

# 多模态视域下外向型AR词汇学习工具的设计与研究

## ——基于HSK词库的实践

祁思宇<sup>1</sup>, 曾祥彬<sup>1</sup>, 兰仁杰<sup>1</sup>, 朱梓源<sup>1</sup>, 吴 铮<sup>2</sup>, 刘 晶<sup>3</sup>

<sup>1</sup>天津理工大学管理学院, 天津

<sup>2</sup>天津理工大学语言文化学院, 天津

<sup>3</sup>天津理工大学艺术学院, 天津

收稿日期: 2025年12月15日; 录用日期: 2026年1月19日; 发布日期: 2026年1月29日

### 摘要

围绕国际中文词汇“形 - 音 - 义”融合学习的难点, 本文依托多模态学习理论研发“纸质 + AR”外向型词汇学习卡片, 并构建“内容规划 - 载体选择 - 技术支撑 - 功能实现”四级体系——依据HSK 1至6级词汇频次重构语境场景, 以纸质载体二维码作为低成本接入渠道, 通过自主研发的轻量化AR引擎实现跨平台识别整合3D立体模型、汉字笔顺动画、场景化例句及手势操控交互功能, 形成四通道协同模式。以动词“打开”为实例设计四步学习流程, 验证后发现该工具可在单次交互中达成“视觉 - 听觉 - 动觉”同步联动, 有效减轻学习者认知压力, 提升词汇记忆效率与知识迁移能力, 为国际中文词汇教学提供可复制且低成本的数字化解决方案。

### 关键词

国际中文教育, HSK词汇, 增强现实技术, 多模态学习, 纸质与AR融合

# Design and Research of an Extroverted AR Vocabulary Learning Tool from the Perspective of Multimodality

## —Practice Based on the HSK Lexicon

Siyu Qi<sup>1</sup>, Xiangbin Zeng<sup>1</sup>, Renjie Lan<sup>1</sup>, Ziyuan Zhu<sup>1</sup>, Zheng Wu<sup>2</sup>, Jing Liu<sup>3</sup>

<sup>1</sup>School of Management, Tianjin University of Technology, Tianjin

<sup>2</sup>School of Languages and Culture, Tianjin University of Technology, Tianjin

<sup>3</sup>School of Art & Design, Tianjin University of Technology, Tianjin

文章引用: 祁思宇, 曾祥彬, 兰仁杰, 朱梓源, 吴铮, 刘晶. 多模态视域下外向型 AR 词汇学习工具的设计与研究[J]. 现代语言学, 2026, 14(2): 134-139. DOI: 10.12677/ml.2026.142123

Received: December 15, 2025; accepted: January 19, 2026; published: January 29, 2026

## Abstract

Focusing on the difficulties in the integrated “form-pronunciation-meaning” learning of international Chinese vocabulary, this study develops “paper + AR” extroverted vocabulary learning cards based on the multimodal learning theory, and constructs a four-level system of “content planning-carrier selection-technical support-function realization”. Situational scenarios are reconstructed according to the frequency of HSK Level 1 to Level 6 vocabulary, with QR codes on paper carriers serving as low-cost access channels. A self-developed lightweight AR engine is adopted to realize cross-platform recognition and integration of 3D solid models, Chinese character stroke animation, contextualized example sentences, and gesture-controlled interaction functions, thus forming a four-channel collaborative mode. Taking the verb “open” as an example, a four-step learning process is designed. Verification results indicate that the tool can achieve synchronous “visual-auditory-kinesthetic” linkage in a single interaction, effectively reduce learners’ cognitive load, enhance vocabulary memory efficiency and knowledge transfer ability, and provide a replicable and low-cost digital solution for international Chinese vocabulary teaching.

## Keywords

International Chinese Education, HSK Vocabulary, Augmented Reality (AR) Technology, Multimodal Learning, Integration of Paper and AR

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

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



Open Access

## 1. 引言

近些年国际中文教育在全球范围内不断发展，随着中国国际影响力的逐步增强与经济交流的日渐频繁，“汉语热”展现出明显的升温趋势，越来越多的学习者把汉语当作重要的第二语言，各国教育系统也逐步引入中文课程。我国通过创办孔子学院、推进中外语言合作项目等一系列措施主动支持中文的国际传播，但在这一繁荣局面背后中文作为第二语言的词汇教学依旧面临不少现实难题，当前广泛应用的词汇教学工具——像传统词卡与多数电子应用程序，普遍存在模态类型单一、语境支撑薄弱、交互体验欠缺等不足，难以达成词汇的深度掌握与长期记忆，特别是对母语为拼音文字的学习者而言，汉字特有的形 - 音 - 义关联关系形成了明显的认知压力，常规教学工具常常不能有效解决这一问题。

