

從動態評量觀點比較教育型 AI 代理人與生成式 AI 對低成就 EFL 英文學習

者英語口說表現之影響

Pedagogical AI Agents versus Generative AI in Enhancing Low-Achieving EFL Learners'
Speaking Performance: A Dynamic Assessment Perspective

林智忠

嘉義大學語言中心

chihchunglin09@gmail.com

【摘要】 近年生成式 AI 廣泛應用於英語口說學習，但其成效高度取決於教學設計。針對低成就學習者，一般生成式 AI 與具教學導向之教育型 AI 的比較研究仍顯不足。本研究以動態評量（DA）為理論基礎，採 16 週準實驗設計，59 名低成就大學 EFL 學習者分別使用教育型 AI 代理人（具遞進式調節功能）與生成式 AI（ChatGPT）進行口說練習。結果顯示，教育型 AI 組在口說複雜度與任務完成度上顯著優於對照組，證實結構化教學支援能有效減少「認知卸載」，是提升低成就者口說表現的關鍵。

【關鍵字】 動態評量、教學代理人、生成式 AI、英語口說表現、低成就學習者

Abstract: While generative AI (GenAI) is widely used for English speaking practice, its effectiveness for low-achieving learners remains underexplored. This 16-week quasi-experimental study compared a pedagogical AI agent—designed with Dynamic Assessment (DA) principles—against a general GenAI system (ChatGPT) among 59 low-achieving university EFL students. MANCOVA results indicated that the pedagogical AI agent led to significantly greater gains in speaking complexity and task completion. The findings highlight that structured, pedagogically mediated support is crucial for fostering deep language processing and preventing cognitive offloading in low-achieving learners.

Keywords: dynamic assessment, pedagogical AI agents, generative ai, English speaking performance, low-achieving EFL learners

1. 前言

低成就 EFL 學習者常因語言資源不足，口說表現呈現片段且任務完成度低。一般生成式 AI（GenAI）雖能提供範例，但易導致學習者產生「認知卸載」（cognitive offloading）（Risko & Gilbert, 2016），直接依賴 AI 輸出而非自主組織語言。相較之下，「教育型 AI 代理人」（Lan & Chen, 2024）依據動態評量（Dynamic Assessment, DA）理論（Lantolf & Poehner, 2014），提供遞進式提示，引導學習者在最近發展區（ZPD）內進行自我修正，而非直接給予答案。本研究旨在比較兩者對口說表現（流利度、正確度、複雜度、任務完成度）的影響差異。

2. 研究方法

本研究採準實驗研究法，參與者為台灣某公立大學的 59 名低成就 EFL 大學生（口說能力低於 CEFR A2 且成績處於班級後 25%）。研究為期 16 週，實驗組（ $n=30$ ）使用依據動

態評量 (DA) 原則設計之教育型 AI 代理人，對照組 ($n=29$) 則使用一般用途之生成式 AI (ChatGPT)。在 AI 介入模式方面，教育型 AI 代理人採遞進式教學調節 (graduated mediation)，從隱性提示開始，視學習者反應逐步提供顯性支架，旨在促進語言延展而非直接替代產出；而 ChatGPT 組則反映自然使用情境，由學生主動請求示例或檢查語法，系統傾向立即提供完整且流暢的答案，未受階段性調節。所有參與者皆須完成六次任務導向口說練習，並接受前後測評量，包含意見表達與情境口說任務。錄音資料由兩位評分者依據流利度、正確度、複雜度與任務完成度進行評分，最後以共變數多變量分析 (MANCOVA) 排除前測差異，比較不同 AI 支援模式之成效。

3. 結果

經共變數多變量分析 (MANCOVA) 結果顯示，在控制前測分數後，AI 支援模式在整體口說表現上達到顯著多變量效果 ($\text{Wilks' } \Lambda = .87, F(4, 53) = 2.03, p < .05, \text{partial } \eta^2 = .13$)，證實兩組學生在口說指標上存在顯著差異。進一步的單變量分析顯示，「口說複雜度」與「任務完成度」為主要的差異來源：教育型 AI 代理人組在口說複雜度上顯著優於 GenAI 組 ($F(1, 56) = 5.62, p < .05$)，顯示動態評量 (DA) 的遞進式引導能有效促使低成就學習者進行語意延展與語言組織，使其產出更具細節的句構；同時，教育型 AI 組在任務完成度亦顯著高於對照組 ($F(1, 56) = 7.41, p < .01$)，反映結構化支架有助於學生維持任務焦點並完整達成要求。相較之下，兩組在「流利度」與「正確度」上雖均較前測進步，但組間並無顯著差異 ($p > .05$)，推測此兩項指標可能需要更長期的自然互動與更明確的形式焦點 (focus on form) 回饋才能展現明顯區別。

4. 討論與建議

本研究結果支持了輸出假說 (Output Hypothesis) 與動態評量的中介效果。教育型 AI 代理人透過結構化引導，促使學習者在產出過程中進行深度語言加工，而非僅是消極模仿 AI 輸出。反觀一般 GenAI 組雖能獲得流暢示例，但因缺乏教學中介，學習者易產生「認知卸載」 (cognitive offloading)，傾向以表層方式快速完成任務，導致語言產出缺乏深度與完整度。基於上述發現，本研究提出以下教學建議：首先，在課室導入 AI 工具時應採策略性引導，結合 DA 的遞進式原則，要求學生先進行自主輸出再參考 AI 建議，以避免過度依賴 AI 產出的完整答案。其次，針對低成就學習者，AI 系統的設計應扮演「中介者」而非「代寫者」，透過精心設計的中介支架引導其覺察語意缺口並延伸內容，從而維持學習者的主體性與認知參與程度。

參考文獻

- Lan, Y. J., & Chen, N. S. (2024). Teachers' agency in the era of LLM and generative AI. *Educational Technology & Society*, 27(1), I-XVIII.
- Lantolf, J. P., & Poehner, M. E. (2014). *Sociocultural theory and the pedagogical imperative in L2 education: Vygotskian praxis and the research/practice divide*. Routledge.
- Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in cognitive sciences*, 20(9), 676-688.

