

A Study on Coordinated Pathways of Gen AI-Assisted Multimodal Cultural Teaching in International Chinese Education

Gen AI 辅助国际中文教育多模态文化教学协调路径探究

| Ouyang Yujuan |

*School of International
Chinese Studies, East China
Normal University,
Shanghai, China*

yujouyang@163.com

ABSTRACT

The rapid advancement of digital intelligence presents unprecedented opportunities and challenges for international Chinese language education, particularly in cultural instruction where learners often struggle with abstract concepts. Therefore, the primary objective of this study explores how Generative Artificial Intelligence (Gen AI) can be utilized to enhance the multimodal discourse competence of international Chinese learners and thereby identify innovative pathways for Chinese cultural teaching. Grounded in social semiotics and multimodal theory, this research employs an empirical inquiry through a task-driven case studies. Specifically, it examines the cognitive mechanisms underlying the cross-modal transformation between learners' "linguistic prompt construction" and "visual semantic generation" during human-machine interactions. Empirical findings from systematically structured, multi-level multimodal cultural teaching cases indicate that the cross-modal transformation facilitated by Gen AI effectively reduce learners' cognitive load during second language acquisition. By providing dynamic visual feedback, Gen AI assists learners in constructing rich and meaningful mental representations of Chinese culture. In conclusion, to mitigate potential algorithmic cultural biases and optimize learning outcomes, this study recommends incorporating a Gen AI interaction literacy module into the curriculum to cultivate learners' critical competencies in human machine collaborative communication. This pedagogical paradigm has the potential to advance the digital-intelligent transformation of Chinese language education and ultimately enhance the effectiveness of global Chinese-language communication.

KEYWORDS

Cultural Teaching; Digital Intelligence; Generative Artificial Intelligence; International Chinese Education; Multimodal Discourse

引言

1.1 研究背景

当今,人工智能技术(Artificial Intelligence,以下简称AI)迅猛发展并广泛普及。在此背景下,国际中文教育正经历着数智化转型。在新媒体语境下,以图像、视频和超链接为核心的多模态语篇已成为信息传播的基本方式(徐燕 & 冯德正, 2023)。传统国际中文教育侧重于语言符号处理,而较难直观呈现语言背后的深层文化意象。随着生成式人工智能(Gen AI)在教学中的广泛使用,促使学习者从传统读写转向多模态交互与创作。因此,当前中文学习不仅要求学生具备语言和文字产出能力,在还需拥有构建多模态转换能力,特别是跨文化语境下的多模态解码与重构能力。用户通过语言指令与Gen AI进行交互,本质上是对社会符号协同与重构的过程(李晓蔚 & 张又雄, 2025)。在这种交互中,国际中文高级学习者需要将抽象的中文概念(如文化语境、诗歌意象和文学作品)转化为多模态描述。这意味着,学习者要从语言文本理解和处理到视觉-听觉符号的输入和输出。该过程不仅涉及到词汇衔接,还关系到文本流利度、语篇衔接、文化审美表达与多维交互模式的复杂动态系统(张超 & 梁文花, 2022)。

尽管多模态话语研究是语言学研究领域热点,但关于国际中文学习者如何在AI交互中进行跨模态转换的研究尚处于起步阶段。以往研究多关注视觉元话语的传播组织功能(徐燕 & 冯德正, 2023),或侧重利用AI构建知识图谱以辅助协作学习分析(郑兰琴等, 2022)。然而,探讨国际中文学习者如何通过撰写中文指令驱动内在语言与多模态文化理解力之间的互动,具有重要意义(徐燕 & 冯德正, 2023)。本研究通过分析学习者在AI生成过程中如何平衡中文表达的流利度与语篇衔接,试图揭示语言指令与视觉结果间的认知映射。为国际中文高级学习者构建多模态语言输入-输出框架,从而帮助其提高国际中文综合运用能力。本研究旨在探讨AI技术能够构建一个结构化的反馈场域:它通过生成可视化的图像结果,为学习者的中文产出提供实时、直观的反馈,从而帮助学习者跨越跨文化交际中的认知鸿沟。本研究探讨的多模态文化教学培养路径,旨在指导教师通过教学激发学习者的创作潜能,训练其精准驱动AI工具,从而提升在数字化全球语境下的中文沟通效能。在达成语言交际目的的同时,深化其对中华文化的感知,培养学习者应对未来多模态社会符号的能力。

