

## 以學習者平行語料庫為本之西班牙語連接詞研究<sup>1</sup>

### Parallel Corpus-based Study of Conjunctions

盧慧娟\*、呂羅雪<sup>+</sup>

Hui-Chuan Lu and Lo Hsueh Lu

#### 摘要

本論文旨在結合語料庫技術建構「西語學習者英西平行語料庫」，並以之為本、輔以應用工具，進行學習者連接詞使用之分析研究。

「西語學習者英西平行語料庫」現階段所收集的語料來源是靜宜大學西文系和成功大學外文系西語組兩不同族群的學習者在課堂內完成的西文作文及與之對應的英文翻譯。上述語料皆經母語者修正與詞類註記之程序，以達有效率系統化分析之目的。以過去相關研究為基礎，我們使用語料庫應用工具對比分析連接詞的使用情形，以探討學習者在第三語連接詞不同語意類型的多用、誤用和少用的情形，及其可能的影響因素。研究結果顯示：(1)學習者在西文和英文整體連接詞的使用比率低於在西語及英語母語者修正文中的連接詞使用率，推論連接詞的少用可能和跟學習者的母語有關。(2)錯誤類型少用的情形集中於「聯繫、補充」類型的連接詞；多用的情形則多屬於「反義、因果」類型的連接詞。(3)在可能的影響變數的檢定中，連接詞的使用錯誤與西語程度間，未呈現顯著的關係性，和第二語（英語）之相關性也未達顯著差異的水準。最後，我們根據分析結果，利用現有之工具，提出連接詞的教學範例以供參考。

**關鍵字：**語料庫為本、學習者語料庫、平行語料庫、連接詞、第二語、第三語

---

<sup>1</sup>本論文是由以「以平行學習者語料庫為本之西班牙語連接詞研究」為題口頭發表於「當代歐洲、文化暨語言發展新趨勢國際學術研討會」（政治大學於 99 年 5 月 15 日舉辦）並被編印於該研討會會前論文集(93-110 頁)的會議論文所修改而成。

\*成功大學外文系，台灣台南市大學路 1 號，Tel: (06)275-7575\*52227 Fax: (06)238-7730  
Department of Foreign Languages & Literature, National Cheng Kung University, Tainan, Taiwan  
E-mail: huichuanlu1@gmail.com

<sup>+</sup>靜宜大學西文系，台灣台中縣沙鹿鎮中棲路 200 號，Tel: (04)2632-8001\*12303 Fax: (04)2462-1281  
Department of Spanish Language and Literature, Providence University, Taichung, Taiwan  
E-mail: slo@pu.edu.tw

### Abstract

This paper aims to introduce the process used to create a “Parallel Learners’ Corpus of Spanish and English” (CPATEI) by integrating the theories and techniques of Corpus Linguistics, and to investigate the conjunctions used by the Spanish learners in Taiwan.

The collected data in CPATEI are contributed by learners of Spanish from two universities in Taiwan, including Spanish compositions and their corresponding English translations written by two different levels of Spanish learners. Besides, we have collected texts revised by native speakers and POS-tagged.

Based on the collected data, we will examine the tendency of learners’ overuse, underuse and misuse of the conjunctions. At the same time, we will discuss the possible factors leading to incorrect usage of conjunctions. The results demonstrate that the frequency of conjunctions used by the learners in their Spanish compositions and English translations are lower than that of the Spanish and English native speakers in their corrections. Among the types of errors of conjunctions, underuse of the copulative and additive conjunctions and overuse of the adversative and cause-consecutive conjunctions are the two types of errors that have been noted in the writing of the Taiwanese learners. Furthermore, there are no relationships found between the wrong use of the Spanish conjunctions and their language level as well as the second language as English and the third language as Spanish. Finally, we hope the result of this research will bring benefits to the curriculum designs for the teaching of Spanish conjunctions.

**KeyWords:** Corpus-based, Learners’ Corpus, Parallel Corpus, Conjunctions, Second Language, Third Language

## 1. 語料庫建構與連接詞分析之目的

在過去幾十年，有關學習者語料庫的研究已引起越來越多的關注，但不可諱言的，其重心仍多集中於英語或單語的建構與開發；而在平行學習者語料庫的建置方面，其中大多數仍聚焦於第二語言(L2)和母語(L1)的平行影響上；以西語學習者為對象，且擴及第二語對第三語影響因素考量的平行語料庫則未見。有鑑於研究的特殊性與實際需求之缺乏，本研究的目標之一即建構「平行雙語學習者語料庫」(CPATEI)。

本論文的語料庫應用部份，則是以連接詞為研究議題。連接詞的主要功能在於連接句中的詞與詞、片語與片語、子句與子句或句與句，扮演著聯繫語意、使表達的語氣通順，以幫助讀者理解文本的關鍵角色。語言分析主題的選定主要來自西班牙語教學的實際觀察，我們發現：(1)台灣西語學習者有寫出具連貫性文章的困難。學習者在書寫句子時，特別在連接詞使用方面，常有少用或誤用的情形，造成語意連貫上的不順暢。(2)他

們所書寫的文本內容過於口語化，對正式書寫文體風格的認知較缺乏。(3)目前台灣所使用西班牙語教科書多以溝通式教學法為主導，內容偏重語言表達之功能，對連接詞的特性、分類及其在不同文體中的使用功能較少涉略，可能因此造成學習者對連接詞的掌握較不足。

本論文旨在結合語料庫技術建構「西語學習者英西平行語料庫」(Corpus Paralelo de Aprendices Taiwanese de Español e Inglés, CPATEI)，並以之為本、輔以應用工具，進行學習者連接詞使用之分析研究。以下內容主要涵蓋語料庫和連接詞兩大部分。我們將由語料庫建構出發，簡述相關工具的應用，呈現現階段建構成果及語料後續處理的過程。接著，分析連接詞的使用與錯誤傾向，並進而檢視可能的影響因素，最後參考母語者使用情形，提出教學範例以供參考。

## 2. 文獻回顧

### 2.1 學習者語料庫建構

#### 2.1.1 單語學習者語料庫

在國外，僅有「以西語為第二語之書寫語料庫」(CEDEL2: Corpus Escrito del Español L2)和「西班牙語學習者口語語料庫」(SPLLOC: Spanish Learner Language Oral Corpus)，且兩者皆以英語為母語之西語學習者為語料收集的對象。「以西語為第二語之書寫語料庫」(CEDEL2)由西班牙馬德里自治大學(Universidad Autónoma de Madrid)暨格拉納達大學(Universidad de Granada)所建構。語料量約 60 萬字，語料來源主要涵蓋超過 1500 位以英語為母語之西語學習者和約 500 位西語母語者。目前未開放檢索使用，因此可利用性有限。另一西語學習者語料庫「西班牙語學習者口語語料庫」(SPLLOC)由英國南安普敦大學(Southampton University)、紐卡索大學(Newcastle University)及約克大學(York University)三所大學暨經濟社會研究院(Economic and Social Research Council)合作所建構。目前共有 60 位以英語為母語之西語學習者參與的錄音檔案約 240 筆。此外，也包含西語母語者的語料錄音檔案約 50 筆。其建構計畫包含兩階段：SPLLOC 1 和 SPLLOC 2。目前 SPLLOC 1 已開放提供研究使用，可選擇限定或不限定條件進行關鍵字的檢索。