在这一现实背景下，增强现实(AR)技术的快速发展与多模态学习理论的广泛应用，为突破现有词汇教学困境提供了新的可能。

## 2. 理论支撑、技术支持及研究现状

多模态话语指运用听觉、视觉、触觉等多种感觉，通过语言、图像、声音、动作等多种手段和符号资源进行交际的现象[1]。这种多模态理论主张认知与意义建构依赖于多种感知通道的协同参与，其扎根社会符号学与认知心理学的核心观点为：人类学习不局限于单一语言符号，而是融合视觉、听觉、触觉、动觉及空间感知等各类模态接收、处理并内化信息。语言学习中词汇的扎实掌握不仅是牢记字形与发音，

更关键在于在丰富感官感受和情境交流中构建形、音、义三者深度联结。动词“打开”的理想学习状态为同步接触文字符号(视觉层面)、聆听标准发音(听觉层面)、观察具体开启动作(动觉与视觉结合)甚至在虚拟或现实场景中亲手完成“打开”操作(触觉与动觉协同)，这种多渠道同步刺激能明显加深记忆印记，助力知识在长时记忆中的储存与调用。

在多模态理论的实践应用探索中，袁志芳、郑艳群(2008)针对对外汉语教学中形容词的释义难题，提出了多媒体释义方法[2]。该方法通过整合文字、图像、音频等多模态资源，构建直观化的释义场景，帮助学习者准确把握词汇意义与用法，为多模态理论在对外汉语词汇教学中的落地提供了早期实践经验，其核心思路与本研究倡导的“多模态协同”理念相契合，均强调通过多元符号资源的整合降低词汇学习的认知难度。

增强现实(AR)技术为真正意义上的多模态学习提供了专属技术保障，可将三维动画、音频、交互式按钮等虚拟数字内容自然融入真实场景，打破现实与屏幕的界限，构建高度沉浸且互动性强的混合式学习环境。在此类环境中，语言不再是抽象孤立的符号，而是被纳入可视、可听且可操作的具体情境。蔡苏等(2017)通过对 AR 在教学中的应用案例进行系统评述发现[3]，AR 技术能够通过虚实融合的呈现方式，将抽象知识具象化，有效调动学习者的多感官参与，提升学习的主动性与沉浸感，其在语言教学领域的应用潜力已得到初步验证。以“西瓜”词卡学习为例，学习者运用 AR 工具扫描词卡时，屏幕不仅会显示三维可旋转的西瓜虚拟模型以强化视觉认知，还能同步播放标准读音以实现听觉信息输入，同时支持通过触屏操作激活“切西瓜”互动动画以获取触觉与动觉体验，此类多模态情境化信息输入可显著增强语境真实性及学习者的现场体验与参与积极性，有效减轻认知负担，对破解汉字学习中“形 - 音 - 义”的认知难题至关重要。

近年来，国际中文智慧教学成为研究热点，相关技术与教学模式的融合探索不断深化。马瑞棱等(2024)基于实践研究深化了对国际中文智慧教学的认识[4]，指出智慧教学工具的核心价值在于实现“教 - 学 - 评”的协同联动，通过技术赋能打破传统教学的时空局限，构建个性化、沉浸式的学习生态，这与本研究“课内外联动”“虚实融合”的设计理念高度一致，为 AR 词汇学习工具与正式教学流程的衔接提供了理论支撑。曹钢等(2023)则基于《国际中文教育中文水平等级标准》，构建了词汇知识图谱与词汇自适应学习平台[5]，强调依据词汇等级与学习者能力特征进行个性化内容推送，其对 HSK 词汇的结构化处理思路与本研究依托 HSK 1~6 级词库进行内容规划的实践形成了呼应，印证了基于标准化词库进行数字化教学工具开发的可行性与科学性。

