

多模态视角下动画短片《浪浪山小妖怪》的话语分析研究

王玉霞, 赵凤宇, 李佳暄

河北大学外国语学院, 河北 保定

收稿日期: 2026年4月13日; 录用日期: 2026年7月8日; 发布日期: 2026年7月21日

摘要

信息高速发展时代的到来, 改变了人们获取和传播信息的方式, 从单一的语言形式转变为包含图像、声音和手势等的多模态形式。鉴于此种情况, 多模态话语分析应运而生, 并被用于分析和解读由多种模态元素组成的复合文本。动画电影《浪浪山小妖怪》是《中国奇谭》系列的首部衍生动画作品, 以《西游记》为背景, 讲述了四只草根小妖怪因意外得罪大王而被迫出逃, 组队假扮唐僧师徒前往西天取经的荒诞冒险故事。该片通过整合文本、图像和声音等各种模态资源, 构建了一个关于“无名之辈”坚持良善、追寻自我价值的连贯叙事。本论文基于Halliday的系统功能语法、Kress和van Leeuwen的视觉语法、van Leeuwen的听觉语法以及张德禄的综合理论框架, 构建了一个研究框架。采用定性与定量相结合的方法, 旨在探讨以下两个问题: 在《浪浪山小妖怪》中, 语言、视觉和听觉模态是如何参与意义建构的; 在意义建构过程中, 语言、视觉和听觉模态之间存在什么关系。通过对动画《浪浪山小妖怪》进行多模态话语分析, 本论文发现, 语言模态通过概念功能、人际功能和语篇功能建构意义, 视觉模态通过再现意义、互动意义和构图意义完成意义建构, 其中影片独特的“仿冒取经”视觉符号系统以戏仿和拼贴为手法, 将传统西游元素进行草根化、喜剧化转译, 构成了其视觉叙事的核心创新。听觉模态通过声音视角、社会距离和情态等方面建构意义, 尤以民族乐器与西方管弦乐的混搭运用, 创造出兼具东方韵味与史诗感的独特听觉风格。在动画中, 不同模态之间存在互补关系下的强化和非强化关系, 以及非互补关系下的重叠和语境交互关系。在这些关系的综合作用下, 共同呈现了完整的情节和“坚持良善、为无名之辈正名”的深刻主题。

关键词

多模态话语分析, 系统功能语法, 视觉语法, 听觉语法

A Discourse Analysis of the Animated Short Film *Nobody* from a Multimodal Perspective

Yuxia Wang, Fengyu Zhao, Jiaxuan Li

School of Foreign Studies, Hebei University, Baoding Hebei

文章引用: 王玉霞, 赵凤宇, 李佳暄. 多模态视角下动画短片《浪浪山小妖怪》的话语分析研究[J]. 现代语言学, 2026, 14(7): 468-478. DOI: 10.12677/ml.2026.147668

Received: April 13, 2026; accepted: July 8, 2026; published: July 21, 2026

Abstract

The advent of the era of rapid information development has transformed the ways in which people acquire and disseminate information, shifting from a monomodal linguistic form to a multimodal form encompassing images, sounds, gestures, and other modes. In light of this situation, Multimodal Discourse Analysis emerged and is utilized to analyze and interpret composite texts consisting of multiple modal elements. The animated film *Nobody*, the first derivative animation of the *Yao-Chinese Folktales* series, is set against the backdrop of *Journey to the West*. It tells the story of four grassroots little monsters who, after accidentally offending their king, are forced to flee and form a fake team impersonating Tang Monk and his disciples on a journey to the West to obtain scriptures. Through the integration of various modal resources such as text, images, and sound, the film constructs a coherent narrative about “nobodies” who insist on kindness and pursue self-worth. This thesis constructs a research framework based on Halliday’s Systemic Functional Grammar, Kress and van Leeuwen’s Visual Grammar, van Leeuwen’s Auditory Grammar, and Zhang Delu’s comprehensive theoretical framework. Employing a combination of qualitative and quantitative methods, it aims to explore the following two research questions: how the verbal, visual, and auditory modalities contribute to the construction of meaning in the animation *Nobody*, and what is the relationship between the verbal, visual, and auditory modalities in the process of meaning construction in this animation. Through a multimodal discourse analysis of the animation *Nobody*, this thesis finds that the verbal modality constructs meaning through the ideational, interpersonal, and textual functions. The visual modality accomplishes meaning construction through representational meaning, interactive meaning, and compositional meaning. The unique “fake pilgrimage” visual symbol system, through parody and collage techniques, translates traditional *Journey to the West* elements into a grassroots, comedic adaptation, forming the core innovation of its visual narrative. The auditory modality constructs meaning through aspects such as sound perspective, social distance, and modality, particularly with the fusion of traditional Chinese musical instruments and Western orchestral music, creating a distinctive auditory style that blends Eastern charm with epic grandeur. In the animation, there exist reinforcing and non-reinforcing relationships under complementary relations, as well as overlapping and contextual interaction relationships under non-complementary relations. Under the combined influence of these relationships, a complete storyline and the profound theme of insisting on kindness and giving recognition to nobodies are collectively presented.

