

大语言模型在语言教学中的应用研究综述

蔡雨晨¹

北京外国语大学中国外语与教育研究中心 / 广州南方学院

【摘要】生成式人工智能的发展推动了大语言模型在教育领域的广泛应用，但其在语言教学中的研究尚缺乏系统总结。本研究采用系统性文献综述方法，从中国知网和 Web of Science 核心数据库中筛选出 57 篇相关文献。研究结果揭示大语言模型能够帮助教育者设计教学任务、生成教学资源、减轻教学负担。同时，它在辅助语言学习者口语与写作练习、个性化辅导、自主学习支持方面展现了显著优势。对于教学管理者，大语言模型通过优化教育决策、促进教育公平，帮助弱势群体获得更多学习机会。然而，其应用也带来了学术诚信、过度依赖等问题，教育者和管理者需具备足够的技术素养以应对这些挑战。

【关键词】大语言模型；语言教育；外语教学；系统性文献综述

1 蔡雨晨，北京外国语大学中国外语与教育研究中心 / 广州南方学院，研究方向：认知语言学。邮箱：1029943523@qq.com；本论文由中央高校基本科研业务费专项资金资助，项目名称：ChatGPT 理论语言学研究能力评估，项目批准号：2024JX062。



A Review of the Applications of Large Language Models in Language Teaching

Yuchen Cai¹

National Research Center for Foreign Language Education, Beijing Foreign Studies University/
Nanfang College · Guangzhou

【 Abstract 】 The development of generative artificial intelligence has driven the widespread application of large language models (LLMs) in the field of education. However, research on their use in language teaching lacks a systematic summary. This study employs a systematic literature review methodology, selecting 57 relevant articles from the CNKI and Web of Science core collections. The findings reveal that LLMs can assist teachers in designing instructional tasks, generating educational resources, and reducing their teaching workload. Additionally, LLMs demonstrate significant advantages in supporting language learners with speaking and writing practice, personalized tutoring, and autonomous learning. For educational administrators, LLMs contribute to optimizing educational decision-making and promoting educational equity, thus providing more learning opportunities for disadvantaged groups. Nevertheless, their application also raises concerns such as academic integrity and over-reliance, highlighting the need for educators and administrators to possess adequate technical literacy to address these challenges.

【 Keywords 】 Large Language Models; language education; foreign language teaching; systematic literature review

¹ Yuchen Cai, National Research Center for Foreign Language Education, Beijing Foreign Studies University / Nanfang College · Guangzhou. Research interest: Cognitive Linguistics; Email: 1029943523@qq.com. This paper is supported by the Fundamental Research Funds for the Central Universities under the project titled “Evaluation of ChatGPT’s Research Competence in Theoretical Linguistics” (Grant No. 2024JX062).

引言

随着生成式人工智能（Generative Artificial Intelligence, GenAI）的快速发展，大语言模型（Large Language Model, LLM）在教育领域的应用逐渐成为学术界与产业界关注的焦点。该技术为教学模式创新与学习体验优化提供了重要契机，同时也引发了一系列新的挑战，促使教育研究者持续探索其在教学中的有效整合路径。

已有研究表明，LLMs 在多学科标准化测试中已表现出接近甚至达到学生水平的能力，能够较好完成选择题与开放性问题（Achiam et al., 2023）。在具体教学场景中，LLMs 已被应用于写作与阅读辅助，并在部分课程中提升学习表现（Malinka et al., 2023）。此外，相关研究指出，ChatGPT 等模型在多学科任务中能够生成逻辑连贯、兼具深度与广度的回答（Susnjak, 2022），并逐渐拓展至师生协作、个性化学习与自动化评估等教学应用场景（Tan et al., 2023）。然而，其应用也伴随着潜在风险，包括学术不端、生成内容偏见、技术依赖以及对非英语学习者的公平性问题（Kasneci et al., 2023）。

尽管现有研究已从多角度探讨 LLMs 在教育中的应用与前景，但针对其在语言教学领域的系统性综述仍相对缺乏。鉴于语言教学在培养批判性思维、表达能力与跨文化交际能力中的核心作用，有必要对相关研究进行系统梳理。因此，本文通过整合现有文献，旨在总结 LLMs 在语言教学中的应用现状，分析其面临的主要挑战，并展望未来研究方向，以推动该领域的深入发展。

