

全球在地化理论视角下的中国乡村文化传播研究

——以贵州“村超”为例

尹含秋, 唐丽君

贵州财经大学外语学院, 贵州 贵阳

收稿日期: 2026年8月2日; 录用日期: 2026年8月26日; 发布日期: 2026年9月3日

摘要

本研究以贵州榕江“村超”国际传播全流程多模态素材为研究载体, 全程依托Python自动化爬虫工具完成YouTube、TikTok、CGTN英文频道、BBC海外报道、贵州官方英文外宣平台多源异构数据采集, 搭建标准化双语多模态语料库; 结合ELAN多模态标注软件、AntConc语料库检索工具、SPSS统计分析软件开展全流程量化实证操作, 形成可落地的“文化补偿翻译 + AR多模态叙事 + 游戏化跨界传播”技术优化方案。

关键词

贵州“村超”, 国际传播, Python数据采集, 多模态语料库, ELAN标注, 量化传播评估

A Study on the Communication of Rural Chinese Culture from the Perspective of Glocalization Theory

—Taking Guizhou “Cunchao” as an Example

Hanqiu Yin, Lijun Tang

School of Foreign Languages, Guizhou University of Finance and Economics, Guiyang Guizhou

Received: August 2, 2026; accepted: August 26, 2026; published: September 3, 2026

Abstract

This study uses the multi-modal materials of the “Cunchao” event in Rongjiang, Guizhou Province

文章引用: 尹含秋, 唐丽君. 全球在地化理论视角下的中国乡村文化传播研究[J]. 交叉科学快报, 2026, 10(5): 1197-1207. DOI: 10.12677/isl.2026.105141

throughout its international dissemination process as the research carrier. The entire process relies on the Python automated web scraping tool to complete the collection of multi-source heterogeneous data from YouTube, TikTok, CGTN English Channel, BBC overseas reports, and the official English external publicity platform of Guizhou Province. A standardized bilingual multi-modal corpus is established. Combined with the ELAN multi-modal annotation software, AntConc corpus search tool, and SPSS statistical analysis software, a full-process quantitative empirical operation is carried out, forming a feasible “cultural compensation translation + AR multi-modal narrative + gamified cross-border dissemination” technical optimization plan.

Keywords

Guizhou “Cunchao”, International Dissemination, Python Data Collection, Multi-Modal Corpus, ELAN Annotation, Quantitative Communication Evaluation

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. 绪论

1.1. 研究实操背景与技术痛点

“全球在地化(Glocalization)”由美国社会学家罗兰·罗伯(Roland Robertson)提出,核心揭示全球化扩张与本土文化互构的双重逻辑,早期研究多聚焦影视、流行媒介的跨国流动,形成两类主流研究路径:一是媒介技术决定论视角,强调海外社交平台算法规则、全球媒介框架会消解地方文化独特性;二是地方能动视角,认为本土主体可依托在地文化资源反向重塑全球媒介叙事。保罗·塞吉斯蒙迪(Paolo Sigismondì)延伸该理论至数字媒介场域,提出数字全球在地化概念,但现有实证研究仍存在两处短板:其一,多数研究仅采用定性文本分析,缺少爬虫、多模态标注工具支撑的大样本量化数据,难以客观衡量全球平台规则与本土文化表达之间的张力;其二,国内文旅、乡村体育相关研究多单向讨论“对外输出”效果,忽视全球受众与地方文化双向平等对话的维度,缺少对民间自发传播、商业市场传播力量的整合分析。

聚焦乡村体育跨文化传播细分领域,谯胜念、石培华等学者围绕村超海外短视频传播展开效度分析,但现有研究存在明显局限:研究素材仅选取少量人工截取片段,未搭建标准化多模态语料库,无法量化视听符号协同失衡问题;同时研究主体单一聚焦政府发布内容,忽略游客、民间创作者、茶企商家等自发生产的传播内容。数字多模态话语领域,高子涵等验证 ELAN 工具用于文化符号量化的可行性,但尚未将该标注体系与全球在地化理论结合,未能解释全球平台叙事逻辑如何造成本土文化符号编码割裂。

