

## 以字頻表為基礎的對外華語初級教材 字彙通用性檢視

黃雅英

初級教材是語文學習的基石，然而兩岸用語的差異卻成為華語學習者的負擔，雖然針對字頻、字彙定量定性、覆現率的學術研究不少，但以字頻研究成果為編定基礎的分級華語教科書卻相當缺乏，本研究以香港地區整理的兩岸三地《現代漢語常用字頻統計》來檢視目前初級華語教材在兩岸三地的通用性，以供教師教學設計以及初級華語教材編輯的參考。最後針對兩岸字彙差異的現象提出教材編寫的建議，可以兩地常用字的頻次為考量，集中編寫初級華語教材的字彙，因文化環境造成的字彙差異，則另外編寫成為該地的文化補充教材或在教材中另作補充說明，讓學習者與教學者可以因地制宜，選擇適用教材，以提升初級教材的功能性與實用性，注重通用性字彙的編採，促進兩岸三地教材資源的流通，避免臺灣華語文教材被邊緣化的危機，期待為臺灣華語文教材競爭力的提升另闢進路。

關鍵詞：字頻、華語文、初級教材、字彙通用性

收件：2008年7月18日；修改：2009年3月17日；接受：2009年4月13日

# Word-Frequency-Based Vocabulary Generalization for Mandarin-as-a-Foreign-Language Texts for Beginners

Ya-Ying Huang

Beginner textbooks play an important role in learning Chinese. However many foreign learners feel confused and bewildered because of cultural and geographical differences. Although in the past lots of researchers contributed towards word-frequency and vocabulary, in the market few textbooks are compiled based on word-frequency. Using the statistics of modern Chinese word-frequency this study aims to find a common denominator in beginner Mandarin textbooks in China, Hong-Kong and Taiwan, and further provide a reference for instructional design and editing materials. In addition, this study comments on differences of vocabulary between China and Taiwan: the beginner textbooks could bring together the most common vocabulary in China and Taiwan, and the differences in vocabulary caused by environment and culture could become supplementary teaching materials. That way learners and teachers can choose suitable textbooks. This method can improve the function and practicality of beginner textbooks. Finally we expect to lift Mandarin textbooks in Taiwan out of their increasingly marginalized role.

Keywords: word frequency, Mandarin as a foreign language, beginner textbooks, vocabulary generalization

Received: July 18, 2008; Revised: March 17, 2009; Accepted: April 13, 2009

## 壹、前言

目前臺灣的對外華語文教材固然不少，但是由於編製單位不同，學習者各自環境背景的差異，仍然無法滿足大家的需求，加上中國大陸漢語教材的蓬勃發展，臺灣對外華語文教材須更具有競爭力。但是，兩岸語用的差異，往往成為華語學習者的負擔，甚至造成日後學習銜接上的漏洞，最令海外華語文教學者感到困擾的，也是兩岸教材的歧異，因此學習者在兩岸都要去的情形下，他們所要學的東西就很多，必須包括兩岸語言、字彙、語法等內。所以教材的編撰，如何能在兩岸之間取得平衡點、交集點，這是今後華語文教學必須審慎研究的問題(李璫，1994：124)。

在華語文教材中，初級教材是語文學習的基石，可說是影響最廣泛的階段，目前初級教材在字頻、字彙定量定性、覆現率的研究已不在少數，但研究成果卻不能說相當豐厚，尤其是漢字教材的編寫和研究仍相當薄弱(李泉，2007：217)，以字頻為編定基礎的分級教科書更是缺乏，導致課文生字選用的科學性不足，學習者也反應生字太多、太難、不實用等問題，於是要在繁忙的現代社會，提供學習者一個學習價值高、又極富溝通交際功能的定量、定性教材，似乎就得考量學習者日後兩岸三地往返的交際環境與學習歷程，審慎選擇適宜的工具字表作為教材編寫的參考。

隨著國際交流的日趨緊密，跨國家、跨地區的常用字趨同性很高，大華語圈的共同使用字比例可說是非常高(王惠，2006：254)，此外，華語文化圈出現的文言化、地方化、方言化，也使漢語的全球化面對很大的困難，從語言學習的立場出發，應該更重視華語文的口語化、大眾化(周清海，2008：5)，胡翔(2005：264)也指出：兩岸在詞語方面存在的差異雖然豐富了現代漢語的字彙量，但是卻在對外漢語教學上造成

極大的困擾。面對全球在地化，在地全球化的趨勢，如何兼顧華語文教育的共性與個性，發展廣為流通，卻又不失地方個性的教材，更是迫切需要的。

因此本研究以香港地區整理的兩岸三地《現代漢語常用字頻統計》中的〈基本字頻統計數表〉，選取足以涵蓋基礎日常用語的兩岸三地常用字，來檢視目前市面上的初級華語教材在兩岸三地字彙選擇的通用性，是否足以提供學習者來往兩岸三地的基本日常用語，並以〈跨地區比較數表〉來檢視兩岸三地通用字間的頻差關係，與地方性字彙的措置方式，以供學習者更有效率的學習以及多一面向的選擇、提供教師在教學應用的效率，以及語文中心分級的便利性，並提供另一面向的初級華語教材字彙編輯的初步規範參考。最後針對兩岸三地字彙差異的現象提出教材編寫建議，以各地常用字的字頻為考量，集中編寫初級華語教材的字彙。而因文化環境造成的字彙差異，則另外編寫成為該地的文化補充教材，或在教材中另作補充說明，讓學習者與教學者可以因地制宜，選擇適用教材，提升初級教材的功能性與實用性，並促進兩岸教材資源的流通，以期建立有系統的多元教材發展。

## 貳、對外華語初級教材字彙編排理念與實踐

目前就對外華語初級教材字彙方面的討論及相關內涵，可就教材的理論規範以及實證研究兩方面來探討。

### 一、理論規範研究

對外華語初級教材是華語學習的重要起點，而生字的選擇與配置，更是對教科書編輯者的一大考驗，研究者先回顧目前對外華語教材生字編排上的問題，再就目前對外華語教材生字編排的原則，來探討對外華語初級教材字彙編輯的規範：