1.2 理论基础与文献综述

1.2.1 社会符号学与多模态话语理论

克瑞斯和勒文(Kress & Leeuwen, 1996)构建的多模式话语社会符号学分析以韩礼德的功能语法理论为基础(李战子, 2003),比照系统功能语法的概念、人际和语篇功能,将视觉图像的意义分为“再现、互动、构图”三大元功能。该方法对认识语言社会符号特征、理解多模态话语生成及指导语言教学具有重要意义。多模态话语分析(Multimodal Discourse Analysis, MAD),又叫系统功能多模态话语分析,是基于系统功能语言学为理论基础。韩礼德提出指出语言是社会符号的理论,关注特定文化社团中的符号实践。意义的产生不仅依赖语言, Thibault (2004)和 Scollon & Levine (2004)“多模态语篇”包含视、听、触等多种意义系统,是

通过符号资源互动实现意义的话语，因此要研究话语和技术的关系以及多模态话语分析。数字化浪潮下，人们对意义的理解已从文字识读，转向对图、文、声、影的综合解码与互动（胡壮麟，2007）。在多模态话语分析中，语言学层面也会通过多维动态进行连接与互补。（张德禄 & 郭恩华，2013）探讨多模态话语的两个视角：基于系统功能语言学的“社会符号学视角”，解析符号资源在社会语境中的协作；基于认知语言学的“概念隐喻视角”，强调多模态隐喻的认知理据，两者结合能促进该理论发展。针对新媒体与数字化语境，多模态话语的分析框架得到了进一步拓展。（吴兰，2021）搭建了包含语境层、信息层和传播层的三维框架，指出信息层由符号与意义构成，传播层揭示了多模态符号在界面反馈中的动态关联。在国际中文教学领域，（李雅，2018）强调多模态理论对教学转型的指导意义，呼吁利用媒介升级改进教学团队和教材建设，以应对跨文化交际需求。

1.2.2 多模态与生成式 AI 环境下的教学现状

谢蓉蓉（2012）曾融合多模态话语分析理论、现代数字技术及虚拟现实，提出“多模态虚拟课程”概念。随着人工智能（AI）技术的介入，多模态研究进入了交互与生成新范式。马小伟（2024）指出，语言教学正向多模态信息化转型，利用 AI、3D 技术构建多维感知、多向互动的环境，解决了传统教学单调的问题，视觉模态作用愈发凸显。朱淑仪等（2022）通过对比研究发现，在线上汉语课堂中视觉模态有着重要的表意功能，线上教学的图像精准度与文字密度普遍高于线下。对语言视听说的教学研究，刘燕（2018）指出多模态话语能有效刺激大脑皮层，通过视听资源的优化组合提升学习效率与自主学习能力。学习者在可以在多模态空间内进行母语、目标语乃至方言的多语言动态切换，利用多模态资源（如生成的图片、诗歌、表情符号）进行意义建构。这种行为不仅降低了认知的吝啬负担，更促成了学习者的各种能动性表现，包括自主规划、批判思考、资源融合及情境感知（方帆 & 吴倩倩，2025）。当然，多模态模式不能完全取代传统教学。实证研究表明，若多模态素材缺乏有效解构引导，过多符号资源反而会削弱语言习得效果。因此，教师在素材选取及语境营造中，仍需保持掌控权与教学张力（周红红 & 范景红，2014）。

Gen AI 已广泛应用于课堂任务设计、学习指导及评价。其赋能教学的研究集中于界定素养内涵、嵌入教学机制与构建人机协作机制。通过对美国 K-12 教育指南的考察指出，AI 素养已成为数智时代学生发展的核心，研究者们多以感知、表示与推理、学习、人机交互和社会影响五大概念搭建主体框架（邓莉等，2025）。在交互实践中，提示词工程是连接人机认知的窗口。进一步证实，提示词框架对生成内容的质量具有决定性影响（史金生 & 葛星辰，2025）。通过对比 ICIO、CRISPE、BROKE 和 RASCEF 四类结构化框架发现，CRISPE 框架在国际中文教学资源生成中表现最优，其通过明确角色、背景、任务、风格和实验等要素，能显著缩小模型性能差异，使教师从单纯的工具使用者转向交互生态的设计者。高校教师如何将 Gen AI 赋能教学也是目前研究关注的重点。通过对麻省理工学院（MIT）案例的研究发现，Gen AI 赋能呈现出技术支撑、教学引领与内容拓展三重逻辑，促使教师从“课程执行者”向“人机协同引导者”转变（夏琪等，2025）。在教学设计上，基于 SAMR 模