基于现有研究空白,本研究以全球在地化为核心分析框架,融合数字多模态话语分析、跨文化翻译、受众接受研究三重维度,一方面依托 Python 自动化采集、ELAN 分层标注完成全流程量化实证,弥补现有研究样本小、数据主观性强的缺陷;另一方面跳出单一政府传播主体视角,整合民间 UGC、商业品牌、境外媒体多元素材,探讨全球平台标准化传播逻辑与榕江本土民族文化表达之间的内在张力,尝试对数字媒介场景下的全球在地化理论进行细节补充。

第一,海外传播多模态素材获取渠道闭塞,人工复制粘贴视频、评论、外宣文本效率极低,海外平台反爬机制、动态加载、分页数据、加密评论接口导致人工采集样本量不足,无法形成大样本实证分析;

现有研究多采用少量人工截取片段作为分析语料, 数据规模受限, 结论客观性不足。

第二, 民族文化多模态符号缺乏标准化量化标注流程, “鱼钩茶”“拦门酒”“苗绣纹样”“芦笙演奏”等视听文化符号分散在数万条短视频中, 人工记录符号出现频次、搭配关系误差极大, 缺少统一软件标注规范与自动化统计工具支撑, 难以量化多模态协同失衡问题。

第三, 海外受众接受度、文化误解程度、品牌记忆度无法通过数据客观测算, 优化策略缺少前后对照实证数据支撑, 也无法回应全球在地化视域下, 全球媒介规则如何扭曲本土文化符号表达这一核心理论问题。

1.2. 研究实操目标与技术任务

1.2.1. 实操目标

(1) 搭建一套适配乡村赛事跨文化交流场景的 Python 自动化多模态数据采集系统, 打通 YouTube、TikTok、海外主流媒体英文报道、国内多元双语内容创作四大数据源(涵盖地方文旅机构、民间创作者、本土茶企商业文案)。

(2) 建立标准化多模态语料库工程化规范, 完成村超跨文化双语内容、赛事短视频、非遗展演音视频素材分类存储。

(3) 依托 AntConc、SPSS 完成双层量化分析: 一是文本层挖掘文化负载词翻译偏差、高频误解搭配; 二是受众层完成海外问卷数据信效度、相关性测算, 搭建三级量化传播效能评估指标体系。

(4) 输出可直接交付文旅企业、本土民间内容创作者使用的标准化跨文化多模态内容制作操作手册。

1.2.2. 核心技术任务拆解

任务 1: Python 爬虫环境搭建与多数据源采集脚本开发;

任务 2: 多模态原始数据自动化清洗、格式统一、去重降噪工程化处理;

任务 3: 双语平行多模态语料库结构化搭建与本地存储架构设计;

任务 4: ELAN 多模态分层标注模板设计、批量标注实操、符号频次统计导出;

任务 5: AntConc 语料库检索, 文化负载词翻译偏差词频、搭配关系提取;

任务 6: 海外受众问卷标准化发放、数据录入、SPSS 信效度与相关性分析;

任务 7: 多模态传播优化技术方案设计、AR 短视频素材制作、小范围试点数据对照测试;

任务 8: 全套技术流程标准化文档、可复用代码、语料样本归档整理。

1.3. 数据来源真实性说明

(1) 海外视频社交平台: YouTube 关键词“Village Super League Rongjiang Guizhou”、TikTok 国际版 #VillageSuperLeague 话题下公开短视频、视频字幕、评论区文本; 采集工具 yt-dlp、TikTokPy, 遵循平台 robots 协议, 仅抓取公开 UGC 内容, 不获取用户隐私信息。

(2) 境外权威媒体英文报道: BBC 官网、China Today 英文频道、华尔街日报官网、新加坡海峡时报英文网页村超相关报道、图文配图说明文本, 通过 requests + BeautifulSoup 静态网页爬虫批量抓取。

(3) 国内官方双语外宣素材: 贵州省政府英文官网、CGTN 村超专题双语稿件、“This is Village Super League”官方海外先导片双语字幕、贵州茶产业出口英文包装文案、村超海外展会宣传册电子版文本, 由榕江县委宣传部提供公开授权素材。

(4) 对照语料: 日本煎茶、印度阿萨姆茶海外电商英文详情页, 作为同类农产品文化翻译对照样本, 统一从亚马逊海外站公开商品页面抓取。

2. 基于 Python 的多模态数据自动化采集与清洗工程

2.1. Python 采集环境标准化搭建

2.1.1. 软硬件基础配置

硬件设备：台式工作站(CPU i7-13700, 内存 32G, 固态 2TB), 满足批量视频下载、多线程爬虫、大语料存储需求;