### （一）字彙編輯的缺失

李泉（2005：184-185）綜合國外學者黎天睦、李曉亮、佟秉正、白樂桑、徐家禎的說法，從國外使用者的角度來探討漢語教科書目前的缺失，研究者歸納其中有關初級教材字彙編輯方面出現的問題為3點，首先，教材中多以詞為單位，單字的解釋並沒有受到重視，容易影響學習者的閱讀與學習興趣；其次，沒有注意到中國文字的特殊性，並針對中國文字特性去編輯教材；最後，不正確的處理中國文字和語言所特有的關係，這些都是字彙編輯上亟需修正檢討的地方。另外，初級者從拼音過渡到漢字的過程也太過短促，這些都是影響教材在海外流通與否的重要因素。

趙金銘（2005：164-169）也提到現行對外漢語教材的缺失，學習者意見普遍集中在以下兩者：一是教材內容沒意思；二是字彙太多。有關教材中的生詞問題，根據高彥德等人的統計，「生詞多、字彙不好（指偏難或不實用）」是現行教材的一大問題，如《速成漢語初級教程》字彙量大，覆蓋了全部甲級詞和大部分乙級詞，共出生詞 3,667 個，卻因詞的數量多，且重複率低，一些詞實用價值不大，使得字彙教學出現「夾生飯」現象，讓學生感到負擔過重。有人認為如果一篇 1,500 字左右的文章，其理想的生詞量應為 40 個，理想的語言難點和內容難點不應超過 10 個（周雪林，1996）。

### （二）字彙編排的原則

字彙編排的一般性原則，不脫由簡而難、由常見到不常見的範疇，趙金銘（2005：121）則進一步指出，當我們把漢字作為一個獨立的系統來進行研究和教學時，要認真考慮教材中應先教哪些漢字？哪些漢字需要重複出現？初級階段教多少漢字？認識多少漢字可以閱讀書報？研究者整理各專家學者的說法討論如下：

Teng（1998：51）明確指出字彙排序的基本原則，是以使用頻率越高、結構語意越簡單、越淺顯易懂、越不需要牽涉其他成分的字彙，為

優先選擇。

黃博儀、孫復瑗、潘羿均及巫宜靜（2008：93）等人在近義詞的研究中也指出，教材編輯應該考量字彙的使用頻率、場合、學習負擔，以及表達需求。在使用頻率上，應同時考量口語和書面語的語料，以高頻者優先。語料庫中的口語語料和書面語料比例，也應審慎檢驗。

傅國忠（2008：307）在華語字彙教學排序法的研究中，指出華語字彙教學排序的9點要素，分別是：1.字彙內部結構的複雜度；2.語意複雜度；3.用法複雜度；4.跨語言差距，也就是母語與目的語間的對應關係；5.使用頻率；6.語素構詞力；7.字彙的文化意涵；8.漢字難度；9.兩岸差異。並認為這9項要素之中，只有使用頻率和語素構詞力兩項與字彙教學排序有直接相關，其餘7項要素皆與學習難度相關，但與字彙教學不一定存在對比關係。

綜合上述，目前對外華語教材字彙編排的原則探討，就實用功能方面，包含使用頻率、兩岸差異、表達需求、文化意涵、語素構詞力等原則；就學習效益方面則須考量結構語意複雜度、漢字筆劃部件困難度、用法複雜度、跨語言差距等面向。但對於華語圈內部所存在的共性與個性的排序問題，卻沒有更進一步的檢討與正視，若就教材本身的實用價值來看，兩岸三地共通的常用字，無論在出現率或流通率上都有一定的份量，這就學習者的學習效益來說，通用性高的常用字比起較於偏狹的地方性字彙，更能擴展溝通交際的範圍以及實用的目的，因此研究者建議初級教材的字彙編寫原則，可以考量兩岸三地字彙的通用性效益作為字彙排序時的參考之一。

## 二、基礎實證研究

新一代教材的改革，需要站在巨人的肩膀上，吸取教訓加以改進，李泉（2005：187）提到舊有教材之所以質量差、不實用的主要原因還是缺乏大綱的指導，因此在教材創新的理念下，更要以基礎性研究為前

提，結合教材大綱、偏誤分析等研究成果，以增加新一代教材的科學性與價值性，以下就目前的基礎研究討論如下：

### （一）字彙數量的探討

在目前已有的漢字統計表基礎上，部分學者已經針對初級教材的字彙量，試圖以量化的統計分析方式，歸納初級教材適用的生字量：

王碧霞、李寧、種國勝、徐葉青（1994）認為留學生在開始記憶漢字時，無論是心理還是視覺方面都是對漢字信息處理最為困難的時候，因此認為，第一學期的漢字量不宜過多，定為 500 字左右較為合適。

孟繁杰與李焱（2006：175）在初級階段對外漢語教材識字的定量與定性分析中發現，覆蓋漢語語料 80% 的 500 個左右的漢字以及覆蓋漢語語料 90% 的 1,000 個左右漢字，可以作為對外漢語教材初級階段一學年漢字輸入量的確定依據。其中覆蓋率為 80% 左右的 500-600 個的漢字可以作為第一學期的教材內容，其餘漢字留給第二學期學習。

史友為（2008：74）以對外漢語教學最低量字彙的研究來探討初級教材的生字量，他提到王還（1986）根據頻率統計的一般規律，認為使用度最高的最初 500-1,000 字彙，其覆蓋率也是最高的。因此許多字彙表都選擇 1,000 詞作為初級階段字彙，但對於最低字彙來說，這個數字仍然太高。大河內康憲（1980）曾試擬漢語入門必要基礎字彙，字彙量是 600 詞，許多日本中國語的教師也認為最好以 500 詞左右為核心字彙的基準。許多研究都顯示最低量字彙大約以 500-600 之間為宜。

綜合上述，在初級階段的總字彙量，應編入頻率較高、較核心的前 500-600 字，即鄧守信在《遠東漢字三千字典》中所言，學完前 500 個最高使用頻率字彙，即可有效提昇一般中文閱讀能力達 70%。

