

國家教育研究院華語字表與其他字表比較研究*

黃淑齡 白明弘 吳鑑城 李詩敏¹ 林慶隆

國家教育研究院語文教育及編譯研究中心

摘要

分級字表包含識字量和字集內容，是評估閱讀及學習成效的指標。由於語言變遷，字表之分級內容會因時、因地、因對象而呈現差異。多年來，各地研究人員為母語者或二語學習者設計了不同字表來因應編寫教材或測驗評量的需求，臺灣目前並沒有共同採用的華語教學字表，國教院執行教育部「邁向華語文教育產業輸出大國八年計畫（102-109）」之子計畫「建置應用語料庫及標準體系」，運用大型語料庫完成符合二語學習者需求的分級字表，為評估此字表的特色，本文選擇具代表性之7份不同地區及適用對象的字表進行比較分析，研究顯示國教院字表在臺灣不同類型的語料庫，如圖書語料、新聞語料、口語語料等覆蓋率上表現最佳，尤其在初級，覆蓋率明顯優於歐洲字表、香港字表、新加坡字表及大陸漢辦字表，平均約高出4個百分點，顯示本字表適用於臺灣教學，也能回應學者「教學字表的制定應充分立足於本地」的建議。此外，從字表的字集內容來看，國教院字表最能代表臺灣當代華語，在收字及分級上均能呈現本地華語樣貌。也是唯一能清楚說明建置步驟、並貫徹以語料庫為本的字表。整體而言，本字表兼具了客觀性、再現性、真實性、字本位及符合二語學習者需求等特色及優勢。

關鍵詞：漢字分級 字表 標準體系 華語文 教學

* 感謝本期刊兩位匿名審查者提供寶貴建議，使本文更臻完善。文中如有疏漏，悉由作者負責。

¹ 本文通訊作者。

1. 前言

識字量是評估學童閱讀及學習成效的指標（王瓊珠等 2008），也是評估二語學習者書面言語活動能力（讀和寫）的依據和條件（白樂桑、張麗 2008）。因此，訂定分級識字量及字集內容是華語教學的重要工作之一。過去，為了教學的一致性及便利性，研究上多著重於字表和詞表的銜接及協調（中國國家漢辦 1992），然而，由於華語不屬於拼音文字系統，漢字的書寫特性需被突顯，不能僅以詞視之，故白樂桑提出「字本位」教學法（1996），帶動了編撰字表的風潮（林季苗 2011）。字表和詞表的主要差異在於功能不同，例如：因應漢字書寫困難的特性，字表在強調組字率、構詞率等等條件下，與重視溝通與使用情境的詞表呈現分級差異，也就是，出現在某一級詞表的單字詞，可能在字表上有不同的級別，例如「其」是構字率高的部件，在字表中屬於 A1 級，但它是偏書面的詞語，故在詞表上納入 B2 級。這也說明字表有獨立於詞表的特性，不能僅以拆詞為字的方式，完全藉由詞表來產出字表。目前，臺灣僅有華測會公布的華語八千詞，該表以詞彙為單位，而非以漢字為單位（張莉萍 2012）。

多年來，各地已訂定之母語學習字表或二語學習字表，包括中國《漢語水平詞彙與漢字等級大綱》（中國國家漢辦 1992）、歐洲漢語能力分級計畫的分級字表（Song 2012）、法國中學漢語學程公布之選修漢語中學生應學漢字字表（Ministry of National Education (France) 2003）、新加坡小學華文字表（新加坡教育部 2007），以及香港小學中文科常用字表（香港浸會大學 2003）等。不同字表除反映出當地的語言使用特色，建置方式也會影響到字表的客觀性及適用性，中國《漢語水平詞彙與漢字等級大綱》（1992）提出以字頻、構詞率及語意聯想（王敏、王曉明 2012）為主要考量，將漢字分為 4 級。歐洲漢語能力分級計畫 (European Benchmarking Chinese Language Project, EBCL) 則以需用性、基本性、常用性及高構詞性 (Allanic and Shu 2012) 等為原則，提出初等漢字（A1 及 A2）的分級方案。黃沛榮（2001）、葉德明（2000）、曾金金（2000）等則重視字的部件，指出整字的部首、能組成較多漢字的部首、與現代生活密切相關的部首，及構詞率高的整字部首有優先學習價值。根據上述研究，漢字分級考量多以字頻、覆蓋率、分布均勻度、構詞率、構字率等統計數值客觀呈現，因此，以語料庫語言學為本是分級字表最好的產出方式。

本文應用國家教育研究院（以下簡稱國教院）建置的華語文語料庫 (Corpus of contemporary Taiwanese Mandarin, COCT)²，以統計法為主，專家法為輔，訂定了四等七級的對外華語教學分級字表。³ 此字表結合語料庫語言學，以最能代表臺灣地區用語的教材語料（詳見 2.3 節）為主，口語語料及書面語語料⁴為輔，先利用統計法產生初步字表，再經專家調校完成定稿，調校幅度在 10%~30% 左右，計有包括能力指標、構字率、構詞率、語義聯想、常用詞群類聚、認寫難易度、專名移除或級別調整、功能詞級別調整及異體字整併等專家調校理由。字表產出後，在口語、圖書、新聞等不同類型的巨量語料庫中，均達到 99% 以上的覆蓋率，顯示此字表並不侷限於教材語料，是能代表臺灣當代語言使用情況的字表。

本文除介紹如何運用語料庫建置字表外，更重要的是藉由對比研究，將本字表與海內外不同的 7 份字表做比較，從教學字量與分級差異、語料庫覆蓋率與字集差異，及各字表的建置方式等三個面向，來說明國教院字表的 5 項特性：客觀性（以統計為主）、再現性（建置步驟透明）、真實性（呈現臺灣華語使用樣貌）、字本位（以認寫為定位），以及符合二語學習者需求（經專家調校）等特色。全文段落安排說明如下：第二節簡介國教院字表之建置流程及成果，第三節簡介對比分析之海內外字表、說明比較面向及分析結果，第四節提出結論。

2. 國教院字表的發展

2.1 分級定位

臺灣目前並無共同採用的分級字表，華測會僅提供詞表（張莉萍 2012），學者遂嘗試以拆詞為字的方式建構字表（張莉萍 2008），HSK 字表收字的原則之一是漢字的總量與分級需對應詞彙總量與分級（王敏、王曉明 2012），因此他們也以拆詞為字的方式建置字表（張莉萍 2008）。然而，越來越多學者認為華語的基本語法單元是「字」而非「詞」（白樂桑 1996；張朋朋 2005），白樂桑和張麗（2008）曾指出：「單個漢字內部所包含的知識和資訊不具有交際性。」因此主張漢字應以「書寫符號」（1996）之特性

² 國教院語料索引典系統網址為 <http://coct.naer.edu.tw/cqpweb/>。

³ 四等七級分別為 Pre-A、A、B、C 四等及 Pre-A, A1, A2, B1, B2, C1, C2 七級。

⁴ 書面語料包含圖書語料及中國時報、國語日報、聯合報等新聞語料。

單獨分級，石定果（1997）也認為漢字應與語音、詞彙、語法等共同組成漢語的語言能力，成為其中獨立的一環。也就是說，「詞」是以「溝通」為目的，根據生活情境的任務導向來進行分級（Council of Europe 2001; American Council on the Teaching of Foreign Languages 2012; Centre for Canadian language benchmarks 2012; 國家漢語國際推廣領導小組辦公室 2016）；而「字」則著重在書寫和識讀上，因為漢字做為書寫符號，拆解開的部件數達 441 個（資策會 1984），而且讀音和書寫關連性不大，對於學習者而言，沒有學過的字不太可能會讀。故唐力行、姚學吾（2001）以問卷調查美國 102 所中文學校的教師，79% 的教師回答「總是」和「經常」教學生認字和寫字，此結果也可以看出認、寫是漢字教學最重要的環節，因此本文的漢字分級定位即以「認寫」為主。而且，由於寫字又比認字難，所以有的字表又細分「被動認讀字」和「主動書寫字」，識字量多於寫字量（法國字表 2003；新加坡字表 2007）；其次，以「認字」為目標所建立的字表會以「日常生活上的常用字」為分級指標之一（黃沛榮 2001；Allanic 2005），這似乎偏向「詞本位」的考量了。事實上，學習與現代生活密切的字或部件，學習者可以因其常見而強化書寫記憶（張欣怡 2009），並不違背「字本位」的宗旨，唯有當生活用字難以書寫時，才會以「字本位」因素調校級別。例如「龜」在教材語料中統計出的級別屬 B1 級，因難寫而調至 C1 級；「喂」字在近 6 億字的新聞語料中，屬於序位 4,000 以上的低頻字，但卻是臺灣生活常用字，加上並不難寫，故調至 A1 級。整體而言，本文之字表分級因綜合考量認、寫之需求，先以字頻高低排出序位，再依書寫難度等專家調校指標來調整字的級數，詳細調校指標說明請見 2.3 節。

2.2 分級字量

根據國教院研究（2017），對母語學習者而言 3,000 個字是識字量指標的最低基準。法國學者 Allanic (2005) 指出漢字不是拼音文字，在認識一定數量的漢字之前，一般學習者連通俗易懂的文章也無法閱讀，因此，他透過分析網路週刊《華夏文摘》的漢字使用頻率，發現如果能認識最高頻的 1,300 至 1,500 字，即能懂得該雜誌的 93%~95% 的內容。法國教育部在 2003 年將中學以漢語為第一外國語的學習字表字數訂為 805 字。歐洲漢語能力分級計畫 (EBCL 2011) 將 A1 級訂定為聽說詞 (oral reception words) 590 個，其中有 320 字為書寫字 (written production characters)、A2 級聽說詞累計為 1,245 個，其中有 630 字為書寫字。新加坡教育部於 2007 年公布小學華文生字表共 2,749

字，也做了識讀和識寫的字量區分，因識寫字（1,078 字）均包含在識讀字中，所以新加坡小學華文生字表其實只有 1,670 個字。中國國家漢辦（1992）為二語學習者編定的字表，則將字集分為四等，分別為甲級 800 字、乙級 804 字、丙級 601 字、丁級 700 字。香港浸會大學（2003）編訂小學中文科常用字表，提出 3,000 字的字表。張莉萍（2008）分析漢語水平考試 (HSK) 大綱、臺灣華語文能力測驗 (TOP-Huayu) 及法國中文字表後，提出對應於 CEFR 的累積字量建議，其中 A2 級為 500 字、B1 級累計 900 字、B2 級累計 1,800 字、C1 級累計 3,000 字。以覆蓋率來看，Laufer (1989) 提出 95% 語料覆蓋率可以達到一般理解，Hu and Nation (2000) 提出 98% 語料覆蓋率可以達到輕鬆理解。國教院 2010 年語詞調查研究則指出 99% 語料覆蓋率為 2,613 字。本文分析以上各種字量建議，邀請 5 位專家學者⁵ 研討，討論後將字表初、中、高級漢字在各種語料（包括新聞語料、圖書語料、教材語料、口語語料等）上的覆蓋率設定為：75%、95% 及 99%，再依此微調張莉萍（2008）對應於 CEFR 的累積字量建議，調整為初級 800 字，中級累計 1,900 字，高級累計 3,100 字。最後，專家依教學經驗，再將各級累計字量細分為 Pre-A 級 200 字、A1 級 500 字、A2 級 800 字、B1 級 1,300 字、B2 級 1,900 字、C1 級 2,500 字、C2 級 3,100 字。

2.3 分級方法

本文以國教院建置的華語文應用語料庫 (Corpus of contemporary Taiwanese Mandarin, COCT) 為語料來源。研究方法以統計法為主，專家法為輔，分級流程如圖 1 所示，以下將簡介各項作法。⁶

⁵ 本文共有 5 位專家學者參與討論字表建置的理論及原則問題，包括語料選擇、字表定位、分級字量、統計指標及調校指標等，其研究領域分別為中文、教育、應用華語及華語教學。另有 5 位資深華語教師參與字表調校，詳見註 12。

⁶ 由於漢字分級需深入說明及討論的細項繁多，考量本文主題為字表成果的比較研究，此處僅簡介分級作法，另有專文深入介紹字表發展過程及成果應用。

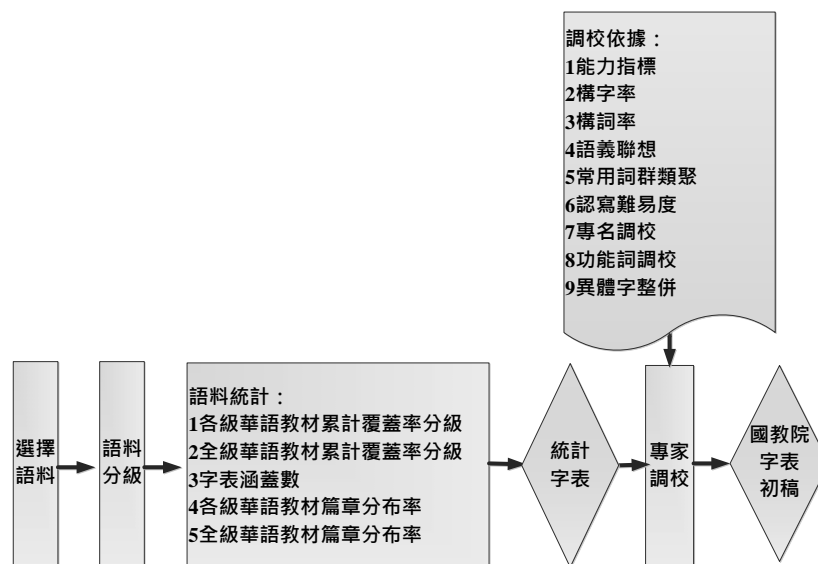


圖 1：國教院字表分級流程圖

(1) 選擇語料

首先針對華語文教材語料、口語語料、新聞語料及圖書語料⁷統計出不同字集，然後觀察字序差，例如：「買、飯、姐」這 3 個初級字，在教材語料的序位約在 35 上下，但在口語語料、新聞語料及圖書語料的序位分別約在 500、800 及 1,300 上下，考量實際教學需求後，乃決定採用教材語料作為統計依據。