在國內，針對「語料庫」一詞的定義，如將之詮釋為「語料彙整」的廣義解釋，則在台灣西語教學界的研究不乏以之為本的相關論著，如白保羅、林姿如等人皆以各自所彙整的書面語語料進行學習語言分析；在口語語料庫方面，有由 Javier Pérez & Miguel Rubio 所建構的台灣西班牙文口語語料庫(Corpus Oral del Español en Taiwán, COET)，建構於 2004 年，內容包含了 82 個以文藻外語學院西語學習者為對象所收集到的會話錄音檔案。上述幾個語料庫的語言資料大部分皆已電子化或經過人工註記，比較可惜的是沒有進一步整合成可互換、再利用或結合科技使之成為可公開檢索的語料共享資源。此外，政治大學建置了以顯示該校外語學院多語種(23 種語種)特色的「外語學習者語料庫」，目前未包含西班牙語。最後，建構中的「成功大學多國語語料庫」，目前包括台灣日語

學習者語料庫和台灣西語學習者語料庫。其中,「台灣西語學習者語料庫」(CATE: Corpus de Aprendices Taiwaneses de Español), 可公開檢索的網址為 <http://corpora.fild.ncku.edu.tw/>。目前計含 1,913 篇學習者書面文本,總字數量約 34 萬字,文本本身標註有錯誤修正、詞類與字根註記,檢索時可設定條件、進行字(串)、該字詞類、字根及下一字詞類的檢索。

### 2.1.2 平行學習者語料庫

在平行的學習者語料庫方面,第一、二語的平行學習者語料庫極為有限,如日-英語(Mark, 2001)、中-日(Shimizu, Du & Dantsuji, 2004)等,而我們所著手建構的「台灣西語學習者英西平行語料庫」(CPATEI)則具有西漢語言組合以及第二和第三學習語平行的特色。

## 2.2 連接詞研究

從寫作表達的角度出發,如同閱讀理論中所強調的閱讀過程,一般而言,無論是說明體,還是敘事體,文章的基礎都是從迷你結構開始,也就是說,從字、句開始,逐步發展到子結構,再形成段落,最後朝向頂尖大結構(Lu & Lu, 2005),而具有連貫字句功能的連接詞在整體架構則扮演著特殊的角色。

### 2.2.1 連接詞分類

根據孫義禎(1993)、陳樹升(1993)、劉啓分(1993)、Montolío (2001)與 Gómez (1997)等對連接詞的屬性分類,整合出以下主要類別及所屬連接詞範例。首先,根據語法屬性,分(1)對等並列、(2)附屬和(3)相關三大類。(1)對等並列連接詞涵蓋以下類型:(a)聯繫連接詞“y/e, ni”。(b)選擇連接詞“o/u”。(c)對稱連接詞“ahora...ahora, bien...bien, ya...ya, ora...ora”。(d)補充連接詞“además”。(e)反義連接詞“pero, mas, sin embargo”。(f)解釋連接詞“es decir, o sea”。(2)附屬連接詞包括(a)一般連接詞“que”。(b)原因連接詞“porque, puesto que”。(c)結果連接詞“por eso, por consiguiente, por (lo) tanto”。(d)目的連接詞“para que, a fin de que”。(e)條件連接詞“si, con tal de que”。(f)讓步連接詞“aunque”。(g)比較連接詞“como”。(3)相關連接詞包括“apenas...cuando, ni...ni..., o...o...”。針對現有語料的文本類型,本文僅就學習者文本中所使用的連接詞屬性進行分析與討論。

### 2.2.2 以母語者語料為研究對象

有關西班牙語連接詞方面, Martínez (1997), Martín & Portolés (1999), Montolío (1998, 2000, 2001), Sánchez (1993), Lu & Lu (2006)等均探討連接詞在西文文章中的功能。Lu & Lu (2006)根據「西班牙語語料庫」(Corpus del Español) 的查詢結果指出自然語中連接詞出現頻率的差異,如:“pero” (43330)高於“mientras que” (1502)<sup>2</sup>等。另外, Lu *et al.* (2009)

<sup>2</sup> 為 2006 年時的查詢結果,因該語料庫語料之擴充,現查詢所得筆數分別為 pero (131,191), mientras que (3,993), 差距亦隨之擴大。

根據所建構的「西、英、中平行語料庫」(CPEIC)分析連接詞的使用情形。由於平行語料資源的缺乏，這個語料庫的多語平行語料主要來自聖經<sup>3</sup>。分析結果顯示在平行的三語語料中，西文連接詞使用的頻率最高，英文次之，中文最少；而根據連接詞的屬性分類，在各個語言中，並列連接詞的數量皆是最多的。

## 2.2.3 以學習語為研究對象

### 2.2.3.1 以英語為學習語

以英語為學習語的連接詞研究在國外如：Granger y Tyson (1996)透過對比分析學習者母語及所學之外語寫作中連接詞的使用情況，以瞭解他們中介語的發展。所分析之連接詞量包括以法語為母語(89,918 字)和以英語為母語者(77,723 字)的英文寫作分別為 916 和 976 個連接詞。研究結果發現：在學習者的英語外語作文中，出現有連接詞使用過多的現象，如：佐證連接詞“indeed, of course, in fact”、舉例連接詞“e.g., for instance, namely”，以及補充連接詞“moreover”；有些連接詞卻使用不足，如：反義連接詞“however, though, yet”和結果連接詞“therefore, thus, then”。Field & Yip (1992)指出補充連接詞多用的現象也出現在香港學生的第二語英語寫作中。上述連接詞使用過多的情形，被推論可能和受母語影響有關。以中文為母語的學習者連接詞使用過多的現象，如要被推論為跟母語有關，則需根據漢英對比或平行對應語料分析的結果，同樣的語意，英文會表達出連接詞的次數應該要少於漢語才算合理。

在台灣，Wu (1996)研究連接詞與英語能力之間的關係，發現連接詞的使用與語言程度的高低有密切關係。他在台灣大學生對英語連接詞的理解與使用之研究中，指出英語能力愈強，使用連接詞的能力也愈強。再者，周玉玲(2002)探討短篇小說運用於英語連接詞教學之情形，研究以事實性文章作為測試，結果發現學生對時間性連接詞掌握較低，其次是補充和反義，對因果連接詞掌握得最好。余昇易(2006)和曾熾竹(2004)探討連接詞教學與大學生寫作表現上的成效及影響，結果顯示連接詞的密集教學能有效改善語言程度能力較低學生之整體寫作表現。

