

大数据时代下湘西苗语短视频传播情况实证研究

江娜，肖霞

吉首大学，湖南 吉首 416000

DOI:10.61369/SE.2025080014

摘要：本次研究所采用的人工抽样方式收集了抖音平台湘西苗语相关的短视频数据，在此基础上从地域分布、内容形式、受众接受度、情感态度以及热门视频等不同角度展开分析，并辅以可视化图例展开说明影响湘西苗语传播效果的因素。通过该部分对湘西苗语短视频的实证研究表明，目前湘西苗语在短视频环境下的传播还存在一定阻碍，应该有针对性地制定科学合理的传承与推广措施。因此，本研究将从政策、平台、内容以及受众等维度出发，提出具体的实现方式方法，以便于将分散、碎片化的民族语言文化遗产进行数字化，达到对少数民族地区民族语言的保护的目的，供少数民族语言的数字化保护提供理论参考与实践道路。

关键词：大数据时代；湘西苗语保护；短视频推广

An Empirical Study on the Dissemination of Short Videos of Miao Language in Xiangxi in the Era of Big Data

Jiang Na, Xiao Xia

Jishou University, Jishou, Hunan 416000

Abstract : This research adopted a manual sampling method to collect short video data related to the Xiangxi Miao language on the Douyin platform. Based on this, an analysis was conducted from different perspectives such as geographical distribution, content form, audience acceptance, emotional attitude, and popular videos. Visual examples were also used to illustrate the factors influencing the dissemination effect of the Xiangxi Miao language. The empirical research on the short videos of the Miao language in Xiangxi in this part shows that there are still certain obstacles to the dissemination of the Miao language in Xiangxi in the short video environment at present. Targeted scientific, reasonable and effective measures for inheritance and promotion should be formulated. Therefore, this study will start from the dimensions of policy, platform, content and audience, and propose specific implementation methods and approaches, so as to digitize the scattered and fragmented ethnic language cultural heritage, achieve the purpose of protecting ethnic languages in minority areas, and provide theoretical references and practical paths for the digital protection of ethnic minority languages.

Keywords : the era of big data; protection of the Miao language in Xiangxi; short video promotion

一、研究背景与目标

某些湘西土家族苗族自治州地区人民使用的语言就是苗语和土家语，而在所有的少数民族语言中湘西苗语也属于十分珍贵的中华少数民族语言之一，在传承民族文化方面有着不可小觑的作用。我国深化数字建设，以短视频为主要的宣传手段，加上新媒体等力量的影响下，数据信息采集技术也逐渐完善成熟，因此利用大数据保存和推广濒危语言成为新的可能，但是在实际运用中，湘西苗语却面临着使用场合缩减、传承断层等问题，^[1]例如，由于政策落实不到位、没有对湘西地区实行本地化的教学安排、年轻人对于语言的态度有所转变等使当地湘西苗语的社会工具作

用逐渐减弱。

根据文献资料显示，目前的学术研究多集中在湘西苗语的语音系统、语法结构和词汇特征等本体研究，但相对缺乏针对大数据时代下其传播效果和受众反应的实证研究，^[2]现在短视频媒介盛行，以短视频为湘西苗语传播的媒介，极其有利于湘西苗语的传播和保护湘西苗语的可持续发展。基于这一研究空白，本研究于2025年2月至5月期间，以国内主流短视频平台抖音及湘西自治州为调研基地，通过人工抽样方式获取183条具有“苗语”标签的相关短视频，并随机抽取2000条用户评论数据，^[3]系统地考察了湘西苗语短视频的传播现状，识别出影响传播效果的关键因素，并最终提出针对性的改进方案，为少数民族湘西苗语语言的数字化

课题信息：课题名称：大数据时代湘西苗语保护路径研究，课题编号：S202510531069，立项单位：湖南省教育厅。

作者简介：

江娜（2004.10-），女，湖南邵阳人，本科，专业：汉语言文学（师范），研究方向：新闻传播学类；

肖霞（2005.07-），女，湖南邵阳人，本科，专业：汉语言文学（师范），研究方向：新闻传播学类。

保护提供一定的数据支撑和科学的理论依据。

对湘西苗语进行探究分析，这里所指的湘西苗语是指覆盖了东部方言区、西部方言区以及中部方言区的主要代表方言。通过对抖音短视频上的数据进行搜集，在大量的短视频里人工筛选出部分短视频，并且从其中视频中随机抽取2000条评论展开进一步分析，综合判断苗语在新媒体时代如何被传播。^[4]