系统环境：Windows 11 专业版, Python 3.12 稳定版本, 配套虚拟环境隔离依赖包, 避免工具版本冲突;

核心依赖库完整清单(全部通过 pip 官方渠道安装, 版本可复现):

网络请求层: requests 2.31.0、httpx 0.26.0、playwright 1.42.0(动态页面渲染);

视频音频处理: yt-dlp 2026.03.22、faster-whisper、FFmpeg 6.1(音视频转码、字幕提取);

数据存储解析: pandas 2.2.0、numpy 1.26、json、csv、os、pathlib(本地文件管理);

反爬与异步优化: threading、asyncio、fake-useragent(随机请求头);

图像辅助处理: Pillow 10.2.0 (图片格式校验、尺寸统一)。

虚拟环境搭建实操命令:

Plain Text

```
python -m venv multimodal_crawl
```

```
multimodal_crawl\Scripts\activate
```

```
pip install requests httpx playwright yt-dlp pandas numpy pillow fake-useragent faster-whisper
```

```
python -m playwright install chrome
```

2.1.2. 反爬机制配套技术方案

fake-useragent 自动生成全球主流浏览器 UA, 每次请求更换标识, 固定添加 Accept-Language: en-US,en 适配海外站点语言识别[1]。asyncio 设置随机请求间隔 1.5 s~3.5 s, 分页抓取每 20 条数据强制休眠 8s, 规避 429 访问限制。通过 playwright 读取 Chrome 本地公开 Cookie 文件, 仅加载游客浏览凭证, 不获取登录隐私数据, 解决 TikTok、YouTube 部分加密接口访问限制。脚本内置 try-except 捕获超时、403、502 异常, 单次失败自动重试 3 次, 连续失败写入日志文件单独标记, 人工复核链接有效性[2]。

2.2. 四大数据源 Python 采集脚本实操逻辑

2.2.1. YouTube 村超视频与元数据批量采集

采集目标: 短视频 MP4 视频文件、自动生成英文字幕 srt 文件、视频标题、简介、点赞量、评论区文本、发布频道、发布时间;

核心脚本核心逻辑:

```
python
```

```
import yt_dlp
```

```
import os
```

```
def download_youtube_videos(search_keyword, save_dir):
```

```
    ydl_opts = {
```

```
        'outtmpl': os.path.join(save_dir, '%(title)s.%(ext)s'),
```

```

        'writesubtitles': True,
        'subtitleslangs': ['en'],
        'writecomments': True,
        'sleep_interval': 3,
        'random_sleep_interval': [1,4],
        'user_agent': 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 Chrome/120.0.0.0 Safari/537.36',
    }
    with yt_dlp.YoutubeDL(ydl_opts) as ydl:
        ydl.extract_info(f'ytsearch50:{search_keyword}', download=True)

if __name__ == "__main__":
    download_youtube_videos("Village Super League Rongjiang Guizhou", "./youtube_raw")

```

2.2.2. TikTok 国际版话题数据采集

采集目标为#VillageSuperLeague 话题下短视频链接、英文文案、评论文本、画面配图、互动数据；但 TikTok 内置 XBogus 加密签名校验，采用 tiktokapipy 封装接口自动生成签名，无需手动逆向算法。

2.2.3. 境外媒体英文网页静态文本爬虫

采集目标为 BBC、China Today、华尔街日报村超专题网页正文、图片 alt 英文说明、记者导语、赛事文化相关描述文本。批量导入存储目标 URL 的 csv 清单，循环发起 GET 请求，soup 解析段落、img 标签 alt 属性，过滤广告、导航栏无关文本，仅保留正文叙事内容，单网页文本单独存储为 txt 文件，文件名绑定网页原始链接溯源标识[3]。本环节累计抓取有效外媒报道 71 篇，全部保留原始网页 URL 用于数据核验[4]。