---

<sup>3</sup> 雖然使用聖經作為平行語料分析的對象一直有相關的爭議性存在，然而不可否認的，它仍是語料庫語言學和自然語言處理平行對應和剖析應用中最常被利用的語言資源，原因不外乎在於它有豐富的多國語言文本、可合法取得的路徑與便利使用的電子化格式。我們使用聖經作為平行語料資料，主要是西英漢三語平行電子檔案格式的語料資源不易取得。以其語料庫所建構的「西英漢平行語料庫」(CPEIC)彙整了 BibleGateway.com 網路資源中的三語平行對應文本，目的是為了便利分析被平行翻譯的三個語言資料。此外，其平行文本的版本採用的是西英文的新國際版(New International Version for English, la Nueva Versión Internacional)和中文的聯合版(Union Version for Chinese)，文本的內容則是聖經的創世紀；其西、英、中語料量分別是 37,407、38,895 和 52,546 字。

### 2.2.3.2 以西班牙語為學習語

以西班牙語為學習語的连接詞研究在國外如：Romero (2005)指出，因為有连接詞連結句子和段落，才能使文章的語意連貫，這對學生的閱讀理解是有助益的。他把 Montolio (2001)對连接詞之分類應用於研究中，分析了 32 篇大學生的作文（含一、四年級兩組），研究發現大學生對反義连接詞的掌握較好。其中，一年級學生在使用「介紹觀點」之连接詞的表現比四年級學生好。至於補充连接詞方面，四年級學生使用「開啓新資訊」之连接詞優於一年級的學生。此外，Poblete (1999)指出學習者所寫的敘事性文章中，具補充屬性之连接詞的使用頻率較高。

在台灣方面，Blanco (2007)對比分析以西班牙語為母語的記者所寫之連續事件為主題的新聞文章 300 篇和 30 位台灣西班牙語學習者的作文，發現同一文章類型中，不同研究對象之相同连接詞使用的頻率不同（如：母語者使用時間连接詞“antes”和“antes de”的頻率分別是 84.94%和 12.04%，而學生在寫作時使用的百分比分別是 64.42%和 34.58%）。研究結果也顯示：學習者在反義连接詞類型有多用的情形，而特定文章類型與特定连接詞之使用間有相關性。而根據我們在「西班牙語語料庫」(Corpus del Español) 的查詢結果顯示“antes”和“antes de”的出現頻率分別是 240417 次(96%)和 10027 次(4%)，如此的結果也說明连接詞在不同文本類型間的使用差異。

## 2.3 第三語相關理論與研究

Leung (2005, 2006)指出第三語言學習者語言知識的轉換來源有兩種：第一語言和第二語言的中介語文法。此外，De Angelis (2007)延續(2005)的研究指出有關跨語言影響或語言知識轉移之因素更可細分成以下幾個主題，包括語言的距離(兩種語言間的同異程度)、學習者目的語和原生語(或第一語)能力、最後使用該語言的時間距離（即：多久以前使用過該語言）、在非母語環境下居住和處於非母語環境的時間長度、習得語言順序以及語言正式度(語言使用於正式或非正式的情境)等。在影響第三語或多國語言習得的因素中，Ringbom (2007)也討論到語言距離和語言能力程度兩個相關的因素。

綜合上述所言，我們瞭解到國內外相關語料庫和连接詞方面的開發與研究情形，由此我們發現在有關以西班牙語為學習語的连接詞研究領域裡，尚缺乏語料庫為本、以西語為第三語或平行語料，探討學習者连接詞使用情形方面的研究。因此，本研究將使用語料庫應用工具，對比分析我們所收集之語料中连接詞使用的情形，並探討學習者多用、誤用和少用的第三語（西班牙語）连接詞的類型，及其可能的影響因素。

## 3. 學習者平行語料之建構與利用

「西語學習者英西平行語料庫」的前期建置工作是由國科會人文學研究中心經費補助之整合型前置規劃案。語料庫前置與現階建構成果有以下幾點差異：(1)過去對於參與的西語程度沒有一致的比較基準，只能以學習時數區分，現階段的參與者則皆統一接受西語程度分級的測試。(2)在前置的建構中，兩個學習者族群完成文本的地點不同，靜宜大學的學生是在課堂中完成，而成功大學是在家中完成；為了達到比較對照的目的，在 2009

年的收集中，所有的文本皆在課堂中完成，以統一控制變數。(3)在 2006 年的建置中，初期使用人工進行 XML 的格式註記，為達系統化與提昇便利性，目前則利用多項工具協助註記與分析的工作。「平行雙語學習者語料庫」(CPATEI) 經過前置階段的檢討、改進，現仍處語料庫的建構初期，因其學習者雙語語料在收集與註記工作之特性，語料量之累積速度較慢，待語料累積至一定數量，計畫將會架設可供公開檢索的語料庫介面。

### 3.1 語料收集

為達到便利分析學習者連接詞使用的情形，我們先行結合語料庫技術，建構「西語學習者英西平行語料庫」(CPATEI: Corpus Paralelo de Aprendices Taiwanese de Español e Inglés)。現階段「西語學習者英西平行語料庫」所收集的語料對象是靜宜大學西文系二年級和成功大學外文系西語組二年級兩族群類型的西語學習者（分別是 33 和 22 人），中文是他們的母語、英文是第一外語，且西班牙文是第二外語。

在 2009 年 3 月執行正式測試前，我們先進行前置測試，由 2 位西語學習者參與，各寫一篇西文作文後翻成英文，西文寫作和英文翻譯的時間各為 25 分鐘，西文作文題目是“Un día en la vida de Taiwan”，最後寫出測試的感想。我們透過前置測試的結果和參與者的感想瞭解正式測試時所該注意的事項。進行正式測試的前一週，參與者先閱讀研究說明書（研究目的與方式）後填寫同意書，同意他們的西文作文和英文翻譯被收錄於「台灣西班牙語學習者西英平行語料庫」，日後僅供學術研究之用途；同時，並收集參與者相關的個人資料。一週後於課堂時間執行正式測試，對象分別是靜宜大學西文系二年級學生和成功大學外文系西語組二年級學生，於課堂中以“Un día en la vida de Taiwan”為題依西文作文、英文翻譯的順序，在不可以使用任何工具書的情況下，於 50 分鐘內完成測試。為有效區分學習者西文程度，每位參與者於測試結束後皆接受威斯康辛西語能力分級測驗(Wisconsin Placement Test)。本研究後續分析所使用的語料來源即是上述兩組學習者在課堂內完成的西文作文（6,185 字）及與之對應的英文翻譯（6,057 字），其來源對象的特質分佈如表一所示。

表一、語料及參與者特質分佈情況

學校系級 (2009 年收集)	人數 (男、女)	學習時數 <sup>4</sup>	西語能力檢定分數 (平均數)	語料庫字數 西：6,185；英：6,057
靜宜西文二	33 (1, 32)	1024	371~588 (445)	西：3,721；英：3,520
成功外文二	22 (9, 13)	160	355~548 (491)	西：2,464；英：2,537

<sup>4</sup> 學習時數的計算方式為： $\alpha$  (每週西文上課時數) x 16(週) x 2(學期數) x  $\beta$  (學習年數)。