Keywords

Multimodal Discourse Analysis, Systemic Functional Grammar, Visual Grammar, Auditory Grammar

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. 引言

1.1. 研究背景

随着信息技术的飞速发展，人们的日常交流方式发生了深刻变化。信息获取不再仅依赖于单一的语言形式，传播媒介也突破了文本的局限。图像、声音、手势等多种符号资源日益成为有效的交流手段，信息的载体正从单一模态向多模态转变。这一趋势催生了多模态话语分析这一新兴研究领域。多模态话语分析突破了传统语言学局限于单模态语篇分析的框架，自二十世纪九十年代兴起以来，已在广告、网

站、纪录片、城市宣传片、电影等多个领域取得丰硕成果[1]。然而, 现有研究多集中于静态语篇, 对动态语篇中听觉模态的关注相对薄弱, 且动态文本的研究也主要聚焦于语言与视觉层面。因此, 对动态视频开展系统的多模态分析具有重要的学术价值。

动画作为一种融合文本、图像、声音等多种元素的动态影像形式, 其创作不仅依赖故事情节, 更深受创作者所处文化背景的影响。动画的传播既依赖内容本身, 也承载着文化输出的功能。作为文化传播的重要载体, 动画在推动中国文化走向世界方面扮演着日益关键的角色。《浪浪山小妖怪》是 2023 年现象级动画短片集《中国奇谭》首集《小妖怪的夏天》的衍生电影, 于 2025 年 8 月 2 日正式上映。该片由上海美术电影制片厂与哔哩哔哩联合出品, 监制陈廖宇、导演於水携原班团队历时两年打造。影片以《西游记》为背景, 讲述了四只草根小妖怪——小猪妖、蛤蟆精、黄鼠狼精和猩猩怪——因意外得罪大王而被迫出逃, 组成“草根仿冒取经团队”, 假扮唐僧师徒前往西天取经的荒诞冒险故事。上映后, 该片累计票房突破十七亿元, 刷新了中国影史二维动画票房纪录, 引发广泛社会讨论。通过整合文本、图像、声音等多模态资源, 《浪浪山小妖怪》构建了一个关于“无名之辈”在困境中坚守良善、追寻自我价值的动人叙事, 延续了短片“打工人”的现实底色并拓展了更深层的主题意涵。基于上述背景, 本论文选取《浪浪山小妖怪》作为研究对象, 运用多模态话语分析理论对其进行系统考察。

1.2. 研究目的及意义

本论文构建了一个利用系统功能语法、视觉语法、听觉语法和张德禄的综合理论框架的分析框架, 以《浪浪山小妖怪》作为案例研究。本论文的目的是准确分析动画中意义的建构, 并使用 AntConc、Elan 和 Praat 软件进行统计分析和标注, 检验不同模态之间的关系。

本论文在理论上和实践上都具有重要意义。在理论重要性方面, 本论文旨在探索多模态话语分析理论在分析动画中的适用性和可行性, 从而扩展其研究领域。首先, 多模态话语分析理论与动画分析相结合, 可以更好地表达动画作品的意义和表达方式, 从而为动态语篇分析和动画分析在各自领域的发展提供支持。其次, 在动态语篇分析方面, 本论文应用多模态话语分析理论来解读动画作品中不同元素所扮演的重要角色, 并进一步推测导演或创作者想要传达给观众的主题和思想。

在实践意义方面, 首先, 《浪浪山小妖怪》以其独特的“草根取经”设定和真挚的情感内核赢得了广大观众的喜爱。通过其荒诞而又充满温情的冒险故事, 它展示了中国文化中关于身份认同、集体协作与道德选择的深刻思考, 引发了观众对普通人如何在困境中保持善良这一普遍性命题的共鸣。其次, 《浪浪山小妖怪》注重群像刻画, 四只小妖各具性格缺陷, 却在冒险中互相扶持、共同成长, 其心路历程具有强烈的感染力。正如导演於水所言, 主创团队就是想通过这四个角色, 争取让每个观众都能在影片中看到自己。最后, 该片作为《中国奇谭》首部衍生电影, 成功探索了国产动画 IP 的衍生开发模式, 其票房与口碑的成功进一步扩大了中国原创动画在国际舞台上的影响力。

1.3. 论文结构

本论文分为五个部分, 第一部分为引言, 介绍了论文的背景、目的与意义以及结构。第二部分是文献综述, 分析了国内外多模态话语分析的研究现状。第三、四部分构成本论文的核心, 第三部分从语言模态、视觉模态和听觉模态三个视角审视视频内容, 第四部分深入探讨这些模态之间的关系。第五部分总结了本研究的主要发现和启示, 讨论了当前研究的局限性, 并提出了进一步探索的方面。

2. 文献综述

本章共分两个部分: 首先, 简要梳理多模态话语分析在国内外的研究现状; 其次, 对现有成果进行

总结与评价,并探讨从多模态话语分析视角分析动画《浪浪山小妖怪》的可行性。

2.1. 多模态话语分析研究综述

随着多媒体技术的不断发展,信息传播的方式和渠道发生了深刻变化。人们的交际方式越来越多样化,意义传递更多地通过多模态的方式实现。单一的语言分析已难以全面解释当代社会交际现象的整体语篇意义,学者们开始将研究视角转向多模态话语分析这一新兴领域。

