

大语言模型赋能多模态交互：数字课堂中汉语二语学习者口语流利度提升路径研究

赵思昂

悉尼科技大学 澳大利亚

【摘要】数字化课堂环境下，汉语二语口语流利度提升面临着教学互动受限和技术应用不足两个核心问题。在教学层次上，互动方式单一，缺少真实情境，缺乏反馈；在技术层面上，语音识别的准确性、多模态的融合程度不高，以及大尺度的语言模型应用不完全。针对以上问题，本文提出以大语种模型为核心驱动力，从多模态沉浸训练场景构建、个性化实时反馈系统优化、分层高级训练框架构建、学习数据监控与优化机制四个方面优化路径，旨在系统提升汉语二语教学水平，为学习者口语流利程度的提升提供一体化解决方案。

【关键词】大语言模型；多模态交互；数字课堂；汉语二语学习者；口语流利度

【收稿日期】2025年11月15日 **【出刊日期】**2025年12月16日 **【DOI】**10.12208/j.sdr.20250283

Large language models empower multimodal interaction: a study on enhancing oral fluency in Chinese as a second language learners in digital classrooms

Siang Zhao

University of Technology Sydney, Australia

【Abstract】 In the digital classroom environment, the improvement of fluency in spoken Chinese as a second language faces two core issues: limited teaching interaction and insufficient technological application. At the teaching level, the interactive mode is single, lacking real situations and feedback; On a technical level, the accuracy of speech recognition, the degree of multimodal fusion, and the incomplete application of large-scale language models are not high. In response to the above issues, this article proposes to use a large language model as the core driving force to optimize the path from four aspects: construction of multimodal immersive training scenarios, optimization of personalized real-time feedback systems, construction of hierarchical advanced training frameworks, and monitoring and optimization mechanisms for learning data. The aim is to systematically improve the level of Chinese second language teaching and provide an integrated solution for enhancing learners' oral fluency.

【Keywords】 Big language model; Multimodal interaction; Digital classroom; Chinese second language learners; Oral fluency

汉语二语口语流利度是衡量学习者语言运用能力的重要指标，数字化课堂为这一能力提供了一个全新的载体，然而现有教学模式的局限性制约着其教学效果^[1]。本文将大数据模型和多模态交互技术相结合，不仅能完善智能教育和二语习得的交叉研究体系，在实践上也可以为数字化课堂下汉语二语口语教学提供切实可行的优化路径，帮助提高教学质量及学生的语言能力。

1 数字课堂中汉语二语口语流利度提升的现存

问题

1.1 教学交互层面的局限

在数字化教学环境下，汉语二语口语教学在互动方式上存在着明显的局限性，难以满足提高流利性的核心需求。已有的交互多为单一语音或文字，缺少多模态要素（如图像、视频等）的深度融合，不能构建真实的言语交流情景，致使学生口语训练陷入机械回答的僵化模式^[2]。同时，受限于课堂规模和时间限制，师生及生生之间的互动不够频繁，个性

化实践机会缺乏，互动方式多为单向输出，缺乏即时双向反馈，难以形成有效的语言交互闭环，更不能模拟真实交流中的话题转折，制约着学习者语用适配和适应性的提升。

1.2 技术应用层面的短板

支撑汉语二语教学的数字化技术体系存在着明显的缺陷，削弱了技术对外语教学的促进作用。已有的语音识别系统在语音识别中存在发音不规范和复杂句式识别不准确的问题，且语义理解模块难以对口语语用逻辑进行深度解析，导致反馈和训练缺乏针对性^[3]。目前，多模态技术融合度不高，语音、图像和文字数据分离，难以构建具有沉浸感的交互场景。另外，大规模语言模型的应用呈现碎片化的特点，没有融入到整个教学过程中，只是作为一个单独的工具来辅助局部的连接，没有形成系统的训练系统，其技术价值没有得到很好地发挥^[4]。

2 大语言模型赋能多模态交互下数字课堂中汉语二语学习者口语流利度提升路径

2.1 构建多模态沉浸式口语练习场景

沉浸式二语习得理论和交互假说指出，真实、可互动的语境是提高口语流利性的核心前提。基于大规模跨模态语料，可以实现文本、语音、图像等多种模态的协同生成，根据学习者的语言水平，动态调整语境难度，并与之相匹配，减少认知负荷，激发积极表达的积极性，为提高口语流利性打下坚实的话语基础^[5]。

在具体实施中，可以大语言模型为基础，构建动态的语境生成系统，并根据不同的 HSK 等级和学习目标，分别建立日常交流、学术研讨、商务洽谈等场景下的语境模板库，并通过模型调用语音合成和图像生成接口，填充动态变化的多模态要素，如针对初级学习者，可形成菜市场问路、校园借物等场景，面向中高级学习者，可生成跨界商务谈判、学术会议演讲等领域的行业素材和术语提示。同时，研发基于大数据模型的虚拟汉语交流助手，能够模拟不同身份和地域的交流对象，支持多模态的语音、文字和表情交互，能够根据学习者的口语表达情况实时调整回复逻辑，增加现实交流中的突发分支，如话题转折、观点反驳等，并定期将节日习俗、社会热点等时效性内容融入其中，使学习者能够在多模态互动中提高口语应变能力和连贯表达能力。

2.2 优化实时化个性化口语反馈体系

在二语习得中，反馈假说强调及时、准确、个性化的反馈对于纠正语言错误和提高语言流利性具有重要意义。该模型具有多维语言分析和多模态输出的能力，可弥补传统课堂反馈滞后，维度单一，同质化严重的弊端。该系统能够同时对口语的速度、停顿频率、句法复杂性、语音语调、语用适配性等核心指标进行量化评价，并根据学习者的个人错误类型生成差异化的反馈，以多模态形式实现内容的传递，满足不同的认知偏好，从而达到反馈的有效性和准确性^[6]。

