

数字媒介下公共话语非正式语体的变迁机制

——基于“AI + 教育”议题的跨平台语料库分析

边新媛, 李维滨

聊城大学外国语学院, 山东 聊城

收稿日期: 2026年6月3日; 录用日期: 2026年6月26日; 发布日期: 2026年7月8日

摘要

本文基于Biber多维度语体分析框架, 整合计算机中介交流(CMC)理论与媒介情境理论, 构建包含14个指标的“非正式语体核心指标编码体系”。通过对比分析2024~2026年期间传统媒体与社交媒体共22万字的平行语料库, 研究发现: 社交媒体的非正式语体正向指标综合频率是传统媒体的1.61倍, 语体从客观陈述转向对话式表达, 情境依赖维度明显增强, 形成情境依赖型表达模式, 表明数字媒介推动公共话语向非正式语体变迁, 并对各指标的媒介效应与文类效应进行了审慎归因。研究进一步构建“公共话语语体变迁的完整因果链”, 揭示数字媒介通过情境融合、情感补偿、效率提升等机制推动非正式语体向公共领域延伸的作用路径。本研究为理解数字媒介下公共话语演变提供了实证证据与理论视角, 对媒体实践与语言规范建设具有启示意义。

关键词

数字媒介, 公共话语, 非正式语体, 语体变迁, 语料库分析, AI + 教育

The Evolutionary Mechanism of Informal Style in Public Discourse in the Digital Media Era

—A Cross-Platform Corpus Analysis on the “AI + Education” Issue

Xinyuan Bian, Weibin Li

School of Foreign Languages, Liaocheng University, Liaocheng Shandong

Received: June 3, 2026; accepted: June 26, 2026; published: July 8, 2026

Abstract

This study, grounded in Biber's multi-dimensional register analysis framework and integrating

Computer-Mediated Communication (CMC) theory with Media Contextualism theory, constructs a “Core Indicator Coding System for Informal Register” comprising 14 indicators. Through a comparative analysis of parallel corpora totaling 220,000 words from traditional media and social media during the period 2024~2026, the research reveals that the comprehensive frequency of positive indicators of the informal register in social media is 1.61 times that of traditional media. The register shifts from objective statement to conversational expression, with a marked enhancement in the context-dependence dimension, thereby forming a context-dependent expression pattern. These findings empirically demonstrate that digital media is driving the transformation of public discourse toward informal register. The study also carefully attributed the media effects and genre effects of each indicator. Furthermore, the study constructs a causal chain of “The Complete Causal Chain of Public Discourse Register Evolution”, elucidating the pathways through which digital media facilitate the extension of informal style into the public sphere via mechanisms such as context integration, emotional compensation, and efficiency enhancement. This research provides empirical evidence and theoretical perspectives for understanding the evolution of public discourse in the digital media era, offering significant implications for media practice and the construction of language norms.

Keywords

Digital Media, Public Discourse, Informal Register, Register Evolution, Corpus Analysis, AI + Education

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

《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》[1]明确提出了“加快数字化发展”“建设数字中国”的目标纲领。在这一背景下,人工智能(AI)与社交媒体的深度融合不仅改变了信息传播的方式,更深刻重塑了公共话语的表达形态。当专业议题进入社交媒体平台,传统正式语体与非正式表达之间的边界开始模糊,这种变迁机制成为数字时代语言研究的重要命题。

关于数字化视域下的语言研究,尤以数字媒介相关的话题,国外相关的研究多围绕四个方向:一是语言多样性与数字媒介使用差异研究[2];二是技术与语言的交互研究[3];三是数字语言教育与应用研究[4];四是数字语言变异与身份研究,如 Di Marco 揭示社交媒体语言简化模式反映了语言接触引发的系统化过程[5], Panjaitan 系统综述研究发现语言经济原则因追求速度和字符限制而成为语言简化的主要因素[6]。国内数字媒介相关研究也已百花齐放,如:戴曼纯阐释数字时代语言生活状况[7];王春辉、程京艳侧重数字语言生态建设[8]与政策研究[9];程润峰关注网络语言,提出网络语言社群化的新组织形式[10];庞俊彩研究数字化转型中媒体用语采纳机制[11];李战子等人则用文献计量法全面分析近十年核心期刊的社交媒体话语研究现状及热点趋势[12];李琳郁则分析互联网语言情感极化现象[13]。

