

# 兒童語音評估：跨領域的語言學研究

曾淑娟\*

## 一、跨領域語言學

「語言學」是一門研究語言的學科；語音學(phonetics)、音韻學(phonology)、構詞學(morphology)、語法學(syntax)、語意學(semantics)與語用學(pragmatics)普遍被視為語言學核心研究領域。各核心領域皆有學者共同認定的理論框架，用以進行語言各階層現象的關聯性探究與分析。針對各種語言，研究的標的可以從個別的音(如ㄅ、ㄆ、ㄇ)、字詞(如「花」、「我」、「蝴蝶」、「葡萄」)、句子(如「我喜歡吃蛋糕」)，到言外之意的解讀(如「你有帶錶嗎？」並非期望「有或沒有」的回答，而是想知道「現在幾點」)。在二十年前，計算與語料庫語言學(computational and corpus linguistics)、心理語言學(psycholinguistics)、社會語言學(sociolinguistics)等研究，與資訊工程、心理學及社會學研究方法結合，以語言特徵審視數據、認知與文化的表現，多被視為跨領域研究，而非核心語言學研究。然而語言學發展到今天，語言涉及的計算、認知、社會、文化等各方面要素都已知是同等重要。以語音學家研究發音構音生理參數與音韻系統抽象特徵為例，不僅要使用儀器(如超音波、電子聲門儀、電磁構音儀等)與統計方法(檢定參數相關性及貢獻度)，設計語料也要考慮語者語言經驗(參照語料庫頻率及詞語／音節共現機率)、語者即時的語言情境(如正式場合或日常生活語境)以及語者背景(如社經及語族差異)。

近年更因語言科技的興起、AI的大量使用、多語社會及社群媒體的語言活力，人類語言習得與機器模型學習語言相互交織，大眾對於語言議題的興趣提升，人人都能成為語言的研究者。在此番浪潮裡，如何界定語言學的核心領域、次領域與跨領域，以及大眾期望語言學家能解答的議題，都在持續地變化。單一領域的基礎理論訓練已不足以回應當今研究生態。作為一門探究語言本質的科學，語言學可以對「學」與「術」都做出什麼貢獻呢？說話、寫字、使用語言是現代社會社交必備的技能，語言科技研發脫離不了語言學專業。但

\* 中央研究院語言學研究所研究員、人文處語言學門共同召集人



織階層直接或間接影響整體韻律配置。評估語者口語表達能力，單詞與連續語音兩種語體都應該納入。然而文獻指出臺灣臨床語音異常評估使用連續語音仍是未普遍使用的評估項目（張晏寧、葉麗莉，2019）。一般而言，音段發音的正確率超過 75%，視為已進入該音段的習得歷程。若達到 90% 即為通過習得門檻。華語、粵語語音習得相關的研究不算少，但數據及詞表設計則各不相同（Li & To, 2017）。以英語為例，較為完整的語音發展習得歷程，很多還是參照 1972 年資料所製作的數據（Sander, 1972）。應用語料庫語言學研究方法，兼顧語者、語料、標記品質的一致性與可信度，所製作的數據資料，是連結跨領域合作重要的媒介。

