

语料库增强的 RAG 大模型在中医汉语教学资源生成中的应用研究

周万勤

朱宇*

厦门大学, 中国

摘要

随着大语言模型与语料库技术的发展, 专门用途汉语教学资源的动态生成逐渐成为可能。本文以中医汉语教学为例, 基于自建专业语料库并结合 RAG 技术, 构建了一套支持知识总结、问答设计、文化分析与教学计划自动化的资源生成框架。该框架以语料库为核心, 充分利用多项语料库来约束和优化生成过程, 体现了基于语料库的语言教学以真实语言证据支撑教学设计的原则。对比双盲打分结果显示, 语料库驱动生成在教学適切性方面显著优于无语料库驱动。教师访谈进一步表明, 该方法能够有效提升备课效率, 提升资源可用性与可信度, 成为强实用性的教学平台。未来, 可通过扩展专门用途汉语语料库, 并结合大语言模型与多模态方法, 生成更高质量的垂直领域教学资源。

关键词

语料库, 教学资源生成, 语言模型, 专门用途汉语

1 引言

语料库驱动的教学资源研究始终是二语习得与应用语言学领域的重要议题 (郑艳群, 2018; Lee et al., 2019; Sultana, 2024)。随着社会发展与现实需要, 专门用途汉语 (Chinese for Specific Purposes, CSP) 教学中的这一需求愈加突出。专门用途汉语教学一直有语言要求高、知识专业性强的特点 (李泉, 2011; 彭湃、胡晓研, 2017; 李艳华、李宇明, 2024), 使得教师在进行教学时难以切实贴合学生需求, 当前多数教学实践仍依赖教师基于有限经验或通用教材进行内容调整, 导致教学材料难以精准对接学习者的专业背景与实际语言需求。

近年来, 大语言模型 (Large Language Models, LLMs) 的兴起为上述困境提供了新的技术可能性 (欧志刚等, 2024)。凭借其在海量文本中习得的语言模式、知识理解以及任务执行的能力, LLMs 理论上可以靠近特定领域语料, 动态合成符合专业语域特征的教学内容 (王新, 2025; 范慧琴等, 2025)。例如, 针对中医、法律或工程等具体场景, 模型可自动生成术语解释、典型对话、案例分析乃至练习任务, 从而在一定程度上缓解人工编写的压力 (祝鸿平等, 2025)。

然而, 将通用大模型直接应用于中医汉语教学面临严峻挑战, 其核心难点在于中医语言的特性。首先, 中医汉语保留了大量古代汉语词汇与句法结构, 且存在显著的古今异义现象, 如“太

* 通讯作者。联系电邮: zhuyu@xmu.edu.cn

阳”指代经络而非天体，“辛苦”指代药味而非劳累，通用模型常因缺乏历时语言学视角而产生语义误读。其次，中医核心概念，如“气”“阴阳”“五行”具有高度的文化负载性，通用模型倾向于使用现代生物医学或通俗生活用语进行浅层解释（张雷平，2019）。因此通用大模型在处理此类高专业度内容时，极易产生看似流畅实则事实错误的幻觉（刘玉屏等，2025），如虚构药方配伍、混淆辨证逻辑等。这种表面流畅但内核失真的输出，对于强调严谨性的中医二语教学而言，是无效且危险的。

单纯依靠模型内部参数记忆难以满足需求，必须引入外部权威知识约束，而检索增强生成（Retrieval-Augmented Generation, RAG）技术为此提供了理想的解决方案，尤其是通过级联检索的方法。首先是事实锚定，通过强制模型依据检索到的权威语料进行生成，从源头上遏制了关于药方、穴位等事实性知识的幻觉；其次是语境还原，利用检索到的原文片段为模型提供准确的古今语义对照和文化语境；最后是难度适配，结合教学法标注的语料库，引导模型模仿专家教师的讲解逻辑，将晦涩的古文转化为符合学习者水平的教学资源。已有研究证实，通过将权威教材、等级标准及行业平行语料向量化以构建领域知识约束层，可使模型在输出前精准锚定专业语义空间（Lewis *et al.*, 2020），这不仅有效规避了垂直领域的幻觉现象，更确保了语言材料的来源可靠性与教学適切性（Smith, 2025）。

本研究聚焦中医汉语这一典型语域，提出融合自建专业语料库与 RAG 技术的解决方案，旨在构建一个可动态生成中医汉语教学资源的智能框架，并围绕三个核心目标展开：

- （1）构建一个可动态生成中医汉语教学资源的系统；
- （2）评估该系统生成的中医汉语教学资源内容质量；
- （3）通过一线教师的试用反馈，考察该系统在真实教学场景中的适用性与改进空间。

2 方法

2.1 数据与语料库建设

为支撑中医汉语教学资源的生成任务，本研究构建了一个多层次、多来源的专业语料库。

语料库包含以下三类材料：第一类为专门用途汉语教材，包括《中医汉语综合教程》《实用中医汉语口语》等针对留学生设计的教材；第二类为权威中医专业著作，如《中医基础理论》《中医内科学》《中药学》等，确保术语体系与概念表述的规范性；第三类为教学指导文件与文化拓展读物，例如《国际汉语教学通用课程大纲》《中医与中国传统文化》等，用以补充语用背景与跨文化解释维度，总计纳入 30 余本出版物，字符数超 1 千万。

表 1. 语料规模说明表

类别	出版物示例	数量(本)	总字符数	字数占比	功能说明
专门用途汉语教材	《中医汉语综合教程》	8	2,480,000	23.40%	教学核心
权威中医专业著作	《中医基础理论》	15	6,450,000	60.90%	专业基石
教学指导与文化	《国际汉语教学通用课程大纲》	10	1,656,000	15.70%	指导补充
合计		33	10,586,000	100%	

