

大语言模型在外语教师教育与发展研究中的应用探索^{*}

张 莲

提要: 随着生成式人工智能技术的快速迭代,大语言模型凭借强大的语言理解、生成和处理能力,在包括外语教师教育与发展研究在内的人文社科研究中的应用潜力和价值日益显现。本研究简要介绍大语言模型应用于自然语言处理任务的特点、优势和外语教师教育与发展研究的主要论题及特点,然后探讨大语言模型在外语教师教育与发展研究中的课堂话语分析、教学支架功能分析等方面的具体应用,最后讨论了大语言模型实际应用可能面临的挑战和问题。

关键词: 大语言模型; 外语教师教育与发展; 探索案例

Abstract: With the rapid iteration of generative artificial intelligence technology, large language models (LLMs) exhibit their application potentials and values in humanities and social science research, including foreign language teacher education and development (FLTED) research, for their powerful language comprehension, generation and processing capabilities. This study briefly introduces the characteristics and advantages of LLMs in NLP tasks, and the main topics and features of FLTED research. It then explores specific applications of LLMs in FLTED research, such as applying LLMs to the analysis of classroom discourse and instructional scaffolding functions in particular. The study also discusses challenges and issues possibly encountered in practical LLMs application.

Key words: large language model; foreign language teacher education and development; exploratory case

中图分类号: H319 文献标识码: A 文章编号: 1004-5112(2024)05-0025-08

1. 引言

随着生成式人工智能(artificial intelligence, AI)技术的快速迭代,大语言模型(large language models, LLMs)以其强大的语言理解、生成和处理能力及智能的交互方式,正在深度塑造人类的生活、工作和理解世界的方式(如 Grossman *et al.* 2023; 肖峰 2023; 程鹏等 2024),逐渐显现出在科学研究和知识生产中的潜力与应用价值。我国外语界部分学者已从不同视角就生成式 AI 应用于知识生产展开了探讨(如秦洪武,周霞 2024)。外语教师教育与发展研究是人文社科领域重要的次领域,相关知识生产事关外语教育理论和实践创新发展,直接影响外语教师队伍建设的这一重大基础性工作。如何将 LLMs 应用于外语教师教育与发展研究具有理论价值和现实意义,值得深入探讨。

2. 关键概念及相关研究综述

2.1 大语言模型和知识生产

从各种信息源可以发现,目前 LLMs 的定义主要包括三方面信息: LLMs 的本质(一种 AI

^{*} 本研究系北京外国语大学双一流建设重大项目“中国特色应用语言学史和应用语言学知识体系创新研究”(编号 2022SYLZD009)的阶段性成果。诚挚感谢徐月梅博士在技术相关表达方面的指导、《外语界》编辑部的宝贵意见和建议。

系统深度学习算法或技术)、构建方式(受训于海量文本数据/参数)、功用或价值(具备理解和生成自然语言的能力)(如 Rytting *et al.* 2023; Van Noorden & Perkel 2023; 肖峰 2023)。2023 年,全国科学技术名词审定委员会联合多家机构将 LLMs 定义为:基于海量文本数据训练的深度学习模型;它不仅能够生成自然语言文本,还能够深入理解文本含义,处理各种自然语言任务,如摘要写作、问答、翻译等。

LLMs 应用于学术研究和知识生产的相关研究目前主要涉及两方面内容。其一,探讨 LLMs 自身技术发展和优化,比如算法优化、数据和算力提升及性能评估(Bengio *et al.* 2021; Chang *et al.* 2024)。其二,探索如何将 LLMs 应用于学术研究,具体围绕三个问题展开:(1)论证 AI 工具驱动科学研究的基本原理及其与人类知识生产模式之间“内在联系的机理”(张越等 2024);(2)报告 LLMs 应用于各类学术研究和知识生产场景的案例及其效果,比如在特定学科专业领域的文献整理、数据标注、编码和分析等(Rytting *et al.* 2023; Rasheed *et al.* 2024; 苏祺 2024);(3) LLMs 应用可能存在的问题、挑战和风险,涉及学术研究和知识生产主体性、隐私保护、数据安全等法律伦理问题(Farquhar *et al.* 2024; 程鹏等 2024)。三个问题内里彼此关联,但第一个问题是核心,即以 LLMs 为代表的生成式 AI 和人类智能(human intelligence, HI)有什么异或同,以至于它能够或很大程度上参与学术研究和知识生产。

