服務中心(ETS)建立了The T2K - SWAL Corpus，運用於TOFEL、GRE、GMAT等考試中；英國劍橋大學英語認證測試考試委員會(UCL)建立的Cambridge Learner Corpus、Business English Text Corpus等，也用於幫助設計劍橋大學英語認證測試。

綜上所述，商務語料庫的建置迄今仍主要以英語為主，相關的教學應用亦多以英語教學為大宗。而隨著商務華語教學的重要性與日俱增，商務華語語料庫的建置及相關教學應用實應有所推展。對此，本研究首先以商務華語語料庫的建置為目標，進而以此研發搭配詞檢索系統，期望透過商務語料在搭配詞方面的應用，推廣商務華語語料庫對商務華語教學帶來的助益。

2.3 搭配詞對詞彙教學的重要性

搭配(collocation)指的是兩個語言單位的組合，如兩個單詞組成詞組。最早提出「搭配」這一概念的是英國語言學家Firth(1957)，他將collocation稱為一種「結伴關係」，其他學者的定義也與其相似，如Halliday和Cowie(1964)都說明搭配是詞彙的「共現」。陸國強(1986)則在《現代英語詞彙學》一書中指出「詞的搭配關係主要指詞與詞之間的橫組合關係，即什麼詞經常與什麼詞

搭配使用。」，他還進一步提出，英語學習到一定階段時，進一步提高的方法之一就是圍繞中心內容進行聯想，釐清詞與詞之間的搭配關係，建立「聯想場」的概念。Sinclair 和 Renouf (1988) 在談到詞彙教學大綱時亦指出，詞彙學習的焦點應該指向：(1) 最常用的詞形；(2) 這些詞形的主要使用模式；(3) 這些詞形所構成的典型組合。換句話說，學習者不但要學會詞彙的發音、拼寫和詞典裡的意思，還要學習它最常用的詞形和搭配。歷來許多研究者皆肯定了搭配詞的重要性，如 Alexander (1984) 認為詞組的學習包括成語和各種固定用法應是詞彙教學的重點，他建議，詞彙教學應重視 3 個 C：Collocation (搭配)、Context (語境)、Connotation (內涵意義)。Woolard (2000) 亦在文章中提出「要學習更多的詞彙並不只是要學習新詞彙，而是要學習舊詞彙的新的搭配用法。」Lewis (2000) 也點出他觀察到許多學生常造出合乎文法但母語者覺得奇怪的句子，原因是學生並不熟悉詞彙的搭配用法。Nation (2001) 同樣指出要使用流利適切的語言，搭配詞的知識是不可或缺的。由此可知，學習者在學習第二語言詞彙的過程中，若不能將詞彙和與之搭配的詞彙合併起來學習，便常會受母語思維的負遷移影響，以母語的方式表達漢語的詞彙搭配，這樣將不利於培養學習者漢語詞彙聯想能力，學習者於是難以習得正確的華語表達方式。

學習者使用搭配詞的困難在海內外皆深受許多詞彙教學研究者關注，如 Bahns and Eldaw (1993)、Farghal 及 Obiedat (1995)、Gitsaki (1996) 皆透過翻譯測驗及克漏字測驗調查了英語非母語的英語學習者使用搭配詞的能力，測驗結果確實指出，英語學習者無論國別，皆有使用搭配詞的困難。詞彙搭配不僅是英語學習者的難點，亦是華語學習者難以克服的障礙，王建勤 (1997)、全香蘭 (2006)、胡明揚 (2006)、馬玉汴 (2006)、高燕 (2008)、彭增安 (2007)、董政、鄭艷群 (2008)、劉亞菲、鄭艷群 (2008)、蕭頻、張妍 (2008)，都在其文章中指出了詞彙搭配不當是華語學習者學習華語時極為常見的問題，張博 (2008) 更在其著作中援引對外漢語教學界多項研究，證實外籍學生不論其母語背景為何，所犯的詞彙方面的偏誤數量遠遠超過其他方面的偏誤。

由上可知，搭配詞教學應成為詞彙教學的重點。在詞彙教學中介紹與核心詞彙搭配的字詞及其用法，藉由豐富學生的詞彙搭配知識，能幫助學習者從語義的理解提升到語用的高度，了解詞彙的文體特徵與功能，是提高學習者語言綜合運用能力的方式之一。

有鑒於搭配詞對詞彙教學的重要性，本研究將以商務華語語料庫為基礎，建置搭配詞檢索系統。藉此平臺將有助於了解商務情境中的常見搭配詞，豐富

學習者在商務語境中的詞彙搭配知識。並透過簡潔的功能設置，增加華語教學相關研究或教學人員使用上的便利和檢索速度，以利商務華語教學或教材編纂時，提供更豐富的商務華語資源。對此，本研究將首先闡述商務華語語料庫以及搭配詞搜尋引擎的建置方式和過程。接著以搭配詞搜尋引擎的應用為主軸，試舉出多項商務語境下的搭配詞實例，並對比中央研究院平衡語料庫之搭配詞檢索結果，論述搭配詞搜尋引擎作為商務華語教學輔助工具所帶來的益處。可望藉由搭配詞搜尋引擎的豐富檢索結果以及多元化的檢索功能，為商務華語教學提供更實質有效的教學參考和幫助。

3. 建置商務華語語料庫智慧搭配詞搜尋引擎

隨著網路科技的日益發達，語料庫建置的方式在近年來有了技術上的革新。早在1999年，Philip Resnik就開始運用網路蒐集語料，在其1999年的研究中，即展示了利用不同語言版本的網頁來建置平行語料庫的方法。Rosie Jones及Rayid Ghani (2000) 亦以建構塔加拉族語料庫為例，展示了一個運用網路建構語料庫的方法。Atsushi Fujii及Tetsuya Ishikawa (2000) 也以網路作為語料來源，揭示了自網頁中蒐集專業科技術語定義的方法。到了2003年，知名語料庫專家Adam Kilgariff和Gregory Grefenstette (2003) 再度強調了「將網路作為語料庫」(Web as Corpus) 的概念。他們指出，因為網路資料量大及取得容易的優勢，善用搜尋引擎，蒐集網路上數量龐大的語料，將是建置語料庫較便捷、有效率的方式。與此同時，愈來愈多的語言科學家及技術專家也傾向將網路作為語言資料的來源。時至今日，將網路作為語料庫語言學的研究來源已成為當今趨勢。學者們在2005年組成了「將網路作為語料庫」(Web as Corpus) 的工作坊，至2014年，已舉辦了9次的研討會，發表了不少應用網路語料庫(Web Corpus) 進行研究的範例。將網路作為語言研究的語料來源，如今已在國外蔚然成風，然國內卻較少相關的研究。現存之華語語料庫也多半採用傳統作法，直接從教科書、書信、備忘錄、新聞、報紙、傳單等蒐集語料。為增進蒐集語料的速度與效率，本研究嘗試引進此種運用網路建置語料庫的方法，以WebBootCat為工具，建造一個網路商務華語語料庫，並用其語料進一步建成商務華語語料庫智慧搭配詞搜尋引擎。

