

人机协同视角下人工智能在中文教育中的应用研究：基于 2020-2026 年文献的内容分析

孔维俭, 陈莹, 吴云东*

韩国东新大学, 罗州

摘要 | 随着生成式人工智能、大语言模型等技术快速发展, 人工智能正在逐步进入中文教育的教学设计、资源生成、语言训练、学习评价与教师发展等环节。本文从人机协同视角出发, 对 2020—2026 年 7 月中文教育领域人工智能应用研究进行内容分析。通过中国知网、万方、维普及 Google Scholar 等数据库进行检索、筛选、去重和全文核验, 最终形成 33 篇中文学术文献作为核心分析样本, 并从技术类型、研究主题、教学内容和研究方法等维度考察该领域的研究特征与发展动向。研究发现: 2023 年是相关研究明显增长的重要节点, 研究关注点由早期的技术应用与智慧教育逐渐转向生成式人工智能支持下的教学实践与人机协同; 生成式人工智能和大语言模型已成为研究最集中的技术类型, 并主要应用于教学材料生成、写作支持、问答反馈、教学设计和个性化学习; 研究主题仍以宏观理论、数字化转型和发展路径为主, 但教学系统、资源、写作和口语等具体应用研究逐步增加, 而阅读、汉字、语法和文化教学研究相对不足; 研究方法仍以理论阐释、综述和策略分析为主, 长期追踪、多中心和对照研究较为有限。基于此, 本文提出“教师定向、AI 辅助、学生加工、教师校准”的人机协同机制, 并建议未来加强中文本体知识约束、学习效果的长期验证、教师 AI 素养培养、专用数据与评价标准建设以及伦理治理, 以推动中文教育人工智能应用由工具辅助向可验证、可持续的人机协同教学发展。

关键词 | 中文教育; 国际中文教育; 人工智能; 生成式人工智能; 大语言模型; 研究动向

Copyright © 2026 by the author(s).

This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

<https://creativecommons.org/licenses/by-nc/4.0/>



1 引言

人工智能进入教育并不是简单地把一种新软件放进课堂, 它会改变教师准备材料的方式, 也会改

变学习者获取解释、练习语言和接受反馈的路径。从教育整体视角看, 人工智能在提升学习效率、资源配置与个性化支持方面具有现实价值, 同时也伴随数据治理、学术诚信、技术依赖和教育公平等问题^[1]。这类判断放到语言教育中尤其值得注意, 因为

作者简介: 吴云东, 东新大学博士研究生, 主要从事教育心理学、创新创业教育、高等教育方向研究。

文章引用: 孔维俭, 陈莹, 吴云东. 人机协同视角下人工智能在中文教育中的应用研究: 基于 2020-2026 年文献的内容分析[J]. 澄启学刊, 2026, 1 (3) : 109-120.

<https://doi.org/10.68070/1b36ft03>

语言学习并不只处理“答案是否正确”，它还涉及语境、文化、交际意图与长期能力形成。

生成式人工智能对语言学习的影响进一步放大了这种复杂性，在人工智能与科技高速发展的今天，生成式人工智能可以服务听、说、读、写，但只有把语言本体特征、学习目标和教学控制结合起来，技术优势才可能转化为稳定的学习收益^[2]。这一认识对中文教育同样具有启发意义，尤其是汉字系统、声调、语序、语用和文化知识都具有较强的语言特异性，通用模型给出的流畅回答并不必然等于适合教学的回答。

中文学术研究强调，生成式人工智能进入教育应坚持以人为本，并将教师判断、内容审核、学习者主体性与技术风险纳入教学设计^[3]。这意味着中文教育研究不能只统计工具数量或罗列功能，还需要继续追问：技术怎样进入真实课堂，教师怎样保留专业判断，学习者的语言能力是否真正得到改善^[4]。

已有人工智能教育研究为这类分析提供了方法参照，学界研究表明，人工智能教育应用正在经历从技术主导向教育赋能的范式演进，智能辅导、自适应学习和自动评测逐步成为核心场景^[5-7]。

大语言模型出现后，研究议题发生了明显变化。相关元分析与综述研究指出，大语言模型（Large Language Models, LLMs）在个性化互动与材料生成方面展现出显著优势，但技术幻觉、过度依赖与长期教学效果验证不足仍是当前研究关注的关键挑战^[8-12]。

2 分析对象与方法

2.1 文献范围与样本构成

本研究文献检索与全文核验截止至 2026 年 7 月。数据源主要基于中国知网（CNKI）、万方数据知识服务平台、维普中文期刊服务平台以及 Google Scholar；针对部分文献，通过期刊官方网站及出版社公开页面进行全文核验。本研究未将外文文献及国际组织报告纳入分析范畴。

检索策略采用主题词与自由词结合的方式，逻辑组合包含两类核心概念：一是中文教育领域，包括“中文教育”、“国际中文教育”、“汉语教学”、“语文教育”；二是人工智能技术，包括“人工智能”、“生成式人工智能”、“大语言模型”、“ChatGPT”、“自然语言处理”、“智能体”、“智能评测”、“智慧学习系统”。

文献筛选严格遵循拟定的纳入与排除标准，纳入标准：（1）研究主题直接聚焦于中文教学、中文学习、国际中文教育或语文教育；（2）研究内容明确涉及人工智能及其相关技术；（3）具备可获取的完整文本用于深度内容分析。排除标准：（1）一般性教育技术研究或未结合中文学科特性的技术应用；（2）重复出版或信息缺失文献；（3）非目标时间窗口内的文献。