AI 是机器(计算机)模仿或模拟 HI 产生的智能。那么,HI 的基础是什么?蔡曙山(2020, 2024)认为,HI 是人类在神经、心理、语言、思维、文化五个层级的认知能力,其中语言既是人类认知和非人类动物认知的分野,又是人类认知的基础。虽然 Chomsky (1968)、Vygotsky (1978)、Pinker(1994)等对人类语言的本质及其与 HI 关系的看法多有不同,但相通且相同的是,他们都认为二者之间的关系天然、深刻且密切。LLMs 之所以被视作 AI 的革命,正因其在于人类自然语言处理(NLP)方面的突破(蔡曙山 2024)。再者,语言在学术研究和知识生产过程中扮演主要角色,文献归纳或数据标记、编码和分析都是对人类自然语言文本的处理(Rytting *et al.* 2023)。正是这种能力使得 LLMs 在知识生产领域具有巨大的应用潜力。

2.2 人工智能视域下外语教师教育与发展研究

外语教师教育与发展研究旨在探究教师个体或群体学习与发展的本质和规律,以及如何应用研究成果指导和规划教师学习与发展实践(王蔷等 2022)。这一领域的研究虽起步较晚,但已历经多次范式转向(Zhang 2022)。学者们从不同理论视角聚焦各种具体问题,从教师行为(包括言语和非言语)到教师认知(如知识、信念、决策、学习),从教师学习的社会文化取向到教师的身份、角色、能动性、情感体验(Borg 2009; Johnson 2009),以及复杂动态系统视角的教师认知(Cochran-Smith *et al.* 2014)等。

从研究方式上看,无论聚焦部分的还原式研究还是涵盖全貌的系统性研究(Davis & Sumara 1997),数据和资料的主体及原型都是语言文本,比如课堂言语或非言语行为的实录及转写、访谈文本、反思日志等。事实上,这也是多数人文社科领域研究基本的数据和资料原型。数据和资料的分析过程之本质亦是 NLP,包括语言文本含义的理解、归纳、分类、生成等。由此可见,外语教师教育与发展研究领域的学术研究和知识生产的基本原理与 LLMs 的机理具有“适配性”(蔡曙山 2020; 王涛 2024)。

笔者在中国知网和 Web of Science 数据库中检索了 2014 年至今人工智能背景下的外语教师教育与发展研究文献。文献计量分析和高被引重点文献细读均显示,该领域目前仍以形而上的研究为主,多围绕 AI 赋能外语教育和教师发展的价值与意义展开思辨式论证探讨,少见

具体展示如何开展研究的实证案例。

3. 外语教师教育与发展研究的 LLMs 应用案例

LLMs 应用于外语教师教育与发展研究案例的选取基于两个原则。(1) 选用笔者相对熟悉的研究场景和数据类型,便于评估 LLMs 生成内容的准确性(Abdurahman *et al.* 2024),或者辨识 LLMs 固有的“幻觉”(hallucination) 现象,即生成“无意义或不忠实于源数据的内容”(Farquhar *et al.* 2024: 625)。可能的情况下,尽量选取经人工标注的数据类型和片段,并将其作为“基准真值”(ground truth) 和 LLMs 生成的内容进行对比,以检验人—机协同(human-LLMs alignment) 效果(Abdurahman *et al.* 2024)。(2) 基于 LLMs 构建方式和特点,选用经典理论和概念的探索应用为主,利于确认 LLMs 应用的適切性和有效性。当前 LLMs 的使用中,特别是处理垂直领域任务时,会用检索增强生成技术(retrieval augmented generation, RAG),以避免或缓解“幻觉”、信息更新不及时与黑匣子现象(Gao *et al.* 2023; Zhao *et al.* 2024)。RAG 技术在 LLMs 的基础上引入并检索外部知识库的信息来增强 LLMs 的生成能力,从而提高模型的准确性和相关性。经典理论和概念相关文献历时积累量大,不仅能为 LLMs 提供丰富的专业词汇和语境,也能借助更多的先验知识提升 LLMs 的应用性能。基于上述原则,笔者选取外语教师教育与发展研究领域两个常见数据场景,分别为课堂话语(classroom discourse) 分析和教学支架(instructional scaffolding) 功能分析。

