

社会语言学视角下的网络流行语与青年身份建构研究

王文博, 张 洁*

新疆师范大学中国语言文学学院, 新疆 乌鲁木齐

收稿日期: 2026年8月15日; 录用日期: 2026年9月11日; 发布日期: 2026年9月20日

摘 要

本文从社会语言学视角梳理网络流行语与青年身份建构研究, 重点考察语言变异、语言态度、群体认同及数字媒介情境之间的关联。既有研究表明, 网络流行语并非孤立的词汇创新, 而是青年在平台互动中表达情绪、协商关系、识别圈层并建构自我位置的符号资源。当前研究已形成语言学、传播学、社会学等多学科交叉格局, 但仍存在概念边界不够清晰、社会语言学理论整合不足、样本与平台覆盖有限、量化分析与情境解释相互脱节以及青年主体经验呈现不足等问题。针对这些不足, 本文提出在拓宽数据来源的基础上建立“计算分析-质性解释-社会语言学验证”的混合研究路径: 利用主题建模识别大规模语料中的议题结构与语义聚类, 借助情感分析描写流行语在不同事件、平台与群体中的情绪倾向及其变化, 再通过网络民族志进入具体网络社群, 结合话语分析解释流行语在互动位置、身份指示、群体边界与意义协商中的作用。该路径可在宏观模式识别与微观语境阐释之间建立联系, 从而更全面地理解网络流行语的社会意义及青年身份建构机制。

关键词

网络流行语, 社会语言学, 青年身份建构, 主题建模, 网络民族志

Internet Buzzwords and Youth Identity Construction from a Sociolinguistic Perspective

Wenbo Wang, Jie Zhang*

School of Chinese Language and Literature, Xinjiang Normal University, Urumqi Xinjiang

Received: August 15, 2026; accepted: September 11, 2026; published: September 20, 2026

*通讯作者。

文章引用: 王文博, 张洁. 社会语言学视角下的网络流行语与青年身份建构研究[J]. 新闻传播科学, 2026, 14(9): 170-176. DOI: 10.12677/jc.2026.149250

Abstract

This paper reviews research on Internet buzzwords and youth identity construction from a sociolinguistic perspective, focusing on the connections among language variation, language attitudes, group identity, and digital-media contexts. Existing studies suggest that Internet buzzwords are not merely lexical innovations; they also function as symbolic resources through which young people express affect, negotiate relationships, recognize social circles, and position themselves in platform-based interaction. Despite increasing interdisciplinary attention, current research still faces several limitations, including blurred conceptual boundaries, insufficient integration of sociolinguistic theory, restricted platform and sample coverage, a separation between quantitative patterns and contextual interpretation, and limited representation of youth agency. To address these issues, the paper proposes a mixed-method framework of “computational analysis-qualitative interpretation-sociolinguistic validation.” Topic modeling can identify thematic structures and semantic clusters in large-scale corpora; sentiment analysis can trace affective tendencies across events, platforms, and groups; netnography can situate linguistic practices within online communities; and discourse analysis can explain how buzzwords index identities, negotiate meanings, and construct group boundaries. Combining macro-level pattern detection with micro-level contextual interpretation can provide a more comprehensive account of the social meanings of Internet buzzwords and the mechanisms of youth identity construction.

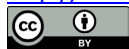
Keywords

Internet Buzzwords, Sociolinguistics, Youth Identity Construction, Topic Modeling, Netnography

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. 研究背景及意义

1.1. 研究背景

互联网与移动社交媒体改变了语言产生、扩散和更新的速度，也改变了青年参与公共表达和日常互动的方式。网络流行语通常在具体事件、平台机制和群体互动中生成，并通过转发、模仿、改写和再语境化不断获得新的含义。王仕勇对我国网络流行语研究的梳理显示，网络热词往往与特定时期的社会生活和媒介事件相伴而生，相关研究也随着网络传播形态的发展持续扩展[1]。从这一意义上说，网络流行语既是词汇层面的语言现象，也是观察数字社会中语言与社会关系变化的重要入口。