数据库初始检索共获得文献记录 472 条；经软件与人工去重后，保留 386 条；通过阅读题名与摘要进行初审，排除不相关文献 283 条，进入全文资格审查阶段 103 条；对全文进行深度评估后进一步排除 70 条；最终精细纳入 33 篇符合标准的中文文献作为本研究的分析样本。文献筛选流程如图 1 所示。

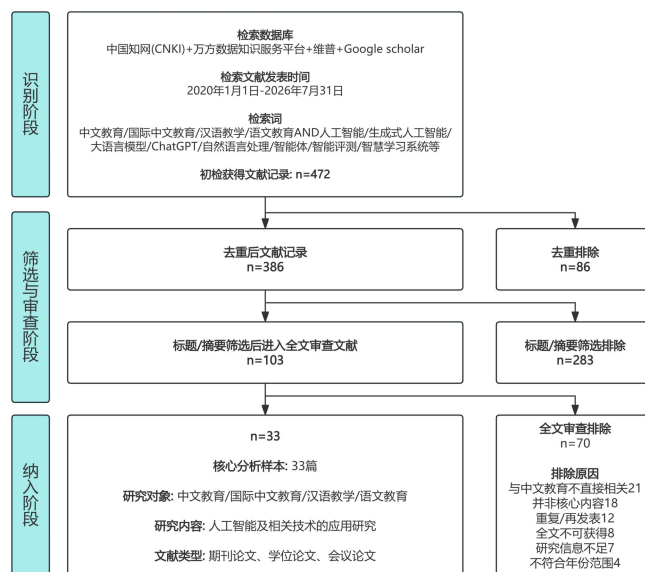


图 1 文献筛选流程图

Fig. 1. Literature Screening Process

基于上述筛选结果，本研究的分析与讨论均建立在最终纳入的 33 篇中文学术文献基础之上。需说明的是，研究结论仅揭示当前中文学术文献所呈现的研究特征与范式，不外推至全球范围内的中文教育研究总体趋势。

2.2 分析维度与编码原则

分析框架保留人工智能技术、教育内容两个基本方向，同时增加年份与研究方法两个维度，为了使“人机协同”成为贯穿全文的分析视角，本文进一步把技术角色、教师决策、学习者活动和 AI 输出控制纳入编码解释：AI 主要承担高重复、可计算、可生成的任务；教师承担目标设定、内容筛选、风险判断和高阶反馈；学习者承担理解、判断、修改、迁移与反思。

本研究由研究者完成两轮编码：第一轮依据预设标签进行初编码，第二轮对技术、主题和研究方法存在交叉的文献进行复核。单标签主题采用“主问题优先”规则，即以论文研究问题、核心研究目的及主要分析对象为依据，而非以文中出现的技术数量决定类别。对于“宏观讨论/教学系统”“设计开发/案例实践”等边界案例，以研究成果最终交付物和证据形态裁决：形成可运行系统、模型或资源原型者归入设计开发；以真实课堂、教学实践或具体应用过程为主要证据者归入案例/实践研究。由于原始研究通过二轮复核和规则化编码提高可重复性。

在技术分类的基础上，本文进一步从中文教育自身的教学内容与应用需求出发构建研究内容分析维度，具体框架如表 2 所示。

表 1 中文教育人工智能应用的技术分析框架
Tab. 1. Analytical Framework for Artificial Intelligence Applications in Chinese Language Education

一级类别	代表性技术	在中文教育中的主要用途
生成式人工智能	大语言模型（LLM）、人工智能生成内容（AI Generated Content, AIGC）、提示工程	教案与材料生成、对话练习、写作反馈、题目生成、解释与问答
语言智能与自然语言处理（NLP）	分词、句法/语义分析、可读性计算、检索增强生成（Retrieval-Augmented Generation, RAG）	文本分级、作文分析、语言特征识别、个性化材料匹配
对话系统与智能体	Chatbot、教学智能体、智能辅导系统	连续互动、任务驱动练习、学习支持、文化与情境教学
语音智能	自动语音识别（ASR）、语音合成（TTS）、语音分析	发音识别、口语训练、听力材料生成、实时纠错
智能评测与学习分析	自动评分、诊断、学习数据分析	HSK 模拟题、作文评分、口语评价、学习路径调整
多模态与智慧环境	多模态生成、虚拟情境、智慧平台	图文音视频资源生成、情境交际、智慧课堂与资源编排

注：技术类别为分析性归纳，一项研究可以同时进入多个技术标签。

表 2 中文教育研究内容的分析框架
Tab. 2. Analytical Framework for Research Contents in Chinese Language Education

分析层级	具体内容	主要观察问题
宏观与治理	数字化转型、智慧教育、伦理与制度	人工智能为何进入中文教育，边界和治理条件是什么
教学系统	课程、教学设计、资源、平台、评价	技术怎样嵌入备课、教学、练习与测评全过程
教师与学习者	教师 AI 素养、专业发展、学习体验	人机分工、教师判断、学习者主体性如何变化
语言知识	语音、汉字、词汇、语法、语篇	技术能否处理中文本体特征及其教学难点
语言技能	听、说、读、写	人工智能对具体技能训练与反馈的作用是否稳定
文化与文学	文化知识、文学阅读、跨文化表达	生成内容能否保持文化准确性、语境感和价值判断

中文教育的内容划分不再照搬韩国语教育的原有结构。本文把教学系统、教师与学习者、语言知识、语言技能、文化与文学等放到同一框架中考察。这样更适合当前中文教育研究的实际情况，因为近年的大量成果已经超出单一的听、说、读、写的训