3.1. 商務華語語料庫建置

為突破以往華語語料庫多從出版物、新聞、報紙等來源，耗時費力地一一蒐集語料的困境，本研究嘗試透過將網路作為語料庫的概念蒐集語料。藉由 WebBootCat 得以在短時間內蒐集大量各種各樣語料的優勢，以此作為語料蒐集的主要工具，將得以大量縮短語料庫建置的時間與人力成本。

WebBootCat 係以 BootCat 軟體為基礎所開發的線上工具，而 BootCat 意指 Bootstrapping Corpora and Terms，為 Baroni 和 Bernardini 於 2004 年所提出的語料蒐集方式。BootCat 軟體主要透過種子字詞 (seed words) 在網路資源中進行搜尋，而後將符合種子字詞的網頁內容集結起來形成語料庫。透過 BootCat 軟體蒐集語料最大的特色在於整個語料蒐集的過程皆可透過電腦自動處理，且語料庫內容可根據自訂的種子字詞而有所調整，更有助於專業領域的語料蒐集。Baroni et al. (2006) 為方便使用者免於軟體下載、安裝等步驟，進而開發了 WebBootCat，也就是線上版的 BootCat 軟體，可供使用者直接在 WebBootCat 網頁上執行語料蒐集工作。

透過 WebBootCat 軟體蒐集語料，首先須選定種子字詞 (seed words)。本研究利用 WordSmith¹ 軟體的詞頻統計功能，分析葉德明教授主編之商務華語教材《遠東商務漢語》，計算出教材字詞的出現頻率，列出由頻率高至低排列的詞表。再經人工篩選，從內挑選出出現頻率最高，且與商業領域相關的專業華語詞彙，做為種子字詞，共計 56 個：公司、產品、市場、合同、客戶、經理、服務、商品、銷售、貿易、貨物、成本、運輸、生產、廣告、包裝、保險、付款、增加、客人、消費者、報價、信用、業務、代理、部門、交易、費用、錢、消費、數量、賣方、交貨、資金、買、買方、品質、秘書、商務、專業、提供、業績、運費、名片、洽談、材料、訂單、促銷、飛機、需求、樣品、工廠、存款、免費、海運、航班。

將 56 個種子字詞鍵入 WebBootCat 軟體，選擇繁體中文為搜尋的語言，並設定需要的資料量和搜尋的檔案類型，WebBootCat 更將這 56 個種子字詞隨機結合成檢索詞在網路中進行搜尋，如：「商量」與「價錢」結合起來搜尋、「促銷」與「商品」結合起來搜尋。搜尋結果會自動列出這些種子字詞出現頻率最高的

¹ WordSmith 由英國利物浦大學的語言學家 M. Scott 所發展之著名語料分析工具，於 1996 年推出 1.0 版，目前已升級至 6.0 版。WordSmith 提供英文、中文等 80 種不同語言的處理，其最主要的功能包括：字詞索引 (concordance)、詞頻統計 (wordlist) 及關鍵詞檢索 (keywords)。

網頁，這些網頁中的所有文字即為本研究商務華語語料庫之語料來源。本研究以 56 個與商業領域有關的種子字詞做為關鍵字，共搜尋到 2677 個網頁位址，共計語料量約為 5 百 55 萬餘字。接者，藉中央研究院中文斷詞系統²對 555 萬商務語料自動斷詞，斷詞處理完成後統整為本研究之商務華語語料庫。

3.2 智慧搭配詞搜尋引擎開發

為提升商務華語語料庫的實用性和便利性，進以作為教師或研究人員一項易於使用的商務華語輔助工具，並有鑑於搭配詞在詞彙教學上的重要性，本研究以商務華語語料庫為研究語料，特別針對搭配詞檢索開發了智慧搭配詞搜尋引擎。

本研究開發之智慧搭配詞搜尋引擎，以前文所介紹的網路商務華語語料庫為檢索語料來源。藉由 Python 程式語言，可計算出語料庫中查詢詞前項及後項詞彙出現的頻率。在電腦伺服器的快速大量運算下，得以即刻篩選出與查詢詞共同出現頻率高的詞彙，並呈現於檢索結果中。

為加快搭配詞檢索速度，並排除使用者不需要的檢索結果。本研究在開發此搜尋引擎時，特別設置查詢詞詞性及搭配詞類型供使用者選擇。前者可針對查詢詞的詞性進行設定，包括動詞、名詞、形容詞及介詞；後者則可挑選與查詢詞相搭配的词性類別，包括動詞和名詞搭配、形容詞和名詞搭配、具修飾成分的名詞和名詞搭配、介詞和名詞搭配。透過這些詞性選項，可方便使用者在一開始即鎖定搭配詞的檢索範圍，減少處理其他較不相關搭配詞的時間和精力。

智慧搭配詞搜尋引擎現已對外開放使用，網址為：<http://140.122.110.84:8000/ICE/Index.htm>。使用者只要輸入查詢詞後可選取查詢詞詞性、欲檢索之搭配詞類型及語料來源後，其檢索結果列於網頁下方。檢索結果亦依搭配詞共現的頻率排列，每組搭配詞後括號中的數字便是其在語料庫中共同出現的次數。點選該組搭配詞時，則會出現其在語料庫中的完整例句。下圖一為智慧搭配詞搜尋引擎檢索「商品」為作名詞之動名詞搭配結果。

² 中央研究院中文斷詞系統介紹及下載連結
<http://rocling.iis.sinica.edu.tw/CKIP/wordsegment.htm>

Intelligent Collocation Engine

商品 Submit 共現頻率 10 Dice值 Normal

語料來源： 中央社新聞 美國之音 財經篇 美國之音
 華爾街日報 商業一 商業二
 查詢詞詞性： 動詞 名詞 形容詞 介詞
 搭配詞類型： 動名 形名 名名1 名名2 介名