### （二）字彙分級的研究

對外華語漢字教學基礎研究的成果，是教材字彙排序上的重要指標，以下整理李泉（2007）於漢字教學研究提及的中國大陸漢語字彙分級情形，以及臺灣學者針對字彙分級的相關研究，按年代先後分別討論

如下：

1. 中國大陸的國家漢語水平考試委員會辦公室考試中心新制訂的(2001)《漢語水平詞彙與漢字等級大綱》(修訂本)共收漢字 2,905 個，其中，甲級字 800 個，乙級字 804 個，丙級字 590 個、丙級字附錄 11 個，丁級字 670 個、丁級字附錄 30 個。

2. 中國大陸的國家對外漢語教學領導小組辦公室編(2002)《高等學校外國留學生漢語言專業教學大綱》共收漢字 2,503 個，一年級 1,491 字，二年級 545 字，三、四年級共收 467 個漢字。分級的主要依據與《漢語水平詞彙與漢字等級大綱》相同，都是以頻率為分級標準。

3. 中國大陸的國家對外漢語教學領導小組辦公室編(2002)《高等學校外國留學生漢語教學大綱(長期進修)》共收漢字 2,605 個，初等階段 1,491 字，中等階段 700 字，高等階段 491 字。

4. 張莉萍(2002)所提出的詞彙分級表，共收 10,155 個辭彙，分為初、中、高級，參考葉德明(1995, 1997)針對臺灣華語文詞彙劃分的研究結果，以及中央研究院五百萬詞平衡語料庫中的高頻詞，以頻率為主要參考，輔以情境因素的考量。

5. 《華語八千詞》是張莉萍 2003-2004 年受國家科學委員會補助的研究計畫成果，共收錄詞彙 8,000 個，分初、中、高 3 級，選詞來源為中央研究院核心詞彙、通用詞彙、參考詞彙表、TOP 詞彙表以及 HSK 詞彙表，以相對頻率加權值的方式計算出每一詞的比重在一總值排序選取，一樣以詞頻為主要依據。

6. 張郁雯(2004)也進行一套華語詞彙的分級工作，共收錄 18,261 個詞彙，並將詞彙分為 5 級，是上述詞彙分級研究中最龐大的一個，分法同樣以詞頻為主要考量。

綜觀上述，目前所見之具規模的詞彙表，其分級排序的方式皆以詞頻為主，其中或增加漢字及情境的考量，但都屬陪襯性質，華語詞彙的特點並未在詞彙排序中佔有任何地位。顯現目前的教材生字排序原則，



若要以漢字形式特點作考量來強化漢字的學習效益，就勢必得捨棄以通用性、溝通交際為考量的教材編排方式，這中間的拉鋸，似乎仍停留在主觀的辯證中，而未有一套科學性的分析法來決定(傅國忠，2008：316)，研究者從學習者的學習目的以及課程運用的角度來看，若以漢字形式特色為編排要素的教材，所強調的是提供學習者根據漢字圖畫性、部件特色以方便記憶為主，較適合於強化漢字認讀記寫的課程中使用，其教材的形式類似坊間英文單字本以字根字首為字群的編排方式，目的是有效率的大量記憶字彙；而以字頻高低做為編排原則，則是強調溝通交際法的學習模式，較適合在初級階段的綜合課程實施，因此本研究站在對外華語初級綜合性教材的分析立場，採取目前較具共識的字頻排序原則做為綜合性教材生字的檢視依據。

從文獻的回顧可以發現，對外華語教材質量的不穩定與課程綱要的欠缺有很大的關係，就在美國的 AP 課程標準、歐洲的 CEFR 語言架構都將中文列入第二外語的範疇時(蔡雅薰，2008：21)，我們自身卻仍缺乏一個與國際接軌的課程綱要，大大影響教材在歐美地區的流通性，在外部支援不力的情形下，與其等待一個國際觀課程標準的完備，不如先就教材內部的檢核與提升做起，教材的競爭力，需要與國際接軌的外部指標，更需要能廣泛適用的內部標準，就目前對外華語初級教材字彙的問題來看，仍存在字彙量過多、過難、不實用等問題，字彙的排序規範，也還不完備，缺乏科學性的系統原則，在基礎研究部分，以字頻為考量的分級規準已經得到共識，關於字彙量的討論也逐漸聚焦成形。

因此，在缺乏一個與國際標準吻合的課程綱要情形下，研究者期望就現有的字頻研究成果來檢視教材中的字彙排序，並以 500 核心詞的觀念來檢視教材中字彙的試用程度，試圖以兩岸三地通用性的原則來解決字彙量過多、過難、不實用等問題，對新一代教材編輯進行內部的反思與檢討，期望另闢一個提升教材競爭力的進路。

## 參、對外華語初級教材生字分析與探討

本研究參考相關文獻並配合研究目的，採內容分析法，一方面探討對外華語文教材在字彙編輯上的現況；另一方面，探討對外華語文教材字彙編撰的通用性模式與效益。茲就研究設計與研究成果，分別敘述與討論。

### 一、研究設計

#### （一）研究方法

本研究以內容分析法為主要的研究方法，運用「量的分析」統計次數，接著再進一步分析內容背後所傳遞的深層意義，並試圖找尋出對外華語初級教材在生字通用性的編排模式。

#### （二）研究對象

根據各華語文中心的語文課程計畫，《新版實用視聽華語（1）》、《遠東生活華語 book（1）》均適用於初級華語課程，而且使用率相當高，因此本研究以 2007 年版的《新版實用視聽華語（1）》、《遠東生活華語 book（1）》為研究對象，各 12 課總共 24 課的生字為研究範疇，探討教材生字在兩岸三地的通用程度。

#### （三）研究工具

香港中文大學何秀煌教授所主持的《現代漢語常用字頻統計》研究計畫，提出了「跨地區」和「跨年代」的理念，在目前的字頻研究成果中，為較符合教材兩岸三地流通性檢驗標準的字頻表，因此以此表作為教材共性與個性的檢驗參考。而該計畫分別統計了兩岸三地各自在不同年代的「跨年代」字頻表，以及在同一時代背景下，兩岸三地間相互對照的「跨地區」字頻表，後者「跨地區」的概念較符合本論文的研究目的，因此，本研究不以各地「跨年代」的字頻表為比較基準，而是以同一時代背景下兩岸三地「跨地區」高頻字的研究結果，作為核心通用字