练，转向智慧课程、资源生成、教师备课、评价、智能体和教育治理等更复杂的教学活动。
按照上述筛选与分类原则，本文最终形成的文献池构成及各层级文献的具体用途如表 3 所示。

表 3 文献池构成与使用方式
Tab. 3. Composition and Utilization of the Literature Pool

文献层级	数量	主要用途
中文理论与背景支撑文献	12	界定人工智能教育的机会、风险、技术范式与语言教育证据
中文教育人工智能应用核心样本	33	用于年份、主题、技术和研究方法的统计与内容分析
合计	33	构成全文参考文献体系

注：核心样本统计范围为 2020-2026 年 7 月公开可核验文献。

3 中文教育人工智能应用研究的总体演进

3.1 从技术试用到生成式人工智能集中进入

从本文核心样本看，2020—2022 年的相关研究主要关注人工智能在中文教育中的技术可用性与智慧教育基础^{[1] [13-15]}；2023 年以后，生成式人工智能尤其是 ChatGPT 成为重要讨论对象^[2-5]；同期关于语言智能教育、国际中文智慧教育、汉语学习与教学以及中文写作辅助的研究进一步丰富了这一方向^[6-11]。

2024 至 2026 年，研究进一步向多模态资源生成、教师数字素养提升、专属系统构建、人机协同口语教学及纵向实证推进^[12-27]。中文学术文献中的相关研究亦呈现由单纯工具辅助向教学指导、人机协同和教师 AI 能力建设发展的趋势^[28-30]。
从本文样本看，2024 至 2026 年出现了多模态资源生成、教师数字素养、系统构建、人机协同口语教学等研究；其中真正具有纵向、持续验证特征的研究仍然有限^[12-27]。
综合核心样本的发表时间可以发现，中文教育人工智能应用研究在不同年份呈现出较为明显的阶段性变化。表 4 合并整合了核心样本文献的年度分布、主要技术类型及研究主题演进概况。

表 4 核心样本文献的年度分布
Tab. 4. Annual Distribution of Core Sampled Literature

年份	文献数	占核心样本比例	主要特征
2020	1	3.0%	以人工智能辅助中文教学的可能性讨论为主
2021	1	3.0%	人工智能与智慧课堂、教学创新开始结合
2022	2	6.1%	技术综述与国际中文智慧教育概念逐步形成
2023	11	33.3%	ChatGPT 带来集中增长，宏观讨论、教学变革和写作应用同步出现
2024	5	15.2%	多模态资源、教师备课、教师素养和具体技能教学增多
2025	8	24.2%	系统构建、模型测评、写作反馈、阅读分级和文化智能体并进
2026	5	15.2%	口语实证、人机协同与教师 AI 素养进一步深化
合计	33	100%	研究由工具试用逐步转向模式、系统与证据建设

3.2 技术类型由单一工具转向复合系统

从本文核心样本的技术标签看，生成式人工智能和大语言模型已经成为最密集的研究对象。语言智能与自然语言处理仍然是中文教育的重要技术基础。阅读分级、写作反馈、语篇分析、文本生成和

智能评测都需要对汉语字词句篇进行计算处理。中文与拼音文字在文字系统和词边界方面存在差异，因此，通用模型即使能生成顺畅文本，也仍需结合中文分级标准、词汇语法规则和学习者语料进行适配。

对话系统和智能体的增长值得单独注意。较早的聊天机器人主要提供问答，现在的智能体更强调任务链、资源调用、持续反馈和个性化支持。它可能成为学习入口，也可能串联练习、反馈、评价等多个步骤。由此带来的研究问题也变了：过去关心机器人能否回答，现在更需要考察智能体能否在教师设定的目标下持续完成教学任务，并且让学习者保持主动思考。但在本文样本中，2025-2026 年的口语研究数量相对前期有所增加；由于语音智能仅涉及 4 篇文献，这一变化仍应视为初步趋势。

表 5 核心样本涉及的人工智能技术类型（多标签统计）
Tab. 5. Types of Artificial Intelligence Technologies Involved in the Core Samples (Multiple-Label Statistics)

技术类型	文献数	主要应用
生成式人工智能/大语言模型	23	材料生成、问答、反馈、备课、写作、题目生成与教学设计
语言智能/NLP	11	中文文本分析、分级、写作分析、语言特征识别与 RAG
对话系统/智能体/智能辅导	10	交互练习、任务支持、连续反馈和文化教学
智慧平台/学习系统	6	课程组织、个性化学习、资源管理与全过程支持
智能评测/学习分析	6	作文评分、HSK 模拟、诊断、文本等级与学习效果分析
语音智能	4	口语训练、发音反馈、语音交互与高级口语系统
多模态生成	3	图文音视频资源、情境材料与文化资源生成

注：一篇文献可以同时涉及多种技术，因此各项相加大于 33。

4 教学技术与教学环节的研究变化

4.1 宏观应用研究主导与微观课堂实践细化

把 33 篇核心文献按主要研究主题单一归类后，宏观理论、发展状态和数字化转型类研究共有 12 篇，占比最高。这一结果并不难理解。生成式人工智能进入教育的速度很快，中文教育首先需要回答的是技术能否使用、怎样规范使用、教师和学习者的角色会不会发生变化。2023 年的集中讨论正是这种先建立问题框架的表现。

不过，宏观研究占比高也意味着实践证据仍然显得尚不足够，到 2024 年以后，研究开始向备课、资源、教学设计、学习系统和评价等具体环节移动。教学系统与资源评价类研究共有 8 篇，已经成为第