目前将 AR 应用于国际中文教学的研究成果并不多见，曲喆(2025)调查了国内外近五年(2021~2025)的研究成果发现[6]，国外将 AR 应用于中文语言教学研究的文章共有 44 篇(其中包括把中文作为第一语言和第二语言教学的)，涵盖词汇、口语、汉字等教学领域；国内相关研究有 47 篇，但主要集中于理论探讨和综述研究。现有研究中，尚未出现将 AR 技术与 HSK 分级字词教学深度融合的实践探索。李卢艳(2019)的研究虽涉及多模态理论在 HSK 中高级词汇教学中的应用[7]，但未引入 AR 技术实现多模态资源的沉浸式呈现。

综上，现有研究已通过多模态理论应用、AR 技术实践、智慧教学探索及 HSK 词汇数字化处理等方面探索，初步验证了相关理论与技术在国际中文教学中的应用价值，但在 HSK 分级词汇的多模态 AR 教学工具开发方面仍存在研究空白。本研究正是基于这一缺口，将 AR 技术与 HSK 分级词库相结合，构建多模态协同的词汇学习体系，旨在解决国际中文教育中词汇“形 - 音 - 义”融合学习的难点。研究基于多模态学习理论，设计并研发了一款结合“纸质卡片”与“增强现实(AR)技术”的外向型词汇学习工具，该工具以 HSK 1 至 6 级词库为内容基础，通过重构语境场景，将二维码作为低成本的 AR 技术接入点，研究团队自主研发的轻量化 AR 引擎能够整合 3D 模型、汉字笔顺动画、场景化例句和手势交互功

能。文章通过动词“打开”的四步学习流程作为实例，论证了该工具能够实现“视觉-听觉-动觉”的同步联动，从而减轻学习者的认知压力，提升记忆效率和知识迁移能力，为国际中文词汇教学提供一个可复制的、低成本的数字化解决方案，研究目标清晰，具有较强的针对性和现实意义。

### 3. “AR 词汇卡”的设计模型与多模态实现

#### (一) 核心设计理念

本设计核心思想聚焦“以人为本”“虚实融合”“课内外联动”三大方向：

1. “以人为本”：以学习者为核心导向、所有功能研发与体验升级围绕二语习得者的认知特点与情感诉求推进工具既关注词汇知识掌握程度，更注重学习过程中动力保持、自主能力培育及策略引导，力求通过友好且具激励性的交互模式减轻认知压力，增强学习主动性与成就感。

2. “虚实融合”：体现在传统纸质词卡与先进 AR 技术的结合运用。纸质媒介发挥基础认知与记忆支撑作用，保障使用便捷性与学习专注力；AR 技术负责拓宽认知边界，通过叠加数字化场景、多模态资源及交互任务，将抽象词汇具象化为可感知、可操作、可探究的立体形态，实现物理载体与虚拟内容的功能互补及体验提升。

3. “课内外联动”：致力于打破学习场景局限。课堂上它可成为教师开展分组活动、展示词汇用法的辅助教学工具，课堂外则成为学习者进行个性化复习、探索性自学及真实语境应用的便携帮手。借助教师端与学习者端的数据互通与任务协同构建连贯统一、相互支撑的混合式学习路径，实现教与学在时空上的无缝对接与高效协同。

#### (二) 系统架构设计

本项目关键在于打造层级清晰、技术整合、体验流畅的学习体系，该体系从上到下划分为内容层、载体层、技术层与功能层，各层级彼此依托、协同发力，达成多模态学习核心目标。

1. 内容层：HSK 词库的结构化构建与情境化设计。内容层作为整套系统的核心基础，其规划安排直接影响学习过程的系统性与实际效果，我们未对 HSK 词库做简单数字化转化，而是开展深度结构化梳理与情境化重塑工作。

1) 以词频为核心导向。严格依据 HSK 词汇的等级划分与词频统计数据，保障学习内容贴合学习者的语言能力及核心应用需求，初级词汇重点融入高频且基础的日常生活场景，中高级词汇则更多出现在复杂抽象的社会文化语境里。

2) 整合语言功能要素。对每个词汇的解析不止局限于词汇本身，还会纳入其常用的句法作用及搭配模式。以动词“建议”的学习为例，会同步展示其典型表达句式“我建议你……”，让词汇学习顺利衔接至语法与语用领域，达成“词-句-境”的有机融合。