比對的基準，統計分析表的設計說明如下：

### 1.通用性生字分析表

以中國大陸、香港為基準，選取兩地累積 50%與 80%的常用字比對臺灣字頻累積 50%與 80%的用字做比較，整理出臺灣教材中容易遺漏的兩地共選字。以附錄 1 為例，研究者將中國大陸累計頻率 80%內共 617 字與臺灣累計 80%內的 619 字進行比對，這總共 1,200 多字當中，共有 70 字是不相同的，其他的 1,100 多字則都在中國大陸與臺灣各自的字頻統計中。因此節選出這 70 幾個字，以中國大陸字頻為序，對照該字在臺灣的頻次，以提醒教材編輯者，可能會忽略掉的中國大陸高頻字（詳見附錄 1 至 4）。

### 2.非中國大陸、香港常用字的臺灣高頻字一覽表

以臺灣為字頻基準，選取累積 50%的 153 個常用字，比對臺灣這 153 個常用字在中國大陸和香港字頻表上的編序，結果共有 27 字不在中國大陸和香港兩地字頻累計的前 50%，換而言之，這 27 個字在臺灣的字頻相當高，但在中國大陸和香港卻不見得是高頻常用字，因此將這些字列為臺灣地方性色彩較濃厚的生字，以檢視臺灣教材在這些用字的選字情形（詳見附錄 6）。

### 3.基本字頻統計數表

以字頻分級的概念檢驗初級教材生字的等級適用性，兩岸三地生字分級標準如表 1。

表 1 基本字頻統計數表參照之累計頻率分級一覽表

臺灣常用字累計頻率				中國大陸常用字累計頻率				香港常用字累計頻率			
153	619	1,086	2,834	156	617	1,081	2,795	158	608	1,048	2,704
體	留	胸	慄	問	副	描	溯	正	號	輸	叩
50%	80%	90%	99%	50%	80%	90%	99%	50%	80%	90%	99%

資料來源：整理自何秀煌（1998）。

## 二、研究成果與討論

### (一)「通用性生字分析表」方面

#### 1. 中國大陸及臺灣常用字頻差表

研究者以中國大陸常用字中累積頻率前 80%與累計前 50%的字為基準，將其與臺灣的字頻序號做比較，試圖整理出臺灣教材中可能忽略的中國大陸核心字群，結果共出現農、企、革、.....、副等（詳見附錄 1），共 70 字落點於臺灣累計頻率 80%的範圍外，而落點在臺灣累計頻率前 50%之外的中國大陸常用字有山、水、展、建、部、活、點、資、它、制、加、濟、氣、品、向、像、身、見等共 18 字（詳見附錄 2）。由於所檢驗的對象總字數介於 300 到 400 之間，所以研究者以中臺累計頻次前 50%的頻差表作為教材通用性檢視的參考，用以分析臺灣教材生字中所出現的中國大陸核心字，以推測該教材所選生字在中、臺兩地的通用程度。其中遠東版生字在前 50%累計表中的 18 個核心字就選錄了 6 個涵蓋率達 33.3%，視聽版則是 5 個涵蓋率為 27.7%。兩本教材在中臺通用性方面的效益差異不大，僅相差 5.6%。

#### 2. 香港及臺灣常用字頻差表

研究者以香港常用字中累積頻率前 80%與累計前 50%的字為基準，將其與臺灣的字頻序號做比較，試圖整理出臺灣教材中可能忽略的香港核心字群，結果共出現份、司、線、.....、號等，共 55 字落點於臺灣累計頻率 80%範圍外（詳見附錄 3），而落點在臺灣累計頻率前 50%之外的香港常用字有：港、香、九、表、點、加、便、見、五、至、頭、度、水、場、機、內、手、市、員、將、四等共 21 字（詳見附錄 4）。研究者以此表作為教材通用性檢視的參考之一，用以分析臺灣教材生字中所出現的香港核心字，以推測該教材所選字彙在香港、臺灣兩地的通用程度。其中遠東版生字在前 50%累計表中有出現的生字有 11 個，涵蓋率達 52.3%，視聽版則出現 8 個，涵蓋率為 38.1%。兩本教材在港臺通用性

的涵蓋率相差 14.2%，比起對中國大陸核心字的涵蓋率，兩者在香港核心字選用的效益差距較大。

比較遠東版和視聽版的中國大陸、香港核心字選用情形，發現在累計 50%的兩岸三地通用性字表中（詳見附錄 5），遠東版涵蓋較多中國大陸、香港地區核心字，共 39 字的中國大陸香港核心字就選了 17 字，涵蓋率將近五成；視聽版則涵蓋 13 字，整體而言較遠東版少了 4 個核心通用字。其中「水」、「點」、「加」、「見」4 個字的落點，都在中國大陸和香港累計頻率 50%以內，可見其重要性，然而此 4 字，在臺灣的累計頻率中，都落在 50%之後，倘若零起點的初級教材生字選擇範圍在 150 至 200 字之間，就容易遺漏掉這 4 個字頻相當高的重要生字。在遠東版和視聽版的檢核中發現，遠東版都有將此 4 字收錄其中，而視聽版第一冊共收 314 個生字，但還是漏掉了一個「加」字，稍嫌可惜。因此就兩本教材在通用性字表收錄字數的比較上，遠東版的涵蓋率是相對較高的。

## （二）「非中國大陸及香港常用字的臺灣高頻字一覽表」方面

此一覽表（見附錄 6）是以臺灣排序為基準選取常用字頻累計前 50%的 153 個字與中國大陸、香港兩地字頻進行比對，整理出座落於中國大陸、香港兩地累計字頻 50%以外的字，列為臺灣區域性考量較濃厚的字，以下以地方性字彙來簡稱，由表中可見，臺灣以累計頻率 50%前為考量所選用的高頻字，與中國大陸和香港比較，出現了：臺、燈、夢、懷、紀、永、亞、痛、韓、衛、聞、陸、園、藥、福、悲、追、善、庭、破、亮、繼、維、限、執、室、復等，共 27 個生字是落點於中國大陸和香港累積字頻前 50%之外的。這些區域性用字較濃厚的字，在考慮初級教材生字量與通用性的前提下，可斟酌學習效益而以補充或替代的形式來做安排，例如初級教材中臺灣的「臺」字的編排，就可以先從兩岸三地通用性較高、筆畫也較少的「台」字作為簡單的入門與基礎，並在後面括弧附註同「臺」的方式讓學習者理解兩字為通同字，較難的「臺」能