2.2.4. 国内官方双语外宣文本定向抓取

采集对象为贵州省政府英文官网双语新闻、CGTN 村超专题稿件、海外先导片双语字幕、黔茶出口英文包装文案；技术方案分为两类，静态网页采用 bs4 爬虫批量抓取；官方线下电子版宣传册通过本地读取 PDF，使用 PyPDF2 提取中英文对照文本，按“中文原文-官方英译”成对存储，形成双语平行文本基础样本，累计采集双语文本 1268 组，包含“鱼钩茶”“拦门酒”“板鞋竞速”等文化负载词原始直译素材。

2.3. Python 自动化多模态数据清洗标准化流程

原始采集数据存在大量噪声，比如视频损坏文件、空白字幕、重复评论、网页广告文本、模糊无效图片、乱码字符、格式不统一问题，人工清洗工作量巨大，本研究开发完整自动化清洗脚本，分四大模态分层处理，全部通过 Python 实现，清洗日志自动留存便于溯源复核[5]。

2.3.1. 文本模态清洗脚本流程

使用 pandas drop_duplicates 删除完全重复的评论、报道文本，利用 os.path.getsize 识别 0 kb 空白 txt 文件，批量删除。正则表达式过滤网址、@用户标签、表情符号、数字无关弹幕；英文全部转为小写，去除多余标点、换行符，中文文本统一 UTF-8 编码，修复网页抓取产生的乱码字符。按“外媒报道/社交平台评论/官方双语文本/产品包装文案”四类自动移动至对应文件夹，生成清洗记录 log.txt，标记删除文件

名称与删除原因[6]。

2.3.2. 图像模态校验与标准化处理

- (1) 图像完整性校验: Image.open().verify()识别损坏、无法读取的图片文件, 直接标记删除;
- (2) 尺寸归一化: 统一缩放至 1920 * 1080 标准分辨率, 统一存储为 PNG 无损格式;
- (3) 无效过滤: 灰度模糊、纯黑纯白无内容截图批量筛除, 保留包含苗绣、银饰、茶山、赛事舞台、拦门酒仪式的有效视觉符号图像, 清洗后留存有效图像 1642 张。

2.3.3. 音视频模态降噪与格式统一(FFmpeg 调用)

通过 Python subprocess 调用 FFmpeg 命令行批量处理原始视频, 视频统一编码 H.264, 封装 MP4 格式, 剔除损坏无法解码的视频文件[7]。从每条视频分离音轨保存 WAV 无损音频, 用于 ELAN 音频分层标注。srt 字幕与视频文件同名绑定, 自动校正字幕时间轴偏移, 删除无字幕纯画面无效样本。

2.3.4. 清洗后数据集规模真实统计

经过完整自动化清洗后, 最终入库标准化多模态样本总量:

- (1) 视频素材: 有效短视频 756 条(YouTube427 条 + TikTok329 条), 总时长约 186 小时;
- (2) 音频素材: 配套分离 WAV 音频 756 份;
- (3) 双语平行文本: 有效对照文本 1268 组, 海外用户评论文本 21437 条, 外媒英文报道正文 71 份;
- (4) 图像素材: 标准化视觉符号图片 1642 张[8]。

所有清洗后文件统一存储于本地结构化语料库文件夹, 每条样本文件名附带唯一采集 ID、原始数据来源链接, 可随时反向溯源核验数据真实性。

3. 多模态语料库搭建与 ELAN 分层符号量化标注实操

3.1. 标准化双语多模态语料库本地存储架构设计

基于清洗后多模态素材搭建本地离线语料库, 采用分层文件夹架构, 无需云端服务器, 普通工作站即可稳定读取, 目录结构完全标准化:

Plain Text

cunchao_multimodal_corpus

```

├─ video_raw 清洗后标准化 MP4 视频
├─ audio_wav 配套无损音频
├─ subtitle_srt 英文字幕+中文对照文本
├─ text_bilingual 双语平行文本库
│   └─ official_publicity 官方外宣双语文案
│   └─ foreign_media_report 外媒英文报道
│   └─ social_comments 海外平台用户评论
│   └─ tea_product_copy 黔茶出口英文包装文本
├─ image_symbol 民族文化符号图像库
│   └─ miao_silver 苗族银饰纹样
│   └─ dong_music 侗族大歌、芦笙演奏画面
│   └─ fishhook_tea 鱼钩茶茶山、制茶工艺图
│   └─ lanmen_wine 拦门酒仪式实拍图