一、文献回顾

大语言模型（LLMs）的发展可追溯至 20 世纪 50 年代的符号主义人工智能研究，但其现代形态形成于 21 世纪初，得益于统计自然语言处理与计算能力的提升，尤其是深度学习的引入。2017 年，Transformer 架构的提出显著提升了模型处理长距离依赖的能力，标志着 LLM 发展的关键转折（Vaswani et al., 2017）。随后，BERT 与 GPT 系列模型推动了大规模预训练范式的兴起，使模型能够通过无监督学习获取语言知识并经微调实现高性能（Douglas, 2023）。

此后，LLMs 进入快速发展阶段，并在文本生成、翻译与对话系统等任务中得到广泛应用。代表性模型包括 OpenAI 的 GPT 系列、Google 的 Gemini 系列、以及 Anthropic 的 Claude。中国亦推出了悟道、文心一言、豆包、Deepseek 等模型。这些模型共同推动了自然语言处理技术的持续演进。



近年来, LLMs 在教学中的应用日益多样, 尤其在二语教学中表现出显著潜力。相关研究表明, 其在语法纠错、翻译、智能写作与口语评估等方面能够提供实时反馈与个性化支持(苏祺, 2024)。在写作教学领域, Fathi 和 Rahimi (2024) 基于社会建构主义视角, 通过跟踪 EFL 学习者与 ChatGPT 的互动写作过程, 发现 AI 辅助对学术写作能力具有显著促进作用。

除语言教学外, LLMs 亦推动了教育模式创新。在翻译教育中, 其被用于优化教学设计与评估体系(王华树、谢斐, 2024)。在高等教育中, LLMs 虽可降低信息检索的认知负荷, 但可能削弱学习者对内容的深度参与(Stadler et al., 2024)。这些研究表明, LLMs 在提升效率的同时, 也对深度学习提出新的挑战。

综上, 尽管现有研究已初步揭示 LLMs 在语言教学中的应用潜力, 但针对其效果、场景与挑战的系统性整合仍较不足。基于此, 本文通过文献综述系统梳理相关研究, 以期在教育实践与后续研究提供更具理论依据的参考。

二、研究方法

本研究旨在通过系统综述并分析现有文献中关于 LLMs 在语言教学中的应用, 评估其教学效果, 并识别研究不足和未来研究方向。具体而言, 本文拟回答如下研究问题: LLMs 在语言教学中的应用有哪些? 这些应用的有效性如何? 未来研究的可能方向有哪些?

本研究借鉴系统性文献综述分析框架(Preferred Reporting Items for Systematic Reviews and Meta-Analyses, PRISMA)(Page et al., 2021)。在文献检索方面, 为保证文献权威性, 选取数据库为中国知网 CSSCI 检索源期刊论文和 Web of Science Core Collection SSCI 期刊论文。由于在以上数据库中直接搜索“大语言模型”和“语言教学”得到的文献数量较为有限, 作者适当扩大了检索关键词的范围。在知网数据库中检索关键词分别尝试了“大语言模型”和“教育”、以及“教育”相关关键词如“教学”或“教师”或“学生”, 在 Web of Science Core Collection 中检索关键词和布尔运算符为 (“large language model*” OR ChatGPT OR GPT-*) AND (“language education” OR teaching OR learning OR teacher OR student)。接着, 确定文献发表时间为 2018 年至 2024 年 8 月, 起始时间设定原因为 2018 年 GPT 系列预训练大语言模型开始兴起并逐步引入教育领域。最终在中国知网检索到 46 篇文献, 在 WoS 中检索到 435 篇文献。由于检索词范围较大, 作者首先根据标题和摘要初步筛选文献, 排除与研究问题无关的研究。随后对初步筛选后保留的文献进行全文本

审阅，确定其是否符合纳入标准。纳入标准包括：研究的焦点是大语言模型在语言教学中的应用；研究中包含实证数据或深入的案例分析，最终保留有效实证研究 57 项（中文 13 项，英文 44 项）。

为了回答上述研究问题，作者提取了当前使用 LLMs 语言教学任务的研究现状的信息。每篇文献中提取了三类核心背景信息：应用场景、面向主体以及大语言模型的应用成效。通过这些背景信息能够全面了解现有研究，并为研究人员和实践者提供指导，以探索 LLMs 的潜在应用方向。