国际网络语言研究较早从计算机媒介交流切入。Crystal 对电子邮件、聊天室、网络文本等多种互联网语言形态的讨论表明，数字媒介会形成具有自身互动条件和表达特征的语言变体，网络空间同时为个体和群体提供新的身份呈现资源[2]。与传统书面语相比，网络交流具有即时性、互动性、可复制性和跨模态性，文字、缩写、表情、语音、视频以及平台符号可以共同参与意义建构。因此，网络流行语的形成不能仅从词义变化或构词方式解释，还需要考察它在何种平台、何种互动关系和何种社会情境中被谁使用、如何被理解。

在国内语境中，网络流行语的传播常表现出明显的事件性和群体性。方毅华、罗鹏对年度网络流行语编码规律的研究指出，网络热词与社会事件、传播情境和网民参与密切相关[3]。康忠德从社会语言学

角度将网络语言视为特定言语群体中的语言变体, 强调网络语言的社会属性及群体差异[4]。这些研究提示, 流行语并不是脱离使用者而独立存在的静态词汇, 而是在互动过程中不断被赋义的社会符号。青年既是网络流行语的重要使用者, 也是其再创造、再传播和意义更新的重要参与者。

因此, 本文采用描述性而非预设价值判断的立场, 把网络流行语视为数字媒介环境中的动态语言实践。研究关注的核心不是判断某一表达“好”或“坏”, 而是解释其形式如何产生、意义如何变化、使用者如何借助它表达立场, 以及这些语言选择如何与青年自我认同和群体关系发生联系。这样的研究立场更符合社会语言学关注“语言变异-社会因素-使用情境”关系的基本取向, 也有助于避免将复杂的网络语言现象简单归结为规范问题。

1.2. 研究意义

网络流行语具有较强的社会指示功能。杨萍将其概括为网民自主话语生产形成的文化景观, 指出网络表达能够集中呈现特定时期的社会经验和群体情绪[5]。对于青年而言, 使用某一流行语不仅是选择一个词, 还可能意味着选择一种语气、立场和互动身份。同一表达在熟人群体、兴趣圈层和公共讨论中的意义并不完全相同, 其幽默、反讽、亲近、疏离等语用效果往往依赖共同背景知识。

从社会语言学角度研究这一现象, 具有三方面意义。第一, 可将网络流行语的形式创新与社会变量联系起来, 考察年龄、平台、群体关系、交往场景等因素如何影响语言选择。第二, 可从身份建构角度解释青年为何主动采用某些标签化、戏仿化或缩略化表达, 并分析这些表达如何形成圈层识别和群体归属。第三, 可推动研究方法由单一文本举例转向多源数据与混合方法, 在保持语境解释力的同时提高研究的可验证性。汤玫英对网络语言形式及其变化的系统讨论, 为理解网络语言的多样性提供了基础[6]; 在此基础上, 将社会语言学与数字研究方法结合, 有助于进一步解释语言形式背后的社会机制。

2. 国内外研究现状

2.1. 国外研究现状及发展趋势

国外相关研究最初多围绕计算机媒介交流展开, 关注电子邮件、聊天室、论坛和网络社区中的语言特征。Crystal 的研究从媒介条件出发讨论网络语言的词汇、拼写、互动方式及身份表达, 说明互联网并非只是传统语言的传输渠道, 而会促成新的语言使用方式[2]。随着社交媒体兴起, 研究对象进一步扩展到社交网络、短文本、网络迷因、标签和多模态内容, 研究重点也从“网络语言有哪些特征”逐渐转向“语言实践如何嵌入数字社群并参与身份建构”。

在方法上, 国外数字语言研究呈现出计算方法与数字民族志并行发展的趋势。一方面, 大规模网络文本为统计建模提供了条件。Blei、Ng 和 Jordan 提出的潜在狄利克雷分配模型(LDA)能够从大量文档中识别潜在主题结构, 为分析网络语料中的议题聚类、主题演变和共现关系提供了经典方法基础[7]。另一方面, Pang 和 Lee 系统总结的情感分析研究表明, 计算方法可以对文本中的意见、情绪倾向和主观性进行识别与分类[8]。这些方法使研究者能够处理传统人工阅读难以覆盖的大规模语料, 但其结果仍需要结合具体语境解释, 尤其是面对反讽、谐音、隐喻、圈层黑话和语义快速漂移时, 单纯依赖自动分类容易损失社会意义。