清洗后的语料经过去除目录、参考文献及图表标注等噪声处理，并对古籍引文进行了自动断句与标点标准化；最终构建为分级检索库：一级语料作为高优先级事实锚点，用于响应语言

教学与专业诊疗查询；二级语料仅在涉及哲学背景或跨文化对比时动态启用。全部文本经阿里云 text-embedding-v4 模型向量化后存入索引库，以支持毫秒级语义检索 (Li et al., 2019)。

2.2 框架设计与技术实现

在系统架构层使用的是语料库 + 大语言模型 + RAG 的三级工作框架。用户输入首先经预处理模块解析，利用 API 提取关键词，进行分类。随后触发级联检索机制 (Lewis et al., 2020)，先在一级语料中执行稠密检索，再根据前面的用户输入与稠密检索内容在二级语料中进行搜索，启用更宽泛的语义匹配，并结合教学適切性规则进行精选，最终保留合适的上下文片段作为 RAG 输入。

生成阶段采用结构化提示工程实现可控输出。系统内置 templates 模块，包含教学计划、知识总结、问答和文化分析四类预设模板，分别用于生成教案、梳理知识点、回答问题和开展文化阐释；每个模板以带占位符的字符串形式定义，包含主题和目标语言水平两个动态字段，并通过模板映射字典实现任务类型到具体模板的匹配；在调用提示生成函数时，系统根据任务类型、主题和学习者水平从映射字典中选取对应模板并完成填充。例如，针对“肝脏”的语音教学任务，系统会调用语音与词汇复合模板，插入“肝 gān”的标准拼音标注、常见误读提醒、以及包含该词的典型中医例句。教学计划对应的 templates 模块示例见附录。随后调用 DeepSeek 模型生成内容，并强制输出为结构化 Markdown 格式，便于教师直接使用。

2.3 平台展示

平台整体采用模块化架构设计。如图所示，系统主界面划分为三大功能区域：左侧为参数配置面板，中部为任务输入与资源生成区，右侧为对话历史记录区。在左侧配置区，用户可设定目标学习者的汉语水平、一级语料最小检索结果数以及最终注入模型的上下文数量，从而动态调控生成内容的难度梯度与信息密度。

在内容生成层面，平台输出严格遵循教学设计逻辑，采用结构化 Markdown 格式呈现，包含教学目标、重点难点分析、术语解释、典型例句、文化注释及教学建议等模块。

图 1. 中医汉语 RAG 教学资源生成系统概览



例如，在处理“肝脏”相关的中医核心概念时，系统不仅准确引用《中医汉语综合教程》中的规范定义，还结合《实用中医汉语口语》中的临床语境例句，并辅以针对 HSK 3 级学习者的简化解释及拼音注释。同时，提出学生学习的重点为“理解‘五脏’作为一个整体概念”，且一个重要的检验指标是说出“肝、脾、肺、肾、心五个脏器名称及其基本功能关键词”，锚定了学习目标，为教师具体教学作出指导。

图 2. 系统所生成的部分教学资源节选（教学目标）

一、教学目标

1. 知识目标

学习者需掌握与“肝”相关的中医核心概念、术语及其在五行系统中的位置和功能。根据《中医汉语综合教程》，“肝”是五脏之一，与脾、肺、肾、心共同构成中医脏腑系统的基本范畴（根据《中医汉语综合教程|优先级: Level_1_Core》）。具体知识点包括：

- “肝”的基本定义：gān n. liver（根据《实用中医汉语口语【基础篇】》），不仅是解剖器官，更是功能系统。
- “肝藏血”：肝具有贮藏血液和调节血量的功能（根据《中医汉语综合教程》表格内容：“肝—藏血”）。
- “肝火”（gānhuǒ）：指肝的病理状态，表现为易怒、头痛、目赤等（根据《实用中医汉语口语【基础篇】》）。
- “五行系统”中肝属“木”，与其他四脏存在“生克”关系，如“肝克脾”（根据《中医汉语综合教程》：“……克……”句式要求）。
- “归属”（guīshǔ v. belong to）、“对应”（duìyīng v. correspond to）、“系统”（xìtǒng n. system）等关键术语，用于描述肝在中医理论体系中的位置（根据《中医汉语综合教程》NEW WORDS列表）。

此外，学生需理解“五脏”作为一个整体概念（“五脏指的是什么？”——根据《中医汉语综合教程》第三课时问题1），并能准确说出肝、脾、肺、肾、心五个脏器名称及其基本功能关键词（藏血、运化、主气、藏精等）。

2. 技能目标

学习者应能：

- 使用“……生……，……克……”句式描述五脏之间的相生相克关系，例如“肝木克脾土”（根据《中医汉语综合教程》教学目标3）。
- 在口语和书面语中正确使用“肝”“肝火”“藏血”等专业词汇进行简单病情描述或健康咨询对话（参考《实用中医汉语口语【基础篇】》中“肝火”的实际应用场景）。

并且在“实战演练”部分还给出了具体的操作流程，涵盖多项真实可感的教学活动。如“角色扮演”游戏，符合中医场景，可以辅助学生以练代学，对设定为 HSK 3 级的学生来说具有很强的锻炼意义。

图 3. 系统所生成的部分教学资源节选（实操演练）

3. 实践操练（20分钟）

活动一：角色扮演——中医诊所

- 情景：患者说“我最近总发脾气，眼睛红”，学生扮演中医，用“你可能是肝火旺，因为肝属木，春季容易这样”回应（整合“肝火”“对应”“归属”）。
- 材料：症状卡片（易怒、失眠、口苦等），五行转盘。

活动二：五行关系排序游戏

- 分组比赛：给五脏卡片（肝、心、脾、肺、肾），让学生按“相克”顺序排列（肝→脾→肾→心→肺→肝），并用“克”造句（根据《中医汉语综合教程》“……克……”句式）。

