

從詞彙覆蓋率比較國際間英語教科書詞彙的使用



圖片來源：Pixta

【語文教育及編譯研究中心副研究員 白明弘】

在閱讀過程中，讀者通常會遇到幾個不認識的單字，這些單字會阻礙理解。如果不認識的單字比例越高，閱讀的阻礙也就越大；反之，認識單字的比例越高，閱讀理解也就越順利。認識單字比例可視為心理詞彙對文本的覆蓋率。例如，若學生在一篇包含 57 個字的文章中標出 5 個不認識的字，那麼覆蓋率為 91.2%。這意味著學生的心理詞彙可以覆蓋文章中 91.2% 的單字。

Alice was beginning to get very tired of sitting by her sister on the bank, and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations in it, "and what is the use of a book," thought Alice "without pictures or conversations?"

節錄自《Alice's Adventures in Wonderland》

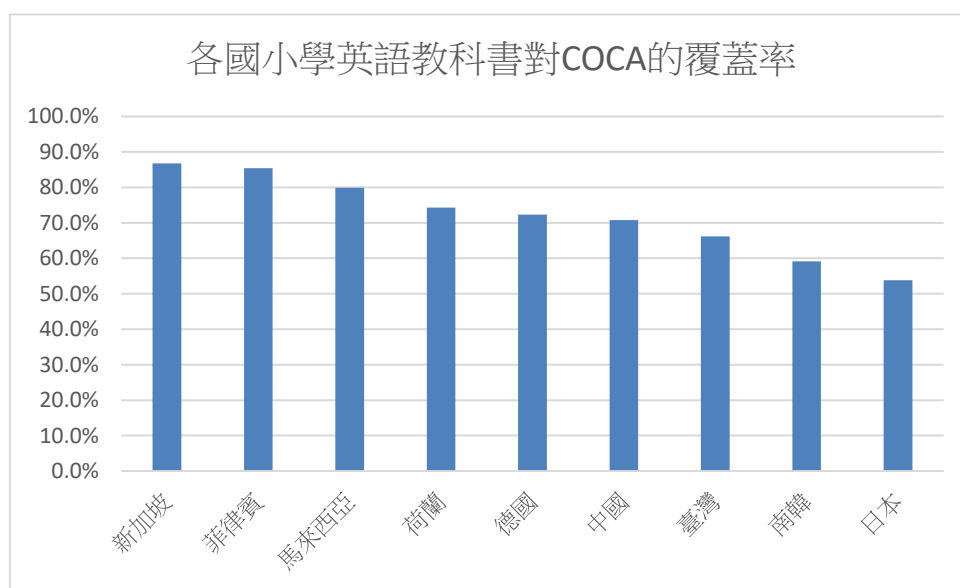
一般而言，覆蓋率在 98% 以上的文本閱讀起來流暢愉快，適合作為學生的課外讀物。覆蓋率在 90%-98% 之間的文本有一定難度，適合作為課堂教材。而覆蓋率低

於 90% 的文本會造成理解困難，使學生感到挫折，故較不適合作為教材。

詞彙覆蓋率還可以作為衡量教科書詞彙與實際語言使用之間差距的指標。當教科書的詞彙能夠覆蓋大部分實際語言使用的情況時，表示教科書所包含的詞彙已經非常接近實際語言使用。這種方式可以評估該課程所包含的內容是否能夠支持相關的教學應用。例如，若要將英語作為其他課程的授課語言，就需要評估目前學生的英語詞彙量是否能夠應對其他課程所需的詞彙。

在我們的研究中分析了 9 個國家的小學英語教科書中的詞彙對 COCA 語料庫的覆蓋率，藉以了解這 9 個國家的小學英語教科書和英語實際使用上的差距（如圖 1）。從圖中可以看出，這些國家的覆蓋率大約可以分成三組，第一組為新加坡、菲律賓及馬來西亞，這三國英語教科書詞彙對 COCA 語料庫的覆蓋率都已經達到 80% 以上。第二組為荷蘭德國及中國（一線城市）¹，覆蓋率大約 70% 出頭。第三組為臺灣、南韓及日本，覆蓋率約在 55%-65% 左右。

圖 1、各國小學英語教科書對 COCA 語料庫詞彙覆蓋率比較



上述的覆蓋率是在未扣除專有名詞、數詞及年份等教學中不必教的詞，這些詞大約佔語料庫的 10%-15% 左右。所以如果扣除這些詞，第一組的實際覆蓋率超過 90%，尤其是新加坡的小學教科書，扣除專有名詞等不需教授的詞彙後估計可達到 95% 以上的覆蓋率，已經和語言的實際使用差不多。我相信這和各國的教育政策有極高的相關

¹ 中國的小學英語教科書版本為一線城市的版本，起始年級為小學一年級。

性。以下是幾項值得深入了解的相關性：

一、覆蓋率和英語程度的相關性

EF EPI (The EF English Proficiency Index) 指標的排名順序幾乎和英語教科書的詞彙覆蓋率排名一致。這代表詞彙的覆蓋率和英語能力有極高的相關度。但這種相關度不能解讀為因果關係，也就是說不能認為單純提高教科書的詞彙量，就能達到提升學生程度的目的。其中還牽涉到起始年級及教學時數。

二、覆蓋率和起始年級及教學時數的相關性

覆蓋率最高的三個國家：新加坡、菲律賓、馬來西亞英語教學的起始年級都是小學一年級。覆蓋率次高的荷蘭及德國，雖然起始年級和中國、臺灣及日本相當，但是每週時數實際上高於中國、臺灣及日本。所以，由此也可以呼應前一段所說，不能認為單純提高教科書的詞彙量，就能達到提升學生程度的目的。起始年級及教學時數可能和英語程度有更直接的相關性。因為當詞彙量增加時，如果沒有增加相應的教學時數，只會增加學生的課業壓力和挫折感。

三、覆蓋率和雙語教育的相關性

覆蓋率最高的新加坡、菲律賓及馬來西亞都是從小學就開始部分課程以英語進行，尤其是新加坡，除母語和品格公民教育課程外，所有課程皆以英語進行。如果學生沒有相應的英語能力，將影響到學生的學習效果。所以新加坡的覆蓋率最高，其次菲律賓及馬來西亞，也都是在小學就實施部分課程以英語進行。所以覆蓋率相對比較高。而德國及荷蘭則是在中學才開始施行雙語教育，所以，在小學對英語詞彙的要求較低，詞彙覆蓋率也就相對較低。至於中國、臺灣、南韓及日本，目前還沒有普及的施行雙語教育，所以對於詞彙覆蓋率的要求相對更低。

資料來源

白明弘 (2023)。應用自然語言處理技術於英語教科書之分析比較研究。國家教育研究院研究計畫，計畫編號：NAER-2022-018-C-1-1-F2-05。執行日期：2022-07-01 至 2023-12-31。連結網址：<https://www.grb.gov.tw/search/planDetail?id=14891235>