与计算路径相对应, 网络民族志强调进入网络社群的互动环境理解文化实践。Kozinets 将网络民族志发展为研究在线社群与数字文化的重要质性方法, 强调研究者需要关注社群规范、参与方式、互动历史以及线上与线下经验之间的联系[9]。这一思路对于网络流行语研究尤为重要, 因为某些表达只有放回具体帖子、评论链、弹幕互动或兴趣社群中, 才能准确理解其指向对象、语气和身份功能。由此可见, 国外研究的发展趋势不是简单地以计算方法替代质性研究, 而是越来越重视不同方法之间的互补。

2.2. 国内研究现状及发展趋势

国内网络流行语研究在 2000 年前后逐步展开, 随后伴随 Web2.0 和社交媒体发展迅速增长。现有成果主要来自语言学、传播学、社会学、教育学和青年研究等领域。早期语言学研究较多关注构词、谐音、缩略、语义迁移和语用特点; 传播学研究重视热点形成、媒介扩散和话题参与; 社会学与青年研究则更多关注流行语所折射的社会心态、代际经验与群体文化。多学科进入使研究对象不断丰富, 但也带来概念口径和分析框架不完全一致的问题。

在语言变异研究方面, 学者通常从形式构成和语义变化描述网络流行语。网络表达可通过缩写、数字替代、谐音、旧词新义、语码混合、格式仿造等方式生成。其共同特点并非“不规范”, 而是对特定媒介条件和互动效率的适应: 短文本环境鼓励压缩表达, 算法推荐加速模仿扩散, 评论区和弹幕的高同步互动又强化了固定句式和模板化表达。社会语言学更关心的是, 这些形式为什么在某一群体中被选择, 以及选择本身传递了什么社会信息。语言形式因此可以被理解为社会意义的索引, 既指向事件, 也指向使用者的群体位置。

在语言态度研究方面, 国内研究逐渐从单一规范判断转向对不同主体态度差异的考察。杨明刚、陈韵超、顾明毅通过大学生调查讨论网络流行语的使用行为, 说明青年对网络表达的接受程度、使用频率与具体交际场景之间存在联系[10]。这类研究的重要价值在于把“社会如何评价网络流行语”的宏观问题转化为可观察的群体差异。不过, 问卷中的自我报告与真实平台行为并不完全等同, 受访者对某一词语的态度也可能随事件热度、交往对象和平台环境变化。因此, 语言态度研究需要与自然语料、访谈和长期观察相结合。

在身份认同研究方面, 何祎金从青少年流行语的意义世界与文化认同出发, 指出青年流行语能够凝结群体经验, 并成为代际文化表达的一部分[11]。近年来, 研究进一步关注青年如何通过“打工人”“社恐”“小镇做题家”等自我指称表达处境、调节情绪和识别同伴。荣婷、左川冀围绕“内卷”展开幻想主题分析, 揭示网络流行语与青年自我身份认同之间的修辞联系[12]。樊小玲等从现代性背景讨论青年自我指称与身份认同的断裂和重构[13]; 任然则强调网络流行语传播中的符号互动、认同与重构关系[14]; 佟文欣从符号互动视角讨论情感共鸣与身份认同机制[15]。这些成果共同说明, 青年身份并非先验固定, 而是在持续互动和自我叙述中被协商出来。

与此同时, 国内研究的跨学科趋势更加明显。孙雪妩从青年亚文化视角讨论网络流行语的传播及社会影响[16], 孙雨萌、李晶晶关注网络空间中的青年身份认同[17], 魏明则从传播主体多元化角度分析网络流行语中的话语权力实践[18]。这些研究将流行语放回社会关系和媒介结构之中, 拓宽了单纯语言形式分析的边界。不过, 不同研究往往分别使用传播、社会心理或文化研究概念, 社会语言学中的语言变体、言语社群、语言态度、身份指示等概念尚未得到充分整合。