3.1 LLMs 应用于课堂话语分析的基础标注和编码

课堂话语指教学过程中师生为特定教育教学目的而开展的言语和非言语交流与互动(Cazden 2001)。课堂话语研究不仅揭示学生学习、认知过程,而且帮助教师反观、检视自身知识与经验,成为认识、理解教师学习和本质与规律的重要方式(张莲 2016, 2022; 张莲,王艳 2014)。Hymes(1972: xviii) 指出,研究课堂中的语言是真正意义上的基础研究,理解课堂中的语言方面取得的进步就是语言学理论的进步。课堂话语的所指范围显然大于“课堂中的语言”,其研究意义自然厚重和深远。

基于文献分析和实证案例研究经验,笔者提出了课堂话语“多重标注与分析方案”的建议(张莲 2016, 2022)。“多重标注与分析方案”是指研究者综合相关概念框架和话语分析方法、技巧与手段,实现目标课堂话语单位多角度切分,多次定义、分析和解释的方法或系统,力求全面、准确描述课堂话语与学习的关系,实现教师有效反思和学习(Rymes 2009; 张莲 2016, 2022),其基础是经典的课堂话语标注和规模计量工具 IRF(Initiation-Response-Feedback) 序列模型(Sinclair & Coulthard 1975)。研究者在量化标注基础上对序列中每一话步的教学功能进行质性标注,建立话步与认知、学习的关系(张莲 2022),为教师理解和反思教学奠定扎实的量—质分析基础。当课堂话语数据量较大时,多重标注和分析费时费力,而 LLMs 的应用似乎带来了提质增效的机会。

课堂话语分析案例选用的 LLMs 是当前广泛使用的 LLMs-a 和 LLMs-b(均为假名),必要时采用 LLMs-c 以核验编码有效性(Farquhar *et al.* 2024)。按照案例选用原则,笔者选用两个英语专业精读课堂片段(片段 1、片段 2) 示例 LLMs 应用过程。LLMs 应用的第一步是设计和改进提示词(prompt)。笔者仔细比对人机对话结果,不断修改完善提示词,整个过程持续近两个月。表 1 是先后六个版本(V1—V6) 中改动较大的三个版本的提示词,其中 V6 是最终采纳生成结果所用的提示词。因篇幅有限,本文仅止步于案例数据的标注和编码,数据挖掘和分析将另文探讨。

表 1 三个版本的提示词示例

版本	提示词
V1	<i>Please use the IRF framework proposed by Sinclair & Coulthard (1975) and tag the given text.</i>
V4	<i>You are an experienced classroom discourse analyst. Here is the transcript of a classroom session from an English Intensive Reading course offered for a group of undergraduates in a program of English language and literature. Please use the IRF (Initiation–Response–Feedback) framework by Sinclair & Coulthard (1975) and tag the transcript. You are required to: 1) tag the transcript with IRF and tell how many there are for each altogether , and 2) summarize the interactive features of this classroom session.</i>
V6	<i>As an experienced classroom discourse analyst , please analyze the following transcript from an English Intensive Reading course for undergraduate English language and literature students. Use the IRF (Initiation–Response–Feedback) framework by Sinclair & Coulthard (1975) to: 1) tag the transcript with IRF labels and provide a count for each category; 2) summarize the key interactive features of the classroom session.</i>

应用过程显示 ,提示词确实会影响对话结果。改进提示词的过程是给 LLMs 提供更多“定点搜索或检索”(grounded search or retrieval) 的机会(Farquhar *et al.* 2024: 625) ,也为使用 RAG 技术时更精准定位并启动外部知识库提供优质线索(Gao *et al.* 2023; Zhao *et al.* 2024) 。表 2、表 3 分别呈现两个课堂话语片段的标注和编码结果及其隐含的互动结构特征 ,所有案例片段均为笔者手动标注、编码的课堂话语样本。