国内学者在 Biber 框架的本土化应用方面已取得重要进展[14],如陈贝颖将 Biber 多维分析范式应用于外交汉语文本研究,提取出八个具有高度解释力的维度[15];牟建霖运用 Biber 多维分析法对比中国式现代化主题下政治新闻英译与原创英语的语域特征差异[16],这些研究为 Biber 框架在中文语境中的应用提供了宝贵经验。尽管已有研究关注 CMC 语言特征及 Biber 多维度语体分析框架,但将二者系统整合,

并聚焦公共议题的研究尚属空白。本文通过自建传统媒体与社交媒体对比语料库, 整合 Biber 多维度分析框架、媒介情境理论[17]与 CMC 理论[18], 构建包含 14 个指标的“非正式语体核心指标编码体系”, 旨在回答以下问题: 在“AI 对语言学习的影响”这一公共议题下, 数字媒介是否推动公共话语向非正式语体变迁? 其影响机制如何?

2. 研究方法

(一) 语料来源

本研究选取 2024~2026 年“AI 对语言学习的影响”相关语料, 构建传统媒体与社交媒体跨平台对比语料库, 总计约 22 万字。为有效控制“文类(Genre)”变量, 剥离媒介形态本身带来的语体差异, 本研究在传统媒体的语料采集中引入了官方互动区文本; 在社交媒体抽样中采取了分层随机抽样策略, 以确保样本的代表性与科学性。

传统媒体语料库(约 12 万字): 选取《科技日报》与《人民日报》作为代表, 兼顾科技议题的专业性与官方媒体的严肃规范性。语料分为两部分: 一是新闻报道正文(10 万字)。通过中国知网检索平台, 限定时间范围 2024~2026 年, 以“人工智能”为主题、“高等教育/语言学习”为交叉学科进行检索, 严格筛选并统一编码格式; 二是官方互动区评论(约 2 万字)。定向采集上述媒体在官方网站及微信公众号相关报道下方的读者评论区内容。此部分语料的引入, 旨在使传统媒体语料库不完全局限于“单向新闻正文”这一单一文类, 部分纳入互动性表达, 以缩小与社交媒体语料库在文类构成上的天然落差。

社交媒体语料库(约 10 万字): 选取小红书与知乎两大平台, 以覆盖“生活化经验分享”与“专业化知识问答”两种不同的数字媒介场景。为避免“高赞抽样”带来的头部效应偏差, 本研究采用分层随机抽样(Stratified Random Sampling)策略。首先, 以 2024、2025、2026 三个年份为时间分层节点, 分别在小红书检索“AI 语言学习”“AI 口语陪练”等话题, 在知乎检索“AI 对语言学习的影响”等问题。其次, 利用 Python 抓取清洗数据后, 在每个时间分层内, 对去重后的有效帖子与评论(剔除广告、无意义刷屏)进行随机抽取。最终, 小红书平台保留有效文本约 5 万字, 知乎平台保留约 5 万字。两类社交媒体语料合计 10 万字, 与传统媒体语料形成具有高度可比性的平行规模。

需要说明的是, 本文的比较单位是“媒介的整体话语生态”而非“单一文类”——即传统媒体平台上的综合话语实践(新闻正文 + 读者互动)与社交媒体平台上的综合话语实践(帖子 + 评论)之间的对比。这种设计反映的是不同媒介环境下公共话语的总体语体倾向, 而非某一特定文类的特征。下文分析中, 将对哪些指标更可能受文类影响、哪些更可能反映媒介属性进行逐一审辨。