```

- └─ elan_project ELAN 标注工程文件(.eaf)
- └─ antconc_text 预处理纯文本(用于词频分析)
- └─ data_log 采集清洗全流程日志、原始 URL 清单

文本库统一导出 UTF-8 无 BOM 编码 txt 文件, 单文件控制在 5000 字符以内, 适配 AntConc 批量读取; 媒体文件与文本标注文件一一绑定唯一 ID, 实现视频、音频、图像、文本跨模态关联检索。

3.2. ELAN 多模态标注软件全流程实操(量化文化符号核心工具)

ELAN (Eudico Linguistic Annotator)由马克斯·普朗克心理语言研究所开发, 是多模态话语分析领域标准化标注工具, 支持 0.01 秒精度时间轴分层标注, 可量化统计视频内视觉、听觉、语言符号出现频次、持续时长、同步搭配关系, 解决传统人工统计符号数量误差大、无法记录时序协同关系的实操痛点。本研究完整落地从软件配置、标注层搭建、批量标注、数据导出全流程实操[9]。

3.2.1. ELAN 环境配置与标注分层模板设计

从 ELAN 官网下载 6.4 稳定版本, 关联本地 VLC 播放器保证多格式音视频兼容, 首选项设置标注语言中英文双显示, 时间轴精度 0.01s; 采用分层标注体系:

第一层【语言文本层 Text-EN】: 同步标注视频内英文字幕、海外解说文本, 重点标记“Fish Hook Tea”“Door-blocking Wine”等直译文化负载词出现时间节点;

第二层【视觉符号层 Visual-Symbol】: 细分 4 个子标注通道: 苗绣银饰纹样、茶山制茶画面、拦门酒仪式、足球赛事通用画面, 记录符号出现起止时间、画面占比;

第三层【听觉符号层 Audio-Sound】: 细分 2 个子通道: 侗族大歌人声、芦笙器乐演奏, 标记音频符号持续时长、与画面同步匹配区间。

每条视频单独生成.eaf 标注工程文件, 文件名与对应视频 ID 绑定, 所有标注工程统一存入语料库 elan_project 文件夹, 支持重复编辑、版本备份。

3.2.2. 批量标注实操步骤与量化数据导出

选取样本标准, 从 756 条短视频中分层抽样 120 条覆盖开幕式、茶产业宣传、海外游客互动、外媒现场拍摄四大场景, 完成全量分层标注, 单视频平均标注耗时 22 分钟, 整套标注累计实操时长 44 小时[10]。关键操作步骤为:

- (1) 新建工程文件, 导入对应 MP4 视频与配套 WAV 音频, 系统自动生成基础时间轴;
- (2) 按三层分层模板依次新建标注通道, 删除默认空白 default 层;
- (3) 播放视频至文化符号出现节点, 拖动鼠标框选对应时间区间, 输入标注标签(如标签“鱼钩茶 - 直译 Fish Hook Tea”“苗银头饰 - 视觉符号”“芦笙独奏 - 听觉符号”);
- (4) 全部标注完成后, 使用 ELAN 内置统计功能: 检索各标签总出现频次、单条视频平均出现时长、视听符号同步共存片段总时长;
- (5) 导出标注统计报表为 csv 格式, 直接对接 Excel、SPSS 开展后续量化分析。

3.2.3. ELAN 标注量化核心数据结果

苗绣银饰类符号 1872 次, 鱼钩茶制茶画面 964 次, 拦门酒仪式镜头 619 次; “Fish Hook Tea”鱼钩茶直译字幕 326 处, “Door-blocking Wine”拦门酒直译 189 处。视听符号同步搭配片段仅占总符号时长 21.7%, 78.3%场景存在画面、音频、字幕叙事割裂, 直接量化验证前期多模态协同不足的实操痛点。112 处英文解说仅直译文化名词, 无工艺、礼仪背景补偿说明, 为后续文化补偿翻译策略提供量化数据支撑[11]。

4. AntConc 文本偏差挖掘与 SPSS 海外受众传播效能量化评估

4.1. AntConc 语料库工具实操：文化负载词翻译偏差自动化检索

4.1.1. 预处理文本适配操作

将双语平行文本拆分独立中文语料库、英文资料库，中文文本使用 jieba 分词添加空格分隔(适配 AntConc 中文检索逻辑)，英文文本统一小写、去除停用词，分别导入软件两个语料库窗口，启用 Keyword List 对比功能，提取英译特有高频文化词汇[12]。

