

大语言模型在小红书双语文案中的多模态翻译质量研究

刘轩彤 齐佩雅 刘正悦

南京师范大学中北学院 江苏 丹阳 212300

【摘要】：在小红书全球化进程加快的背景下，中英双语文案成为跨文化传播核心载体。研究运用语料库分析法，采用量化可读性评分 + 人工多维评估方法，以自建小红书中英双语文案与符号语料库为基础，通过对日常类、文化负载类、情感类文案的实例分析，考察大语言模型在社交媒体多模态翻译中的表现。研究发现，现有大语言模型多模态翻译存在语义偏差、文化错位等问题，其主要影响因素为文化知识、多模态协同、可读性适配与语境缺失。

【关键词】：大语言模型；多模态翻译；小红书；双语文案；语料

DOI:10.12417/2982-3846.26.02.020

1 引言

2025年初，小红书（RedNote）迎来Tiktok用户迁移潮，平台双语文案激增，跨文化需求上升。小红书的一键翻译功能难以处理中英混合及图片文字，无法满足多模态翻译需求。GPT-4、NoteLLM-2等大语言模型虽已初步实现文本到符号的跨模态转译，但现有研究多聚焦模型构建与通用评估，对社交媒体中网络热梗、文化负载词、表情符号引发的翻译偏差缺乏系统性实证分析，存在专用评估体系缺失和成因挖掘不足的缺口。为此，本研究以小红书双语文案为对象，聚焦大语言模型的多模态翻译质量与文化适配性，探讨其翻译效果、典型偏差、成因及优化策略；通过构建语料库并采用“量化评分+人工统计”路径，全面评估模型表现，为提升跨文化传播效能提供理论与实证支持。

2 文献综述

2.1 多模态翻译理论研究

胡壮麟（2007）谈语言学研究的跨学科倾向研究中指出多模态翻译理论突破了传统翻译研究仅关注纯文本语言转换的局限，将文本、图像、表情符号、排版符号等多种符号模态纳入研究范畴，认为翻译是不同模态符号协同共建意义的跨语言交际活动。研究发现在小红书这类社交媒体场景中，文案、表情符号与配图形成了紧密的表意整体。多模态翻译理论为分析这类复合文本的跨语言转换提供核心框架，也帮助研究者理解不同模态信息在翻译中的协同关系与表意偏差（Zappavigna&Doran,2025）。

2.2 翻译质量评估研究

翻译质量评估理论是衡量译文效果的核心依据。经典的翻译质量评估框架围绕准确性、流畅性、忠实性、文化适配性等

核心维度构建，能够从多个层面判断译文对源文本信息的传递效果。王华树和谢亚（2023）指出随着大语言模型在翻译领域的普及，社交媒体翻译场景对评估体系提出新的要求，在传统框架基础上补充情感一致性、风格贴合度、可读性等维度，构建面向小红书多模态文案翻译的综合评估标尺，为人工评估与量化分析提供标准依据。

2.3 大语言模型辅助翻译研究

2018年，GPT系列模型推出后，国内外机构陆续研发Flamingo、GPT-4等多模态模型，相关研究例持续推进。国内多模态研究起步较早，早期偏向数理与控制方向，之后逐步与语言学研究结合（程佩，2024）。大模型普及后，多模态翻译研究逐渐展开，李德凤等人（2025）在现有模型搭建与通用翻译基础上，提出要对垂直领域与场景化文本进行探索。

3 研究设计与语料库构建

3.1 语料来源与筛选

本研究语料均来自小红书平台公开的用户原创图文笔记，覆盖美妆、穿搭、美食、旅行、生活感悟、文艺分享等多个领域。语料均为用户自主创作的原生内容。研究优先选取含表情符号、特殊符号或图文语义呼应的多模态笔记，确保适配多模态翻译场景。研究兼顾文本类型多样性，避免样本同质化。经人工校验，确保最终语料库兼具代表性与可用性，为后续研究提供可靠支撑。

