

基于人工智能的多模态“视听说”高等教育设计教学模式应用研究

陈佳¹ 黄蕙²

1 绵阳城市学院，四川绵阳，621000；

2 重庆医科大学附属儿童医院宜宾医院，四川宜宾，644609；

摘要：当前，人工智能技术发展迅速，高等教育中的“视听说”教学面临着新的机遇和挑战。传统教学模式普遍存在资源更新缓慢、互动不足、评价单一等问题，难以满足培养学生综合语言能力的需求。本文以多模态理论为基础，探讨人工智能和“视听说”教学深度融合的可行性和必要性，从教学资源、教学过程和评价体系这三个方面构建了基于人工智能的多模态教学模式。进一步提出了高等教育设计的教学路径，以期高校语言教学提供新的解决方案和实践方向，为提升学生综合语言能力和学习积极性提供重要参考。

关键词：人工智能；高等教育；设计教学

DOI：10.64216/3080-1494.25.10.026

引言

语言学习的核心关键是交流，“视听说”课程承担着培养学生综合运用语言能力的重要任务^[1]。在当前高等教育实践当中，“视听说”课程存在教材内容陈旧、课堂形式单调以及反馈滞后等问题，使得学生学习效果和实际需求存在一定差距。随着人工智能技术不断发展，教育领域正发生着深刻变革，语音识别、自然语言处理以及智能推荐等技术给课堂带来了更多的可能性，让“视听说”教学能够实现多模态融合与智能化升级。多模态学习理论强调多种感官通道的协同作用，而人工智能的技术支持为这一理论的落实提供了现实基础。

1 高等教育中“视听说”教学的现状与问题

1.1 教学资源更新缓慢，学习体验单一

当前，许多高校的“视听说”课程中，教材和配套资源沿用多年，更新速度明显滞后，教材中的听力材料大多是传统录音，语速和内容与真实语境存在差距，让学生的课堂听力体验无法接轨现实语言环境。与此同时，视频资源数量有限且内容偏重知识点讲解，忽视情境再现，导致学生的学习缺乏沉浸感。

除了资源本身陈旧之外，课程设计也有较强模式化倾向，很多课堂依旧采用教师播放音频、学生跟读或回答问题的形式，教学过程机械重复，学生体验单调。在多模态学习盛行的当下，缺乏视觉、听觉和交互等多元元素的“视听说”教学较为乏力，难以充分激发学生的

感官参与和思维投入，学生学习时更多是被动接受，互动性和创造性不足。长此以往，学生学习兴趣逐渐减弱，甚至出现课堂注意力分散的问题。

1.2 学习评价模式单一，效果反馈滞后

在“视听说”课程中，普遍存在评价体系单一性的弊端，目前学生学习成果大多靠期末考试或者阶段性测试衡量，测试形式主要是书面答题和听力选择题，这种方式忽视了学生口语表达、实时交流、语境适应等实际能力^[2]，许多学生考试成绩虽然较高，但是在真实对话场景中表现较差，这表明评价体系无法真实反映学习效果。

更突出的问题是反馈环节存在滞后性。学生完成课堂练习之后，很难马上获得有效的个性化指导，教师通常要在课后批改录音或者作业，这就导致反馈周期变长，学生在等待过程中容易错失反思和改进的时机。部分高校虽然引入了录音提交与评分软件，然而多数系统只能给出量化分数，缺少针对发音、语调、语义理解等层面的深度建议，这种滞后浅显的反馈方式限制了学生学习进步，增加了教师的教学负担。

1.3 师生互动深度不足，课堂氛围低迷

“视听说”课程本应重视师生之间的交流与互动，然而在现实中，互动形式局限于教师提问、学生回答的传统模式，课堂氛围通常偏向单向传递，导致学生参与度普遍不高，很多时候只有个别学生与教师进行对话，

其他学生处于被动聆听的状态,这种互动无法形成良好的群体学习氛围,使得学生普遍缺少参与感和成就感。

在互动内容方面,教师提出的问题大多是教材例题或者有固定答案的问题,缺少开放性和探究性,让学生很难在交流过程中发挥创造性思维。课堂的时间分配偏向于教师讲解,留给學生自主表达的机会非常有限。由于互动深度不够,课堂气氛显得十分沉闷,缺乏语言学习所需要的活跃环境。随着学习热情不断下降,学生逐渐对英语课程产生距离感,进而影响到“视听说”能力的提升。