(二) 非正式语体核心指标界定

本研究整合 Biber 多维度语体分析框架、CMC 理论与媒介情境理论, 构建包含 14 个指标的“非正式语体核心指标编码体系”, 涵盖四个维度, 见表 1。

(三) 数据统计与分析步骤

语料数据的清洗与分层管理依托 Python 环境(调用 pandas 库进行结构化处理)完成, 并利用 NLTK 等自然语言处理工具包进行初步的分词与词性赋码。最后计算每类指标每千字的出现频率。统计分析采用独立样本 t 检验比较传统媒体与社交媒体在各非正式语体指标上的差异, 显著性水平设定为 $\alpha=0.05$, 效应量采用 Cohen's d [19]进行评估。编码信度检验采用 Cohen Kappa 系数, 结果显示 $\text{kappa}=0.87$, 表明编码一致性良好。需要特别指出的是, 本研究表 2 中的前三个维度属于“非正式语体正向指标”, 而维度四(篇章衔接)属于“正式语体特征的逆向指标”, 即该数值越低, 表明文本碎片化程度越高、语体越倾向于非正式。

Table 1. Four-dimensional indicators of the informal style core metric coding system**表 1.** 非正式语体核心指标编码体系四维度指标

| 维度 | 指标 | 定义与语体功能说明 |
|----------------|------------|-----------------------------------|
| 维度一 互动性与信息性 | 第一/二人称 | 突破单向信息传递, 建立“对话者”身份, 拉近传受双方心理距离。 |
| | 情态动词、问句 | 激发读者思考与反馈, 模拟面对面的人际协商过程。 |
| | 感叹句、口语化句式 | 降低表达的严肃性, 增强公共话语的亲合力与即时互动感。 |
| 维度二 情境依赖度 | 语境补全型省略 | 依赖共享背景知识的零指代, 用于提升沟通效率。 |
| | 语义表达模糊 | 缺乏明确指称对象, 体现思维的碎片化特征。 |
| | 技术缩写(功能区分) | 区分专业交流功能与身份标签功能。 |
| 维度三 情感立场 | 表情符号 | 作为文本交流的“非语言补偿机制”, 弥补语音语调等情感线索的缺失。 |
| | 评价性形容词/副词 | 直接凸显主观态度与价值判断, 使语体从客观陈述转向主观评价。 |
| | 网络新词、立场标记 | 展现表达者的个性化情感色彩, 强化情绪感染力。 |
| 维度四 篇章衔接 | 逻辑连词密度 | 测量文本组织形式, 反映非正式语体在宏观逻辑上的严密程度。 |
| | 跨句指代呼应 | 衡量语篇连贯性, 反映碎片化表达对传统长语篇结构的解构。 |

3. 研究结果与分析

(一) 非正式语体核心指标的量化分布特征

通过对两类语料库进行统计, 研究发现社交媒体在“互动性、情境依赖、情感立场”三大非正式维度上的综合频率(61.43 次/千字)是传统媒体(38.27 次/千字)的 1.61 倍; 而在代表正式语体严密性的“篇章衔接”维度上, 社交媒体的频率仅为传统媒体的 50%。具体量化数据如表 2 所示:

Table 2. Comparison of informal style indicators frequency between two types of corpora (per 1000 characters)**表 2.** 两类语料库非正式语体指标频率对比(次/千字)

| 维度 | 指标 | 传统媒体平均 | 社交媒体平均 |
|----------------|-----------------|--------|--------|
| 维度一 互动性与信息性 | 人称代词(第一/二人称) | 3.65 | 13.72 |
| | 情态动词 | 27.90 | 28.86 |
| | 问句 | 1.55 | 2.07 |
| | 感叹句 | 0.10 | 1.24 |
| | 口语化句式 | 0.07 | 0.92 |
| | 小计 | 33.27 | 46.81 |
| 维度二 情境依赖度 | 语境补全型省略 | 1.20 | 8.66 |
| | 语义表达模糊 | 1.90 | 1.79 |
| | 技术缩写 | 0.45 | 1.41 |
| | 小计 | 3.55 | 11.86 |
| 维度三 情感立场 | 表情符号(情感型 + 互动型) | 0.00 | 0.26 |
| | 评价性形容词/副词 | 1.34 | 1.79 |
| | 网络新词 | 0.02 | 0.10 |
| | 立场标记 | 0.09 | 0.61 |
| | 维度三小计 | 1.45 | 2.76 |