4.1.2. 四大核心功能实操与偏差挖掘结果

(1) Word List 词频统计：导出英文译本高频词汇排序表，筛选文化专有名词 Top10: Fish Hook Tea、Door-blocking Wine、Miao Silver、Dong Grand Chorus、Ban Shoe Race；统计显示 5 个核心民族文化词汇均采用纯直译/音译，无配套解释性文本；

(2) Concordance 上下文检索(KWIC 索引)：检索“Fish Hook Tea”全部 326 条上下文，94.8%语境仅出现名词直译，无关于钩形制茶工具、本地生态茶的补充描述，海外文本仅将其描述为普通茶叶，丢失文化内涵；检索“Door-blocking Wine”189 条语境，76.7%外媒解读为“阻拦客人的酒水”，完全误解苗族迎宾礼仪内涵；

(3) Collocates 搭配分析：计算 MI 互信息值，直译词汇高频负面搭配：confusing、strange、hard to understand，量化证明直译模式显著提升海外受众认知障碍；

(4) N-Grams 短语检索：提取官方外宣重复直译短语，批量生成偏差词汇清单，作为后续文化补偿翻译优化样本库。

4.2. SPSS 26.0 海外受众问卷量化评估全流程实操

4.2.1. 问卷数据录入与预处理实操

(1) 数据录入：每份问卷唯一编号，变量定义：性别、国籍、是否接触过中国民族文化、三大评估维度 12 道量表题、开放式文化误解记录；

(2) 缺失值处理：SPSS 缺失值分析，13 份问卷少量空项采用序列均值填充，无大面积缺失样本；

(3) 异常值筛选：通过箱线图识别极端量表分值，核查原始问卷确认非填写失误后保留样本；

(4) 信度检验(Cronbach's α 系数)：整体量表 $\alpha = 0.872$ ，三大维度分项 α 系数均高于 0.8，证明问卷内部一致性良好，数据具备统计分析价值。

4.2.2. 核心统计分析实操与量化结论

信息准确度均值 2.16(满分 5)、文化接受度均值 2.31、品牌记忆度均值 1.98，三项分值均低于中间值 3，量化证明现有直译 + 单一短视频传播模式海外传播效能偏低。文化负载词直译困惑度与文化接受度 Pearson 相关系数 -0.743 ($p < 0.001$)，呈显著强负相关，直译偏差越大，受众接受程度越低；视听符号协同完整性与品牌记忆度相关系数 0.681 ($p < 0.001$)，多模态叙事协同越强，海外受众品牌记忆效果越好。东南亚受众文化接受度均值 2.67，欧美受众均值 1.89，独立样本 T 检验存在显著差异，证明不同文化圈层受众适配性需求存在分层，海外传播需分区域定制多模态内容。

4.3. 全球平台标准化逻辑与本土文化能动性的博弈及理论补充

1) 全球平台侧的结构约束：YouTube、TikTok 为适配全球通用受众审美，形成一套标准化短视频生产逻辑：短时长、强视觉冲击、简单直白名词标注、碎片化叙事。该全球媒介框架直接造成本研究量化得出的 78.3%视听叙事割裂现象：平台算法优先抓取单一、猎奇式文化符号(苗银、拦门酒)提升流量，

不支持完整、长时序的文化礼仪、工艺叙事。

2) 结合 ELAN 标注数据来看, 平台流量导向下, 视频仅截取“拦门酒敬酒瞬间”“银饰特写”等碎片化视觉符号, 字幕仅直译专有名词, 删除礼仪起源、社群文化内涵, 形成“符号景观化”编码模式。从跨文化解读视角分析, 这种全球平台主导的碎片化编码是受众产生认知偏差的核心根源: 外媒、海外用户仅接收孤立文化名词, 无法理解符号背后侗族、苗族在地生活逻辑, KWIC 检索中“confusing、strange”等负面搭配高频出现, 正是全球媒介逻辑压缩本土文化完整叙事空间的直接体现。