(2) 語料分級

教材以臺灣教材為主，經調查，105 年臺灣 19 所華語中心使用之教材，初級及中級初、中等（以下簡稱初中級）多使用實用視聽華語、當代中文課程、遠東生活華語、遠東商務漢語及迷你廣播劇等教材；中級高等及高級以上（以下簡稱中高級）比較不一致，整體而言，包含較多新聞類、文化類及文言文教材。本文選用 34 套教材⁸，包含上述初中級所有教材，中高級部分

⁷ 本文所採用之語料庫規模（去除標點符號、數字、英文字），華語教材語料庫約為 64 萬字，口語語料庫約為 650 萬字，圖書語料庫（遠流出版社各類圖書）約為 1 億字，中國時報語料庫約為 10 億字，國語日報語料庫約為 550 萬字，聯合報語料庫約為 5 億 8,000 萬字。

⁸ 34 套教材為：《當代中文課程》第 1~3 冊，《實用視聽華語》第 1~5 冊，《遠東

則選用僑委會出版的文化教材，因為根據統計，發現僑委會的教材難度較高，和其他教材呈現難度互補，如表 1 所示。

表 1：教材難度分布統計

	僑委會教材以外教材的難度分布	僑委會教材的難度分布	總計
初級篇數	341 (35.0%)	82 (9.6%)	423
中級篇數	542 (56.0%)	523 (61.2%)	1065
高級篇數	81 (8.4%)	250 (29.2%)	331

表 1 顯示教材分級後，非僑委會教材難度偏低，僑委會教材的難度偏高。根據本文字表建置流程，教材自動分級後，較難的教材會被分到字表較高的級數（例如：中級或高級），可避免難字因字頻高而排到字表初級的可能性。依據對語料實際的觀察，難字本身多為低頻，需要比易字有更多的語料，才能取得有效的統計值，如果難度較高的教材不足，將在低頻字重要性之評估上失去準確度，因此本文最後斟酌採用了部分僑委會教材。

選取了 34 套教材後，為避免難度較高之教材影響初級字表的內容，本文先將教材分級，再依教材級別取字。本文利用師大文本可讀性分析系統 (CRIE) 進行教材語料自動分級 (宋曜廷等 2013)。為驗證自動分級的正確性，邀請 5 位專家針對部分教材作人工分級，實驗結果顯示，不管是 5 位專家之間的一致性，或 CRIE 與 5 位專家間的一致性均非常高，在 Weighted Kappa (Cohen 1968)⁹ 的計算中分別為 .71 及 .79，皆屬於「高度吻合」等級。

生活華語》1~3，《遠東商務漢語》，《實用中文讀寫》1~2，《迷你廣播劇》，《Advanced AP》，《幼童華語讀本》1~4，《快樂學華語》1~2，《Hello 華語》1~12，《青春華語》，《初中華文》1~6，《高中華文》1~6，《中文讀本》1~6，《中級華語》，《500 字學華語》，《一千字說華語》，《兒童華語課本》1~12，《華語》初 / 中 / 高級課本，《臺灣鄉土教材》第 1~4 冊，《兒童文化教材》，《小豆豆學華語》1~2，《安弟與小甜甜》，《華語閱讀課本》，《中國唐朝的詩》，《中國書法》，《中國民間的節日》，《中華民國的故事》(一)(二)，《僑教雙週刊一字的故事》，《僑教雙週刊—閱讀專刊》，《寓言故事》1~5，《成語故事》3001、5005，《民間故事》1~2，《海華文庫》。

⁹ Weighted Kappa 計算標記一致性，值落在 0~1 間，分為五組來表示不同等級的吻合度：.0~.20 極低的吻合度 (slight)、.21~.40 一般的吻合度 (fair)、.41~.60 中等的吻合度 (moderate)、.61~.80 高度的吻合度 (substantial) 和 .81~1 幾乎完全吻合 (almost perfect)。

34 套教材中共有 1,820 篇文章，統計出的字數如表 2，刪除初、中、高級重複出現的漢字後，教材語料共含 4,188 個漢字，作為本文分析的底本。

表 2：教材自動分級級數及字數

CEFR 等級	篇數	包含字數
初級	424	1,602
中級	1,065	3,491
高級	331	3,679

(3) 語料統計

接下來，本文以各級漢字累計覆蓋率（字頻）、全級漢字累計覆蓋率（字頻）、字表涵蓋數、各級篇章分布率及全級篇章分布率等 5 項統計法指標，統計出初步字表，以下分別說明。累計覆蓋率分級是採計字頻常見的作法，本文採用張郁雯（2003）所提出的累計頻率分級法，將 75%、85%、90%、95% 設立為字的累計頻率值，分別針對初級、中級、高級及全部教材語料庫中的字各自分成 1 到 5 級的區間。選取初級字集時，先看初級累計覆蓋率值，低者優先。兩字同一區間時，再參考全級累計覆蓋率值的區間，低者優先¹⁰。中級及高級字表作法相同，只是統計的語料來源不同。採計字表涵蓋數是為了兼顧質性與量化分析，本文參照了 7 份其他字表¹¹，標註每一字在各字表出現的次數，作為取字及分級的指標之一。採計篇章分布率是為了避免高頻字只集中在特定的篇章或領域中，而失去統計「常用性」的意義。本文提出頻率正規化及整合多面向均勻度的計算方法 (Bai et al. 2016)，同時考慮篇章、主題、時間、語式、媒體等不同面向的均勻度，能精確地預估字詞在不同類別中均勻分布的性質。以上各項數值的實際呈現請見附錄 1。

(4) 專家調校

根據統計指標排出字表序位後，依分級字量產出各級統計字表，再由 5

¹⁰ 每一級別取字及排序時，除了採計該級語料的累計覆蓋率及篇章分布率外，均會再採計全部語料（稱為全級）統計出的累計覆蓋率及篇章分布率，原因是一個級別的教材語料數量有限，無法細分字和字之間的使用次數差異。

¹¹ 包括華語八千詞表拆詞為字、常用國字標準字體表、師大全漢字檢索系統字表、大陸新 HSK 詞表拆詞為字、香港小學中文科常用字表、法國對外漢語字表等。

位資深華語教師¹²，根據能力指標、構字率、構詞率、語義聯想、常用詞群類聚、認寫難易度、專名移除或級別調整、功能詞級別調整及異體字整併等調校原則做進一步調整，以下分別說明。**能力指標**為國教院「建置應用語料庫及標準體系」計畫所訂定的四等七級華語能力描述。如：**Pre-A 級**「讀」「寫」中要求「能讀懂及書寫簡單的數量名稱」，因此將「一～十」等數字均調整為 **Pre-A 級**。**構字率**參考中研院《中文字基礎部件及部件屬性》、資策會五大字集 (Big5) 部件表、大陸《GB13000.1 部件規範》及《現代常用部件及部件名稱規範》等部件表，取交集數多之獨體字部件，並取資策會提供的構字頻率 .03 以上者，共收錄 140 個獨體部件。構字率的計算方式為「構字數 × 字頻」，例如：「鬼」可構成「鬼、魅、魂、魄、魃、魍、魎、魔……」等 x 個字，這些字出現的總頻率為 y，則構字率為 x 乘以 y。這個數值能代表獨體字部件的重要性，因為它不僅考量了一個部件能組成多少字，同時也考量了被組成字是不是常用。**構詞率**指單字組成複合詞的比率，如「的」很少和其他語素結合成詞，所以構詞率較低（6.33），「子」則常常和其他語素結合成詞，所以構詞率相對較高（169.17），本文採用張凱（1997）的計算公式，算式如下：「單字構詞數 + 雙字構詞數 / 2 + 三字構詞數 / 3 + 四字構詞數 / 4 + 五字構詞數 / 5 + 六字構詞數 / 6」。構字率、構詞率在調校過程中的實際呈現請見附錄 1。

語義聯想乃將語義相關的字以聯想方式群聚，如數目字一至十原本分屬 **Pre-A** 及 **A1** 等不同級次，最後以人工調校方式均挪移至 **Pre-A 級**。相同的，常用詞群類聚是指學了「左、右」後才會常用「邊」字，因此將「邊」由 **Pre-A** 移到 **A1**，與「左、右」同級，「旁」則由 **A2** 移到 **A1**，與「邊」同級。根據認寫難易度調整級別係指將「難以書寫」、「構詞後的詞較難」，及「構詞後罕用且構式及語義固定」的字以專家法調整位置，如「遙」字難寫，由 **A2 級** 移到 **B2 級**；「首」字多構成難詞，如「首都、首飾、匕首、斬首…」等，故由 **A1** 移至第 **B1 級**。**專名移除或級別調整**指將專名用法頻率偏高的字移除或後移，以維持字表的通用性，如：丁、欣由 **A1 級** 移至 **B1 級**。**功能詞級別調整**是因為這類字常是分布均勻的高頻字，但某些語法字不會太早學習，因此配合教學需求而調整其級別，如「把」字多數為介詞用法，由 **Pre-A 級**

¹² 5 位資深華語教師均為現任教師，分別任教於大學之華語文教學系所或華語教學中心，教學資歷 3 位 20 年以上，2 位 15 年以上，均有豐富之華語文教學經驗，經諮詢學者推薦參與本字表研發。

移至 A1 級；常用量詞「袋」則由 B1 級移至 A1 級，以配合教學需求。異體字整併是將異形同義且用法完全相同之漢字整併為一字，如：裡／裏、台／臺、群／等，以收納更多漢字。

由於專家調校指標共有上述 9 項，當兩個以上指標衝突時，專家以討論取得共識的方式決定級別，如：「木」字為基本構字部件，但構詞類型少；「機」由「木」「幾」組成，書寫難度較高，但構詞率很高，經討論後以識讀為優先，將「機」列為 Pre-A 級，「木」列為 A1 級。

(5) 專家調校信度

5 位專家調校除個別調校外，也召開 7 次會議討論級別不一致的漢字並取得共識。本文針對專家個別調校結果做了 Weighted Kappa 一致性統計¹³，詳見表 3。結果顯示專家調校的信度平均為 .82，達到「幾乎完全吻合」等級。

表 3：專家調校信度

	專家一	專家二	專家三	專家四	專家五
專家一	1.00	.72	.75	.73	.75
專家二	.72	1.00	.88	.86	.86
專家三	.75	.88	1.00	.89	.88
專家四	.73	.86	.89	1.00	.86
專家五	.75	.86	.88	.86	1.00

2.4 分級結果

本文提出對外華語教學字集共計四等七級，3,100 個漢字。其特色如下：

a. 客觀 (objective)：本字表全面經由語料庫統計產出，專家調校也以統計數據為參考，每一項調校均說明理由，力求科學、客觀並確保字集具備基本性及通用性。

b. 可再現 (reproducible)：過去，字表的字集內容是最常被探討的對象，字表的產生過程易被忽略，除了有礙經驗累積與交流，使用者也無法得知各地字集差異的原因，因此，再現性為本字表的建置目標之一。本字表產出所使用的語料庫及建置流程均如 2.3 節所述，除可供驗證外，每隔一段時間，

¹³ 凡於會議現場共同調校的漢字級別均未納入 Weighted Kappa 一致性統計。如字表完成後共同整體調校的漢字。

當代表新時代語言的語料產生時，即可利用相同研究方法快速建置新的字表。另外，異地（如不同地區字表）與不同對象（如母語者學習字表）字表，亦可依循相同程序建置。

c. 真實性 (authentic)：本字表以臺灣用字為主，所選用的教材語料為臺灣採用教材，專家調校也以臺灣用語為主要考量，例如：因應臺灣生活常見的捷運、珍珠奶茶等用語，將原本應屬於中級字的「珍」（口語語料序位 890，新聞語料序位 1,018）、「捷」（口語語料序位 1,984，新聞語料序位 1,032）等字納入初級。這兩個字在漢辦字表列為丙級，法國字表、EBCL 字表均未收。

d. 字本位 (Character-based)：本字表以「認寫能力」為定位，突顯字（結構為主）、詞（溝通為主）之不同特性（白樂桑 1996）；同時，為顯示漢字不同於拼音系統文字之特色 (Allanic et al. 2012)，本字表標註每一字在口語語料庫和書面語（包含圖書及新聞）語料庫的序位，藉由序位差，可觀察出聽說字和讀寫字的等級差異，例如：在 A1 字集中「喔、它、哇、啦、剛」等字在口語語料庫和書面語語料庫中每百萬字的字頻比是最高的，顯示它們傾向口語，相反的，「局、票、司、市、郵」字頻比的值較低，表示它們偏書面語。然而，由於這項資訊受限於口語語料是否完善，¹⁴ 目前僅能作為教材編撰者或評量出題者的參考，尚不能完全成為區分書面字與口語字的依據。

e. 符合二語學習者需求 (meet the needs of L2 learners)：本字表以二語學習者的需求為主，在語料庫的統計基礎上加入資深華語教師的經驗，進行字表調校。EBCL 曾提出常用性 (frequently-used)、需用性 (necessary) 及基本性 (essential) 等原則是對外華語教學中漢字分級的重要依據 (Allanic and Shu 2012)。一般而言，「常用性」以語料庫字頻來表現；學習者應先學習哪些字則代表「基本性」，例如由漢字書寫特性所衍生出的重要部件字、高構詞率的字等，都是學者認為最基本的字（黃沛榮 2001; 葉德明 2000; 曾金金 2000）；「需用性」則指出生活上需使用到的字（黃沛榮 2001; Allanic 2005），有些常用字不見得是最需用的字，例如功能詞「把、且」等等。本文藉語料庫統計出字的「常用性」，再由專家根據教學考量，搭配「需用性」及「基本性」，調校出完善的教學字表。例如，在初級字表中盡量將名詞性