二、数据来源与研究方法

(一) 数据来源

研究中的所有数据都是来源于抖音短视频这一平台，包括人工提取的苗语类短视频内容及短视频的元数据共183条，其中包括苗语短视频的点赞数量、制作情景等，还从其中随机采样了2000条用户的评论。此外，在搜集数据时需要严格遵守平台规定和科研伦理的要求，对于收集到的数据来源于平台上的公开信息，并且经过多个步骤验证后选取，具有真实性和代表性。

(二) 分析工具与维度

本课题结合人工编码与人工智能分析技术方法，基于评论数据开展视频发布的主观感受分析、评论主题挖掘，主要包含六方面：①视频发布地域分布②视频发布的内容场景分类③观众对视频的理解程度④评论的情感倾向性⑤高热度视频的内容特点⑥评论类别高频词，确保过程严谨系统化，结果准确可参考。^[5]

三、调研数据分析

(一) 地域与内容场景分布特征

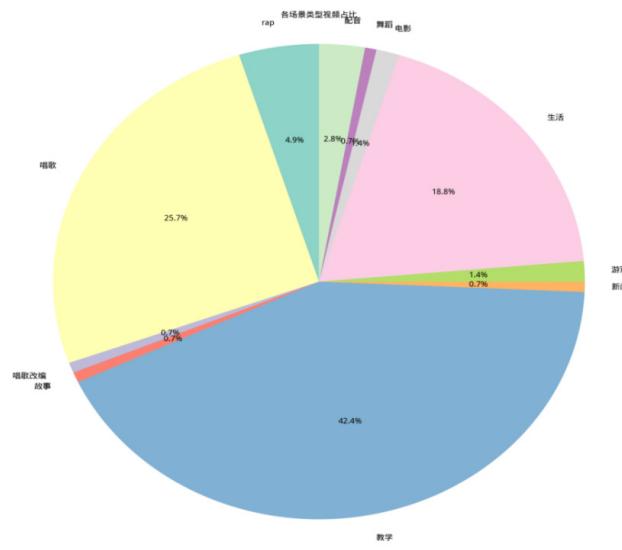


图1 苗语短视频各场景类型占比图

1. 其中语言教学类视频为数量最多的类型，占比达42.4%;民族歌舞表演类视频位居其次，占比达25.7%;日常生活场景类视频占比为18.8%;民间故事讲述、新闻资讯类、游戏娱乐类及其他类型的视频占比均小于5%。从以上不同场景类型中的数量分布可以发现，一是苗语是属于像其他一种文化技能需要学习的，而不是

像一般的生活工具，所制作的内容都是被假定主要是学语言的，因此内容制作者比较愿意制作相对简单的教学类内容；^[5]二是在人们的社会生活中不需要广泛地使用苗语交流，因此对于苗语的期待只是因其语言内容的特殊性而产生的了解和学习愿望，使得表演性的歌唱类成为一种流行形式。



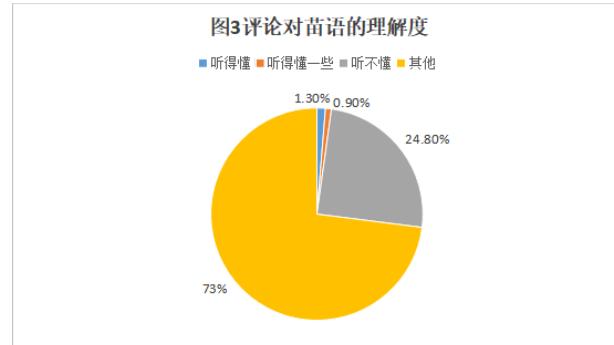
图2 湘西苗语教学类视频在所有地区教学类苗语短视频占比

2. 地域传播对比分析

图2是2020—2025年苗语教学类视频所占的比例图，对比湘西、贵州两地苗语教学类视频在平台上所占的比例，可以得知湘西从2020年的4.0%逐渐下降至2025年的3.7%，整体呈现出边缘化的趋势。

这是由多方面因素造成的：一是湘西与贵州苗族的人口数量差异，贵州苗族人口占据全国苗族人口总数的40%以上，^[6]而湘西苗语使用者不足百万人，在贵州就具备天然传播规模的优势；二是短视频平台本身具有流量特性，“贵州苗语”相关话题总播放量为3.2倍于“湘西苗语”相关话题总播放量，且短视频算法本身的流量聚集特性使“自带流量”的强势方言展现出了更为突出的马太效应。