## 3.2 語料後續處理

### 3.2.1 錯誤修正

語料的後續處理包括：學習者語料之西文作文和英文翻譯分別經由西語母語者修改學生西文作文<sup>5</sup>，以及由英語母語者修改學生英文翻譯之程序完成修正文的語料彙整。該語料庫因此涵蓋學習者西文作文原始文、西語母語者西文修正文、學習者英文翻譯原始文和英語母語者英語修正文共四類語料。

### 3.2.2 可利用之學習者語料分析工具

為達後續有效率系統化分析之目的，本研究所利用之語料庫工具和與之配合的數個小程式分述如下。

(1) Michael O'Donnell 所開發的「語料庫工具」(UAM Corpus Tool)：可便利外籍母語者針對學習者錯誤進行修正、註記者分類註記，結果輸出、檢索與統計的語言分析工具。

(2) Helmut Schmid 的「詞類註記器」(Tree Tagger of University of Stuttgart)：可依西文、英文不同語言別類，自動標註文本之詞類與字根。

(3) Mike Scott 的「WordSmith」詞彙分析工具(WordSmith Tool, 第五版)：可進行詞表高低頻排序、兩頻率表關係性檢定和搭配詞檢索。

(4)「西語搭配詞工具」(Spanish Collocation Tool)<sup>6</sup>：對西語語料進行前後鄰近搭配詞的統計檢定，提供搭配詞與其詞類統計及排序結果，並可進一步比較兩語料搭配詞表之差異性。不同於工具(3)所獲取的結果是詞頻表，「西語搭配詞工具」所提供的則是搭配詞表。

(5) Michael Barlow 的「平行檢索器」(ParaConc)：方便對不同語言的平行檢索。

在過去，就算沒有相關工具或小程式可協助或利用的前提下，研究分析的工作仍可以人力手動的方式一一被完成，但善用工具的優勢，則讓我們節省時間與精力，以投注於更多電腦科技尚無法取代的語言分析工作。

## 4. 連接詞分析應用

### 4.1 研究問題

在以學習者平行語料庫為本之語言分析方面，我們根據 1.研究目的所陳述有關連接詞學習探討的重要性與必要性以及 2.相關文獻的論證，我們設定以下的研究問題：

<sup>5</sup> 為避免母語修改者的風格影響分析結果，本論文所分析的西文原始文經過四位母語者(Rita Swade, Cindy Barquero, Jesús Robla 和 Saul Aviles)的修改，不同母語者間的差異度如下：在成功大學的學習者作文修改是 2.94%，而靜宜大學則是 4.17%，皆在 5%的區間範圍內。

<sup>6</sup> 由成功大學資工所王廷軒助理開發。

- (1) 學習者作文和母語者的修正文中，連接詞使用的分佈情形有何差異？
- (2) 連接詞主要的錯誤類型是什麼？
- (3) 語言程度和第二語是否對第三語的連接詞錯誤使用有顯著影響？

## 4.2 研究對象

以所建構「台灣西語學習者英西平行語料庫」中現階段靜宜 33 名和成大 22 名兩組參與者在課堂內所完成的西文作文（6,185 字）、與之對應的英文翻譯（6,057 字）<sup>7</sup>及由母語者所修正共四類型語料進行研究分析。

## 4.3 研究方法

以語料庫法，充分利用語料庫分析工具的多樣化與便利性，透過詞類與錯誤註記、語料檢索、分類、計算與統計檢定變數間之相關性。與之配合的工具與執行步驟敘述如下。

- (1) 透過「語料庫工具」，母語者修正學習者錯誤，註記者將錯誤分類（錯誤使用、多用、少用）的訂定，註記錯誤類型（根據母語者的修改結果分：選擇錯誤的連接詞（含標點符號）、該用連接詞而未用、不該使用連接詞而使用）和語意屬性（轉折、結果、補充等）。分析者再根據錯誤類型的檢索結果計算，並使用 SPSS (15 版)統計檢定變數間（西語能力檢定、英文連接詞錯誤數、西文連接詞錯誤數）的相關性。
- (2) 使用「詞類註記器」對四類語料（西文作文和英文翻譯的原始文和修正文）依語言分別進行詞類自動註記，將檢索的層次由字詞本身提昇到詞類的範圍。採用互現訊息統計法，針對靜宜和成大兩校學習者的西文作文中連接詞的前後搭配成分進行分析，並以 KL (Kullback-Leibler divergence)法比較學習者原始作文和母語者修正文的差異。
- (3) 採用「WordSmith 詞彙分析工具」協助我們從詞類註記過的西文和英文語料，擷取不同語料中連接詞的使用頻率，原始文和修正文間的相關性，以及連接詞所出現的位置等訊息。
- (4) 透過「西語搭配詞工具」統計檢定，進一步瞭解連接詞前後緊鄰搭配成分的詞彙與詞性。
- (5) 將處理完的語料載入「平行檢索器」，透過連接詞的詞類檢索，獲取西文所平行的英文句中是否有連接詞對應，並重複由英文檢索，查詢計算西文平行對應的結果。

---

<sup>7</sup> 針對論文審查曾提出有關語料量「少」可能影響語言分析結果之代表性的問題，我們以相對於本論文所分析的學習者程度為對象、透過檢索「台灣西語學習者語料庫」(CATE, 其文本類型包含敘述、論說類型，以及藝術、歷史、休閒、健康等九大主題)的語料並逐一進行觀察、對照與檢視本論文所提出的分析結果與論點，藉此增加文本類型的多樣性並彌補現有平行學習者語料量的不足所做的推論結果。其檢索結果如「使用最多者仍是並列連接詞，補充連接詞“además”有少用的情形，轉折及反義連接詞有多用等的情形」和本論文的檢索、分析結果類似。

## 4.4 結果與討論

### 4.4.1 學習者原始文和母語者修正文的連接詞使用比較

表二顯示兩族群學習者之原始文和修正文中連接詞之使用率（使用出現次數/總字數）及其差異<sup>8</sup>。

表二、學習者原始文和修正文之連接詞使用情形

連接詞/總字數	靜宜大學	差距	成功大學	差距
西文原始文	5.03%	0.32%	4.63%	0.24%
西文修正文	5.35%		4.87%	
英文原始文	6.51%	0.51%	6.54%	0.04%
英文修正文	7.02%		6.58%	

從表二，透過比較原始文和修正文中連接詞的使用比率差異，我們發現在西文作文（靜宜：5.03%；成大：4.63%）和英文翻譯（靜宜：6.51%；成大：6.54%）的原始文中連接詞的使用比率皆低於在西語（靜宜：5.35%；成大：4.87%）及英語（靜宜：7.02%；成大：6.58%）母語者修正文中的連接詞使用率，也就是學習者低於母語修改者的使用率。此外，我們也觀察到：在成大西文(0.24%)的使用率差異比英文大(0.04%)，而在靜宜的使用率差異則是英文(0.51%)比西文大(0.32%)。整體而言，兩族群的學習者在連接詞的使用率都還要再提高，才能更接近母語者的使用情形。此外，主修的不同也反應在不同語言連接詞使用率的差異上。亦即：外文系學習者西文連接詞使用率差異大於英文；反之，西文系的學習者英文連接詞使用之差異大於西文連接詞。外文系的在英文原始文和修正文中連接詞使用率的差異小於西文系的表現，第二語（英語）的程度差異可以合理地解釋此一結果。而另一方面，西文系和外文系在西文原始文和修正文中的連接詞使用率的差異雖然不大，但西文系在西文連接詞使用率的差異表現上有些微高於外文系學習者的連接詞使用。我們透過進一步觀察學習者所使用的連接詞發現<sup>9</sup>：連接詞使用的傾向可能是導致此

