

第三章 研究方法

本章之目的在說明本論文之研究方法與研究過程，共分成四節：第一節為研究方法之選定；第二節為本研究對象及工具；第三節為本研究之程序；第四節則介紹本研究資料蒐集、整理及分析方式。

第一節 研究方法之選定

本研究運用「內容分析法(content analysis)」探討「在」的語法及語義內容，並考察英語為母語的華語學習者使用「在」的情形以及偏誤現象，最後進一步編擬教學語法內容原則及排序，同時檢驗教材中語法點之說明與編排方式。

一、內容分析法之內涵

所謂內容分析(content analysis)，亦稱資訊分析(informational analysis)或文獻分析(documentary analysis)。內容分析以檔案記錄、文件等資料為分析對象，主要在解釋某特定時間某現象的狀態，或在某段期間內，該現象的發展情形。(王文科，1995)。

Frankfort-Nachmias & Nachmias(2003: 406-407)提到「內容分析的程序牽涉到兩種過程的互動：研究者指出所要研究的內容特性，並且當它們在被分析物體出現時，應用規則來確認與紀錄這些特性。很顯然，內容所被登錄的類別將隨研究問題與資料的本質而有所不同。」

內容分析是一種量化的分析過程，但並不是一種純粹的「定量分析」，他是一種「質」與「量」並重的研究方法。由於注重角度的不同，有偏向定量分析，也有著重定質分析的(楊孝燦，1989)，見表 3-1：

表 3-1 內容分析法：量的研究 vs. 質的研究

	量的研究	質的分析
研究內容	重視「顯著內容」即構成文件物質非推論的資料。	探討「潛在內容」即構成文件材料中推論、內隱或潛在意義。
研究樣本	選樣範圍既大且廣，注重類目統計，以頻率多寡為主。	選樣範圍較小。
研究分析	系統的、客觀的	彈性的、主觀的
研究方向	機械的、系統的、類目選擇與分類極為重要	低頻率或非頻率研究

其中，根據吳明清(1998)及 Frankfort-Nachmias & Nachmias (2003)，量的內容分析包含抽取文獻樣本、界定類別、內容、界定紀錄的單位、界定脈絡單位，以及界定計數系統等幾項主要任務，茲分別說明如下：

(一) 抽樣：內容分析的抽樣程序最先步驟為編列抽樣架構，或列出將可能被抽取當樣本的所有單位表。抽樣的架構確定之後，以隨機抽樣最為普遍。

(二) 類別：類別與研究目標有關，它是進一步分析內容的依據。分類必須要有窮盡性(exhaustiveness)及互斥性(mutual exclusivity)。所謂的窮盡性是確保與研究有關的每一個紀錄單位，皆可加以分類；互斥性則表示沒有一個紀錄單位，會被系統中超過一個以上的特定類別所舍括。

(三) 紀錄的單位：類別的選擇未必能決定合適的紀錄單位或分析單位。亦即就某些文獻以及某些類別來說，並無獨一無二的紀錄單位。在內容分析中常用的五個紀錄單位分別是單字(words)或術語(terms)、主題(themes)、人物(character)、段落(paragraphs)以及項目(item)。

(四) 計量系統：

1. 時間—空間系統(time-space system)：是基於測量空間或時間單位，以描述被分析物質不同類別的相對重要性。

2. 出現系統(appearance system)：是登錄者用來尋找物質某種屬性的出現。脈絡單位的規模將決定發生的次數，而記錄單位在鄰近地方的重複發生是被分開計數的。

3. 次數系統(frequency system)：是將每次特定屬性的發生予以記錄。
4. 強度系統(intensity system)：一般用於處理態度和價值的研究。

二、選用內容分析法之原因

本研究選擇「內容分析法」為研究方法有以下五點理由，分述如下：

(一) 內容分析法適用於研究對象無法接觸時。

本文之研究對象幾乎已無法接觸，捨內容分析法之外，別無其他替代法。如採用已建置而成之母語者語料庫或學習者中介語料庫，皆有研究對象無法接觸的侷限。

(二) 內容分析經常採用較大的樣本。

本文乃以語料庫為本(corpus-based)的研究，不管是分析各類「在」之母語者使用頻率或是學習者之使用及偏誤情形，都需要較大的樣本數，以求提高所得值之客觀性及準確度。