1.4 学习动机普遍不足,能力提升受限

许多学生在进入大学以后,学习英语的内在动机明显下降,将“视听说”课程当作必修科目,而非自我发展的途径。在课堂当中,这些学生学习积极性明显不足,部分学生甚至产生了“应付学习”的心理。缺乏强烈学习动机的学生通常不愿意在课下进行自主训练,导致课堂学习效果大打折扣。

外部激励措施的缺失也是导致动机不足的重要原因。传统课程缺乏即时奖励和正向反馈,学生付出努力之后往往得不到及时的认可,进一步削弱了他们的学习热情^[3]。再加上课堂活动形式比较单调,学生很难从中获得成就感和乐趣,久而久之,这些学生逐渐丧失了对“视听说”能力提升的信心。学习动机不足直接限制了学生的学习成果,使得语言综合能力的发展受到阻碍,即便在考试成绩上有所体现,但实际交际能力依旧难以突破,无法满足未来学术研究或者职业发展的需求。

2 多模态“视听说”与人工智能融合的可行性与必要性

2.1 适配教学革新诉求

高等教育改革一直强调课堂教学的互动性和创新性,而传统“视听说”课程的单一模式已经无法满足新的教学需求^[4]。人工智能技术介入之后,语音识别、语义分析以及图像处理等多模态元素能够被有机整合,进而可以为学生提供更接近真实环境的语言输入与输出条件。借助智能算法,教师能够依据学生的表现及时对课堂内容做出调整,让教学活动变得更有针对性和动态性。

在这一过程中,课堂从原来的单向传递逐渐转变为多向交互,学生不只是能通过听和说参与,还可以借助

视频、图像、动作等多模态资源获得多感官刺激,以此提升语言学习的深度。教学方式更新以后,课程在设计方面会更具灵活性,也能让学生的学习体验变得更加丰富,这就形成了适配高等教育革新方向的现实可能。

2.2 契合互动跃升需求

语言学习的关键在与交流和互动,但是现实当中的课堂常常存在师生互动不足、同伴交流缺乏等问题。人工智能和多模态技术的结合给互动质量和数量都带来了新的可能,基于智能语音识别的对话系统能够模拟真实的场景,学生与系统进行交流时,不但可以锻炼口语表达能力,还能在出现错误时马上得到修正。

互动的范围不只局限在师生间,还包括人机互动和同伴之间的协作。在多模态学习环境中,学生能够通过虚拟情境开展角色扮演,依靠视觉线索和语音提示推动交流,这种方式能够激发学习者的参与热情。互动水平提高以后,课堂从被动接受知识转变为主动探究知识,教学的吸引力和有效性也因此得到增强。

2.3 贴合能力培育诉求

“视听说”教学的核心目标是提升学生的语言综合能力,人工智能与多模态的结合正好契合这一诉求。利用多模态资源进行教学,能让学生在听、看、说、思等多环节接受训练,如AI提供语音纠错和语调分析,可帮助学生改善发音情况;图像与视频的引入能让学习内容变得更加直观,有助于学生加深对语义和语境的理解;智能推荐系统会依据学生表现推送个性化学习任务,让不同水平学习者都获得适合的提升路径。这种能力的培育不仅局限于语言技能方面,还扩展到批判性思维、跨文化交际等方面。学生在多模态环境中不断进行综合训练,逐步形成适应未来学习与职业场景的能力结构,教学模式的应用价值因此凸显。

2.4 呼应评价优化需求

评价体系是否科学决定着教学改进的具体方向。传统“视听说”课程评价通常看重成绩,忽视了过程性表现和能力发展动态变化,人工智能技术可以实时收集学习数据,从语音质量、互动频率到情绪表现,都能形成可量化指标。借助这一优势,多模态“视听说”教学能建立更加全面的评价机制。在课堂中,AI系统能为学生口语表达提供即时反馈,涵盖发音、语速、语调等细节^[5]。教师分析数据时不再局限于最终成绩,而是追踪学