3. 文献评述

3.1. 理论与概念层面的不足

综合现有研究, 首先需要解决的是研究对象和理论层级的衔接问题。“网络语言”“网络热词”“网络流行语”“网络迷因”等概念在部分研究中交替使用, 但它们在形式范围、传播强度和研究单位上并不完全相同。若缺少操作性定义, 研究者在样本选择时容易把短期热点词、长期网络词汇、固定句式和多模态迷因混合在一起, 从而影响研究之间的可比性。未来研究可将“在一定时间段内经由数字平台高频传播、被群体反复模仿或再语境化, 并具有可识别社会意义的语言单位”作为工作性定义, 再依据研究目的限定词、短语、句式或标签等具体单位。

其次, 理论使用存在“多而散”的问题。现有研究已涉及身份、亚文化、符号互动、传播机制等多个视角, 但部分研究只是借用概念解释个案, 没有建立语言形式与社会变量之间的稳定分析链条。社会语言学可以在此提供中层框架: 语言变异理论用于描述不同群体和场景中的形式差异; 言语社群或实践共同体视角用于界定共享规范的互动群体; 语言态度理论用于解释使用者和观察者的评价差异; 身份建构理论则用于分析语言选择如何索引自我位置和群体边界。将这些概念放在同一分析框架中, 才能避免把身份认同仅理解为词语内容所“反映”的结果。

3.2. 数据与研究设计的不足

现有研究的第二个突出问题是数据来源与研究设计仍较单一。案例研究能够细致解释某一流行语的意义生成, 但如果只选择少量高知名度词语, 容易忽略未进入年度榜单却在特定圈层高频使用的表达。问卷调查能够比较群体态度, 却可能受到回忆偏差和社会期许影响。单个平台语料又容易把平台特有的用户结构和推荐机制误认为普遍规律。因此, 未来的数据拓展不应只是简单增加文本数量, 而应在“平台-时间-群体-互动类型”四个维度上进行分层采样。

具体而言, 可同步采集微博等公共讨论平台、短视频评论区、弹幕社区和兴趣社群中的公开文本, 并记录发布时间、互动量、回复关系、话题标签等元数据; 对同一流行语进行月度或季度追踪, 观察其从出现、扩散到语义变化的过程; 同时通过问卷和半结构访谈获取青年使用者对词语意义、使用对象和场景限制的自我解释。这样形成的多源数据能够相互校验: 平台语料回答“实际怎样使用”, 问卷回答“总体如何评价”, 访谈回答“为什么这样使用”。

3.3. 计算语言学与质性研究的结合路径

仅仅提出“拓宽数据来源”仍不足以解决大规模语料与社会意义解释之间的矛盾。更可行的方向是建立计算语言学方法与质性研究相互嵌套的混合研究设计。第一步是语料构建与预处理。研究者可按照时间窗口和平台类型采集公开语料, 对用户名等可识别信息进行去标识化处理, 并保留必要的上下文、回复链和时间信息。随后进行分词、同形异义消歧、表情符号与网络缩写标准化, 同时建立流行语变体词表, 避免因拼写变体导致统计遗漏。

第二步可使用主题建模识别大规模语料中的议题结构。LDA 等主题模型能够根据词项共现关系发现潜在主题[7]。在网络流行语研究中, 不宜把主题模型结果直接等同于“真实社会议题”, 而应将其作为发现模式的工具。例如, 可比较某一流行语在就业、学习、消费、情感关系等主题中的出现概率, 观察其语义关联是否随时间发生迁移; 也可比较不同平台的主题分布, 判断同一表达在不同社群中是否承担不同功能。模型给出的高概率词和代表文本应由研究者返回原始语境进行人工解释, 从而避免算法标签替代社会语言学分析。