(三) 內容分析法適合縱貫式分析。

透過語料庫並利用內容分析法，可觀察各程度的學習者語言使用及偏誤現象。

(四) 內容分析的記錄資料是有自然發生性。

母語者語料庫和學習者語料庫中的語料都是自然語料。

(五) 內容分析比大規模的調查，成本較低，且較省時。

內容分析比大規模的調查成本低、且較省時：因為許多文獻大多集中在圖書館、報館資料室，均易查詢。大部分語料來源自既有的、開放性的語料庫，僅需上網取得即可。

綜合以上幾點，因此本論文採用「內容分析法」做為研究之方法。

三、內容分析法之限制

採用內容分析法有其適切性及優點，但亦有幾點限制，如下所述：

(一) 內容分析易有抽樣偏差發生。

例如：使用語料庫檢索而得之分析材料即可能產生抽樣誤差。

(二) 文獻提供的僅是語文的行為。

例如：語料庫僅提供非語文的行為，則未提供直接的資料。語料庫之內容能包含書面的語料，或是口語語料轉寫成的檔案，無法提供非語言信息，如肢體語言、聲調等。

(三) 內容或材料殊異或格式缺乏標準化。

例如：使用語料庫檢索而得之分析材料，其最原始的來源文本之內容分佈可能不相同，也可能是缺乏標記或標記格式不一。前者沒有解決方法；至於後者，筆者在統整語料時，將以統一格式分析處理。

第二節 研究流程

本研究之研究流程如圖 3-1 所示。

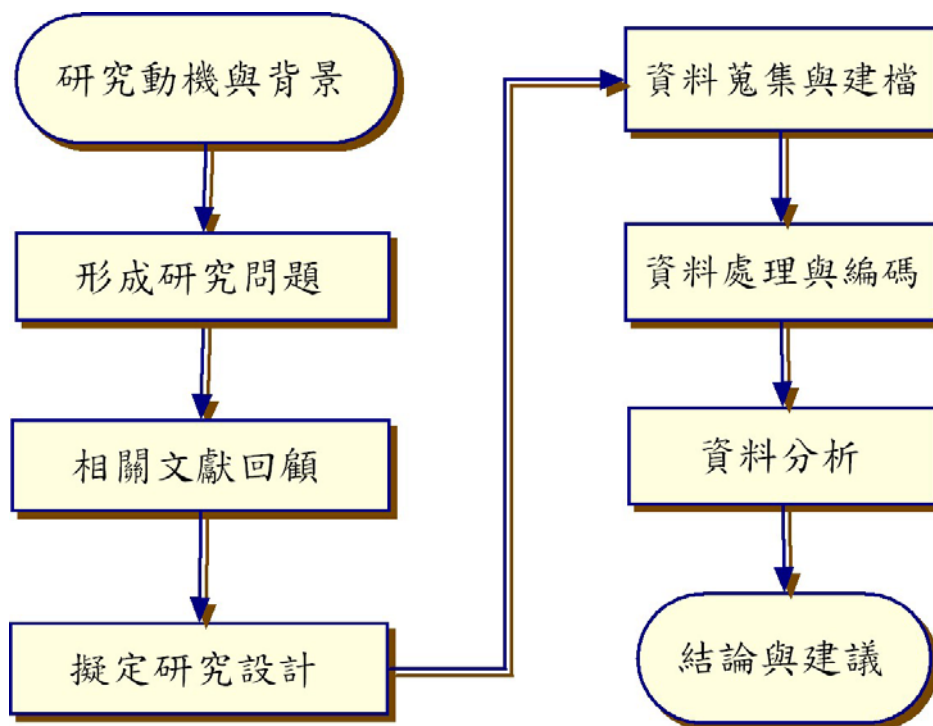


圖 3-1 研究流程圖

第三節 研究對象、工具及資料之分析處理

本文研究之對象為現代漢語中的「在」，文獻考察取向包含以下幾個個面向：(一) 相關文獻資料、(二) 母語者語料庫、(三) 學習者語料庫、(四) 第二語言教材。以下再分別針對(二)、(三)、(四)，說明其內容涵蓋項目及資料處理、分析方法。

一、母語者語料庫

(一)、語料來源、規模及內容

本研究使用兩岸的母語者語料庫，以北京大學 CCL 現代漢語語料庫及台灣中央研究院現代漢語平衡語料庫來考察各類「在」的使用情形。語料庫的規模及內容分述如下：

1. 北京大學 CCL 現代漢語語料庫

CCL 現代漢語語料庫乃北京大學中文系漢語語言學研究中心 CCL 所建置，語料規模合計約兩億六千萬字 (264,444,436)。語料來源包含報刊、文學、翻譯作品、應用文、電視電影、戲劇、網路文章、口語等資料。