三、大语言模型在语言教学中的应用场景及成效

有关具体应用方向，许家金和赵冲（2024）将 LLMs 在英语教学中的角色总结概括为语言顾问、语伴和语言测评专家三种角色，并展示了在听、说、读、写、译教学案例中 LLMs 发挥的作用。郑春燕等（2024）从教学、学习、管理和评价四个维度，提炼了人工智能支持语言教学的典型应用，具体包括自动问答与聊天机器人、机器翻译、信息检索与语料库等。本研究借鉴了以上分类方法，从面向语言教育者、语言学习者和语言教学管理者三个视角，总结梳理了 LLMs 在语言教学中的具体应用场景。下表总结概括了这些应用场景的现状。

表 1 大语言模型在语言教学中的应用场景及成效

面向主体	应用场景及成效	部分文献举例
语言教育者	教学设计与实施（设计教学任务、生成教学资源、改变教学策略与方法）	冯庆华和张开翼（2024）， Yeh（2024）
	教学测评与反馈	Ghafouri et al.（2024）， Godwin-Jones（2024）
语言学习者	辅助语言练习与反馈（口语、写作等）	Boudouaia et al.（2024）， 杨连瑞（2024）
	提供个性化辅导	Guo & Li（2024）， Al-Emran（2024）
	促进自主学习与自我调节	Karatas et al.（2024）， Liu, et al.（2024）
教学管理者	支持教育决策（帮助弱势群体、分配教学资源、识别并减少偏见）	Prinsloo et al.（2024）， Kwak & Pardos（2024）



3.1 大语言模型服务语言教育者的应用场景及成效

在教学设计与实施方面，LLMs 在任务设计、资源生成及教学策略优化中展现出优势。冯庆华和张开翼（2024）指出，教师可利用模型的推理与生成能力设计复杂文本分析和高级写作任务，从而提升学生语言能力并激发创造性思维。Yeh（2024）通过一学期对 13 名在职 EFL 教师的跟踪研究发现，ChatGPT 有助于生成互动性与适应性学习材料（如绘本与卡拉 OK 练习），从而促进以学生为中心的课堂环境。Kartal（2024）基于叙事探究与访谈分析 12 名英语师范生的实践经验，指出人机共创、批判性分析与协作在教学中的重要性。

在教学测评与反馈方面，LLMs 同样具有应用潜力。Ghafouri et al.（2024）通过三阶段教学方案验证，ChatGPT 可在头脑风暴、修改、反馈与测试等环节提升教师效能并促进学生写作能力发展。Godwin-Jones（2024）也指出，GenAI 有助于考试设计与反馈生成。

然而，相关研究亦指出潜在风险。Liu（2024）基于对中国十所大学 17 名教师的访谈发现，GenAI 对学术诚信与评估有效性构成挑战，且教师普遍缺乏系统指导，多依赖自发策略应对。Derakhshan 和 Ghiasvand（2024）对 30 名伊朗 EFL 教师的访谈表明，ChatGPT 具有双重效应：一方面可提升学习自主性、减轻教学负担并支持评估设计，另一方面可能引发作弊、虚假信息传播及抄袭等问题。

总体而言，LLMs 在支持高阶任务设计、优化教学资源及减轻教师负担方面具有显著潜力，但其在学术诚信与信息可靠性方面的风险，以及教师支持不足的问题，仍需进一步关注。

3.2 大语言模型服务语言学习者的应用场景及成效

李康康等（2024）讨论了 LLMs 赋能的智能学伴在促进学习互动、提供个性化辅导、辅助智能测评等方面的应用。本研究将智能学伴的角色细分为辅助语言练习与反馈（如口语、写作等）、提供个性化辅导、促进自主学习和自我调节。

LLMs 辅助学生口语、写作练习及反馈相关的研究较多（Wan & Moorhouse, 2024; Guo, 2024; Boudouaia et al., 2024; Moorhouse, B.L., 2024; 杨连瑞, 2024; Kim et al., 2024; Fathi & Rahimi, 2024），总体研究结果显示，人工智能赋能的语言练习对学生相关技能的提升有显著优势，但教师的参与在确保语言学习效果方面仍然不可或缺。在口语练习方面，Wan 和 Moorhouse（2024）探讨了使用 GenAI（如 Call Annie）作为语言学习者的口语伙伴的潜力。Call Annie 是一个视频聊天机器人，能够通过与用户的沉浸式视频通话支持他们的语言发展。在写作学习方面，Boudouaia et al.（2024）通过实验设计，探讨了 GPT-4 对阿