¹⁴ 國教院 COCT 語料庫的口語語料多為電視節目字幕，部分為口語，但部分為準口語（偏書面）。

的字前移，如：「哥、妹、女」等由統計出的 A1 級移至 Pre-A 級；高頻的功能字後移，如：「把」由統計出的 Pre-A 級移至 A1 級、「且」由 A1 級移至 A2 級。惟統計出的「常用性」與專家認定的「基本性」及「需用性」應如何達成平衡，亦即字表的專家調幅應該多少才是適當的呢？表 4 顯示本文調校幅度¹⁵，各級約在 10% 至 30% 左右。本文曾比較臺灣常用教材¹⁶，發現各教材取字的差異性很大，故仍須一定幅度的調校，才能契合教學使用。

表 4：國教院字表的專家調校幅度

級別	調整字數 / 總字數	調校幅度 (%)
Pre-A	20/200	10
A1	70/300	23
A2	96/300	32
B1	145/500	29
B2	178/600	30
C1	170/600	28
C2	178/600	30

完整國教院分級字表請見附錄 2。

3. 國教院字表與其他字表的比較分析

3.1 其他字表簡介

本文選定 7 份官方或常見字表作為分析比較的對象。其中，二語學習者字表包括：法國教育部訂定之中學以漢語為第一外國語的學習字表、歐洲漢語能力分級計畫初級字表，及中國國家漢辦訂定之《漢語水平詞彙與漢字等級大綱》字表。另外，為分析本字表與母語學習者字表的差異，也納入國教院《2010 年常用語詞調查計畫》字表及中國國務院編撰之《通用規範漢字表》。最後，介於二語學習者與母語學習者間之華人地區學童字表亦為比較

¹⁵ 調校幅度係指依統計指標產出統計字表後，再經專家調整每一漢字級別的幅度。

¹⁶ 本文比較四部教材及其詞彙量，無交集者高達 6 成。教材分別是當代中文課程一、二、三冊 2,035 詞，新版視聽實用華語一、二、三冊 2,080 詞，遠東生活華語一、二冊 (A,B) 1,731 詞、華語教學基礎詞庫 1.0 版基礎核心、基礎情境 2,038 詞。

對象，包括：新加坡教育部公布之《小學華文生字表》、香港浸會大學編訂之《小學中文科常用字表》。¹⁷ 採納比較的 7 種字表字數及分級情況詳如表 5：

表 5：其他字表

地區 / 名稱	年份	字數	字體	級別
法國教育部字表 (以下簡稱法國字表)	2003	805	簡體	初級
歐洲漢語能力分級計畫字表 (以下簡稱歐洲字表)	2011	950	簡體	初級
中國漢辦《漢語水平詞彙與漢字等級大綱》 字表(以下簡稱漢辦字表)	1992	2,905	簡體	甲乙丙 丁四級
臺灣《2010 年常用語詞調查計畫》字表 (以下簡稱臺灣 2010 年常用語詞字表)	2017	4,961	正體	未分級
中國《通用規範漢字表》 (以下簡稱中國通用一級字表)	2013	3,500	簡體	未分級
香港《小學中文科常用字表》 (以下簡稱香港字表)	2003	3,000	繁體	一~六 年級
新加坡《小學華文生字表》 (以下簡稱新加坡字表)	2007	2,749	簡體	一~六 年級

上列字表簡述如下：

a. 法國字表：白樂桑（1996）提出漢語教學應以字本位為主，法國教育部（2003）的課程規劃也明定以漢字為教學重心，並清楚區分識字量與寫字量，識大於寫，如以華語為第一外語學習者，在高中一年級結束時其要求為被動認讀字 300 個，主動書寫字 505 個，¹⁸ 共 805 字。（張金蘭 2014）

b. 歐洲字表：歐洲漢語能力分級計畫 (European Benchmarking Chinese Language Project, EBCL) 是 2011 年由歐洲學者發起，中國學者及華語教師共同參與的漢語分級研究計畫，旨在依循 CEFR 架構，為歐洲的漢語使用與學

¹⁷ 因考慮年代及編撰目的，民國 71 年教育部公告之《常用國字標準字體表》不納入比較。

¹⁸ 能書寫表示能認讀，另列認讀字是指在此階段尚不需學會書寫的字。

習建立初級 (A1, A2) 指標說明。Allanic et al. (2012) 指出漢字收字原則有二：一為需用性 (necessary) 及基本性 (essential)，二為高頻使用 (frequently-used)，因為根據統計，頻率在前 500 名的漢字可覆蓋 74.8% 的詞彙 (Allanic et al. 2012)。EBCL 漢字分級的具體作法為：A1 級先以白樂桑的 255 字表為基礎，再依 CEFR A1 能力等級所需用詞來增補漢字，參考了法國中等學校初級字和中國漢辦建議的初學字表等，總計收 320 字。A2 級的作法類似，先採白樂桑的中學生初級主動書寫字 505 字為核心字表，再依其他字表增補，累計收 630 字。

c. 漢辦字表：中國 1992 年出版《漢語水平詞彙與漢字等級大綱》，定位為適用對外漢語教學領域的教學設計、教材編寫和成績測試，及作為少數民族漢語教學和中小學語文教學的參考，他們認為字表的分級與詞彙分級必須能銜接、協調，以求常用詞和常用字之間可建立一種科學的有機的聯繫。該書收字共分為四個級別。甲級字 800 個，乙級字 804 個，丙級字 601 個，丁級字 700 個，分別對應四個等級的詞彙表。根據王敏和王曉明（2012）分析，《漢語水平詞彙與漢字等級大綱》分級和收字的指導原則有三：一、漢字的總量與分級需對應詞彙總量與分級；二、分級主要參考《現代漢語常用字表》，即以字頻統計為主；三、分級時除了考慮字頻，可在一定範圍內進行必要的語義聯想添加。《漢語水平等級標準與語法等級大綱》（劉英林 1996）則進一步提到漢字、詞彙及語法分級都必需搭配漢語水平等級標準，但僅提到數量上的對應關係，對於漢字、詞彙及語法與能力等級如何對應則未加說明。

d. 臺灣 2010 年常用語詞字表：教育部於 1982 年出版《常用國字標準字體表》，內收 4,808 字，多年來一直是臺灣母語者的用字標準參照表。因年代久遠，國家教育研究院承接教育部計畫，進行 2010 年常用語詞調查，選擇 2010 年出版的書籍、新聞、雜誌、口語、教科書等共計 190 萬字的語料，分析出 4,961 字，是臺灣母語者用字的真實呈現。其中，前 2,613 字即覆蓋該語料庫 99% 語料。

e. 中國通用一級字表：自 2001 年起，中國教育部、國家語言文字工作委員會聯合研製漢字使用規範，該字表整合了 1955 年《第一批異體字整理表》、1986 年《簡化字總表》、1988 年《現代漢語常用字表》以及 1988 年《現代漢語通用字表》，並根據中國大陸用字現狀加以修補。於 2013 年正式頒布，成為社會一般應用領域的漢字規範，共收字 8,105 個，其中一級字表（常用字集）3,500 個，主要滿足基礎教育和文化普及的基本用字需要；二級字表

3,000 個，使用度次於一級字；三級字表 1,605 個，由姓氏人名、地名、科技術語和中小學教材文言文用字構成。

f. 香港字表：香港 1990 年編定的《小學常用字表》被列為中文教科書的編撰指南，但其中包含不少罕用字，亦非根據文字學原則編排，造成識字教育上應用的不便，故香港浸會大學語文中心於 2001 年開始進行小學常用字研究，把小學中文科常用字按級別排序，2003 年完成《小學中文科常用字表》3,000 字，主要供小學中文科老師、課程設計者和教材編撰者使用，為識字教學提供具備學術基礎的參考資料。

g. 新加坡字表：新加坡實施雙語教育政策，英語是主要的教學語言，華語作為第二語言只在華語課上使用。2002 年新加坡教育部曾頒訂的《中小學華文課程字表》，選錄 3,500 字，主要來自中國語委於 1988 年頒佈的《現代漢語通用字表》，但因新加坡與中國文化環境有異，採為底本的字表年代久遠，故於 2007 年重新修訂《小學華文生字表》，收錄 2,749 字，以考量字頻和口語為主要原則，以夠用、實用為目標（王惠 2010）。

3.2 比較面向

華語字表內涵隨著使用對象、目的、地域不同而有差異，例如字集內容、分級與否等區別；然也有共同的部分，例如語料覆蓋率高的字表在教學效率上亦較高，而建置方式若透明、可重建、而且是基於語料庫的字表，則較具客觀性。這些面向均是字表特色及優勢的表徵，因此以下分別針對教學字量與分級差異、語料覆蓋率與字集內容差異、建置方式等 3 個面向，進行國教院字表與其他字表的比較。

（1）教學字量與分級差異

除母語者字表未分級外，二語學習者字表或華人地區學童字表均分級，各字表分級字量如表 6：

表 6：各字表分級字量

字表名稱	總字數	分級		字數
國教院對外華語教學字表 (以下簡稱國教院字表)	3,100 (含異 體字共 3,131 字)	Pre-A		200
		A1		300
		A2		300
		B1		500
		B2		600
		C1		600
		C2		600
法國字表	805	初級	被動認讀字	300
			主動書寫字	505
歐洲字表	630	A1	書寫字	320
		A2	書寫字	310
漢辦字表	2,905	甲		800
		乙		804
		丙		601
		丁		700
臺灣 2010 年常用語詞字表	4,961	未分級		4,961
中國通用一級字表	3,500	未分級		3,500
香港字表	3,000	一年級		508
		二年級		716
		三年級		632
		四年級		558
		五年級		412
		六年級		174

表 6：各字表分級字量（續）

字表名稱	總字數	分級		字數
新加坡字表	2,749 (扣除 讀、寫 重複字 後為 1,670 字)	一上	識讀字 (含識寫字)	153
		一下	識讀字 (含識寫字)	157
		二上	識讀字 (含識寫字)	170
		二下	識讀字 (含識寫字)	155
		三上	識讀字 (含識寫字)	166
		三下	識讀字 (含識寫字)	135
		四上	識讀字 (含識寫字)	167
		四下	識讀字 (含識寫字)	153
		五上	識讀字 (含識寫字)	148
		五下	識讀字 (含識寫字)	120
		六上	識讀字 (含識寫字)	90
		六下	識讀字 (含識寫字)	56

以教學字量來看，專為二語學習者所設計的漢字表，初級約為 800 字，國教院字表（800）、法國字表（805）、漢辦字表（800）均在此範圍，中級以上因國教院參考 CEFR 指標分為 B、C 兩等，漢辦分為乙、丙、丁三等而難以比較。華人地區學童字表的分級字量則差異甚大，香港小學三年級累計教學字量為 1,856 字，新加坡僅約其一半 936 字。以總字量來看，新加坡小學生的習字量 1,670 字亦約為香港小學生 3,000 字的一半，可見香港的漢字教育較接近母語學習的要求，與臺灣 2010 年常用語詞字表 4,961 字，大陸 3,500 字等母語者字表較接近。

(2) 語料覆蓋率與字集內容差異

語料覆蓋率指字表能覆蓋的語料量比率，從整體字表來看覆蓋率，可以看出字表是否具有代表性；從分級字量來看覆蓋率，可以推估字表的習得效率。評估語料覆蓋率時應以相同的語料庫為基準，然而，因為字表的產出地各有不同，雖然都是漢字字集，也會有地域性和時間性的差別，本文比較的 7 份字表皆為當代字表，沒有時間性的問題，但區域的不同仍會造成語料覆蓋率的落差，也就是說，覆蓋率低可能是字表有偏頗，也可能是字表與統計所使用的語料庫有地域上的區別，所以用字較不相應，這是進行比較時必須考慮的因素。此外，語料覆蓋率也可以用來驗算字表的代表性，例如國教院字表採用教材語料作為統計依據，字表完成後，再利用口語語料、書面語語料庫等來驗證其是否能代表臺灣母語使用現況。最後，由於多數字表為簡體字集，必需先經簡繁轉換後才能在國教院語料庫中做覆蓋率統計，由於一個簡體字可對應不止一個正體字，如：「发」可對應到「發」和「髮」，因此轉換後的字量會多於原字表字量，例如法國簡體字表 805 字，簡繁轉換後為 886 字，不論採國教院 805 字或 886 字覆蓋率與之對比，都不能得到完全公平的比較。根據推估，數值差距若在 1 個百分點內，表示覆蓋率表現是相當的。依照以上覆蓋率特性，接下來即針對國教院字表與 7 份其他字表的覆蓋率及內容差異進行討論：

a. 法國字表 vs. 國教院字表

法國字表 805 字，經簡繁轉換後為 886 字，取國教院字表前 886 字於 COCT 中不同語料庫進行覆蓋率統計，得到以下結果：

表 7：法國字表與國教院字表在各語料庫的覆蓋率比較（%）

語料庫類型	口語	圖書	華語教材	中國時報	國語日報	聯合報
法國字表覆蓋率	86.71	81.64	85.06	78.00	80.81	78.13
國教院字表 886 字覆蓋率	87.16	81.94	86.14	79.63	81.85	79.82