2.1.1. 国外多模态话语分析研究

Harris 于 1952 年首次提出话语分析这一概念。然而,随着社会与科技的发展,仅对单一语言模态进行话语分析已显现出诸多局限,学者们开始探索对其他模态的分析方法。

Barthes 于 1977 年发表的论文《图像的修辞》被视为多模态话语分析早期的奠基性成果[2]。他在该文中深入探讨了图像在意义表达上与语言的相互作用,揭示了图像与语言之间并非简单的并列关系,而是存在着深层次的交互影响。他进一步提出了三种核心的图文关系,即图像为文字提供具体指代或背景的锚定关系、图像对文字内容进行视觉化解释或补充的说明关系、以及图像与文字在意义上相互衔接共同构建更完整信息的接递关系[2]。Barthes 的这些洞见为多模态话语分析的发展奠定了坚实基础。

二十世纪九十年代,多模态话语分析理论迎来了迅猛发展。O'Toole 于 1994 年出版的《展示艺术的语言》率先将系统功能语法应用于绘画、建筑及雕塑等艺术形式的研究,提出视觉模态具有与语言相似的三种功能,即再现功能、情感功能与构图功能[3]。此后,Kress 与 van Leeuwen 于 1996 年合著了《阅读图像:视觉设计语法》一书,该书以 Halliday 的系统功能语言学为理论基础,将语言的三大元功能扩展到视觉图像领域,构建了包括再现意义、互动意义和构图意义在内的视觉语法理论框架[4]。这一著作被广泛视为多模态话语分析领域确立的标志性成果,为后续研究奠定了坚实的理论基础。

与此同时,van Leeuwen 于 1999 年出版了《言语、音乐与声音》一书,构建了用于分析听觉符号的理论框架[5]。他将声音符号和语言符号一样视为具有意义潜势的资源,从声音视角、社会距离、音质、旋律等方面阐释声音符号的功能,奠定了听觉话语分析的基础,弥补了多模态话语分析在听觉模态研究方面的不足。

进入二十一世纪后,Norris 于 2004 年发展出一个综合性的多模态交互分析理论框架,主张所有交互行为均呈现出多模态的特性[6]。LeVine 与 Scollon 于同年提出了一种用于收集、整理与分析多模态话语数据的新型交际工具[7]。Jewitt 于 2009 年探讨了除语言之外是否存在其他可用于交际的符号资源[8]。O'Halloran 于 2004 年和 2005 年基于系统功能语言学视角,分析了数学与科学教科书中的语言、符号语法与视觉图像的结合,探讨了它们在意义构建中的作用[9][10]。Bateman 与 Schmidt 于 2011 年合著了《多模态电影分析》一书,从三大元功能意义的建构角度探讨电影语篇[11]。

近年来,国外多模态话语分析的研究呈现出显著的增长趋势,理论和应用研究并驾齐驱,研究领域和范畴不断拓展。从理论上,以系统功能符号学为基础的理论一直是研究重点;从实践上看,研究对象从书籍封面、海报等静态语篇扩展到电视、电影等动态语篇,跨学科研究趋势日益明显。

2.1.2. 国内多模态话语分析研究

与国际研究相比,国内对多模态话语分析及其相关领域的研究起步相对较晚。2003 年,李战子首次将多模态话语分析理论引入中国[12]。她在《多模式话语的社会符号学分析》一文中引用了 Kress 与 van Leeuwen 提出的视觉语法理论,系统阐述了多模态话语理解与生成的过程,标志着中国学者开始对多模态话语分析领域展开系统探索。

胡壮麟于 2007 年区分了多模态符号学和多媒体符号学,强调了多模态化是社会符号学发展的新趋

势,提出应重视培养人的多模态识读能力[13]。朱永生于同年对多模态话语的产生及定义进行了论述,分析了模态、模式和媒介的区别,提出了识别多模态话语的两个标准,即至少有两个感觉器官参与交流或话语中至少涉及两种符号系统[14]。

张德禄于2009年在系统功能语言学的基础上,尝试建立了多模态话语分析的综合理论框架,将模态间的关系分为互补关系与非互补关系两大类[1]。互补关系包括强化与非强化两类,非互补关系包括交叠、内包和语境交互三种形式。这一框架为多模态话语分析提供了重要的理论参考。此后,张德禄于2018年又提出了一种更加注重符号关系考察的改进路径,进一步推动了该领域的发展[15]。此外,韦琴红于2009年深入剖析了视觉模态中的图像、情态及印刷版式等要素[16]。冯德正于2014年和2017年系统梳理了多模态语篇分析的基本问题,介绍了视觉叙事分析框架,并指出了多模态话语分析未来的发展方向[17][18]。潘艳艳与李战子于2017年分析了2003年至2017年间发表于CSSCI来源期刊的相关论文,识别了多模态话语分析研究中存在的问题,并讨论了其发展轨迹[19]。潘艳艳于2020年对从多模态话语分析到多模态认知批评分析的理论演进及研究成果进行了系统回顾[20]。

