

基于多模态语料库的网络流行语演变研究

包鹏飞

香港大学文学院，香港

收稿日期：2025年2月7日；录用日期：2025年2月28日；发布日期：2025年3月13日

摘要

本研究聚焦于网络流行语这一互联网时代极具特色的语言现象，依托多模态语料库展开深度剖析。通过精心构建涵盖多模态数据的网络流行语语料库，综合运用社会语言学、认知语言学和传播学等多学科理论，全面且深入地探究网络流行语的多模态特征、演变路径及其背后的复杂影响因素，旨在为网络语言规范、舆情监测以及社会文化研究提供具有重要价值的参考依据，推动相关领域的学术发展与实践应用。

关键词

多模态语料库，网络流行语，演变研究

A Study on the Evolution of Internet Buzzwords Based on Multimodal Corpus

Pengfei Bao

Faculty of Arts, The University of Hong Kong, Hong Kong

Received: Feb. 7th, 2025; accepted: Feb. 28th, 2025; published: Mar. 13th, 2025

Abstract

This study centers on the unique linguistic phenomenon of internet buzzwords, a hallmark of the internet age, and performs an in-depth analysis utilizing a multimodal corpus. By meticulously constructing a comprehensive corpus of internet buzzwords enriched with multimodal data, and by integrating theories from sociolinguistics, cognitive linguistics, and communication studies, this research delves into the multimodal attributes, evolutionary trajectories, and the intricate factors influencing the emergence and dissemination of internet buzzwords. The ultimate goal is to offer substantial insights for internet language standardization, public sentiment monitoring, and sociocultural investigations, thereby advancing both academic research and practical implementations in pertinent fields.

Keywords

Multimodal Corpus, Internet Buzzwords, Evolutionary Analysis

Copyright © 2025 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

在数字化浪潮的席卷下，互联网已深度融入社会生活的每一个角落，成为信息传播与人际交流的关键平台。网络流行语作为互联网时代语言发展的独特结晶，以其迅猛的态势不断涌现，如“凡尔赛文学”“内卷”“破防”等，这些词汇不仅在网络空间中广泛传播，更逐渐渗透至人们的日常口语表达与书面交流中。网络流行语的生成、传播及演变过程，绝非单纯的语言层面的变化，而是深刻映射出社会文化的动态变迁、民众心理认知的转变以及技术发展的日新月异[1]。传统的语言研究方法，在面对网络流行语这种融合了文本、图像、视频、音频等多种模态信息的复杂语言现象时，常常面临诸多局限，难以全面、深入地揭示其内在规律。多模态语料库的兴起，为语言研究领域带来了全新的视角与方法，它能够整合多种模态的数据资源，为深入剖析网络流行语提供了有力的支持。因此，开展基于多模态语料库的网络流行语演变研究，具有极为重要的理论探索价值与实践指导意义。

在社会语言学领域，通过对网络流行语多模态特征和演变规律的研究，可以深入探究语言与社会变迁之间的紧密联系，揭示社会文化因素如何影响语言的发展和演变。在认知语言学方面，分析网络流行语的语义构建和理解过程，有助于深化对人类认知模式和语言认知机制的认识。从传播学角度出发，研究网络流行语的传播路径和影响因素，能够为信息传播理论提供新的实证依据，丰富传播效果研究的内容。

此外，网络流行语的快速传播和广泛使用，不可避免地会对语言的规范性产生一定影响。一些网络流行语可能存在语法错误、语义模糊等问题，通过深入研究其演变规律和使用特点，可以制定更加科学合理的网络语言规范策略，引导网络语言健康发展。对于舆情监测而言，网络流行语往往是社会舆情的重要风向标。它们能够敏锐地反映出民众对社会热点事件、政策法规等的态度和看法。及时掌握网络流行语的演变动态，有助于相关部门更准确地把握社会舆论走向，为制定有效的舆情应对策略提供参考。在社会文化研究方面，网络流行语作为社会文化的生动载体，蕴含着丰富的社会文化信息。研究其演变过程，能够深入洞察社会文化的变迁趋势，为社会文化研究提供新的视角和研究素材。