根據上表，國教院字表與法國字表在覆蓋率上表現相當，表示法國字表的通用性高。然而，若取國教院字表前 886 字與法國字表做對照，交集 699 字，佔 78.89%，表示有兩成以上的字不同。分析國教院字表前 886 字有而法

國字表無者共 187 字，包含一些低頻而重要的部件，例如：「巴、木、玉」等，也包含一些生活用字，例如：「妹、咖、啡、鼻、喂」等，反觀法國字表有而國教院無者共 187 字，包含「懷、財、繼、黨」等字，這些字在國教院字表中均列為 B 等，據此，發現法國字表覆蓋率高的另一原因，在於法國初級字表中其實包含許多中級的高頻字，可見法國字表的定位與國教院字表有異，他們不以構字部件等書寫考量為主，而是希望中學生盡可能在少量的字集中就達到閱讀理解的能力，故初級字表中未包含較低頻的日常用字，而納入高頻的中級漢字。

b. 歐洲字表 vs. 國教院字表

歐洲字表也是簡體字表，經簡繁轉換後，原 630 字可對應到 694 個正體字，取國教院字表 694 字於 COCT 不同語料庫進行覆蓋率統計，結果如表 8：

表 8：歐洲字表與國教院字表在各語料庫的覆蓋率比較（%）

語料庫類型	口語	圖書	華語教材	中國時報	國語日報	聯合報
歐洲字表覆蓋率	79.93	72.64	78.42	66.40	70.15	67.10
國教院字表 694 字覆蓋率	82.62	76.43	82.14	71.35	74.82	72.01

歐洲字表的建置方式乃採舊字表為底，再主觀增補當代字集中的字，程序中未經語料驗證，因此常用度不夠平衡，導致覆蓋率較低。若取國教院字表相同字數來做比較，兩表交集為 542 字，佔 78.10%，國教院字表有而歐洲字表無者共 152 字，包括：「哭、跳、腳、麵」等初級常用字，國教院字表無而歐洲字表有者亦 152 字，包括：「博、藝、鞦、韆」等中高級字。

c. 漢辦字表 vs. 國教院字表

漢辦字表經簡繁轉換後，2,905 字可對應到 3,163 個正體字，分級字量與國教院相同字量做覆蓋率對比¹⁹後，結果如表 9：

¹⁹ 漢辦字表的丁級累計字經簡繁轉換後，共有 3,163 字，超出國教院字表的 3,131 字，故丁級部分國教院無法以相同字量進行比較。

表 9：漢辦字表與國教院字表在各語料庫的覆蓋率比較（%）

語料庫類型	口語	圖書	華語教材	中國時報	國語日報	聯合報
漢辦字表甲級 881 字覆蓋率	83.26	77.02	81.35	71.16	75.14	71.64
國教院字表 881 字覆蓋率	87.02	81.76	86.05	79.36	81.53	79.53
漢辦字表乙級 1,771 字覆蓋率	94.53	92.10	93.52	90.95	92.38	91.05
國教院字表 1,771 字覆蓋率	96.43	94.09	95.87	94.21	95.15	94.29
漢辦字表丙級 2,421 字覆蓋率	97.14	95.82	96.33	95.37	96.15	95.43
國教院字表 2,421 字覆蓋率	98.41	97.19	98.11	97.35	97.82	97.43
漢辦字表丁級 3,163 字覆蓋率	98.72	98.24	98.51	98.25	98.51	98.27
國教院字表 3,131 字覆蓋率 ²⁰	99.22	98.61	99.09	98.65	98.91	98.72

漢辦字表也是二語學習者字表，由上表可看出它的初級字在 COCT 語料庫的覆蓋率較低，這顯示出不同地區的華語用字差異大，如語氣詞的不同（臺灣：咦、嗨 vs. 大陸：喂、吶），食物用字的不同（臺灣：蚵、魷、滷 vs. 大陸：閘、腊、梁），常見專名用字的不同（臺灣：莉、婷 vs. 大陸：馮、耿），生活用字的不同（臺灣：飆、嗆 vs. 大陸：甬、姥）等等，故各地有建置個別字表的必要性。然而，從兩份字表的整體交集來看，共有 2,781 字是相同的，交集率高達 88.82%。國教院字表有而漢辦字表無者共 350 字，除上述所列不同用字外，尚包括「姊、朶、喔、莓、牠」等臺灣常用字，國教院字表無而漢辦字表有者則為 382 字，包括「襖、滬、馱、輓、哆」等臺灣罕用字。

d. 臺灣 2010 年常用語詞字表 vs. 國教院字表

國教院 2010 年語詞調查計畫字表特色有二：一為代表臺灣當代母語使用者之用字現況，二為以語料庫為本、未經人工調校之字表。由於是純粹統

²⁰ 國教院字表 3,100 字，若包含異體字則為 3,131 字。

計得出的字表，該字表 2,613 字即可覆蓋 190 萬字語料庫²¹ 的 99% 語料，取此 2,613 字於 COCT 語料庫做覆蓋率對比，結果如表 10：

表 10：臺灣 2010 年常用語詞字表與國教院字表在各語料庫的覆蓋率比較（%）

語料庫類型	口語	圖書	華語教材	中國時報	國語日報	聯合報
2010 年常用語詞字表 2613 字覆蓋率	98.90	98.08	98.03	98.49	98.66	98.50
國教院字表 2613 字覆蓋率	98.73	97.79	98.49	98.05	98.38	98.09

國教院字表與臺灣 2010 年常用語詞字表覆蓋率表現相當，可見國教院字表適用對象雖為二語學習者，但與母語者用字相去不遠，尤其經專家搭配能力指標調校後，仍能維持統計字表的高覆蓋率。比較 2010 年常用語詞字表 4,961 字與國教院字表 3,131 字（含異體字），交集部分共 3,109 字。國教院字表 3,131 字中扣除異體字，只有 8 字「雁、篆、甥、鉢、鱗、乒、乓、痔」不在 2010 年常用語詞字表中，可見國教院字表與臺灣母語者字表相當一致，是足以代表臺灣用字且符合臺灣二語教學需求之簡化字表。

e. 中國通用一級字表 vs. 國教院字表

中國通用一級字表 3,500 字，經簡繁轉換後為 3,775 字，因其未按字序排列，故僅能以 3,775 字與國教院 3,131 字做覆蓋率比較，結果如下：

表 11：中國通用一級字表與國教院字表在各語料庫的覆蓋率比較（%）

語料庫類型	口語	圖書	華語教材	中國時報	國語日報	聯合報
中國通用一級字表 3,775 字覆蓋率	99.27	99.20	99.15	99.17	99.30	99.22
國教院字表 3,131 字覆蓋率	99.22	98.61	99.09	98.65	98.91	98.72

中國通用一級字表是代表大陸母語使用者的字表，比國教院字表多出 644 字，但兩者覆蓋率表現相當。檢視兩字表交集共 3,026 字，佔國教院字集 96.65%，國教院字表有而中國通用一級字表無者共 105 字，包括許多語氣

²¹ 2010 年常用語詞調查計畫使用包含約 190 萬字的語料庫，內容包括書籍、新聞、雜誌、口語、教科書及其他等語料，是臺灣母語者用字的真實呈現。

詞，如「喔、嗯、噢、嗨」等，及臺灣生活用字「粧、粽、鮪、燉、姪、螂」等；國教院字表無而中國通用一級字表有者共 749 字，包括「瞅、氈、饑、伊、戍、廬」等難度較高的書面用字，較不適合二語教學使用。

f. 香港字表 vs. 國教院字表

香港字表為母語者字表，無簡繁轉換問題，與國教院字表分級累計字量的覆蓋率比較如表 12：

表 12：香港字表與國教院字表在各語料庫的覆蓋率比較（%）

語料庫類型	口語	圖書	華語教材	中國時報	國語日報	聯合報
香港字表一年級 508 字覆蓋率	67.83	60.84	68.78	53.39	56.75	53.09
國教院字表 508 字覆蓋率	74.56	66.14	74.54	58.38	63.12	59.47
香港字表二年級 1,224 字覆蓋率	88.21	84.12	88.29	80.50	82.57	80.73
國教院字表 1,224 字覆蓋率	92.28	88.36	92.51	86.74	88.76	87.05
香港字表三年級 1,856 字覆蓋率	94.19	91.82	94.13	90.50	91.65	90.71
國教院字表 1,856 字覆蓋率	96.67	94.43	96.24	94.42	95.37	94.53
香港字表四年級 2,414 字覆蓋率	96.82	95.68	96.67	95.15	95.67	95.26
國教院字表 2,414 字覆蓋率	98.40	97.16	98.10	97.34	97.82	97.42
香港字表五年級 2,826 字覆蓋率	97.87	97.22	97.61	97.21	97.33	97.27
國教院字表 2,826 字覆蓋率	99.06	98.27	98.77	98.48	98.76	98.55
香港字表六年級 3,000 字覆蓋率	98.18	97.71	97.87	97.91	97.84	97.95
國教院字表 3,000	99.19	98.50	98.97	98.61	98.88	98.68

香港字表初級字的覆蓋率偏低，表示香港用字與臺灣不同外，其低年級所習字相較於二語學習者而言仍稍有不足。觀察國教院字表有而香港字表無者共 233 字，包括「喔、哇、颯、菇、閩、滷、蚵」等語氣詞及臺灣常用字，國教院字表無而香港字表有者亦為 233 字，包括「晉、螺、鼎、蓉、傢、滔、鳴、兌」等字。若取國教院字表前 3,000 字與香港字表交集，共有 2,767 字相同，佔 92.23%，可見整體而言兩字表是相近的，只是常用性的排序落差甚大。

g. 新加坡字表 vs. 國教院字表

新加坡字表 2,749 字，扣除識讀及識寫的重複字後為 1,670 字，經簡繁轉換後共 1,831 字，與國教院字表分級累計字量的覆蓋率比較如表 13：

表 13：新加坡字表與國教院字表在各語料庫的覆蓋率比較（%）

語料庫類型	口語	圖書	華語教材	中國時報	國語日報	聯合報
新加坡字表一年級 343 字覆蓋率	54.97	47.91	55.95	38.38	42.73	39.17
國教院字表 343 字覆蓋率	70.09	61.65	68.69	54.38	58.51	54.97
新加坡字表二年級 693 字覆蓋率	75.51	67.05	75.41	58.57	63.24	59.52
國教院字表 693 字覆蓋率	82.61	76.41	82.13	71.34	74.80	72.00
新加坡字表三年級 1,027 字覆蓋率	83.80	77.70	84.09	71.88	75.55	72.68
國教院字表 1,027 字覆蓋率	90.29	85.64	88.72	84.39	86.32	84.56
新加坡字表四年級 1,379 字覆蓋率	90.09	86.29	90.35	81.93	84.87	82.48
國教院字表 1,379 字覆蓋率	93.76	90.14	93.08	89.96	91.33	90.04
新加坡字表五年級 1,672 字覆蓋率	93.27	90.28	92.89	88.13	90.00	88.49
國教院字表 1,672 字覆蓋率	96.03	93.52	95.23	93.79	94.73	93.82
新加坡字表六年級 1,831 字覆蓋率	94.56	92.13	94.18	90.42	92.06	90.69
國教院字表 1,831 字覆蓋率	96.62	94.34	96.13	94.37	95.33	94.47

新加坡字表為華裔學童字表，介於母語者與二語學習者字表之間，根據統計，它的初級部分覆蓋率偏低，與國教院字表覆蓋率差平均達 10% 以上，推測是主觀選字或人工調校的幅度較高，又未考量所收字的常用度所致。經取國教院字表前 1,831 字與新加坡字表對比，共有 1,525 字相同，佔 83.29%。觀察國教院字表有而新加坡字表無者共 306 字，包括「簽、捷、妻、姊、浴、寓、叉」等，國教院字表無而新加坡字表有者亦為 306 字，包括「蘿、瞪、眯、拇、感、麴、獻」等，部分是國教院的高級字，如「蘿、獻」；部分如「感、麴」等字新加坡列為小學用字，但未收錄於國教院 3,100 字表中，屬偏地域性的用字。

（3）各字表建置方式

上述多數字表僅提出收字原則及字集內容，對於字表產出的方式未加以說明，在國教院字表提出之前，歐洲字表是少數解說建置流程的字表，細節可參考 3.1 節對歐洲字表的簡介，然而，歐洲字表的覆蓋率表現稍低，原因在於他們採用其他字表作為底稿或增訂依據，但所採用字表的產出方式並未說明，因此歐洲字表的流程雖清楚，字集的產出卻仍是不透明的。

本文認為建置字表的過程應與完成的字集同樣受到檢視，因為字集內容的不完善常可歸因於建置流程的不合理，經過討論改善的流程，是與字集一樣重要的經驗傳承和未來續建字表的依據。國教院字表是目前已公布字表中建置流程最詳明的字表，包括所選用的語料內容、語料分級方法、收字原則、統計指標、調校理由、調校信度、驗證方式等均清楚陳述，除了可驗證外，更重要的是希望這套流程能建立再現字表的架構，未來可隨語料庫的更新、字表使用對象的變更，及不同華語地區的需求而據以建立不同的字表。