型（替代、增强、修改、重新定义），引入 CRAFT 与 SHIFT 框架，解释了 Gen AI 如何重构课程、课堂与评价体系（邓莉等，2025）。Gen AI 也塑造了新的课堂话语空间。语言教学正从单一文字输出，转向人机交互的意义协商。学习者撰写指令，本质是利用中文与 AI 进行符号协同与意义重构，要求学习者具备将“中文意象”进行跨模态映射的能力。该过程具复杂动态系统特征，文本流利度、语篇衔接与多维交互间存在协同或竞争，此逻辑同样适用于 AI 指令输出分析。然而，Gen AI 也带来算法偏见、信息幻觉、教师能力焦虑等伦理风险，引发对技术中立性的批判反思。

尽管多模态资源已广泛应用于国际中文教育，但 Gen AI 辅助下的中文学习相关研究仍存不足：首先，现有研究多集中于多模态资源的呈现（刘燕，2018；朱淑仪等，2022），而较少探讨学习者作为生成者时，如何构建跨模态指令。其次，多模态在国际中文教育领域具有较大潜力，但仍然缺乏实证研究支持 Gen AI 驱动的产出受限（如撰写 Prompt）对学习者长期语言能力的影 响（李雅，2018）。此外，当前的评价体系仍依赖传统读写能力，未能整合智慧教育与多模态识读所需的指标（胡壮麟，2007；马小伟，2024）。在中文文化意象生成、汉字视觉逻辑转换等领域，如何评估学习者在 AI 交互中的语义解码与重构能力仍是空白。

综上所述，Gen AI 在通识教学和语言习得中有赋能价值，但在国际中文教育，如何通过高效提示词策略提升资源质量、培养多模态能力，仍待探究。

研究方法

Gen AI 为国际中文教育带来深刻启示。教学模式由“媒介辅助”向“人机协同”的转变，可适时审视学习者的语言能力构成：学习者不仅需具备基础的听说读写能力，更需具备跨越语言符号、视觉图景与机器逻辑的多模态话语理解和处理能力。为探究生成式人工智能（Gen AI）在跨模态转换中对二语认知负荷及文化意象建构的影响，本研究采用任务驱动的案例研究法展开实证分析。本节将阐述具体的研究设计与教学目标重构路径。

2.1 研究设计与参与者

国际中文教育的传统课堂评估局限于语言表达的正确性，评价标准通常由语法规则、词汇等级和发音构成。人机交互涉及复杂的意义协商过程，在 Gen AI 辅助的文化教学中，单纯的语法正确无法实现深层文化意义的传递。因此，本研究采用任务驱动的案例研究法，便于在真实教学情境中观察干预过程，并通过收集指令文本、生成图像及互动修改等多模态数据，形成实证互证。

在案例选择上，本研究采用目的性抽样，选取了中国西部高校中来华留学生的传统文化课程作为研究对象。案例中的国际中文学习者汉语水平处于 HSK2-5 级不等（分为初、中、高级），具备基础日常交际能力，但在处理抽象文化议题时普遍存在认知负荷。在任务执行环节，本研究记录了学习者及教师利用 Gen AI 进行文化图像创作的交互全过程。

2.2 教学目标与路径建构

国际中文教学模式正逐步由媒介辅助转向人机协同。在这一背景下，学习者的语言能力构成需要重新审视：除了基础的听说读写，学习者还需具备跨越语言符

号、视觉图景与机器逻辑的多模态话语处理能力。为此，本节通过重构教学目标，探索从传统语言习得向多模态应用转变的路径。

多模态话语协调指学习者输入中文指令，调动 AI 生成并输出匹配的文化语义图景，确保二者在语义和情感上保持一致。与追求标准句式相比，多模态协调更关注语言的“效能感”。在跨文化交际中，由于词汇量不足或受母语认知负迁移影响，中文表达往往倾向于空泛。例如，学习者写出语法完美但缺乏文化特征的句子“一张美丽的画”。这对 Gen AI 是低效指令，生成结果往往模糊。反之，如果学习者能够使用“中国画”“小雨”“远山”“小桥”等精准词汇，即使句式简单，也能触发 Gen AI 生成契合的中华视觉图景。这种语言输入与输出的匹配体现了多模态话语协调能力，能有效弥补中文学习者文化产出能力的不足。