購買 商品(206)	提供 商品(153)	是 商品(130)	有 商品(87)	種 商品(61)
使用 商品(55)	收到 商品(49)	拍賣 商品(42)	訂購 商品(42)	含 商品(40)
促銷 商品(33)	兼銷 商品(33)	退回 商品(32)	包括 商品(31)	刊登 商品(30)
確保 商品(26)	出售 商品(26)	選擇 商品(24)	說明 商品(22)	從事 商品(22)
發現 商品(19)	享有 商品(18)	指 商品(18)	作為 商品(18)	庫存 商品(18)
進行 商品(18)	內無 商品(17)	保護 商品(16)	有關 商品(15)	涉入 商品(15)
需要 商品(15)	到 商品(14)	請 商品(14)	檢驗 商品(14)	確認 商品(14)
讓 商品(14)	保證 商品(13)	交付 商品(13)	計算 商品(12)	適用 商品(12)
得 商品(12)	創新 商品(12)	具 商品(12)	具有 商品(12)	使 商品(11)
負擔 商品(11)	提高 商品(11)	屬於 商品(11)	為 商品(11)	分析 商品(11)
利用 商品(11)	推銷 商品(10)	開展 商品(10)	願意 商品(10)	通過 商品(10)
製造 商品(10)	代理 商品(10)	完成 商品(10)	選購 商品(10)	包含 商品(10)

本網頁目前支援 Firefox & Chrome & Safari

圖1 智慧搭配詞搜尋引擎

特別值得注意的是，此智慧搭配詞搜尋引擎提供多項具專業性語料庫供使用者選用，包含：「中央社新聞」、「美國之音 財經篇」、「美國之音」、「華爾街日報」、「商業一」、「商業二」³等。使用者可依需要選擇欲使用之專業華語語料，輸入欲檢索的字詞後，便可得到該檢索詞在選用之語料下的搭配詞情況。例如：使用者想瞭解新聞性文章中「商品」的搭配詞使用情形，可選取「中央社新聞」或「華爾街日報」作為語料來源。而本研究主要探討網路商務華語之搭配詞使用情型，故此次以「商業二」作為研究語料。

4. 商務華語語料庫在智慧搭配詞搜尋引擎中的應用

本研究建置之智慧搭配詞搜尋引擎，可提供華語教師進行多方面專業性搭配詞檢索功能。而本次研究主要針對商務類語料進行搭配詞探究，以提供商務華語教學及教材編寫者編寫商務華語教材的參考。首先在語料內容方面，來源

³ 「商業一」及「商業二」語料皆是以 WebBootCat 語料蒐集軟體所建置之商業語料，然而兩者之差異為「商業一」的種子字詞取自商務類期刊，而「商業二」的種子字詞為《遠東商務漢語》。

為商業性質的網站文本，比起一般綜合性語料庫之檢索結果，更能凸顯出商務類用語的特色亦較具有針對性。為使搜尋結果精簡化、避免一般語料庫搜尋結果過於龐雜、數量龐大、詞性不一的缺點，本搜尋引擎能讓使用者先界定查詢詞詞性、進而界定想要的搭配詞詞性，由系統首先將不符合使用者需求的語料排除。

搜尋結果以共現頻率排序，共現次數愈高，表示關鍵詞與搭配詞的相連性越高。一般而言，共現次數高的搭配詞比起共現次數低的搭配詞，較適合作為教師及教材編寫者選用的對象。此外，在檢索結果下點擊所檢索關鍵詞(紅字)，便能查看該字詞在篇章中的使用情況，如下圖2所示。此精簡實用的搭配詞搜尋引擎可協助教師及教材編輯者找出更多的搭配詞組合及大量搭配的範例，同時亦可協助教師及學習者觀察出該關鍵字詞在語句中的呈現方式，提供更貼近母語者語感的使用情形。

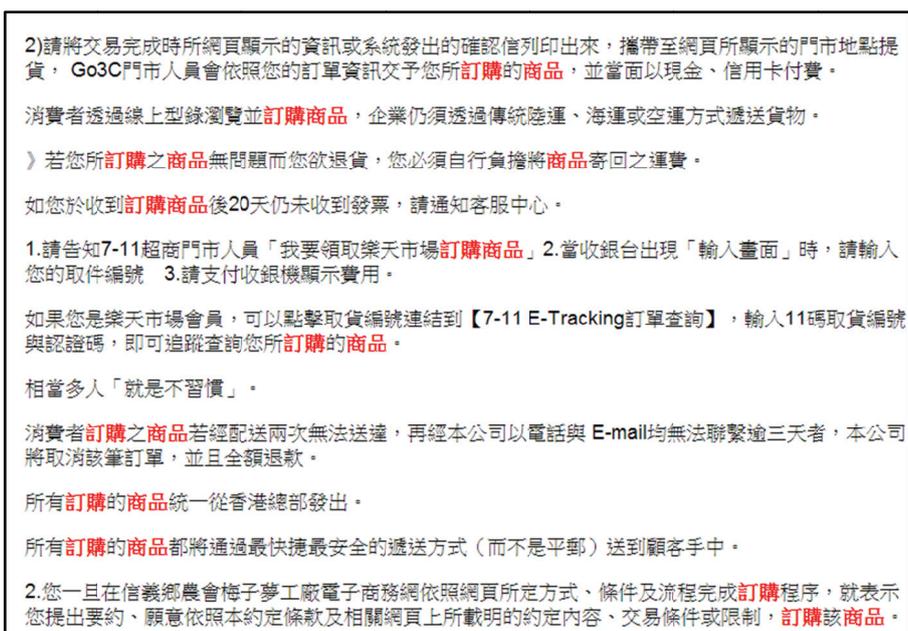


圖2 搭配詞在篇章中的使用情況

李泉(2004)認為編寫教材必須遵循十大基本原則，其中的定向原則及目標原則強調必須加強教材的針對性，確定教材的基本走向及性質，根據教學對象需求來編寫。商務華語教材是專業華語教材的一類，是以從事商務相關行業的學習者為對象編纂的，故比起通用華語教材，教材內容的選擇須具有較強的

針對性，須涵蓋商業場合所運用到的對話和詞彙。因此，相較於一般語料庫，本研究建置之商務華語語料庫以網站商務類文本為語料，得以更準確地反映出網路商務類詞彙和用語，對於編寫商務華語教材而言，可提供更具切合的商務情境。

