

# 人工智能辅助下外语教学中的情绪研究

陈思蓓 董静希 刘桂玲

吉林大学 外国语言文化学院, 吉林省长春市, 130012;

**摘要:** 生成式人工智能在语言教学中的应用引起了越来越多的学术关注。本研究分别对 AI 辅助外语教学和情绪研究进行文献综述, 前者结果显示, 以 ChatGPT 为代表的聊天型机器人被广泛用于写作、口语、翻译等外语技能训练, 主要集中于高等教育阶段和英语, 但亦存在过度依赖导致创造力下降和伦理风险等问题。后者显示, 生成式人工智能能够增强积极情绪, 降低消极情绪。两个方向的研究方法都以混合研究为主, 重主观数据, 总体呈现跨学科化、情境多样化与工具多元化趋势。

**关键词:** 生成式人工智能; 外语教学; 情绪研究; 数字化学习

**DOI:** 10.64216/3080-1494.26.02.024

## 引言

外语教育(foreign language education, 以下简称 FLE)多发生在非母语环境中, 存在语言环境缺失、互动不足等问题。人工智能所创设的虚拟环境能够弥补这一缺憾, 提高学习成绩, 且 AI 课堂能够为学习者提供情绪价值。但是在最新的生成式人工智能技术(generative AI, 以下简称 GenAI)辅助 FLE 中, 学习者的情绪状态如何, 情绪状态对学习效果产生多少影响? 这些都是外语教学研究的新兴领域和重要课题。基于此, 本文综述近年来有关 GenAI 辅助 FLE 和情绪作用的研究进展, 为相关教学与研究提供参考。

## 1 人工智能辅助外语教学研究

### 1.1 传统 AI 辅助外语教学研究

自然语言处理技术是语言教学领域历史最悠久、最具生命力的人工智能技术, 人工智能是一项利用机器模仿或执行通常需要人类智能的认知功能的技术。传统 AI 辅助外语教学遍布各个教学领域, 其中应用最多的领域是写作, AI 评估并指导学生作文, 将教师从修改语病等简单工作中解脱出来, 但教师仍需承担讲解等主体作用。而传统 AI 在翻译中的主要应用是预翻译、构建语料库和术语库, 虽然能够减少认知负荷、提高翻译质量, 但实践中主要集中于译前和译后, 人类翻译被认为优于机器翻译。自动语音识别(ASR)是利用计算机将语音信号转换成文本的技术方法, 基于 ASR 的语音测评系统被广泛用于口语学练, 然而其主要遵循“模仿跟读一打分”的机械重复模式, 对于提升语言组织能力质量效果不强。

### 1.2 生成式人工智能辅助外语教学系统性研究

#### 1.2.1 简介

传统 AI 模型建立在预定义的对话场景和有限的语料库数据之上, 不能适应脚本交互之外的意外响应, 限制了 AI 在外语教育中更广泛的运用。美国 OpenAI 推出 ChatGPT, 使 GenAI 技术大众化。要了解 GenAI, 就要先了解 ALM 的原理。高级语言模型( ALM )是使用深度学习技术构建的, 通过计算机处理和生成人类语言, 从而增强人与机器之间的人工智能系统, 具体应用包括 GPT-4、Google 的 BERT 和 Facebook 的 RoBERTa 等。随着 ALM 的发展, 近年来出现了 ChatGPT 、 Deepseek 等大型语言模型( LLM ), LLM 是 ALM 的一个子集, 其中“大”是指庞大的训练数据和参数。此外, GenAI 以多模态的形式输出, 可以实现自发多样的交互, 为外语教育提供新范式。

#### 1.2.2 研究方法

将范围精确到 GenAI 大规模投入使用后, 本研究对 GenAI 辅助外语教学的文献进行系统综述, 我们采用 PRISMA 指南, 使用 Web of Science、Scope 和知网梳理 2023 年 1 月到 2025 年 9 月的文章。