为实现学习者交际能力从语言描述向多模态话语协调转变，文化教学目标需在以下三个层面进行重构：一是在语言描述上，从表意明确到细节具象。文化教学不再只是让学生“把话说对”，而是要学会用语言“把文化画出来”。针对中文学习者难以具象描绘文化场景的特征，教学需重点训练空间方位词（上、下、内、外）、色彩饱和度及中国传统色、质感词汇的使用。学习者需掌握如何通过这些具体词汇，将抽象的跨文化思维转化为 Gen AI 可识别的空间逻辑与文化符号。二是在教学路径上，全球范围内的中文学习者往往缺乏真实的中华文化沉浸语境。因此，课程需将提示词策略纳入文化教学目标。学习者需学会在中文语境下，通过文化身份设定、历史上下文约束和逻辑链引导，与 Gen AI 进行多轮对话。这种策略互动能修正 AI 生成结果的文化偏差，同时锻炼学习者在跨文化交际中的意义协商能力。三是学习者在初期通常缺乏深层的中国文化图式。在处理文化内容时，教学目标应转向培养对特定中文意象的视觉化生成。这要求学习者在理解词汇字面意义的基础上，感知其背后的历史意境，并将其运用于多模态话语协作中。由此，帮助学习者完成从被动理解中国文化到主动产出跨模态文化视像。

研究结果及探讨

在课堂教学中，任务型教学最利于培养学习者的多模态话语能力。将 Gen AI 引入课堂，可以设计出既符合教学目标，又兼具挑战性和交互性的学习任务。针对国际中文学习者特征，AI 的视觉辅助能有效降低他们的跨文化表达焦虑，激发交际动机。本研究基于真实中文文化课堂环境中的多模态数据收集，考察了不同汉语水平的国际中文学习者在 Gen AI 辅助下的任务执行及教师指导修正的指令生成表现，以下将分别按照初、中、高级阶梯式任务的实施结果，对此进行深度探讨。

3.1 基于生成式 AI 的课堂任务实施结果

在国际中文教育阅读课程中出现的诗歌和文化俗语蕴含着深厚的文化意象，这些意象往往具有抽象性和多义性。对于缺乏中国深层文化图式的国际中文学习者而言，传统的语言讲解很难让他们真正体会其中的意境。对于蕴含文化意象的作品创作表现的课堂旨在引导学习者将抽象的文化意象转化为 AI 绘画指令，培养其跨模态的文化转换能力，本研究记录了学习者在初、中、高级三个层次的文化意象转化任务中的具体行为表现。

案例一（初级任务）：“亭中饮茶”文化场景

教学挑战：初级学生在描述中国文化场景时，常受母语负迁移影响，遗漏方位词（如“在……里/上”）或错用存现句，导致空间逻辑表达不清。

任务流程：

- (1) 语义拆解。教师引导学生分析场景的视觉及空间要素：文化地点（中式亭子）、人物动作（老人喝茶）、细节道具（桌子、茶壶）。
- (2) 指令构建。若学习者直接输入受母语影响的初级指令“老人在亭子喝茶”（图 1），AI 会画出缺乏包含中国文化元素的图像（图 1）。教师需启发学生使用脚手架拆解提示词：文化地点与方位（在哪儿）：“在一个亭子里面”；人物与动作（谁+动词+着）：“一位老人坐着喝茶”；细节定位（名词+方位词+放着）：“石桌子上放着一个中国的茶壶”（图 2）。



图 1. 初级指令生成图像

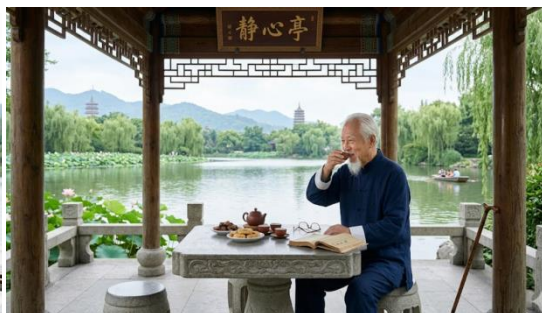


图 2. 修正指令后的生成图像

- (3) 模态验证。若 AI 画面中茶壶飘在空中，或者老人没有坐在亭子内部，学习者应直观地意识到方位词“上”或“里面”的缺失，进而修正指令，向存现句的标准语序对齐。
- (4) 能力提升。该任务利用视觉反馈纠正了基础方位词和存现句结构，同时让初级学习者在多模态重构中，自然接触并内化“亭子”、“茶壶”、“饮茶”等基础文化元素。这一过程证实多模态重构有效帮助初级学习者内化了基础文化元素与空间语法。

