

基于多模态话语分析视域下的云南文旅宣传片意义构建研究

陈 杰

西南民族大学外国语言文学学院, 四川 成都

收稿日期: 2026年5月14日; 录用日期: 2026年7月13日; 发布日期: 2026年7月23日

摘 要

本研究聚焦于一则文旅宣传片 *Embrace the “Yunnan lifestyle” at a leisurely pace* 中的多模态话语分析, 旨在探讨文本与图像模态如何协同构建意义。研究基于系统功能语法和视觉语法理论框架, 采用定量与定性相结合的方法, 借助 Elan 6.9 软件对宣传片进行转录和标注, 分析其文本模态的概念功能、人际功能和语篇功能, 以及图像模态的再现意义、互动意义和构图意义。研究发现, 文本与图像模态之间存在互补和非互补关系。各模态相互配合, 形成相应的意义建构策略, 以此展现云南的文化特质与生活理念。

关键词

“悠享云南生活”, 多模态话语分析, 文本模态, 图像模态

A Study on the Construction of Meaning in Yunnan's Cultural and Tourism Promotion Videos from the Perspective of Multimodal Discourse Analysis

Jie Chen

School of Foreign Languages and Literatures, Southwest Minzu University, Chengdu Sichuan

Received: May 14, 2026; accepted: July 13, 2026; published: July 23, 2026

Abstract

This study focuses on a cultural and tourism promotional video *Embrace the “Yunnan lifestyle” at a*

leisurely pace and conducts a multimodal discourse analysis to explore how textual and visual modes work together to construct meaning. Based on the theoretical frameworks of Systemic Functional Grammar and Visual Grammar, this research adopts a combined quantitative and qualitative method. With ELAN 6.9, the video is transcribed and annotated. It analyzes the ideational, interpersonal and textual metafunctions of the textual mode, as well as the representational, interactive and compositional meanings of the visual mode. The findings indicate that there exist complementary and non-complementary relationships between the two modes. The coordination of different modes forms specific meaning construction strategies to present the cultural characteristics and life philosophy of Yunnan.

Keywords

Embrace the “Yunnan Lifestyle” at a Leisurely Pace, Multimodal Discourse Analysis, Textual Modality, Image Modality

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

在全球化与数字技术快速发展的背景下，跨文化文旅交流日益频繁，文旅宣传片已成为地域文化对外传播的重要载体。作为典型的多模态文本，文旅宣传片融合文字、图像、音频等多种符号资源，其传播效果依赖于不同模态的协同互动与意义共建，如何优化多模态配置、提升跨文化传播效果，已成为多学科共同关注的议题。

多模态话语分析以系统功能语言学与视觉语法为核心理论，为多媒体文本的意义解读提供了完整框架，当前相关研究成果丰富，但针对英文对外文旅宣传片，结合定量与定性方法系统探究图文模态互动机制的研究仍有待深化。基于此，本文以多模态话语分析为视角，探究文旅宣传片中文本与图像的意义共建逻辑，及其对受众认知的影响。

本研究具备理论价值以及实践价值，放大了多模态话语分析的应用场景范围，验证了相关理论于对外传播文本中的适用情况，丰富多模态互动研究体系的内容范畴；能为文旅宣传片创作起到参考作用，助力地域文化高效开展对外传播；研究融合了多学科视角，为相关领域研究给出了新点子。

本文将3分41秒的云南英文文旅宣传片作为研究对象来开展研究，采用定量跟定性融合的办法，依靠Elan 6.9软件开展符号标注和编码的统计，结合典型片段阐明模态协同的机制，本研究打破以往重视定性、忽视定量的局限，做到理论和实证的融汇，同时立足跨文化传播方面的需求，为文旅对外传播提供能落地的策略举措，既有着理论创新所带来的价值，又具备现实应用所产生的价值。

2. 文献综述

本章将围绕视觉语法以及多模态话语分析展开分析。

2.1. 视觉语法国内外研究现状

国外的视觉语法最先由Kress和Van Leeuwen [1]提出。随后，视觉叙事理论是视觉语法理论的进一步发展，由Painter、Martin和Unsworth [2]等学者提出。该理论框架不仅关注单个图像的意义，还扩展到多个图像组成的复杂视觉叙事。学者冯德正[3]基于语料分析对视觉语法理论进行修正、补充，提出新的