2. 载体层：物理与数字接口的精细设计。作为连接实体空间与虚拟环境的核心衔接载体，载体层选用纸质词卡形式以达成“低技术门槛”与“高认知亲和”的双重诉求，词卡正面从上到下依次设置规范拼音(音韵信息输入)、核心词汇(字形核心)、英文解释(跨语言对应)、艾宾浩斯遗忘曲线图(元记忆管理)、NFC 音频触发区(即时听觉巩固)构成“音-形-义-时”四维度信息集合；词卡背面依照“提示-拓展-交互-触发”逻辑顺序，右上方设置高概括度插图(激活先验认知图式)，中部展示短语实例与图形化词汇延伸(句法-语义网络)，下方布设简答型互动问答(促使语言输出)，右下角设置二维码(AR 数字接入端口)。正背两面共用 NFC 感应区以确保“即触即听”的无缝使用体验，此版式设计借助“正-背”功能划分与“码-频”双通道接入方式达成纸质载体的低成本、可迭代的数字化升级，同时符合学习者对实体学习材料的熟悉度偏好，大幅降低技术接受阻力。

3. 技术层：AR 核心的无缝整合。作为整套系统动力核心的技术层，其核心构成是团队自主开发的

AR 应用工具，这款工具借助前沿计算机视觉技术可快速稳定辨识词卡图像或二维码达成精准三维定位功能(即 3D Registration，确保虚拟物体能稳定“置于”现实场景之中)。其底层搭载高性能数字资源管理体系，可依据识别结果实时调取并渲染相应的多媒体资源集合。该应用需适配市面上主流移动设备包括智能手机与平板电脑，力求实现低功耗且高流畅的渲染表现，避免技术卡顿破坏沉浸体验，为功能层的优质呈现筑牢稳固技术支撑。

4. 功能层：多模态沉浸式学习体验的落地。功能层是用户直接接触的模块，是多模态理论的实际应用载体，核心目标是充分调动学习者的各类感官路径。

1) 多模态内容呈现。视觉模态：构建多层次视觉输入体系，涵盖词卡自带的静态配图、AR 激活的高清晰度 3D 模型及动画演示(例如学习“地震”一词时能展示地壳运动的动态场景)，此外词卡正面还用不同颜色的笔画方向箭头辨识偏旁部首和其余组成部分，在静态状态下展示汉字结构与书写次序，切实减轻非汉字文化圈学习者“形”方面的认知压力。听觉模态：打造沉浸式听觉场景，所有词汇发音均由标准普通话播音员录制，保障输入的准确性，核心亮点是提供例句诵读把词汇融入完整句子之中增强语流感知与语感培养。动觉模态：打造针对性交互操作设计，学习者不再是被动的观赏者而是主动的参与主体，他们能够通过手势与 AR 模型展开互动，比如“点击”选项、“拖拽”归类、“缩放”查看细节、“旋转”观察整体形态，借助身体动作加深记忆印象。

2) 教师端支持系统。本项目着重凸显教师的重要作用，特意研发教师端管理平台，教师可通过该平台创建班级、管理学生信息、查看班级整体学习数据面板(例如词汇掌握比例、高频错误点)，还能灵活布置 AR 学习任务、组织依托 AR 资源的课堂小组活动，实现该工具与正式教学流程的无缝衔接，助力教师开展个性化教学工作与精准化教学干预。

5. 多模态学习路径设计——一套完整的学习流程需实现多种模态的有序配合与逐步递进具体路径如下(以“打开”为例)：

1) 视觉感知与预期构建。学习者查看纸质词卡获取“打开”的文字形态、拼音“dǎ kāi”及门类提示图片，构建初步语义猜想；

2) 实体触发与 AR 激活。先用手机触碰词卡 NFC 区域即时收听“dǎ kāi”的标准读音，完成听觉准备之后扫描二维码，虚拟抽屉便叠加在现实桌面之上，系统同步语音播报目标词汇、提示交互启动；

3) 多模态融合与意义确立。学习者在屏幕上点击抽屉把手，抽屉动画流畅展开，同步播放例句“请打开抽屉”。视觉(文字形态 + 动画)、听觉(读音 + 句子)、动觉(点击动作)在同一时间、区间内高度配合，让动作、语音与文字形成紧密关联，即时完成语义固定；

4) 输出巩固与迁移深化。界面发送提示音“请打开书”的学习任务，学习者要在真实场景里找到书并在 AR 画面中完成虚拟“打开”操作，实现从理解到应用的单步转化，用闭环方式强化学习成效。