案例二（中/高级任务）：“江南水乡”的文化景观风格

教学挑战：中高级学生面临词汇瓶颈，在表达中国特有的文化美学时，往往全靠“漂亮、好看”等基础词汇，缺乏深层文化语汇的调用能力。

任务流程：

- (1) 语义拆解。教师引导学生分析江南水乡的核心文化视觉要素：特色建筑（旧房子/徽派建筑）、典型水景（木船/乌篷船）、特定氛围（下雨/烟雨朦胧）。
- (2) 指令构建。若学习者输入基础写实指令“春天，墙是白的，屋顶是黑的，河上有一条木船，正在下小雨”（图 3），画面往往像普通的现代农村下雨风景照，缺乏特定的文化底蕴。教师需启发学生进行词汇升级，构建文学意境指令：如“烟雨江南”、“白墙黑瓦的传统徽派建筑错落有致”、“一条

乌篷船在碧波中荡漾”、“岸边的青石板路十分湿润”；文化审美限定：如“具有中国诗意的美学风格，色调淡雅，充满古典意境”（图4）。



图3. 基于基础词汇的生成图像



图4. 添加文化审美限定词的生成图像

- (3) 模态验证。学习者通过对比两组指令生成的画面差异，直观感受到“文化词汇”带来的画风突变。若初次生成的画面依然偏向现代写实，学习者需意识到基础词汇的局限，主动增加“诗意”、“古典意境”等审美限定词来修正偏差。
- (4) 能力提升。这种基于视觉差异的强烈反馈，极大提升学习者对中文特有文化词汇（如“烟雨江南”、“乌篷船”）的认可和了解，帮助他们突破日常交际词汇的瓶颈，顺利进入文化审美的语境。

案例三（高级任务）：“大漠孤烟直”的视觉重构

教学挑战：高级阶段的学习者需要理解“孤”与“直”在古汉语中传达的独特空间美学。

任务流程：

- (1) 语义拆解。教师引导学生分析诗句的视觉要素：环境（大漠）、主体（烟）、动效（直）。
- (2) 指令构建。若学习者直接复制粘贴“大漠孤烟直”，生成效果图5。输入初级指令“沙漠里的烟”（图6），画面往往缺乏美感。教师需启发学生分解提示词：环境描写（形容词+名词）：如“广阔的黄色沙漠”“落日”；主体状态（数量词+形容词+名词+动词趋向）：一道笔直的烟雾直冲天空……；文化与风格限定（固定短语）：中国传统水墨画风格，留白，孤独感（图7）。



图5. 古诗原句的生成图像



图6. 基于基础词汇的生成图像

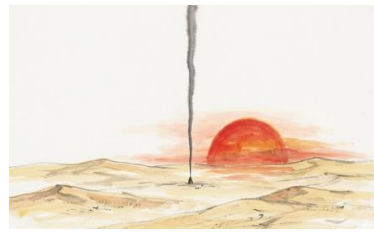


图7. 融合文化限定词的生成图像

(3) 模态验证。若 AI 画面中出现了杂乱的村庄或多处烟火，学习者应意识到对“孤”字的理解存在偏差。进而将指令修正为“沙漠里只有这一道孤零零的烟”。

(4) 能力提升。该任务促使高级学习者从美学角度审视汉字（如动词“直”、形容词“孤”）的张力，实现了中文空间逻辑的可视化生成。

3.2 基于 AI 反馈的描述语更正与润色

在多模态话语语篇中，AI 生成的图像误差实际上构成了二语习得过程中的负相关证据。在传统课堂中，学生写错一个方位词或趋向补语，教师的修改往往是枯燥的。而在本研究中，视听反馈带来的语义落差成为了引导学习者进行自主的语言能力修正的动力。