理论框架——视觉叙事。国外文献对于视觉语法的分析应用大多涉及到跨领域,比如将视觉语法同教育学结合起来,学者发现视觉语法有助于学生更好地理解和分析多模态文本,还能提高他们的视觉和多模态识字技能[4]。还有学者将视觉语法应用于英语语法教学方面的指导,鼓励学习者利用视觉图像的不同内涵,更好地理解英语语法[5];在关于儿童眼球运动的实验中,学者运用视觉语法理论对数据进行了分析[6]。

毋庸置疑的是,关于视觉语法的研究是具有意义的。通过在 CNKI 平台上以关键词“视觉语法”进行检索,不难发现,关于视觉语法的期刊以及论文等数量大致呈逐年上升的态势。而国内视觉语法相关的研究也多以 Kress 和 Van Leeuwen [1] 的所提出的概念为理论基础,从再现、互动以及构图意义三个维度考察多模态意义实现和语法构成,并与 Halliday [7] 提出的语言的三种元功能相对应。视觉语法理论为学者分析语篇提供了新的途径,以往被广泛应用于分析影视作品,或是教学话语以及广告[8]等等多种语篇形态的意义构建[9]。但是关于视觉语法理论的研究,较少聚焦于文旅宣传片题材。

2.2. 多模态话语分析国内外研究现状

国外多模态话语分析研究在国际学术界得到了广泛关注和深入发展。随后,胡壮麟[10]进一步解释社会符号学的多模态化和多模态话语分析。国外多模态话语分析关注教育领域的应用,例如 Li, Liu, Xie, Chang 和 Shi [11] 采用多模态话语分析框架来描述衡量混合式同步课堂中人际互动的动态协调。还有学者利用多模态话语分析框架来考察教科书种的文本图像[12]。除此之外,国外多模态话语分析还被广泛应用于分析多模态语篇。

国内多模态话语分析研究近年来发展迅速,主要集中在理论框架的构建和应用领域的拓展上。李战子[13]首次引入了视觉语法的概念,开创了国内多模态话语分析研究的先河。此后,张德禄[14]在吸收海外学者研究成果的基础上,提出了一个全新的多模态话语分析框架,强调不同模态形式特征及其相互关系的重要性。近年来,多模态话语分析的研究主要集中于应用方面。据馨[15]通过视觉语法理论分析广告的图像意义。在儿童文学与视觉传播交叉领域,陈冬纯和陈芝敏[16]从宏观和微观两个视角,分析国内外儿童绘本的多模态叙事结构特征。

相较于现有研究,本文在研究对象与分析视角上形成有效补充。当前国内多模态文旅传播研究多聚焦传统纪录片、旅游海报与短视频,针对外宣文旅宣传片的专项多模态话语分析较少。

3. 理论框架

本研究以系统功能语法、视觉语法、模态间的协同关系为理论基础,系统功能语法通过其三大元功为文本模态分析提供理论框架,视觉语法揭示图像模态的表意功能,模态间协同关系理论探究模态间的关系。

3.1. 系统功能语法

三大元功能是 Halliday[7]系统功能语法的核心概念,分别为概念功能、人际功能以及语篇功能。关于概念功能,语言可以用来表达新信息,传达未知的内容,并且语言作为一种符号编码,是联结人和外部世界或人的内心世界的中介。概念功能的及物性主要包括六种过程:物质过程、心理过程、关系过程、行为过程、言语过程以及存在过程。基于主题契合性的考虑,本文将重点关注心理、言语和关系过程。关于人际功能,旨在用语言构建人与人之间的社会关系以及交流互动,而语篇功能则更加关注语言在语篇中的组织功能,即语言如何将口头的或书面的话语组织成连贯统一的语篇。

3.2. 视觉语法

Kress 和 Van Leeuwen [1] 以 Halliday [7] 的系统功能语法中的三大元功能的相关概念作为理论基础,

向外延伸至视觉领域从而发展成后来的视觉语法理论。在视觉语法框架下，图像意义大概可以分为三部分，即再现意义、互动意义以及构图意义，它们分别对应着系统语法框架下的概念功能、人际功能以及语篇功能。除此之外，参与者分为两类，再现参与者和互动参与者。简单说，再现参与者指的是图像中出现的元素，而互动参与者通常指的是创造或观看图像的人。关于图像三种意义的分类，本文后续已作详细阐述。本研究针对动态文旅宣传片，将单帧图像分析延伸至镜头序列、画面叙事，关注连续画面中互动关系、构图逻辑的动态变化，完善视觉语法在动态多模态文本中的应用。