## 4. 应用前景、挑战与反思

### (一) 预期应用成效与核心价值

此设计针对国际中文词汇教学的核心问题，应用价值体现在多个方面：学习动机与沉浸体验方面依托游戏化且交互性强的 AR 场景，这一工具彻底改变传统词汇学习的枯燥情况，学习者不再被动记诵词汇，转而成为虚拟场景的探索者与实际操作者，这种高强度的参与感受与趣味性能有效激发长期学习积极性，尤其适合青少年与初级水平成人学习者群体；词汇记忆与认知理解层面依据多模态编码理论，此设计为每个词汇构建起立体的认知系统；教学赋能方面支持教师灵活运用 AR 资源设计课堂任务、跟踪学习情况数据，达成更精准的教学指导与更高效的课堂管理，推动以学习者为核心的混合式教学模式发展。

## (二) 面临的难题

本项目应用前景可观，但在实际落地与推广阶段仍需应对诸多挑战，技术层面难题重点聚焦 AR 应用体验提升，需在不同配置移动设备上维持渲染流畅度与识别精准度并在复杂真实光照条件下保障虚拟模型高精度融合与自然展示，内容层面难题难度更高，为 HSK 数千个词汇分别打造既契合词汇意义又具备交互属性的优质 AR 场景及媒体素材是耗时费力的创造性任务且对团队内容策划、美术设计及技术实现能力有极高标准。

## (三) 未来研究推进路径

后续研究将从多维度向纵深方向推进，短期核心任务为重点开展严谨的实证探究并通过组建对照组开展教学实验收集词汇习得效率、记忆保持率及学习者反馈等维度的量化与质性资料，借助科学方法验证该工具的实际应用效能，中期计划着重开展人工智能(AI)技术的融合探索且可基于整合式学习分析技术依据学习者的互动数据及学习成绩记录实时调整 AR 内容的难度等级与展示序列，构建真正意义上的个性化自适应学习系统，从长期发展视角研判本研究搭建的“多模态-AR”框架拥有良好的拓展潜力，后续可尝试将其应用于成语、惯用语及抽象语法知识的教学场景，探索以可视化、情境化形式破解中文高阶语言学习的难点问题，最终建成覆盖各类语言要素的全方位综合型智能学习平台。

本研究成果目前集中于理论框架构建与系统原型设计，属于设计开发与初步探讨阶段。希望本研究能最终为国际中文教育的数字化、智能化转型升级贡献切实可行的路径与方案。

## 5. 结论

本项研究融合多模态理论与 AR 技术，设计开发面向 HSK 的纸质——AR 词汇学习卡，以“打开”为实例的四步学习路径验证显示四种通道的协同作用能在单次交互中实现“形 - 音 - 义”绑定，为分级词汇的情境化呈现与低成本数字化转化提供可借鉴模式，后续通过实证调研、AI 技术融合及框架拓展有望构建涵盖成语与语法知识的智能学习平台，为国际中文教学的数字化变革贡献可行方案。

## 基金项目

天津理工大学市级大学生创新创业训练计划资助项目——《启明 AR 汉字外向型 HSK 分级词汇卡》；项目编号 202510060039。

## 参考文献

- [1] 张德禄. 多模态话语分析综合理论框架探索[J]. 中国外语, 2009, 6(1): 24-30.
- [2] 袁志芳, 郑艳群. 面向对外汉语教学的形容词多媒体释义方法研究[C]//中文教学现代化学会. 数字化汉语教学进展与深化. 北京: 北京语言大学网络信息与教育技术中心, 北京语言大学对外汉语研究中心. 2008: 583-588.
- [3] 蔡苏, 张晗, 薛晓茹, 等. 增强现实(AR)在教学中的应用案例评述[J]. 中国电化教育, 2017(3): 1-9, 30.
- [4] 马瑞棱, 蔡建永, 曹钢, 等. 基于实践的国际中文智慧教学认识深化[J]. 天津师范大学学报(社会科学版), 2024(6): 9-16.
- [5] 曹钢, 董政, 徐娟. 基于《国际中文教育中文水平等级标准》的词汇知识图谱与词汇自适应学习平台构建[J]. 国际汉语教学研究, 2023(1): 21-30.
- [6] 曲喆. 增强现实(AR)在国际中文写作教学中的应用研究[D]: [硕士学位论文]. 天津: 天津理工大学, 2025.
- [7] 李卢艳. 基于多模态理论的对外汉语词汇教学设计——以 HSK 中高级词汇为例[D]: [硕士学位论文]. 杭州: 浙江科技学院, 2019.