续表

| | | | |
|--------------|--------|-------|-------|
| 维度四: 篇章衔接 | 逻辑连词密度 | 12.50 | 6.30 |
| | 跨句指代呼应 | 8.20 | 4.10 |
| | 维度四小计 | 20.70 | 10.40 |

基于 Biber 多维度语体分析框架, 本研究揭示了在“AI + 教育”这一科技教育类公共议题下, 社交媒体相比传统媒体的语体呈现三大特征:

1、互动性维度显著跃升, 语体从客观陈述向对话式表达转变。社交媒体在第一/二人称代词的使用上表现出极端的倾斜(13.72 对 3.65)。传统媒体为维持严肃性, 极力弱化作者与读者的主体显现; 而社交媒体则通过高频的“你”“我”“大家”, 将公共议题的探讨转化为拟态的“面对面人际协商”。感叹句(1.24 对 0.10)的大量引入, 则为这种对话机制注入了即时互动的情绪价值。语料显示, 小红书用户频繁使用“家人们”“宝子们”“姐妹们”等群体化称谓, 知乎用户则多采用“我”“你”的直接互动, 而传统媒体几乎完全避免第一、二人称使用, 仅以第三人称客观陈述, 例如: “国家应加快推动 AI 进课堂的政策落地”“关键在于教师信息素养的提升”等以第三人称或无主语句式呈现的政策性讨论。这一发现与 Meyrowitz 媒介情境理论的核心观点相吻合[17], 同时契合 CMC 理论的“情感补偿机制”: 计算机中介交流的文本属性缺失语音语调等情感线索, 用户通过人称代词等符号化互动标记弥补情感传递的不足。从文类敏感度来看, 人称代词和口语化句式的跨平台差异可能同时受到媒介属性与文类规范的双重驱动: 传统媒体新闻正文的“去人称化”和“书面化”本身是该文类的结构要求, 社交媒体帖评混合体的“对话式”表达亦与互动评论这一文类的交际功能天然相关。因此, 人称代词差异中相当部分可能源于文类固有属性, 不宜完全归因于媒介变迁。相比之下, 感叹句的使用在两种媒介的评论场景中均属非必须成分, 其跨平台差异更可能反映媒介即时交互属性对表达情绪化的独立驱动。

感叹句与口语化句式的差异同样凸显社交媒体的互动导向。小红书中的“太牛了!”“绝了!”“真心推荐!”等感叹句式, 通过情绪强化实现广泛共情传播。小红书中的“靠谱”“拿捏了”“稳了”等网络口语化表达, 降低了表达门槛, 增强了亲切感。

2、情境依赖度内部发生结构性重组: 语境补全成为沟通核心机制。细化后的数据显示, 社交媒体在“情境依赖”上的高分, 主要并非由于逻辑混乱导致的“语义表达模糊”(两类媒介在此指标上差异不显著, 分别为 1.90 与 1.79), 而是源于极高频的“语境补全型省略”(8.66 对 1.20)。在小红书等平台, “这个好用, 快冲”等零指代现象, 高度依赖平台社群共享的背景知识。这表明, 数字媒介将公共话语从“高语境跨越”转变为“低语境圈层内交流”, 以降低语法完整性换取了极高的沟通效率。而传统媒体评论区即便在互动场景中, 仍倾向于补全指代信息。例如: AI 辅助教学可以提高效率, 但教育部门要加强监管, 防止学生过度依赖人工智能工具; 大模型在语言教学中前景广阔, 但需要更多实证研究验证是否长远有效。这表明, 即便脱离了新闻正文的文类约束进入互动评论场景, 传统媒体读者仍倾向于遵循“陌生人传播”的精确性规范, 而非社交媒体式的“类社区传播”中依赖共享背景知识的高效省略。