表 2 课堂话语片段的标注和编码结果

片段	IRF 计数(LLMs-a)	片段	IRF 计数(LLMs-a)
1	<i>** Summary of Tagged Exchanges **</i> <i>— ** Total Initiations (I) : ** 10</i> <i>— ** Total Responses (R) : ** 5</i> <i>— ** Total Feedback/Evaluations (F/E) : ** 9</i>	2	<i>** Summary of Tagged Exchanges **</i> <i>— ** Total Initiations (I) : ** 20</i> <i>— ** Total Responses (R) : ** 16</i> <i>— ** Total Feedback/Evaluations (F/E) : ** 18</i>

表 3 课堂话语片段中师生话语互动结构特征及描述

片段	师生话语互动结构特征(LLMs-a)
1	<i>** Summary of the Interactive Features of the Classroom Session **</i> <i>1. ** Teacher-Centered Instruction: **</i> <i>— The session is predominantly led by the teacher , who provides detailed explanations of vocabulary , cultural nuances , and context related to the text “A Drink in the Passage.”</i> <i>— The teacher often initiates topics and directs the flow of the lesson.</i> <i>2. ** Use of Questions: **</i> <i>3. ** Student Participation: **</i> <i>4. ** Feedback Provided: **</i> <i>5. ** Code-Switching: **</i> <i>6. ** Cultural References: **</i> <i>7. ** Linguistic Focus: **</i> <i>8. ** Classroom Dynamics: **</i> <i>9. ** Attention to Language Nuances: **</i> <i>10. ** Encouraging Critical Thinking: **</i> <i>---</i>

(续表)

2	** Summary of the Interactive Features of the Classroom Session **
	1. ** Teacher-Centered Dialogue: ** — The teacher guides the discussion , frequently posing questions to elicit student participation. — Maintains control over the lesson flow , ensuring that key topics are covered.
	2. ** Student Participation: **
	3. ** Use of the IRF Framework: **
	4. ** Focus on Language Development: **
	5. ** Bilingual Instruction: **
	6. ** Literary Analysis: **
	7. ** Cultural Sensitivity and Context: **
	8. ** Developing Inferential Skills: **
	9. ** Encouragement of Critical Thinking: **
	10. ** Positive Learning Environment: **

注： 两个片段的互动结构特征都有具体描述 限于篇幅仅保留各自第一条的具体描述作为示例。

从表 3 可以看出 ,LLMs 从不同方面概括了两个片段的师生话语互动特征 ,包括教师中心、学生参与、语码转换(或双语教学)、语言关注、文化敏感、课堂动态、学习氛围、思辨能力发展等 ,回应了课堂话语分析的各个层面(Cazden 2001; 张莲 2022) 。 LLMs 能就具体互动情形作出较为精细且有区别性的定性概括 ,认为两个课堂都是“Teacher-Centered Instruction” ,但对片段 1 的定性描述是“The session is predominantly led by the teacher” ,而对片段 2 的描述则是“The teacher guides the discussion , frequently posing questions to elicit student participation; maintains control over the lesson flow” 。 由此可见 ,LLMs 能够较好完成类似文本处理任务。 需要指出的是 ,上述过程的多个阶段也是人机协同对齐(alignment) 的过程 ,即研究者将 LLMs 生成的结果和基于人类手动标注分析的结果反复比对的过程(Abdurahman *et al.* 2024) 。

3.2 LLMs 应用于教学支架功能分析的基础标注和编码

支架概念用于描述学习或解决难题情境中 ,教师(或一个能力更强的伙伴) 给学生(或一个能力稍弱的伙伴) 提供的支持和帮助(Wood *et al.* 1976; Swain 2000; Li & Zhang 2022) 。 Wood *et al.*(1976) 提出的教学支架功能框架尤为常用 ,细节如表 4 所示。

表 4 教学支架及其功能

序号	名称	功能
1	兴趣唤起(Recruitment , R)	引发学习者的好奇心和探索欲 形成积极学习态度
2	降低自由度(Reduction of Degrees of Freedom , RDF)	降低学习难度 提升学习者的自我效能感、学习动力和成就感
3	兴趣维持(Direction Maintenance , DM)	帮助学习者保持学习动力 防止他们遇到困难时放弃
4	标记显著特征(Marking Critical Features , MCF)	帮助学生更快掌握学习内容要点 提高学习效率 解决学习问题
5	情绪控制(Frustration Control , FC)	帮助学习者建立积极学习心态 增强抗挫能力
6	示范(Demonstration , D)	帮助学习者掌握学习方法和策略 提高自主学习和问题解决能力