活动三：对应连线

- 工作表：左边列“肝、心、脾…”，右边列“春、夏、长夏…”“青、红、黄…”，学生连线并说“肝对应春季和青色”（强化“对应”用法）。

由无语料库驱动的大语言模型生成的对应练习内容如下,相较于本系统生成的内容,显得较为笼统、缺乏具体性,难以直接应用于实际教学活动中。

2.4 评价设计

教学资源生成质量的评估采取双盲打分的形式。依据田旭升、黄小圆(2024)的分析,我们将教学任务拆解为语言、文化、跨文化三类,使用相同的提示词分别对有语料库驱动与无语料库驱动两类方法进行教学资源生成,语言水平统一设定为 HSK 4 级。提示词示例如下:

“请输出一份中医汉语教学资源,主题为 {“肝脏”的语法和文化教学}。”

图 4. 对应的无语料库驱动的大模型生成内容

七、课后作业

- 写一篇短文:《我的朋友最近总是发脾气,作为中医爱好者,我会给他什么建议?》(要求使用“因为...所以...”“容易...”“伤肝”等结构)
- 查找一种养肝的中药或食物(如枸杞、决明子),用中文介绍其功效。

生成共计 60 份教学资源后,本研究邀请 5 位具有一线汉语教学经验的教师进行百分制盲评。为确保评分的信度与客观性,本研究参考了 von der Gracht (2012) 关于专家共识测量 (Consensus Measurement) 的质量控制框架,采用离散程度指标 (Measures of Dispersion) 中的标准差 (SD) 作为评判依据。遵循该框架中的原则,结合百分制量表的跨度,本研究将标准差小于等于 8 设定为临界值。若同一份材料的评分 $SD \leq 8$, 则认为专家间已达成共识,反之则对该材料进行重评。

试用反馈的采集使用访谈。招募 10 名一线汉语教师参与为期 7 天的平台试用,试用前提供 5 分钟操作培训,结束后进行半结构化访谈。访谈提纲围绕两大主题:(1) 本系统是否减轻其在专业内容准备上的负担;(2) 与其他通用 AI 相比,本系统在有何优势或不足。对访谈录音进行逐字转录,并使用主题分析进行分析。分析遵循了 Braun 和 Clarke (2021) 的六步方法,其中包括熟悉数据、生成初始代码、识别主题、完善主题、定义和命名主题以及撰写最终分析,具体步骤如下。

主题分析过程由两名具有应用语言学背景的研究者协同完成。首先,两名研究者分别对前一部分的转录文本进行独立开放式编码,通过对比与讨论形成初步编码手册,明确各节点的定义与边界。随后,两名研究者基于该手册对全部文本进行背对背编码。编码完成后,采用 Holsti 信度系数对编码一致性进行检验,经计算,本研究的编码者间平均信度系数为 0.87,表明编码结果具有较高的一致性,主题编码框架如下。

表 2. 教师访谈主题编码框架及编码频次

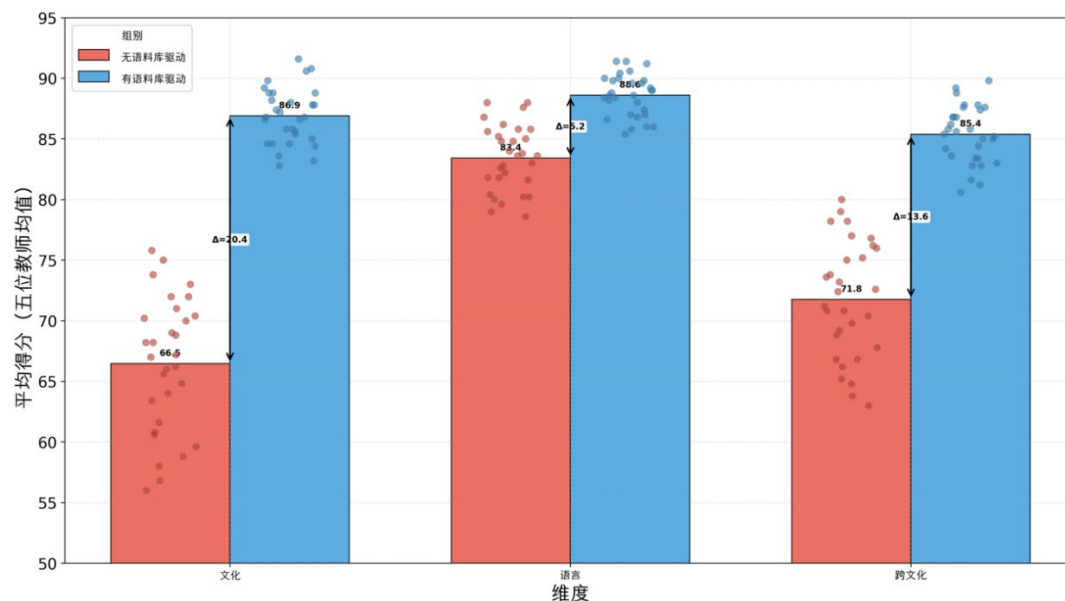
核心主题	子节点	提及人数	编码频次
教学效率与备课负荷的显著优化	跨学科知识整合耗时减少	10	21
	专业术语来源可信度高	9	15
生成资源的教学適切性与认知匹配度	难度层级符合学习者水平	7	11
	文化负载词解释具情境性	7	12
教学活动模式的创造性不足	活动类型较为单一	6	8
	细分领域语料覆盖不足	5	6
对专门用途汉语教学的战略意义	提供规范化教学脚手架	4	5
	助力教师角色向教学设计转型	7	10