為驗證此商務語料庫具有有別於一般語料庫的特性，本研究以較無商務色彩的一般性質詞彙「網站」作為關鍵字，分別在中央研究院現代漢語平衡語料庫3.0版（約5百萬字綜合性語料）及本研究所建置的商務華語語料庫進行檢索，比較兩者間的不同。

「中央研究院現代漢語平衡語料庫」簡稱「中研院平衡語料庫」，是第一個經過完整詞類標記的漢語平衡語料庫。為求語料的平衡性，語料的蒐集以六大主題為主軸，其收錄內容比例為：哲學占10%、科學占10%、社會占35%、藝術占5%、生活占20%、文學占20%，可說是現代漢語無窮多語句中一個代表性的樣本⁴。對於華語教學研究、漢語語言研究者而言，中研院平衡語料庫可說是目前最具代表性的綜合語料庫，當代學者無論在詞表研擬、詞彙分析或教材編纂等方面，皆以此為首要的參考依據。有鑑於中研院平衡語料庫之代表性和重要性，本研究以此作為對比分析的語料依據，以此驗證商務語料庫具有有別於一般語料庫的特性。

「網站」一詞在中研院平衡語料庫的搭配詞檢索結果共360筆。但由於筆數甚多，故僅擷取詞頻高於5的前39筆左方搭配詞羅列於下表1。

表1 「網站」一詞在中研院現代平衡語料庫的搭配詞檢索結果

搭配詞	詞頻	搭配詞	詞頻	搭配詞	詞頻
的(DE)	78	女性(N)	9	相關(V)	6
在(P)	49	透過(P)	9	公司(N)	6
個(N)	26	或(Conj)	8	及(Conj)	6
上(V)	13	從(P)	8	大(V)	6
入口(N)	13	學習(V)	8	國(N)	6
一(N)	12	電腦(N)	8	許多(N)	6

⁴ 參見中央研究院資訊所、語言所詞庫小組所編技術報告第 95-02/98-04 號「中央研究院漢語料庫的內容與說明」。

色情 (N)	11	美國(N)	8	目前(N)	5
該(N)	11	將(D)	7	進入(V)	5
有(V)	11	於(P)	7	同志(N)	5
是 (SHI)	10	不法(A)	6	與(Conj)	5
官方 (N)	10	等(Cab)	6	至(P)	5
這(N)	9	業者(N)	6	臺灣(N)	5
架設 (V)	9	可以(D)	6	全(N)	5

中研院平衡語料庫由於收錄的語料主題與面向眾多，檢索出的搭配詞並不僅限於商業領域相關的詞彙如「公司」，而是各類領域相關搭配詞皆有，例：「色情」、「架設」、「女性」、「透過」、「美國」、「學習」、「臺灣」等，反應出一般生活中的常用搭配。

本研究接著於智慧搭配詞搜索引擎中選用「商業二」之商務華語語料庫檢索「網站」一詞之「動詞」、「形容詞」及「名詞」搭配，將所得結果錯誤的搭配詞刪除⁵，將與「網站」一詞的左方搭配組合列於下表 2。括號內數字為搭配頻率。

表2 「網站」一詞在商務華語語料庫的搭配詞檢索結果

左方搭配詞 詞性	動詞	形容詞	名詞
搭配詞檢索 結果	使用 網站 (29) 通過 網站 (15) 瀏覽 網站 (13)	相關 網站 (17) 知名 網站 (12) 可靠 網站 (10)	公司 網站 (59) 商務 網站 (29) 官方 網站 (29) 甲方 網站 (24) 入口 網站

⁵ 由於中文斷詞系統的準確率約 90%，約 10%的詞彙詞性被錯誤標記，故搭配詞搜尋引擎有時會呈現錯誤的搭配詞。

			(21) 工廠 網站 (17) 物流 網站 (12)
--	--	--	--

從表 2 可看出，總體而言，「網站」左方多與搭配商業領域相關的詞彙，如：「知名」、「可靠」、「公司」、「商務」、「物流」等，較多與商業領域相關的詞彙。由此可看出，本研究建置的商務華語語料庫以商務類文本為語料，能反映出商務類詞彙和用語，對於商務華語教材編寫者及商務華語教學者來說，選擇商務華語語料庫作為輔助工具，將更具有針對性。

本研究所建置之商務華語搭配詞搜尋引擎，除了具有可檢索大量商務類搭配詞之特性，更可設定查詢詞本身及欲搜尋的搭配詞詞性，例如：動名搭配、形名搭配、名名搭配等，有助於使用者快速查找特定的搭配詞類別。然而使用此智慧搭配詞搜尋引擎需注意的是，因在對語料庫文本標註詞類時，受到詞性標註軟體功能上的限制，一個詞彙只能賦予一個詞性，故所得搭配結果會因為上述限制造成某些搭配錯誤情形，需再以人工方式加以篩選。

歷來許多關於學習者搭配詞偏誤的研究皆指出，學習者的搭配詞偏誤大多都集中在動詞與名詞的搭配使用上(Nesselhauf 2003、2005；Chen 2002；Shih 2000；Liu 2002)。Nesselhauf (2005) 曾抽選了 318 篇德籍進階級英語學習者的英語作文，從中觀察學習者使用動名搭配詞的情形，結果顯示，學習者共使用了 2082 組動名搭配詞，然其中就有 507 組是錯誤用法，偏誤率大約是四分之一；Liu(2002) 也觀察了大量臺灣高中生及大學生的英文作文，結果顯示：學習者使用的錯誤搭配用法共有 265 組，然其中 233 組都屬於動詞與名詞搭配的錯誤。可見相較其他詞類的搭配，動詞與名詞的搭配用法是學生最易犯錯的類型。因此在詞彙教學中，動詞與名詞的搭配應是教學的重點之一，教師在教學中或教材編輯者在編寫教材時，應適時補充詞彙的動、名詞搭配用法。對此，本研究利用智慧搭配詞搜尋引擎作為動詞、名詞搭配詞的檢索工具，可快速、便利地列舉出豐富的商務類動詞、名詞搭配，雖然因語料庫詞性標註功能的限制，所得搭配詞之詞性判別仍需由人工進一步檢查，但不影響其搭配詞較具商務性質之結果。本研究期望此智慧搭配詞搜索引擎可作為教師及研究人員在搭配詞教學時或教材編輯時的輔助工具。