#### 四、中研院學齡前兒童語音習得常模語料庫

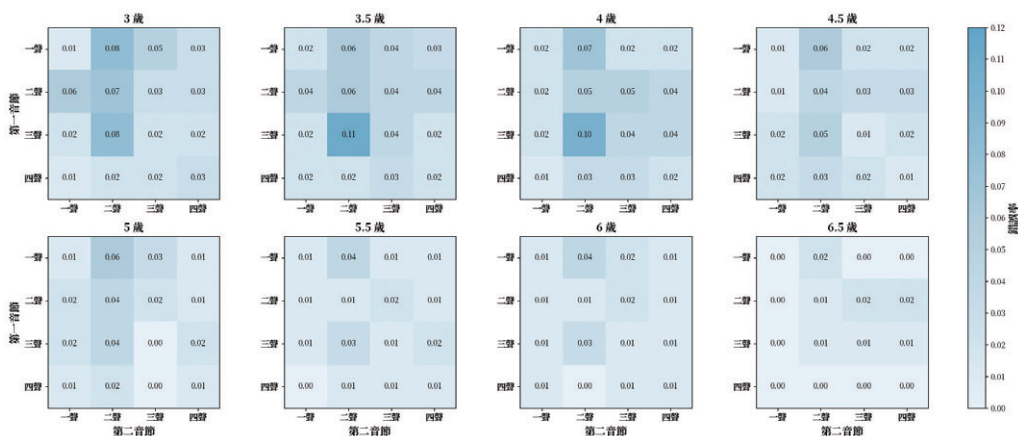
本文介紹中研院語言所建立的大型語音習得語料庫。收集臺北市及新北市 3 至 6 歲兒童的臺灣華語發音樣本計畫始於 2017 年。首先參照成人對話內容所計算的音節組合頻率，<sup>1</sup> 製作「中研院兒童語音平衡詞表」（見表一），兼顧詞語的音韻結構、聲調組合、語意、詞性等特徵，以利系統性習得數據的計算。最終取得 798 名通過本計畫聽力篩檢兒童，合計共約 12 萬字的語音樣本（曾淑娟，2024）。經過聲調聽感正確性以及詞語實際發音的標註與可信度檢驗後，建立分齡習得歷程。這份資料既可作為習得研究素材，也能提供學習模型作為訓練資料。

表一：中研院兒童語音平衡詞表（曾淑娟，2024）

動物	食物	交通工具	身體部位	動作	物品	遊戲	地點	景觀	節日
天鵝	牛奶	火車	牙齒	吃飯	皮鞋	大富翁	花園	月亮	生日
母雞	牛排	汽車	耳朵	走路	吸管	吹泡泡	客廳	白雲	
老虎	芒果	飛機	嘴巴	爬山	玩具	足球	噴水池	斷崖	
老鷹	果汁			穿衣服	時鐘	拼圖	廚房		
兔子	草莓			掃地	茶杯	積木	學校		
刺蝟	甜甜圈			淋雨	彩色筆				
恐龍	蛋糕			游泳	窗戶				
烏龜	葡萄			買菜	鈕扣				
蜘蛛	壽司			睡覺	溫度計				
蜜蜂	熱狗			說話	筷子				
螃蟹	饅頭			寫字	電視				
醜小鴨	蘋果			騎馬	盤子				
				關燈	輪胎				