在应用研究方面,国内学者将多模态话语分析理论广泛应用于海报、漫画、教材、纪录片、电影、电视节目、网络直播等多种语篇类型。在静态语篇研究领域,学者们主要应用视觉语法理论分析图文关系及其社会功用;在动态语篇研究领域,研究对象涵盖了纪录片、电影、电视节目、网络直播等多种形式。研究者们从多模态视角分析这些动态语篇中语言、图像、声音等模态的意义构建及其协同关系。

2.2. 小结

通过对多模态话语分析既往研究的系统回顾可以发现,国外学者在理论研究与实证研究方面均做出了重要贡献。从Barthes的图文关系理论,到Kress与van Leeuwen的视觉语法,再到van Leeuwen的听觉语法,多模态话语分析的理论框架不断完善。国内研究者在借鉴国际研究成果的基础上,结合本土语料开展了一系列实证研究,既验证了多模态话语分析理论的可行性与适用性,也拓展了其研究范围。

在梳理现有动画及影视作品研究文献时可以发现,多模态话语分析已被成功应用于电影、纪录片、电视节目等多种动态语篇的分析中。Kress与van Leeuwen的视觉语法理论和van Leeuwen的听觉语法理论为分析动态多模态语篇提供了坚实的理论工具。然而,目前从多模态话语分析视角对动画短片进行系统研究的文献相对较少,尤其缺乏对《浪浪山小妖怪》这一具有代表性的国产动画作品的多模态话语分析。

《浪浪山小妖怪》是2023年爆款动画短片集《中国奇谭》首集《小妖怪的夏天》的衍生电影,于2025年8月2日正式上映。该片以《西游记》为大背景,讲述了四只草根小妖怪因意外得罪大王而被追杀,被迫组成“草根仿冒取经团队”,假扮唐僧师徒前往西天取经的荒诞冒险故事。该片融合了语言、图像、声音等多种模态资源,通过多模态协同构建了关于“无名之辈”在困境中坚持良善、追寻自我价值的连贯叙事,具有丰富的多模态话语特征。因此,本论文旨在探索该动画中多种模态的意义建构方式,并考察在意义建构过程中各模态之间的关系,以填补相关研究的空白。

3. 《浪浪山小妖怪》多模态话语分析的模态功能分析

3.1. 语言模态的叙事构建与人际互动功能

根据Halliday的系统功能语法,语言模态通过概念功能、人际功能和语篇功能三大元功能驱动叙事发展[21]。与一般动画中语言模态仅承担情节推进功能不同,《浪浪山小妖怪》的语言模态在构建“草根逆袭”这一当代中国价值观主题上具有独特的叙事策略。在动画《浪浪山小妖怪》中,语言模态主要指角色对话的台词以及视频下方显示的字幕。本节将从及物性系统、语气与情态系统、主位结构三个维度,结合具体台词进行深入分析。

在概念功能层面,及物性系统是编码角色经历和外部世界事件的核心机制[21]。通过对影片台词的细致分析可以发现,物质过程、关系过程和心理过程三类小句使用最为频繁。影片大量使用了物质过程来推动情节,但其独特之处在于,这些物质过程的发出者始终是“小妖怪”这一边缘群体,而非传统西游故事中的英雄主角。物质过程是驱动情节发展的核心动力。影片开场,大王发现铜鼎上的字被擦掉后愤怒质问是谁把本王的字擦掉了,这一物质过程小句不仅交代了事件起因,更直接引发了后续的追查与惩罚,成为整个故事的导火索。随后,狼护卫命令小妖们去追捕并一定要把他们抓回来,其中的“追”和“抓回来”两个物质过程将叙事迅速推向逃亡阶段。在假扮取经段落,小猪妖说咱们分头行动,自己去找钉耙,让同伴去找木棒,这一物质过程小句交代了角色分工,推动了凑齐装备这一关键情节的发展,也标志着角色从被动受害者向主动行动者的转变,是草根主体意识觉醒的语言表征。

关系过程用于定义事物之间的属性或归属关系[21]。在这部影片中,关系过程承担着建构“身份政治”的独特功能。传统西游故事中,身份是固定不变的——孙悟空永远是齐天大圣,妖怪永远是反派。但《浪浪山小妖怪》通过关系过程的反复使用,质疑了这种身份的本质主义。小猪妖说“我们这种小妖怪,连大王洞都进不去”,这一归属式关系过程揭示了身份的社会建构性——不是因为他们生来就是底层,而是因为社会结构将他们排斥在外。妖王说“你们不过是几只蝼蚁”,则展示了权力者如何通过话语将阶层差异自然化、本质化。

心理过程表达角色的感觉、认知和情感状态[21]。心理过程在影片中也承载了“自我认同”的当代命题。小猪妖在集市上听说书人讲孙悟空的故事后喃喃自语“我也想成为那样的英雄”,这一心理过程标志着他从安于现状到萌发改变渴望的转折,是角色成长的情感起点。当蛤蟆精第一次穿上僧衣时,他对着镜子问自己是否真的像唐僧,这一疑问式心理过程表达了他对自我身份的怀疑与不安。而猩猩怪在最终对峙前内心独白“我不想再躲了”,这一认知过程预示了他即将爆发的反抗。而这些心理过程表明,他们并不是传统英雄叙事中的“我要变强大”,而是“我要停止自我否定”,这更贴近当代年轻人面对职场压力、社交焦虑时的心理状态。

