

解碼語文學習：課堂到生活的大數據



圖片來源：Pixta

【語文教育及編譯研究中心副研究員兼中心主任 吳鑑城】

隨著數位時代的來臨，學生的語文學習不再僅依賴於教科書，課外讀物已成為他們獲取知識的重要來源。然而，教科書與課外讀物在語詞使用上是否存在差異？這是一個值得深入探討的問題。

閱讀理解的關鍵：詞彙覆蓋率

研究顯示，詞覆蓋率對閱讀理解有重要影響。詞覆蓋率是指文章中所有詞彙中，讀者已經學過的比例。當學生在閱讀材料中遇到過多未學習過的詞彙時，閱讀難度會顯著增加，甚至可能導致挫折感。一般而言，當學習者已掌握文本中 95% 的詞彙時，便能夠進行一般理解；若掌握 98% 的詞彙，則能輕鬆理解。相對地，僅掌握 80% 詞彙時，只能大略了解重點。這意味著，當學生在閱讀課外讀物時，若遇到許多在教科書中未曾學過的字詞，理解和掌握這些讀物的難度將大幅提升，進而影響學習效果。

教科書與課外讀物的語料分析

為了了解教科書與課外讀物的語詞差異，本研究蒐集了豐富的語料進行分析。教科書部分涵蓋了國語文、數學、社會和自然科學領域的教材；課外讀物則包含了文化部好書推薦、教育局優良讀物及各大圖書館推薦的書籍，以及國家教育研究院臺灣華語文口語語料庫和網路影音平臺上的節目。

課內外語詞使用差異與挑戰

由於十二年國教課綱教科書尚未完全編制完成，因此本次分析以九年一貫課程綱要的教科書為主。隨著十二年國教課綱的逐步實施，未來相關研究將可涵蓋更完整的教材語詞內容，為教科書與課外讀物的對比分析提供更全面的依據。

研究分析顯示，九年一貫課程綱要中累積至六年級的相異字數如下：國語文約 2,650 字、數學 1,800 字、社會 2,600 字、自然科學 2,200 字。相比之下，在課外讀物中，以第三學習階段（國小五、六年級）為例，國語文領域包含 4,150 個相異字與 23,399 個相異詞，社會領域有 4,443 個相異字與 32,992 個相異詞，自然科學則達到 3,981 個相異字與 24,990 個相異詞。這些數據顯示，課外讀物的用字量遠遠超過教科書，並呈現出更為豐富的語詞組合。

在字詞覆蓋率方面，研究結果指出，第三學習階段時，小學教科書所學字詞對課外圖書的字覆蓋率超過 99%，詞覆蓋率約為 85%；而在媒體內容中的詞覆蓋率甚至超過 90%。這表明，在這一階段，教科書已涵蓋了學生日常閱讀所需的大部分常用字詞，能基本支持他們的課外閱讀需求。然而，在第一學習階段的字詞覆蓋率仍存在不足。字覆蓋率僅約 90-94%，而社會和自然科學領域的詞覆蓋率甚至不足 70%。這意味著低年級學生在閱讀課外讀物時，可能因未學過足夠的詞彙而面臨較大的理解挑戰。

多策略應對語詞學習挑戰

基於研究發現，為有效提升小學生的語文學習效果，可以採取多層面的策略來應對課內外語詞差異問題。首先，可編製常用分級字詞參考表，根據學生不同年級和學習階段的需求，精確反映其詞彙掌握情況，並作為教科書和教材編寫的參考依

據。此舉能幫助學生在課外閱讀中理解更多字詞，減少困難，增強閱讀信心與興趣。同時，設計補充詞彙教材，例如詞彙卡或簡易讀物，幫助學生吸收課外讀物中的高頻詞彙，逐步建立穩固的詞彙基礎。

此外，可推廣導讀與共讀活動，透過學校與家庭的協作，讓學生逐步接觸更具挑戰性的課外讀物，並在互動中提升理解能力。為進一步支持語文學習，還應運用數位學習工具，例如語言學習應用程式，提供遊戲化的詞彙學習模組，吸引學生自主學習，促進詞彙量的擴展。

最後，建議導入教材編輯輔助系統，結合自然語言處理和數據分析技術，精確分析教材內容並將結果應用於教材編寫過程中，以確保內容更加符合學生的學習需求並提升整體教材品質。這些相輔相成的措施將幫助學生克服語文學習的挑戰，促進閱讀能力的提升與全面發展。

資料來源

吳鑑城（2023）。小學生課外讀物及學習表現之語料庫建置技術發展及常用字詞分析研究。國家教育研究院研究計畫，計畫編號：NAER-2022-022-C-1-1-F1-07。執行日期：2022-03-01 至 2023-12-31。連結網址：

<https://www.grb.gov.tw/search/planDetail?id=14896215>