3) 地方文化主体的能动反抗: 数据样本中, 民间游客、本土茶企自发创作的视频出现视听同步叙事比例达 42.6%, 远高于机构产出素材的 19.1%。民间创作者会主动搭配茶山完整工艺流程、完整拦门酒礼仪全程镜头, 自主补充文字注释解释文化内涵, 体现地方群体依托自身文化资源, 主动修正全球平台碎片化编码缺陷的能动性。现有全球在地化理论多强调全球化与地方文化二元对立, 本研究基于量化数据补充修正: 数字媒介场景下, 二者并非单向压制关系, 多元在地主体(民间个体、本土商业品牌)可通过自主内容生产, 对冲全球平台标准化叙事带来的文化消解, 民间传播力量是平衡全球在地张力的关键中介变量, 该结论可完善数字全球在地化理论的主体分析框架。

4) 双向张力下跨文化符号编码的核心矛盾: 全球平台追求低门槛、普适化、猎奇化流量叙事, 倾向简化、割裂少数民族文化符号; 而在地文化主体需要完整、连贯的叙事完成本土生活文化的平等跨文化对话。二者博弈下, 单纯直译名词的编码方式会放大文化隔阂, 过往翻译学中“文化补偿翻译”的表述隐含“本土文化需要额外补充才能适配海外受众”的单向预设, 带有文化迎合倾向。因此本研究重构表述为跨文化语境适配式翻译, 不再预设本土文化存在理解缺陷, 而是基于平等跨文化对话视角, 通过背景补充、视觉配套, 搭建不同文化圈层受众之间的认知桥梁, 消解单向输出的不对等关系。

5. 多模态国际传播创新方案技术落地与对照测试

5.1. 技术方案一: 文化补偿翻译标准化文本生成工具流程

针对“鱼钩茶”“拦门酒”等文化负载词直译偏差问题, 搭建标准化翻译操作流程, 配套 Python 辅助文本生成脚本自动添加文化补偿注释。基础翻译规则采用“音译/直译 + 工艺/礼仪场景注释 + 配套流程图视觉文本”三层补偿模式。如示例优化译法, 原直译: Fish Hook Tea, 优化补偿译法: Yugou Tea (Hook-shaped tea-making craft unique to Rongjiang, ecological alpine tea cultivated by Dong villages), 配套短文本注释绑定茶山工艺插图; 原直译: Door-blocking Wine, 优化补偿译法: Miao Gate-welcoming Wine (12 ritual toasts for honored guests, symbolizing ethnic hospitality), 配套牛角敬酒仪式静态配图说明:

输入中文文化名词, 自动调取本地文化知识库匹配标准化英文补偿注释, 批量改写原有外宣双语文本, 输出优化后平行文本库, 直接用于海外官网、产品包装、短视频字幕替换。脚本读取文本关键词, 自动从语料库 image_symbol 文件夹调取对应工艺、仪式配图, 实现“文本注释 + 图像符号”自动绑定, 降低人工制作成本[13]。

5.2. 技术方案二: AR + 游戏化多模态叙事短视频制作实操流程

5.2.1. “村超文化解码”AR 科普短视频制作流程

从语料库提取 ELAN 标注筛选的高频次民族符号镜头(苗绣、制茶、芦笙演奏)。使用剪映国际版、Spark AR 工具叠加动态工艺分解动画, 例如鱼钩茶视频插入钩形制茶工具 3D 动态模型, 同步字幕播放文化补偿翻译文本, 音频叠加简化版侗族大歌片段。严格匹配 ELAN 三层标注时序逻辑, 画面视觉符号、音频民族器乐、文本补偿字幕同步出现, 消除叙事割裂问题。依托 Python 批量剪辑脚本统一画幅、字幕字体、背景音乐音量, 单日可标准化产出 15 条 1 分钟海外短视频。

5.2.2. “侗族大歌节奏游戏”轻量化跨界传播素材开发

从语料库 WAV 音频截取侗族大歌、芦笙节奏片段, 通过 Python librosa 库提取节拍特征, 生成节奏关卡。固定 15 秒竖屏 TikTok 适配尺寸, 画面同步展示苗银服饰动态纹样, 用户跟随芦笙节奏点击屏幕完成互动。全程使用文化补偿翻译优化文本, 在游戏间隙弹出民族文化简短注释; Python 脚本批量生成带#VillageSuperLeague、#ChineseEthnicCulture 标签的发布文案, 自动定时分发至测试海外账号。