<sup>8</sup> 由於我們所比較的兩個族群的作文篇數不同，故採以使用率（連接詞數/總字數）的計算方式較為客觀公平。如和其它相關文獻在探討連接詞的使用率（以所有連接詞為分母，去計算不同類型或個別的連接詞的使用比例）做比較，我們上述的計算方式所得出的平均值會顯得比較小。

<sup>9</sup> 學習者連接詞使用情形如下表所示。

連接詞類型(連接詞/總字數)	靜宜大學	成功大學
y（聯繫）	2.95%	3.08%
o（選擇）	0.86%	0.73%
cuando（時間）	0.24%	0
porque（因果）	0.27%	0.12%

一結果的因素之一，因為被視為較易使用的西文並列連接詞(y/o)，在靜宜西文系的學習者方面，所使用的比率佔所有連接詞的 75.94%，低於成大西文系學習者的使用率(82.46%)。亦即：學習時數較少的學習者（成功大學學生）在較簡單的聯繫和選擇連接詞上使用得比較頻繁。相對的，在其它連接詞使用的傾向，如反義連接詞(如：“pero/sin embargo”「但是」)或是附屬類型中的因果連接詞(如“porque”「因為」)等，靜宜西文系的學習者所使用的比率則高於成大西文組學習者的使用率。並列的聯繫連接詞是較易使用的連接詞類型，成大的學生使用得比較多，正確率相對容易提高；而靜宜在較難使用的連接詞上使用得比較多，也容易出現錯誤，所以在與母語者修改的結果比較後，成大學生在原始文和修正文的差距反而略微小於靜宜學生的表現。此外，我們可以推論出連接詞學習的發展階層為並列的聯繫連接詞“y/o”早於反義連接詞“pero/sin embargo”。

接著，我們進一步去觀察並列連接詞使用在第二語和第三語間的細部對應關係，我們以平行句的對應關係去瞭解在學習者原始文裡，西文對應到英文時連接詞於兩個語言出現的比例數，並與母語修改者的修正文裡西譯英連接詞使用比例做比較，結果如表三所示。

**表三、西譯英並列連接詞使用之比較**

學校	靜宜大學		成功大學	
	原始文	修正文	原始文	修正文
西譯英連接詞使用	1.75:1	2.01:1	1.40:1	1:53:1

從表三，我們可以觀察到西文對應至英文的並列連接詞在學習者原始文和母語者修正文的差異，其結果驗證成功大學外文系西語學習者(1.40:1.53)在連接詞使用的表現確實比靜宜大學西文系學習者(1.75:2.01)接近母語者修正文的西譯英比例。

有關西文學習時間長度較長的學習者和母語者的使用率差異高過西文學習時間較短的學習者的情形，在後續 4.4.3 章節中，我們將特別針對西文能力程度（而非西文學習時間長度）進行統計檢定驗證語言程度與連接詞使用正誤的相關性。

進一步，我們分析學習者所使用的連接詞特性時發現：學習者在連接詞的使用多集中於“y(e)/and”和“o(u)/or”並列連接詞的類型，當我們限定第一字之詞類為連接詞，且同時兼顧互現訊息分數和結果筆數皆高者的範例如“y luego, y yo, y me”。以右側被搭配詞為附屬性連接詞的“luego”（“y luego”）為例，表四前兩列是兩組學習者在原始文和修正文

sin embargo（反義）	0.05%	0
pero（反義）	0.21%	0.12%
si（條件）	0.11%	0
aunque（讓步）	0.03%	0

中“y luego”兩詞搭配出現的統計結果。我們進一步查詢“y luego”一詞在「西班牙語語料庫」中的出現情形（表四第四列）<sup>10</sup>。

**表四、「y luego」搭配詞於原始文和修正文出現的情形**

y luego	搭配次數	MI 值	“y”出現個數	“luego”出現個數
原始文	3	3.76	188	6
修正文	5	83.56	202	6
西班牙語語料庫	4,239	6.6	1,249,549	19,283

從表四，我們可以觀察到：連接詞與其後被搭配成分“luego”共同出現之機率在原始文的搭配次數是 3，而在修正文是 5，“y luego”這個搭配詞在修正文的互現訊息值(83.56)高過原始文(3.76)，“luego”此一附屬性的連接詞在母語者的修正文中幾乎只要出現就跟“y”做搭配組合，但此一被搭配成份在學習者的原始文本中則僅有 50%的機率是與連接詞“y”共同出現，我們進一步查詢到“y luego”在「西班牙語語料庫」中的出現個數(4,239)與互見訊息值(6.6)。結果顯示其互見訊息值雖不如修正文高，但仍約為原始文的兩倍，類似這樣的連接詞搭配範例，可以作為教學過程中提醒學習者特別注意的重點之一。

#### 4.4.2 錯誤類型

學習者在西文作文和英文翻譯的錯誤率(連接詞錯誤數/連接詞使用數)結果如表五所示。

**表五、錯誤率分佈情況**

語言	總錯誤率	誤用	多用	少用
西文	18.60%	4.32%	4.65%	10.30%
英文	10.63%	2.53%	1.72%	5.82%

由表五，我們觀察到學習者在西文連接詞使用的錯誤率高於英文。此外，我們也注意到根據誤用、多用和少用的不同錯誤類型分類：在學習者西文作文中，錯誤率所佔比例由高到低分別是少用、多用和誤用；而在其英文翻譯中，則依序為少用、誤用和多用。因此，不管在學習者的西文作文或英文翻譯的文本中，少用的情形都是三類型中錯誤率最高者。如果再細分三類錯誤的連接詞屬性，我們特別注意到誤用和少用的情形多集中於補充類型屬性的連接詞；而多用的情形則偏屬於反義類型的連接詞。上述的分析觀察和 Granger y Tyson (1996)的研究發現剛好相反，因為他們指出以英語為外語之學習者的英語作文中，反義屬性的連接詞有使用不足的情形，而補充屬性的連接詞則有使用過多的現象。此外，Granger y Tyson (1996)和 Field & Yip (1992)指出在以非英語為母語的學

<sup>10</sup> 感謝論文審查者建議加入「西班牙語語料庫」的檢索結果（表四第四列），輔助修正文與原始文之對比，更能凸顯該連接詞搭配的重要性。

習者在其第二語的英語寫作中，皆出現連接詞使用過多的情形，他們推論可能和受母語影響有關。上述研究結果的主要差異在於目的語不同，Granger y Tyson 和 Field & Yip 的研究對象是以學英文為外語的學習者，而我們的研究對象則是以中文為母語、英文為第一外語、西文為第二外語的台灣西語學習者。而 Blanco (2007)和我們一樣以台灣的西語學習者為對象所收集語料的分析，其獲致的結果與我們有類似之處：學習者在反義連接詞有多用的現象。以上的比較結果顯示不同目的語間存在著連接詞使用的差異現象。