生成成长轨迹和学习习惯,这样的动态评价方式既为学生带来了个性化指导,也推动教师不断优化教学设计,形成良性循环。

3 基于人工智能的多模态“视听说”教学模式构建要素

3.1 教学资源设计

教学资源是“视听说”课堂的核心载体,资源设计情况决定了学生的学习深度与广度。在多模态和人工智能技术的支持之下,资源不再局限于传统音频与文字形式,而是融合了视频、图像、语音合成、交互动画等多元形式,这样的资源能够从视觉、听觉和语言等多个感官通道同时作用,帮助学生形成立体化语言理解和表达能力。人工智能的介入让资源具备动态更新与个性化推送的功能,学习者在平台上完成任务之后,系统会依据其答题正确率、语音准确度和学习时长等数据,自动筛选出更适合学习者的资源,比如对发音存在明显问题的学生,平台会推送标准发音的跟读材料并生成个性化语音练习包,对于听力理解能力比较弱的学生,会提供逐级递进的听力素材,保证学习目标逐步达成。

资源设计不仅要关注形式是否多样,还要注重内容的真实程度和时代特性。利用人工智能的文本以及语料库检索功能,能够从海量真实语境当中提取适合教学的素材,如新闻报道、影视片段和国际会议演讲等。这些材料和学生的实际需求高度契合,让学习过程不再脱离现实生活场景。课堂和现实之间的距离因此被大大缩短,学生在学习过程当中能获得更强的迁移与应用价值。

多模态资源的整合还可以激发学生的学习兴趣。传统课堂中单调的音频播放很难长时间吸引学生的注意力,而视频、动画和互动练习等形式能让学生始终保持投入的状态。教师在设计资源时,可以将人工智能生成的交互式情境融入到课堂当中,让学习活动变得更加灵活多样。资源的全面优化不仅能满足不同水平学生的学习需求,还能推动教学目标从知识传授向能力培养方向转变。

3.2 教学过程设计

教学过程是实现资源价值的核心环节,涉及目标设定、课堂活动组织以及学习反馈的全链条内容。在人工智能和多模态技术的支撑作用下,教学过程不再是线性的知识传递模式,而是呈现出循环往复、动态调整的显

著特征。课堂从导入一直到总结的每一个环节,都能够通过智能化技术提升效率和效果。

在课堂导入阶段,教师可以借助智能语音助手或者多模态展示工具,为学生营造沉浸式的语言学习环境。比如,系统能够通过视频片段搭配互动问答的方式,引导学生快速进入到学习的情境中,增强他们的参与感。

进入知识讲解环节之后,人工智能能够依据学生的实时反应提供数据支持,教师可以根据系统生成的反馈调整讲解的深度或者速度,进而实现个性化的教学。

在组织学习活动时,学生不只通过听和读完成任务,还能借助语音输入、图像识别和角色扮演等多样形式参与课堂,这些活动依靠人工智能平台的支撑,能够实现即时记录和自动分析。例如,系统可以记录学生口语表达过程,识别其中的语音错误,并且及时提供纠正建议,让学生在活动中边练习、边改进。

反馈环节的优化同样是教学过程设计的重要组成部分。传统反馈通常存在滞后性,而人工智能能即时生成个体学习报告,包含语音准确率、语法错误率、词汇使用频率等多项指标。这些数据不但有助于学生自我反思,还为教师提供直观参考,方便在课堂内外实施针对性指导。教学过程在这种动态循环当中不断完善,最终让学习变得更高效且更具针对性。

3.3 教学评价体系设计

教学评价是检验教学质量和学生发展水平的重要工具,人工智能赋能的多模态“视听说”课程评价体系强调全面性与动态性,突破了以往单一成绩考核存在的局限^[6]。这一评价体系不仅包括结果性考查,还注重学习过程中的表现和能力的持续提升。