中研院平衡語料庫中的搭配詞檢索功能，其檢索出之搭配詞除無法依據詞性分類，檢索結果呈現得較一般性搭配情況。對於教師或教材編輯者而言，藉中研院現代漢語平衡語料庫檢索搭配詞不僅難以區分不同詞性的搭配詞，學習者常出現學習困難的動名搭配詞數亦較少，在教學及教材編輯時可提供的資源於是就相對不足。為顯示兩語料庫之搭配詞檢索結果確實有數量上的差異，本研究選用了 19 個較具商務性質之名詞作為關鍵字，將共現頻率（詞頻）設定為 10 以上，分別在此兩個語料庫中檢索這 19 個關鍵詞的動詞搭配詞。本研究將錯誤的搭配詞及非關鍵搭配詞人工篩選並刪除後，所得檢索結果如表 3 所示。灰階標記該詞彙同時出現在這兩個語料庫所檢索出來的搭配組合中。

表 3 商務華語語料庫及中研院現代漢語平衡語料庫之動名詞搭配詞檢索結果

關鍵詞 (名詞)	商務華語語料庫檢索結果 (動詞)	中研院現代漢語語料庫檢索結果
商品	共 19 筆 購買、使用、收到、拍賣、訂購、促銷、兼銷、退回、刊登、出售、檢驗、確認、交付、適用、利用、推銷、製造、代理、選購	共 1 筆 仿冒
成本	共 24 筆 計算、控制、支付、造成、提高、預計、投入、支出、付出、節約、影響、收購、監督、花費、耗用、扣除、計入、反映、負擔、管理、算、測算、分析	共 3 筆 降低、節省、增加
資金	共 14 筆 提供、 募集 、取得、需要、 投入 、閒置、獲得、利用、運用、超募、提高、吸引、補充、籌集	共 3 筆 投入 、 募集 、需
市場	共 26 筆 擴大、 進入 、 成爲 、 開拓 、促進、拓展、退出、推動、 進軍 、 投入 、 開發 、形成、 打開 、 掌握 、 加入 、佔領、發展、擁有、 打入 、 看好 、維護、引入、主導、創造、引導	共 20 筆 進入 、 成爲 、 開放 、 投入 、 拓展 、 開拓 、

		進軍、打入、因應、掌握、打開、經營、爭取、影響、開發、退出、看好、加入、形成、壟斷
訂單	共 10 筆 確認、接受、取消、收到、接到、獲得、處理、發出、下、取得	共 1 筆 下
運費	共 8 筆 支付、含、加、負擔、預付、收取、包括、包含	共 0 筆
貨物	共 23 筆 交付、運送、裝載、外銷、提供、收到、承保、接受、檢驗、包含、購買、購進、出售、收取、提領、描述、提取、訂購、拒收、投保、計算、領取、處理	共 0 筆
廣告	共 10 筆 做、刊登、看、購買、從事、進行、投放、看到、選擇、利用	共 3 筆 刊登、拍、做
費用	共 20 筆 支付、收取、負擔、承擔、扣除、支出、繳付、需要、查、自負、繳、提供、計算、造成、處理、含、收、賠付、交付、預付	共 5 筆 收取、支付、負擔、收、需
品質	共 8 筆 提升、影響、提高、改善、確保、判斷、保證、注重	共 3 筆 提升(提昇)、提高、提昇、改善

業績	共 2 筆 提升、影響	共 0 筆
樣品	共 2 筆 寄、提供	共 0 筆
消費者	共 18 筆 保護、提供、規定、吸引、提醒、供、瞭解、告知、成爲、幫助、建議、服務、欺騙、帶給、利用、重視、要求、賣給	共 5 筆 吸引、提醒、建議、供、保護
合約	共 9 筆 簽訂、履行、訂立、取消、符合、違反、解除、簽署、取得	共 1 筆 簽訂
業務	共 18 筆 辦理、從事、代理、開展、進行、管理、提供、拓展、申請、委託、擴展、經營、處理、承保、負責、收購、說明、供應	共 5 筆 負責、經營、處理、辦理、從事
保險	共 13 筆 買、辦理、投保、提供、承保、強制、支付、享有、申請、安排、簽訂、領取、持有	共 0 筆
航班	共 2 筆 搭乘、預定	共 0 筆
交易	共 13 筆 進行、完成、從事、達成、辦理、接受、訂立、參與、模擬、處理、停止、開展、批准	共 2 筆 進行、從事
包裝	共 3 筆 提供、採用、描述	共 0 筆

由表 3 可看出，總體而言，兩語料庫檢出的動詞搭配詞在筆數上有明顯的差異，中研院現代漢語平衡語料庫共現頻率為 10 以上的搭配詞較少，搭配詞組合筆數多半都在 5 筆以下。19 個關鍵詞中，有 7 個在中研院平衡語料庫中查無其動詞搭配詞。相較之下，本研究所建置的商務華語語料庫搭配詞搜尋引擎則呈現出較為豐富多樣的動詞搭配詞檢索結果，可提供教師及教材編輯者豐富大量的搭配範例。

以「商品」一詞為例，在中研院現代漢語平衡語料庫中，共現頻率在 10 以上的搭配詞僅有「仿冒商品」。然而與「商品」搭配的词彙應涵蓋各個面向，藉由智慧搭配詞搜尋引擎可查找到與買賣相關的搭配詞包括：「購買」、「訂購」、「選購」、「促銷」、「出售」、「推銷」；與商品使用相關的搭配詞包括：「使用」、「適用」、「利用」；與商品流通相關的搭配詞則有：「收到」、「退回」、「交付」。又如「訂單」在中研院現代漢語平衡語料庫中，與之共現頻率在 10 以上的動詞僅有「下」一詞。而在商務語境中，「訂單」的一系列處理過程皆有經常搭配使用的動詞，包括從「接收訂單」、「獲得訂單」、「取得訂單」到「確認訂單」，再到「處理訂單」或是「取消訂單」，這些動詞搭配皆出現在智慧搭配詞搜尋引擎檢索結果中，建議教師在講授「訂單」一詞時可予以補充相關的動詞搭配。

本研究所採用的語料內容來源為網路商務文本，研究成果主要呈現網路商務搭配詞使用情形，因此所得搭配結果或多或少帶有網站用語特色。但本研究特別歸納整理此智慧搭配詞搜尋引擎四大特色供華語教師、學習者及研究人員參考：（1）以網路商務類文本為語料，較具有針對性。對於編寫商務華語教材而言，此網路商務語料庫能更準確地反映出網路商務類詞彙和用語，更加切合真實之網路商務情境。（2）精簡實用的搭配詞檢索功能。其搭配詞檢索功能，是針對商務華語教材之編寫者或商務華語教師找尋商務語料所設計，能讓使用者設定查詢詞本身及欲搜尋的搭配詞之詞性，以將不符合需求的語料排除，使搜尋結果精簡化。（3）能查看搭配詞在篇章中的使用情況，可以成為編寫課文及例句的語料來源。（4）豐富多樣的搜尋結果。商務性質語料的搭配詞搜尋結果較一般語料庫豐富，教師及教材編輯者能參考搭配詞的共現頻率資料，從豐富的檢索結果中選擇共現頻率高的搭配詞作為優先教學、編寫教材的參考。