确定搜索关键词为“生成式人工智能”或“GenAI”或“ChatGPT”或“Copilot”或“Claude”或“Gemini”或“Midjourney”和 Foreign Learning Education(“外语教育”或“外语学习”或“FLE”或“SL”或“ESL” )。使用“标题”“摘要”和“关键字”作为过滤标准, 我们最初从三个数据库中找到了 1075 篇中英文文章。在排除重复或本研究无关的文章后, 我们纳入了以下研究:

(1) 采用实证研究设计, (2) 使用特定的 GenAI 工

具, (3) 关于二语或外语学习的研究。在本次筛选阶段我们获得了 134 篇文章。接下来, 进一步全文检索,

排除 (1) 案例研究 (2) 学科发展的整体性看法, 最终纳入了 59 项研究。

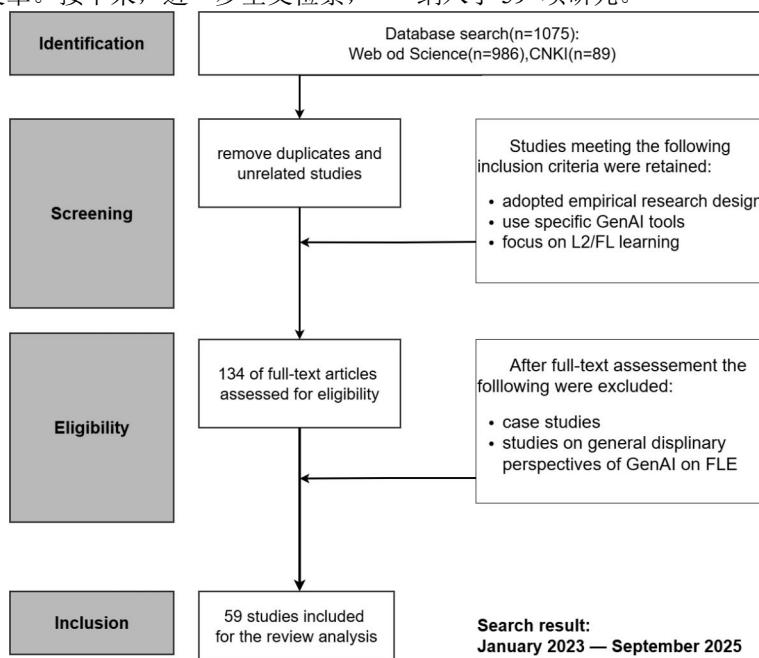


图 1. PRISMA 流程图

### 1.2.3 研究结果

#### (1) 研究设计

研究地区分布呈现以亚洲为中心, 北多南少的格局, 其中 35 项研究来自于亚洲, 在欧洲进行了 6 项研究, 非洲进行了 1 项研究, 体现出对 GenAI 辅助外语教学的全球性关注。

研究的目标语言多为英语, 非英语研究共有 3 项, 2 项汉语 1 项日语。可能是因为 ChatGPT 目前仍以英语为主导语言, 以 GPT-3 模型为例, 93% 的训练数据主要是英语, 其余 7% 是其他语言, 而语言模型的数据量越大, 机器人的性能就越好。

教育层次多为高等教育的低阶段, 关注硕博阶段的研究各有一项, 随着研究的发展有向学前和初中等教育扩散的趋势。

大多数研究采用混合方法 (n=23), 纯粹的定性方法 (n=16) 和定量研究方法 (n=17) 占比相当。数据来源方面, 传统的访谈 (n=16)、调查问卷 (n=20) 仍为主流, 也包括测试成绩 (n=3)、日志 (n=2) 和口语录音等多模态数据 (n=3), 一项研究使用了眼动追踪数据。可得出, 目前的实证研究对访谈、实验日志等主观数据的依赖较为严重, 容易陷入主观偏见, 且缺乏深入的循证方法和客观的观察性研究。

#### (2) 研究内容

对写作的研究仍占绝对多数 (n=32)。研究发现, GenAI 通过提供实时反馈, 能帮助学习者扩大词汇量、改善句子结构, 提高写作技能。此趋势离不开传统 AI 的大量研究, 也因为 LLM 模型以文本为基础, 有利于培养写作技能。