3.3. 模态间的协同关系

张德禄[14]将模态之间的关系分为两大类：互补关系和非互补关系。互补关系是指一种模态不能充分表达交际者的意义，需要借助另一种模态来强化或协同。互补关系被分为强化关系和非强化关系。强化关系包括突出、主次和扩充三种关系；非强化关系包括协调、联合和交叉。非互补关系是指模态之间不需要依靠对方补足缺失信息，彼此意义不形成互补增益。非互补关系包括交叠、内包和语境交互。交叠包括冗余、抵消和排斥；内包包括整体与部分、抽象与具体；语境交互包括独立和依赖。

4. 多模态文本解构

这一章节将主要围绕文本模态和图像模态进行分析，深入探讨两种模态的符号特征、互动机制及其文化语境下的意义建构。

4.1. 文本模态

文本模态包括自然语言文本、语音识别文本等等。本研究的研究对象选取的是文旅宣传片 *Embrace the “Yunnan lifestyle” at a leisurely pace* 中具有代表性的官方英文字幕，即聚焦于视频中人物对话的原始语料。韩礼德将语言的功能分为三大元功能，即概念功能、人际功能和语篇功能。

4.1.1. 概念功能

概念功能通过六大过程来实现[7]，因其中的关系过程、心理过程和言语过程在视频中的出现频率较高，且表达的意思同此则文旅宣传片的主题较为接近，故文本选取这三大过程中具有代表性的语句各 2 句置于表 1 中，如下：

Table 1. Examples of ideational function
表 1. 概念功能实例

英文字幕	所属过程
1) In today’s increasingly fast-paced life, slowing down is the best way to feel alive.	关系过程
2) Dali is very artistic in its own way.	关系过程
3) I felt like I was, you know, walking around one of those Chinese paintings.	心理过程
4) I gotta say I really love Yunnan.	心理过程
5) I mean it’s really pretty.	言语过程
6) Now follow me and let’s experience life in Yunnan.	言语过程

关系过程在概念功能方面起到描述事物状态及属性的作用，例句 1 和例句 2 都呈现出关系过程，例句 1 说明了“slow down”行为和“feel alive”状态两者之间的关系，这种关系过程不光传达了一种生活哲学，也呈现出说话者对现代生活状态的认识与评判。例句 2 描述了主体“Dali”所具有的“artistic”属

性，这种描述又传达出大理的文化内涵，也体现了说话者对大理独特艺术风格的认知及欣赏，助力我们知晓和审定现实世界里的各类现象与事物。

于概念功能中，心理过程和表达说话者的内心感受及情感有关，表 1 中的例句 3 跟例句 4 都表现出心理过程，例句 3 和 4 都直接抒发了说话者对云南的喜爱之感，此情感的抒发传达出说话者对云南的炽热情感，也说明了说话者对云南文化和生活方式的认可与好评，这些心理过程的表达有利于我们理解、体悟说话者的内心世界，由此达成情感方面的共鸣和交流。

言语过程于概念功能里关联着说话者的观点与言语行为，表 1 里例句 5 及例句 6 都体现了言语过程，例句 5 体现出说话者针对某物的评价，这样的评价不光传递了说话者观点，也体现出说话者对于事物的审美判断，例句 6 是一种邀请加上提议，此言语行为已经把说话者意图予以传达，也体现出说话者希望跟听众分享云南生活的愿景，促成观点和信息的有效交流传播。

4.1.2. 人际功能

人际功能关心语言怎样建立说话者和听话者之间的关系，以及说话者运用语言表达自身态度、意图和情感的做法，语气是达成该功能的主要语法办法之一[7]。语气的一般种类有：祈使、告知以及疑问，本文挑选三则同宣传片主题比较符合的有代表性语句进行剖析。