二大主题。这个变化说明人工智能正在从课外辅助工具进入教师工作的内部流程，例如生成教案、改写文本、设计任务、生成多模态材料、形成模拟试题和提供作文评分。

教师研究目前只有 3 篇被归入主要主题，但其重要性在快速上升。早期研究常把教师视为工具使用者，近年的研究则开始讨论教师的数字素养、人工智能素养、专业判断与人机协作能力。生成式人工智能越能“自动做事”，教师就越需要判断哪些内容可信、哪些反馈适合学习者、哪些任务不能交给模型。这不是技术替代教师，反而使教师的专业筛选能力变得更关键。也正因为如此，2025 年以后样本中出现了若干系统构建和教学模式研究，但现有样本尚不足以证明其已经成为整个领域的主导方向。核心样本的研究主题及其数量分布如表 6 所示。

表 6 核心样本的主要研究主题分布（单一归类）
Tab. 6. Distribution of Main Research Themes in the Core Samples (Single-Category Classification)

主要主题	文献数	比例	研究特征
宏观理论、状态与数字化转型	12	36.4%	讨论机会、风险、智慧教育、技术变革与发展路径
教学模式、系统、资源与评价	8	24.2%	面向备课、资源生成、系统构建、评价和课堂流程
教师发展与 AI 素养	3	9.1%	关注备课能力、数字素养和学科化 AI 素养

续表 6 核心样本的主要研究主题分布（单一归类）
Tab. 6 (Continuation) . Distribution of Main Research Themes in the Core Samples (Single-Category Classification)

主要主题	文献数	比例	研究特征
写作教学	4	12.1%	涉及辅助写作、反馈、智慧模式与学习支持
口语教学	3	9.1%	集中在口语训练、实时支持和课堂效果
阅读/汉字教学	2	6.1%	包括文本分级识别与人机协同汉字教学
文化教学	1	3.0%	以文化资源与教学智能体为代表
合计	33	100%	从宏观讨论逐步向具体技能和课堂系统扩展

4.2 年度主题的变化：由宏观思辨走向深度应用

从本文年度样本分布来看，不同研究主题存在阶段性变化。2023 年 11 篇样本中，宏观类有 6 篇，表明该年度样本首先集中讨论 ChatGPT 带来的影响、风险与转型。

这种变化是较为积极的，但同时也要避免把案例越来越多直接理解成证据已经充分，很多研究依

然使用单次任务、短期课堂或少量学习者数据，模型版本变化又很快，研究结果的可重复性面临现实困难。接下来更需要稳定的任务设计、可比的评价标准和跨学期追踪，而不是不断换一种模型再做一次功能展示。进一步从时间维度考察可以发现，不同研究主题在各年份的分布并不均衡，其具体变化情况如表 7 所示。

表 7 核心样本主要主题的年度变化
Tab. 7. Annual Changes in the Main Research Themes of the Core Samples

年份	宏观 /转型	教学系统 /资源	教师发展	写作	口语	阅读 /汉字	文化	合计
2020	1	0	0	0	0	0	0	1
2021	1	0	0	0	0	0	0	1
2022	2	0	0	0	0	0	0	2
2023	6	3	0	2	0	0	0	11
2024	0	1	2	1	0	1	0	5
2025	2	2	0	1	1	1	1	8
2026	0	2	1	0	2	0	0	5
合计	12	8	3	4	3	2	1	33

5 中文知识与语言技能研究

5.1 写作和口语最先形成较稳定的应用场景

在 33 篇核心样本中，明确把某一种中文教学内容作为中心问题的研究共有 10 篇。

口语研究有 3 篇，数量虽然少于写作，但近两年的增长更快。口语场景对人工智能提出的要求更高，因为系统不仅要理解内容，还要处理语音、发音、流利度、互动节奏和语境。现有研究开始显示实时纠错、话题引导与个性化练习能够增加学习者开口机会，不过学习者对模型的信任程度、人机互动质量与真实交际迁移仍需要更长时间观察。

阅读和汉字各有 1 篇代表性研究。阅读研究开始把中文分级标准、自然语言处理、大语言模型与 RAG 结合，用于文本特征量化和等级识别，这一方向很可能继续增长。汉字教学的人工智能研究则还比较薄弱，现有成果更多是教学设计案例，尚缺少对字形识别、书写过程、部件意识和长期记忆效果的系统实验。

文化教学目前只有 1 篇进入主要主题。生成式人工智能能够快速生成文化说明、情境任务和多模态资源，但文化内容恰恰最容易出现语境失真、刻板化或事实错误。未来文化教学不应只看生成速度，还要建立来源核验、文化解释层级和学习者跨文化理解的评价机制。总体来看，目前明确聚焦某一具

体中文教学内容的研究主要分布于写作、口语、阅读、汉字与文化等领域，其数量及人工智能应用方式如表 8 所示。

若进一步从发表年份进行考察，可以看出人工智能用于具体中文教学内容研究在近几年逐步由写作向口语、阅读、汉字和文化等方向扩展，其年度分布如表 9 所示。

表 8 明确指向具体中文教学内容的研究分布
Tab 8 Distribution of Studies Focusing on Specific Chinese Language Teaching Contents

教学内容	文献数	主要人工智能应用
写作	4	写作支持、纠正性反馈、智慧教学模式、生成式学习支持
口语	3	实时纠错、话题引导、口语训练系统、AI 支持项目学习
阅读	1	文本语言特征量化、等级识别、NLP+LLM+RAG
汉字	1	ChatGPT 背景下人机协同教学设计
文化	1	文化教学智能体、海外资源整合与生成式支持
合计	10	具体内容研究仍明显少于宏观、系统与教师类研究