再者，針對母語這個影響因素的探討，如就連接詞全面性加總的結果而言，我們發現台灣的西語學習者傾向少用連接詞，連接詞整體性的使用傾向主要是由於聯繫連接詞在所有連接詞中佔有極高的比例。我們根據 Lu *et al.* (2009) 在平行語料中連接詞使用排序（西文>英文>中文）的研究結果(如表六第四列連接詞的總和與第一列之聯繫及補充之並列連接詞檢索結果所示)可以解釋台灣西語學習者連接詞少用的現象，即：在相對平行的語意中，中文少用連接詞的情形影響學習者第二外語西語的學習和使用結果。

表六、「西英漢平行語料庫」之檢索結果

	西文(37,407)	英文(38,895)	中文(52,546)
聯繫、補充	2,768	2,431	328
反義、因果	386	567	129
總和	3,154	2,998	557

如進一步對連接詞語意屬性進行較為細項分類，我們發現不同類型連接詞間的差異：台灣的西語學習者在聯繫和補充的並列連接詞中有少用的情形，而反義和因果連接詞則有多用的情形。一方面，針對不同類型的連接詞比較，台灣西語學習者在並列和補充連接詞少用的情形可以被這兩類連接詞在平行對應語料中(如表六第二列所示)所呈現的分佈傾向排序（西文>英文>中文）所解釋：因為對等語意中，中文傾向不用並列的聯繫或補充連接詞，所以台灣西語學習者在進行西文寫作時，受到母語影響的結果會導致少用的情形。另一方面，對於反義和因果的多用情形，我們則以中西文的對比分析來解釋：因為中文「雖然...可是/但是」和「因為...所以」的表達，在西文的使用上，屬反義和因果的連接詞成分“*aunque, sin embargo, no obstante, pero*”以及“*como, porque, por eso, por lo tanto*”只能擇一出現，我們根據 Anderson (1990, 1993)「學習原則」中的「一對一原則」解釋學習者在反義和因果連接詞多用的結果。此外，針對反義和因果兩類連接詞在平行語料中出現頻率的計算，我們獲致「英文>西文>中文」之使用頻率的高低排序(如表六第三列所示)，這個階層排序顯示西文和英文間的使用情形與前述的補充並列類型有異，我們因此推論台灣西語學習者的第一外語英文和第二外語西文間可能存有某種程度的互動關係，導致影響了學習者的使用結果，這部分則有待後續進一步的研究與探討。

#### 4.4.3 語言程度和第二語與第三語的關係

西語能力、英文翻譯和西文作文的連接詞錯誤數之間的相關性檢定結果如表七所示。

表七、相關性檢定結果

學校	西語能力		英文錯誤數	
	靜宜大學	成功大學	靜宜大學	成功大學
西文錯誤數	0.570	0.220	0.265	0.488

透過表七，我們可以瞭解到西語能力檢定、英文連接詞使用錯誤數和西文連接詞使用錯誤數彼此之間的相關性。SPSS 的檢定結果顯示：三變數間的相關性皆未達顯著水準 ( $P>0.05$ )者：西語程度和英語連接詞的錯誤使用數 ( $P=0.002$ )皆分別與西語連接詞使用的錯誤數的多寡無顯著關係存在；亦即西語連接詞的使用錯誤與西語程度無顯著的關係性 (靜宜  $P=0.570$ ；成大  $P=0.220$ )，且和第二語 (英語) 之相關性亦未達顯著差異的水準 (靜宜  $P=0.265$ ；成大  $P=0.488$ )。這和 Wu (1996)研究學習者之英語連接詞使用與學習者第二語英語能力之間有關係的結果不同，我們發現西語連接詞的使用錯誤和第三語西語語言程度的高低沒有相關性存在。此外，我們的檢驗結果也顯示第二語和第三語連接詞錯誤使用彼此間亦無必然之關係。

### 5. 教學因應之道

余昇易(2006)和曾嬾竹(2004)在探討連接詞教學對學生寫作表現成效的影響的研究，其結果顯示連接詞的密集教學能有效改善語言程度較低學生之整體寫作表現。根據上述對學習者錯誤的分析和討論結果，我們提出以下的教學因應之道。而我們所參考的是依據母語者的「西班牙語語料庫」(Corpus del Español<sup>11</sup>，四千一百萬字)、「現代美語語料庫」(Corpus of Contemporary American, COCA<sup>12</sup>，兩千萬字)和「現代漢語平衡語料庫」(Academia Sinica Balanced Corpus of Modern Chinese<sup>13</sup>，八百萬字)的檢索結果，按照連接詞語意分類所做的排序。這三個極具代表性的母語語料庫的共同特色是語料量豐富且檢索功能強，特別值得一提的是前二者，因建置、註記與程式系統模式類似，對西英雙語的研究非常有幫助。我們根據學習者連接詞使用的分析研究結果，在教學因應策略上針對相關連接詞透過上述三個母語語料庫進行檢索、分類與排序，以作為教學內容設計的參考，而這裡僅舉出補充和反義連接詞類型的常用連接詞作為範例，結果如表八所示<sup>14</sup>。

<sup>11</sup> Corpus del Español: <http://www.corpusdelespanol.org/x.asp>

<sup>12</sup> <http://www.americancorpus.org/>

<sup>13</sup> <http://dbo.sinica.edu.tw/SinicaCorpus/>

<sup>14</sup> 西文的數據是在「西班牙語語料庫」(Corpus del Español)設定 19、20 世紀為查詢範圍的檢索結果。

表八、連接詞於西、英和中文三語言之頻率高低排序

連接詞	西文	英文	中文
補充 連接詞	además (15,376) a la vez (2,361) asimismo (2,184) por otra parte (2,118) a su vez (1,902) por otro lado (962) por un lado (731) por una parte (517)	on the other hand (17,986) besides (16,747) moreover (16,166) on one hand (10,444) furthermore (10,351) on the one hand (4,159)	而且 (2,637) 並且 (995) 此外 (994) 一方面 (478) 還有 (338) 另一方面 (338) 除此之外 (113)
反義/ 讓步 連接詞	pero (131,191) aunque (28,397) sin embargo (14,566) mas (13,740) a pesar de (6,364) sino que (5,690) no obstante (3,588) empero (634)	but (1,811,191) though (164,532) however (142,356) yet (135,705) although (107,965) rather (100,144) even though (35,150) even if (33,802) nevertheless (13,204)	但是 (4,953) 雖然 (2,976) 可是 (2,532) 然而 (1,372) 而是 (932) 儘管 (451) 就算 (321)