此外, 技术缩写在两类媒介中展现出截然不同的使用场景和语用功能。语料显示, 科技日报与小红书均高频使用“GPT”“AI”“NLP”等缩写, 但使用情境和语用动机存在差异: 传统媒体多在技术阐释场景使用(如“GPT 模型在语言生成方面的优势”), 旨在追求科技报道的精确与简洁, 履行专业交流功能。例如: 科技日报评论区中“GPT 在 NLP 领域的突破值得持续关注, 但应警惕其生成内容的可靠性问题”等评论, 缩写仍嵌套在论证结构中, 服务于对技术本身的评价与讨论; 社交媒体则在经验分享场景或个人成就叙事时使用(如“用 GPT 练口语, 三个月从四级压线过到六级 550+”), 技术缩写常与“高效”“上岸”等词汇共现, 成为用户构建“数字化生存能力”和“圈层认同”的符号工具, 更多被用作身份标签。用户

借此构建圈层认同, 这是一种典型的社会语言学维度的变异。两类评论区的缩写使用场景差异, 进一步印证了技术缩写的语用功能并非由词汇本身决定, 而是受媒介场景与交际目标的共同塑造。

社交媒体中高频出现的指代现象, 84%属于语境补全型省略。这种“不言自明”的表达模式, 反映了数字媒介下公共话语从“陌生人传播”向“类社区传播”的转变, 语体功能从单纯的信息传递转向了情感联结。语料显示, 小红书中的“这个 APP 真好用, 姐妹们快去冲”“试了三个, 这个最稳”等高度情境依赖的表达大量涌现, 而传统媒体则几乎无此类表述。这一现象印证了媒介情境理论的核心命题: 社交媒体的即时性、对话性特征降低了表达的精确性要求, 用户依赖共享的情境背景知识进行沟通, 形成“情境依赖型表达”范式。

文类敏感度方面, 语境补全型省略的高频出现与社交媒体的“帖子 + 评论”混合文类有天然亲缘性——评论本身就是对前文语境的承接与回应, 零指代在该文类中具有功能合理性。但需指出的是, 即便在传统媒体评论区(同样具有回应性文类特征), 零指代的使用频率仍远低于社交媒体评论区, 这提示媒介的即时交互属性对语境补全型省略具有独立于文类的增量效应。技术缩写的功能分化则较少受文类影响, 更直接地反映了不同媒介场景中用户的交际目标与身份建构需求。

3、情感立场极化, 抽象价值判断向具象经验分享迁移。社交媒体的情感立场维度数值接近传统媒体的两倍。传统媒体依然依赖评价性形容词(如“革命性意义”“深远影响”)来建构议题的专业公信力。例如: 人民日报微信公号读者评论中可见“人工智能与教育深度融合是大势所趋, 对提升教育公平和质量具有深远意义”“这项政策将推动教育数字化迈上新台阶”等抽象价值判断型评价, 较少出现具象化的个人经验数据。反观社交媒体, 用户大量调用表情符号作为“非语言补偿机制”, 且评价表达呈现出强烈的具象化特征(如“听力从 15 分提到 25 分(点赞)”), 这种量化的经验分享模式弱化了传统宏大叙事的距离感。

评价性形容词的使用差异在很大程度上受文类规范制约: 新闻正文以客观陈述为标准, 自然弱化主观评价; 而社交媒体的帖子与评论均以态度表达为交际核心。表情符号的跨平台差异则更可能反映媒介属性的影响——表情符号作为 CMC 环境下的非语言补偿机制, 其使用与屏幕交互的媒介条件直接相关, 文类因素的解释力相对有限。