從語料內容的層面來看，中研院現代漢語平衡語料庫有其優點，其語料來自哲學、文學、生活、社會、科學、藝術六大領域，面面俱到，豐富多樣。然也是因為取材廣泛，平衡語料庫裡某一特定領域的語料在種類及數量上則沒有專業語料庫多，對專業華語學習及教學來說是一大限制。近年來海內外學習者對專業華語的需求逐漸擴增，而現存專業華語語料庫寥寥可數，專業語料庫的發展有其必要性。雖然本研究語料來源僅來自於網路商務文本，與「一般商務文本」在使用習慣上可能具有些許差異，但引進國外「將網路作為語料庫」(Web as corpus) 的技術，以一個更有效率的途徑及方式建立商務華語語料庫，期望能為國內更多相關研究提供參考。若要得到更多元的商務華語搭配詞，實需大量的網路商務語料及一般商業文書語料進行對比探討，期望在未來研究中可更

進一步收集更多一般商務文書及商業教科書內容，進以建置更全面之商務華語語料庫進行提供更為詳盡之商務搭配使用情形，並期許未來能有更多不同專業領域的華語語料庫陸續落成。

5. 結語

近年來海內外學習者對專業華語的需求逐漸增加，專業華語相關的課程、教材與研究已成為華語教學界關注的領域。目前專業華語教學、教學理論研究、教材編寫及研究雖已取得了一些成果，但在專業華語語料庫方面，相較於國外各種專業英語語料庫的建設成果，目前專業漢語語料庫的建設仍顯得不足。而語料庫不僅是語言研究和自然語言處理研究的資訊來源，其提供的大量的實際語料，對語言教材的編纂及課堂教學也具有相當重要的價值，故各類專業華語語料庫的建置已勢在必行。

而 21 世紀以來，由於中外經貿交流擴大，商務華語教學研究及教材編寫在專業華語領域中呈現一枝獨秀的趨勢，因此相較於其他各種專業華語語料庫，商務華語語料庫的建設顯得迫切許多。而目前國際上的商務語料庫不多，商務華語語料庫更是寥寥可數。有鑒於此，本研究以「網路作為語料庫」的概念，取《遠東商業漢語》中之高頻華語商務常用詞彙作為種子字詞，利用 WebBootCat 語料蒐集軟體有效率地蒐集大量網路商務華語語料，完成網路商務華語語料庫的建置工作。並基於搭配詞對詞彙教學的重要性。本研究進一步開發智慧搭配詞搜尋引擎以便於檢索商務華語語料庫之語料，方便於商務華語教師或教材編輯人員將檢索出的搭配詞運用於華語教學及教材編寫上。

相較於一般綜合性語料庫之搭配詞檢索功能，此智慧搭配詞搜尋引擎不但在語料上較有針對性，在商用詞彙的搜尋結果上也較豐富多樣，且有便利使用者的設計。此精簡實用的搭配詞搜尋引擎能協助教師及教材編輯者找出更多不同詞性的搭配詞，提供大量搭配的範例，以彌補搭配詞學習上常見的語感侷限，對華語教師、學習者及教材編輯者而言是極佳的輔助工具。

本研究仍有許多侷限及受限之處需加以改進，如：因在標註語料庫文本之詞類時，受詞性標註軟體功能上的限制，一個詞彙只會被賦予一個詞性，故所得搭配結果會因此限制出現搭配錯誤情形，需再以人工方式加以篩選。另一方面，由於取材為網路商業文本，研究結果帶有「網路用語」特色，而網路商務文本與一般商務文本在內容及使用方法上仍有所差異。期許在未來的研究中，可增加更多一般商務文書及華語商務教科書內容於語料中，進以建置更全面且

更多元之商務華語語料庫，以提升研究成果之廣度。

而本研究期許此華語商務語料庫的建置及智慧搭配詞檢索方法為拋磚引玉之姿，提供國內外專業科目之華語語料庫建置及研究工作作為參考。期望可促成更完備之網路商務中文語料庫之建置並其他專業領域華語語料庫之落成。另一方面，本研究應用網路商務華語語料庫所開發之智慧搭配詞搜尋引擎，具有專業之商務搭配詞資源，提供使用者便捷的檢索功能和介面。此搜尋引擎可應用教學現場、教材編纂和搭配詞研究等方向，對於商務華語教學、搭配詞教學、教材編輯及辭典編纂等都將有所助益。也期望相關研究和應用得以以此搜尋引擎為基礎，拓展商務華語教學的深度與廣度。

引用文獻

- Alexander, R. J. 1984. Fixed Expressions in English: Reference Books and the Teacher. *English Language Teaching Journal* 38. 2: 127-134.
- Bahns, J. & Eldaw, M.. 1993. Should we teach EFL students collocations? *System* 21:101-114.
- Baroni, M. & Bernardini, S. 2004. BootCaT: Bootstrapping Corpora and Terms from the Web. In Proceedings of LREC.
- Baroni, M., Kilgarriff, A., Pomikálek, J. & Rychlý, P. 2006. WebBootCaT: instant domain-specific corpora to support human translators. In Proceedings of EAMT, 247-252.
- Chen, P. C. 2002. A corpus-based study of the collocational errors in the writings of the EFL learners in Taiwan. Unpublished master's thesis, National Taiwan Normal University, Taiwan, R.O.C.
- Farghal, M. & Obiedat, H. 1995. Collocations: A neglected variable in EFL. *IRAL*, 33/4: 315-333.
- Firth, J. 1957. *Papers in Linguistics, 1934-1951*. Oxford: Oxford University Press.
- Fujii, Atsushi & Ishikawa, Tetsuya. 2000. Utilizing the World Wide Web as an encyclopedia: Extracting term descriptions from semi-structured text. In Proceedings of the 38th Meeting of the ACL, 488-495.
- Gitsaki, C. 1996. The development of ESL collocational knowledge (Unpublished