表 9 具体中文教学内容研究的年度分布
Tab 9 Annual Distribution of Research on Specific Chinese Language Teaching Contents

年份	写作	口语	阅读	汉字	文化	合计
2020-2022	0	0	0	0	0	0
2023	2	0	0	0	0	2
2024	1	0	0	1	0	2
2025	1	1	1	0	1	4
2026	0	2	0	0	0	2
合计	4	3	1	1	1	10

5.2 语言知识本体研究的局限与滞后

与写作、口语相比，把中文语言知识本体作为核心问题的研究在本文样本中相对较少。中文语法是尤其需要补上的方向。大语言模型对高频常规表达往往表现不错，但在近义结构、语体限制、篇章衔接和语用条件上仍可能给出“语法上看似成立、教学上却不够合适”的解释。未来可以把学习者语料、语法项目分级和错误类型库接入检索增强系统，让模型输出不再只依赖通用训练数据。

语音研究也不宜只停留在能识别声音，中文声调、轻声、变调、语流音变以及句调都与交际效果

有关，自动语音识别的转写正确率并不能直接代表发音评价有效。比较理想的路线，是把声学特征、人工评分和课堂任务结合，形成适合中文二语学习者的口语诊断指标。

6 研究方法 with 证据形态

从主要研究方法看，理论阐释、综述和策略讨论共 14 篇；原稿将设计开发与案例研究合并为 10 篇。考虑到两类研究的认识论基础不同，本次修订按题名、研究目的和方法描述进行重新编码，其中设计开发/系统构建 6 篇，案例/实践研究 4 篇。

这一结构与技术发展速度有关，模型版本频繁变化，研究者往往先做概念判断或快速案例，再进入效果研究。然而，中文教育若要真正形成可推广的人工智能教学方案，证据建设必须追上技术更新。尤其需要增加三类研究：一是跨学期甚至跨学年的纵向研究，观察技术辅助能否转化为稳定语言能力；二是包含教师反馈、同伴反馈和非 AI 条件的对照研究，分辨效果究竟来自模型还是来自额外练习时间；三是跨水平、跨母语背景的比较研究，避免用

少量便利样本概括所有中文学习者^[31-33]。

此外，模型输出本身也应成为可重复的研究材料，研究报告需要记录模型名称、版本、提示词设计、采样参数、调用日期和人工审核方式。若这些信息缺失，同一研究很难复现，而生成式人工智能的快速迭代又会进一步放大差异。从整体样本来看，不同研究方法的使用比例存在较为明显的差异，核心样本所采用的主要研究方法及其分布情况如表 10 所示。

表 10 核心样本的主要研究方法分布

Tab. 10. Distribution of Major Research Methods Used in the Core Sampled Studies

主要方法	文献数	比例	说明
理论阐释/综述/政策与策略分析	14	42.4%	用于界定概念、机会、风险、路径与教育转型
设计开发/系统构建	6	18.2%	系统、智能体、教学模式与资源原型的设计、开发或构建
案例/实践研究	4	12.1%	真实课堂、教学实践、应用过程与实践反思
量化/准实验/任务效果测评	5	15.2%	比较模型表现或考察具体学习效果
调查/访谈/混合方法	4	12.1%	关注教师、学习者、应用产品和使用体验
合计	33	100%	案例研究属于实证研究形式；但纵向、多中心和长期对照证据仍不足

7 研究动向、现实问题与后续方向

7.1 人机协同的功能分工与教学机制

从本文样本文献可以进一步提炼出“目标设定—AI 生成/诊断—学习者判断与加工—教师审核与干预—形成性评价—下一轮任务调整”的协同链条。其关键不在于增加一个 AI 工具，而在于重新配置教学控制权：教师掌握课程目标、任务难度、内容真实性和最终评价权；AI 承担材料生成、语言分析、即时提示、重复性反馈和学习数据整理；学习者则必须保留对 AI 输出的验证、选择、修改和迁移责任。由此形成的协同机制可概括为“教师定向、AI 辅助、学生加工、教师校准”。这一机制能够解释为什么写作、口语等高频互动任务较容易形成应用场景，也解释了为什么中文语法、文化和高阶语用任务仍需要更强的教师介入。

以写作教学为例，现有样本支持的协同路径可以具体化为：教师根据学习者水平设定写作目标与语言限制；AI 生成示例、提供初步反馈或指出可能的语言问题；学习者比较 AI 建议与自身表达并完成修改；教师审核关键错误、解释语用差异并进行

高阶反馈；最后通过独立写作或迁移任务检验学习效果。口语教学则可形成“AI 情境对话—语音/表达即时提示—学习者自我修正—教师针对性干预—真实交际任务迁移”的路径。上述路径是对样本文献中已有实践的机制化归纳，并不意味着所有研究已经完整实现了这一闭环。

到了 2025-2026 年，样本中出现智能体、智慧教学模式和系统构建研究，但这些成果的数量仍然有限，尚不能据此判断其已成为整个领域的增长主轴。这种变化意味着教学设计不能继续以简单地在教、学的过程中加一个 AI 环节为目标，更有价值的做法，是重新安排教师、学习者和人工智能各自承担什么任务。例如，模型可以快速生成多版本材料，教师负责确定难度、筛查事实和组织任务；模型可以提供第一轮反馈，学习者先判断并修改，教师再处理高层次语言和内容问题，这样的人机分工比学生直接问模型、接受答案更具有相当的教学意义。