2. 多模态网络流行语语料

2.1. 语料来源

本研究精心选取微博、微信、抖音、B 站等主流社交平台作为主要语料来源。微博作为信息传播的核心平台之一，拥有庞大的用户群体和极高的话题热度。在这里，各种热点事件、明星动态、社会话题等都能引发大量用户的讨论，从而催生众多网络流行语。例如，在某明星的热门事件中，网友们创造并传播了“吃瓜群众”这一流行语，用来形容那些围观事件、不参与其中但乐于了解情况的人群。微信作为人们日常社交的重要工具，其多样化的交流场景为网络流行语的产生和传播提供了丰富的土壤[2]。在微信群聊中，针对工作压力大的情况，大家可能会讨论“996”工作制，使得这一词汇逐渐成为网络流行语，反映了职场中的工作时间问题。抖音以其独特的短视频形式，吸引了大量用户参与内容创作和传播。

许多网络流行语通过抖音上的热门视频迅速扩散,如“好嗨哟,感觉人生已经到达了巅峰”这一流行语,就源自抖音上的一个搞笑视频,随后在网络上广泛传播。B站作为年轻人喜爱的文化社区,充满了各种亚文化元素和创意表达。在这里,“二次元”相关的流行语不断涌现,如“番剧”“弹幕”等,这些词汇反映了B站独特的文化氛围和用户群体的兴趣爱好。

为了全面、准确地收集语料,我们综合运用了爬虫技术和人工采集两种方式。通过爬虫技术,能够快速、大规模地获取网络平台上的公开数据,提高语料收集的效率。同时,为了确保语料的质量和代表性,对于一些重要的、具有特殊意义的网络流行语及其相关内容,我们采用人工采集的方式,进行细致筛选和整理。

2.2. 语料类型

收集的语料类型丰富多样,涵盖文本、图片、视频、音频等多模态数据。文本数据主要来源于网络平台上用户发布的评论、帖子、弹幕等。这些文本是网络流行语的基本载体,直接反映了用户的语言表达和交流内容。例如,在微博的热门话题讨论中,用户的评论里常常包含各种网络流行语,如在讨论环保话题时,可能会出现“碳中和”“绿色生活”等流行词汇。图片数据在网络流行语的传播中起着重要的辅助作用,其中表情包和梗图最为常见。表情包以其简洁、生动的形象,能够直观地表达各种情绪和含义,与网络流行语搭配使用,增强了表达效果。比如,“微笑”表情包原本是表示友好的笑容,但在网络语境中,常常被用来表达无奈、嘲讽等复杂情绪,与相关的网络流行语一起,传达出丰富的语义。梗图则是通过将图片与文字相结合,创造出具有特定含义和传播力的图像内容。例如,“真香”梗图,源自一档综艺节目中嘉宾的前后态度转变,配上“真香”这一流行语,用来形容某人先拒绝某事,后来又打脸接受的情况,在网络上广泛传播[2]。

视频数据包含各类短视频、网络直播片段等。在短视频中,网络流行语往往与画面、声音、字幕等多种元素相互配合,共同传达信息。例如,在抖音的一些舞蹈短视频中,会配上当下流行的音乐和网络流行语的字幕,如“跟着节奏摇摆起来,奥利给”,通过视频的传播,这些流行语也被更多人熟知。网络直播片段中,主播与观众的互动交流也会产生大量的网络流行语,如“老铁”“666”等,这些词汇在直播场景中广泛使用,并逐渐传播到其他网络平台。音频数据包括网络流行歌曲、搞笑音频段子等。网络流行歌曲的歌词中常常包含一些流行语,随着歌曲的传唱,这些流行语也得到了更广泛地传播。例如,某首流行歌曲中唱道“燃烧我的卡路里”,这一歌词迅速成为网络流行语,被人们用来鼓励自己减肥、健身。搞笑音频段子则以幽默风趣的语言和独特的音效,传播各种网络流行语,如一些模仿名人声音的音频段子,会加入当下流行的网络流行语,引发用户的共鸣和传播。