尔及利亚 EFL 写作学习的影响。研究发现使用 GPT-4 的实验组在写作技能、态度和行为意图等方面显著优于对照组。站在学生的角度上, Kim et al. (2024) 通过与 20 名中国高等教育学生的深入访谈, 探讨了学生对使用 GPT-4 嵌入写作系统进行学术写作的看法。研究发现, 学生期望 AI 在写作过程中担任多任务写作助手、虚拟导师和数字同伴的角色, 认为 AI 辅助写作能够改善写作过程以及写作表现。

除了辅助口语与写作练习, LLMs 还可以提供个性化辅导。Guo 和 Li (2024) 探讨了 EFL 学生使用自制的检索增强生成 (Retrieval Augmented Generation, 简称 RAG) 聊天机器人来辅助其写作的效果。结果显示, 学生通过使用这些聊天机器人来生成写作想法、制定写作大纲和识别语法拼写错误, 显著提高了写作动机和自信心, 并且更明确了写作目标。

促进自主学习主要体现在 LLMs 可以激发学生的学习动机和提高学习参与度 (Li et al., 2024; Karatas et al., 2024)。Karatas et al. (2024) 以语言教师的身份, 在四周内引导学生使用 ChatGPT 进行学习。通过访谈收集的定性数据经主题分析得出结论: ChatGPT 对学生们的学习体验具有积极影响, 并通过其多功能性和在不同学习活动中的便捷性提升了学生的学习动机和参与度。此外, LLMs 还可以通过多维提问策略帮助学生有效解决复杂的学习问题, 提升其自主学习能力 (翟雪松等, 2024)。自我调节主要涉及调节学习策略与方法。Liu et al. (2024) 提出了一种基于 LLMs 支持的认知学术语言学习模型 (Cognitive Academic Language Learning Model, 简称 CALLA-LLM), 通过对照试验验证了使用该模型的学生在小学 EFL 写作表现、自我调节学习 (Self-Regulated Learning, 简称 SRL) 策略和写作动机等方面得到了显著提升。

尽管 LLMs 在服务语言学习者上展现了诸多优势, 在具体使用过程中也存在不足。魏小东和习鹏飞 (2024) 指出 LLMs 和虚拟数字人也存在生成内容偏差、语用情感缺失、学习反馈不当等问题。为了解决上述问题, 他们通过实例验证了人机协同内容生成构建可信学习资源、实现多元反馈方式以及实时学习行为分析支持情感识别与表达等应对策略的有效性。此外, 尽管 LLMs 在降低认知负担方面表现出色, 但其可能在深入理解和批判性思维的培养上存在不足 (Stadler et al., 2024)。

综上所述, LLMs 在服务语言学习者方面展现了广泛的应用前景, 特别是在促进学习互动、个性化辅导、智能测评等方面。其在口语和写作练习中的应用显著提升了学生的语言技能和学习动机, 同时也支持了自主学习和自我调节。然而, 模型在生成内容偏差、无法培养批判性思考能力、反馈不当等方面仍存在挑战, 需通过人机协同等策略加以解决, 以确保学习效果 and 体验的优化。



3.3 大语言模型服务教学管理者的应用场景及成效

LLMs 对教学管理者的辅助主要体现在优化教育决策，促进教育公平上。具体而言，它可以帮助弱势群体、公平分配教育资源以及识别并减少偏见。Prinsloo et al. (2024) 探讨了弱势学生的数字福祉 (digital well-being) 问题及如何利用人工智能教育决策支持系统 (AI-based Educational Decision Support Systems, 简称 AI-EDSS) 更好地识别和理解他们。研究表明，设计和实施公正、道德、负责且透明的 AI-EDSS 时，应明确考虑弱势学生的需求。弱势学生的数字福祉应与其整体福祉一同被整体看待，教学设计者、教育者、政策制定者以及 AI 工具和应用的开发者应理解这些系统的潜在益处和相关风险，并强调需要让学生参与此类系统的设计和评估。