智能体的发展还会把若干原本分散的功能串起来。一个中文学习智能体可以读取学习者水平、生成练习、记录错误、提供反馈，再根据表现调整下一轮任务。技术上这已经越来越可行，真正难的是

教育规则：什么错误应该立即纠正，什么问题应该让学习者自己发现，哪些内容必须由教师解释，学习数据保存多久。未来研究要把这些规则写进系统设计，而不是让模型自行决定。

7.2 研究热点与中文教育核心需求之间仍有距离

目前研究热点高度集中于大语言模型，但中文教育的核心难点并没有因此自动解决，大语言模型擅长自然语言生成，所以写作、问答和材料生成自然成为高频场景；可是中文教育还包含声调、汉字、词汇分级、语法结构、语篇组织与文化语用，这些内容需要更细的语言学约束。

目前阅读分级与文本识别的研究已经展示了一种值得延伸的方式：把教育标准、规则系统、NLP 和大语言模型结合，而不是让模型单独工作。同样的思路可以用于语法、词汇和文化教学。比如，面向某一级别学习者生成例句时，可以先用课程标准限定词汇与语法范围，再由模型生成，最后通过自动检测和教师审核完成筛选。这样会比开放式生成更接近教学实际。

第三条动向是本文样本中的具体技能研究呈现“输出先行”的趋势，写作和口语研究数量高于阅读、汉字和文化。该结果应理解为样本结构，而非一线课堂应用规模的直接测量。

7.3 效果、伦理与主体性需要同时进入评价框架

人工智能教育研究不能只记录效率，先前教育研究已经提醒，人工智能可能提高资源获取和个性化支持，同时也会带来技术依赖和学习质量风险[34]。UNESCO 的相关指南强调，数据隐私和人的主体地位应成为教育应用的基本前提^[3]。对于中文教育而言，这两点需要落实到具体课堂指标中。

一方面，效果评价应从模型回答得好不好转向学习者是否学会，一份由模型生成的高质量作文并不能证明学习者写作能力提高；一次顺畅的人机对话也不能证明学习者能在真实交际中完成同样任务。比较合理的评估应至少区分即时任务表现、迁移表现与延迟保持效果。

另一方面，Zhang 等人的研究显示，在 AI 支持的中文学习中，技术信任过高可能与较弱的学习自主性相关^[28]，这说明喜欢用 AI 和因 AI 而学得更好并不是同一件事。后续研究可以把验证行为、修改行为、提示词迭代、主动拒绝错误建议等纳入学习过程数据，观察学习者是否真的在与人工智能协作。评价还应关注文化与伦理。生成式人工智能可能给出流畅但失真的文化解释，也可能在不同地区中文表达、身份称谓和价值议题上呈现偏差^[35]。文化教学尤其需要来源可追溯、人工审核和多视角对照，不能只依据模型的语言流畅度判断质量。综合前文对技术类型、教学内容、研究方法及现实问题的分析，当前中文教育人工智能应用所表现出的主要趋势、存在的问题以及后续需要重点推进的研究任务如表 11 所示。

表 11 中文教育人工智能应用的主要趋势、问题与研究任务

Tab. 11. Major Trends, Challenges, and Future Research Tasks in the Application of Artificial Intelligence in Chinese Language Education

观察到的趋势	当前问题	后续研究任务
生成式 AI 成为核心技术	研究易被模型热度牵引	加强中文本体、课程标准和学习者语料约束
工具向智能体与系统演进	教学规则常被交给模型默认处理	明确教师、学习者、AI 的任务边界与控制权
写作和口语研究增长	阅读、汉字、语法、文化证据偏少	补足理解技能与语言知识的长期实验
资源生成与备课效率提高	生成内容可能出现幻觉与难度失配	建立可追溯来源、质量审核与分级检测机制
自动评价开始进入课堂	评分效度、公平性和可解释性不足	建设中文专用评价数据集并与人工评分持续校准
教师 AI 素养受到重视	培训容易重工具操作、轻专业判断	发展学科化 AI 素养、伦理意识和教学决策能力
实证研究数量增加	短期、小样本、版本不可复现问题仍多	开展纵向、多中心、对照和可复现研究

7.4 研究边界与解释范围

本文的样本来自公开可核验文献,并以主题边界清楚的 33 篇核心样本开展统计,因此结果更适合解释“进入本文文献池的中文教育 AI 应用研究呈现何种结构”,而不适合估计整个中文教育领域的真实应用规模。特别是关于 2023 年转折、写作与口语相对集中、语音技术上升及系统构建增加等判断,均属于样本内趋势。与此同时,伦理问题在现有样本中更多以理论讨论呈现,专门针对伦理风险、技术依赖和学习者主体性的实证调查仍较少。

8 结论与启示

从本文核心样本看,2023 年是生成式人工智能研究集中出现的明显节点;2025-2026 年样本中出现更多系统建设与实证研究,但其数量和验证周期仍有限,因此不能将其直接概括为整个领域已经完成转向^[31-33]。

从技术侧看,大语言模型已经成为主线,但中文教育真正需要的是多技术协同。自然语言处理、语音智能、检索增强、自动评测和多模态资源都应根据教学问题组合使用。把所有任务都交给通用大模型既不经济,也不一定可靠。尤其涉及中文分级、语音评价、语言规范和文化内容时,专用数据、规则约束与人工审核仍不可缺少。