3 结果

3.1 教学资源质量的评估

收集共计 300 份教学资源的评分数据后进行汇总，进行统计分析并绘制柱形散点图，如下所示。

图 5. 教师双盲评分结果散点 - 柱形图



进一步地，为验证显著性，对数据进行配对 T 检验，并发现各个维度在有无语料库驱动对比的评分结果都为非常显著。

表 3. 配对 T 检验结果表

维度	无语料库 ($M \pm SD$)	有语料库 ($M \pm SD$)	t	p	显著性
语言	83.41 ± 2.72	88.60 ± 1.71	-23.07	< .001	***
文化	66.46 ± 5.54	86.90 ± 2.30	-32.24	< .001	***
跨文化	71.75 ± 4.83	85.38 ± 2.39	-29.21	< .001	***

结果显示，语料库驱动条件下生成的教学资源在各维度均显著优于对照组 ($p < .001$)。其中，文化適切性维度差异最为突出 ($t = -32.24, p < .001$)：语料库组平均得分为 86.90 ($SD = 2.30$)，而无语料库组仅为 66.46 ($SD = 5.54$)。通过考察输出内容可见，无语料库驱动的生成资源常将中医特有的概念如“气”“经络”简化为日常语义或类比性解释，仅仅停留于概念的介绍；而有语料库生成的资源则能有效引用教材中的规范表述，在保持学术严谨性的同时兼顾学习者的认知负荷。例如，在关于“肝主疏泄”的教学片段中，语料库驱动输出不仅准确使用了“情志”等关联术语，还嵌入了典型临床语境下的例句，体现出较强的语域意识，如下。

图 6. 所嵌入的临床语境下例句与练习

情境对话 (医患问诊)

医生: 你最近心情怎么样?

患者: 经常生气, 心里不开心。

医生: 这可能是肝的问题。中医说, 肝主疏泄, 管人的心情。心情不好, 身体就不舒服。

患者: 那我应该怎么办?

医生: 吃点药, 让肝舒服一些。也要多休息, 保持好心情。

术语填空 (选词填空)

词语库: 肝 心情 气 舒服 生气

中医说, ____ 主疏泄, 能够调畅 ____, 调节 ____。

如果肝不好, 人容易 ____, 身体不 ____。

参考答案: 肝 / 气 / 心情 / 生气 / 舒服

在语言准确性方面 ($t = -23.07, p < .001$), 两组差距相对较小。值得注意的是, 无语料库组的语言点表述上具有较好的表现, 但在术语一致性与教学语言层级控制上存在波动。例如, 同一概念在不同段落中可能交替使用各种术语表达, 易造成初学者混淆; 而语料库组通过检索一级语料中的术语定义与搭配范式, 有效规避了此类问题, 并能根据任务类型动态调整语言复杂度。

跨文化解释力维度亦呈现显著分化 ($t = -29.21, p < .001$)。无语料库驱动的教学资源多依赖通用文化对比策略, 如中医和西医对比, 解释流于表面; 而语料库驱动的教学资源能结合《国际汉语教学通用课程大纲》及文化拓展读物中的教学指引, 提供更具情境敏感性的中介话语。例如, 在解释“上火”这一文化负载概念时, 系统不仅说明其在中医理论中的病理机制, 还补充了留学生常见的误解及教学应对建议, 体现出对学习者的前概念的关照。

3.2 教师试用访谈的反馈

通过对转录文本进行细致的主题分析, 我们提炼出四个核心反馈主题, 涵盖了该生成框架在实际教学场景中的优势、不足及其在专门用途汉语教学中的潜力。

3.2.1 教学效率与备课负荷的显著优化

所有受访教师均对该系统在专门用途汉语教学备课流程中的效率提升与认知负荷减轻作用给予了高度肯定。教师们一致认为, 该框架有效解决了中医汉语教学中跨学科知识转换的痛点。

在传统模式下, 非中医专业背景的汉语教师需要投入大量时间成本进行知识整合与术语规范化。教师 T4 提到: “以前, 处理一个核心概念, 可能同时对照中医学教材和对外汉语教学法资料, 让语言难度符合对应的 HSK 水平, 这个过程非常花时间。”

教师对系统基于 RAG 机制提供的教学素材的专业语域来源可靠性表示认可。平台通过对一级语料库的密集检索，能够提取出真正契合中医语境的典型例句与搭配范式，而非通用大模型倾向生成的泛化语料。教师 T7 指出：“系统提供的关于‘气虚’的例句，内容直接来源于临床或教材并且有标注，也方便我们回头检查，更具教学价值。”这印证 RAG 技术在保证专业术语使用的一致性和准确性方面的优势。

3.2.2 生成资源的教学适切性与认知匹配度

教师们普遍认可本框架生成的教学资源在教学适切性方面优于通用大模型。这种优势主要体现在难度梯度控制和跨文化解释策略的有效应用。

教师们普遍认为想要精准调教通用的大型语言模型难度极高，且需要投入巨大的精力。而本系统能够依据预设的教学规则和学习者语言水平，在内容输出中实现术语难度的动态调整。例如，教师 T5 指出：“在解释‘阴阳’时，系统能有意识地避免引用那些很难懂的古典哲学文本，又能在中医学的功能性描述和对比中有所侧重，这点我觉得很好。”

系统在生成文化解释内容时展现出对学习者的自身基础知识的关照。通过检索二级语料库中的《国际汉语教学通用课程大纲》等指导文件，系统能够生成更具适用性的教学资源。例如，在解释“上火”等文化负载词汇时，系统不仅提供中医病理机制的阐释，还补充了针对西方学习者常见误解的教学应对建议，实现了从知识介绍到跨文化理解引导的深化。