在人工智能技术的支持下,评价维度得以显著扩展。传统课堂中,教师依靠笔试或口试来评判学生水平,而智能系统则能够细化到发音清晰度、语调自然度、语义连贯度等具体指标。通过对语音、视频互动和学习行为数据的多维采集,系统能够建立学生的学习画像,动态展示其语言能力的变化趋势。评价由静态的终点结果转向动态的过程跟踪,这种方式更符合语言学习的规律。

评价体系具有显著的个性化特征,不同学生在学习过程中表现出的优势与不足各不相同。人工智能平台依靠算法分析为学生推送个性化评价反馈,如有的学生需要在词汇使用方面获得更多指导,有的学生则需要在语调自然度上开展训练,系统会针对这些差异给出精准改进路

径,教师可以依据这些数据进行因材施教,避免“一刀切”的弊端。

此外,教学评价体系还担负着激励功能,通过即时反馈和可视化的进步曲线,学生能直观感受自己的成长,这无形中增强了学生的学习动机。智能化评价结果不仅服务于学生个体,也能给教师提供课堂整体诊断数据,帮助教师优化教学设计。从单一分数到多维画像的转变,代表了教学评价体系从粗放管理走向精细化、科学化的趋势。

4 基于人工智能的多模态“视听说”高等教育设计教学路径

4.1 资源整合:多模素材智能推荐

在人工智能支持下,教学资源的整合不再局限于单一渠道,而是通过算法实现动态优化与精准分配,教师 and 系统应一起建立资源数据库,将影视片段、新闻音频、互动动画、语音合成材料和实时对话场景进行分类整理。人工智能可以依据学生的学习水平、兴趣偏好以及以往任务表现,推送差异化资源来确保学习过程个性化、动态化。系统还可以通过数据挖掘预测学生可能会遇到的瓶颈,并且提前准备好补充材料。

在实际应用时,学生登录到学习平台以后,系统会根据他们的口语流畅度和听力理解情况自动生成学习任务,发音不准确的学生能收到配有口型演示的语音训练视频,并且在跟读环节获得即时评分;词汇掌握不足的学生则会被推送语义场景化的单词学习模块,如餐厅、交通或者学术讨论场景,通过短片和互动问答扩展词汇量。教师在课堂中利用系统推荐的资源安排分组活动,每个小组接收到的素材都是不一样的,以此保证任务难度和学生水平相匹配。通过这种方式,资源从静态储备转变为动态流动,学生不会再被单一教材所束缚,而是在多模态智能推荐当中形成持续的学习路径,这种整合策略不但提升了课堂效率,还明显增强了学生的学习动力与自主性。

4.2 环境营造:沉浸课堂互动增强

多模态教学强调情境化和体验化,人工智能给沉浸式环境提供了坚实技术基础。课堂要通过虚拟现实、智能投影以及语音交互系统营造真实语境,让学生在模拟场景中完成相关的语言任务。人工智能能够实时识别学生的语言表达情况,还能调整场景进展,让课堂氛围保

持高度互动和沉浸感。环境设计不再只是单纯依靠教师进行讲授,而是借助技术赋能实现多方位的参与。

在教学中,教师可以设计一个模拟商务谈判任务,让学生分组进入虚拟会议场景。系统会依据学生的表达内容生成即时回应,如对观点提出质询或者给出反馈,推动交流持续深入。另一部分学生进入虚拟新闻演播室,系统要求他们完成播报任务,并且会对其语速重音和语调进行实时评分,学生在场景中不断调整表达,形成语言运用和情境反应的紧密结合。课堂之外,学生还能通过移动端进入虚拟市场,与系统角色进行交易对话来练习日常口语。教师在后台实时观察数据并且及时调整课堂设计。沉浸式环境让学生打破单调的课堂局限,主动投入语言实践,这不仅提升了语言能力,还增强了跨文化交际意识。

4.3 学习支持:智能反馈动态优化

学习支持环节直接决定学生能否持续进步和学习体验的好坏,教师要建立一个基于人工智能的动态反馈机制,对学生的语音、语法、语义理解以及课堂互动进行多维度监测。反馈既需要做到即时又要做到个性化,人工智能可以在学生练习过程中捕捉关键数据,并且生成错误提示与改进路径让学生在学时时不断进行自我修正,教师通过这些反馈开展针对性指导,真正实现教与学的双向优化。