口语是非英语母语者通过托雅考试的重难点, 且传统教学具有费用昂贵、资源分布不均等问题, 因此现实需求成为研究的热点。且口语领域出现了市场先行的特点, 我们在 Google Play 里搜索 “speaking”、“AI”, 出现下载量超过千万的应用产品三款 (Praktika、ELSA、Learna), 超百万的应用十余款, 除了这些专用的口语学练软件, 更多人使用 Gen AI 软件自带的语音功能, 通过特定指令训练其扮演专业口语陪练, 帮助自己润色表达。

#### 1.2.4 总结

基于此, 本文提出使用 GenAI 辅助 FLE 的四个建议。

鉴于 ChatGPT-4o 等最新 AI 模型的多模态能力, 应该扩展到写作技能之外, 并加速对其它语言的开发。很多非英语国家或小语种专业因为科研能力欠缺、市场狭小, 无力发明使用本国语言的 GenAI 或软件, 各国研究者也要加快本国语料库构建和技术开发。

其次, 应进一步探索 GenAI 在语言学习中的综合

工具能力, 加强与教育理论与使用平衡, 发掘 GenAI 在各方面的作用。传统 AI 虽未能彻底变革师生-课堂教学模式, 但 GenAI 可扮演多种角色, 拓展教育边界, 打破时空、师资壁垒。

但是, 针对 GenAI 是否能彻底取代人类教师的观点, 本文认为 GenAI 不能完全替代同学、教师在社交等多方面的影响。GenAI 虽逻辑通顺, 但缺乏实质内涵, 无法生成真正情感化的文本。一项研究通过对 64 项实证研究的 11 项统计结果进行分析, 发现虽然 AI 在语言复杂性等机械方面与人类相匹配, 但人类教师能通过互动、介入等社交行为进行监管, 维持长期兴趣。然而目前的多数研究没有明确定义人类教师的角色或让教师参与人工智能语言学习。

而当前科技在反馈质量、学习方法和策略指导方面还存在明显不足, 未来研究应侧重于培养学生批判素养, 提高识别错误、理清谬误的能力。教师也应该提高创造力、协作和批判性的 AI 素养, 深入对学习方法和反馈的研究。

最后, 一些调节因素会影响 GenAI 的效果, 如个体差异, 包括知识水平、学习动机和认知风格等主观因素, 以及不同的设计特点、交互质量以及环境等客观因素。今后研究应提出针对不同场景下针对不同学生的因材施教的教学策略。

## 2 情绪研究综述

### 2.1 实体教学中的情绪研究

传统二语习得研究主要采用纯认知研究路径, 对情绪等非认知因素不甚关注。但学习者并非被动机器, 应重视其情绪体验和主体地位。1998 年情绪在学生学习与成就中的作用学术大会在美国成功召开, 吸引了各国学者对学生情绪的广泛关注。Krashen 创建的情感过滤假说奠定了情绪研究的基础, Swain 认为认知与情绪在二语学习中起着重要作用, 强调重视情绪。中国也于 2001 年提出的新课程改革中加大了对学生情感态度的关注。近年来, 二语学习情绪研究不断从邻近学科引入新的研究方法与理论基础, 已然成为应用语言学前沿领域。

早期情绪研究以消极情绪为主, 后发生了“积极转向”。Horwitz 提出的外语焦虑的概念和《外语课堂焦虑量表》的开发将消极情绪研究引向深入。中国学者对于外语焦虑研究兴起于 20 世纪 80 年代后期, 研究从多角度分析了影响成人外语学习的恐惧和焦虑的成因及

其表现, 在中国非英语专业大学生中展开了混合研究, 开发了系列问卷。研究表明, 无聊与语言焦虑等负面情绪呈现正相关, 而与一系列积极心理因素呈现不同程度的负相关。即无聊情绪会通过抑制学生积极情绪、放大负面情绪影响学习结果。“积极心理学”概念最早由人本主义心理学家亚伯拉罕·马斯洛提出。此后中文版《外语愉悦量表》和二语坚毅量表陆续得到开发, 视角也拓展到其他主体, 如来自中国 11 所大学的 2002 名英语学习者感知到的教师热情、情绪和社会行为学习参与之间的复杂关系。