Table 2. Examples of interpersonal function
表 2. 人际功能实例

英文字幕	语气
1) Now follow me and let's experience life in Yunnan.	祈使
2) But here in Yunnan, slowing down isn't about stopping, it's about savoring.	陈述
3) So, when will you start slowing down?	疑问

表 2 当中的例句 1 是典型祈使句，它直接向听话者发出邀请以及建议，希望听话者跟说话者一道体验云南生活，本句表明了说话者对听话者的尊重和邀请，想要凭借共享体验搭建更紧密的社会关系，例句 2 以陈述语句存在，呈现了有关云南生活的特点，并且呈现了说话者对这种生活方式的领会和评定，以陈述语气为手段，说话者跟听话者可建立起基于信息共享和观点交流的互动关联。例句 3 直接给听话者提出问题，询问听话者何时开启放慢生活的节奏，从表象上看是询问，深层要义是大力邀请听话者到云南体验“慢生活”，疑问语气对说话者和听话者之间互动与对话的建立有促进作用，凭借提出问题引导听话者思考和分享见解。

借分析不同的语气，说明不同语气在人际功能中起到不一样的作用，这些语气不光对说话者和听话者的互动情况有影响，也反映了语言在构建、维护社会关系时所具备的重要作用。

4.1.3. 语篇功能

作为语篇功能中的两个核心概念，是主位和述位，它们对组织信息和构建连贯语篇起到了十分重要的作用，所以本文主要对主位和述位开展分析。

对于主位和述位的分析主要以表 1 中的例句为展开点，首先对于例句 1 来讲，“In today's increasingly fast-paced life”被用作主位，为读者搭建了讨论的背景，而“慢下来是感受生活本真的最好方式”作为述位，体现了明确的观点跟信息，这种结构使读者清楚句子重点是体现快节奏生活当中慢下来的重要价值。表 1 包含的例句 2 里面，以“Dali”作为主位，直接引入了所讨论的主题，而“is very artistic in its own way”被作为述位，刻画了大理别具一格的艺术性，该结构让信息传递直白又一目了然，例句 4 和例句 5

当中的主位分别是“I gotta say”和“I mean”，这些表达方式领进了说话者的个人观点，述位“really love Yunnan”和“it’s really pretty”切实体现了说话者对云南的喜爱及赞美。这种结构不光传达了说话者的情感态度，还以直接表达强化了信息的力度，例句6中所涉及的主位“Now follow me”是祈使句，直接给听者发出邀请，述位“let’s experience life in Yunnan”具体讲清了邀请的相关内容，即一块儿感受云南的生活，把主位和述位做这样的安排，不但传达了说话者的意图，还以祈使语气加强了互动的程度。

4.2. 图像模态

图像模态是通过视觉符号的系统性组织，以直观、感性的方式传递意义，并在多模态文本中承担独立或协同表意功能的符号资源体系。本研究选取云南文旅宣传片 *Embrace the “Yunnan lifestyle” at a leisurely pace* 作为研究对象，分析宣传片图像模态的意义。本节将先从微观视角出发，分析视频的图像意义；其次，将根据 Elan 6.9 标注的具体数据，宏观展现视频中各部分占比。

4.2.1. 再现意义

视觉语法的再现意义由两个部分构成：叙事再现和概念再现，二者的根本区别在于是否存在矢量。关于叙事再现，分为三类情况：行动过程、反应过程以及言语和心理过程。关于概念再现，同样存在三种情况的分类：分类过程、分析过程以及象征过程。一帧图像中可能不止包含一个过程，可根据图像元素的尺寸大小和位置判断主次[1]。

选取代表性的图片分析其意义。首先是行动过程，如图1中，孩子拿着水枪对准大人。此时，拿着水枪的儿童是行动的发出者，矢量是他所表现出来的肢体动作，而目标则是旁边的成年人。在图示所示的过程中，存在行动者、矢量和目标，故该过程属于行动过程。此外，图1中另一个白色衣服儿童虽不构成行动过程，但也展现出一个无形的矢量——即视线。左边白色衣服的儿童是反应者，右边黑色衣服的成年人是现象呢，通过视线连接起来，故构成反应过程。图2中反应的是言语和心理过程，图中记者正在述说着自己的感想感悟，提倡“慢生活”，而这也是为了接下来引入对于云南的介绍与宣传作铺垫。