在人际功能层面,语气系统和情态系统是实现人际意义的关键手段。语气系统决定了对话中的角色关系。影片中,大王和狼护卫大量使用祈使句,如“给我追”“抓回来”“闭嘴”等,这些祈使句没有主语、语气强硬,体现了绝对权威。在擦鼎段落,大王对小妖们说“明天之前,把铜鼎擦干净,否则……哼”,祈使句后的停顿与威胁语气将上层对下层的压迫感推向极致。相反,四只小妖之间多使用疑问句和陈述句。例如,小猪妖问同伴是否真的要西天,蛤蟆精反问同伴确定这样行不行,这些疑问句体现了平等协商的关系。在小雷音寺段落,猩猩怪第一次主动开口询问自己能否说一句话,这一疑问句以极其卑微的方式发出,却标志着他从沉默到发声的突破。

情态系统表达说话者对命题的确定程度。低值情态表示不确定,高值情态表示肯定。影片中,蛤蟆精问同伴“你确定我们这样行吗”,其中的“确定”虽属中高值情态,却以疑问形式出现,反映了角色内心的不安。而猩猩怪说“我信你”,这一低值情态的真诚表达,在高度原子化的现代社会背景下,传递了对信任与团结的呼唤,构成了对“精致利己主义”的文化批判。在影片结尾,小猪妖说“我们一定能走出去”,“一定”这一高值情态词出现在他口中,标志着他从最初的迷茫到最终坚定的成长。妖王说“你们必须死在这里”,“必须”作为高值情态与“死”搭配,将权力压迫推向极致,也为后续的反抗提供了张力。

在语篇功能层面,主位结构通过主位和述位的配置来组织信息。影片中,蛤蟆精说“唐僧……我演过”,主位是作为已知信息的“唐僧”被置于句首,述位“我演过”则是新信息,制造了喜剧效果——一个蛤蟆精竟然演过唐僧。又如,猩猩怪说“木棒……我拿得动”,主位“木棒”是前文刚出现的道具,述位“我拿得动”则传递了新信息,展现了他的决心。

3.2. 视觉模态的表意与情感渲染功能

根据 Kress 与 van Leeuwen 的视觉语法,视觉模态通过再现意义、互动意义和构图意义建构信息[4]。本节将结合具体镜头,深入分析影片在视觉设计上的独到之处,揭示影片视觉设计中区别于常规动画的民族美学特色与文化转译策略。

在再现意义层面,视觉模态分为叙事再现和概念再现两类。叙事再现依赖矢量表现动作与变化,概念再现则无矢量,表现稳定状态与分类[4]。影片大量运用叙事再现中的动作过程来讲述事件。在擦鼎段落中,小猪妖趴在巨大的铜鼎上,用鬃毛拼命摩擦,背上的皮肤逐渐磨破,血迹渗出。这一系列动作过程通过手臂的来回运动、血滴的下落等矢量,直观地展现了底层小妖的艰辛与自我牺牲。在逃亡段落,小妖们被狼护卫追赶时,他们奔跑、跌倒、爬起、再奔跑,这些动作过程的连续呈现营造了紧张感。反应过程在影片中也很常见,当猩猩怪第一次举起木棒时,其他三只小妖的目光齐刷刷地看向他,眼睛瞪大、嘴巴微张,这一反应过程展现了他人的震惊与期待。而妖王凝视猩猩怪的目光与他恐惧的表情构成了反向反应过程,揭示了权力关系的微妙变化。

影片最具独创性的是概念再现中的“戏仿式西游符号系统”。这不是对原著形象的忠实再现,而是一种后现代式的“降格”处理。影片多次呈现四只小妖笨拙地模仿唐僧师徒行走的远景镜头:蛤蟆精穿僧衣扮唐僧,猩猩怪拿木棒扮孙悟空,小猪妖扛钉耙扮猪八戒,黄鼠狼精闭嘴扮沙僧。这一视觉符号系统并非简单复制传统西游形象,而是采用了戏仿与拼贴的后现代手法。猩猩怪手中的木棒与孙悟空的金箍棒形成巨大反差,金箍棒重达一万三千五百斤、可大可小,是力量的象征,而猩猩怪手中的木棒是随手捡来的树枝,弯曲、粗糙。这种降格处理嘲弄了传统英雄叙事的不可及性,同时赋予小妖怪以“我也能当英雄”的可能性。蛤蟆精的僧衣与唐僧的锦斓袈裟形成反差,唐僧的袈裟镶嵌珠宝、光芒四射,而蛤蟆精的僧衣是破旧的粗布,穿在身上后被大肚子撑破,纽扣崩开。这一视觉符号既制造了喜剧效果,也消解了唐僧的庄严,暗示“成佛之路”的荒诞性。小猪妖的钉耙与猪八戒的九齿钉耙形成反差,九齿钉耙是神兵利器,而小猪妖的钉耙则是农家用的普通农具,齿已生锈、手柄断裂,这一视觉符号强调了小妖的草根身份,他们连像样的武器都没有,却要去取经。黄鼠狼精本是话痨,却为了扮演沙僧而强行闭嘴,嘴巴紧抿、腮帮鼓胀,这一视觉细节本身就是喜剧。