- doctoral dissertation) . Center for Language Teaching and Research, University of Queensland, Brisbane, Australia.
- Halliday, M., MacIntosh, A. & Stevens, P. 1964. *The linguistic sciences and language teaching*. London: Longman.
- Jones, Rosie & Rayid Ghani. 2000. Automatically building a corpus for a minority language from the Web. In Proceedings of the Student Workshop of the 38th Annual Meeting of the Association for Computational Linguistics, 29-36.
- Kilgarriff, A. & Grefenstette, G. 2003. Introduction to the Special Issue on the Web as Corpus. *Computational Linguistics*, 3/29:333-347.
- Leech, G. 1997. Teaching and Language Corpora: A Convergence. In A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles (Eds.) , *Teaching and Language Corpora*:1-23.
- Lewis, M. 2000. *Teaching collocation: Further development in lexical approach*. London:Language Teaching.
- Liu, L. E. 2002. A corpus-based lexical semantic investigation of verb-noun miscollocations in Taiwan learners' English. Unpublished master's thesis. Tamkang University, Taiwan.
- Mindt, D. 2001. English corpus linguistics and the foreign language teaching syllabus . In: Jenny Thomas & Michael H.Short (eds.). *Using Corpora for language Research*. Beijing: Foreign Language Teaching and Research Press.
- Nesselhauf, N. 2003. The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*24:223-242.
- Nesselhauf, N. 2005. *Collocations in a Learner Corpus*. Amsterdam:John Benjamins.
- Resnik, P. 1999. Mining the Web for bilingual text. In Proceedings of the 37th Meeting of ACL, 527-534.
- R.O.C. Nation, I.S.P. 2001. Learning vocabulary in another language. Cambridge: Cambridge Press.
- Sinclair, J. M. & Renouf, A. (Eds.) . 1988. A Lexical Syllabus for Language Learning. In Carter & McCarthy (Eds.) . *Vocabulary and Language Teaching*:40-158.
- Shih, H. H. 2000. The influence of synonymy over collocation learning. Selected Papers from the Ninth International Symposium on English Teaching, 516-520.
- Woolard, G. 2000. Collocation: Encouraging learner independence. Teaching

- collocation: Further developments in the lexical approach :28-46.
- 王硯農。1992。〈談談中醫漢語系列教材〉。《世界漢語教學》。第4期。301-303。
[Wang, Yan-nong. 1992. Talk about Chinese medicine textbook series. *Chinese Teaching in the World*4:301-303.]
- 王建勤。1997。《漢語作為第二語言的習得研究》。北京：北京語言大學出版社。
[Wang, Jian-qin. 1997. *Studies in Second Language Acquisition of Chinese*. Beijing: Beijing Language and Culture University Press.]
- 石金才。1990。〈關於編寫外交人員初級漢語教材的設想〉。《世界漢語教學》。第3期。188-190。[Shi, Jin-cai. 1990. Ideas about Elementary Chinese diplomats writing textbooks. *Chinese Teaching in the World*3:188-190.]
- 全香蘭。2008。〈韓語漢字詞對學生習得漢語詞語的影響〉。《基於中介語語料庫的漢語詞彙專題研究》。132-139。[Quan, Xiang-lan. 2008. The impact of Korean Hanja on the acquisition of of Chinese words. *Interlanguage corpus-based Chinese vocabulary thematic studies*:132-139.]
- 呂必松。1990。《對外漢語教學發展概要》。北京：北京語言大學出版社。[Lü, Bi-song. 1990. *A Summary of the Development of Teaching Chinese as a Foreign Language*. Beijing: Beijing Language and Culture University Press.]
- 李泉。2004。〈第二語言教材編寫的通用原則〉。《第三屆全國語言文字應用學術研討會論文集》。447-460。[Li, Quan. 2004. The General Principle of Second Language Textbook Compilation. *The Third National Symposium on Applied Linguistics*:447-460.]
- 李泉。2011。〈論專門用途漢語教學〉。《語言文字應用》。第3期。110-117。[Li, Quan. 2011. Teaching Chinese for Specific Purposes and Its Textbook Compilation. *Applied Linguistics* 3:110-117.]
- 李育娟。2011。〈商務華語學習者需求分析〉。《華語文教學研究》。第8卷第3期。23-46。[Li, Yu-juan. 2011. Learner's Needs Analysis in Business Chinese. *Journal of Chinese Language Teaching*8.3:23-46]
- 李曉琪。2010。〈系列商務漢語教材編寫探索〉。《對外漢語教學與研究》。第1期。40-49。[Li, Xiao-qi. 2010. The Compilation of a Business Chinese Textbook Series. *Teaching Chinese as a Foreign Language* 1:40-49.]
- 杜厚文。1981。〈漢語科技文體的語言特點〉。《語言教學與研究》。第2期。87-101。
[Du, Hou-Wen. 1981. The Language Features of Chinese technical literary style.

*Language Teaching and Linguistic Studies*2:87-101]

- 杜昭玫。2011。〈商務漢語教材之內容分析與建議〉。《華語文教學研究》。第 8 卷第 2 期。81-98。[Tu, Chao-mei, 2011. Teaching Materials for Business Chinese: An analysis and Suggestions. *Journal of Chinese Language Teaching*8.2:81-98.]
- 周玉林。2010。〈試論“商務英語語料庫”的建設及其應用〉。《長沙鐵道學院學報（社會科學版）》。第 11 卷第 1 期。148-151。[Zhou, Yu-lin. 2010. On Building and Applying the "Business English corpus". *Journal of Changsha Railway University (Social Science)* 11.1:148-151.]
- 金惠淑。2005。《觀光華語導遊教材設計—針對韓國的專門大學觀光中國語系之用途》。臺北：國立臺灣師範大學碩士論文。[Jin, Hui-Shu. 2005. Chinese Language Training Materials for Tour Guides—A Course for Korean College Students major in Tour Chinese. Taipei: National Taiwan Normal University Master thesis.]
- 胡明揚。2006。〈詞彙教學理論〉。《對外漢語詞彙及詞彙教學研究》。203-225。[Hu, Ming-yang. 2006. Vocabulary teaching theory. *Research in Chinese vocabulary and vocabulary teaching*:203-225.]
- 馬玉汭。2006。〈詞彙教學方法〉。《對外漢語詞彙及詞彙教學研究》。263-291。[Ma, Yu-bian. 2006. Vocabulary teaching method. *Research in Chinese vocabulary and vocabulary teaching*: 263-291.]
- 高燕。2008。《對外漢語詞彙教學》。上海：華東師範大學出版社。[Gao, Yan. 2008. *Teaching Chinese as a foreign Language*. Shanghai: East China Normal University Press.]
- 陸國強。1986。《現代英語詞彙學》。上海：上海外語教育出版社。[Lu, Guo-qiang. 1986. *Modern English Lexicology*. Shanghai :Shanghai Foreign Language Education Press.]
- 陳淑惠。2008。《法律華語教材設計與研究》。臺北：國立臺灣師範大學碩士論文。[Chen, Shu-Hui. 2008. The Design and Research of Legal Chinese Teaching Materials. Taipei: National Taiwan Normal University Master thesis.]
- 陳麗宇、李欣欣。2012。〈當代初級商務華語教材之文化研究〉。《華語文教學研究》。第 9 卷第 4 期。41-73。[Chen, Li-yu & Lee, Hsin-Hsin. 2012. Culture Study of Fundamental Business Chinese Teaching Material. *Chinese Language*