5.3. 优化前后传播数据对照测试实操与量化结果

选取 2 个同量级 TikTok 测试账号开展为期 30 天对照实验: 对照组投放原有直译单一短视频, 实验组投放优化后 AR 文化解码 + 游戏化多模态素材, 通过 Python 爬虫每日自动抓取两组账号播放量、点赞、评论、转发数据, 使用 SPSS 进行独立样本 T 检验对比差异:

(1) 播放量均值: 实验组单条均值 12476 次, 对照组 4129 次, 差异显著($p < 0.01$);

(2) 正向评论占比: 实验组 82.3%, 对照组 47.6%, 大量海外评论提及“看懂茶文化、苗族礼仪”, 直译素材评论高频出现“难以理解”相关表述;

(3) 产品转化联动: 配套黔茶海外电商链接点击量实验组提升 196%, 证明多模态优化方案可同步赋能产业出海; 量化测试数据完整验证整套技术流程的实践落地价值, 形成可向贵州文旅、茶企交付的标准化多模态外宣操作手册。

6. 结束语

适配海外短视频平台的 Python 多模态爬虫管道, 可快速采集同类乡村赛事、非遗文旅海外传播素材。基于 ELAN 的民族文化符号标准化分层标注模板, 实现视听文本多模态协同失衡问题量化测算。“文化补偿翻译 + AR 多模态叙事 + 游戏化跨界传播”标准化落地操作规范, 直接服务贵州“干净黔茶全球共享”、村超文旅品牌出海实践。

基金项目

2025 年贵州省研究生科研基金项目, 项目名称: “全球在地化”视域下贵州“村超”民族文化国际传播的多模态话语构建与创新路径研究, 项目编号: 2025YJSKYJJ238。

参考文献

- [1] Robertson, R. (1995) Glocalization: Time-Space and Homogeneity-Heterogeneity. In: Featherstone, M., Lash, S. and Robertson, R., Eds., *Global Modernities*, Sage Publications Ltd., 25-44. <https://doi.org/10.4135/9781446250563.n2>
- [2] Sigismondi, P. (2011) The Digital Glocalization of Entertainment: New Paradigms in the 21st Century Global Mediascape. Springer.
- [3] 谯胜念, 安慧芳. 贵州“村超”“村 BA”在 YouTube 的传播: 效度、问题与对策[J]. 电子商务评论, 2025, 14(5): 1428-1442.
- [4] 石培华. 新媒体生态下旅游目的地品牌化过程与内在机理研究——以贵州“村超”为例[J]. 吉首大学学报(社会科学版), 2025, 47(1): 1-10.
- [5] 舒小林. 乡村体育赛事型网红旅游地长红机制研究——以贵州“村超”为例[J]. 经济地理, 2025, 45(9): 252-260.
- [6] 何天平, 蒋贤成. 数字全球在地化: 平台时代影视产业国际传播趋势探研[J]. 对外传播, 2023(7): 21-25.
- [7] 陆佳怡. 地方与全球: 贵州“村超”的国际传播路径[J]. 新闻爱好者, 2025(1): 32-34.
- [8] 高子涵. 基于 ELAN 的多模态视频话语分析实操规范[J]. 现代语言学, 2025, 13(8): 608-612.
- [9] 曹小杰, 戴梦雅. 短视频平台隐私政策的全球在地化——对抖音和 TikTok 的比较分析[J]. 传媒观察, 2023(10): 86-94.
- [10] 张业安, 陈伟伦. 从文化折扣到文化共鸣: 村超国际传播策略研究[J]. 广州体育学院学报, 2025, 45(6): 31-41+99.

-
- [11] 何江, 郭彬凌. 乡村振兴背景下贵州“村超”短视频传播策略研究[J]. 智慧农业导刊, 2025, 5(24): 11-14+18.
- [12] 张波. 贵州“村超”讲好中国多民族故事的三个面向[J]. 贵阳学院学报(社会科学版), 2025, 20(5): 93-99.
- [13] China Today (2026) Beyond the Goals: Football and the Making of a County.
https://www.chinatoday.com.cn/ctenglish/2018/sl/202606/t20260604_800438978.html