(二) 评论分析与受众认知研究



1. 通过针对评论数据量化的统计可见，只有1.3%的网友认为自己完全听懂了视频中苗语的内容；0.9%的网友认为自己“能部分听懂”；有24.8%的网友认为自己完全听不懂；剩余73%的评论则是不具体说懂或者不懂的语言方面论述，^[7]而是一些普通文化的点评。因此“高播放量—低理解度”这种本该互相促进的传播关系在事实作用下产生了背离。

导致这一现象的主要原因有三方面：一是湘西苗语短视频的受众结构错位，超过95%明确表达理解情况的用户是“外部围观者”，平台算法更倾向于将内容推送给对少数民族文化感兴趣但

不具备苗语基础的人群；二是湘西苗语的使用人群参与不足，其中熟练掌握苗语的年长者大多极少使用智能手机和不熟悉短视频平台的操作使用，而年轻一代的苗语能力又普遍衰退，^[8]无法创作湘西苗语短视频，造成了传承主体与传播平台的脱节。

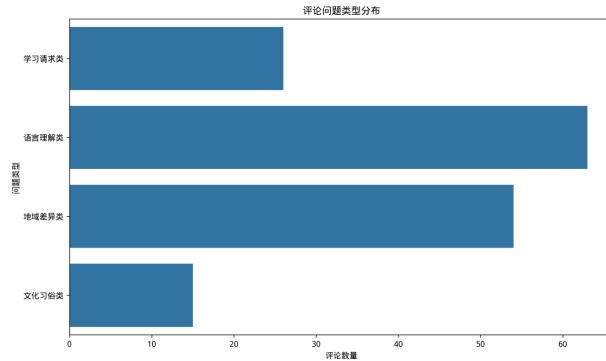


图4评论问题类型分布

2. 由上述情感倾向及关注重点可知，对湘西苗语短视频持正面态度的人群占39.7%，主要集中在以下三个方面：一是苗族文化审美体验。在苗族服饰、苗族舞蹈等民族视觉文化符号上看到更多吸引点，采用“音画分离”的方式欣赏苗族视觉文化符号而不必依靠语言理解；二是方言对比的兴趣。^[9]有25%以上的评论谈到苗语与其他的语言或方言之间进行了比较和对比，这是人们总是把自己的所知所见作为衡量和判断新的事物的心理习惯；三是苗族语言的认知表达。有27%的评论提到对苗语的认知情况，“听不懂”占了苗族语言认知表示数量的58.1%，但“学”“音发”“是什么意思”等高频词（如图6所示），表示出了有想学或者想了解的群体。

（三）高热度视频内容特征

高热度视频的内容类型比较集中，歌舞表演类的比例远高于其他类别；但是从“苗语”被标记为“热度高”的几个视频来看，这些抖音短视频或只用极少的语言碎片片段拼凑出残破的语言表象符号，并无真情实感地使用苗语。之所以出现这样的情况，一方面是基于用户的消遣动机更多是图“新奇性”，其次是为了激发对方的好奇心理在异域风情的语境之下完成“文化打卡”或者“点赞刷屏”，而且这种方式也不需要以了解这种语言为前提条件；另一方面，正是基于平台机制与创作环境的原因，在强烈的算法推荐体系下，“强表现力”等于是天然存在在“热度”之中，而歌舞视频因其“动”态性强、节奏明快、视觉震撼，因而更容易吸引用户的瞬间目光。

四、主要研究结论

湘西苗语短视频具有一定的传播障碍，传播障碍分析如下：

1. 不过从整体上来看，湘西苗语短视频内容的数量一直在上升的，但是湘西本地方的传播力以及话语权偏弱，甚至整个西南地区苗语本地化传播相比贵州等地都处于非常边缘的地位，苗语地域文化的传播差异性越来越突出，湘西苗语的传播力还不够强。

2. 对湘西苗语短视频用户而言，积极的、充满肯定性的、以及觉得有趣的信息占绝大多数，苗语及苗族文化的内核被许多观众所把握，苗族方言与普通汉语之间迥异的巨大差别得到了广泛关注，然而相较于普通语种来说，一般观众对苗语本身的认识明显不足，“高观看量 - 低理解度”的对立观念较为常见，应着重提高用户的理解度以及提高内容的可接近度。

3. 少数民族语言苗语高热度短视频的题材都存在一定的场景局限性，大多数都是以歌舞表演为主的，更有一些与苗语无关的内容，这反映出了大数据时代传播中语言的价值、文化传播的表现方式已经出现了偏差。而且，一部分人接收的也是经过算法判定推荐出来的一些类似的内容，因此也导致内容趋于同质化。