<sup>1</sup> 中研院中文口語語料庫檢索系統：<https://tmc.ling.sinica.edu.tw/index/>。

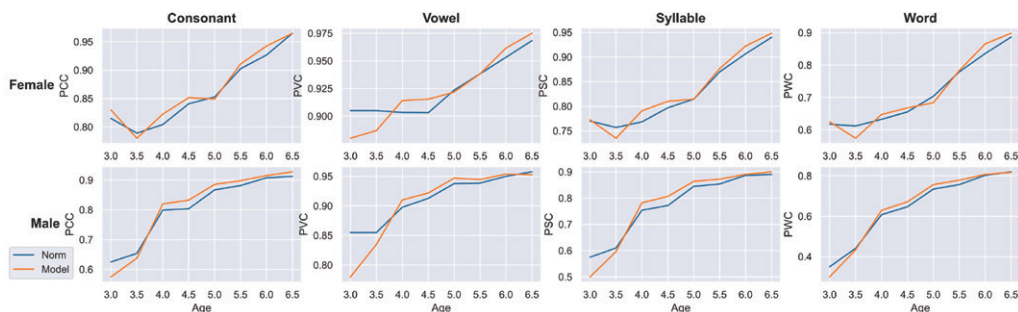




圖三：臺灣華語學齡前兒童雙字詞聲調錯誤率一分齡習得歷程

## 七、自動音段辨識模型與語音評估系統

常模語料庫作為語音習得數據，也可製作類似身高、體重百分位的常模對照圖表。再以精準記音資料，利用開源預訓練的大語言模型予以細緻微調，透過語言學標記的聲學物理特徵強化預測模型效能。我們以 Whisper 與 Wav2vec2 訓練的學齡前兒童語音音段辨識模型，與人工標記答案比對，已可達到低於 5% 的預測錯誤率 (Lin et al., 2024)。中研院兒童語音習得語料庫的聽感標記包含「詞語」(words)、「音節」(syllables) 及「聲調」(tones)。個別的音可再區分為元音 (vowels) 與輔音 (consonants)。以這些語言學單位的正確率作為量化指標，可以評估兒童語音習得的進程，文獻多以 Percentage of consonants correct (PCC)、Percentage of vowels correct (PVC) 等表示。如圖四所示，音段辨識模型預測結果與人工聽感標記所計算出的評估指標 (PCC、PVC、PSC、PWC)，已經達到統計顯著的相關性。



圖四：模型預測 (Model) 與人工聽感 (Norm) 評估指標相關性 (Lin et al., 2024)

這意味著，我們有可能可以利用模型預測加速最初期的語音評估篩檢，讓專業從業人員將更多時間與心力用於更需要人力的語言評估及介入治療。由中研院及國科會經費支持，目前正測試線上工具，將串聯錄音、音段自動辨識、參照常模數據、整合評估模型預測結果，提供給大眾測試使用。也規劃在下一階段將詞表錄音擴增至語句錄音。以前面提及的音段與聲調為例，若口語溝通是以讓對方瞭解語意為主要目的，那麼評估自然語音的「正確性」，就不一定是要「完全符合標準發音」，而是要「能讓人無礙的瞭解語意」。建置此類語料庫可進一步探索如何適當地以自然的連續語音語體，評估孩童的口語能力。

## 八、結語

本文介紹的兒童語音評估跨領域研究，是架接在多年研究成果之上，才得以進行。這包含了成人對話語料庫、聽常聽損兒童語句與說故事語料庫的建置、核心口語詞頻表與音節結構頻率組合表的製作、自動語音文字對齊工具與兒童習得詞表收集系統工具的開發等。耐心的分階段探索人類自然語言的口語溝通議題，兼及語音形式、語言結構及語境等面向。從少量，但具有良好標註的語料庫，逐步進行細緻的語言學研究假設試驗。將語音學與語言治療、資訊工程結合，透過語料標記、聲學分析、機器學習等研究方法，擴大研究目的與應用的設定，將語言學理論、現象與社會活動連結。語言學仍然可以是探索人類語言本質的科學，只是我們需要考慮與不同領域的專家合作，而且應該更看向語言是作為溝通的這個本質。

## 參考文獻

- 張晏寧、葉麗莉 (2019)。〈台灣語言治療師執行語音異常評估之現況調查〉，《台灣復健醫學雜誌》47 卷 1 期，頁 31-47。
- 曾淑娟 (2024)。〈台灣華語 3 至 6 歲兒童語音習得跨域研究〉，收錄於曾淑娟、齊莉莎等著，《中央研究院語言學研究所成所 20 週年慶祝論文集》，頁 475-502，臺北：中央研究院語言學研究所。
- Li, X. X., & To, C. K. (2017). A review of phonological development of Mandarin-speaking children. *American Journal of Speech-Language Pathology*, 26(4), 1262-1278.
- Lin, Y.-S., Tseng, S.-C., & Jang, J.-S. R. (2024). Leveraging phonemic transcription and Whisper toward clinically significant indices for automatic child speech assessment. *INTERSPEECH 2024*, Kos Island.
- Sander, E. K. (1972). When are speech sounds learned? *Journal of Speech and Hearing Disorders*, 37, 55-63.