图3中属于分类过程，分类过程通常关注的是上下级或同级之间的关系。图片中展示的是一些形状各异的耳饰，这些耳饰同属于手工艺品，故此图展现的是分类过程。图4属于分析过程，分析过程涉及两类再现参与者：表示整体的承载者和表示部分的特征，它主要研究整体与部分的关系。图4是关于云南的一张图片，而文字中提及“云南的生活”，这构成整体与部分的关系，图中的景象只是云南的一部分。图5属于象征过程，而象征过程，顾名思义，研究的问题就是“图像参与者是什么或意味着什么”。



Figure 1. Children playing with water guns
图1. 儿童玩水枪



Figure 2. Journalist on site
图 2. 记者进行现场报道



Figure 3. Handicrafts
图 3. 手工艺品



Figure 4. Yunnan life
图 4. 云南的生活

象征过程中通常运用具有象征性的图像元素去表达抽象的概念或情感，图 5 画面中一位女士轻松地坐在草地上面，并抚摸着依偎在她身边的狗狗，而她的另一半拿着手机记录着这一个幸福的画面，并且视频配文“既惬意又多样”。故该图像属于象征过程，图中的画面象征着一种惬意愉快的生活。



Figure 5. Cozy life
图 5. 惬意生活

分析单帧图片的意义是从较为微观的视角出发，下文用工具 Elan 6.9 对整个视频标注后可得出一定的结果，并进行相应的较为宏观视角的分析。

Table 3. Annotation results of narrative meaning and representational meaning
表 3. 叙事意义和再现意义的标注结果

层	标注数量	最小时长	最大时长	平均时长	总标注时长	恢复时间
AcP	7	1.839	10.193	4.314286	30.2	31.476
RP	12	1.226	33.182	11.964833	142.578	0.0
SMP	4	4.51	172.0	48.650002	194.6	0.0
AnP	3	4.4	8.097	5.769	17.307	11.26
CIP	3	3.58	23.238	11.842667	35.528	20.72
SP	4	1.286	4.286	2.75025	11.001	31.571

从表 3 可见能观察到，不同过程标注数量和时长有较为显著的差异，这呈现各类过程在叙述中的占比和所起作用，虽说行动过程所标注的数量只有 7 次，然而其总的时长达到 30.2 秒，表明行动过程不是最为频繁的，但每次出现都比较周全，对推动情节发展起到了比较重要的作用，反应过程的标注数是最多的，总时长达到的时长为 142.578 秒，这体现了于该视频叙述而言，反应过程占据重要地位，借人物的反应呈现情感变化与内心的想法，这正是叙述里不可或缺的一块，言语与心理过程标注的数量只有 4 次，但其加起来的时长有 194.6 秒，这说明言语和心理过程在叙述里出现的频率较低，但每次出现都十分重要。

分析过程的标注有 3 次，合计时长有 17.307 秒，分析过程于叙述过程里出现并不频繁，每次现身都比较精炼，分类过程的标注次数是 3 次，整体时长是 35.528 秒，出现的频率比较低，象征过程有 4 次标注情况，时长整体是 11.001 秒，但有着深刻的象征寓意，视频用象征手法去表达深层次意思。

以对不同过程标注数量和时长的分析为途径，我们可了解各类过程于叙述中的比重和作用，这有利于我们更好地领会视频叙述的结构和要点，于是更高效地开展内容研究，各类过程标注时长反映了不同符号资源对纪录片所产生的影响程度，就例如言语和心理过程标注数量不多，可它平均标注时长是最长，体现此视频专注深度内容的告知，不依赖高频次“碎片化符号”，而是利用长时期的言语互动，整个视频里面，反应过程的占比是最大的，利用展示人物对事件的反应，视频可更高效地与观众建立起情感上的关联，观众可借助人物的反应感受相仿的情感，进而产生共鸣、共情的情况。

4.2.2. 互动意义

互动意义是图像意义的重要组成部分之一，体现了图像和互动参与者之间的关系。下文将通过分析互动意义的接触、态度、社会距离、情态四个方面去分析图像所含的互动意义[1]。