3.2.3 教学活动模式的创造性不足

尽管系统在内容生成方面表现突出，但在教学法多样性和特定班级个性化需求的响应上仍存在局限，这是未来优化的主要方向。

课堂活动多样化缺乏：当前框架生成的资源形式以结构化文本为主。教师们期待系统能进一步支持复杂、有创造性教学活动的设计，以提升课堂互动性。教师 T2 建议：“如果系统可以直接生成基于临床案例的情景模拟对话脚本、角色扮演任务或辩论任务大纲，并且给我们更多相关的创造性任务的启发会更好。”

框架目前的核心语料资源集中于中医基础理论。教师们在尝试生成跨专业应用场景，如中医美容、针灸、现代药理学等 RAG 语料库中没有的内容时，效果就会有所下降。此外，平台尚未提供教师对语料进行个性化权重调整的接口，难以完全满足特定教学班级的差异化需求。

3.2.4 对专门用途汉语教学的战略意义

受访教师普遍认为本研究框架具有重要的战略和示范意义，它不仅是一个工具，更是专门用途汉语教学资源规范化与智能化的建设性工作。

教师们将该系统视为一个专业知识的筛选与组织工具，为非专业背景的汉语教师提供了一个权威且规范的教学脚手架，使得教学内容不再依赖于教师的个人经验或有限的通用教材，而是基于高质量专业语料库的系统化输出。

教师 T10 提到，通过工具将知识查找和术语规范化等耗时任务自动化，教师可以把更多精力聚焦于教学策略、课堂互动等其他更高阶的多教学环节，促进教师的精力集中，也有助于这类专门用途汉语教学在实践过程中的策略转变。

4 讨论

4.1 理论与实践意义

本研究在专门用途汉语教学领域中,对大语言模型的应用进行了垂直性探索,并在此基础上深化了 RAG 方法在汉语教学场景下的运用。传统上,知识架构理论强调知识的组织方式、层级结构及其在特定任务中的可调用性 (Nonaka & Takeuchi, 1995)。然而,在生成式人工智能迅速发展的背景下,该理论亟需回应“如何将外部权威知识有效嵌入模型生成过程”这一新命题 (陈嘉靖, 2025; 万力勇等, 2025; 祝鸿平等, 2025)。本研究通过引入 RAG 构建了一个双层语料库驱动的知识耦合框架,不仅验证了外部知识注入对模型输出可控性的提升作用,也拓展了知识架构理论在人机协同教学场景中的适用边界。

实验结果明确显示,仅依赖通用大语言模型的通识知识难以满足专门用途教学对文化适切性、术语规范性和学科准确性的高要求。这与 (Kim & Vajravelu, 2025) 在医学英语教学中的发现高度一致:当模型缺乏领域约束时,其生成内容虽语法流畅,却常出现事实偏差或文化误读。相比之下,本研究提出的 RAG 系统通过一级语料库确保专业术语的一致性,通过二级语料库 (辅助性教学文件、跨文化注释、案例脚本等) 提供情境化解释策略,实现了多维度的动态结合。本系统采用级联检索策略,即先检索核心概念再基于上下文需求调用解释性资源的方法,在提升生成质量的同时提高了系统响应速率,可推广至其他现有 RAG 应用于教育的研究中。

在实践层面,本研究直接回应了专门用途汉语教学长期面临的问题,即优质教学资源稀缺、跨学科知识整合难度高、非专业背景教师教学难的问题。访谈数据一致表明,该框架显著降低了教师在术语核查、文化解释和内容适配上的成本,这也是 RAG 系统独有的特点。教师们指出,这种来源可追溯、内容可信赖的设计,不仅提升了教学内容的权威性,也增强了教师对 AI 工具的信任感。这点对提升人机合作效率、增强教育教学适切性尤为重要 (王新, 2025; 祝鸿平等, 2025)。

更重要的是,该系统促使教师角色从内容生产者向教学设计师转型。当原本查询资料、匹配语言等基础性、重复性的知识整理工作由 AI 承担后,教师得以聚焦于高阶教学活动,如差异化指导、互动任务设计与学习者情感支持。这种专业化转型趋势,与刘邦奇等 (2025) 和王争录、张博 (2023) 所倡导的 AI 赋能而非替代教师理念高度契合。

4.2 技术优势与局限

本系统的优势集中体现在三个方面,即内容可控性、教学适切性保障与认知匹配能力。首先, RAG 机制通过外部语料库对 LLM 生成过程施加软约束,有效抑制了通用模型在专业领域常见的幻觉问题。尽管在本次实验中尚未发现严重的幻觉现象,但 Lee *et al.* (2024) 说,在专业领域中,大语言模型很可能不知道自己表达的逻辑漏洞。如本系统虽能检索到《中医基础理论》中“肝开窍于目”的规范表述,但在生成面向 HSK3 级学习者的简化解释时,输出为:“肝的疏泄功能正常,眼睛才能明亮。如果肝火旺盛,就会出现眼睛发红、肿痛。所以同学们要少熬夜,保护肝脏才能保护眼睛。”这一内容将“肝开窍于目”简化为肝火旺就导致眼睛红的线性因果,忽视了肝血不足会导致“目失所养”这一关键原因,在具体的教学适应过程中还存在疏漏。

其次,通过结构化提示工程系统实现了对教学难度的动态调控。例如,当输入学习者水平为初级时,模型对五行概念的解释侧重类比;而对高级学习者,则引入生克关系与临床应用。

这种分层解释策略在教师访谈中获得高度认可，反映了一线教师的真实需求。更重要的是模型不再一刀切地进行输出，而是依据预设或实时反馈的学习者特征调整汉语教学资源的输出难度（丁安琪、蒙小凤，2025）。并且，二级语料库的引入显著丰富了跨文化解释的维度。不同于早期 AI 教学系统仅提供字面翻译或简单对比，本系统能调用包含历史语境、哲学基础、常见误解及教学建议的多维注释，使跨文化教学超越简单中西对比的二元对立，走向情境化、反思性理解。