表八呈現了台灣西語學習者的第一語中文、第二語英文和第三語西文三個語言的母語者在連接詞使用上高低頻的順位，如：在西文補充屬性的同義連接詞中依序以“además, a la vez”和“asimismo”的出現頻率排序最高，同義的英文連接詞依序為“besides, moreover”和“furthermore”，中文則是「而且、並且、此外」。而我們也注意到中文表達中的「一方面」和「另一方面」的高低頻排序(前者高於後者)不同於西文的“por otra parte, por otro lado, por un lado, por una parte”和英文的“on the other hand”和“on the one hand”；且西文中的“por otra parte, por otro lado”高於“por un lado, por una parte”。又如西文的反義連接詞高低頻依序為“pero, sin embargo, no obstante, empero”。教學者若善用上述表格中連接詞排序之跨語言的同異特質，根據聽、說、讀、寫不同需求的語言技能設計教案，預期學習者應可從中獲益。

接下來，即是由真實的文本中學習。這一個程序分兩方面進行，先是透過「語料庫工具」直接檢索出由西語母語者所修改的修正文，以便與學習者原始作文進行對比，分「多、少和誤用」類別以及根據「讓步、結果和補充」語意分類提醒學習者注意(如句子(1)a-b, (2)a-b 所示)。接著再間接搭配「Google 翻譯器」(Google Translator)<sup>15</sup>的英、中翻

<sup>15</sup> <http://translate.google.com/#>

譯(如句子(1)c-d, (2)c-d 所示), 讓學習者理解其母語(中文)和第一外語(英語)的語意。雖然自動翻譯的結果和人工翻譯仍有相當的改善空間, 但不可諱言, 「Google 翻譯器」仍是目前自動翻譯工具中成效較佳者。對於因應教學的大量範例, 考量效率與便利的因素, 我們採用人工修正自動翻譯結果的模式來節省時間與人力。

(1) 補充連接詞少用

- a. habitualmente leer en la cama hasta me duermo\_luego apaco la luz (學習者西文作文原始文)
- b. habitualmente leo en la cama hasta que me duermo **y** luego apago la luz (母語者修改文)
- c. usually I read in bed until I fall asleep and then turn off the light (英文翻譯)
- d. 通常我會躺在床上看書, 一直到我睡著, 然後關燈。(中文翻譯)

(2) 反義連接詞多用

- a. Aunque es un poco sosa, pero me gusta la. (學習者西文作文原始文)
- b. Aunque es un poco aburrido, **me** gusta. (母語者修改文)
- c. Although it is a bit boring, I like it. (英文翻譯)
- d. 雖然有點無聊, 但是我喜歡。(中文翻譯)

最後, 則是藉由西語母語者語料庫「西班牙語語料庫」的關鍵字居間(Keywords in Context: KWIC)功能, 透過同一類型的一連串連接詞範例, 強化學習者的印象。我們進一步觀察在相關連接詞前後高頻率出現的成分特色, 藉此教導學習者學習西語母語者在連接詞使用的慣用模式, 以之前所討論過的搭配詞“y luego”為例, 搜尋結果如範例((3)a-c)的句子所示。

(3)

- a. Cuba a los platillos volantes, de los que se habla durante mucho tiempo **y luego** pasan años sin que se vuelva a decir nada de ellos.<sup>16</sup>
- b. y mi concentración no rinden al máximo. Si yo dirigiera una sola vez **y luego** nada, otra vez y luego nada, entonces no conseguiría buenos resultados.<sup>17</sup>
- c. máquinas comerciales cada vez más. En primer lugar porque es muy práctico, **y luego** porque nosotros no somos una fábrica de ordenadores.<sup>18</sup>

<sup>16</sup> Title: Entrevista (ABC). Author: Corral Pedro. Source: <http://www.abc.es>.

<sup>17</sup> Title: Entrevista (ABC). Author: Rubio José Luis. Source: <http://www.abc.es>.

<sup>18</sup> Title: Entrevista (ABC). Author: Rubio José Luis. Source: <http://www.abc.es>.

根據上述範例，我們發現主要包含“...(,)y luego...”和“Y luego...”兩種模式。如逐字解釋，“y”是連接詞，“luego”表時間「之後」之語意，當兩詞搭配時，則具有前後語意連接、口氣延順，以及進一步說明之意。舉例來說：3(a)和(b)句中的“y luego”屬於連接前後語意和口氣；而 3(c)句中的“y luego”則有進一步說明原因之意。

## 6. 結論

本論文結合語料庫技術建構「西語學習者英西平行語料庫」(西文作文 6,185 字及英文翻譯 6,057 字)，並以之為本、輔以應用工具，進行學習者連接詞使用之分析研究。在語料庫建構方面，語料的後續處理包括修正與錯誤的註記。此外，為達後續有效率系統化分析之目的，本研究也介紹了目前可利用的語料庫工具及與之配合的開發程式。

在以語料庫為本的應用分析方面，和過去文獻之研究結果不同，我們獲致以下結論：(1)在西文作文和英文翻譯的原始文中，連接詞的使用比率皆低於在西語及英語母語者修正文中的連接詞使用率，學習者在連接詞的使用率還需加強。我們進一步透過母語者平行語料的比較，對連接詞的少用可能和跟學習者的母語(中文)有關提出驗證說明。(2)在錯誤類型上，我們發現少用的情形集中於「補充」類型的連接詞；多用的情形則多屬於「反義」類型的連接詞。(3)在可能的影響變數檢定中，連接詞的使用錯誤與西語程度間，未呈現顯著的相關性，且和第二語(英語)之相關性也未達顯著差異的水準。最後，我們根據分析結果，利用現有之工具，提出連接詞的教學範例以供參考。

本論文的限制主要和學習者語料的數量有關。受制於雙語平行語料處理本身的複雜度、現有人力與資源分配等外在因素的限制，在語料數量上無法快速擴增，是本研究的一大限制。此外，因為目前所分析的筆數有限，日後也需要較大量的語料來進一步衍生、驗證與強化本論文與過去研究結果不同的結論，使之更具說服力。在未來工作方面，除了擴增語料庫的語料量外，我們還將從應用工具與分析議題兩方向努力。在應用工具方面，如「西班牙語詞語搭配工具」目前僅限於前後緊鄰成分的搭配關係(bi-gram)，我們會朝理想化的多詞搭配關係(N-gram)邁進。另外，我們將以相同的研究模式與路徑，對不同學習議題進行分析研究，以朝逐步涵蓋第三語之研究廣度與深度的目標努力。

## 感謝

首先，感謝成功大學頂尖計畫(R021)以及國科會人文學研究中心前置規劃案(「中介語的發展及變異性」整合型(2007.4~2008.6)計畫之西班牙語子計畫)的經費補助。再者，感謝 UAM 所開發的 Corpus Tool 這個免費工具的資源共享。最後，對協助語料庫建構、作文修改(巴辛蒂、史瓦帝、艾索爾、羅布拉、德瑞克)及語料工具開發和分析的助理群們(王廷軒、仇志良、吳奕賢、張展毓、劉亭佑、張偉婷、葉倩廷、張以姍、林子涵、陳姿均、吳彥融)致上最誠摯的謝意。