認讀辨識即可，對初級學習者而言，筆畫較少、部件單純的「台」不僅通用性高，也較方便書寫以進行更多元的字彙練習。

遠東版第一冊在第六課出現的「燈」字，相較於中國大陸與香港地區是屬於初級後段到中級階段的選字範圍，然而遠東版的分級方式是分成初、中、高三級，這裡的初級包含從入門到基礎的整個初級階段，因此以初級階段常用字 1,000 的規準來看，這樣的編排在選字的控制上還是合理的。而新版的視聽華語將原先的第一冊拆成兩冊，因此研究所選的新版視聽華語第一冊等同於初級前段的教科書，然而在「台」、「臺」的選用上，則是以「臺」字作為課文生字，並且出現在第一課，然而「臺」字在香港與中國大陸都是累計頻率 99% 之後的低頻字，等同於高級階段的生難字，雖然以「臺」（台）的補充方式呈現，但課文例句中的用字仍以「臺」字為主，對於起點階段的學習者而言，似乎會有學習負擔上的疑慮。

### （三）「基本字頻統計表」方面

遠東版第三課與視聽版第八課所出現的生字「噢」，在臺灣、中國大陸、香港的字頻中，分別為 3,410、3,206、3,321，都屬於累計頻率 99% 之後的低頻字，但卻出現在初級教材當中，此外，遠東版第十二課出現的生字「欸」在臺灣、中國大陸、香港的字頻中，分別為 4,443、0、4,260，都屬於累計頻率 99% 之後的低頻字，在中國大陸的字頻表中甚至沒有出現，在教材生字精簡的需求下，每個生字的選用應該要更有效益。

## 肆、結論

在缺乏課程綱要指導的情形下，本研究參照以跨地區、跨年代為重點的統計字頻表，就教材中字彙的通用性做一初步的檢視，試圖以兩岸三地通用性的原則來解決字彙量過多、過難、不實用等問題，最後針對教材字彙的安排，提供淺薄的建議以供參考。

## 一、初級華語教材以通用性為前提的常用字建議

初級華語教材在通用性前提下，不只要考慮自身環境適用性，更要考量日後學習效益和價值，因此，同屬兩岸三地的高頻字應列入教材選字的考量範圍內。尤其在通同字的選擇上，可以考量兩岸三地的用字頻率，以較常見的生字為主。

## 二、初級華語教材以通用性為前提的區域性生字配置建議

以各地常用字的字頻為考量，集中編寫初級華語教材的字彙，而因文化環境造成的字彙差異，除了可以另外編寫為文化補充教材外，也可以考慮將兩岸三地中的用語差異，放在教材附錄中比對參考。

## 三、對未來研究建議

本研究僅就字頻的面向進行探討，並試圖找尋有效結合知識結構與實用功能的字彙編排方式，日後研究還可以就教材的外部編排方式、內部結構、選字的文化性、構詞的強弱性等來探討。

華語文初級教材的編制是一份艱鉅且重要的工作，如何兼顧各項能力的訓練並且在結構——功能——文化的基準上，進行字彙的安排與選擇，這之間的拉鋸與完備，還需要教材編輯者與教學者持續的關注與努力。

## 參考文獻

- 王惠（2006）。新加坡華文教育的焦點之一：漢字——基於中小學教材語料庫的計量分析。《漢語教學學刊》，2，242-255。
- 王碧霞、李寧、種國勝、徐葉青（1994）。從留學生記憶漢字的心理過程探討基礎階段漢字教學。《語言教學與研究》，3，21-33。

- 王還 (1986)。漢語詞匯的統計研究與詞典編纂。辭書研究, 1986(4), 106-108。
- 史友為 (2008)。對外漢語教學最低量基礎詞彙試探。語言教學與研究, 1, 73-81。
- 何秀煌 (1998)。「現代漢語的邏輯研究」計畫之現代漢語常用字頻率統計。2008年6月7日, 取自 <http://humanum.arts.cuhk.edu.hk/Lexis/chifreq/>
- 李泉 (2005)。對外漢語教學理論思考。北京市: 教育科學。
- 李泉 (2007)。對外漢語漢字教學綜觀。載於崔希亮 (主編), 漢語教學: 海內外的互動與互補 (頁 197-221)。北京市: 商務。
- 李鑒 (1994)。海外華語文教育工作之展望。載於黃沛榮 (主編), 當前語文問題論集 (頁 103-127)。臺北市: 國立臺灣大學中國文學系。
- 周清海 (2008, 3月)。亞太區華語文與華語文教育的發展。載於中原大學應用華語文學系、世界華語文教育學會聯合舉辦之「2008 亞洲太平洋地區華語文教學與發展」國際學術研討會論文集 (頁 i-x), 臺北市。
- 孟繁杰、李焱 (2006)。初級階段對外漢語教材漢字的定量與定性分析。載於世界華語文教育學會舉辦之「第八屆世界華語文教學」研討會會前論文集第四冊教學研究 (2) (頁 175-182), 臺北市。
- 周雪林 (1996)。淺談外語教材評估標準。外語界, 62, 60-63。
- 胡翔 (2005)。海峽兩岸詞語差異及其原因——基於《兩岸現代漢語常用詞典》的詞彙對比研究。載於趙金銘 (主編), 對外漢語教學的全方位探索——對外漢語研究學術討論會論文集 (頁 254-264)。北京市: 商務。
- 傅國忠 (2008, 3月)。華語詞彙教學排序法初探。載於中原大學應用華語文學系、世界華語文教育學會聯合舉辦之「2008 亞洲太平洋地區華語文教學與發展」國際學術研討會會議論文集 (頁 307-318), 臺北市。
- 黃博儀、孫復瑗、潘羿均、巫宜靜 (2008, 5月)。從近義詞詞組的語意分析看華語教材的詞彙安排。載於國立高雄師範大學華語文教學研究所舉辦之「2008 華語文教材編寫」國際研討會會議論文集 (頁 93-111), 高雄市。
- 張郁雯 (2004)。詞彙分級研究。載於柯華葳 (主編), 華語文能力測驗編製——研究與實務。臺北市: 遠流。
- 張莉萍 (2002)。華語文能力測驗理論與實務。臺北市: 師大書苑。
- 趙金銘 (2005)。漢語與對外漢語研究文錄。北京市: 外語教學與研究。
- 蔡雅薰 (2008, 5月)。以共同歐洲語言架構 CEFR 為基礎之華語文能力指標暨課程綱要規劃設計。載於國立高雄師範大學華語文教學研究所舉辦之「2008 華語文教材編寫」國際研討會會議論文集 (頁 21-25), 高雄市。
- Teng, Shou-Hsin. (1998). Sequencing of structures in pedagogical grammar. *Journal of Chinese Language Teachers Association*, 33(2), 41-52.