从教学侧看,现有研究仍有明显的不均衡^[36],宏观讨论和系统设计较多,具体语言内容研究较少;在具体内容中,写作和口语走在前面,阅读、汉字、语法、词汇和文化研究还需要扩展。下一阶段不宜只追逐新模型名称,更应围绕真实学习目标持续积累证据^[37]。

一是加强学习效果的长期验证,研究应把即时生成质量与学习者能力增长区分开来,并通过延迟测验、迁移任务和对照条件判断人工智能的真实贡献。

二是建设面向中文教育的专用数据与评价标准,不同水平学习者语料、中文分级材料、常见偏误库、语音样本和文化知识资源,可以为模型训练、检索增强和自动评价提供更可靠的基础。

三是把教师专业判断写入人机协同机制里面,人工智能可以承担高重复、可标准化的工作^[38],但课程目标、内容适切性、复杂错误解释和文化价值判断仍需要教师负责。教师 AI 素养的核心也不应只是会不会使用某个模型,而应包括验证、取舍、解释与教学决策^[29]。

四是完善全过程评价体系,未来系统需要覆盖备课、学习任务、过程反馈、形成性评价和总结性评价,并对模型版本、提示设计、数据处理和人工

审核方式留下记录。只有这样,研究结果才更容易复现,也更便于不同机构比较。

人工智能正在为中文教育提供新的教学条件,但技术越强,越需要清楚的教学目标和边界。真正值得期待的不是“人工智能替代中文教师”,而是通过可验证、可解释、可持续的人机协作,让教师有更多精力处理高层次教学问题,让学习者得到更及时又不过度依赖技术的支持。

参考文献

- [1] MIAO F, HOLMES W, HUANG R, et al. AI and education: Guidance for policy-makers[R]. Paris: UNESCO, 2021.
- [2] MIAO F, HOLMES W. Guidance for generative AI in education and research[R]. Paris: UNESCO, 2023.
- [3] 刘利, 刘晓海. 关于国际中文智慧教育的几点思考[J]. 语言教学与研究, 2022(5): 1-9.
- [4] 刘利, 史中琦, 崔希亮, 等. ChatGPT 给国际中文教育带来的机遇与挑战——北京语言大学与美国中文教师学会联合论坛专家观点汇编[J]. 世界汉语教学, 2023, 37(3): 291-315.
- [5] ZAWACKI-RICHTER O, MARÍN V I, BOND M, et al. Systematic review of research on artificial intelligence applications in higher education — where are the educators?[J]. International Journal of Educational Technology in Higher Education, 2019, 16: 39.
- [6] OUYANG F, JIAO P. Artificial intelligence in education: The three paradigms[J]. Computers and Education: Artificial Intelligence, 2021, 2: 100020.
- [7] CHEN L, CHEN P, LIN Z. Artificial intelligence in education: A review[J]. IEEE Access, 2020, 8: 75264-75278.
- [8] KASNECI E, SESSLER K, KÜCHEMANN S, et al. ChatGPT for good? On opportunities and challenges of large language models for education[J]. Learning and Individual Differences, 2023, 103: 102274.
- [9] TLILI A, SHEHATA B, ADARKWAH M A, et al. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education[J]. Smart Learning Environments, 2023, 10: 15.
- [10] KOHNKE L, MOORHOUSE B L, ZOU D. ChatGPT for language teaching and learning[J]. RELC Journal, 2023, 54(2): 537-550.
- [11] LYU B, LAI C, GUO J. Effectiveness of chatbots in improving language learning: A meta-analysis