4、篇章衔接断裂, 逻辑链条让位于碎片化呈现。维度四的数据直观展现了数字媒介对长篇幅逻辑的解构。传统媒体的逻辑连词密度(12.50)和跨句指代呼应(8.20)均远高于社交媒体。传统媒体评论区虽为互动性文本, 但仍保留了较高的篇章衔接度。例如, 科技日报评论区中“虽然 AI 在知识传授方面具有效率优势, 但是培养学生的批判性思维仍需教师引导, 因此技术应服务于育人而非替代育人”这样的评论, “虽然……但是……因此……”的逻辑链条完整, 论证结构清晰。而社交媒体评论区的典型表达如“AI 讲得太好了, 颠覆认知, 必须安利”, 则呈现为碎片化的情绪宣泄, 缺乏显性逻辑词衔接。在社交媒介环境中, 受限于屏幕阅读习惯与注意力机制, 长复句和严密的因果论证被替代。文本呈现出“去连接化”的特征, 段落之间不再依赖显性的逻辑词(如“因此”“综上所述”)进行硬链接, 而是依赖读者的联想与认知跳跃完成意义缝合, 呈现了向非正式网络语体的形态跨越。

篇章衔接指标的跨平台差异具有最高的文类敏感度。新闻正文作为结构完整的独立文本, 天然要求高密度的逻辑连词和跨句指代; 而社交媒体中评论占比较大, 评论多为碎片化的短回应, 文类本身即不要求严密的篇章衔接。因此, 维度四的差异中文类效应的占比可能高于其他维度, 不宜简单解读为“媒介导致逻辑退化”, 更准确的表述为: 媒介环境中文类结构的重组(从以长文本为主转向以短互动为主)导致了整体篇章衔接密度的下降。

(二) 社交媒体推动非正式语体变迁的关联机制

为揭示社交媒体推动非正式语体变迁的作用机制, 本研究构建“媒介属性→表达规范→用户策略→语体特征”的因果链, 整合 CMC 理论与媒介情境理论的核心命题。

1、媒介属性驱动。社交媒体的低准入门槛、即时反馈、碎片化传播、去中心化互动等特征, 与传统媒体的严肃性、专业性、单向性形成对比。这些属性为非正式语体延伸提供了技术条件和环境基础。

2、表达规范调整。社交媒体打破公私场景边界(媒介情境理论), 允许私人领域的非正式表达进入公共讨论; 同时强调互动导向, 将公共讨论从“信息传递模式”转向“互动对话模式”。

3、用户表达策略。为适应媒介环境, 用户采取三项核心策略: 1) 情感补偿(CMC 理论): 通过表情符号、感叹句弥补文本交流的情感缺失; 2) 效率提升: 使用缩写词、省略句式简化表达; 3) 互动强化: 频繁使用人称代词、问句激发回应。

4、语体特征显现。上述策略集体实践呈现为可观测语体: 前三维度指标综合频率是传统媒体的 1.61 倍; 其中人称代词、感叹句、情感型表情符号等指标显著提升, 验证了因果链的有效性。

(三) 媒介效应与文类效应的审慎归因

本研究比较的是传统媒体与社交媒体的综合话语实践, 而非单一文类, 因此所观察到的语体差异中不可避免地混合了媒介属性效应与文类结构效应。基于上文对各指标文类敏感度的逐一分析, 本文将 14 项指标按“媒介驱动为主”“媒介与文类交互”“文类驱动为主”三个层级进行归因区分(见表 3):

Table 3. Three-level attribution table

表 3. 三层级归因表

| | | |
|---------|------------------------------|--|
| 媒介驱动为主 | 表情符号、技术缩写 (功能分化) | 与屏幕交互条件和 CMC 补偿机制直接相关, 文类解释力有限 |
| 媒介与文类交互 | 语境补全型省略、 感叹句、立场标记 | 媒介即时性有独立增量效应, 但文类功能需求亦构成必要条件 |
| 文类驱动为主 | 人称代词、口语化句式、逻辑 连词密度、跨句指代呼应 | 新闻正文 vs 互动评论的文类规范本身即要求不同, 媒介属性的作用为间接的 |