影片的视觉美术风格融合了中国传统水墨画的元素。背景中的山峦、云雾采用淡墨晕染技法,远山近树层次分明,留白处理富有诗意,这是对上海美术电影制片厂经典动画《大闹天宫》《小蝌蚪找妈妈》民族美学传统的直接传承。与写意背景形成对比的是,前景中的小妖造型采用写实风格,毛发根根分明、皮肤纹理清晰、表情夸张细致。这种写意背景与写实前景的视觉设计,使观众在东方诗意与喜剧张力之间获得独特的审美体验。

在互动意义层面,视觉模态通过接触、社会距离和视角三个维度实现与观众的关系建构[4]。接触方面,当四只小妖第一次集体面对真正的老鼠精时,他们惊恐地望向镜头,形成了索取类图像。观众与角色形成直接的目光对视,仿佛角色在向观众求助,瞬间拉近了心理距离。这种设计突破了传统动画中观众与角色的距离感,使“草根叙事”获得了情感共振的合法性。另一个典型的索取图像出现在猩猩怪准备喊出“我是齐天大圣”之前,他深吸一口气,直视镜头,眼神从恐惧转为坚定,这一索取类图像邀请观众见证他的蜕变。社会距离方面,影片多用远景和中景展现小妖们在荒野中前行的场景。例如,在一望无际的荒野上,四个小点缓慢移动,身后是连绵的群山,这种远景镜头让观众以旁观者的视角审视他们的冒险,产生距离感与渺小感。而在表现小妖们相互鼓励的时刻,则使用近景甚至特写。例如,篝火旁,四张小脸挤在画面中,彼此的呼吸声可闻,这种亲密距离强化了团队温情。视角方面,影片在展现小妖们时多采用平视视角,观众与他们处于同一高度,体现了平等关系。而在展现妖王时,多采用仰视视角,

观众必须抬头仰望妖王，暗示其权力高位。当猩猩怪最终举起木棒反抗时，镜头突然切换为仰视猩猩怪，这一视角的转变在视觉上宣告了权力关系的倒置。

在构图意义层面，视觉模态通过信息值、显著性和取景实现意义的组织[4]。信息值方面，在小雷音寺对峙的高潮场景中，妖王始终位于画面中心最高处，身形被光影放大，而四只小妖被置于画面下方且相互依偎，体型渺小。这种中心与边缘、大与小的构图格局直观地再现了权力结构中的核心与边缘地位。当猩猩怪喊出“我是齐天大圣”后，构图发生了戏剧性变化，猩猩怪走向画面中心，妖王被挤到边缘，这一视觉叙事无需言语便宣告了反抗的胜利。显著性方面，妖王的宫殿以深红色和金色为主色调，光影厚重，营造威严与压迫感。小妖们的场景则以青灰色、草绿色为主，光线柔和，传递朴素与生命力。在变装段落，小妖们穿上唐僧师徒的服装后，画面突然变得明亮，色彩饱和度的提升暗示了他们从无名小妖到角色扮演者的身份转变。取景方面，当小妖们被妖王的手下包围时，画面常常被妖王麾下妖怪的躯体所分割，粗壮的胳膊、巨大的阴影形成取景框架，将小妖们框在狭小的空间内，暗示了他们身陷囹圄的无助。而在突围段落，镜头拉远，取景框架消失，小妖们奔跑在开阔的原野上，这一视觉变化象征了他们从束缚中挣脱。

3.3. 听觉模态的氛围营造与主题强化功能

根据 van Leeuwen 的听觉语法，声音模态通过声音视角、社会距离、音质和旋律建构意义[5]。《浪浪山小妖怪》的听觉设计超越了常规动画“配乐烘托气氛”的功能层级，创造性地将民族乐器与西方管弦乐混搭，构建了一种“文化混响”的听觉美学。

在声音视角层面，声音分为角色音、地面音和场景音三类。角色音是听众必须注意的声音，地面音是背景声音，场景音是环境声音。影片开场，小猪妖刷锅的摩擦声作为角色音成为听觉焦点，洞穴中滴水的回声作为地面音提供了环境背景，远处若隐若现的妖怪交谈声作为场景音延伸了空间的纵深感。当大王的声音响起时，它迅速从地面音强化为角色音，而摩擦声则退为背景。这一层次变化引导观众将注意力从日常劳作转移到即将到来的危机。在社会距离层面，四只小妖在篝火旁商议计划时的对话音量适中，音质清晰且自然，体现了个人距离，观众感觉像在听朋友聊天。妖王在小雷音寺中讲话时，声音响亮且伴有低沉回声，模拟了公共距离，仿佛从高处传来的审判。

在音质与音色层面，声音可以传递情绪与性格。蛤蟆精的音色尖锐、语速快，体现了他作为吐槽役的急躁性格。猩猩怪的音色低沉、沙哑，伴有气息声，体现了他口吃社恐的怯懦。当猩猩怪最终喊出“我是齐天大圣”时，音质突然变得饱满、洪亮，气息声消失，共振增强，这一瞬间的音质转变象征了他从自卑到自信的蜕变。

