

人工智慧浪潮下的語言研究與語言教學

陳浩然*

一、前言

打開手機、輸入問題、等待回答——這個動作，如今對許多人來說已和查字典、問朋友一樣自然。AI 不是未來式，它早就在我們的日常裡安靜落地：幫我們回信、修改文法、翻譯語言，甚至在我們還沒想清楚問題之前，就先給了一個答案。

人工智慧 (Artificial Intelligence, AI) 與生成式人工智慧 (Generative AI) 的快速發展，近年來已對人文社會科學的研究方法與高等教育實踐產生深刻影響。語言作為人類認知、社會互動與文化傳承的核心媒介，不僅是人工智慧技術最早介入的研究對象之一，也成為生成式模型能力與限制最集中顯現的場域。隨著自然語言處理 (NLP)、大型語言模型 (LLM) 與語言科技應用逐步進入研究與教學現場，語言學門亟需一個能同時回應技術發展、教育實務與理論反思的對話平臺。

基於此一學術與實務需求，語言學門於 2025 年規劃辦理「AI 在語言學習及語言學之應用」八場工作坊，¹ 邀集來自資訊工程、語言學、學習科學與語言教學等不同領域之學者，依其研究專長從技術、教學、研究與理論層面，系統性探討人工智慧對語言研究與語言教育所帶來的衝擊與可能性。

二、主題探討與成果

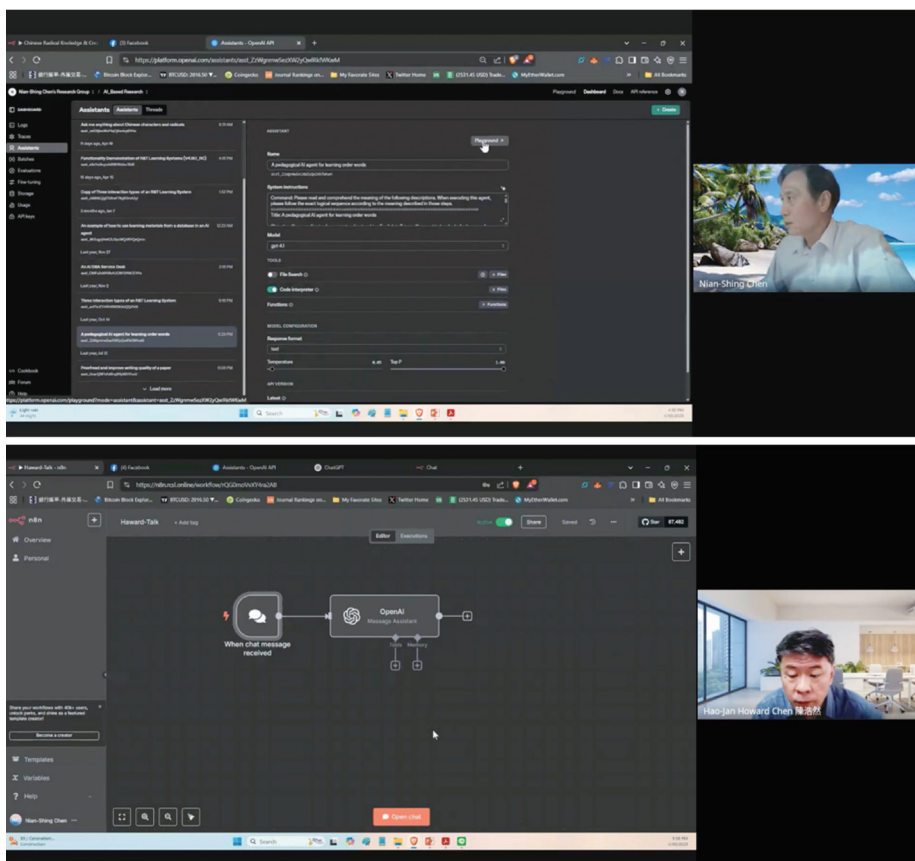
系列規劃中，首先從語言科技的基礎與發展脈絡談起，為後續教學與理論層面的討論建立共同語言。此一技術取向的起點，由國立清華大學資訊工程學系高宏宇教授主講。高教授長期投入自然語言處理、資料探勘與機器學習相關