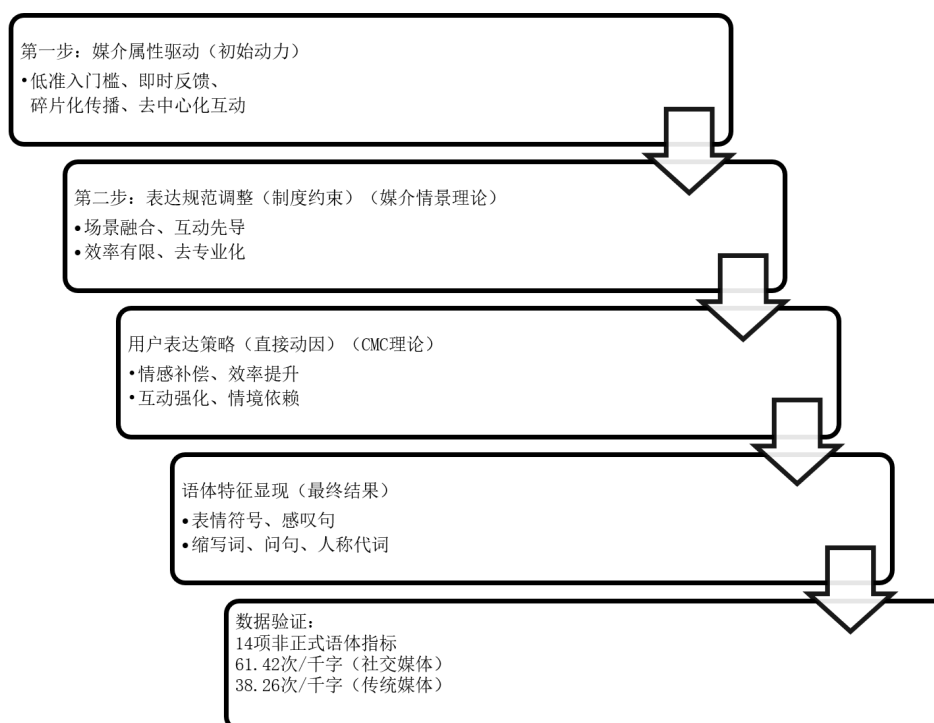


Figure 1. The complete causal chain of the transformation of public discourse style

图 1. 公共话语语体变迁的完整因果链

这一区分并不意味着“文类驱动为主”的指标与媒介变迁无关——文类结构本身的重组(传统媒体以新闻正文为主体, 社交媒体以互动评论为主体)恰恰是媒介变迁的产物。但严格而言, 这些指标的跨平台差异反映的是“媒介改变了公共话语的文类构成”, 而非“媒介直接改变了同一文类的表达方式”。本文的核心结论——数字媒介推动公共话语向非正式语体变迁——在“媒介驱动为主”的指标上获得了最直接的证据支撑, 在“交互”和“文类驱动”层级的指标上则需附加上述条件性理解, 见图 1。

4. 结论与讨论

本研究聚焦“AI+教育”议题, 通过对比传统媒体与社交媒体语料, 揭示社交媒体对公共话语非正式语体的具体影响。核心发现包括: 数字媒介加速了公共话语的非正式化趋势; 评价性表达呈现平台分化现象, 传统媒体通过抽象价值判断建构议题严肃性, 社交媒体通过量化经验分享传递具体态度; 语境补充型省略的显著增长, 体现了情境依赖型表达范式的形成。

本研究的理论贡献在于, 整合 Biber 多维度语体分析框架、CMC 理论与媒介情境理论, 构建“描述-解释-机制”三层次分析模型以及公共话语语体变迁完整因果链, 系统科学地佐证解释了数字媒介推动公共话语向非正式语体变迁; 方法贡献在于, 构建包含 14 个指标的“非正式语体核心指标编码体系”, 实现非正式语体研究的可操作化与系统化, 为后续研究提供了可复制的分析工具。

本研究存在一定局限: 语料来源仅选取四大平台, 覆盖面有待拓展; 时间跨度限定为 2024~2026 年, 时效性有待延伸; 研究仅聚焦单一议题, 结论的普适性需进一步验证。未来研究可拓展至 AI 医疗、AI 职场等多元议题, 建立长期追踪机制, 深入探究非正式语体对语言表达规范、公共讨论质量及用户语言能力的长期影响, 为 AI 时代的语言规范与教育策略提供前瞻性建议。必须指出的是, 本研究比较的是不同媒介平台上的综合话语生态, 所观察到的非正式化倾向中混合了媒介属性效应与文类结构效应。本文已通过各项指标的文类敏感度逐一辨析, 区分了“媒介驱动为主”“媒介与文类交互”“文类驱动为主”三个归因层级。未来研究可进一步采用同文类跨媒介对照设计, 以更精确地量化媒介变量与文类变量的主效应及交互效应。