2. 台灣中央研究院現代漢語平衡語料庫

平衡語料庫是由中央研究院資訊所、語言所詞庫小組所建置完成的。語料蒐集始自一九九〇年前後，於一九九五年七月完成第一版(兩百萬詞)，一九九六年十一月規劃完成 WWW 版，一九九七年完成 3.0 版 (約五百萬詞)，並在同年十月上線使用，本研究所採用的語料庫乃是最新的 4.0 版，字數總計約七百九十萬字 (7,949,851)。

語料內容包含書面及口語，涵蓋哲學、科學、社會、藝術、生活、文學等主題，取得來源包括報紙、雜誌、學術期刊、教科書、工具書、學術論著、一般圖書、視聽媒體、會話訪談等。

針對母語者語料的部分，本研究利用 CCL 現代漢語語料庫及平衡語料庫作為研究材料，所使用之語料不另行篩選主題或文體，而是囊括所有範圍的語料。然而，由於目標樣本數量龐大，再加上平衡語料庫線上版有檢索限制：院外檢索限兩千行資料，因此筆者在兩個語料庫中，鍵入關鍵字「在」，

採抽樣方式篩選語料來分析。其中，對比兩岸「在+VP」使用情形，各取兩岸語料庫中「在+VP」語料前 2000 筆分析；計算各類別「在」之使用頻率時，則於各兩岸語料庫中取前 1000 筆作為分析材料。

(二)、語料儲存/管理、語料標注方式

本研究將從兩個語料庫中篩選出的語料，以 Microsoft Office Excel 作為儲存與管理之工具，分別建檔。所蒐集之語料均依句鍵入，語料庫所記錄之格式項目包含本研究之所分之類別（在_x）與例句。

二、學習者語料庫

透過學習者中介語料庫，考察英語為母語的華語學習者之中介語中各類「在」的使用及偏誤情形，進而歸納其偏誤類型並提出解釋，此為本研究重心之一。以下針對學習者語料庫之來源、規模、內容、語料儲存/管理所使用之應用程式、語料篩選、語料標注/分析方式等項提出說明。

(一)、語料來源、規模及內容

本研究所使用的學習者語料庫中的語料來源有三，分別是（1）國立台灣師範大學華語文教學研究所鄧守信教授專案研究建置之中介語料庫，以及（2）國立台灣師範大學華語文教學研究所曾金金教授於 94 及 95 學年度第二語言習得課中蒐集的語料。本研究在前面兩個來源中，利用輸入關鍵字「在」檢索所有句子後，找出屬於本文研究範圍者，其他如「實在」、「現在」、「在乎」等包含「在」的詞均於第一次篩選中剔除。最後，再加上（3）筆者另行採集的語料，以上述三者統合而來之語料庫作為本研究之分析對象。

本研究「在」之語料總數為 2128 筆，語料來源以鄧守信教授的中介語料庫為主，所用語料出自此語料庫的筆數計 1788 筆，佔語料庫所有語料的 84.02%，表 3-2 之內容即為本研究語料庫中各來源資料之筆數，以及各來源資料在語料庫所有資料中所佔的比例。

表 3-2 各來源語料數及其在語料庫中所佔之比例

學習者語料庫	筆數	所佔比例
資料來源 (1)	1788	84.02%
(2)	96	4.51%
(3)	244	11.47%
總計	2128	100%

註：關於資料來源編號(1)、(2)、(3)之內容，請參見上頁。

至於語料的內容，本研究所使用之語料絕大部分是書面語料，書面語料佔所有語料的 98.50%，蒐集自教師命題作文、作業、習作本與測驗卷等；口語語料為課後訪談紀錄所得，僅有 1.50%。書面與口語語料在語料庫及各語料來源中的分佈詳見表 3-3。

表 3-3 書面及口語語料數及其在語料庫中所佔之比例

學習者語料庫	書面(筆數/所佔百分比)	口語(筆數/所佔百分比)
資料來源 (1)	1788/84.02%	0/0.00%
(2)	64/3.01%	32/1.50%
(3)	244/11.47%	0/0.00%
總計	2096/98.50%	32/1.50%