案例四：语法描述的风格迁移

教学挑战：国际中文学习者面临词汇瓶颈，表达事物时全靠“很漂亮、很好”。需要通过两组指令的对比，让其体验文学词汇带来的画风突变。

场景示例：同样的场景（如：上海的春天），要求学习者撰写两组提示词。第一组使用写实的、教科书式的语言“春天，上海有很多漂亮的花。街上有很多人和车，天气很好。”（如图 8）；第二组使用富有文学气息的、具有特定情感色彩的语言“上海的春天。温暖的阳光照在路边的梧桐树上，树上开满了美丽的花。街上有很多车和人，非常热闹，两边的老房子很有历史感。请画出电影的感觉，看起来很浪漫。”



图 8. 基于客观写实语的生成图像



图 9. 基于文学语体润色后的生成图像

多模态话语对比：观察两组指令生成的画面差异。学习者会直观地感受到两种语体的指令产生的画面有本质不同。

观察表明，当生成的图像（图 9）呈现出强烈的浪漫与历史交织的氛围，图像视觉差异的反馈使学习者深有感触。这种体验不仅促使他们摆脱了依靠“很好、很漂亮”进行表达的惯性，更能激发学习者对中文词汇微小语义差别学习动机，促进高级汉语学习突破在学习中出现的近义词瓶颈，开始有意识地在写作中进行修辞与语体风格的润色。

结论

4.1 实践意义

上述实证观察与案例表明, Gen AI 在国际中文教学中的应用并不是简单的工具替代, 而是重新构建了意义协商的教学模式。本研究深入剖析了在 Gen AI 交互环境中, 多模态话语能力在国际中文文化教学与意义重构中的实践路径。

首先, 实现从语言符号向语义空间转换的认知。在传统教学模式中, 国际中文学习者主要依赖课本上的文字符号进行语言的输入和输出。由于绝大多数学习者的母语与中文存在巨大的形态差异, 这种单一的符号记忆模式容易造成认知断层, 使得汉字背后的深层意象难以被直观理解。引入 Gen AI 后, 其“提示词转换机制”能将抽象的语言指令瞬间转化为具象的视觉画面。学习者在向 Gen AI 撰写提示词时, 为了让机器“画对”, 必须反复进行精准的词义辨析与逻辑重组。这一过程, 促使他们从被动的“接受语言知识”走向主动的“多维空间建模”。这种能力的训练不仅促进了学习者对中文深层语义的提取, 还在语言输入与视觉输出之间形成了双向反馈路径, 大幅提升了二语习得过程中的意义协商效率。

其次, 提升跨文化图式构建的能力。受限于有限的二语词汇量, 国际中文学习者在处理复杂的中文长句时, 往往承受着极高的“认知负荷”。通过 AI 生成的图像反馈, 学习者能够将原本零散、堆叠在短时记忆中的语法细节, 迅速转化为长时记忆中稳定的视觉图式。

最后, 多模态能力的培养不能脱离实际的中文语用环境, 应将其视为人机协同的意义协商。在教学目标上, 教师需要突破语法标准的评教规则, 进入多模态话语协调的教学模式; 在教学任务设计上, 仍要充分利用生成式 AI 的创造性生成能力与偏差反馈功能, 将语言训练转化为多跨模话语实践。这种多模态的介入显著降低了二语学习中的认知损耗, 使学习者能将省下来的认知资源投入到深层的语用逻辑与内涵理解中。在原有的文化理解层面, 多模态话语能力使学习不再停留在表层知识, 而是深入意境解读。例如, 通过对松、竹、梅或古典诗歌意象的数字化重构, 学习者能在视觉逻辑中直观识读其中的空间美学与价值取向。这有助于学习者跨越母语图式的限制, 在母语与目标语图式之间建立起有效的映射关系。

4.2 教学建议与反思

基于上述讨论, 针对当前 Gen AI 在国际中文教育中的应用现状, 本研究提出以下两点建议与反思。

首先, 警惕算法偏见, 引入批判性多模态素养教育。生成式模型主要基于海量的跨文化数据进行训练。在处理特定民族元素时, 模型往往存在算法偏见, 容易输出刻板化或过度简化的中国符号。国际中文学习者由于对目标语的辨别力尚在发展阶段, 若过度依赖这些样板化图像, 极易对中文表达背后的细微差异产生误读。因此, 教师在应用 AI 工具时, 必须开展“批判性多模态素养”教育。教学中应强调对 AI 输出结果进行语义校验, 引导学生主动辨析: 算法生成的图像是否真正准确地传达了中文文本的内核? 其次, 完善教学大纲, 融入多模态人机协同模块。当前的国际中文教学大纲主要聚焦于“听、说、读、写、译”五项语言技能, 缺乏对数字化环境下多模态沟通能力的明确界定。随着人机协同成为全球沟通的主流模式, 国