然而本研究也存在一些局限。语料库覆盖广度不足。尽管核心语料库涵盖中医基础理论，但在教师尝试生成针灸操作、中药配伍或中医美容等细分领域内容时，系统因缺乏相应语料而表现下降。这印证了 Li *et al.* (2025) 的调查：其一，领域专用 AI 系统的效能高度依赖于底层知识库的完备性与更新机制。其二，创造性的教学活动生成能力有限。当前系统主要输出说明性文本，难以生成高互动性任务，这类更高阶的思维方式或许是 RAG 弥补不了的（Guțu & Popescu, 2024），后续的微调、后训练、强化学习等阶段可能可以改善。其三，个性化调节接口缺失。系统目前采用统一参数设置，无法根据文化背景或教学风格进行调整，这类任务需要更多的参考资源，也限制了其在真实多元课堂中的适应性。

上述局限指向了本系统未来的三个优化方向。首先，语料库建设应向多模态、动态演化方向发展（欧志刚等，2024），整合资源，构建一体化知识图谱以回应教师对细分领域内容生成的诉求（赵凤娇、王世友，2025）。其次，应发展教师与 AI 协同的教学活动生成范式，通过构建教学活动模板库，将 RAG 检索到的知识点自动注入模板节点，并开发可视化干预界面允许教师拖拽调整语料权重，实现人机共创（王争录、张博，2023；王新，2025）。最后，需探索基于学习者模型的自适应生成机制，通过追踪学生的练习错误、提问模式等行为数据（马志强等，2025），动态推断其认知状态与文化背景，使系统能够针对“阴虚”“阳虚”混淆等具体问题精准干预，推动专门用途汉语教学从标准化供给迈向精准化服务。

5 结语

本研究通过构建语料库驱动的检索增强生成框架，验证了语料库与大语言模型融合在中医汉语教学资源生成中的有效性。结果显示，语料库组在文化適切性、跨文化解释力和语言准确性三个维度上均显著优于无语料库组，表明语料库作为知识锚点能够有效克服通用大语言模型在专业领域中的知识幻觉与语域偏差问题。从理论层面看，本研究拓展了知识架构理论在人机协同教学场景中的应用，当外部知识通过 RAG 机制被系统性地嵌入模型生成过程，知识成为动态可调用的智能资源，回应了专门用途语言教学对知识组织与提取的要求；实践层面，本系统减轻了教师在专业内容准备上的认知负荷，通过语料库的权威性保障使教学内容从个人经验走向规范化参照。本研究亦存在局限，如语料库的广度与深度限制了系统在细分领域的表现，创造性教学活动的生成能力仍有待提升，个性化调节机制的缺失制约了其在多元课堂中的适应性。未来研究可聚焦于构建多模态动态演化的语料库体系、发展教师与 AI 协同的教学活动设计范式，以及探索基于学习者模型的自适应生成机制。

附录

TEACHING_PLAN_TEMPLATE = """"你是一位经验丰富的中医汉语教学专家，可以充分地指导学生，依据他们的语言水平进行学习。请根据提供的上下文知识，为{topic}主题生成一份非常详细和全面的教学计划。

重要要求：

1. 必须引用上下文：在生成内容时，必须充分引用和整合上下文知识中的具体信息、概念、术语和案例。对于每个要点，都要尽可能引用上下文中的相关内容。
2. 标注引用来源：当引用上下文内容时，请标注来源（如：根据《来源文档名》...），确保信息的可追溯性。
3. 内容详细程度：每个部分都要详细展开，不要只写标题或简单概述，要提供具体、深入的内容。

详细内容要求：

1. 教学目标：

知识目标：详细列出学习者需要掌握的具体知识点

技能目标：明确说明学习者需要培养的具体能力

情感目标：阐述学习者应建立的态度和价值观

2. 教学重点：

详细列出核心概念和关键术语

对每个重点概念进行详细解释

说明这些重点之间的逻辑关系

3. 教学难点：

识别学习者可能遇到的具体困难

对每个难点提供详细的分析和解决方案

引用上下文中的相关解释和示例

4. 教学步骤：

导入环节：设计具体的导入活动，引用上下文中的相关背景知识

新知呈现：详细讲解核心内容，引用上下文中的概念、定义、原理和案例

实践操练：设计具体的互动活动，结合上下文中的实际应用场景

总结巩固：系统梳理知识点，整合上下文中的关键信息

5. 教学活动：

设计 5-8 个具体的课堂活动

每个活动都要引用上下文中的相关内容作为素材

提供活动的具体步骤、所需材料和预期效果

6. 评估方式：

详细说明形成性评估的具体方法和标准

提供总结性评估的详细方案

给出具体的评估示例和评分标准

7. 拓展资源：

推荐相关学习材料

提供深入学习的路径和方向

列出相关的参考文献和资源链接

目标学习者水平： {target_level}"""

参考文献

- Braun, V., & Clarke, V. (2021). Doing coding. In *Thematic analysis: A practical guide* (pp. 103–128). SAGE.
- Chen, Jiajing (陈嘉靖). (2025). 人工智能技术背景下的职业教育学术资源库建设路径探析 [Exploration of the construction path of vocational education academic resource library under the background of artificial intelligence technology]. *中国职业技术教育* [Chinese Vocational and Technical Education], (20), 96–103.