* 國立臺灣師範大學英語學系特聘教授、人文處語言學門召集人

¹ 本系列並非單點式講座，而是以研究專長為軸線，逐步形成人機協作、教學應用與理論制衡並進的完整學門對話結構。

研究，對語言模型架構與語言理解能力的演進具有深厚研究基礎。他回顧自然語言處理自規則式與統計模型，逐步發展至深度學習與大型語言模型的歷程，指出生成式人工智慧並非突如其來的技術斷裂，而是長期研究累積的成果。他也梳理了自然語言處理的技術演進脈絡，從早期的詞袋模型、TF-IDF 與詞向量表徵 (Word Embedding)，一路推演至 Transformer 架構及 BERT、GPT 等大型語言模型的爆發性發展，深入剖析生成式 AI 的核心機制，包括「上下文學習」(In-Context Learning) 與多模態理解能力。最後以 TAIDE 引擎為例，展望在地化、可信任且特定任務導向 (Task-oriented) 的 AI 模型發展趨勢。

建立語言科技的基本理解之後，系列轉向人工智慧在教學設計中的角色與可能性，由國立臺灣師範大學學習科學學士學位學程陳年興教授從學習科學的角度，探討人工智慧如何在教學情境中被有意識的設計與使用。陳教授長期研究數位學習、同步教與學以及科技輔助學習歷程，特別關注科技如何支持而非取代教師的專業判斷。

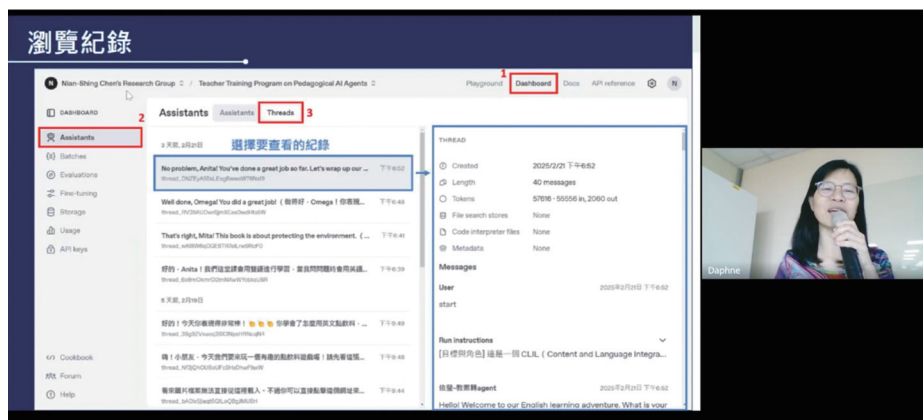


圖一：陳年興教授示範教學型 AI 代理人平臺操作介面

講座中，他提出「教學型 AI 代理人」(Pedagogical AI Agent) 之核心概念，主張教育應從依賴學生提問技巧的被動式聊天機器人，轉向主動式代理人的典範轉移，強調 AI 代理人應作為教師的「數位分身」(Digital Twin)，其目的在於賦能教師 (Enable Teachers) 以實現全天候的一對一輔導，並致力於培養具備自主產出能力的「生成式學生」(Generative Student)。在實踐技術上，他展示了「No-Code」的開發路徑，教師僅需掌握清晰的自然語言「指令」(Instruction) 撰寫技巧——涵蓋「目標與角色」、「環境與內容」及「學習流程與任務」三大維度——即可構建出遵循教師專業設計、且能避免模型幻覺的教學代理人，真正落實「工作流程自動化」的精準教學願景。