3.3 綜合分析

從上述比較可知，多數字表具有通用性，也有地域性，圖 2 統計不同字表的交集字數，8 份字表中全部都收的字僅佔 11%；考量部分字表僅有初級字，故放寬到 5 份字表包含的字，約佔 53%，因此各地確有自訂字表的必要性，王惠認為「教學字表的制定應充分立足於本地」（2010），與本文的觀察相同。

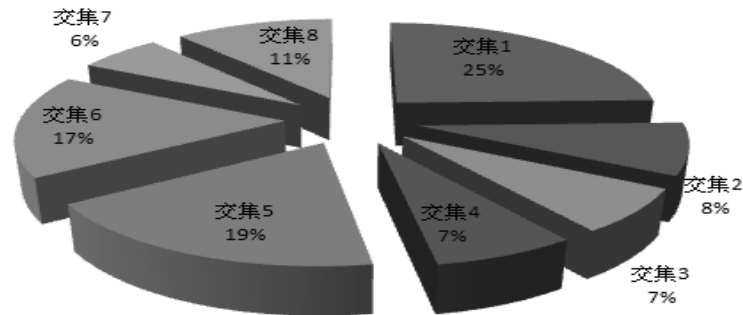


圖 2：八份字表的漢字交集率

其次，母語者字表和二語學習者字表最大的不同就是分級和字量，接近母語者的字表通常收納 3,000 字以上（如：臺灣 2010 年常用語詞字表、中國通用一級字表），二語學習者字表則多為 800 至 1,600 字（如：法國字表、歐洲字表、新加坡字表），一方面是由於他們認為初、中級字表對學習者而言是最重要的，部分學者則認為是因為漢字的書寫難度高，而華語教學的時間有限，故以「夠用」和「實用」為最佳標準（王惠 2010），法國字表、歐洲字表及新加坡字表甚至都再區分出主動書寫字與被動認讀字的分級字數，可見書寫難度對二語學習者字表而言十分重要。此外，未分級的母語字表通常不是教學字表，而是某地用字之調查，這類字表不建議直接拿來教學使用，除了包含大量低頻字，如臺灣 2010 年常用語詞字表的「洞、圇、訔、擻、斌」等，中國通用一級字表的「卜、匕、刁、汛、吮」等，且字量過多外，沒有分級資訊，也會對教學造成不便。本文建置的字表字量接近母語者字表，又搭配分級資訊，便於不同程度的學習者使用，是最完整的教學字表之一。

第三、根據上節統計，不同字表在初級的覆蓋率表現差異最大，原因是為了符合教學需求，初級字表常是專家調幅最大的一級。²² 然而，不同的教學目標及主觀想法，也會促使初級字表的內容差異大，例如前述法國字表及歐洲字表列入初級的字在國教院字表中約有兩成以上屬於中高級字或不收錄的字，詳見圖 3。因此，每一字表的建置原則需確立，並加以說明，才能讓使用者明白優先或延後習得某些字的原因。例如國教院字表的建置原則為 2.3 節所提到的 5 項統計指標及 9 項專家調校理由。其中，專家調校的理由及調校比重分布如圖 4。

²² 國教院字表的專家調幅各級均在 3 成上下，初級沒有顯著的差異，請參考表 4。

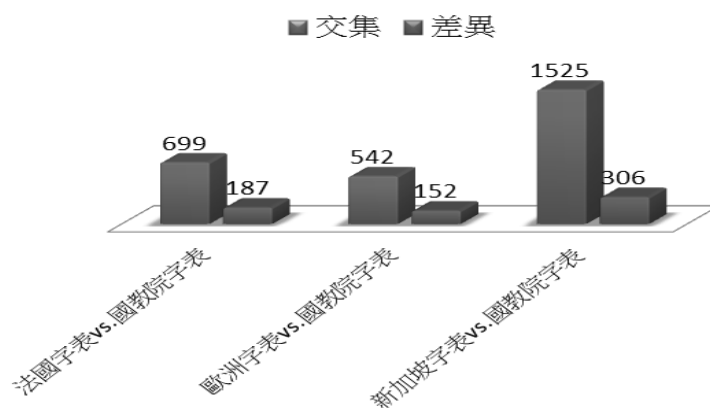


圖 3：三份初中級字表與國教院字表交集數的差異

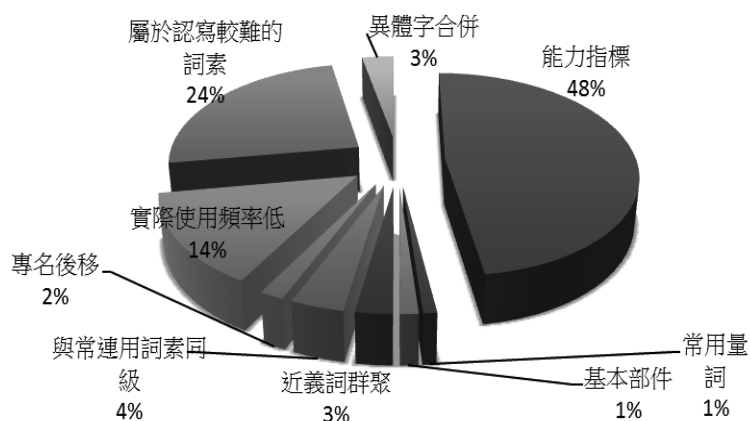


圖 4：專家調校理由分布

第四，字表的定位不同會造成字集內容的差異，例如國教院字表以「認寫」為主，不排除語氣詞，觀察其他字表，如中國通用一級字表、香港字表等則較少收錄語氣詞，一方面可能是他們語氣詞用得少，但也可能是字表定位的關係。又如異體字，簡體字表沒有異體字問題（如：法國字表、歐洲字表、新加坡字表、中國通用一級字表都只收「床」，不收「牀」），香港字表則不收「床」，僅收「牀」，國教院字表兩者兼採，並定位為同級字。

第五，採用舊字表作為底本或增訂依據來建立新的字表，往往會導致常用度失衡、覆蓋率偏低的現象，如歐洲字表。因此，選用當地語料庫，先統計出初步字表再加以調校，應該是最好的字表產出方式。

4. 結論

本文分析國教院字表的建置流程及成果，並藉由與其他 7 份字表的對比研究來說明國教院字表的特色：客觀、可再現、真實性、字本位及符合二語學習者需求。以語料庫為本的教學字表建置有以下重要步驟：選擇語料、決定字表定位、設定分級字量、決定統計及調校指標、分析調校信度及幅度，以及字表驗證，上述每一方面在實作時都會遭遇困難，過去字表因未說明建置流程，因此也較少關於實作上的討論，本文根據實際建置經驗，將採取的作法歸納如下：本文在「語料選擇」上歷經兩個步驟，其一，比較母語者書面語語料、母語者口語語料、母語者新聞語料及二語學習者華語教材語料，觀察各語料產生字表的字頻序，選取了最適合二語學習者的華語教材語料。其二，比較目前常用教材和僑委會教材的難度和取字廣度，最後選取包含僑委會教材在內的 34 套教材為統計底本。在「分級定位」上，採字本位，以認寫功能為主，與詞本位、溝通為主的詞表做出區隔。然雖如此，在不違背字本位的原則下，選取與現代生活密切的字也是常見的調校理由。在「分級字量」的設定上，分 3 階段進行，首先，以對應 CEFR 提出的字量訂出 3,000、1,800、800 的標準（張莉萍 2008）。再參考 Laufer (1989) 及 Hu and Nation (2000) 等關於閱讀理解的覆蓋率需求分析，將字表初、中、高級漢字在各種語料上的覆蓋率設定為：75%、95%、99%，最後，專家依教學經驗，將各級累計字量分別定為 Pre-A 級 200 字、A1 級 500 個字、A2 級 800 字、B1 級 1,300 字、B2 級 1,900 字、C1 級 2,500 字、C2 級 3,100 字。

在「統計及調校指標」方面，本文訂定 5 項統計指標，分別是各級華語教材累計覆蓋率分級、全級華語教材累計覆蓋率分級、字表涵蓋數、各級華語教材篇章分布率與全級華語教材篇章分布率，及 9 項專家調校指標，分別是能力指標、構字率、構詞率、語義聯想、常用詞群類聚、認寫難易度、專名移除或級別調整、功能詞級別調整及異體字整併等，以此 14 項指標來兼顧字表之常用性、需用性及基本性。在「調校信度及幅度」上，本文邀請 5 位資深華語教師參與調校，並針對個別調校結果作了 Weighted Kappa 一致性統計，結果顯示專家調校的信度達到「幾乎完全吻合」等級。調校後字表與統計字表相較，各級調校幅度均在 3 成左右，經分析，有可能是因為各種教材的字詞交集率偏低的緣故。此外，調校前後的字集語料覆蓋率，包含口語、新聞及圖書等語料庫，均下降不到 .2%，表示本字表可兼及「常用性」及「基本性」、「需用性」的要求。在「字表驗證」部分，與漢辦字表、歐洲字表、

新加坡字表及香港字表的初級比較起來，本字表在各語料庫的覆蓋率平均高出 4 個百分點，可見各地字表確有地域性的限制，而本字表最能代表臺灣當代華語的樣貌。字表需經實際教學及評量使用後方能顯出效能，因此在字表內容得到回饋後也可以在上述步驟中進行調整；未來也可進一步運用計算語言學技術，思考專家調校指標數據化的方法，如以機器協助判斷「語義聯想、常用詞群類聚、認寫難易度」等面向的分數等等，使語料庫發揮更大的效益。

總體而言，本字表的建置以語料庫統計法為主，專家調校亦根據統計數據，故能達到客觀的目標。客觀亦有賴於建置流程透明（即「可再現」），因此本文在字表建置的每一階段都清楚紀錄流程，希望研發經驗能累積至下一字表。此外，本文所使用的語料庫都是臺灣語料，專家調校時亦重視在臺二語學習者的生活需求，因此收字、分級均考量真實性，如將「珍（珠奶茶）、捷（運）」兩字納入初級。本文也呼應白樂桑教授的「字本位」教學法，將部件構字率、構詞率、書寫難易度等多項字本位原則納入專家調校指標。最後，本字表為符合二語學習者的需求，故以教材語料為統計來源，再輔以其他類型語料之驗證，並且予以分級。過去受限於缺乏巨量語料支持，很難客觀評估字的級別，因此多數字表僅把字集以整體呈現的方式表列，不符合教學需求；有了語料庫之後，不僅有客觀的分級依據，也能提供學習者相應的例句，提升學習的效能及品質。

引用文獻

- Allanic, Bernard. 2005. "Yiji yuedu zibiao" ji shuoming (« Le "premier seuil d'accès à la lecture" et son explication » (The first threshold for reading and its explanation)). *International Chinese Language Teaching & Learning Report* 4: 14-24.
- Allanic, Bernard, and Changying Shu. 2012. The division and connections between lexical items for oral and for written activities. *EBCL Project Seminar Report*. Belgium Brussels.
- American Council on the Teaching of Foreign Languages. 2012. *ACTFL Proficiency Guidelines*. Accessed online, June 15, 2017. <https://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012>.
- Bai, Ming-hong, Jian-cheng Wu, Ying-ni Chien, Shu-ling Huang, and Ching-lung Lin. 2016. A study on dispersion measures for core vocabulary compilation.

- Computational Linguistics and Chinese Language Processing* 21.2: 1-18.
- Centre for Canadian language benchmarks. 2012. *Canadian Language Benchmarks: English as a Second Language for Adults*. Accessed online, June 15, 2017. <http://www.cic.gc.ca/english/pdf/pub/language-benchmarks.pdf>
- Chinois Classe de Seconde Générale et Technologique. 2003. Ministry of National Education (France): Chinese for General and Technological High School. Accessed online, June 15, 2017. <http://www2.cndp.fr/archivage/val-id/55324/55324-9742-18137.pdf>
- Cohen, Jacob. 1968. Weighed kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin* 70.4: 213-220.
- Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- Hu, Marcella, and Paul Nation. 2000. Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language* 23.1: 403-430.
- Laufer, Batia. 1989. What percentage of text-lexis is essential for comprehension? *Special Language: From Humans Thinking to Thinking Machines*, eds. by Christer Lauren and Mar Nordman, 316-323. Clevedon: Multilingual Matters.
- Song, Lianyi. 2012. European benchmarking Chinese Language-overall introduction. Accessed online, May 15, 2017. <http://ebcl.eu.com/wp-content/uploads/2012/05/EBCL-Berlin-Seminar-Overall-Introduction.pdf>.
- 中華民國教育部。1982。《常用國字標準字體表》。臺北：正中書局。[Ministry of Education, R.O.C. 1982. *List of Commonly Used Characters in Chinese*. Taipei: Cheng Chung Book Company.]
- 中國國家漢辦。1992。《漢語水平詞彙與漢字等級大綱》。北京：北京語言學院出版社。[Office of Chinese Language Council International. 1992. *The Syllabus of Graded Word and Characters for Chinese Proficiencies*. Beijing: Beijing Language and Culture University Press.]
- 王敏、王曉明。2012。〈《漢語水平詞彙與漢字等級大綱》字表的收字分級考察例析〉，《河北大學學報（哲學社會學科版）》，第37卷第4期，103-107。[Wang, Min, and Xiao-ming Wang. 2012. The analysis for the classification of characters in the table of “words and characters for Chinese ability outline”. *Journal of Hebei University (Philosophy and Social Science)* 37.4: 103-107.]
- 王惠。2010。〈新加坡華文用字量與教學研究〉，《語文建設通訊》，第95期，23-29。[Wang, Hui. 2010. The study of Chinese teaching and Chinese words