3.2 语料体系构建

结合小红书文案的内容特征与语言风格，本研究将筛选所得语料划分为三大类型，分别为侧重口语表达的日常生活实用类文案，涵盖好物测评、美食探店、旅行攻略等内容；融入网络热梗、地域方言及文学典故的文化负载表达类文案。

基金项目：本文系2025年南京师范大学中北学院大学生创新创业训练计划项目：从文本到符号：大语言模型在小红书双语文案中的多模态翻译实践与研究（项目编号：202513906009Y）研究成果之一。

作者简介：刘轩彤（2004-），女，汉族，江苏徐州人，本科，南京师范大学中北学院，研究方向：英语语言文学；齐佩雅（2005-），女，汉族，江苏淮安人，本科，南京师范大学中北学院，研究方向：英语语言文学；刘正悦（2005-），女，汉族，安徽阜阳人，本科，南京师范大学中北学院，研究方向：英语语言文学。

依靠表情符号、特殊符号与精简句式烘托情绪的情感与符号表达类文案。本研究结合可读性网站量化评分与人工评估开展综合分析，依托专业检测工具对文本的句子长度、词汇难度与句法复杂度进行客观测量，并由研究者人工校准网络流行表达、符号用语与文化内容的评估偏差。

4 大语言模型多模态翻译表现实证分析

4.1 整体翻译效果与可读性统计

通过对自建语料库进行汇总整合，数据显示日常生活类文案整体可读性平均在 85 分左右；文化负载类文案可读性得分大部分在 70~85 之间，平均分接近 80；而情感与符号类可读性得分最高，基本在 90 分以上。情感一致性方面，三类文案的源语言与目标语言情感基本适配。

翻译准确率方面，人工评估发现日常生活类文案与情感与符号类文案的准确率较高，未出现语义偏差或信息遗漏。但人工评估发现文化负载类中一些网络热词以及特定的文化词汇的翻译出现了一些问题。

4.2 分类型实例分析

表 1 日常生活类文案翻译分析

源语言	目标语言	易读指数
I sipped the milk tea slowly,savoring the sweet taste on my tongue.	我慢慢啜饮奶茶，细细品味舌尖的甜味。	88.2

如表 1 所示，在日常生活类文案翻译中，大语言模型对场景化、生活化表达的处理表现出较高的稳定性。以本条语料为例，从翻译准确率来看，译文完整传递了原文的核心信息：“sipped the milk tea slowly”被译为“慢慢啜饮奶茶”，既保留了动作的轻柔感，也符合中文表达习惯；“savoring the sweet taste on my tongue”对应为“细细品味舌尖的甜味”，精准还原了原文对味觉体验的细腻刻画。从情感一致性来看，原文通过“slowly”“savoring”等词传递出一种悠闲、享受的松弛感，译文选用“慢慢”“细细品味”两个叠词，在节奏与语感上与原文形成呼应，自然传递出同样的惬意氛围，实现了情感层面的等效转换。该译文句子简短、用词通俗，可读性表现优异，说明小红书在处理这类低文化负载、高场景化的日常文案时，能够同时兼顾信息准确性、情感传递与阅读流畅度，整体翻译质量较高。

表 2 文化负载类文案翻译分析

源语言	目标语言	易读指数
江湖气	a sense of chivalry	75.88
人们到江湖中游泳。	People swim in the Jianghu.	100.24

然而，在翻译文化负载类文案时（表 2），小红书对典型文化意象的处理存在明显缺陷。第一条语料从翻译准确率来看，“江湖气”承载着中式语境中重情义、懂变通的市井侠气与社会交往特质，而“chivalry”特指西方骑士精神，二者文化内涵差异显著，译文窄化了源语语义，属于文化误读；情感一致性上，“江湖气”自带洒脱、烟火气的松弛感，而“chivalry”带有庄重、刻板的骑士文化色彩，未能传递原文的情感基调；第二条语料从翻译准确率来看，源语“江湖”指江河湖海等自然水域，译文直接音译为“Jianghu”，易让不了解中文文化的读者联想到武侠语境中的“江湖”，造成严重语义误解；情感一致性上，原文是生活化的日常表达，译文反而增添了不必要的文化神秘感，偏离了原文平实的语气。