教育资源的公平分配在大语言模型时代也十分重要。Kwak 和 Pardos (2024) 的研究探讨了如何通过优化 LLMs 来缩小不同国家和语言在教育相关知识表示上的差异，特别是在 K-12 国家课程分类法 (K-12 state curricula taxonomies) 的应用中，这对于将开放教育资源和辅导内容与国家课程标准对齐至关重要。研究发现，当前的 LLMs 在非英语和非美国内容上表现不佳，但通过对模型进行微调，即使使用少量标记数据，也能显著提升其在这些语言和国家中的技能标记能力。这一研究表明，通过微调 (fine-tuning)，LLMs 有潜力在全球范围内更公平地支持教育内容的对齐和标记，促进教育资源的公平分配和使用。

LLMs 可能会在数据偏见的影响下产生不公平的结果，从而加剧教育中的不平等。然而，如果有效地检测和纠正这些偏见，LLMs 也有潜力通过个性化学习资源和教学辅助来促进教育公平。Lee et al. (2024) 设计了一个全面的框架，该框架识别了 LLMs 生命周期中每个步骤的潜在偏见，包括代表性偏见、测量偏见和学习偏见。他们指出从初始开发到在教育环境中定制预训练模型的各个阶段，理解并识别偏见来源的重要性，如训练、微调和部署 LLMs 时可能引入的各种偏见。为了确保更公平的教育结果，教育管理者需要与开发人员合作，确保数据集具有代表性，并为特定的教育环境定制 LLMs 以减轻偏见，从而使模型在教育应用中更为公平和有效。

总的来说，LLMs 在服务教学管理者方面表现出显著成效，主要应用于优化教育决策和促进教育公平。通过 AI-EDSS 系统，LLMs 能够识别和理解弱势学生的需求，提升他们的数字福祉。研究表明，设计这些系统时需关注公正性和透明度，确保包括弱势学生在内的所有学生得到公平对待。此外，LLMs 还能通过微调提升对全球教育内容的标记和对齐能力，尤其是在非英语和非美国内容的处理中，从而更公平地分配教育资源。然而，LLMs 的偏见问题仍需重视，教育管理者需与开发者合作，减少偏见对教育结果的负面影响。

3.4 小结

通过对近五年 LLMs 应用于语言教学的文献梳理，模型在服务语言教育者、语言学习者和教学管理者上展现了诸多优势，也带来了相应挑战。在支持语言教育者方面，LLMs 通过优化教学任务设计、丰富互动性学习材料，并改善教学评估与反馈机制，显著提升了教学质量。这些技术工具不仅提升了教师的教学效能，还在减轻其工作负担方面发挥了关键作用。然而，LLMs 的使用也带来了新的挑战，特别是在学术诚信和评估公正性方面，教师们亟需更完善的支持和培训。

对于语言学习者，LLMs 的应用极大地促进了个性化辅导、自主学习能力的提升，并在语言练习中展现了强大的支持作用。学生通过与模型的互动，不仅在口语和写作技能上取得了显著进步，还增强了学习的积极性和参与度。然而，LLMs 在内容生成的准确性、批判性思维的培养等方面仍有改进空间，需要结合教师的指导来实现更全面的学习效果。

在教学管理领域，LLMs 在优化教育决策、促进教育公平方面展现了独特的优势。通过 AI 教育决策支持系统，模型能够更好地理解和响应弱势学生的需求，推动教育资源的公平分配。然而，它在处理数据偏见时仍存在风险，需要教育管理者与开发者密切合作，以确保其在教育中的应用更加公正和有效。总体而言，尽管 LLMs 在语言教学中的应用展现了广阔前景，但其潜在问题也需要在进一步的研究和实践中得到解决。

四、研究局限与未来展望

尽管 LLMs 在语言教学中的应用研究取得了显著进展，但当前研究仍存在不足，这为未来研究的进一步拓展和深入提供了方向。首先，现有研究的焦点主要集中在写作和口语训练领域，对听力和阅读等关键语言技能的探索明显不足。听力和阅读在语言学习中同样至关重要，未来研究应致力于探讨 LLMs 如何支持个性化的听力练习、帮助学习者理解复杂文本，以及如何在阅读任务中通过提供实时释义、个性化阅读路径和互动性支持，实时反馈提升他们的听力和阅读能力。

其次，现有的大多数研究主要集中在高等教育环境中，对基础教育阶段的应用研究相对较少。基础教育是语言学习的关键期，对于学生未来的语言能力发展至关重要。如何利用 LLMs 激发学生的学习兴趣、提供适龄的语言支持，并在教学中发挥积极作用，亟需更多的实证研究。未来的研究应更多地关注 LLMs 在中小学教育中的应用，探索其在早期语言学习、词汇积累、语法理解以及跨学科语言支持中的作用。此外，还应考虑如何在基础