际中文学习者的核心素养应当扩展至“数字媒介的操控力”与“多模态信息的整合力”。建议在课程设置中增设专项模块，系统训练学生掌握中文提示词的工程化构建方法、跨模态信息的解码策略，以及基于 AI 反馈的语言迭代修正能力。同时，教学评估体系也应随之进行结构化修正：不仅要考查学生语言表达的准确度，更要评估其利用中文指令驱动 AI 处理复杂任务的综合效能。未来的国际中文教育应通过构建多维能力培养路径来适应全球媒介变迁，其目标不仅是培养掌握语言工具的人才，更应致力于培养具备数智化交际能力的复合型学习者。

参考文献

- Kress, G., & van Leeuwen, T. (1996). *Reading images*. Brown Prior Anderson.
- Scollon, R., & Levine, P. (2004). *Multimodal discourse analysis as the confluence of discourse and technology*. Georgetown University Press.
- Thibault, P. (2004). *Body dynamics and the multimodal integration of meaning across space-time scales*. Paper presented at the 9th International Pragmatics Conference.
- 邓莉, 赖建桥, & 高妍. (2025). 生成式 AI 重构课堂教学: 基于美国经验的考察. *教育发展研究*, 45(22), 66-75.
- 方帆 & 吴倩倩. (2025). 生成式人工智能背景下英语媒介教学课堂中的超语实践与学习者能动性探究. *外语教育研究前沿*, 8(4), 13-23.
- 胡壮麟. (2007). 社会符号学研究中的多模态化. *语言教学与研究*, (1), 1-10.
- 李晓蔚 & 张又雄. (2025). 社会符号学视域下用户的跨模态 AI 交互行为及多模态素养教育研究. *情报资料工作*, 46(4), 97-105.
- 李雅. (2018). 多模态话语分析理论对国际汉语教学的启示. *民族教育研究*, 29(5), 139-144.
- 李战子. (2003). 多模式话语的社会符号学分析. *外语研究*, (5), 1-8.
- 刘燕. (2018). 多模态话语分析理论在大学英语听说教学中的有效应用. *教育理论与实践*, 38(27), 55-57.
- 马小伟. (2024). 人工智能时代大学英语多模态信息化教学创新研究. *湖北开放职业学院学报*, 37(11), 9-10, 13.
- 史金生 & 葛星辰. (2025). 基于提示词策略的生成式人工智能在国际中文教学中的应用研究. *陕西师范大学学报 (哲学社会科学版)*, 54(5), 135-145.
- 吴兰. (2021). 新媒体编辑多模态话语的研究范畴与理论框架. *中国编辑*, (1), 31-36.
- 夏琪, 杨一鸣, & 田贤鹏. (2025). 生成式人工智能如何赋能高校教师教学创新——基于麻省理工学院的案例考察. *中国大学教学*, (12), 83-91.
- 谢蓉蓉. (2012). 多模态虚拟课程——基于多模态话语分析理论的未来网络课程新模式. *中国电化教育*, (4), 74-79.
- 徐燕 & 冯德正. (2023). 多模态语篇中的视觉元话语: 以新媒体传播为例. *外语学刊*, (2), 27-33.
- 张超 & 梁文花. (2022). 二语写作文本流利度与语篇衔接的多维交互模式. *外语教学与研究*, 54(4), 557-568, 639-640.

- 张德禄 & 郭恩华. (2013). 多模态话语分析的双重视角——社会符号观与概念隐喻观的连接与互补. 外国语（上海外国语大学学报）, 36(3), 20-28.
- 郑兰琴, 范云超, & 牛佳玉. (2022). 基于在线协作学习交互文本的跨领域知识图谱构建技术. 电化教育研究, 43(12), 70-77.
- 周红红 & 范景红. (2014). 论多模态话语分析理论视域下的英文影视欣赏课教学. 北京交通大学学报（社会科学版）, 13(1), 124-128.
- 朱淑仪, 李苑祯, & 钟婉颖. (2022). 线上线下汉语课堂教学中视觉模态运用的对比分析. 华文教学与研究, (1), 72-78.