语言学习情境中,教师提供有效支架的关键在于准确判断。首先,教师需评估学生的真实水平,以提供恰当的支架。其次,教师需准确判断何时逐步“撤回”支架,使学生更好地开展自主学习。最后,教师需懂得如何组织小组活动,促成组内“社会互动”,便于学生相互提供支持 (Swain 2000)。如此,学生学习的问题转化为教师学习和发展的问題,研究的焦点随之转向对教学支架及其功能的分析,而分析的前提是对教学支架进行标注和编码。按照案例选用原则,笔者选取语言学专业“话语分析”和“认知语言学”课的两段课堂话语片段 3、片段 4 进行支架标注和编码,其中前者为主案例,后者为辅助案例,以形成对比较验。以片段 3 为例,终版提示词如下所示:

As an experienced researcher on instructional scaffolding, analyze the following excerpt of a classroom interaction from a discourse analysis session with MA students in linguistics and applied linguistics. Using the six scaffolding functions framework by Wood, Bruner, and Ross (1976), identify, tag, and label all instances of scaffolding provided by the teacher in the excerpt. Note that there may be multiple scaffoldings within a single line.

支架识别和标注牵涉复杂的语义判断和上下文逻辑推理,所以研究者一般会邀请 2—3 名编码员核验编码一致性,确保编码准确率。在 LLMs 应用语境中,核验方法是将同一数据放入不同 LLMs 中生成结果或同一 LLMs 中多次生成结果,既观察 LLMs 的有效性和适用性,也初步判断数据处理的准确性和稳定性。表 5 呈现了片段 3 的 LLMs 支架标注和编码结果。

表 5 片段 3 应用 LLMs-a 和 LLMs-b 重复标注和编码结果

标注和编码汇总	LLMs-a 支架数		LLMs-b 支架数		人工识别支架数
Recruitment (R)	第一次 4	第二次 4	第一次 5	第二次 5	4
Reduction of Degrees of Freedom (RDF)	第一次 7	第二次 7	第一次 12	第二次 12	6
Direction Maintenance (DM)	第一次 11	第二次 11	第一次 7	第二次 7	11
Marking Critical Features (MCF)	第一次 11	第二次 11	第一次 7	第二次 7	12
Frustration Control (FC)	第一次 5	第二次 5	第一次 3	第二次 3	5
Demonstration (D)	第一次 3	第二次 3	第一次 3	第二次 3	3

笔者从表 5 归纳出三点。第一,LLMs 处理案例数据的结果喜忧参半。在提示词清晰、恰当且一致的前提下,不同 LLMs 的表现存在差异。LLMs-a 似乎能更有效处理案例数据类型,其生成结果和人工结果具有明显一致性。课堂话语研究实践中,降低自由度(RDF)和标记显著特征(MCF)两个支架类别的确较难区别定性,人工识别时也常令研究者举棋不定。独立生成结果时,LLMs-a 和 LLMs-b 各自表现稳定,且重复生成结果一致性高。从 LLMs 评估视角来看,这一点对 LLMs 个案是有价值的 (Abdurahman *et al.* 2024; Chang *et al.* 2024)。第二,两个 LLMs 生成结果差异较大,可能与 LLMs 受训数据不同有关,这给数据分析和后续研究结果带来不确定性和风险 (如 Rytting *et al.* 2023)。根据相关文献建议,笔者采用 LLMs-c 进行核验,发现 LLMs-c 和 LLMs-a 的生成结果高度一致。当然,此种情况不排除 LLM-c 和 LLMs-a 属于同一 LLMs 系列产生的一致性。第三,整个 LLMs 应用探索过程中,人机协同或者研究主体多元协同始终在场 (Rasheed *et al.* 2024),在基于 LLMs 的学术研究和知识生产中可能都是必选项。笔者使用辅助案例片段 4 重复相同 LLMs 应用过程,结果也支持上述发现。