教育环境中保证 LLMs 的有效性和安全性，避免过度依赖技术带来的潜在风险。

此外，当前研究的地域和语言范围相对狭窄，主要集中在英语及其相关文化背景，对其他语言和多语言学习环境的关注较为有限，与中文 LLMs 相关的研究较少。语言学习本质上是多语言、多文化的过程，LLMs 的普及应用需要考虑不同语言和文化背景下的适用性与有效性。未来的研究应更多地关注 LLMs 在全球不同语言背景，特别是在非英语国家和地区下的表现，进而探讨其在促进多语言教育公平性和包容性方面的潜力。研究可以围绕如何利用 LLMs 支持非英语语言的学习、如何应对不同文化背景下的语言习得挑战、如何开发或将现有本国语言的 LLMs 融入教学，以及如何通过技术手段缩小全球教育资源的不平等等问题展开。

未来的研究和实践还应围绕如何通过 LLMs 设计更有助于深度理解和批判性思维培养的学习体验展开，同时确保教育工作者和管理者具备足够的技术能力，以在不断变化的技术环境中有效应对挑战。当前的 LLMs 往往擅长生成表层信息和解决简单任务，但在培养学习者的批判性思维方面仍有较大改进空间。为此，未来在教育领域配置的模型在设计上应注重引导学习者进行更复杂的认知过程，如分析、推理、评估和创造性思考。比如 LLMs 可以通过生成多视角的讨论和引导性问题，鼓励学习者在理解文本时考虑不同的解释和观点，进而深化他们的理解。此外，随着 LLMs 和其他 AI 技术的不断进步，教育工作者需要持续更新他们的技术知识，培训计划应当涵盖从基础的技术操作到高级的教学设计应用，确保教师能够灵活运用这些工具。同时，教育管理者应具备足够的技术理解，以有效评估和选择适合的技术解决方案，并指导教育政策的制定与实施。

总之，未来的研究应在拓宽应用领域、扩大研究对象的同时，继续深化对 LLMs 在不同教育阶段、语言背景和文化环境中的潜力和局限性的理解。通过多维度、多层次的研究，推动 LLMs 在语言教学领域的全面应用，并为教育公平和质量的提升贡献更多的理论和实践依据。

五、结语

本研究系统回顾了 LLMs 在语言教学中的应用，揭示了其在支持教育者、促进学习者发展以及优化教学管理中的潜力和局限性。研究采用了系统性文献综述的方法，现有研究成果分析显示，在支持语言教育者方面，LLMs 通过优化教学设计、生成教学资源、以及改进教学评估与反馈，显著提升了教学效果。其次，对于语言学习者，模型在个性化辅导、

自主学习支持、以及口语和写作技能练习中表现出色。此外, LLMs 还在优化教育决策、促进教育公平方面展现了重要作用, 能帮助教学管理者更好地理解并回应学生需求, 合理分配教学资源。然而, 模型也可能存在偏见、学生过度依赖和学术诚信危机等局限性, 教师、学生和管理者应具备技术素养, 尽量避免其带来的潜在风险。

当前研究也存在不足之处。研究对象仅选择了核心期刊中的文献, 可能忽略了其他来源的相关研究, 导致研究覆盖面受限。许多创新性和实践性较强的研究成果可能发表在非核心期刊或会议论文中, 这些研究可能提供了不同的视角或应用场景。一些尚未被广泛认可但具有潜在影响力的研究可能因此未被纳入综述, 进而限制了对 LLMs 在语言教学中应用的全面理解。因此, 未来研究应考虑纳入更广泛的文献来源, 涵盖多样化的教育背景和语言环境, 以期全面评估 LLMs 在不同教育体系中的应用成效。这将有助于识别更多潜在的应用场景, 并更全面地了解 LLMs 在语言教学中的优势与挑战。