- used in Singapore. *Newsletter of Language Construction* 95: 23-29]
- 王瓊珠、洪儷瑜、張郁雯、陳秀芬。2008。〈一到九年級學生國字識字量發展〉，《教育心理學報》，第39卷第4期，555-568。[Wang, Chiung-chu, Li-yu Hung, Yu-wen Chang, and Hsiu-fen Chen. 2008. Number of characters school students know from grade 1 to grade 9. *Bulletion of Education Psychology* 39.4: 555-568.]
- 白樂桑。1996。〈漢語教材中的文、語領土之爭：是合併，還是自主，抑或分離？〉，《世界漢語教學》，第4期，100-102。[Bellassen, Joël. 1996. The territorial dispute between the spoken and written Chinese over the textbooks: Integration, self-rule or separation? *Chinese Teaching in the World* 4.]
- 白樂桑、張麗。2008。〈《歐洲語言共同參考框架》新理念對漢語教學的啟示與推動—處於抉擇關頭的漢語教學〉，《世界漢語教學》，第3期，58-73。[Bellassen, Joël, and Li Chang. 2008. The enlightenment and the impetus of the new approach of the Common European Framework of Reference for language on the Chinese language teaching. *Chinese Teaching in the World* 3: 58-73.]
- 石定果。1997。〈漢字研究與對外漢語教學〉，《語言教學與研究》，第1期，29-41。[Shi, Ding-guo. 1997. Chinese character studies and teaching chinese as a foreign language. *Language Teaching and Linguistic Studies* 1: 29-41.]
- 宋曜廷、陳茹玲、李宜憲、查日齋、曾厚強、林維駿、張道行、張國恩。2013。〈中文文本可讀性探討：指標選取、模型建立與效度驗證〉，《中華心理學刊》，第55卷第1期，75-106。[Sung, Yao-ting, Ju-ling Chen, Yi-shian Lee, Jih-ho Cha, Hou-chiang Tseng, Wei-chun Lin, Tao-hsing Chang, and Kuo-en Chang. 2013. Investigating Chinese text readability: Linguistic features modeling, and validation. *Chinese Journal of Psychology* 55.1: 75-106.]
- 林季苗。2011。〈漢語教學四大原則與法國經驗〉，《華語文教學研究》，第8卷第2期，65-79。[Miao, Lin-zucker. 2011. Four principles of Chinese language teaching: Feedback from a France experience. *Journal of Chinese Language Teaching* 8.2: 65-79.]
- 林慶隆、吳鑑城、陳毓欣。2018。《九十九年常用語詞調查報告書》。臺北：國家教育研究院。[Lin, Ching-lung, Jian-cheng Wu, Yu-xin Chen. 2018. *2010 Commonly Used Words Survey Project Report*. Taipei: National Academy for Educational Research.]
- 柯華葳、林慶隆、張俊盛、陳浩然、高照明、蔡雅薰、張郁雯、陳柏熹、張莉萍。

- 2016。《華語文八年計畫「建置應用語料庫及標準體系」105 年工作計畫期末報告》。臺北：國家教育研究院。[Ko, Hwu-wei, Ching-lung Lin, Jason S. Chang, Hao-jan Chen, Zhao-ming Gao, Ya-hsun Tsai, Yu-wen Chang, Po-see Chen, and Li-ping Chang. 2016. *The 2016 Research Report of The 8-Year Project of Construction and Application of Mandarin Chinese Corpus and Standard Systems*. Taipei: National Academy for Educational Research.]
- 香港浸會大學。2003。《香港小學中文科常用字表》。2016 年 11 月 26 日取自 http://lcprichi.hkbu.edu.hk/table_search.html。[Hong Kong Baptist University. 2003. *The Chinese Characters Recommended for the Subject of Chinese Language in Primary Schools*. Accessed online, November 26, 2016. http://lcprichi.hkbu.edu.hk/table_search.html.]
- 唐力行、姚學吾。2001。〈美國中文學校教師的需求：教學提高及教師培訓〉，*Journal of the Chinese Language Teachers Association*, 第 32 卷第 1 期，81-106。[Tang, Frank, and Xue-wu Yao. 2001. Demand for teachers in American Chinese schools: Teaching improvement and teacher training. *Journal of the Chinese Language Teachers Association* 32.1: 81-106.]
- 莊德明、鄧賢瑛。2009。《漢字構形資料庫的研發與應用》。臺北：中研院資訊所文獻處理實驗室。[Juang, Der-ming, and Sian-ying Deng. 2009. *Development and Application of Chinese Character Formation Database*. Taipei: Sinica Academia.]
- 國家漢語國際推廣領導小組辦公室。Chinese Proficiency Test (HSK)。2016 年 8 月 3 日取自 http://english.hanban.org/node_8002.htm#nod。[Office of Chinese Language Council International. Chinese Proficiency Test (HSK). Accessed online, August 3, 2016. http://english.hanban.org/node_8002.htm#nod]
- 國家教育研究院。2017。〈如何進行教學閱讀 Q&A〉。2017 年 1 月 26 日取自 <http://www.naer.edu.tw/ezfiles/0/1000/img/63/54777878.pdf>。[National Academy for Educational Research. 2017. How to carry out teaching and reading. Accessed online, January 26, 2017. <http://www.naer.edu.tw/ezfiles/0/1000/img/63/54777878.pdf>]
- 黃沛榮。2001。《漢字教學的理論與實踐》。臺北：樂學書局。[Huang, Pei-jung. 2001. *Theory and Practice of Chinese Character Teaching*. Taipei: Lexis Book Company.]
- 曾金金。2000。〈德國學生學習漢字的情況及其學習策略〉，《語言研究—國際漢語教學學術研討會論文集》，327-331。[Tseng, Chin-chin. 2000. Ger-

- man students' strategy of learning Chinese characters. *Linguistic Studies- Proceedings of the International Symposium on Chinese Language Teaching* 327-331.]
- 張金蘭。2014。〈法國漢字教學現況及其對臺灣華語文教學的啟示〉，《中原華語文學報》，第13期，29-52。[Chang, Ching-lan. 2014. Chinese character teaching in France and the revelation of TCSL in Taiwan. *Chung Yuan Journal of Teaching Chinese as a Second Language* 13: 29-52.]
- 張欣怡。2009。《針對中級華語學習者之漢字課程設計》。臺北：國立臺灣師範大學華語文教學研究所碩士論文。[Zhang, Xin-yi. 2009. *The Design on the Chinese Characters Course in Intermediate Level for Learners of Chinese as a Foreign Language*. Taipei: National Taiwan Normal University MA thesis.]
- 張朋朋。2005。〈談「字本位」的內涵〉，《漢字文化》，第4期，9-12。[Chang, Peng-peng. 2005. The study of sinogram-based theory of Chinese language. *Sinogram Culture* 4: 9-12.]
- 張郁雯。2003。〈詞彙分級研究〉，《華語文能力測驗編製：研究與實務》，柯華葳（主編），83-102。臺北：遠流出版社。[Chang, Yu-wen. 2003. The study of vocabulary grading. *Chinese Language Proficiency Test Preparation: Research and Practice*, ed. by Ko, Hwa-Wei, 83-102. Taipei: Yuan-Liou Publishing.]
- 張莉萍。2008。〈對外漢語字集〉。發表於「台灣華語文教學年會暨研討會」，2008年11月。花蓮：慈濟大學。[Chang, Li-ping. 2008. Chinese character sets. Paper presented at *Taiwan Chinese Language Teaching Annual Conference*, November 2008. Hualien, Taiwan: Tzu Chi University.]
- 張莉萍。2012。〈對應於歐洲共同架構的華語詞彙量〉，《華語文教學研究》，第9卷第2期，77-96。[Chang, Li-ping. 2012. The study of the vocabulary size at the CEFR levels for CFL/CSL learners. *Journal of Chinese Language Teaching* 9.2: 77-96.]
- 張凱。1997。〈漢語構詞基本字的統計分析〉，《語言教學與研究》，第1期，42-51。[Chang, Kai. 1997. Statistical analysis of basic words in Chinese word formation. *Language Teaching and Linguistic Studies* 1: 42-51.]
- 新加坡教育部。2007。《新加坡小學華文字表》。2016年12月13日取自 <https://www.moe.gov.sg/.../character-list-primary-chinese-2007.pdf>。[Ministry of Education of Singapore. 2007. The Chinese Characters for the Subject of Chinese Language in Primary Schools. Accessed online, December 13, 2016. <https://www.moe.gov.sg/.../character-list-primary-chinese-2007.pdf>]

- 葉德明。2000。〈外籍生漢字書寫策略探討〉，《第六屆世界華語文教學研討會論文集（三）》，311-320。臺北：世界華語文教育學會。[Yeh, Teh-ming. 2000. Foreign students' writing strategy of Chinese characters. *The 6th World Chinese Language Teaching Seminar Proceeding* 311-320. Taipei, Taiwan: World Chinese Language Association.]
- 劉英林（主編）。1996。《漢語水平等級標準與語法等級大綱》。北京：高等教育出版社。[Liu, Ying-lin. (ed.) 1996. *Chinese Proficiency Scales and Graded Grammar Levels*. Beijing: Higher Education Press.]

[審查：2017.9.14 修改：2018.2.14 接受：2018.3.1]

黃淑齡

Shu-Ling HUANG

10644 臺北市和平東路一段 179 號

國家教育研究院語文教育及編譯研究中心

Research Center for Translation Compilation and Language Education

National Academy for Educational Research

No.179, Sec. 1, Heping E. Rd., Taipei City 10644, Taiwan

slhuang@mail.naer.edu.tw

白明弘

Ming-Hong BAI

10644 臺北市和平東路一段 179 號

國家教育研究院語文教育及編譯研究中心

Research Center for Translation Compilation and Language Education

National Academy for Educational Research

No.179, Sec. 1, Heping E. Rd., Taipei City 10644, Taiwan

mhbai@mail.naer.edu.tw

華語文教學研究

吳鑑城

Jian-Cheng WU

10644 臺北市和平東路一段 179 號

國家教育研究院語文教育及編譯研究中心

Research Center for Translation Compilation and Language Education

National Academy for Educational Research

No.179, Sec. 1, Heping E. Rd., Taipei City 10644, Taiwan

wujc@mail.naer.edu.tw

李詩敏

Shih-Min LI

10644 臺北市和平東路一段 179 號

國家教育研究院語文教育及編譯研究中心

Research Center for Translation Compilation and Language Education

National Academy for Educational Research

No.179, Sec. 1, Heping E. Rd., Taipei City 10644, Taiwan

smli@mail.naer.edu.tw

林慶隆

Ching-Lung LIN

10644 臺北市和平東路一段 179 號

國家教育研究院語文教育及編譯研究中心

Research Center for Translation Compilation and Language Education

National Academy for Educational Research

No.179, Sec. 1, Heping E. Rd., Taipei City 10644, Taiwan

cllin@mail.naer.edu.tw

附錄 1 運用華語文語料庫計算「漢字」統計範本

序位	字	初級教材	初級教材	初級教材	初級教材	中級教材	中級教材	中級教材	中級教材	高級教材	高級教材	高級教材	高級教材	全級教材	全級教材	全級教材	全級教材
		字頻	累積字頻	累積字頻	累積字頻	字頻	累積字頻	累積字頻	累積字頻	字頻	累積字頻	累積字頻	累積字頻	字頻	累積字頻	累積字頻	累積字頻
1	的	1054	2066	0.0544	1	12736	12736	0.034087	1	7987	7987	0.037082	1	21777	21777	0.034287	1
2	不	663	6553	0.1422	1	5924	33220	0.088911	1	2923	17484	0.081176	1	9510	63282	0.099636	1
3	好	688	5890	0.1278	1	2354	94988	0.254227	1	764	61197	0.284128	1	3808	147054	0.231533	1
4	有	611	7164	0.1554	1	4611	53410	0.142947	1	2299	25173	0.116874	1	7521	86870	0.136775	1
5	個	352	12654	0.2745	1	3061	76879	0.20576	1	1112	46854	0.217536	1	4525	122142	0.192321	1
6	天	341	13689	0.2947	1	2153	103730	0.277625	1	769	60433	0.280581	1	3263	164309	0.258701	1
7	上	285	15207	0.3299	1	2248	97236	0.260244	1	1172	44598	0.207062	1	3705	150759	0.237366	1
8	小	426	10386	0.2253	1	1686	122555	0.328008	1	436	84764	0.393546	1	2549	205449	0.323474	1
9	去	360	11946	0.2592	1	2180	99416	0.266079	1	696	65524	0.304218	1	3236	170792	0.268908	1
10	多	218	18049	0.3916	1	1335	135858	0.363613	1	745	61942	0.287587	1	2298	212590	0.334718	1
11	就	207	18474	0.4008	1	2751	82609	0.221096	1	1242	39879	0.185152	1	4200	135129	0.212757	1
12	家	249	15998	0.3471	1	1774	120869	0.323496	1	1012	50960	0.2366	1	3035	183328	0.288646	1