参考文献

- [1] 中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要[EB/OL]. https://www.gov.cn/xinwen/2021-03/13/content_5592681.htm, 2026-03-22.
- [2] Helm, P., Bella, G., Koch, G. and Giunchiglia, F. (2024) Diversity and Language Technology: How Language Modeling Bias Causes Epistemic Injustice. *Ethics and Information Technology*, **26**, Article No. 8. <https://doi.org/10.1007/s10676-023-09742-6>
- [3] Alvero, A.J., Sedlacek, Q., Leon, M. and Peña, C. (2025) Digital Accents, Homogeneity-by-Design, and the Evolving Social Science of Written Language. *Annual Review of Applied Linguistics*, **45**, 50-68. <https://doi.org/10.1017/s0267190525000042>
- [4] Lee, S., Jeon, J., McKinley, J. and Rose, H. (2025) Generative AI and English Language Teaching: A Global Englishes Perspective. *Annual Review of Applied Linguistics*, **45**, 85-108. <https://doi.org/10.1017/s0267190525100184>
- [5] Di Marco, N., Loru, E., Bonetti, A., Serra, A.O.G., Cinelli, M. and Quattrociocchi, W. (2024) Patterns of Linguistic Simplification on Social Media Platforms over Time. *Proceedings of the National Academy of Sciences*, **121**, 15-17. <https://doi.org/10.1073/pnas.2412105121>
- [6] Panjaitan, L.L. and Patria, A.N. (2024) Social Media and Language Evolution: The Impact of Digital Communication on Language Change. *International Journal of Linguistics, Literature and Translation*, **7**, 53-57. <https://doi.org/10.32996/ijllt.2024.7.12.8>
- [7] 戴曼纯. 数字时代的语言生活及其研究[J]. 外国语(上海外国语大学学报), 2024, 47(1): 17-25.
- [8] 王春辉. 语言数据安全论[J]. 语言战略研究, 2022, 7(4): 15-25.
- [9] 程京艳. 数字时代语言失范的界定研究[J]. 外国语(上海外国语大学学报), 2025, 48(4): 62-69.

- [10] 程润峰, 谢晓明. 论网络语言的社群化[J]. 语言战略研究, 2022, 7(3): 23-33.
- [11] 庞俊彩. 语言接触区的数字重构: 缅甸语社交媒体用语中外来词的采纳机制与文化认同研究[J]. 中国科技术语, 2025, 27(5): 123-127.
- [12] 李战子, 韩泽婷. 近十年 CSSCI 期刊社交媒体话语研究现状及热点趋势分析[J]. 外国语(上海外国语大学学报), 2025, 48(5): 102-111.
- [13] 李琳郁. 互联网语言通货膨胀现象的表征与动力机制[J]. 新媒体研究, 2025, 11(13): 70-74+79.
- [14] Biber, D. (1991) *Variation across Speech and Writing*. Cambridge University Press.
- [15] 陈贝颖. 外交汉语文本语言特征差异的多维度分析[D]: [硕士学位论文]. 厦门: 厦门大学, 2023.
- [16] 牟建霖. 中国政治新闻英译的语域特征多维分析——以中国式现代化主题新闻为例[D]: [硕士学位论文]. 北京: 北京外国语大学, 2025.
- [17] Meyrowitz, J. (1985) *No Sense of Place: The Impact of Electronic Media on Social Behavior*. Oxford University Press.
- [18] Baron, N.S. (2008) *Always On: Language in an Online and Mobile World*. Oxford University Press.
- [19] Cohen, J. (2013) *Statistical Power Analysis for the Behavioral Sciences*. Routledge.