在具體實作層面，陳教授提出了一套指令撰寫的「黃金準則」，即包含「目標與角色」、「環境與內容」及「學習流程與任務」的三段式模板，以確保數位分身能精準執行教學意圖。在角色定義上，必須明確區分「真人教師」(內容提供者)、「AI 教師」(執行者) 與「學生」三者，並嚴格遵守「內容由教師提供」的原則，以根絕 AI 產生幻覺的風險。透過此一結構化流程，教師能將原本耗時的重複性教學工作，轉化為可規模化、全天候運作的自動化流程，有效擴增自身的教學影響力。

隨著人工智慧逐步被納入教學設計，系列進一步關注教師如何在實務層面參與人工智慧系統的建構與調整，國立中央大學資訊工程學系張嘉惠教授以其在資訊工程、自然語言處理與對話系統方面的研究經驗，介紹教育型對話代理人設計工具 EduACT。張教授長期投入教育情境中的對話分析與學習歷程研究，其跨域背景使該工具能兼顧技術可行性與教學實用性。



圖二：張嘉惠教授示範 EduACT 平臺對話紀錄與學習歷程儀表板介面

她介紹了專為教育工作者打造的「EduACT」平臺，其核心優勢在於提供「No-Code」(無程式碼)的開發環境，旨在降低技術門檻，賦能教師透過自然語言對話或模組化設定，即可輕鬆建構專屬的課程 AI 助教。該平臺不僅支援多樣化的教學任務設計，更具備強大的後臺分析功能，能將學生與 AI 的互動歷程轉化為視覺化儀表板 (Visual Dashboard)。透過分析對話數據，教師能精準追蹤並掌握每位學生的學習狀態與難點，具體實踐「Human-in-the-loop」的教育理念。

EduACT 平臺也引入了高度直覺性的「對話式代理人建構」(Creating agents via chatting with teachers) 功能。教師無需具備程式背景，僅需透過與系統進行自然語言對話，即可定義出代理人的行為模式與教學目標，這極大化了「賦能教師」的設計初衷，透過平臺機制促進了教師間的經驗分享與協作，為教育現場建立起一個可持續發展的 AI 教學資源生態系。本場在系列中扮演關鍵橋接角色，將「人機協作」的理念落實到教師可實際操作的教學設計層面。



圖三：EduACT 工作坊結束後與會者合影留念

當人工智慧逐漸進入教學與研究實務，相關的學術責任與倫理議題亦隨之浮現。國立陽明交通大學英語教學研究所孫于智教授從研究方法與學術寫作，探討生成式人工智慧在研究歷程中的應用界線。孫教授長期研究電腦輔助語言教學、學術寫作與研究方法，特別關注研究者在實際研究過程中所面臨的困境與需求。她說明人工智慧在研究主題發想、文獻整理與初步分析上的輔助潛力，同時強調原創性、引用規範與責任歸屬的重要性，提醒研究者在運用新工具的同時，仍須維持學術誠信與方法自覺。

在評量與學習回饋層面，系列聚焦語音科技於語言教學中的實際應用。國立臺灣師範大學資訊工程學系陳柏琳教授以其在語音辨識、自然語言處理與資

訊檢索方面的研究專長，首先回顧人工智慧與自然語言處理的發展脈絡，說明語言能力在人工智慧中的核心地位。LLM 可視為高度擴展的語言模型，透過向量語意與上下文嵌入，展現出強大的生成、推理與指令理解能力，並已深刻影響機器翻譯、對話系統、資訊檢索與語言學習等應用場域。

在此基礎上，陳教授說明其在電腦輔助發音訓練 (CAPT) 與口說能力評量 (SPA) 中的實際應用。內容涵蓋錯誤發音偵測與診斷 (MDD)、自動發音評量 (APA)，以及多面向口說能力評估模型，強調系統如何對學習者的發音、流暢度、語言使用與內容表現進行多層次分析。透過實際系統與研究案例，展示當前語言科技如何支援口說練習、回饋與評量，並為未來更智慧化、即時化的語言學習與評估模式提供方向。