- Ding, Anqi (丁安琪), & Meng, Xiaofeng (蒙小凤). (2025). 从模仿到重构: AI生成教案对国际中文教育新手教师教学设计能力的影响 [From imitation to reconstruction: The impact of AI-generated lesson plans on the instructional design ability of novice teachers in international Chinese education]. *贵州师范大学学报(社会科学版)* [Journal of Guizhou Normal University (Social Sciences)], (5), 42–52.
- Fan, Huiqin (范慧琴), Han, Peiling (韩沛玲), Cao, Fang (曹芳), et al. (2025). 基于人工智能生成内容的国际中文读写一体化教学模式构建与应用 [Construction and application of an integrated reading and writing teaching model for international Chinese based on AIGC]. *语言教学与研究* [Language Teaching and Linguistic Studies], (5), 1–13.
- Guțu, B. M., & Popescu, N. (2024). Exploring data analysis methods in generative models: from Fine-Tuning to RAG implementation. *Computers*, 13(12), 327. <https://doi.org/10.3390/computers13120327>
- Kim, J., & Vajravelu, B. N. (2025). Assessing the current limitations of large language models in advancing health care education. *JMIR Formative Research*, 9(1), e51319. <https://doi.org/10.2196/51319>
- Lee, H., Joo, S. J., Kim, C., Jang, J., Kim, D., On, K. W., & Seo, M. (2024, June). How well do large language models truly ground?. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)* (pp. 2437-2465).
- Lee, H., Warschauer, M., & Lee, J. H. (2019). The effects of corpus use on second language vocabulary learning: A multilevel meta-analysis. *Applied Linguistics*, 40(5), 721-753.
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, 33, 9459–9474.
- Li, Quan (李泉). (2011). 论专门用途汉语教学 [On teaching Chinese for specific purposes]. *语言文字应用* [Applied Linguistics], (3), 110–117.
- Li, W., Zhang, Y., Sun, Y., Wang, W., Li, M., Zhang, W., & Lin, X. (2019). Approximate nearest neighbor search on high dimensional data—experiments, analyses, and improvement. *IEEE Transactions on Knowledge and Data Engineering*, 32(8), 1475-1488. <https://doi.org/10.1109/TKDE.2019.2909204>
- Li, Yanhua (李艳华), & Li, Yuming (李宇明). (2024). 新世纪专门用途中文的发展动力、现状与趋势 [Development dynamics, status quo, and trends of Chinese for specific purposes in the new century]. *首都师范大学学报(社会科学版)* [Journal of Capital Normal University (Social Sciences Edition)], (2), 87–96.
- Li, Z., Wang, Z., Wang, W., Hung, K., Xie, H., & Wang, F. L. (2025). Retrieval-augmented generation for educational application: A systematic survey. *Computers and Education: Artificial Intelligence*, 8, 100417.
- Liu, Bangqi (刘邦奇), Nie, Xiaolin (聂小林), Wang, Yafei (王亚飞), et al. (2025). 生成式AI赋能教育: 技术框架、应用场域及价值——2024智能教育发展研究报告 [Generative AI empowering education: Technical framework, application scenarios, and value—2024 intelligent education development research report]. *中国电化教育* [China Educational Technology], (3), 61–70.
- Liu, Yuping (刘玉屏), Ou, Zhigang (欧志刚), & Wu, Xiaoqin (武晓琴). (2025). 生成式人工智能赋能国际中文教学的效果测评——以教学设计、HSK模拟试题编写及作文评分为例 [Effectiveness evaluation of generative AI empowering international Chinese teaching: Taking instructional design, HSK mock test compilation, and essay scoring as examples]. *民族教育研究* [Journal of Research on Education for Ethnic Minorities], 36(1), 156–166.

- Ma, Zhiqiang (马志强), Zhang, Bing'er (张冰儿), & Ning, Jing. (2025). AI 驱动的智能适应学习系统设计何以有效? —— 基于对近十五年来国际前沿研究成果的元分析 [Why is AI-driven intelligent adaptive learning system design effective? —A meta-analysis based on international frontier research results in the past fifteen years]. 现代教育技术 [Modern Educational Technology], 35(10), 99–107.
- Nonaka, I., & Takeuchi, H. (1995). Theory of organizational knowledge creation. In *The knowledge-creating company: How Japanese companies create the dynamics of innovation* (pp. 56–94). Oxford University Press.
- Ou, Zhigang (欧志刚), Liu, Yuping (刘玉屏), Qin, Ke (覃可), et al. (2024). 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 [Generation and evaluation of AI multimodal teaching resources: Based on the application of AIGC in international Chinese education]. 现代教育技术 [Modern Educational Technology], 34(9), 37–47.
- Peng, Pai (彭湃), & Hu, Xiaoyan (胡晓研). (2017). 专门用途汉语教师的培养模式及自我发展 [Training models and self-development of teachers of Chinese for specific purposes]. 东北师大学报 (哲学社会科学版) [Journal of Northeast Normal University (Philosophy and Social Sciences)], (3), 109–114.
- Smith, S. (2025). Enhancing English language education with retrieval-augmented generation (RAG): A new paradigm for AI-powered teaching and learning. *International Journal of Education Humanities and Social Science*, 8(4), 235–252.
- Sultana, S. (2024). Corpus-Driven Approaches for Advancing L2 Grammar: An Empirical Investigation. *Journal of Language and Culture in Education*, 1(2), 50–58. <https://lce.dsigma.gr/index.php/lce>
- Tian, Xusheng (田旭升), & Huang, Xiaoyuan (黄小圆). (2024). 基于传统哲学范式的中医汉语课程跨文化教学思考 [Reflections on cross-cultural teaching of TCM Chinese courses based on traditional philosophical paradigms]. 医学与哲学 [Medicine & Philosophy], 45(19), 73–76, 81.
- Von Der Gracht, H. A. (2012). Consensus measurement in Delphi studies: review and implications for future quality assurance. *Technological forecasting and social change*, 79(8), 1525–1536.
- Wan, Liyong (万力勇), Zhang, Xinyi (张心怡), & Wang, Yue (王月). (2025). 生成式人工智能背景下人机共创数字教育资源质量控制: 框架与策略 [Quality control of human-machine co-created digital educational resources in the context of generative AI: Framework and strategies]. 远程教育杂志 [Journal of Distance Education], 43(3), 54–63.
- Wang, Xin (王新). (2025). DeepSeek 等生成式人工智能赋能国际中文数字教材研编 [DeepSeek and other generative AIs empowering the research and compilation of international Chinese digital textbooks]. 青海民族大学学报 (社会科学版) [Journal of Qinghai Minzu University (Social Sciences)], 1–8.
- Wang, Zhenglu (王争录), & Zhang, Bo (张博). (2023). 从同存走向共生: AI 赋能教师教育的实践理性 [From coexistence to symbiosis: Practical rationality of AI-empowered teacher education]. 高教发展与评估 [Higher Education Development and Evaluation], 39(3), 21–29, 108, 120.
- Zhang, Leiping (张雷平). (2019). 中医语言研究的百年回顾与思考 [A centennial review and reflection on TCM language research]. 江淮论坛 [Jianghuai Forum], (5), 166–172.
- Zhao, Fengjiao (赵凤娇), & Wang, Shiyu (王世友). (2025). 基于知识图谱的国际中文数智教材概念模型构建 [Construction of a conceptual model for international Chinese digital-intelligent textbooks based on knowledge graphs]. 语言教学与研究 [Language Teaching and Linguistic Studies], (4), 13–23.
- Zheng, Yanqun (郑艳群). (2018). 汉语教学资源研究的新进展与新认识 [New progress and new understanding in the research of Chinese teaching resources]. 语言文字应用 [Applied Linguistics], (3), 106–113.