3. 网络流行语的多模态特征分析

3.1. 文本特征

通过对多模态语料库中的文本数据进行深入分析,我们可以全面了解网络流行语的文本特征。在词频方面,统计结果显示,一些与社会热点、娱乐八卦、生活方式相关的词汇在网络流行语中出现的频率较高。例如,在某一段时间内,由于一部热门电视剧的播出,剧中的经典台词“我emo了”成为网络流行语,在微博、抖音等平台的文本语料中频繁出现,反映了当下年轻人对情绪低落状态的一种表达。分析词长发现,网络流行语多以简洁、短小的词汇或短语为主,这符合网络交流快速、便捷的特点。例如,“yyds”(永远的神)、“绝绝子”等流行语,简洁明了,易于记忆和传播。

在词性分布上,名词、动词、形容词在网络流行语中都占有一定比例。名词类网络流行语常常指代特定的事物、概念或人群,如“打工人”指代从事体力或脑力劳动的上班族,反映了一种社会群体的自我认同和调侃。动词类流行语则表达各种行为和动作,如“种草”表示推荐、分享自己喜欢的事物,体现

了网络社交中的信息传播行为。形容词类流行语用于描述事物的特征或人的状态，如“凡尔赛”用来形容那些以低调的方式炫耀自己的人，生动地刻画了一种特定的行为风格。

对句法结构的研究表明，网络流行语常常突破传统语法规则，采用简洁、灵活的表达方式。例如，“无语子”这一流行语，在句法上打破了传统的语法结构，“子”字作为一种语气词后缀，增强了表达的情感色彩，使表达更加生动、口语化。从语义特征来看，网络流行语往往具有隐喻、转喻等语义特点。例如，“柠檬精”并非真正指一种化学物质，而是通过隐喻的方式，用来形容那些羡慕、嫉妒别人的人，以柠檬的酸味来比喻嫉妒的心理，形象生动地传达了丰富的语义。

3.2. 图像特征

在图像模态的范畴之中，针对色彩与网络流行语相关图片里的运用展开深入且全面的剖析。不同的色彩往往承载着丰富多元且独具特色的情感与意义，例如红色通常象征着热烈激昂、喜庆欢愉，而绿色或许意味着清新宜人、环保低碳等。对构图特征进行细致入微的探究，其涵盖了画面的布局规划、元素的排列方式等诸多方面。诸如部分表情包凭借极度夸张的人物表情以及简洁明了的构图，强有力地凸显出某种特定的情绪。针对图片中的人物和场景展开深度透彻的剖析，从而明晰其与网络流行语之间的紧密关联，人物的动作、表情以及场景的设置往往能够对网络流行语的含义起到良好的辅助阐释作用。此外，表情包作为网络流行语传播的关键载体，拥有独树一帜的图像特征，其简洁清晰、生动有趣的形象能够迅速吸引用户的注意力，在极大程度上增强网络流行语的传播成效。

3.3. 视频特征

就视频模态而言，时长无疑是一项极其关键的特征。不同类型的网络流行语视频，其时长可能存在显著的差异。一些旨在快速吸引用户目光的搞笑短视频或许较为短促，而某些用于深度解读网络流行语的视频则可能相对较长。对视频画面的内容进行全方位地分析，包括画面的主题设定、情节的发展走向等。画面中的场景切换、人物互动等元素均与网络流行语的传播和理解紧密相联。声音在视频中发挥着举足轻重的作用，背景音乐、音效、人物语音等能够营造出特定的氛围，增强视频的感染力，从而更高效地传播网络流行语。字幕作为视频中文字信息的重要组成部分，其字体、颜色、出现的时机等均会影响用户对网络流行语的接收和理解。另外，视频中的特效运用，诸如滤镜、动画效果等，能够为视频增添趣味性和吸引力，进一步有力地推动网络流行语的传播。

