

## 《臺灣華語文語料庫—華語文教與學的必備工具》

### 教學效能再進階



臺灣華語文語料庫為華文正體字中建置類別最完備、語料量新且最多的語料庫。

本書為該語料庫建置的部分成果及繼中研院詞頻辭典的通用詞彙後最新的詞頻表，讀者可了解語料庫所使用的資料。

本語料庫結合臺灣華語文能力基準，建置「臺灣華語文語料庫與能力基準整合應用系統」讓大家更方便運用，華語文教學效能再進階。

圖片製作：策略溝通辦公室

【語文教育及編譯研中心 林慶隆、盧昱勳】

《臺灣華語文語料庫—華語文教與學的必備工具》主要的內容在介紹「臺灣華語文語料庫」(Corpus of Contemporary Taiwanese Mandarin, COCT)，以及應用語料庫所建置的通用詞頻表。

「臺灣華語文語料庫」是目前華語文正體字建置類別最完備、語料最新且數量最多的語料庫。通用詞頻表則是繼中研院詞頻辭典的通用詞彙後，最新的詞頻表。教育部為了持續提升華語文教育的品質及效能，2013 年將華語文教育八年計畫(2013-2020 年)的「建置應用語料庫及標準體系」計畫，請本院執行。本書就是建置華語文語料庫的部分成果。

華語文漢字、詞語及語法點分級標準，若能植基於語文實際使用的情境，則更能符合華語文教學的需要，而且必須是有適合的語料庫可使用。然而，「中央研究院漢語平衡語料庫」以語言學研究為主，且語料僅維護至 2007 年；國內其他單位的語料庫

也是以學術研究為主，較缺乏針對華語文教學的需求。

因此，本院以「美國當代英語語料庫」(COCA)為目標，進行臺灣華語文語料庫的建置。至於通用詞頻表，中研院曾於 1998 年研訂中研院詞頻辭典，內容包括中文核心詞彙及通用詞彙，不過至今也有 20 年，亟需運用臺灣華語文語料庫中，最新且大量的語料來研訂新的通用詞頻表。

這本書的內容主要摘錄自該計畫的歷年成果報告，再經研究團隊多次討論修訂才定稿。其成果除了可應用於建置華語文標準體系，更可使用在華語文、國語文、英語文等教學設計、教材編撰、測驗評量、辭典編撰、人工智慧及相關語文研究，提升語文教育的效能。

此外，本語料庫也結合「臺灣華語文能力基準」(TBCL)，建置成「臺灣華語文語料庫與能力基準整合應用系統」(網址：<https://coct.naer.edu.tw>)，是科技整合應用於華語文教學的新發展，歡迎有興趣的民眾可上網使用。