## 參考文獻

- 余昇易 (民 95)。連接詞教學及其對大學生寫作影響之探討。元智大學應用外語學系碩士論文。
- 周玉玲 (民 91)。短篇故事在英語連接詞教學法效益之研究。台灣師範大學英語系在職進修碩士班論文。
- 孫義禎 (民 82)。西班牙語實用語法新編。台北市：建宏出版社。
- 陳樹升 (民 82)。西班牙文文法。台北市：中央圖書出版社。
- 曾熾竹 (民 93)。線上連接詞教學對大學生英文寫作之效能研究。清華大學外國語文研究所碩士論文。
- 劉啓分 (民 82)。實用西班牙文文法。台北市：中央圖書出版社。
- 盧慧娟、呂羅雪 (民 98)。語料庫為本之中介語研究。淡江外語論叢，13，30-47。
- Anderson, R. (1990). Models, processes, principles, and strategies: Second Language acquisition in and out of the classroom. In B. VanPatten and J. Lee (Eds.), *Second Languages Acquisition: Foreign Language Learning* (pp.45-78). Clevedon, UK: Multilingual Matters.
- Anderson, R. (1993). Tour operating principles and input distribution as explanations for underdeveloped and mature morphological systems. In Kenneth Hyltenstam and A. Viborg (Eds.), *Progresion and Regresion in Language* (pp.309-339). Cambridge: Cambridge University Press.
- Blanco, P. D. (2007). Los conectores en noticias de sucesos, retórica contrastiva en textos de periodistas españoles y alumnos taiwaneses. *Círculo de Lingüística Aplicada a la Comunicación*, 30, 24-44.
- Davies, M. (2002-) Corpus del Español (100 million words, 1200s-1900s). Available online at <http://www.corpusdelespanol.org>.
- Davies, M. (2008-) The Corpus of Contemporary American English (COCA): 410+ million words, 1990-present. Available online at <http://www.americancorpus.org>.
- De Angelis, G. (2005). Interlanguage transfer of function words. *Language Learning*, 99(3), 379-414.
- \_\_\_\_\_. (2007). *Third or Additional Language Acquisition*. Clevedon: Multilingual Matters.
- Field, Y. & Yip, L. M. O. (1992). A comparison of internal cohesive conjunction in the English essay writing of Cantonese speakers and native speakers of English. *RELC Journal*, 23, 15-28.
- Gómez Torrego, L. (1997). *Gramática Didáctica del Español*. Madrid: SM.
- Granger, S. & S. Tyson. (1996). Connectors usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes*, 15(1), 17-27.
- Leung, Y-K. I. (2005). L2 vs. L3 initial state: A comparative study of the acquisition of French DPs by Vietnamese monolinguals and Cantonese-English bilinguals. *Bilingualism: Language and Cognition*, 8(1), 39-61.

- \_\_\_\_\_. (2006). Full transfer vs. partial transfer in L2 and L3 acquisition. In R. Slabakova, S. Montrul, & P. Prévost (Eds.), *Inquiries in Linguistic Development: In Honor of Lydia White* (157-188). Amsterdam: John Benjamins.
- Lu, Hui-Chuan, Cheng-Yu Chang & Lo Hsueh Lu. (2009). Creation and application of CPATEI, a parallel Spanish-English learner corpus with Mutual Relevance Analysis between multiple interlanguages. Abstract accepted by AACL in Alberta, Canada.
- Lu, Hui-Chuan & Lu Lo Hsueh. (2005). Estudio sobre la interacción entre las estructuras textuales y oracionales. *Glosas Didácticas*, 15, 145-156.
- \_\_\_\_\_. (2006). Estudio del léxico sobre un corpus de textos de lectura. *TamKang Studies of Foreign Languages and Literatures*, 28, 35-50.
- \_\_\_\_\_. (2009). Creación del Corpus Paralelo de Aprendices Taiwanese de Español e Inglés (CPATEI). *Foreign Language Studies*, 9, 135-156.
- Mark, K. (2001). A parallel learner corpus: Using computers in a humanistic approach to language teaching and research. *Japanese Journal of Language and Society*, 4(1), 5-16.
- Martínez, R. (1997). *Conectando Texto. Guía para el Uso Efectivo de Elementos Conectores en Castellano*. Barcelona: Octaedro.
- Martín, M. & J. Portolés. (1999). Los marcadores del discurso. In I. Bosque & V. Demonte (Eds.), *Gramática Descriptiva de la Lengua Española* (pp.4051-4213). Madrid: Espasa Calpe.
- Montolío, E. (1998). La teoría de la relevancia y el estudio de los marcadores discursivos. In Martín Zorrquino, M. & E Montolio (Eds.), *Los Marcadores del Discurso. Teoría y Análisis* (pp.93-119). Madrid: Arco Libros.
- \_\_\_\_\_. (2000). *Manual Práctico de Escritura Académica*. Vol III. Barcelona: Ariel.
- \_\_\_\_\_. (2001). *Conectores de la Lengua Escrita*. Barcelona: Ariel.
- Poblete, M. T. (1999). La cohesión de los marcadores discursivos en distintos tipos de discurso. *Estudios Filológicos*, 34, 165-180.
- Romero, M. G. (2005). Análisis de los conectores en ensayos argumentativos escritos por estudiantes universitarios. Ponencia presentada en XXIV ENDIL de San Cristóbal, Venezuela.
- Ringbom, H. (2007). *Cross-linguistics Similarity in Foreign Language Learning*. Clevedon: Multilingual Matters.
- Sánchez, I. (1993). ¿Por qué son tan incoherentes los ensayos que escriben los estudiantes? *Tierra Nueva*, 1, 87-93.
- Shimizu, M., Du, F., & Dantsuji, M. (2004). Construction of the Chinese Learners' Parallel Corpus of Japanese and its preliminary analysis. *K@ta*, 6(2), 91-105.
- Wu, Jwo-Shien. (1996). The Relationship between conjunction use and language proficiency. *Studies in English Language and Literature*, 1(8), 50-57.

## 語料庫與工具

Academia Sinica Balanced Corpus of Modern Chinese, <http://dbo.sinica.edu.tw/SinicaCorpus/>

Google Translator, <http://translate.google.com/#>

Corpus de Aprendices Taiwaneses de Español (CATE), <http://corpora.flld.ncku.edu.tw/>

Corpus del Español, Mark Davies, <http://www.corpusdelespanol.org/x.asp>

Corpus Escrito del Español L2 (CEDEL2), UAM,  
<http://www.uam.es/proyectosinv/woslac/cedel2.htm>

Corpus of Contemporary American (COCA), Mark Davies, <http://www.americancorpus.org/>

ParaConc, Michael Barlow, <http://paraconc.com/>

Spanish Learner Language Oral Corpus (SPLLOC), <http://www.splloc.soton.ac.uk/>

Tree Tagger, University of Stuttgart,  
<http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/>

UAM Corpus Tool, Michael O'Donnell, <http://www.wagsoft.com/CorpusTool/>

WordSmith Tool v.5, Mike Scott, <http://www.lexically.net/wordsmith/>