3.4. 音频特征

在音频模态之下，深入且系统地探究音高、音强、音长、音色和节奏等特征。音高的变化能够传递丰富多样的情感和语气，高音可能代表着惊讶、兴奋等情绪，低音则可能传递出沉稳、严肃等情感。音强的大小能够有效地突出重点，显著强化音频的表现力。音长的长短能够对节奏和韵律产生影响，例如一些网络流行歌曲通过规律有序的音长变化来营造出独特的节奏。音色的特点也与网络流行语的表达存在密切关联，不同的音色能够给予人截然不同的听觉感受，如清脆的音色可能让人感觉活泼、欢快。节奏是音频的核心特征之一，富有节奏感的音频更易于被用户铭记于心，进而有力地促进网络流行语的传播。

4. 网络流行语的演变路径研究

4.1. 生命周期

网络流行语具备显著的生命周期，通常可以划分为萌芽期、爆发期、稳定期和衰退期。在萌芽期，网络流行语通常由少数个体创造或使用，可能源自某个小众圈子或特定事件，其传播范围相对有限。随

着相关事件的持续发酵或某个契机的出现,进入爆发期,网络流行语在网络上迅速广泛传播,被大量用户运用和讨论。在稳定期,其使用频率相对稳定,成为网络语言的组成部分,被广泛应用于各类网络交流场景。然而,伴随时间的推移以及新的网络流行语的涌现,原有的流行语逐步进入衰退期,使用频率下降,逐渐被人们淡忘。

4.2. 传播路径

网络流行语的传播路径主要是从线上拓展至线下,从特定群体蔓延至大众群体。首先在网络平台崭露头角,通过社交媒体的传播,例如微博的话题探讨、抖音的短视频传播等,在网络用户中迅速扩散开来。随着其影响力的不断扩大,逐渐从线上延伸至线下,出现在人们的日常对话、广告宣传、影视节目等场景之中。同时,网络流行语往往先在特定群体中盛行,如年轻人、某个兴趣爱好群体等,然后通过这些群体的传播,逐渐被大众所接纳[3]。

4.3. 语义演变

网络流行语在演变进程中会出现语义变化。常见的语义演变涵盖词义扩大、缩小、转移和褒贬化[4]。词义扩大指原有的词汇意义范畴扩展,例如,“粉丝”一词原本指的是食品,现今则主要指崇拜某个明星或事物的人,词义范围显著扩大。词义缩小则与之相反,如“勾当”一词,原本有事情、活动的含义,现在主要指坏的事情。词义转移表示词汇的意义发生转变,如“种草”原本指的是种植植物,如今在网络语境中意味着推荐、分享自己喜欢的事物。褒贬化是指词汇的褒贬色彩发生变化,例如,“奇葩”一词原本是褒义词,指奇特而美丽的花朵,现在在网络语境中常被用于形容行为或想法怪异的人,带有一定的贬义色彩[5]。

4.4. 形式演变

网络流行语的形式持续演变,从最初的文字形式逐步发展为融合图片、视频、音频等多种模态。早期的网络流行语主要以文字形式传播,伴随互联网技术的进步和社交媒体的普及,图片、表情包等开始与文字相结合,增强了表达效果[5]。之后,短视频和音频的兴起,促使网络流行语的传播形式更加丰富多样,从单一模态向多模态融合的方向发展。例如,一些网络流行语通过短视频的形式,将文字、画面、声音等多种元素融合一体,以更为生动、形象的方式进行传播[6]。