### 2.2 虚拟教学中的情绪研究

除了传统课程外, 二语习得情绪研究在 AI 辅助外语教育方面也具有巨大前景:

结合技术接受理论, 研究者们采用 UTAUT 或 TAM、TAME 等模型研究学生对这一新技术用于外语教学的接受度, 结果大都体现接受意愿, 受到感知有用性、易用性、较低焦虑度等因素影响。

另外, 教学效果影响因素研究是也主要方向之一, 大多通过对照或随机分配的干预设计, 采用混合方法, 探讨学习者即时情绪、个人特质、环境等因素的中介效应或调节作用。例如: 来自伊朗的 492 名英语学习者填写了关于 ChatGPT 辅助外语学习情况、幸福感和情绪调节的问卷, 结构模型结果显示, 情绪调节在使用 ChatGPT 与幸福感之间具有强烈的中介效应。此外, 在使用 ChatGPT 与幸福感和情绪调节之间发现了显著的正相关关系。

近年还有很多类似研究, 得出结论也较为一致: AI 辅助英语学习效果与积极情绪呈正相关、能够降低消极情绪。实验组成绩大多相较于对照组有提升, AI 在外语成绩提升中的中介效应较显著, 与性别、年龄、地区等因素相关性较小。未来研究应寻求新突破、如在不同教育层面进行长期实验, 开发基于情绪变化机制的 AI 辅助外语教学的方法, 或基于心理学理论, 升级原有系统、情绪响应式教学等情绪调节策略, 使其适应 GenAI 背景。例如, Xin (2025) 提出了情绪感知-适配教学模型, 即利用学习者情绪实时检测(面部表情、语音情感)动态调节任务难度、反馈频率与激励策略。

目前 AI 辅助外语学习中的情绪研究的关注重点主要在高等教育人群, 这可能是因为大学生对情绪具有更高的认知能力。但也提醒了未来的研究者应该不拘一格,

向K-12、留守儿童群体、欠发达地区、非课堂等领域扩张。在一些发展中国家，社会观念发展较慢，学生心理需求长期得不到关注，且面对AI人们更容易揭露自我。因此，推广AI辅助FLE的在提升教育公平的同时还能够为孩子提供成长陪伴，缓解不良情绪。

### 3 总结

综上所述，我们梳理了何为人工智能如何辅助外语教学，对AI领域突破性技术生成式人工智能聊天机器人的研究进行了系统性综述，得到了关于其研究层次、设计、内容的结论，提出了四个建议。我们注意到情绪对外语学习、教育的影响，分别对实体课堂和虚拟教学中的情绪研究进行了综述。

### 参考文献

- [1] Barrot JS. Exploring ChatGPT in EFL teaching: Potentials and pitfalls[J]. TESOL Quarterly, 2023, 57(4): 1229–1245.
- [2] 蒋国武. 基于AI测试的宏微观音系语训实证效应研究[J]. 中国电化教育, 2023(8):125–133.

[3] 李成陈, 李嵬, 江桂英. 二语学习中的情绪研究: 回顾与展望[J]. 现代外语, 2024, 47(1): 63–75.

[4] 梁君英, 刘益光. 人类智能的翻译能力优势——基于语料库的人机翻译对比研究[J]. 外语与外语教学, 2023(3): 74–84, 147–148.

[5] 孙海琴, 李可欣, 陆嘉威. 人工智能赋能语音识别与翻译技术对同声传译的影响: 实验与启示[J]. 外语电化教学, 2021(6): 75–80, 86, 12.

作者简介: 陈思蓓 (2005.04-), 女, 汉族, 四川省成都市, 吉林大学, 本科在读, 研究方向: 应用语言学, 外语教育。

董静希 (2005.06-), 女, 汉族, 湖南省衡阳市, 吉林大学, 本科在读, 研究方向: 应用语言学, 外语教育。

基金项目: 本文为吉林大学2025年大学生创新创业训练计划《基于混合方法的AI辅助英语口语训练的情绪及效果研究》编号: S202510183042、吉林大学研究生教学改革项目“数智赋能语言学课程群PBL教学中数字素养与科研能力提升研究”(2025RGZNP006)成果。