*Teaching*9.4:41-73.]

- 張黎。2006。〈商務漢語教學需求分析〉。《語言教學與研究》。第3期。55-60。
[Zhang, Li. 2006. Need Analysis in Teaching Business Chinese. *Language Teaching and Linguistic Studies*.3:55-60.]
- 張博等。2008。《基於中介語語料庫的漢語詞彙專題研究》。北京：北京大學出版社。 [Zhang, bo. (eds.) . 2008. *Interlanguage corpus-based Chinese vocabulary thematic studies*. Beijing: Beijing Language and Culture University Press.]
- 張惠雯。2002。《中醫華語教材編寫之規畫設計》。臺北：國立臺灣師範大學碩士論文。 [Zhang, Hui-wen. 2002. The design of second language textbook for Chinese Medicine. Taipei: National Taiwan Normal University Master thesis.]
- 程樂樂、朱德君。2008。〈醫學漢語教材編寫芻議〉。《海外華文教育》。第2期。40-46。 [Cheng, Le-le&Zhu, De-jun. 2008. Discussion on Chinese medicine teaching materials. *Overseas Chinese Education*. 2:40-46.]
- 彭增安。2007。《跨文化的語言傳通—漢語二語習得與教學》。上海：學林出版社。 [Peng, Zeng-an. 2007. *Cross-cultural Language Communication— Chinese Second Language Acquisition and Teaching*. Shanghai : Academia Press.]
- 路志英。2006。〈商貿類漢語教材編寫和研究的基本情況述評〉。《雲南師範大學學報(對外漢語教學與研究版)》。第4卷第5期。28-33。 [Lu, Zhi-ying. 2006. A Summary Review of the Compilation and Study of Business Chinese Textbooks. *Journal of Yunan Normal University (Teaching and Research on Chinese as A Foreign Language Edition)* 4.5:28-33.]
- 董政、鄭艷群。2008。〈歐美學生漢語量詞的使用情況〉。《基於中介語語料庫的漢語詞彙專題研究》。89-98。 [Dong, Zhen & Zheng, Yan-qun. 2008. The Chinese quantifiers usage of students in Europe and America. *Interlanguage corpus-based Chinese vocabulary thematic studies*: 89-98.]
- 劉亞菲、鄭艷群。2008。〈韓國學生漢語量詞的使用情況〉。《基於中介語語料庫的漢語詞彙專題研究》。79-88。 [Liu, Ya-fei. & Zheng, Yan-qun. 2008. The Chinese quantifiers usage of Korean students. *Interlanguage corpus-based Chinese vocabulary thematic studies*: 79-88.]
- 蕭頻、張妍。2008。〈印尼學生漢語單音節動詞語義偏誤〉。《基於中介語語料庫的漢語詞彙專題研究》。62-72。 [Xiao, ping & Zhang, Yan. 2008. The

華語文教學研究

monosyllabic verb semantic errors of Indonesian students. *Interlanguage corpus-based Chinese vocabulary thematic studies*: 62-72.]

[審查：2014.08.14 修改：2015.03.12 接受：2015.05.06]

王敬淳

Ching-Chun Wang

106 臺北市和平東路一段 162 號

臺灣師範大學華語文教學研究所

Graduate Institute of Teaching Chinese as a Second Language, National Taiwan Normal University

No. 162, Sec. 1, He-ping E. Rd., Taipei, Taiwan R.O.C. 10610

matrillwan@yahoo.com.tw

陳浩然

Howard Hao-Jan Chen

106 臺北市和平東路一段 162 號

國立臺灣師範大學國語中心暨英語學系

Department of English, National Taiwan Normal University

No. 162, Sec. 1, He-ping E. Rd., Taipei, Taiwan R.O.C. 10610

hjchen@ntnu.edu.tw

潘依婷

I-ting PAN

106 臺北市和平東路一段 162 號

國立臺灣師範大學國語中心暨英語學系

Department of English, National Taiwan Normal University

No. 162, Sec. 1, He-ping E. Rd., Taipei, Taiwan R.O.C. 10610

Leopardpan@gmail.com

The Application of Web Business Chinese Corpus: A Case Study on Collocation

Ching-Chun Wang **Howard Hao-Jan Chen** **I-ting Pan**
Graduate Institute of Teaching **Mandarin Training Center &**
Chinese as a Second Language **Department of English**
National Taiwan Normal University

Abstract

The increase of Sino-foreign trades nowadays has aroused the urgent need for business Chinese teaching resources. To meet this need, the current study applied the 'Web as Corpus' method to build a Web Business-Chinese Corpus (WBCC) by collecting a large number of online business Chinese texts. To learners of business Chinese, the correct use of Chinese collocations plays a crucial role in the success of Chinese learning. Hence, this study uploaded the WBCC onto the Intelligent Collocation Engine (ICE), which allows users to search for business Chinese collocations via the ICE tool. To show that the WBCC and the ICE indeed provide more business-related search outcomes than other Chinese corpora, this study compared the collocation information of 19 business-related nouns retrieved from the WBCC and the ICE with that from the Academia Sinica Balanced Corpus of Modern Chinese, and the results showed that information generated from WBCC and ICE was more related to business Chinese. For example, for the noun 'shangpin' (product), 19 verb collocates can be found in WBCC, whereas there is only one verb collocate found in the Academia Sinica Balanced Corpus of Modern Chinese. Findings of this study thus suggest that the WBCC and the ICE can serve as good business Chinese teaching resources.

Keywords : Business Chinese, Web as Corpus, Collocation, Corpus Application, Chinese for Specific Purposes