5. 网络流行语的影响因素分析

5.1. 社会文化因素

社会热点事件是网络流行语产生和传播的重要推动力。例如,新冠疫情期间,“清零”“核酸检测”等与疫情防控相关的词汇成为网络流行语,反映了社会对疫情的关注和应对举措。流行文化也对网络流行语的形成和发展产生影响,如电影、电视剧、音乐等作品中的经典台词、角色形象等常常成为网络流行语的来源。此外,亚文化群体的文化特色和语言习惯也会催生网络流行语,如二次元文化中的“萌”“御宅族”等词汇,展现了二次元文化的独特魅力。

5.2. 心理认知因素

从众心理致使人们在网络交流中倾向于使用流行的语言表达,以融入群体。当看到周围的人都在使用某个网络流行语时,个体往往会随之使用,从而推动了网络流行语的传播。求异心理则促使人们追求独特、新颖的语言表达方式,创造出富有个性的网络流行语。例如,一些年轻人为了彰显自身的与众不同,会创造一些独特的词汇或表达方式,这些词汇在小圈子中流行后,有可能扩散成为大众流行语。娱

乐心理也是网络流行语传播的重要因素，人们在网络交流中期望通过使用幽默、风趣的网络流行语来增添交流的趣味性和娱乐性。

5.3. 技术发展因素

互联网技术的发展为网络流行语的产生和传播提供了基础条件。高速的网络连接、便捷的信息发布平台，使得网络流行语能够迅速传播。移动终端的普及，让人们能够随时随地在网络上交流，进一步促进了网络流行语的传播。社交媒体平台的不断创新，如短视频的兴起、直播的发展等，为网络流行语的传播开辟了新的渠道和形式。同时，人工智能技术的应用，如智能推荐系统，能够依据用户的兴趣和行为，推送相关的网络流行语内容，加快了其传播速度。

6. 结论

通过构建多模态网络流行语语料库，并基于该语料库进行深入研究，本研究揭示了网络流行语的多模态特征、演变路径和影响因素。网络流行语在文本、图像、视频和音频等多个模态上展现出独特的特征，这些特征相互配合，共同传达其丰富的意义。其演变路径呈现出明显的生命周期、传播路径和语义、形式演变规律。社会文化、心理认知和技术发展等多方面因素共同影响着网络流行语的产生、传播和演变。

未来的研究可以进一步扩大语料库的规模和范围，涵盖更多的网络平台和语言类型，以提高研究结果的普遍性和可靠性。在研究方法上，可以结合更多的学科理论和技术手段，如机器学习、深度学习等，对网络流行语进行更深入地分析和预测[7]。此外，还可以关注网络流行语对语言教学、文化传承等方面的影响，为相关领域的实践提供更多的指导和建议。

参考文献

- [1] 耿敬北, 陈子娟. 网络社区多模态话语分析[J]. 外语教学, 2016, 37(3): 35-39.
- [2] 肖珺. 多模态话语分析: 理论模型及其对新媒体跨文化传播研究的方法论意义[J]. 武汉大学学报(人文科学版), 2017, 70(6): 126-134.
- [3] 高倩琳. 网络流行语“拿捏”的词义演变探析[J]. 现代语言学, 2023, 11(9): 4018-4025.
- [4] 潘淑婷. 当代网络语“麻了”的语义演变及语法化[J]. 现代语言学, 2024, 12(5): 266-274.
- [5] 倪昕. 流行语“X 门”新解[J]. 现代语言学, 2023, 11(10): 4732-4738.
- [6] 石磊, 杜军平, 周亦鹏, 等. 在线社交网络挖掘与搜索技术研究[J]. 智能系统学报, 2016, 11(6): 777-787.
- [7] Thibault, P.J. and Baldry, A. (2001) Towards Multimodal Corpora. *Corpora in the Description and Teaching of English: Papers from the 5th ESSE Conference*. Helsinki, 29 August-2 September, 2000, 1000-1016.