这两条语料的系统可读性评分尚可，但表意偏差使其无法实现跨文化交际的等效传递，凸显了大语言模型对文化负载词多义性的语境识别不足。

表 3 情感与符号类文案翻译分析

源语言	目标语言	易读指数
{ \ / } (. - .) / \ 🍰	祝你生日快乐!	75.88

在情感与符号类文案案例中（表 3），模型对多模态符号的解读与转换表现良好。源语由颜文字与生日蛋糕 emoji 构成，译文精准还原了祝福场景的核心表意，翻译准确率较高。在情感一致性上，译文延续了源语轻松俏皮的祝福氛围，传递出温暖真诚的情感基调，与多模态符号的表意功能高度契合。译文句式简短、用词通俗，可读性良好，整体实现了符号文本向自然语言的顺畅转换，体现了模型对这类社交场景多模态情感符号的有效理解与适配处理。

5 翻译问题成因与典型偏差总结

基于对小红书双语文案语料库的系统分析，结合量化评分与人工评估结果，本研究归纳出大语言模型在多模态翻译中的四大典型偏差及其成因。

5.1 文化知识缺失

模型训练数据以主流语言文化为主，对非主流及亚文化覆盖不足，难以准确传递“江湖气”等文化内涵深厚的词语，反映出跨文化语义对齐的缺失。

5.2 多模态协同不足

在“文本+符号/表情”类文案中，模型常忽略符号的语义补充功能，仅翻译文本，如将“竹子”意象配以四叶草符号，导致

图文语义割裂与文化符号错位。

5.3 可读性适配不足

文化内涵丰富的长文本译文可读性偏低，模型处理复杂句法、网络热梗时输出流畅度下降，且传统指标难以适用于表情符号、缩略语等社交媒体表达，易生成语义模糊或风格不协调的译文。

5.4 语境缺失

模型缺乏对社交语境、用户情感及文化背景的深度理解。处理短文本、强情感化内容时表现稳健，但面对长文本或多模态内容时，易出现语义偏差、情感弱化或语调失当。用户问卷显示近半数受访者偏好“归化翻译”，而模型默认采用直译策略，忽视了语境对翻译策略选择的影响。

综上，四类偏差相互交织，共同制约了模型在小红书多模态翻译中的整体表现，为后续优化策略提供了明确方向。

6 结语

本研究以小红书双语文案为对象，构建近千条“文本—符号”语料库，采用量化评分与人工统计相结合的方法，系统考察了大语言模型的多模态翻译表现。结果表明，模型处理短文本与强情感化内容时语义传递较为稳健，但在高文化负载词、多模态协同表达及复杂语境下，普遍存在文化意象丢失、符号选择偏差及可读性适配不足等问题。因此，在未来特定场景中，大语言模型辅助翻译需要更多关注文化知识、多模态协同、可读性适配与语境缺失这几个影响因素。

参考文献：

- [1] 程佩.(2024).基于预训练模型的神经机器翻译关键技术研究[博士学位论文].郑州轻工业大学.
- [2] 胡壮麟.(2007).谈语言学研究的跨学科倾向.外语教学与研究,39(6),403 - 408.
- [3] 李德凤,王华树,刘世界.(2025).国家翻译技术能力与大语言模型.上海翻译,(2),18 - 24.
- [4] 孙逸群,周敏康.(2025).机器翻译错误类型与译后编辑方法研究——以《中国新经济:创新与规则》(节选)汉译英为例.翻译实践研究,(7),45 - 58.
- [5] 王华树,谢亚.(2023).ChatGPT 时代翻译技术发展及其启示.外国语言与文化,7(4),80 - 89.
- [6] 赵军,曹鹏飞.(2023).大语言模型——人工智能发展史上的里程碑.新兴科学与技术优势,2(1),80 - 88.
- [7] Zappavigna,M.,&Doran,Y.J.(2025).Emoji as interpersonal resources in LLM chatbot conversations:A social semiotic analysis of tenor and affiliation in human - AI interaction.Social Semiotics.