附錄 1 中國大陸及臺灣累計前 80%的常用字頻差表

編碼	單字	中國大陸			臺灣			頻差 (%)
		序號	頻次	頻率 (%)	序號	頻次	頻率 (%)	
1	農	162	810	0.12220%	735	168	0.02554%	-79.11%
2	企	237	631	0.09522	784	155	0.02356	-75.25
3	革	318	473	0.07137	661	196	0.02980	-58.25
4	藝	342	438	0.06609	641	203	0.03086	-53.31
5	省	345	435	0.06564	685	186	0.02828	-56.92
6	技	354	423	0.06383	759	162	0.02463	-61.42
7	石	360	415	0.06262	648	201	0.03056	-51.21
8	村	362	412	0.06217	738	166	0.02524	-59.41
9	勞	363	411	0.06202	838	139	0.02113	-65.93
10	河	380	396	0.05976	692	184	0.02797	-53.19
11	型	389	387	0.05840	791	153	0.02326	-60.17
12	階	391	387	0.05840	686	186	0.02828	-51.58
13	城	399	376	0.05674	672	191	0.02904	-48.82
14	營	411	367	0.05538	743	165	0.02508	-54.71
15	支	418	363	0.05478	666	194	0.02949	-46.16
16	料	419	363	0.05478	766	159	0.02417	-55.87
17	志	422	360	0.05432	956	113	0.01718	-68.38
18	隊	438	341	0.05146	1,016	102	0.01551	-69.87
19	益	442	339	0.05115	624	209	0.03177	-37.89
20	源	446	333	0.05025	767	159	0.02417	-51.90
21	投	454	323	0.04874	634	204	0.03101	-36.37
22	算	459	317	0.04783	690	185	0.02812	-41.21
23	職	461	316	0.04768	637	204	0.03101	-34.96
24	派	462	314	0.04738	684	186	0.02828	-40.32
25	評	472	305	0.04602	704	180	0.02736	-40.55
26	銀	497	288	0.04346	747	164	0.02493	-42.63
27	廠	505	280	0.04225	1,008	103	0.01566	-62.94
28	查	507	277	0.04180	629	207	0.03147	-24.72
29	份	508	276	0.04165	646	201	0.03056	-26.63

附錄 1 中國大陸及臺灣累計前 80%的常用字頻差表 (續)

編碼	單字	中國大陸			臺灣			頻差 (%)
		序號	頻次	頻率 (%)	序號	頻次	頻率 (%)	
30	幹	510	275	0.04150%	1,337	61	0.00927%	-77.65%
31	養	512	274	0.04135	772	158	0.02402	-41.91
32	客	516	270	0.04074	622	210	0.03192	-21.64
33	劃	520	268	0.04044	951	114	0.01733	-57.15
34	積	522	268	0.04044	725	170	0.02584	-36.10
35	律	525	266	0.04014	751	163	0.02478	-38.27
36	站	526	266	0.04014	656	199	0.03025	-24.63
37	款	528	265	0.03999	1,216	75	0.01140	-71.49
38	萱	529	265	0.03999	0	0	0	-99.99
39	爾	534	260	0.03923	645	202	0.03071	-21.73
40	供	540	257	0.03878	674	190	0.02888	-25.52
41	毛	542	257	0.03878	1,055	95	0.01444	-62.76
42	批	546	252	0.03803	925	119	0.01809	-52.43
43	織	547	250	0.03772	847	138	0.02098	-44.39
44	緊	554	246	0.03712	654	200	0.03040	-18.10
45	護	555	245	0.03697	714	175	0.02660	-28.04
46	臉	558	243	0.03667	746	164	0.02493	-32.01
47	酒	560	243	0.03667	818	146	0.02219	-39.47
48	克	567	237	0.03576	691	184	0.02797	-21.79
49	模	568	237	0.03576	830	141	0.02143	-40.06
50	靈	569	237	0.03576	633	205	0.03116	-12.86
51	味	570	236	0.03561	711	177	0.02691	-24.44
52	唱	580	232	0.03501	803	149	0.02265	-35.30
53	承	581	232	0.03501	635	204	0.03101	-11.42
54	裝	582	232	0.03501	768	159	0.02417	-30.96
55	範	583	231	0.03486	861	133	0.02022	-42.00
56	葉	584	230	0.03471	764	160	0.02432	-29.92
57	飛	589	228	0.03440	651	201	0.03056	-11.19
58	館	590	228	0.03440	723	171	0.02600	-24.44

附錄 1 中國大陸及臺灣累計前 80%的常用字頻差表 (續)