在具体实施中，可构建大语言模型多维度口语测评反馈系统，首先采用高精度语音识别技术采集学习者口语语音并将其转换成文本，从流畅度维度对其进行分解分析，再融合多模态反馈形式，针对语音语调错误，生成标准发音和嘴型动态图片对照；针对句法逻辑问题，给出可视化的句法结构图，标注错误节点和优化方案；针对语用不合适的情况，对场景化的语用案例和文化背景进行分析。在此基础上，构建学习者口语错误数据库，根据历史错误记录定制个性化的强化反馈，例如针对频繁使用介词的学习者，增加多模态介词用法辨析练习，为停顿过多的学习者提供节奏控制训练素材，并将反馈内容与后续课堂训练任务进行深度关联，形成“评价-反馈-强化训练”闭环，保证反馈真正转化为口语流利的动力。

2.3 搭建分层式流利度进阶训练框架

二语习得的渐进原则和差异化教学理论揭示了不同水平学习者在口语流利度短板上存在着明显的差异，需要有一个梯度的培训系统来匹配^[7]。该方法能够根据预测试数据判断学习者的熟练程度，并根据每个阶段的核心训练目标，生成梯度的多模式训练内容，在确保训练针对性的前提下，实现从基础向高层次流畅过渡的平滑过渡，避免训练难度不平衡而导致的学习失败和能力停滞不前。

在具体实施中，需要构建三个层次的流畅进阶训练框架，第一阶段通过大语种模型产生多模态跟随阅读材料，同步提供汉字的拼音、声调标注、口型动态图，实时对比学习者的发音数据，纠正发音偏差和咬字不清的问题；并通过日常语音图像的快速回答训练，强化基本表达流畅性。中级阶段模型向用户推送多模态主题素材，如短视频片段或相关图片，要求学习者完成内容描述、观点表达或情节

串联等任务，实时监测篇章逻辑连贯性；针对句子衔接生硬的情况，推荐过渡词的多模态使用示例，并设计多人协作的多模态话题讨论场景，指导学习者互动回答，提高会话连贯性。高阶阶段为学术报告、跨境商务谈判等特定场景，需要学习者完成多模态口语任务，强化专业术语和正式语体的运用，并对跨文化语用规范进行补充，以提高高阶口语流利度。

2.4 完善多模态学习数据监测与优化机制

学习分析理论和闭环教学理论都强调对学习数据进行实时监控和动态优化是保证学习效果的关键。对大数据进行多模态数据处理和分析，可突破传统数字化课堂数据收集单一、分析肤浅的局限，融合语音、文本和交互行为等多模态数据，构建完整的学习者学习画像，深入挖掘其中蕴含的学习规律和能力缺陷，为动态调整训练计划提供数据支持，形成“数据收集-分析诊断-方案优化-效果反馈”的闭环，实现精准、长时效地提高口语流利程度^[8]。

在具体实施中，需要构建大语言模型驱动的多模态学习数据监控系统，首先将数字课堂各交互终端之间的数据接口打通，全面收集学习者的语音、文本转写、唇语、互动日志等多模态数据，通过模型的整合和结构化处理，剔除无用的信息；其次，从流利度、速度、句法、语用三个方面构建流畅度的量化评价模型，形成分阶段的学习报告。针对发音能力较差的学习者，通过对其进行多模态语音矫正训练，针对语篇连贯性较差的学习者，对其进行多模态语音训练；对篇章连贯性较差的学习者，进行多模态叙事训练；同时，通过汇总同类学习者的共性问题，反向优化教学资源。此外，还需要建立长效数据跟踪模块，定期对学习者口语流利度变化曲线进行跟踪，并对训练梯度进行调整，为教师提供可视化的数据看板，达到技术辅助和人工教学相结合的目的。

3 结束语

综上所述，大语言模型赋能多模态交互的数字

化口语课堂，正将“流利”由应试指标还原为一种思维与情感同步流动的生活体验。基于大数据的汉语二语口语流利性提升路径，可实现技术和教学深度融合，不仅能有效解决传统数字课堂存在的诸多问题，还能为学习者构建一套科学、有效的口语能力提升体系，对于汉语国际教育数字化教学改革具有重要意义。

参考文献

- [1] 陈惠静,黄冠,王倩.数字化叙事在国际中文口语课堂中的应用研究[J].创新教育研究,2025,13(2):196-203.
- [2] 陈绮琪,姚双云.外语焦虑、外语愉悦与交际意愿对二语口语互动流利度的影响[J].国际汉语教学研究,2025(3):36-46.
- [3] 梁永宁.多模态视域下的国际中文课堂语用教学模式探析[J].汉字文化,2023(8):95-97.
- [4] 江宇豪,万齐昕.高级水平汉语二语学习者应答语研究[J].河池学院学报,2022,42(5):52-59.
- [5] 王希竹,金晓艳.二语学习者口语表达流利度的影响因素研究——基于特定情境焦虑视角[J].汉语学习,2021(6):85-93.
- [6] 高思畅,邹申.汉语二语口语韵律能力评估维度的任务类型效应[J].华文教学与研究,2023(4):59-67+91.
- [7] 徐锦芬,冯晓丽.任务复杂度对二语口语产出影响的元分析[J].现代外语,2025,48(1):90-102.
- [8] 张文贤,袁明珠,姚骏.汉语二语学习者互动行为表达需求的调查与分析[J].国际汉语教学研究,2025(3):26-35.

版权声明：©2025 作者与开放获取期刊研究中心(OAJRC)所有。本文章按照知识共享署名许可条款发表。

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



OPEN ACCESS