第三步可引入情感分析考察流行语的情绪和立场变化。Pang 和 Lee 所总结的情感分析方法为文本主观性和情绪倾向的计算处理提供了基础[8]。对于网络流行语, 研究重点不应停留在简单的正负二分类, 而可结合情绪类别、立场对象和语用功能进行细化。例如, 同一流行语可能在自我调侃时表现为表面负向、实际亲近, 在讽刺他人时又表现为排斥。研究者可以先用自动方法筛选情绪分布和异常变化, 再对高频或高争议语境进行人工编码, 并通过抽样复核报告分类准确率。这样既能利用计算方法处理规模, 又能承认网络反讽和语义漂移对自动识别造成的限制。

第四步应通过网络民族志恢复流行语的社群情境。网络民族志强调持续观察在线社群的互动规则、参与结构和文化背景[9]。研究者可以选择具有代表性的公开社群, 记录某些表达在何种话题下出现、谁首先使用、其他成员如何接续、纠正或戏仿, 以及新成员是否需要掌握特定表达才能顺利参与互动。与

一次性抓取相比, 持续观察能够揭示“词语为什么在这里有这种意思”, 也能发现算法统计难以识别的圈层规范、幽默规则和关系距离。研究过程中应遵循研究伦理, 对可识别信息进行匿名化, 并谨慎处理半公开或具有明确社群边界的内容。

第五步以话语分析解释身份建构机制。话语分析可把流行语放在完整互动序列中, 关注称谓、代词、评价词、引语、反讽、重复和回应方式等语言资源, 分析说话者如何定位自己、定位他人以及协商群体边界。例如, 当某一自我称谓被群体成员重复使用时, 需要区分它是单纯描述、幽默式自我定位, 还是用于建立共同经验; 当圈外成员使用同一表达时, 还要考察其意义是否发生变化。通过比较“谁说-对谁说-在什么语境说-他人如何回应”, 才能把身份建构从抽象结论落实到可观察的语言证据上。

最终, 可形成“计算发现-质性解释-再验证”的循环: 主题建模和情感分析从大样本中发现主题聚类、情绪变化和平台差异; 网络民族志与话语分析对关键案例进行深描, 提出关于身份指示和群体边界的解释; 再返回大规模语料检验这些解释是否具有跨时间、跨平台的稳定性。必要时还可加入访谈, 让参与者对研究者的解释进行补充或修正。该设计并非把多种方法机械叠加, 而是让每一种方法回答不同层次的问题, 从而提高研究的解释力和可检验性。

在实际操作中, 还应把方法之间的对应关系预先写入研究方案。例如, 以“某一流行语如何参与青年身份建构”为总问题时, 可设置三个相互衔接的子问题: 其一, 该词在不同平台和时期与哪些主题共同出现; 其二, 不同语境中的情绪倾向、评价对象和互动方向有何差异; 其三, 使用者如何解释该词与自身经历、群体归属及他者区分之间的关系。前两个问题主要由计算分析和人工编码共同回答, 第三个问题则依靠网络民族志、话语分析和访谈深入解释。随后再把质性研究形成的类别转化为可检索的语言特征, 返回语料库进行频率和分布检验。通过这种往返式设计, 研究结论既有大样本证据, 也保留具体语境中的意义层次。

3.4. 青年主体视角与动态研究

青年主体视角仍需进一步加强。现有研究常从研究者立场概括某个词“代表”何种青年心态, 但青年群体内部存在年龄、教育经历、职业、地域、兴趣圈层和平台使用习惯等差异。同一流行语可能被不同青年赋予不同意义, 也可能出现从自我称谓到泛化标签、从圈内表达到账户营销用语的变化。因此, 研究应把青年视为意义协商的参与者, 而不是被动承载某种社会心态的同质群体。半结构访谈、焦点小组和参与式观察可用于了解使用者如何解释自己的语言选择, 并与实际语料进行对照。

此外, 网络流行语具有明显的生命周期。若只在某一时点截取语料, 容易把短期事件效应当成稳定意义。未来可采用纵向设计, 对同一批词语进行连续追踪, 记录频率、搭配、主题和情绪分布的变化, 再结合关键事件节点解释语义转折。对于身份研究而言, 这种动态视角尤其重要, 因为身份本身也是在持续互动中形成和调整的。只有同时观察语言形式和社会语境的变化, 才能解释某一表达为何在不同阶段具有不同的身份功能。