【参考文献】

- 冯庆华 & 张开翼. (2024). 人工智能辅助外语教学与研究的能力探析——以 ChatGPT-4o 和文心大模型 4.0 为例 [On the Competencies of AI-Assisted Foreign Language Teaching and Research: A Case Study of ChatGPT-4o and ERNIE 4.0]. 外语电化教学 [Computer-Assisted Foreign Language Education] (03), 3-12.
- 李康康, 卢颖翔 & 杨现民. (2024). 大语言模型赋能智能学伴: 系统架构与风险防控 [Large Language Models Empowering Intelligent Learning Companions: System Architecture and Risk Prevention]. 现代远程教育研究 [Modern Distance Education Research] 36 (03), 20-28.
- 苏祺. (2024). 大语言模型在二语教学中的应用效能解析 [An Analysis of the Effectiveness of Large Language Models in Second Language Teaching]. 外语界 [Foreign Language World] (03), 35-42.
- 王华树 & 谢斐. (2024). 大语言模型技术驱动下翻译教育实践模式创新研究 [Innovating Translation Education Practices Driven by Large Language Model Technologies]. 中国翻译 [Chinese Translators Journal] 45 (02), 70-78.
- 魏小东 & 习鹏飞. (2024). 大语言模型驱动下的虚拟数字人第二语言教育应用研究 [Application of Virtual Digital Humans in Second Language Education Driven by Large Language Models]. 电化教育研究 [E-Learning Research] 45 (05), 75-82.
- 许家金 & 赵冲. (2024). 大语言模型在英语教学中的角色 [The Role of Large Language Models in English Language Teaching]. 外语教育研究前沿 [Frontiers in Foreign Language Education Research] 7 (01), 3-10.



- 杨连瑞 . (2024) . ChatGPT 大语言模型背景下的二语习得 [*Second Language Acquisition in the Context of ChatGPT-Based Large Language Models*]. 现代外语 [*Modern Foreign Languages*] 47 (04), 578-585.
- 翟雪松, 季爽, 焦丽珍, 朱强 & 王丽英 . (2024) . 基于多智能体的人机协同解决复杂学习问题实证研究 [*An Empirical Study on Human-AI Collaborative Problem Solving for Complex Learning Tasks Based on Multi-Agent Systems*]. 开放教育研究 [*Open Education Research*] 30 (03), 63-73.
- 郑春萍, 于淼 & 郭智妍 . (2024) . 人工智能在语言教学中的应用研究: 回顾与展望 [*Artificial Intelligence in Language Teaching: A Review and Prospect*]. 外语教学 [*Foreign Language Education*] (01), 59-68.
- Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., Avila, R., Babuschkin, I., Balaji, S., Balcom, V., Baltescu, P., Bao, H., Bavarian, M., Belgum, J., Bello, I., ... Zoph, B. (2023). *GPT-4 Technical Report*. arXiv reprint arXiv: 2303.08774.
- Boudouaia, A., Mouas, S., & Kouider, B. (2024). A Study on ChatGPT-4 as an Innovative Approach to Enhancing English as a Foreign Language Writing Learning. *Journal of Educational Computing Research*, 07356331241247465. <https://doi.org/10.1177/07356331241247465>
- Derakhshan, A., & Ghiasvand, F. (2024). Is ChatGPT an evil or an angel for second language education and research? A phenomenographic study of research-active EFL teachers' perceptions. *International Journal of Applied Linguistics*, 1-19. <https://doi.org/10.1111/ijal.12561>
- Douglas, M.R. (2023). Large Language Models. *Communications of the ACM*, 66, 7 - 7. <https://doi.org/10.1145/3606337>
- Fathi, J., & Rahimi, M. (2024). Utilising artificial intelligence-enhanced writing mediation to develop academic writing skills in EFL learners: a qualitative study. *Computer Assisted Language Learning*, 1-40. <https://doi.org/10.1080/09588221.2024.2374772>
- Ghafour, M., Hassaskhah, J., & Mahdavi-Zafarghandi, A. (2024). From virtual assistant to writing mentor: Exploring the impact of a ChatGPT-based writing instruction protocol on EFL teachers' self-efficacy and learners' writing skill. *Language Teaching Research*, 13621688241239764. <https://doi.org/10.1177/13621688241239764>
- Godwin-Jones, R. (2024). Distributed agency in second language learning and teaching through generative AI. *Language Learning & Technology*, 28(2), 20-31. <https://doi.org/10.64152/10125/73570>
- Guo, K. (2024). EvaluMate: Using AI to support students' feedback provision in peer assessment for writing. *Assessing Writing*, 61, Article 100864. <https://doi.org/10.1016/j.asw.2024.100864>
- Guo, K., & Li, D. L. (2024). Understanding EFL students' use of self-made AI chatbots as personalized writing assistance tools: A mixed methods study. *System*, 124, Article 103362. <https://doi.org/10.1016/j.system.2024.103362>