4. 结语

以 LLMs 为代表的 AI 应用于学术研究和知识生产渐成普遍现象。本研究基于相关概念和现有文献梳理,实证探索了 LLMs 在外语教师教育与发展研究中的应用,拟提出如下看法。首先,外语教师教育与发展研究和 LLMs 的方法应用,特别是数据标记、编码之间具有一定的逻辑合理性,问题焦点与数据类型及其处理之间具有较好的适切性,且大规模数据处理具备高效、经济的预期。其次,通用 LLMs 在外语教师教育与发展研究领域的应用表现存在差异。第三,LLMs 用于数据标记和编码仍有偏误和失误,会因研究领域和问题不同而出现不稳定的处理结果,人机协同或多元主体协同应是必选项。最后,研究者应用 LLMs 的技巧有待学习和优化,但错误和失误的风险并不会降低或消失(Van Noorden & Perkel 2023),不谨慎的使用和应用可能带来错误的知识生产和危害。

AI 是对 HI 的模拟,但作为 AI 基础的语言不只是一些字词和语句及其统计或概率计算,而是一种文化和传统、一个文化或文明共同体的历史。人类语言表达的是人与真实世界之间的关系,在应用模拟 HI 的语言模型认识、理解世界之时,所有谨慎都不为过。

参 考 文 献

- [1] Abdurahman S , Atari M , Karimi-Malekabadi F *et al.* Perils and opportunities in using large language models in psychological research [J]. *PNAS Nexus* , 2024 , 3(7) . <https://doi.org/10.1093/pnasnexus/pgae245>.
- [2] Bengio Y , Lecun Y & Hinton G. Deep learning for AI [J]. *Communications of the ACM* , 2021 , 64(7) : 58-65.
- [3] Borg S. Language teacher cognition [A]. In Burns A & Richards J C (eds.) . *The Cambridge Guide to Second Language Teacher Education* [C]. Cambridge: Cambridge University Press , 2009. 163-171.
- [4] Cazden C B. *Classroom Discourse: The Language of Teaching and Learning* [M]. Portsmouth , NH: Heinemann , 2001.
- [5] Chang Y P , Wang X , Wang J D *et al.* A survey on evaluation of large language models [J]. *ACM Transactions on Intelligent Systems and Technology* , 2024 , 15(3) : 1-45.
- [6] Chomsky N. *Language and Mind* [M]. New York: Harcourt Brace Jovanovich , 1968.
- [7] Cochran-Smith M , Ell F , Ludlow L , Grudnoff L & Aitken G. The challenge and promise of complexity theory for teacher education research [J]. *Teachers College Record* , 2014 , 116(5) : 365-378.
- [8] Davis B & Sumara D J. Cognition , complexity , and teacher education [J]. *Harvard Educational Review* , 1997 , 67(1) : 105-125.
- [9] Farquhar S F , Kossen J , Kuhn L & Gal Y. Detecting hallucinations in large language models using semantic entropy [J]. *Nature* , 2024 , 630(8017) : 625-636.
- [10] Gao Y F , Xiong Y , Gao X Y *et al.* Retrieval-augmented generation for large language models: A survey [OL]. *arXiv* , 2023. arXiv: 2312.10997.
- [11] Grossmann I , Feinberg M , Parker D C *et al.* AI and the transformation of social science research [J]. *Science* , 2023 , 380(6650) : 1108-1109.
- [12] Hymes D. Introduction [A]. In Cazden C B , John V & Hymes D (eds.) . *Functions of Language in the Classroom* [C]. New York: Teachers College Press , 1972. xi-lvii.
- [13] Johnson K E. *Second Language Teacher Education: A Sociocultural Perspective* [M]. New York: Routledge , 2009.
- [14] Li D & Zhang L. Exploring teacher scaffolding in a CLIL-framed EFL intensive reading class: A classroom discourse analysis approach [J]. *Language Teaching Research* , 2022 , 26(3) : 333-360.