- of comparative studies[J]. *International Journal of Applied Linguistics*, 2025, 35(2): 834-851.
- [12] LO C K, YU P L H, XU S, et al. Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies[J]. *Smart Learning Environments*, 2024, 11: 50.
- [13] HUANG S. Research on the application of artificial intelligence in teaching Chinese as a foreign language[C]//*Proceedings of the 2020 International Conference on Social Sciences and Big Data Application*. Paris: Atlantis Press, 2020: 201-204.
- [14] SUMO D S, BAH M L. Chinese language education in the era of artificial intelligence: Innovation development, pedagogy & the smart classroom[J]. *Education Quarterly Reviews*, 2021, 4(4): 264-270.
- [15] MAKSIMOVA A. A systematic review of research on the use and impact of technology for learning Chinese[J]. *International Journal on Cybernetics & Informatics*, 2022, 11(4): 67-78.
- [16] 徐娟, 马瑞凌. ChatGPT 浪潮下国际中文教育的技术变革[J]. *国际汉语教学研究*, 2023(2): 41-52.
- [17] 袁羲, 吴应辉. ChatGPT Plus 给国际中文教育带来的机遇、风险及应对策略[J]. *云南师范大学学报(对外汉语教学与研究版)*, 2023(3): 53-62.
- [18] 马瑞凌, 梁宇. 国际中文教育数字化转型的三重逻辑——从 ChatGPT 谈起[J]. *河南大学学报(社会科学版)*, 2023, 63(5): 112-118, 155.
- [19] 周建设, 薛嗣媛. 论语言智能教育[J]. *语言战略研究*, 2023, 8(4): 30-43.
- [20] 马瑞凌, 徐娟. 语言智能赋能国际中文智慧教育: 现实境况与未来路向[J]. *国际中文教育(中英文)*, 2023(2): 43-52.
- [21] 蔡薇. ChatGPT 环境下的汉语学习与教学[J]. *语言教学与研究*, 2023(4): 13-23.
- [22] 金旋. ChatGPT 深度融入国际中文教育的应然功能、实践困境和应用策略[J]. *云南师范大学学报(哲学社会科学版)*, 2023, 55(4): 52-58.
- [23] 陈柯宇. ChatGPT-4 在中文写作中的辅助性定位[J]. *云南师范大学学报(对外汉语教学与研究版)*, 2023(3): 81-92.
- [24] 李建涛, 孔明, 钟英华. “表达驱动”教学理论在数智技术赋能国际中文教学中的设计探讨——以 Second Life 和 ChatGPT-4 为例[J]. *河南大学学报(社会科学版)*, 2023, 63(6): 103-109, 155.
- [25] 欧志刚, 刘玉屏, 覃可, 等. 人工智能多模态教学资源的生成与评价——基于 AIGC 在国际中文教育的应用[J]. *现代教育技术*, 2024, 34(9): 37-47.
- [26] 余江英, 杨晓珊, 郑开文. 人机共教: AIGC 赋能国际中文新手教师智慧备课的新生态[J]. *云南师范大学学报(对外汉语教学与研究版)*, 2024, 22(2): 49-59.
- [27] 范慧琴, 王璐璐, 曹芳, 等. 生成式人工智能赋能国际中文教师数字素养提升研究: 评估框架、现状及建议[J]. *天津师范大学学报(社会科学版)*, 2024(3): 72-83.
- [28] ZHANG J, ZHOU X, GOH Y S. Unpacking AI-supported Chinese as a foreign language learning: How beginner-level learners' cognitive and motivational factors predict speaking proficiency[J]. *Acta Psychologica*, 2025, 260: 105703.
- [29] HAN M, DU L, ZHANG Z, et al. Developing an AI literacy competency model for primary school Chinese language teachers: A pilot study[J]. *Frontiers in Education*, 2026, 11: 1831415.
- [30] 刘威童, 石佳佳, 李佩泽. 生成式人工智能赋能国际中文教育的基本框架与实践探索——写作智能体构建与教学实证研究[J]. *国际中文教育(中英文)*, 2025, 10(2): 102-117.
- [31] 李琦, 史金生. ChatGPT 赋能产出导向型国际中文口语课的实践路径与应用效果研究[J]. *海南师范大学学报(社会科学版)*, 2026, 39(3): 92-104.
- [32] 杨亦鸣, 王宇婷, 李顺发, 等. 人工智能赋能高级阶段汉语口语训练系统和教学模式研究[J]. *世界汉语教学*, 2026, 40(1): 3-21.
- [33] 马瑞凌, 徐娟. 生成式人工智能赋能国际中文教与学的智慧化——基于实证研究的思考[J]. *世界汉语教学*, 2026, 40(1): 129-142.
- [34] WU Y. Integrating Artificial Intelligence into Education: Opportunities, Challenges, and Response Strategies[J]. *AI Innovations and Applications*, 2025, 1(1): 115-124. DOI:10.63944/sj51.AIA.
- [35] WU Y. The Role of Generative Artificial Intelligence in Korean Language Learning: Applications, Challenges, and Instructional Strategies[J]. *Review of Resp Ai*, 2026, 1(2): 54-65. DOI:10.17613/g7mbr-1s626.
- [36] LI X, LI B, CHO S J. Empowering Chinese language learners from low-income families to improve their Chinese writing with ChatGPT's

- assistance afterschool[J]. Languages, 2023, 8(4): 238.
- [37] 穆肃,陈孝然,周德青. 生成式人工智能赋能教学设计分析:需求,方法和发展[J]. 开放教育研究, 2025, 31(1): 61-72.
- [38] 肖锐,宋靖雯,张邝弋. 大语言模型赋能国际中文教育数智化转向与发展探析[J]. 海南师范大学学报(社会科学版), 2025, 38(5): 84-91.

Application of Artificial Intelligence in Chinese Language Education from a Human-AI Collaboration Perspective: A Content Analysis (2020–2026)

Kong Weijian, Chen Ying, Wu Yundong
Dongshin University, Naju, Republic of Korea

Abstract: With the rapid development of generative artificial intelligence (AI) and large language models, AI is increasingly integrated into Chinese language education, including instructional design, resource generation, language practice, assessment, and teacher development. From a human–AI collaboration perspective, this study conducts a content analysis of AI-related research in Chinese language education from 2020 to July 2026. Literature was retrieved, screened, deduplicated, and verified through CNKI, Wanfang Data, VIP, and Google Scholar, resulting in 33 Chinese-language publications as the core sample. The studies were analyzed in terms of AI technologies, research themes, instructional content, and research methods. The findings show that 2023 marked a significant increase in research, with the focus shifting from technological applications and smart education toward generative AI-supported teaching and human–AI collaboration. Generative AI and large language models have become the most prominent technologies, mainly supporting teaching-material generation, writing, question answering, feedback, instructional design, and personalized learning. Research remains concentrated on macro-level theories, digital transformation, and development pathways, while studies on instructional systems, resources, writing, and speaking are increasing. Research on reading, Chinese characters, grammar, and culture remains relatively limited. Methodologically, theoretical and review studies continue to dominate, whereas longitudinal, multicenter, and controlled studies are insufficient. Accordingly, this study proposes a human–AI collaboration mechanism characterized by “teacher-directed, AI-assisted, learner-processed, and teacher-calibrated” instruction. Future research should strengthen Chinese-specific linguistic constraints, long-term learning evaluation, teacher AI literacy, dedicated datasets, assessment standards, and ethical governance.

Keywords: Chinese Language Education; International Chinese Language Education; Artificial Intelligence; Generative AI; Large Language Models; Research Trends