序位	字	初级教材	中级教材	高级教材	全级教材	初级教材	初级教材	中级教材	中级教材	中级教材	高级教材	高级教材	高级教材	全级教材	全级教材	全级教材	字表覆盖数		
		篇章分布	篇章分布	篇章分布	篇章分布	构词率	构词	构词特性	构词率	构词	构词特性	构词率	构词	构词特性	构词率	构词		构词特性	
1	的	0.60562	0.949273	0.935621	0.866232	4	的, 1020	5.166667	的, 12346	6	的, 7777	6.333333	的, 21143	7	的, 21143	7			
2	不	0.54063	0.876044	0.873461	0.793627	18	16667	10.5667	不, 3157	12	11246	10.5667	不, 1377	161	7833	4908	4908		
3	有	0.520256	0.647991	0.582851	0.552006	18	166667	39.439	3.24	58	83333	34.7	130	76	91667	1887	1636		
4	好	0.409592	0.75917	0.86878	0.760548	12	11667	40	4.203	37	033333	1134	2834	49	22619	1059	1453		
5	惜	0.330221	0.741699	0.714061	0.628546	1	352	3.33333	290	1	29051	5.66667	100	6.99999	6.333333	4262	41254		
6	上	0.327608	0.618562	0.571517	0.522868	18	166667	73	M.205	6	11246	10.5667	299	28	25238	1383	4183		
7	小	0.319643	0.657712	0.747479	0.585496	20	16667	116	162	58	676667	1358	1249	56	83333	1707	6455		
8	上	0.316847	0.401787	0.392832	0.354659	19	41667	30	19	292	5	616	1	799	30	16667	142	183	
9	去	0.30065	0.633536	0.542292	0.513561	10	083333	303	103	147	1.7	556667	34	159	92	59	66667	435	1286
10	多	0.279018	0.513764	0.614451	0.495829	10	083333	303	103	147	1.7	556667	34	159	92	59	66667	435	1286
11	就	0.275106	0.781266	0.799708	0.651523	2	205	205	205	205	205	205	205	205	205	205	205	205	205
12	家	0.235165	0.566654	0.549918	0.469611	13	083333	303	103	147	1.7	556667	34	159	92	59	66667	435	1286

資策會基礎部件組字頻率表（摘錄自莊德明、鄧賢瑛，2009）

註：本文僅取用獨體字部件

編號	序 號	部 件	組字字數	組字頻率(%)	累計頻率(%)
1	98	口	2526	6.29%	6.29%
2	1	一	1732	6.03%	12.32%
3	38	人	955	2.56%	14.88%
4	172	日	895	2.51%	17.39%
5	39	亻	551	2.47%	19.86%
6	5	丶	308	2.04%	21.90%
7	82	土	795	1.82%	23.72%
8	281	白	191	1.72%	25.44%
9	52	勹	292	1.64%	27.08%
10	40	八	713	1.57%	28.66%
11	18	十	572	1.48%	30.14%
12	153	木	956	1.42%	31.56%
13	373	言	364	1.24%	32.80%

編號	序 號	部 件	組字字數	組字頻率(%)	累計頻率(%)
39	183	才	515	0.64%	54.67%
40	144	王	330	0.64%	55.30%
41	137	冫	306	0.63%	55.94%
42	184	攴	263	0.63%	56.57%
43	73	厶	278	0.62%	57.19%
44	27	丁	149	0.59%	57.78%
45	391	門	128	0.57%	58.35%
46	353	彡	324	0.57%	58.92%
47	58	冫	398	0.56%	59.47%
48	68	冫	206	0.54%	60.02%
49	3	丨	90	0.54%	60.56%
50	83	士	188	0.54%	61.09%
51	35	人	243	0.53%	61.63%

附錄 2 國教院四等七級字表初稿

Pre-A 級 (200 字)

的,不,好,有,個,天,上,小,去,多,就,家,起,得,兒,子,他,太,學,中,
問,再,三,喜,今,友,老,知,給,道,見,買,水,公,氣,飯,姐,最,開,覺,
四,別,早,車,師,電,寫,頭,花,六,書,果,機,弟,我,一,你,麼,了,這,
嗎,們,要,在,看,來,什,很,可,以,呢,那,都,大,生,人,啊,下,想,沒,
對,請,裡/裏,先,點,時,方,明,吃,兩,哪,怎,吧,和,歡,幾,會,真/真,
媽,五,謝,字,西,十,朋,您,面,常,東,話,為,做,快,少,跟,日,候,用,
過,後,能,國,現,打,前,地,二,玩,事,期,文,比,心,出,年,坐,走,所,
新,晚,月,從,同,名,回,誰,錢,手,路,帶,她,安,山,口,八,塊,爸,外,
喝,樂,愛,房,拿,飛,是,還,也,到,王,說,樣,美,些,聽,叫,找,平,如,
又,星,哥,住,百,半,每,因,門,七,九,本,馬,妳,午,妹,姓,昨,元,千,
女,工

A1 級 (300 字)

邊,色,眼,畫,習,長,動,亮,張,物,服,李,林,把,華,玉,菜,田,拍,等,
像,關,客,問,著/着,衣,包,次,歌,牛,春,定,情,行,意,然,便,教,放,成,
海,近,自,高,分,經,幫,位,白,種,進,才,法,加,北,正,久,照,風,忙,
作,接,聲,念,遠,當,火,跑,信,錯,只,空,台/臺,表,金,光,親,母,狗,
灣,隻,校,唱,試,筆,店,站,球,醫,境,漂,練,紅,司,課,貴,鐘,黃,單,
運,館,熱,圖,慢,片,睛,肉,雨,準,較,香,毛,爺,換,影,餐,右,左,號,
奶,票,游,瓜,迎,巴,泳,貨,郵,笑,題,陽,樹,第,重,非,己,興,思,相,
向,應,難,身,力,往,節,望,語,南,辦,腳,穿,發,音,全,讓,該,已,孩,
通,部,差,園,始,燈,茶,紙,草,係,睡,息,詞,特,記,送,共,視,算,場,
禮,歲,旅,計,青,交,跳,轉,英,木,伯,洗,腦,休,班,顏,冷,累,米,臉,
舒,雙,杯,糕,嘴,廳,樓,貓,宜,牙,枝,搬,汽,綠,雪,蛋,鼻,喂,麵/麪,
租,豆,糖,認,容,直,件,怕,或,條,必,賣,河,體,滿,許,結,故,化,完,
更,離,但,其,父,活,各,言,原,處,越,石,收,秋,院,識,希,易,萬,忘,
市,雞/鷄,魚,病,男,羊,旁,功,景,停,屋,網,夠,夏,室,街,藥,床/牀,
黑,刀,豬,冬,末,週,桌,卡,咖,啡,內,袋,餓,哭,懂,吵,局,筷,冰,渴,
斤,飽,椅,址

A2 級 (300 字)

立,員,棒,猜,哇,蘋,告,變,它,合,連,形,備,數,趣,示,具,展,派,麻,器,約,圓,珍,捷,喔,感,實,理,害,流,而,且,代,利,目,留,報,假,參,船,健,程,刻,訴,鳥,剛,費,需,遊,夜,雲,入,油,觀,舞,味,象,式,食,品,阿,裝,步,保,鄉,領,奇,量,賽,務,乾,聰,呀,啦,戲,叔,煩,里,皮,簡,祝,注,騎,康,足,鐵,京,線,燒,鞋,切,價,低,支,附,丁,戶,鬧,盤,座,排,折,巧,餅,婆,獅,哈,短,朵/朶,聞,耳,級,窗,甜,湯,藍,紹,銀,髮,介,寄,查,爬,概,封,洋,冒,超,腿,克,陪,厭,逛,貼,沙,闖,煙/烟,眉,刷,矮,踢,汁,蕉,淺,寸,套,碼,姊,覽,簽,舅,莓,世,主,清,城,讀,夫,酒,輕,答,居,舉/擧,由,無,傳,總,神,農,楚,雖,於,受,苦,死,被,軍,命,民,掉,此,龍,及,商,考,決,急,選,除,業,願,提,求,極,慣,終,解,助,句,界,根,初,周,精,拜,度,怪,類,統,管,深,演,改,整,隊,反,論,壞,何,顧,消,背,例,社,義,遇,婚,傷,取,性,般,破,殺,府,救,村,妻,布,虎,寓,趕,偷,敢,緊,失,另,隨,弄,孫,丟,究,痛,土,舊,努,借,存,研,戴,珠,溫,帽,盒,拉,飲,碗,鮮,肚,付,訂,籃,汗,笨,架,廚,畢,姑,晴,箱,輛,胖,瘦,宿,澡,酸,瓶,齒,舍,浴,餃,褲,裙,櫃,臭,辣,叉,罐,匙,鹹

B1 級 (500 字)

甲,欣,乙,牠,首,陳,古,德,娃,佳,強/强,偉,漢,者,指,寶,稱,順,蓋,聖,誕,恭,宵,閃,噁,噢,丙,謎,晶,嚐,使,建,養,將,詩,並,族,之,戰,與,卻/却,史,術,造,落,治,擔,至,達,任,富,政,志,斷,麗,齊,令,兵,武,曾,士,產,孔,竹,倒,忽,甚,竟,抓,逃,蛇,烏,鼠,猴,鹿,吳,料,底,鄰,野,騙,捉,引,則,勢,責,爭,守,團,江,祖,群/羣,段,講,靜,護,調,細,續,幸,談,推,技,適,響,驗,環,議,藝,章,財,夢,修,材,賞,繼,替,待,靠,危,聊,討,圍,宮,敬,善,險,迫,窮,際,勇,亂,止,嚴,持,權,敗,敵,投,犯,葉,警,鬼,植,驚,勸,仍,恐,既,忍,躲/躲,唐,漸,刺,煮,搖,辛,吹,縣,兔,狼,井,招,丈,嚇,模,兄,吸,漁,尾,咬,匹,滾,雷,棋,廣,角,育,須,確,據,幹,呼,資,職,勞,福,省,眾,庭,獨,科,州,格,婦,湖,肯,互,勝,負,紀,專,集,壓,乎,永,堂,質,掛,態,罪,若,似,豐,密,慶,盛,未,復,增,察,聯,採,移,鏡,素,設,宣,陸,惜,絕,即,免,俗,翻,堅,

端,穎,廟,鼓,祥,嘛,季,層,突,困,迷,訪,熟,規,雜,制,血,證,暖,衛,
績,印,劇,姨,系,橋,況,抱,減,頂,值/值,禁,歷,疑,否,尊,寒,浪,涼,
醒,效,按,散,仔,牆,悲,誤,遍,彩,退,缺,輪,激,普,篇,弱,預,夕,淨,
岸,份,源,誠,尺,臨,策,案,固,狀,仁,博,申,銅,型,曲,娘,池,舌,洲,
針,織,孝,歐,毒,童,島,傘,鬆,尤,碰,享,讚,抬/擡,露,罵,伸,避,躺,
仗,搶,陣,貪,淚,智,微,閒,剩,衝,壯,頓,喊,憐,堆,充,惡,昏,棄,聚,
厲,挑,耐,膽,拔,禍,脫,獎,洞,惱,尖,晨,絲,途,敲,賺,裸,貌,踏,撞,
搭,雄,偏,摸,牌,吉,沖,嫁,砍,悶,吞,乖,娶,剪,泡,郎,鎮,烤,扔,挖,
購,曆,款,羽,敏,壺,嘗,鴨,項,紫,桃,埧,圾,唉,嘆,熊,醜,區,陰,典,
標,束,基,組,暗,乘,訓,厚,律,升,私,速,泉,載,醉,淡,零,疼,森,麟,
暑,肥,寬,亞,呆,廠,粉,貝,秒,染,港,鹽,庫,麥,欠,錶,央,啤,乳,億,
冊,粗,鍋,巷,彎,棟,川,炸,吐,扮,薪,填,煎,狂,濃,礎,齡,瘋,震,頁,
腐,枕,膚,拌,診,廁,蝦,鎖,唸,炎,漿,丸,燙,蚊,巾,炒,醋,扇,橘,椒,
颯,醬,濕/溼,帥,污/汙,饅

B2 級(600 字)

芳,怡,莉,稻,趙,積,愁,魔,遙,奔,官,朝,皇,帝,靈,莊,楊,孟,箭,盡,
射,攻,獵,仙,雅,翁,仲,羅,宋,劉,弓,劍,珊,玲,婷,祭,依,致,紛,評,
承,懷,探,忠,胡,亡,編,判,虛,慌,罷,欺,藏,凡,佩,遭,妙,傲,滅,猛,
尋,寧,驕,威,歸,幸,獸,捕,施,燕,尚,滴,雀,盜,柳,廉,蟻,嬰,螞,盾,
儒,沈,莫,匠,爪,杜,錦,鄭,徐,鳳,斑,甘,芬,慕,贏,繩,皺,愚,默,窩,
賊,扁,鷹,斧,矛,苗,梨,巫,驢,豹,菇,革,黨,良,抗,閩,板,濟,導,補,
擇,率,防,奮,橫,顯,番,災,益,貧,獲,損,勤,操,降,僅,配,列,餘,供,
賀,製,宗,營,掃,屬,築,序,梅,創,辭,蟲,揮,授,限,骨,墓,奏,榮,毫,
阻,握,烈,擺,擊,悟,壁,含,維,透,喪,侵,緣,藉,伴,旦,輩,鬥,慮,飾,
廢,符,谷,怒,穩,鞭,逐,遷,爛,慧,哀,仰,掌,愈,彈,曉,謂,徵,融,儀,
幽,墨,偶,秀,爆,蘭,雕,磨,僑,哎,曬,沉,朱,壑,泰,浩,昌,蓮,複,檢,
暴,錄,輸,批,圈,登,沿,腸,縱,燃,恨,糧,副,優,閉,捨,縮,芽,映,胸,
競,稅,遲,灰,置,割,牧,劃,柱,納,恩,幕,述,諒,延,混,宇,異,略,飼,
幼,遺,朗,徒,松,席,倍,波,殘,閣,穴,擴,貿,潮,寺,潔,帳,准,裁,賓,
測,拾,卷,委,倉,勉,純,姿,繪,梯,版,額,予,蒸,筋,誌,宅,譯,佛,返,
盟,協,瞭,域,胃,脈,宙,鋼,徑,均,岩,屆,構,荀,葡,航,襯,綿,鍛,礦,