註：關於資料來源編號(1)、(2)、(3)之內容，請參見上頁。

以上語料來源學生背景為英語為母語之人士，學生華語能力涵蓋了初、中、高三種等級，學習者學習時間最短為一個月，最長為五年。

(二)、語料儲存/管理、語料標注方式

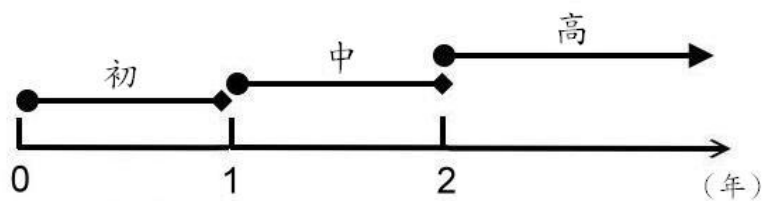
本研究所建語料庫沿用語料來源(1)及(2)的處理方式，以 Microsoft Office Excel 作為儲存與管理之工具，所蒐集之語料均依句鍵入。語料庫所記錄之格式項目包含語料來源、學習月數、程度等級、類別(在_x)、正誤、偏誤類型、語義分類、結構分類、例句。

以下再進一步說明上述幾個項目之內容：

一、語料來源：來自國立台灣師範大學華語文教學研究所鄧守信教授

專案研究建置之中介語料庫的標 1；來自國立台灣師範大學華語文教學研究所曾金金教授於 94 及 95 學年度第二語言習得課中蒐集的語料，標注 2；若為筆者另行蒐集之語料則標 3。

二、程度等級：筆者依學習者學習華語時間長短，將語料分為三級(圖 3-2)：初級為學習時間為一年以下，不含零程度，語料庫標記為 A；中級為學習時間為一年到兩年，不含一年，語料庫標記為 B；高級為學習時間為兩年以上，不含兩年，語料庫標記為 C。



● 表不含此數值 ◆ 表含此數值

三、類別 (在_x)：以本文的分類為依據，也就是將「在」分為「在₁」、「在₂」、「在₃」。

四、正誤：各類「在」之正確與否，正確標 O，錯誤標 X。

五、偏誤類型：標注語料之偏誤類項，詳細的類別請見本文第五章之討論。

六、語義分類：語義分類包含存在、取決於、具體處所、抽象處所、時間、範圍、內容媒介、條件等項。

七、結構分類：結構分類適用於「在₂」，包含以下四類：

- (1) 句首，如：**在中國**，使用超音波是不合法的可是人們還用。
- (2) 動前，如：我天天**在那裡**吃飯…。
- (3) 動後，如：一個人時可以站**在商店前**慢慢欣賞…。
- (4) 定語，如：中國經濟的崛起與它軍力的擴張以及**在國際舞台上的**政治影響力的增加存在著不可分隔的關係…。

語料紀錄格式如下表 3-4 (頁 38) 所示。

三、第二語言教材

市面上，針對特定背景的學習者所設計之華語教材尚不普遍。一般語言中心也往往採用同一本教材來教授不同國籍、不同母語背景的學生。然而，對不同母語背景的學習者而言，同樣的教學內容在難易程度上是有所差異的。一套理想的、針對性的教材不管是在內容的選定或編排順序上必須因人制宜，而不只是採用學習者母語作為說明語言。

本研究為求了解現行教材對於英語為母語的學習者之適用度為何，因此，除了建構出「在」的教學語法之外，還進一步以研究成果檢驗教材中「在」之語法點解釋與編排方式是否切合學習者的需求。本研究所檢驗的對象包含目前台灣及美國地區最為廣泛使用的華語教材，分別是《實用視聽華語》、《遠東生活華語》、《實用漢語課本》及《中文聽說讀寫》四套。

表 3-4 語料庫記錄格式

語料來源	來台學習月數	程度等級	類別 (在 x)	正誤	偏誤類型	語義分類	結構分類	例句
2	18	B	3	O		時間(進行)	NA	她在念第二個碩士。
3	24	B	2	O		時間	句首	在九十年代，網路公司出現了。
1	4	A	1	X	誤加	具體處所	NA	一九九六我在了法國。
1	2	A	2	X	錯序	具體處所	動前	我學在國語中心。
1	2	A	1	O		存在	NA	要是媽媽不在，我管不了我的兩個妹妹。
1	3	A	3	X	誤用「在」	時間(進行)	NA	我的同學在不好看，胖跟笨。
1	12	A	2	O		時間	動前	我想在畢業以後先去流浪一段時間，
1	8	A	2	X	缺漏「在」	範圍	動前	因為這個博物館世界上非常有名，所以除了台灣來的參觀的人也有很多外國來的人要參觀這個博物館。
1	8	A	2	O		範圍	動前	中國的興起還不只於經濟而已，它在軍事、政治、等各方面的成就層出不窮，使得它能夠在國際舞台上拌演主導角色