在声音情态层面，影片最具独创性的设计是民族乐器与西方管弦乐的混搭运用。影片在温馨段落使用笛子、琵琶、古筝等民族乐器，营造东方田园诗意；在紧张段落加入西方管弦乐，营造史诗感；而在高潮段落，猩猩怪喊出“我是齐天大圣”时，背景音乐骤然转为激昂的鼓乐与铜管乐的宏大交响——其中鼓点节奏模仿了中国传统戏曲中的“急急风”，将东方打击乐的爆发力与西方交响乐的恢弘气势融为一体。这种“中西合璧”的声音设计，既不是简单的民族元素堆砌，也不是对西方配乐模式的照搬，而是一种有意识的“文化协商”，隐喻了当代中国在全球化语境中寻求文化主体性的努力。

4. 《浪浪山小妖怪》多模态话语的协同与升华

根据张德禄的多模态话语分析综合理论框架，模态间存在互补关系与非互补关系两大类[1]。互补关系指不同模态相互配合，共同表达完整意义，又可细分为强化关系与非强化关系。非互补关系指第二种模态对第一种模态的意义建构没有大的贡献，但仍作为一种模态出现，包括交叠、内包和语境交互等形

式。本节将从“新西游故事”的当代重构角度，分析多模态协同如何共同讲述一个具有当代中国价值观的故事。

4.1. 模态强化：共鸣与强调

多模态协同最直接的方式是强化关系，即不同模态传递相同或相似的信息，彼此印证，产生强烈的共鸣效应[1]。擦鼎受罚段落是这一关系的典型体现。在这一段落中，视觉模态呈现小猪妖趴在铜鼎上用鬃毛摩擦、背部磨破、血渗出的劳累动作；听觉模态提供鬃毛摩擦金属的尖锐声、小猪妖压抑的喘息声、水滴落下的回声；语言模态几乎为零，只有几声压抑的呻吟。三种模态共同指向“底层小妖的艰辛与无奈”这一核心意义，产生强烈的共情效果。这种强化不是一般意义上的情绪渲染，而是对“劳动异化”的视觉化批判——小妖用自己的鬃毛作为工具擦鼎，身体成为劳动的工具，这与当代中国对“体面劳动”的价值追求形成隐性对话。

变装登场段落同样体现了强化关系。小妖们首次穿上唐僧师徒的服装时，视觉模态展示他们的滑稽造型，蛤蟆精的僧衣被撑破，猩猩怪的木棒弯曲；听觉模态配上轻快戏谑的民间小调，以笛子和琵琶为主；语言模态中小猪妖说“咱们现在就是取经团队了”。三者的协同强化了荒诞喜剧的效果，使观众在视觉、听觉和语言三个层面同时感受到幽默。

4.2. 模态互补：延伸与建构

互补关系指不同模态提供不同类型的信息碎片，像拼图一样组合在一起，共同建构出一个更为完整、立体和复杂的意义整体[1]。假扮取经团队段落中，语言提供荒诞设定、视觉呈现滑稽模仿、听觉奠定喜剧基调，三者互补建构了“身份重构”的意义。在这一系列场景中，语言模态提供荒诞的角色分配，蛤蟆精不情愿地说“我演唐僧”，小猪妖说“猩猩壮，你扮孙悟空”，这些台词提供了角色分工的荒诞设定。视觉模态呈现笨拙的模仿动作，猩猩怪拿起木棒笨拙地挥舞，因口吃而说不出台词，蛤蟆精穿上僧衣却因大肚子撑破衣服。这提供了角色模仿失败、反差滑稽的视觉信息。听觉模态的背景音乐是一段轻快、戏谑的民间小调，旋律跳跃而幽默，这提供了喜剧与轻松的情感信息。三种模态分别承担设定、呈现和定调的功能，彼此延伸补充，缺一不可，立体地建构了小妖们在绝境中苦中作乐、逐渐建立团队认同的成长过程。

这一互补关系的深层功能自我肯定与实现。通过扮演英雄，小妖们不是在模仿，而是在“成为”。这与当代青年文化中“角色扮演”作为自我探索和自我确认的方式形成呼应，是对“草根能否逆袭”这一时代命题的乐观回答。

4.3. 模态冲突与主题升华：反衬与反思

最高级的多模态协同有时通过刻意制造模态间的冲突或反衬，创造出深刻的反讽、矛盾或悬念，从而引导观众进行深度反思。影片最高潮的小雷音寺对峙段落堪称模态冲突的典范。

在这一段落中，妖王的语言模态传递诱惑与威胁，他威逼四只小妖假扮四位尊者欺骗唐僧师徒，并以唐僧肉为诱饵。与此同时，妖王身边的妖怪们发出低沉的笑声与附和声。然而，一向沉默寡言、患有口吃的猩猩怪，此时却缓缓站直身体，举起了手中的木棒。视觉模态呈现的是他颤抖却坚定的身躯、直视妖王的目光。听觉模态上，背景音乐从妖王讲话时的低沉压抑逐渐减弱，转为一片短暂的寂静，随后响起一声轻微却无比坚定的木棒敲击地面的声响。此刻，妖王的语言威逼与猩猩怪的视觉和听觉所呈现的无声反抗之间产生了巨大的冲突和矛盾。这种模态间的不和谐让所有观众感到震撼，产生了强烈的戏剧张力。这种冲突并非失误，而是精心设计的叙事策略，它模拟了现实中最卑微的个体在面对绝对强