图 6 中的图像参与者直接看向互动参与者，存在明显的目光交流，故该图像中存在索取类接触。除此之外，该图采用正面平视、近景的镜头视角，提供了一种真实和直接的观看体验，使得互动参与者感觉身临其境。与被拍摄对象处于同一水平线上，这有助于增强内容的真实感和沉浸感，并且近景的镜头视角可以营造出一种较为亲密的社会距离，拉近图像参与者同互动参与者之间的关系。图 6 的色彩饱和度看起来适中，颜色鲜明但不过于强烈。色彩区分度良好，人物的服装颜色与背景的绿色植物和天空的蓝色形成明显区分，使得人物与背景之间的界限清晰，增强了视觉的层次感。色彩协调度也相对较高，图像的深度感主要通过背景的模糊处理来实现，这种浅景深的效果使得人物更加突出，同时也增加了图像的立体感和层次感。总体而言，这张图像在色彩、亮度、照明等方面都处理得当，情态值较高，故该图像较为清晰、生动和真实。

利用工具 Elan 6.9 剖析文旅宣传片，数据汇成下表 4:



Figure 6. Interviews with Chinese Americans
图 6. 对美籍华人的采访

Table 4. Annotation results of interactive meaning
表 4. 互动意义的标注结果

层	标注数量	最小时长	最大时长	平均时长	总标注时长	恢复时间
De	13	0.364	33.143	7.691	99.983	0.0
Of	14	1.143	11.454	4.782357	66.953	24.455
ELS	10	4.455	43.455	16.884199	168.842	0.014
LAS	3	2.0	4.819	3.697333	11.092	104.545
HAS	8	2.909	8.728	4.318125	34.545	11.364
CU	4	2.091	4.273	3.25025	113.001	20.636
MCU	8	0.727	33.728	7.05675	56.454	0.909
MS	10	1.181	14.455	6.7363	67.363	1.727
MLS	4	0.272	5.182	2.47725	9.909	24.273
LS	13	1.273	8.637	4.923077	64.0	11.455

表 4 的数据主要来自文旅宣传片中互动意义方面距离、视角以及社会接触的标注数量和标注时长，着眼于标注数量来看，远景以及提供类接触的镜头出现频率比较高，平视镜头最长的最大时长出现，有 43.455 秒，这意味着在展示场景以及动作时，视频更乐于运用平视镜头来给出稳定清晰的视角，且在镜头标注时长里，平视镜头是最长的，时长有 168.842 秒，这说明平视镜头在视频里占比比较大，是视频阐释的主要视角。

平均时长的数据进一步说明出不同镜头类型的使用习惯，平视镜头平均时长属于最长的一类，是 16.884199 秒，这说明于视频之中，平视镜头较多用来详细地展示场景或动作，给观众提供充分信息的时候。

以对不同视角标注数量和时长的分析为途径，提供类接触的镜头占比不及索取类接触，这可使观众更好地理解与共鸣图像参与者的情感情形，凭借观察图像参与者的反应，观众会更有力地感受场景情感深度，视频当中平视、特写镜头的占比比较高，较多平视镜头可以清晰地捕获对话以及动作的细节，有益于信息的无误传递，而较多的特写镜头对理解人物情感状态以及行为意图意义非凡。

4.2.3. 构图意义

构图意义主要包括信息值、显著性以及取景方面的分析[1]。如图 6，置于图像中间位置的人物相对较大，且周围的背景虚化。由此可知，该图像想要突出强调的就是这位被采访的美籍华人，故将其置于较为显著的位置。其次，元素在不同“区域”的位置会赋予它们特定的信息值，这里的“区域”主要是指左边和右边、顶部和底部、中心和边缘。一般来说，居左是已知信息，居右是新信息，图 6 中人物左边有着关于人物的简介，由此可知，该部分是作为已知信息出现的。

取景就是通过线条或框架对图像中的元素进行连接或割断，如图 7，道路在画面中延伸至远方，形成了一个引导线，引导观众的视线向画面深处移动。这种构图技巧不仅增加了画面的深度感，也象征着旅程和探索的过程，与字幕中的“探索的过程”相呼应。