4. 结语

网络流行语研究已经从早期的形式描写逐渐走向社会意义和身份问题, 但从社会语言学角度看, 仍需要进一步打通语言形式、平台语境、群体差异与身份实践之间的联系。现有成果已经证明, 青年会通过网络语言表达经验、调节关系、识别同伴并协商自我位置; 与此同时, 流行语的意义并不固定, 而是在传播、模仿和再语境化中持续变化。因此, 对网络流行语的研究应坚持描述性、情境化和动态化原则, 避免以单一价值尺度预设结论。

未来研究的关键不只是获得更多数据, 而是形成能够连接宏观模式与微观意义的研究设计。主题建

模可以帮助研究者从海量文本中发现议题结构, 情感分析可以描写情绪与立场的分布和变化, 网络民族志能够还原社群规范和参与经验, 话语分析则可以解释具体互动中身份如何被语言行动建构。将这些方法与社会语言学的语言变异、言语社群、语言态度和身份建构理论结合, 并通过访谈、人工编码和跨平台比较进行验证, 可形成更完整的证据链。由此, 网络流行语不再只是需要被列举和释义的“热词”, 而成为理解数字时代青年语言实践、群体关系和社会意义生成机制的重要材料。

参考文献

- [1] 王仕勇. 近十年我国网络流行语研究综述[J]. 重庆工商大学学报(社会科学版), 2012, 29(5): 7-12.
- [2] Crystal, D. (2006) *Language and the Internet*. 2nd Edition, Cambridge University Press.
- [3] 方毅华, 罗鹏. “年度十大网络流行语”编码规律解析[J]. 现代传播(中国传媒大学学报), 2011(12): 77-80.
- [4] 康忠德. 网络语言的社会语言学考察[J]. 石家庄学院学报, 2011, 13(2): 93-97.
- [5] 杨萍. 网络流行语: 网民自主话语生产的文化景观[J]. 新闻前哨, 2010(4): 87-89.
- [6] 汤致英. 网络语言新探[M]. 郑州: 河南人民出版社, 2010.
- [7] Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003) Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- [8] Pang, B. and Lee, L. (2008) Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, 2, 1-135. <https://doi.org/10.1561/15000000011>
- [9] Kozinets, R.V. (2015) *Netnography: Redefined*. 2nd Edition, Sage Publications.
- [10] 杨明刚, 陈韵超, 顾明毅. 大学生网络流行语使用行为的调查与分析——以上海部分高校为例[J]. 中国广告, 2011(12): 118-122.
- [11] 何祎金. 意义世界与文化结晶: 青少年流行语的社会学考察[J]. 中国青年研究, 2018(8): 10-16.
- [12] 荣婷, 左川翼. 偏差与归一: 青年自我身份认同的话语修辞——基于“内卷”网络流行语的幻想主题分析[J]. 现代传播(中国传媒大学学报), 2024, 46(6): 158-168.
- [13] 樊小玲, 陈祉睿, 邓颜. 断裂与重构: 现代性背景下青年的自我指称与身份认同[J]. 中国青年研究, 2023(7): 34-40.
- [14] 任然. 符号、认同与重构: 网络流行语传播中的青年身份认同研究[J]. 新闻论坛, 2025, 39(3): 86-89.
- [15] 佟文欣. 情感共生与身份认同: 符号互动视角下网络流行语的话语分析[J]. 传播与版权, 2023(6): 93-95.
- [16] 孙雪姩. 青年亚文化视域下网络流行语的传播及其社会影响研究[J]. 新闻传播科学, 2025, 13(4): 485-492.
- [17] 孙雨萌, 李晶晶. 青年群体网络空间的身份认同研究——基于网络流行语的传播分析[J]. 新闻潮, 2025(3): 34-37.
- [18] 魏明. 传播主体多元化与网络流行语的话语权力实践[J]. 江汉论坛, 2022(6): 129-135.