在课堂口语展示环节,学生进行一分钟陈述,系统实时捕捉其语音特征。若出现发音不清晰的单词,系统立即标注并在屏幕上显示标准发音,学生可以立刻进入跟读训练,比较语音曲线并进行调整。另一位学生完成听力任务时,系统根据其回答内容分析理解偏差,并推荐难度更低的补充材料,例如提供带关键词提示的音频,再次完成练习以加深记忆。课后,平台为每位学生生成个性化学习报告,包括语速、语调、语法错误率以及课堂互动频率。学生根据反馈自主调整学习计划,而教师通过数据了解班级整体情况,决定是否在下一课程中加入专项训练。这种动态反馈机制不仅优化了学习路径,还提升了学生的自我调节能力,使学习更具连贯性和有效性。

4.4 评价改进:多维测评促进发展

评价体系对于教学模式落地来说非常关键,应建立以人工智能为支撑的多维测评体系,涵盖语言产出质量、

学习过程参与度和能力发展趋势等方面。系统会利用语音识别、自然语言处理和学习行为分析等技术,为学生建立能够反映阶段性成果和记录长期成长轨迹的动态学习画像。评价目标不只是为了甄别学生的水平,更重要的是促进学生发展,帮助他们明确自身优势与不足。

在课堂互动环节,系统会实时采集学生的发音准确率、语速流畅度和词汇多样性,生成即时可视化曲线展示其能力分布。小组讨论结束之后,平台会根据发言次数、逻辑性和互动深度生成多维评价,并且结合学生的语音表现进行综合评分。学生能够在报告当中直观看到自己在口语自然度、语法正确率或者听力理解等方面的强弱点,课后系统还会生成个人能力画像,指出如语音自然度偏低或者词汇使用不足的问题,并且推荐改进任务。教师可以利用这些数据设计有针对性的课程活动,从而使教学更具科学性和精准性。通过这种多维测评方式,学生逐渐摆脱对单一分数的依赖,转而在持续反馈与动态监测中建立成长意识。这种评价模式不仅提升了学习动力,也推动了高等教育“视听说”课程的整体质量提升。

5 结束语

多模态和人工智能相结合为高等教育“视听说”课程增添了新的活力。借助资源的多样化与智能化、过程的互动化与个性化、评价的动态化与精细化,教学模式逐渐突破了传统课堂的局限。高校要在政策支持和技术保障的前提下,逐步推进人工智能赋能的多模态教学模式,让其真正成为促进教育创新和学生发展的有效途

径。

参考文献

- [1] 曾庆敏. 多模态视听说教学模式对听说能力发展的有效性研究[J]. 解放军外国语学院学报, 2011(6): 5.
- [2] 王晓玲. 大学英语“视听说”多模态教学路径构建研究[J]. 牡丹江教育学院学报, 2023(8): 63-67.
- [3] 王晓星方兰. 多模态初级日语视听说教学模式设计与实践[J]. 西部学刊, 2024(8).
- [4] 丛丽君. 基于产出导向法的大学英语视听说课程多模态教学机制构建研究[J]. 当代教研论丛, 2025, 11(2): 21-24.
- [5] 郭新雨, 刘晶波, 杨雯嘉, 李松晓, 周悦楠. 视听说课程多模态混合式学习模式研究[J]. 英语广场: 学术研究, 2023(19): 90-94.
- [6] 黄倩妮. 多模态视域下英语专业视听说课程教学研究[J]. 英语广场: 学术研究, 2023(21): 83-86.

作者简介: 陈佳, 1990 年 11 月, 女, 汉, 四川德阳, 绵阳城市学院, 本科, 副教授, 文创设计, 四川省绵阳市安州区花荪镇绵阳城市学院。

黄蕙, 1997 年 8 月, 女, 汉, 四川自贡, 重庆医科大学附属儿童医院宜宾医院, 硕士, 设计心理学。

项目基金: 四川省民办教育协会 2024 年重点科研项目《基于人工智能的多模态“视听说”高等教育设计教学模式应用研究》(项目编号: MBXH24ZD15)。