Figure 7. Road views
图 7. 道路景象

5. 模态协同的实践呈现

在张德禄的图文协同关系理论中，语言模态和图像模态之间的两种关系是互补关系和非互补关系[14]。

图 8 中的女士正在讲话，她的手臂呈张开状态，开放的手势往往用来邀请听众参与到对话或活动中来，表现出一种包容和欢迎的态度。她的面部看起来温和而友好，嘴角微微上扬，且目光直视镜头，为索取类接触，她试图同视频观看者建立一种联系。图像中人物的手势以及状态与字幕中的问题“*So, when will you start slowing down?*”相呼应，增强了信息的传达效果。而对于英文字幕的中文翻译更是对视频图

像意义的补充说明，英文字幕中仅提及“slow down”，但是中文翻译中却译为“云南生活”，直接将英文字幕的深层意义翻译出来，表达一种邀请，更贴合了这则文旅宣传片的主题。所以图 8 中的图文模态属于互补关系中的强化关系，文本模态是对图像模态的一种强化，通过扩充手段加以实现。

关于图 9 我们可以看到图像中展示了一系列色彩鲜艳、图案复杂的手工艺品，而字幕则提供了对图像内容的解释和补充。图像中展示的手工艺品具有浓厚的民族特色和文化多样性，这与字幕中提到的“Because it's very diverse somehow.”相呼应，揭示图像和文本在信息上是相同的，一起凸显了文化多样性的主题，这种一致性体现出图文之间的互补跟一致联系，图像以视觉元素把文化的多样性直观呈现，而文本给出针对这一视觉信息的解释和强调，使观众更全面地理解图像所表达的信息，观众不能透彻领会图片所蕴含的含义，图 9 属于互补关系当中的非强化关系，两种模态都不可缺少其一，互为补充。

图 10 体现出一组竹制的乐器，乐器有序地放置在竹席上面，旁边还可见一些陶制的容器，图像里的这些元素和音乐创作、演奏直接有一定联系，这跟字幕当中提到的“I think Dali is a great place for writing m.”这一信息紧密相依。图像依靠呈现具体的音乐器具物品，直接地传达了大理和音乐创作的联系，而文本进一步夯实了这一概念，表明大理是一个可促使音乐创作灵感产生的地方，所以图 10 里的图文模态属于互补关系里的强化关系，文本模态起到强化图像模态的作用。



Figure 8. Journalist's speech
图 8. 记者的讲话

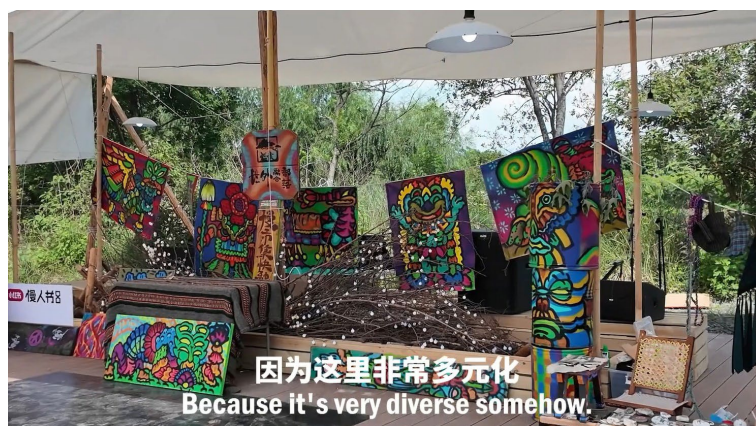


Figure 9. Diverse scenes
图 9. 多样化的景象



Figure 10. Bamboo musical instruments
图 10. 竹制乐器

6. 结论

本文着力就文旅宣传片 *Embrace the “Yunnan lifestyle” at a leisurely pace* 内的多模态文本进行研究，分析了文本与图像模态协同起来构建意义的模式，研究结果显示出文本跟图像之间的非强化及强化关系对于有效传达信息的重要性。

在文本模态分析的操作里，概念、人际、语篇功能以不同的语法结构和语气类型为途径，切实地传达信息并增进了说话者和听话者间的交流，尤其关系过程、心理过程和言语过程频繁出现，跟宣传片的主题紧密关联着，加强了信息传达以及情感共鸣，图像模态分析揭示了依靠视觉符号系统组织去传递意义的办法，探讨再现意义、互动意义和构图意义，揭示了图像于叙述当中的关键作用，利用人物的反应和互动呈现情感变化以及内心活动，进而与观众搭建起情感的联系。