貸,衫,鉛,肺,旗,階,櫻,障,搞,署,昔,液,刊,羨,瞧,罰,羞,悔,憑,趁,怡,召,抖,陷,膀,某,抵,痕,塞,糟,郊,託,愉,泥,誇,謀,仇,擁,塗,虧,溜,迫,贊,憂,拒,兇,碎,倆,釋,孤,稍,幅,戚,碌,氛,悠,勵,扶,攤,沾,趟,翅,牽,扣,柔,飄,穫,慰,擋,戒,佔,掩,溪,硬,隔,歉,摔,溝,糊,辨,夾,瞪,赤,括,埋,滑,逼,踩,怨,謙,閱,噴,抽,縫,牢,懶,鋒,插,跌,儉,淹,諧,撿,疲,腰,唯,鑽,邀,繞,繫,控,寵,咐,吩,擦,撲,睜,盆,籠,妨,淋,跪,觸,宴,儘,爐,伙,焦,彼,播,援,漫,拼,跡,輔,綁,駕,豪,逢,漲,哲,詳,仿,症,串,捲,琴,絡,炭,炮,距,搜,巨,執,描,謹,欲,灘,悄,違,蜜,慎/愼,棍,渡,軟,陶,揚,刮,皆,划,擾,添,嫌,盲,爽,撈,隆,潑,妝/粧,恢,訝,凍,療,冠,挨,拳,銷,斜,拖,棉,伍,癱,擠,幣,遵,培,繁,企,摩,托,檣,粥,臂,啣,鍊,哼,粽,肩,暫,乏,匆,瓦,贈,薄,俱,塑,粒,戀,捐,姻,膠,臥,怖,傻,鵝,咳,蔬,犬,胎,拆,抄,孕,估,倦,屁,喉,嚙,勿,窄,媒,賭,鈔,尿,皂,核,券,襪,闖,肌,駛,腫,謊,膩,鬍,暈,檔,坪,餵,哦,舟,凸,峰/峯,尼,凹,鑰,晦,刪,簿,菸,叮,酥,几,碟,茄,蹟,楣,葦,痠,飆,滷

C1 級(600 字)

君,蒂,妮,蓓,秦,臣,獻,魯,龜,喻,雁,曹,氏,耕,奉,賢,柴,賦,蒙,鴉,頗,蘇,陵,祈,侯,梁,呂,蜀,斯,昭,韓,郭,范,萱,彥,鳩,禪,勁,屈,廷,乃,欸,鵲,蔣,垂,鳴,枚,株,卵,械,愧,慚,撲,恍,耀,諷,憤,辱,耍,侮,慈,壽,惑,奪,饒,灌,袍,酬,哄,牲,徹,殿,翠,慨,伏,荒,丘,奴,擅,疏,葬,堪,妒,撫,扎,掙,誘,闊,浮,惹,潛,奈,奢,侈,翔,嫉,懇,惠,賤,抹,亦,碧,楷,促,範,驅,督,籍,襲,倡,犧,索,貢,患,役,艱,秩,隱,賴,循,駐,挺,緒,崇,倫,轟,泣,啟/啓,霜,憲,燭,疊,啼,碑,璃,隸,臘,奧,捧,謠,玻,黏,轎,肖,揀,蟬,譬,畏,采,伐,佈,盪,叛,紮,坡,卑,凱,粵,芝,韻,竿,僕,閔,淑,椿,雛,蔔,蘿,吏,埔,吾,洛,盧,篆,曰,兮,烽,硯,董,褒,播,鉢/鉢,殊,旋,措,遣,振,懸,脅,逮,忌,股,毅,貫,騰,悅,佑,衡,兼,旺,斥,津,逆,裂,巡,拋,暢,奠,慘,狹,肅,擬,滋,撥,妄,魂,毀,頌,儲,拓,柄,聘,賠,泛,荷,辯,栽,占,涉,傍,趨,祕,迅,渾,猶,逝,挫,撒,嚮,涵,揭,殖,墳,衷,胞,緩,訊,監,矩,懈,躍,側,懼,憶,撐,礙,挽,呈,凝,腔,詢,鬱,顛,紋,灑,審,誼,刑,撓,披,喚,浸,陌,傾,催,洪,勻,蓄,輝,脾,筒,緞,漏,寂,寡,耗,鈴,咱,歪,蕩,掏,瑰,蓬,霧,糾,衰,契,晃,

搏,釀,疾,坦,榜,庸,塵,盼,誓,潰,匯,娛,頒,允,覆,蹤/踪,譽,劑,縛,
掀,喘,銳,踐,撒,攔,銘,稚,譜,併,悉,裳,誦,燦,掘,廊,袖,丹,湧,勾,
宏,彙,歧,眠,劣,辰,顫,湊,倚,躬,脂,迴,朽,幟,惶,煞,寢,禦,懊,茂,
肝,澤,遞,憾,濫,煤,償,嶼,截,墊,脩,沮,淘,欄,兆,逗,酷,析,釘,龐,
蹲,熬,綜,販,塌,玫,歹,蠟,摘,焰,碩,濁,卓,夥,黎,扯,壇,滲,債,駝,
狠,軌,芒,汰,晝,梳,裔,鞠,邁,惋,赴,飢/饑,騷,嗜,締,攪,攝,募,押,
吻,曝,蹈,淪,揉,躁,茫,妥,燥,辜,腹,竊,豫,攔,疫,頑,晰,塔,旬,崩,
滯,鍵,枉,蠻,桶,衍,敷,跨,籬,耽,禿,颯,逍,磚,俊,摯,凌,沫,勒,怠,
艘,惕,盜,膏,歇,糙,蠱,旨,獄,澈,嫂,堡,抑,僱,弊,諾,輯,乞,豔/艷,
拐,遮,趴,註,脹,膨,磅,漆,剔,眨,峽,洽,抉,脆,盯,削,冤,菌,瀑,寞,
屍,裕,釣,瞞,噪,蜂,薯,罩,攀,沸,砂,踴,漠,售,殼,魄,剝,毯,嗓,幻,
裏,澆,樞,尫,偵,瀟,蛙,履,魅,癢,涕,笛,撕,廂,嗽,瞎,臟,溶,癌,堵,
啞,肪,囉,后,繳,咦,頻,枯,媳,吼,脖,稀,蒜,旱,甩,頸,嫩,秘,嗦,叨,
嘮,恆,聾,霉,雇,氧,斗,薑,詐,廈,艙,痴/癡,蛀,棕,甥,檬,檸,菠,楓,
蔥,婿,蚵,唇,柚,咪,碳,皎,芹,魷,餡,睏,痣,鈍,鯢,噁,遛,糗,賄,虐,
框,沐,痘,靴,嬌,賂

C2 級(600 字)

狐,狸,魏,諸,叢,捆,賜,昂,蒼,祀,遜,遼,彌,嚷,梭,曠,蠶,冥,瞬,馴,
陡,峭,糞,摧,矚,卒,鑼,盞,慷,淆,潤,穎,虔,繡,竭,蔽,囊,溉,艦,醇,
纖,驟,闕,擎,煌,暮,稠,鞏,勃,巢,爍,渠,蝕,籤,囑,惚,廓,鈍,瀟,萎,
蝶,淒,袂,皿,摟,髓,壤,扭,搓,軸,醞,昆,魁,瑣,穀,剎,邏,蝴,喇,擲,
濺,撰,伺,緝,哺,瓣,紡,墅,綱,墟,翹,傢,捏,鑲,姆,崖,厖,睦,吊,叭,
挪,丐,剖,覓,乒,乒,坊,侶,墮,隧,籲,凋,鯨,篷,蝸,爾,恥,焚,征,籌,
豈,逸,纏,庇,彷彿,拘,豎,哩,禽,藤,怯,拂,濤,吟,樑,玄,貞,凶,俯,苛,
屑,撇,漓,邪,呵,邦,襟,紗,睹,輿,嘿,嘻,窺,禱,彰,沛,匪,侍,垢,屢,
蕭,蔑,殷,鳴,嘩,沃,韌,淵,矯,搗,帖,涯,塘,傑,汲,咕,扒,扛,嬉,偽,
暇,鶴,煥,棚,黯,墜,杖,嚼,脊,嘉,熄,嶄,攏,囚,譏,蔗,咽,攜,壘,槩,
溢,棗,葛,輾,溯,弘,潭,媚,屏,荊,泊,礁,杼,唾,霸,簾,畜,凰,覓,聳,
絞,羹,蔭,埃,灼,瘡,斬,秉,咒,奸,繭,猿,丑,綻,坑,駁,蟹,訣,蘆,伶,
俐,敘,瘟,鍾,竄,罕,瑞,腕,昧,澎,軀,鶯,撩,嶺,霞,窟,歎,矢,嬌,蹺,
炫,畔,粹,喧,咀,萍,藻,卸,懦,鄙,汪,翩,隙,轟,湃,壁,矇,呻,耘,砸,

膜,俠,陋,僵,亭,肆,毆,膝,綴,頰,嘍,崇,烹,屎,恕,叩,綵,俘,蛛,迪,
齋,頰,劫,兜,誨,懲,僻,翼,隴,朦,拱,濱,慾,拭,澄,輻,鬚,渣,秤,褪,
鑄,嗅,錫,虜,澳,紳,逞,齣,敦,顛,柏,疆,栗,肢,畸,棘,猥,沼,堤,芋,
狡,眨,吝,馨,菊,瑕,萌,徊,徘徊,猾,酌,洵,敞,疵,絮,婉,竅,溺,蹄,摺,
帕,疚,傭/佣,豚,荔,甄,杏,敞,憬,濛,舵,豁,懂,樵,篩,菲,啄,嘲,綽,
佐,芭,拙,簪,桑,弦,螃,寧,鑑/鑒,聆,筍,詭,刨,滔,姦,鋁,壘,髦,鈣,
紊,撼,賊,薦/荐,劈,瞄,莖,銜,稿,馳,烱,箏,螢,棺,帆,兌,爵,愆,愚,
繃,掠,憩,瀕,辮,眺,棲,瀉,膊,弧,倖,槽,桿,悼,眷,攬,烘,盈,淤,坤,
弛,蚬,喬,斃,絆,紐,淇,阱,岔,螂,虹,嗇,傅,鏟,缸,苔,妖,蹦,霍,疤,
纜,鱗,琢,屠,趾,眯,雌,瞋,櫛,奕,揣,跛,嘯,蕾,續,娟,裸,莽,瘠,疙,
獺,猖,嘔,挾,姪,釐,灸,鉅,卦,淫,拯,橡,氓,刃,橙,茁,瑩,艇,函,藹,
諮,恤,耶,駭,俄,柵,崗,釁,砲,糯,祿,砥,礪,瀾,癱,瘓,癒,訕,徨,徬,
坎,坷,誣,洩,鴛,鴛,噓,汜,燉,鰾,傀,忿,唆,窘,斟,贅,茱,轍,噠,癖,
訟,悴,焦,輟,寨,潦,勘,覬,覷,赦,殃,餒,舔,諂,腎,彬,疹,飪,舖/鋪,
紆,饋,侑,餽,饗,逾,磁,鮑,鮭,罹,轄,瘤,迄,嗆,甦,醇,匿,輒,鈕,睫,
賃,腥,諜,譴,蠅,謗,蚤,誹,痊,妓,蟬,咎,痰,窰,醢,誨,膿,燴,餌,涮,
僥,癢,渲,痔

A Comparative Study of the NAER Chinese Character List and Other Chinese Character Lists

**Shu-Ling HUANG Ming-Hong BAI Jian-Cheng WU
Shih-Min LI Ching-Lung LIN**

**Research Center for Translation Compilation and Language Education,
National Academy for Educational Research**

Abstract

A Chinese character list contains proficiency levels and corresponding character sets, which are regarded as indicators of one's reading and learning ability. The character list will tend to vary depending on the needs of learners. Over the years, officials and researchers have published different Chinese character lists for native speakers as well as Chinese L2 learners in order to meet the needs of designing textbooks, teaching or grading. As there is no official Chinese character list in Taiwan, in 2013 the National Academy for Educational Research (NAER) implemented an 8-year project with the Ministry of Education to develop corpora and a standard system, which included creating a graded Chinese character list for L2 learners in Taiwan. To evaluate the character list, this paper looks at 7 representative character lists for the purpose of comparative analysis. The study shows that the NAER's character list ranks highest in performance in terms of corpus coverage, especially at the primary level. Its average score was 4 percentage points higher than that of the European, Hong Kong, Singapore, and Chinese Hanban character lists. In terms of the character set contents, the NAER list is the most representative Taiwanese character list, and here the locality is taken into account in evaluating and grading the character selection. The final grade is based on the all of the above factors, included the details of each step in the creation of the character list. On the whole, the NAER character list is objective, reproducible, local, word-based, and meets the needs of second language learners.

Keywords: Chinese character grading, Chinese character list, standard system,
Chinese language, pedagogy