在技術與應用逐步成形之際，系列回到教師專業與教育價值的核心問題。國立臺中教育大學英語學系廖美玲教授從師資培育與跨文化學習的研究視角，探討 Authentic Intelligence 與 Artificial Intelligence 在語言教學中的交會。廖教授長期關注教師專業發展與電腦輔助語言學習，強調理解技術、批判使用與倫理判斷並重，並指出制度支持與教師培力在 AI 時代的重要性。

她從全球、美國與臺灣高等教育機構的 AI 政策談起，指出多數政策提供的是方向性指引，而非現成答案，真正的關鍵在於教師如何於課堂中落實 AI 素養 (AI literacy)。透過整合多個 AI 素養框架，以大學英語閱讀與寫作課程為例，展示如何引導學生理解 AI、有效使用與提問、批判檢核生成內容，並在倫理前提下將 AI 作為協作夥伴而非替代者。

系列最後由中央研究院語言學研究所王聖富助研究員主講兩場，從語言學理論角度回顧類神經網路與預訓練語言模型在語言研究中的角色與限制。王助研究員專長於語音學、音韻學、計算語言學與類神經網路模型，長期關注語言學理論與計算模型之間的互動。

在批判性反思與方法論重構方面，他主張學界應揚棄「人工智慧」此一模糊且掩蓋「AI 帝國」剝削本質的流行術語，回歸科學嚴謹性，改以「類神經網路」或「表徵學習」稱之。他進一步提出語言學家介入技術的兩條具體路徑：一是「解析模型表徵」，即透過分析向量空間中的語意分布（如複數類別、同音詞發音關聯）來揭示潛藏的語言結構；二是「以模型驗證理論」，將模型視為測試語法（如刺激貧乏說）或音韻理論（如可學習性）的實驗場域，並透過設計「挑戰集」來揭露模型依賴捷思而非真正理解的缺陷。這兩場在系列中扮演關鍵的理論制衡角色，使整體討論在技術熱潮中保持學術自覺。

55

Form-meaning relationship:
(Gahl and Baayen, 2024)

	#ta	ta	aim	im#	#la	la
time	1	1	1	1	0	0
(2) C = time	0	0	1	1	1	1
thyme	1	1	1	0	0	

- They start with (vector) representations for form and meaning
 - Form: segment trigrams
 - Meaning: word embeddings (fasttext)
- From these, they get two new predictors
 - Semantic similarity: correlations between the embeddings of 'time' and 'thyme'
 - Semantic support: How strongly is a word (with this embedding) likely to be pronounced this way (with this segment trigram vector)
 - Learned through an algorithm that is kind of like neural nets, but only linear

10
3

Thank you!
I hope you have
enjoyed the ...

AI
talk

Computational
linguistics talk

Sheng-Fu Wang

Shu-Chuan Tseng

圖四：王聖富助研究員講解形式與意義關係的向量表徵分析

三、結語

AI 技術仍在高速演變，今天學會的工具，也許明天就已更新換代。但這恰說明，持續的跨領域對話與反思，遠比追趕任何一個版本還更重要。本系列工作坊透過多元場次與跨領域視角，協助與會者從「使用者」角色逐步建立對 AI 運作邏輯與限制的基本認識。多場工作坊不約而同指出，無論是教學設計或研究詮釋，核心判斷仍需仰賴人類的專業知識與價值思考——AI 扮演的是「延伸」與「輔助」的角色，而非取代者。語言教學與語言學研究雖面對各自的差異化需求，卻也共同面臨評量規範、倫理指引與方法論重整等課題。

在此背景下，單次技術導入遠不如持續的跨領域對話更具意義。2025 年語言學門 AI 系列工作坊所體現的，正是一種以「理解、試探、反思」為核心的學術態度，為語言學與語言教學在科技變遷中重新確認自身的學科價值，奠定了對話的起點。