

語詞也需要做普查！瞭解語詞常用度的用處

【語文教育及編譯研究中心助理研究員 吳鑑城】

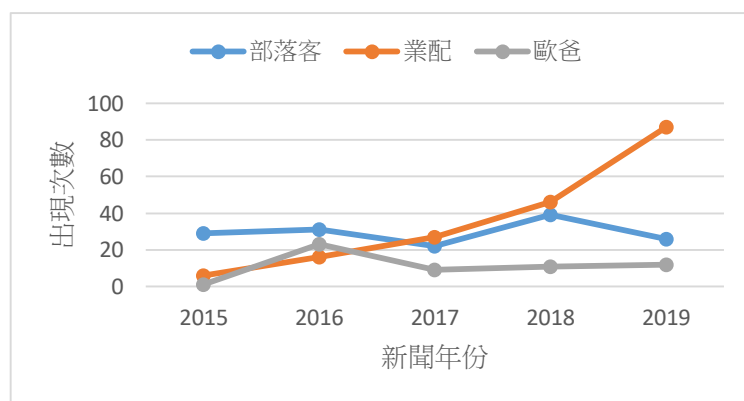
「哇，這個產品，有位部落客推薦過，說 CP 值很高耶！」「那應該是他接的業配，不要太信以為真啦！」

「今天在機場，聽到一大群人超熱情地大喊歐爸，那個姓歐的人好受歡迎，他是誰呀？」「拜託！那個韓星最近超火，有著一大票的粉絲，重點是，他可不姓歐，真的很 Orz 耶！」

以上的對話，可能就出現在你我的日常生活周遭。然而，「部落客」、「CP 值」、「業配」、「歐爸」、「粉絲」、「Orz」等可能從未見過，或僅有依稀印象的語詞，究竟是自己「孤陋寡聞」還是它們本來就是罕用的特殊用語呢？換個角度思考，個人熟悉的語詞是真的廣泛被使用或恰巧僅是自己常用呢？

其實，語言是用來溝通的一種媒介，除了如連接詞、介係詞等較為穩定的功能詞（function words）外，隨著不同時代新興概念、事物的出現，常常使得如名詞、動詞等內容詞（content words）會出現新語詞或是新詞義的誕生，也促使語言本身不斷地演變。這類新興的詞彙、詞義除了出現於日常生活之中，例如「業配」、「歐爸」（新興詞彙）、「打卡」、「粉絲」（新詞義），當然也存在於各個專業領域之中，包括各個新興術語。而這些新興的詞彙雖然有些會「曇花一現」，被淹沒在大量的詞海之中，但也有許多語詞因為能確切地被應用於生活或是工作情境之中，並透過各式媒介如電視、報紙、網路、教材等廣為傳播，無形中成為當代語言中常用的語詞。

圖一呈現了「部落客」、「業配」，以及「歐爸」在各年度新聞中出現的次數情形，跟你想像的是否一致？而瞭解語詞的使用度，究竟有甚麼用處呢？



圖一、各詞彙於各年度新聞（含《中國時報》、《聯合報》、《蘋果日報》、《大紀元時報》）中出現的次數

常用語詞調查就像進行語詞界的人口普查一般，透過瞭解語詞常用性、在各種文本中的分布情形等，可作為許多研究或實務相當重要的參考資訊。以下列舉兩種常見之應用：

一、**教材編撰及課程設計**：教學者或教材編纂者可利用語詞常用資訊，協助教材編撰以及課程設計等任務，例如以語詞常見程度作為教學階段學習內容的規劃參考，提升教材的合適度。

二、**辭典編輯**：隨著語詞的持續誕生，在難以盡收的情況下，辭典編纂者可根據語詞常用性，作為收詞優先性的考量。試想，當有兩個都需要辭典提供釋義的語詞，優先提供常用度較高的語詞當能更符合使用者查詢語詞的需要。

除上述應用外，若累積各年代常用語詞資訊，我們還可藉以發現語詞的興衰，進而探究其背後之原因，或許為語言文字使用習慣的改變，也或許為社會變遷之影響；有特定學科領域的常用語詞資訊，學習者就可以瞭解該領域之常見專業術語或是陳述時的習慣用語，使其更能儘快融入該領域的表述方式，展現其專業度，語詞調查的應用，實可謂是不勝枚舉。

那要如何進行常用語詞調查呢？現今研究者大都採用語料庫為本的方式。語料庫是一或多個為了特定目的而收集，具有代表性、結構性，且有一定規模的語言資料庫。由於可以呈現較接近真實的語言使用情境，語料庫被視為記錄語言使用的重要資料庫。而常用語詞調查基本上可分為三個階段：（一）確認常用語詞調查之目標；（二）依目標設計適合的語料庫建置目標；（三）語料收集、處理及分析。而在進行語詞調查時，若僅靠人力進行語料處理，明顯地耗時費力，並會限制調查所使用的語料量，進而影響調查成效。所幸隨著科技之進步，傳統以人工方式進行文本分詞（將連續的中文字串切分為個別有意義的語詞），以進行後續統計分析的過程，現已可透過電腦自動化處理，明顯地提升了調查效率，也使得調查的範圍可以大幅地提高。

有鑑於常用語詞調查的重要性，本院近年來也陸續完成《九十九年常用語詞調查報告書》、《一〇〇至一〇四年常用語詞調查報告書》、以及《解讀新聞字詞：新聞媒體常用語詞調查（105-107年）》等調查成果供各界參考。冀能透過厚實知識基礎建設，為各式研究或實務奠定良好的基石。

資料來源

吳鑑城（2018）。常用語詞調查之研究初探。國家教育研究院研究計畫，計畫編號：NAER-107-12-F-2-10-00-1-10。執行日期：2018-05-01 至 2019-04-30。連結網址：<https://rh.naer.edu.tw/handle/gq845>