五、推广与保护策略建议

（一）降低语言理解门槛，提升内容可达性

大力推进双语字幕标准化应用，提供“苗语 - 普通话”精准对照字幕，并对文化专有词汇和表达方式添加必要的注释说明；根据用户的语言基础和认知水平建立内容分级体系，开发“基础会话”“情景对话”等系列化入门课程，如经常使用的词“吃饭”，结合图像、动画、情景剧等多样化表现形式，可以借鉴其他情境剧的视频，了解用户感兴趣的点，增强内容的可理解和学习趣味性。

（二）根据用户兴趣焦点来发掘苗语的传播价值。

创作者根据用户群体喜欢“方言对比”这样的特点，创造“苗汉词汇对照”“方言特色解说”“语言接触现象”等等有趣和高价值的知识类科普文章，在湘西苗语短视频方面还能够与用户感兴趣的湘西苗语短视频中的部分非遗项目结合起来，比如将“湘西苗语”视频中所讲到的“湘西赶尸”一起加入短视频推广。根据抖音上的相关数据显示：截至2025年9月17日为止，抖音上一条介绍湘西赶尸的抖音短视频收获了905.9万赞，可见大多数短视频的用户都非常喜欢湘西赶尸文化。

（三）进一步促进湘西地区的文化与语言相融合，让湘西苗语得到更好的传承

湘西苗语创作人员可以通过上述数据向人们展示苗族歌舞、苗族节庆仪式、苗族工艺制作等内容，在教授唱苗歌时加上对歌词的意义说明、语法解释和文化背景的讲述，在教授苗话的同时也教会了听众语言的规律；在抖音平台的视频内容或个人主页的相关内容中为网友提供线上的视频教学链接、线上课程的链接以及一些学习的社区入口、词典的入口等，使读者有更深的学习欲望。

（四）建立互动反馈的传播生态体系来打造湘西苗语传播的新局面

对于湖南省有关方面来说，在支持创作者认真分析短视频上弹幕与积极回复评论里关于湘西话理解上的问题，并收集整理为常见的问答集合，比如制作“苗语问答专栏”“常见误区解析”等内容，有助于加深了解者对于苗语的理解。除此之外，进一步培育当地政府和文化部门、教育部门等与平台紧密合作，结合各地

实际需求培养一批既懂苗语、又擅长新媒体的人才，包括线上线下结合的方式举办创作培训班，或者落实创用金等方式共同扶持优质作品产出，拓宽从业群体与受众群体之间鸿沟。

（五）抖音短视频平台上线创新的算法推荐机制，助推了湘西苗语短视频的内容的差异化发展

本文提出短视频平台可以在流量分配时多给予诸如教学类、日常生活类、故事类等功能性的内容一定的倾斜，在平台中设置

湘西苗语的专项推荐渠道，让算法更多地为有湘西苗语短视频内容的用户提供流量扶持；短视频平台建立用户兴趣标签系统，由用户根据自身对湘西苗语的内容喜好进行筛选，为用户提供相应的湘西苗语学习内容以及难易程度供其选择；短视频平台建立湘西苗语的内容质量评价体系，将湘西苗语使用准确性及文化表达深度纳入评价标准中，倒逼创作者提高作品内容质量及深度。

参考文献

- [1] 湘西土家族苗族自治州人民代表大会.湘西土家族苗族自治州非物质文化遗产保护条例 [A/OL].2025.
- [2] 张世方,沈丹萍.中国语言资源保护的理念与实践——以汉语方言为视角 [J].语言学研究,2017,(01):6-16.
- [3] 朱萍.全球一体化背景下的城市语言生活现状、语言治理及其对文化国际传播的影响研究——以义乌市为例 [J].文化学刊,2025,(02):221-224.
- [4] 庄善洁,孟淑梅.大数据环境下赫哲族语言的数字化挖掘与保护 [J].边疆经济与文化,2017,(11):44-45.
- [5] 李乾丙,王相飞,周榕.中国民族传统体育文化国际传播效果提升的影响因素与组态路径——基于 YouTube 热门视频的实证分析 [J].武汉体育学院学报,2024,58(05):35-42. DOI:10.15930/j.cnki.wtxb.2024.05.007.
- [6] 曹志耘.湘西方言概述 [J].语文研究,2007,(01):42-47.
- [7] 中国互联网络信息中心(CNNIC).第56次中国互联网络发展状况统计报告 [R].北京：中国互联网络信息中心.2025.
- [8] 严妍.我国少数民族语言文化数字化保护与研究 [J].中国民族博览,2024,(15):228-230.
- [9] 张静,李醉海.少数民族语言文字的“数字化”保护探索——以西北部分民族语言为例 [J].贵州民族研究,2018,39(09):200-203.DOI:10.13965/j.cnki.gzmzyj10026959.2018.09.045.