- [15] Pinker S. *The Language Instinct: How the Mind Creates Language* [M]. New York: William Morrow, 1994.
- [16] Rasheed Z, Waseem M, Ahmad A *et al.* Can large language models serve as data analysts? A multi-agent assisted approach for qualitative data analysis [OL]. *arXiv*, 2024. arXiv: 2402.01386.
- [17] Rymes B. *Classroom Discourse Analysis: A Tool for Critical Reflection* [M]. New York: Hampton Press, 2009.
- [18] Rytting C M, Sorensen T, Argyle L *et al.* Towards coding social science datasets with language models [OL]. *arXiv*, 2023. arXiv: 2306.02177.
- [19] Sinclair J M & Coulthard R M. *Towards an Analysis of Discourse: The English Used by Teachers and Pupils* [M]. Oxford: Oxford University Press, 1975.
- [20] Swain M. The output hypothesis and beyond: Mediating acquisition through collaborative dialogue [A]. In Lantolf J P (ed.). *Sociocultural Theory and Second Language Learning* [C]. Oxford: Oxford University Press, 2000. 97-114.
- [21] Van Noorden R & Perkel J M. AI and science: What 1 600 researchers think [J]. *Nature*, 2023, 621(7980) : 672-675.
- [22] Vygotsky L S. *Mind in Society: The Development of Higher Psychological Processes* [M]. Cambridge, MA: Harvard University Press, 1978.
- [23] Wood D, Bruner J S & Ross G. The role of tutoring in problem solving [J]. *Journal of Child Psychology and Psychiatry*, 1976, 17(2) : 89-100.
- [24] Zhang L. Theoretical paradigm shifts in second language teacher education and development: A historical perspective [A]. In Kimura Y, Yang L, Kim T-Y & Nakata Y (eds.). *Language Teacher Motivation, Autonomy, and Development in East Asia* [C]. Cham: Springer, 2022. 55-74.
- [25] Zhao S Y, Yang Y Q, Wang Z L *et al.* Retrieval augmented generation (RAG) and beyond: A comprehensive survey on how to make your LLMs use external data more wisely [OL]. *arXiv*, 2024. arXiv: 2409.14924v1.
- [26] 蔡曙山. 论语言在人类认知中的地位和作用 [J]. 北京大学学报(哲学社会科学版), 2020, (1) : 138-149.
- [27] 蔡曙山. 我言 故我在: 兼论人类和计算机两种语言智能 [J]. 外国语文, 2024, (4) : 11-32.
- [28] 程鹏, 童松, 崔鹤, 孙泽昌. AI 时代下科学知识生产的伦理困境与策略 [J]. 科学学研究, 2024. <https://doi.org/10.16192/j.cnki.1003-2053.20240709.001>.
- [29] 秦洪武, 周霞. 大语言模型与语言对比研究 [J]. 外语教学与研究, 2024, (2) : 163-176.
- [30] 苏祺. 大语言模型在二语教学中的应用效能解析 [J]. 外语界, 2024, (3) : 35-42.
- [31] 王蔷, 张莲, 顾佩娅. 中国外语教师教育与发展研究 [C]. 北京: 清华大学出版社, 2022.
- [32] 王涛. 生成式人工智能之于历史研究的机遇与挑战 [J]. 史学理论研究, 2024, (3) : 15-23.
- [33] 肖峰. 生成式人工智能与知识生产新形态——关于三阶知识生产的探析 [J]. 学术研究, 2023, (10) : 50-57.
- [34] 张莲. 基于课堂话语研究的外语教师学习与发展: 理据、议题和方法 [J]. 山东外语教学, 2016, (3) : 47-54.
- [35] 张莲. 外/二语课堂话语研究: 理论、方法和问题 [A]. 王蔷, 张莲, 顾佩娅. 中国外语教师教育与发展研究 [C]. 北京: 清华大学出版社, 2022. 221-234.
- [36] 张莲, 王艳. 通过课堂话语分析促进外语教师学习: 一项实证案例研究 [J]. 外语与外语教学, 2014, (3) : 36-41.
- [37] 张越, 郭玥, 余江. 通用人工智能驱动的科研新范式: 理论与实践 [J]. 科学学研究, 2024. <https://doi.org/10.16192/j.cnki.1003-2053.20240829.001>.

作者单位: 北京外国语大学 北京 100089