編碼	單字	中國大陸			臺灣			頻差 (%)
		序號	頻次	頻率 (%)	序號	頻次	頻率 (%)	
59	剛	593	226	0.03410%	827	141	0.02143%	-37.15%
60	江	594	225	0.03395	852	136	0.02067	-39.11
61	誰	595	225	0.03395	719	172	0.02615	-22.99
62	哪	597	224	0.03380	1,439	52	0.00791	-76.61
63	篇	599	224	0.03380	681	188	0.02858	-15.45
64	股	602	222	0.03350	1,065	94	0.01429	-57.34
65	塊	605	218	0.03290	856	134	0.02037	-38.08
66	曲	609	217	0.03274	679	188	0.02858	-12.72
67	率	610	217	0.03274	653	200	0.03040	-7.15
68	舞	611	217	0.03274	875	130	0.01976	-39.65
69	秋	615	216	0.03259	730	169	0.02569	-21.18
70	副	617	215	0.03244	933	118	0.01794	-44.71

註 1：編碼 38 之「萱」字，在現代漢語常用字頻率統計表中的臺灣區累計次數為 0，因此序號、頻次皆標示為「0」。

註 2：此表節選中國大陸累計頻率前 80%的 617 個常用字，並一一比對相對應的臺灣字頻，結果共出現 70 字落於臺灣累計頻率前 80%之外，也就是初級教材字彙編輯總生字量為 500-600 字時，可能會因為頻率序號較後面而忽略的中國大陸常用字，研究者加以整理成此表，作為臺灣初級教材字彙編排時的提醒與參考。

資料來源：整理自何秀煌 (1998)。

附錄 2 中國大陸及臺灣累計前 50%的常用字頻差表

編碼	單字	中國大陸			臺灣			頻差 (%)
		序號	頻次	頻率 (%)	序號	頻次	頻率 (%)	
1	山	103	1,199	0.1809%	207	670	0.10190%	-43.70%
2	水	111	1,159	0.1749	188	724	0.11010	-37.07
3	展	114	1,138	0.1717	321	456	0.06932	-59.63
4	建	121	1,065	0.1607	252	584	0.08878	-44.76
5	部	122	1,050	0.1584	222	649	0.09866	-37.73
6	活	123	1,048	0.1581	156	847	0.12880	-18.58
7	點	131	1,009	0.1523	169	781	0.11870	-22.02
8	資	133	998	0.1506	356	407	0.06187	-58.92
9	它	136	989	0.1492	220	651	0.09896	-33.69
10	制	137	980	0.1479	298	489	0.07434	-49.73
11	加	142	945	0.1426	216	655	0.09957	-30.17
12	濟	146	928	0.1400	375	380	0.05777	-58.75
13	氣	150	893	0.1348	263	568	0.08635	-35.92
14	品	151	892	0.1346	309	480	0.07297	-45.79
15	向	152	890	0.1343	178	753	0.11450	-14.76
16	像	153	867	0.1308	183	740	0.11250	-14.01
17	身	154	867	0.1308	166	795	0.12090	-7.62
18	見	155	862	0.1301	187	729	0.11080	-14.80

註：此表節選中國大陸累計頻率前 50%的 156 個常用字，並一一比對相對應的臺灣字頻，結果共出現 18 字落於臺灣累計頻率前 50%之外，也就是初級教材字彙編輯時可能會因為頻率序號較後面而忽略的中國大陸常用字，研究者加以整理成此表，作為臺灣初級教材字彙編排時的提醒與參考。

資料來源：整理自何秀煌（1998）。

附錄 3 香港及臺灣累計前 80%的常用字頻差表

編碼	單字	香港			臺灣			頻差 (%)
		序號	頻次	頻率 (%)	序號	頻次	頻率 (%)	
1	份	333	464	0.06994%	646	201	0.03056%	-56.31%
2	司	340	447	0.06737	643	202	0.03071	-54.42
3	線	343	438	0.06602	763	161	0.02448	-62.93
4	支	363	408	0.0615	666	194	0.02949	-52.04
5	算	408	356	0.05366	690	185	0.02812	-47.59
6	買	416	353	0.05321	722	171	0.02600	-51.14
7	派	420	351	0.05290	684	186	0.02828	-46.55
8	克	421	350	0.05275	691	184	0.02797	-46.98
9	客	430	341	0.0514	622	210	0.03192	-37.89
10	投	432	340	0.05125	634	204	0.03101	-39.48
11	股	448	326	0.04914	1,065	94	0.01429	-70.92
12	皮	466	312	0.04703	762	161	0.02448	-47.95
13	溫	469	309	0.04657	693	183	0.02782	-40.27
14	銀	471	303	0.04567	747	164	0.02493	-45.41
15	羅	478	297	0.04477	892	126	0.01915	-57.21
16	衣	485	295	0.04446	713	175	0.02660	-40.17
17	請	489	293	0.04416	664	195	0.02964	-32.88
18	假	490	291	0.04386	627	207	0.03147	-28.26
19	跟	492	290	0.04371	639	204	0.03101	-29.05
20	壓	508	283	0.04265	733	168	0.02554	-40.13
21	誰	513	279	0.04205	719	172	0.02615	-37.82
22	顧	514	278	0.04190	705	180	0.02736	-34.70
23	巴	527	270	0.04070	1,052	95	0.01444	-64.51
24	料	528	270	0.04070	766	159	0.02417	-40.61
25	職	529	269	0.04054	637	204	0.03101	-23.51
26	裝	530	269	0.04054	768	159	0.02417	-40.38
27	團	532	268	0.04039	688	185	0.02812	-30.38
28	飛	539	264	0.03979	651	201	0.03056	-23.21
29	姐	540	263	0.03964	1,478	50	0.00760	-80.83

附錄 3 香港及臺灣累計前 80%的常用字頻差表 (續)