Zhu, Hongping (祝鸿平), Zhu, Ruizhe (祝睿喆), Ye, Hongwu (叶宏武), et al. (2025). 职业教育数字资源开发困境与突破：人工智能技术优势下的模式重构 [Dilemmas and breakthroughs in the development of digital resources in vocational education: Model reconstruction under the advantage of artificial intelligence technology]. 远程教育杂志 [Journal of Distance Education], 43(3), 64–72.

投稿：2025年12月14日；接受：2026年3月8日；出版：2026年4月9日

作者简介

周万勤，厦门大学国际中文教育学院 / 海外教育学院语言学及应用语言学硕士研究生，研究方向为信息技术辅助的汉语作为第二语言习得、计算语言学。

朱宇，美国普渡大学外语教育博士（Ph.D.），厦门大学国际中文教育学院 / 海外教育学院华文系系主任，厦门大学教授、博士生导师、博士后合作导师。福建省语言学会常务理事，国家社科基金通讯评审专家及成果鉴定专家，*Chinese for Specific Purposes* 期刊执行主编、《海外华文教育》编委。主持国家社科基金、教育部规划基金、国家语委等课题多项，已出版学术专著两部，发表国内外核心学术期刊论文数十篇。研究兴趣为信息技术辅助的行业汉语教育与测评、基于语料库的语域语言差异分析。

Application Research on Corpus-Enhanced RAG Large Language Models in the Generation of Teaching Resources for Chinese as a Foreign Language in Traditional Chinese Medicine

Wanqin Zhou

Yu Zhu

Xiamen University, China

Abstract

With the advancement of Large Language Models (LLMs) and corpus technologies, the dynamic generation of teaching resources for Languages for Specific Purposes (LSP) has gradually become feasible. Taking the teaching of Chinese for Traditional Chinese Medicine (TCM) as a case study, this paper constructs a resource generation framework supporting knowledge summarization, question-and-answer design, cultural analysis, and automated instructional planning. This framework is built upon a self-constructed specialized corpus and integrates Retrieval-Augmented Generation (RAG) technology. Centered on the corpus, the framework leverages multiple corpora to constrain and optimize the generation process, embodying the principle of corpus-based language pedagogy where instructional design is underpinned by authentic linguistic evidence. Results from a double-blind scoring comparison indicate that corpus-driven generation significantly outperforms non-corpus-driven methods in terms of pedagogical appropriateness. Furthermore, teacher interviews reveal that this approach effectively enhances lesson preparation efficiency, improves the usability and credibility of resources, and serves as a highly practical teaching platform. Future research may focus on expanding LSP Chinese corpora and integrating LLMs with multimodal methods to generate higher-quality teaching resources for vertical domains.

Keywords

Corpus, teaching resource generation, language models, Chinese for specific purposes

Wanqin Zhou is a Master's candidate in Linguistics and Applied Linguistics at the School of International Chinese / Overseas Education College, Xiamen University. Her research interests include Technology-Assisted Chinese as a Second Language Acquisition and Computational Linguistics.

Yu Zhu holds a Ph.D. in Foreign Language Education from Purdue University, USA. He currently serves as the Chair of the Department of Chinese Language and Culture, Professor, Doctoral Supervisor, and Postdoctoral Collaborative Supervisor at the School of International Chinese / Overseas Education College, Xiamen University. Dr. Zhu is an Executive Director of the Fujian Provincial Society of Linguistics, a Communication Review Expert and Achievement Appraisal Expert for the National Social Science Fund of China, the Executive Editor-in-Chief of the journal *Chinese for Specific Purposes*, and

an Editorial Board Member of *Overseas Chinese Education*. He has presided over multiple research projects funded by the National Social Science Fund of China, the Ministry of Education Planning Fund, and the State Language Commission. He has authored two academic monographs and published dozens of papers in core academic journals both domestically and internationally. His research interests focus on Technology-Assisted Industry-Specific Chinese Education and Assessment, as well as Corpus-Based Analysis of Register Variation.