模态协同的实践开展进一步展现了文本模态跟图像模态的互补情形，在强化关系所存在时，文本模态对图像模态起到了补充与强调作用，使信息丰富和明确的程度增加，在非强化关系的情形当中，两种模态互补互需，都必不可少，一起促成了完整的意义。

本研究不光揭示了多模态文本内模态协同的复杂性和多样性，还为理解、评估文旅宣传片的传播效果提供新的视角。图像模态承担情感营造、场景再现以及氛围渲染，文本模态承担概念解读、文化释义以及价值提炼，二者各司其职，缺失任意一者都会造成意义断层，而非简单信息叠加。未来的研究可以进一步探索不同文化背景下多模态文本的模态协同关系，以及这种关系如何影响跨文化传播的效果。通过这种分析，我们可以更好地理解 and 利用多模态文本在文旅宣传中的潜力，以更有效地传达信息和吸引观众。

基金项目

2026 年研究生创新型科研项目(项目编号: YB2026003)。

参考文献

- [1] Kress, G. and van Leeuwen, T. (2006) *Reading Images: The Grammar of Visual Design*. 2nd Edition, Routledge.
- [2] Painter, C., Martin, J.R. and Unsworth, L. (2013) *Reading Visual Narratives: Image Analysis of Children’s Picture Books*. University of Toronto Press.
- [3] 冯德正. 视觉语法的新发展: 基于图画书的视觉叙事分析框架[J]. 外语教学, 2015, 36(3): 23-27.
- [4] Papademetriou, E. and Makri, D. (2015) The Metalanguage of “Visual Design” into the Classroom for the Construction

- of Intermodal Meanings. *International Journal of Education and Literacy Studies*, **3**, 32-43.
- [5] Qadha, A.M. and Al-Wasy, B.Q. (2022) The Impact of Implementing Visual Grammar on Learning Participle Adjectives by EFL Learners. *PSU Research Review*, **8**, 455-466. <https://doi.org/10.1108/prr-02-2022-0017>
 - [6] Li, L., Martin, C., Baogen, L. and Xiaomei, G. (2020) Moving Eyes on Pictures Following Visual Grammar Benefits Meaning Making: Evidence from the Independent Reading of Chinese Preschool Children. *Journal of Chinese Writing Systems*, **4**, 57-70. <https://doi.org/10.1177/2513850219886109>
 - [7] Halliday M.A.K. (1995) *An Introduction to Functional Grammar*. Edward Arnold.
 - [8] 刘丹. 视觉语法视域下竖屏微电影广告的多模态隐喻构建——以华为广告《悟空》为例[J]. 华侨大学学报(哲学社会科学版), 2020(1): 154-160.
 - [9] 潘艳艳, 李战子. 国内多模态话语分析综论(2003-2017)——以 CSSCI 来源期刊发表成果为考察对象[J]. 福建师范大学学报(哲学社会科学版), 2017(5): 49-59, 168-169.
 - [10] 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.
 - [11] Li, X.J., Liu, Q.T., Xie, K., Chang, Y.B., Shi, Y.F. and Ma, J. (2023) Understanding Interpersonal Interaction Characteristics in a Blended Synchronous Classroom: A Multimodal Discourse Analytic Perspective. *Asia Pacific Journal of Education*, **45**, 725-748. <https://doi.org/10.1080/02188791.2023.2206550>
 - [12] Alfred, B. and Olayinka, I. (2025) A Multimodal Discourse Analysis of In-Text Images of Nigerian Literature Texts for Primary Schools. *Language and Semiotic Studies*, **11**, 233-268. <https://doi.org/10.1515/lass-2024-0040>
 - [13] 李战子. 多模态话语的社会符号学分析[J]. 外语研究, 2003(5): 1-8.
 - [14] 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.
 - [15] 琚馨. 视觉语法视角下公益广告的多模态话语分析——以《没有医院就没有希望》公益广告为例[J]. 今传媒, 2020, 28(2): 88-90.
 - [16] 陈冬纯, 陈芝敏. 基于新视觉语法的英汉儿童绘本叙事结构比较研究[J]. 西安外国语大学学报, 2019, 27(4): 36-41.