編碼	單字	香港			臺灣			頻差 (%)
		序號	頻次	頻率 (%)	序號	頻次	頻率 (%)	
30	討	547	259	0.03904%	695	182	0.02767%	-29.13%
31	劃	550	257	0.03874	951	114	0.01733	-55.26
32	票	553	257	0.03874	897	125	0.01900	-50.94
33	席	555	256	0.03859	891	126	0.01915	-50.36
34	藝	558	254	0.03828	641	203	0.03086	-19.39
35	署	567	251	0.03783	1,407	55	0.00836	-77.90
36	擔	568	250	0.03768	753	163	0.02478	-34.24
37	阿	570	250	0.03768	642	203	0.03086	-18.10
38	講	572	249	0.03753	667	194	0.02949	-21.42
39	怕	576	245	0.03693	628	207	0.03147	-14.78
40	供	579	243	0.03663	674	190	0.02888	-21.14
41	查	581	243	0.03663	629	207	0.03147	-14.08
42	益	582	242	0.03648	624	209	0.03177	-12.89
43	豐	583	241	0.03632	1,096	89	0.01353	-62.75
44	味	584	240	0.03617	711	177	0.02691	-25.62
45	京	591	237	0.03572	1,237	72	0.01095	-69.36
46	遊	593	235	0.03542	837	140	0.02128	-39.91
47	吧	594	234	0.03527	657	198	0.03010	-14.66
48	布	595	234	0.03527	736	167	0.02539	-28.02
49	獲	596	233	0.03512	676	190	0.02888	-17.75
50	站	597	233	0.03512	656	199	0.03025	-13.86
51	護	599	232	0.03497	714	175	0.02660	-23.92
52	店	603	228	0.03437	883	128	0.01946	-43.38
53	石	606	228	0.03437	648	201	0.03056	-11.08
54	膚	607	228	0.03437	2,000	26	0.00395	-88.50
55	號	608	228	0.03437	755	163	0.02478	-27.89

註：此表節選香港字頻累計頻率前 80%的 608 個常用字，並一一比對相對應的臺灣字頻，結果共出現 55 字落於臺灣累計頻率前 80%之外，也就是初級教材字彙編輯總生字量為 500-600 字時，可能會因為頻率序號較後面而忽略的香港常用字，研究者加以整理成此表，作為臺灣初級教材字彙編排時的提醒與參考。

資料來源：整理自何秀煌（1998）。

附錄 4 香港及臺灣累計前 50%的常用字頻差表

編碼	單字	香港			臺灣			頻差 (%)
		序號	頻次	頻率 (%)	序號	頻次	頻率 (%)	
1	港	52	1,840	0.2773%	610	217	0.03299%	-88.11%
2	香	117	1,136	0.1712	475	295	0.04485	-73.81
3	九	124	1,061	0.1599	213	661	0.10050	-37.17
4	表	126	1,053	0.1587	165	801	0.12180	-23.28
5	點	128	1,042	0.1571	169	781	0.11870	-24.40
6	加	130	998	0.1504	216	655	0.09957	-33.81
7	便	137	952	0.1435	234	628	0.09547	-33.47
8	見	140	941	0.1418	187	729	0.11080	-21.86
9	五	141	939	0.1415	198	709	0.10780	-23.85
10	至	143	932	0.1405	179	753	0.11450	-18.51
11	頭	144	932	0.1405	176	763	0.11600	-17.43
12	度	145	927	0.1397	175	764	0.11610	-16.88
13	水	147	921	0.1388	188	724	0.11010	-20.71
14	場	148	915	0.1379	247	599	0.09106	-33.97
15	機	150	907	0.1367	230	633	0.09623	-29.61
16	內	151	906	0.1366	172	772	0.11740	-14.06
17	手	152	902	0.1360	201	703	0.10690	-21.39
18	市	153	900	0.1357	269	547	0.08315	-38.70
19	員	154	897	0.1352	221	649	0.09866	-27.03
20	將	156	895	0.1349	167	784	0.11920	-11.65
21	四	157	893	0.1346	164	808	0.12280	-8.74

註：此表節選香港累計頻率前 50%的 158 個常用字，並一一比對相對應的臺灣字頻，結果共出現 21 字落於臺灣累計頻率前 50%之外，也就是初級教材字彙編輯時可能會因為頻率序號較後面而忽略的香港常用字，研究者加以整理成此表，作為臺灣初級教材字彙編排時的提醒與參考。

資料來源：整理自何秀煌（1998）。

附錄 5 「中國大陸、臺灣」及「香港、臺灣」累計前 50%的通用性生字  
統計表與教材字彙分析結果

中國大陸、臺灣累計前 50%常用字頻差表				香港、臺灣累計前 50%常用字頻差表			
編碼	單字	遠東生活華語 (出現課次)	實用視聽華語 (出現課次)	編碼	單字	遠東生活華語 (出現課次)	實用視聽華語 (出現課次)
1	山	九	四	1	港		
2	水	六	八	2	香		
3	展			3	九	一	四
4	建			4	表	三	
5	部			5	點	三、四、十一	十一
6	活			6	加	十	
7	點	三、四、十一	十一	7	便	二	六
8	資			8	見	四	二
9	它			9	五	一	四
10	制			10	至		
11	加	十		11	頭		
12	濟			12	度		
13	氣			13	水	六	八
14	品			14	場		
15	向			15	機		六
16	像		五、六	16	內		
17	身	十、十二		17	手	五	
18	見	四	二	18	市	十	
				19	員		
				20	將		
				21	四	一	四
共計		6/18	5/18	共計		11/21	8/21

資料來源：研究者自行整理。



附錄 6 非中國大陸及香港常用字的臺灣高頻字一覽表與教材字彙檢視結果

編碼	單字	臺灣 字頻序號	中國大陸 字頻序號	香港 字頻序號	出現課次（替換通同字）	
					遠東版	視聽版
1	臺	185	2,529	2,682		一（台）
2	燈	429	815	657	六	
3	夢	454	691	819		
4	懷	483	787	766		
5	紀	486	681	631		
6	永	495	908	745		
7	亞	499	935	644		
8	痛	508	743	727		
9	韓	514	1,645	807		
10	衛	515	1,313	997		
11	聞	531	860	687		
12	陸	542	784	617		
13	園	548	813	722		
14	藥	559	1,501	878		
15	福	566	803	618		
16	悲	572	924	1,162		
17	追	577	757	826		
18	善	579	654	694		
19	庭	585	817	776		
20	破	590	627	672		
21	亮	595	639	933		
22	繼	597	763	647		
23	維	601	632	676		
24	限	602	684	748		
25	執	615	840	832		
26	室	616	713	823		
27	復	618	731	955		

資料來源：研究者自行整理。