- Karatas, F., Abedi, F. Y., Gunyel, F. O., Karadeniz, D., & Kuzgun, Y. (2024). Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners. *Education and Information Technologies*, 1-24. <https://doi.org/10.1007/s10639-024-12574-6>
- Kartal, G. (2024). The influence of ChatGPT on thinking skills and creativity of EFL student teachers: a narrative inquiry. *Journal of Education for Teaching*, 50(4), 627-642. <https://doi.org/10.1080/02607476.2024.2326502>
- Kasneçi, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneçi, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, Article 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kim, J., Yu, S., Detrick, R., & Li, N. (2024). Exploring students' perspectives on Generative AI-assisted academic writing. *Education and Information Technologies*, 1-36. <https://doi.org/10.1007/s10639-024-12878-7>
- Kwak, Y., & Pardos, Z. A. (2024). Bridging large language model disparities: Skill tagging of multilingual educational content. *British Journal of Educational Technology*, 55, 2039–2057. <https://doi.org/10.1111/bjet.13465>
- Lee, J., Hicke, Y., Yu, R. Z., Brooks, C., & Kizilcec, R. F. (2024). The life cycle of large language models in education: A framework for understanding sources of bias. *British Journal of Educational Technology*, 55, 1982–2002. <https://doi.org/10.1111/bjet.13505>
- Li, B., Bonk, C. J., Wang, C. R., & Kou, X. J. (2024). Reconceptualizing Self-Directed Learning in the Era of Generative AI: An Exploratory Analysis of Language Learning. *IEEE Transactions on Learning Technologies*, 17, 1515-1529. <https://doi.org/10.1109/TLT.2024.3386098>
- Liu, X. H. (2024). Navigating Uncharted Waters: Teachers' Perceptions of and Reactions to AI-Induced Challenges to Assessment. *Asia-Pacific Education Researcher*, 1-12. <https://doi.org/10.1007/s40299-024-00890-x>
- Liu, Z. M., Hwang, G. J., Chen, C. Q., Chen, X. D., & Ye, X. D. (2024). Integrating large language models into EFL writing instruction: effects on performance, self-regulated learning strategies, and motivation. *Computer Assisted Language Learning*, 1-25. <https://doi.org/10.1080/09588221.2024.2389923>
- Malinka, K., Peresíni, M., Firc, A., Hujnák, O., & Janus, F. (2023). On the Educational Impact of ChatGPT: Is Artificial Intelligence Ready to Obtain a University Degree? *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 1*, Turku, Finland. <https://doi.org/10.1145/3587102.3588827>



- Moorhouse, B. L. (2024). Generative artificial intelligence and ELT. *ELT Journal*, Article ccae032. <https://doi.org/10.1093/elt/ccae032>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery*, 88, 105906. <https://doi.org/10.1016/j.ijsu.2021.105906>
- Prinsloo, P., Khalil, M., & Slade, S. (2024). Vulnerable student digital well-being in AI-powered educational decision support systems (AI-EDSS) in higher education. *British Journal of Educational Technology*, 55, 2075–2092. <https://doi.org/10.1111/bjet.13508>
- Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior*, 160, Article 108386. <https://doi.org/10.1016/j.chb.2024.108386>
- Susnjak, T., & McIntosh, T. R. (2024). ChatGPT: The End of Online Exam Integrity? *Education Sciences*, 14(6), 656. <https://doi.org/10.3390/educsci14060656>
- Tan, K., Pang, T., Fan, C., & Yu, S. (2023). *Towards Applying Powerful Large AI Models in Classroom Teaching: Opportunities, Challenges and Prospects*. arXiv reprint arXiv: 2305.03433.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30, 5998-6008.
- Wan, Y. W., & Moorhouse, B. L. (2024). Using Call Annie as a Generative Artificial Intelligence Speaking Partner for Language Learners. *RELC Journal*. <https://doi.org/10.1177/00336882231224813>
- Yeh, H. C. (2024). The synergy of generative AI and inquiry-based learning: transforming the landscape of English teaching and learning. *Interactive Learning Environments*, 1–15. <https://doi.org/10.1080/10494820.2024.2335491>