权时那种沉默却不可摧毁的反抗意志。

最终,猩猩怪用尽全身力气喊出“我是齐天大圣”,这句打破口吃的话语与随后其他小妖纷纷站到猩猩怪身边的视觉动作,以及背景音乐骤然转为宏大激昂的管弦乐与鼓乐,瞬间和解了先前所有的模态冲突。正是通过先制造冲突再完成和解这一系列复杂的多模态操作,影片才如此有力地将主题升华:真正的英雄不在于出身与能力,而在于在绝境中坚持良善、选择反抗的勇气。

猩猩怪用尽全身力气喊出“我是齐天大圣”,这一生理障碍实现的——语言障碍的克服象征着“被压制者重新获得话语权”。这一复杂多模态操作,将影片主题升华为对“每一个平凡人都拥有为自己命名的权利”的肯定,这正是当代中国“奋斗幸福观”的文化表达。

5. 结论

本研究以系统功能语法、视觉语法、听觉语法及张德禄的多模态话语分析综合理论框架为指导,对动画电影《浪浪山小妖怪》进行了深入的个案研究,旨在探讨其多模态话语如何协同构建意义并传播文化内涵。

本研究发现,《浪浪山小妖怪》在多模态设计上具有区别于常规动画的三大独特性。第一,视觉模态构建了“戏仿式西游”符号系统,通过草根化、喜剧化的造型对传统英雄形象进行“降格”处理,解构了英雄主义的不可企及性。第二,听觉模态采用民族乐器与西方管弦乐混搭的策略,在全球化语境中探索文化主体性。第三,视觉美术风格融合水墨写意与造型写实,传承了上海美影厂的民族美学传统。在文化符号学层面,影片通过语言模态的叙事、视觉模态构图、听觉模态的配乐,以及多模态协同的机制,共同讲述了一个具有当代中国价值观的新西游故事——平凡人可以通过坚持良善、团结协作,实现自我价值的确认与成长。

研究最终验证了张德禄的理论框架,模态间存在显著的互补关系与非互补关系,共同服务于整体意义与电影主题的表达。该动画通过语言、视觉、听觉三大模态的协同运作,有效建构了意义并传递了“坚持良善、为无名之辈正名”的深刻主题,为国产动画的多模态叙事提供了有益的参考。

参考文献

- [1] 张德禄. 多模态话语分析综合理论框架探索[J]. 中国外语, 2009, 6(1): 24-30.
- [2] Barthes, R. (1977) *Rhetoric of the Image*. Fontana.
- [3] O'Toole, M. (1994) *The Language of Displayed Art*. Leicester University Press.
- [4] Kress, G. and van Leeuwen, T. (2006) *Reading Images: The Grammar of Visual Design*. 2nd Edition, Routledge. <https://doi.org/10.4324/9780203619728>
- [5] van Leeuwen, T. (1999) *Speech, Music, Sound*. Macmillan. <https://doi.org/10.1007/978-1-349-27700-1>
- [6] Norris, S. (2004) *Analyzing Multimodal Interaction: A Methodological Framework*. Routledge. <https://doi.org/10.4324/9780203379493>
- [7] LeVine, P. and Scollon, R. (2004) *Discourse and Technology: Multimodal Discourse Analysis*. Georgetown University Press.
- [8] Jewitt, C. (2009) *The Routledge Handbook of Multimodal Analysis*. Routledge.
- [9] O'Halloran, K.L. (2004) *Multimodal Discourse Analysis: Systemic Functional Perspectives*. Continuum.
- [10] O'Halloran, K.L. (2005) *Mathematical Discourse: Language, Symbolism and Visual Images*. Continuum.
- [11] Bateman, J.A. and Schmidt, K.H. (2011) *Multimodal Film Analysis: How Films Mean*. Routledge.
- [12] 李战子. 多模式话语的社会符号学分析[J]. 外语研究, 2003(5): 1-8, 80.
- [13] 胡壮麟. 社会符号学研究中的多模态化[J]. 语言教学与研究, 2007(1): 1-10.
- [14] 朱永生. 多模态话语分析的理论基础与研究方法[J]. 外语学刊, 2007(5): 82-86.
- [15] 张德禄. 系统功能理论视阈下的多模态话语分析综合框架[J]. 现代外语, 2018, 41(6): 731-743.

- [16] 韦琴红. 论多模态话语中的模态、媒介与情态[J]. 外语教学, 2009, 30(4): 52-55.
- [17] 冯德正, 张德禄, Kay O'Halloran. 多模态语篇分析的进展与前沿[J]. 当代语言学, 2014, 16(1): 88-99, 126.
- [18] 冯德正. 多模态语篇分析的基本问题探讨[J]. 北京第二外国语学院学报, 2017, 39(3): 1-11, 132.
- [19] 潘艳艳, 李战子. 国内多模态话语分析综论(2003-2017) [J]. 福建师范大学学报(哲学社会科学版), 2017(5): 49-59, 168-169.
- [20] 潘艳艳. 多模态话语分析到多模态认知批评分析的发展综述[J]. 外国语文, 2020, 36(1): 35-42.
- [21] Halliday, M.A.K. (1994) An Introduction to Functional Grammar. 2nd Edition, Arnold.