

基于Elan的宣传视频多模态话语分析

——以第22届上海国际艺术节先导片为例

李智慧

上海大学外国语学院, 上海

收稿日期: 2025年6月25日; 录用日期: 2025年8月18日; 发布日期: 2025年8月29日

摘要

近年来, 多模态话语分析一直是语言学研究的热点, 是实现话语意图的有效工具, 也是检验意义传递效果的客观证据。本文选取第二十二届上海国际艺术节先导片作为研究语料, 以多模态话语分析理论为基础, 对视频构建的语境进行分析, 同时借助多模态分析软件Elan, 对各模态进行分层标注和统计, 阐释不同模态之间的相互关联和意义建构, 旨在呈现上海艺术文化的全新内涵, 提升宣传片的制作质量, 帮助更好促进中国艺术形象的对外传播。

关键词

多模态话语理论, Elan, 宣传视频

Multimodal Discourse Analysis of Promotional Videos Based on Elan

—A Case Study of the 22nd Shanghai International Arts Festival Trailer

Zhihui Li

School of Foreign Languages, Shanghai University, Shanghai

Received: Jun. 25th, 2025; accepted: Aug. 18th, 2025; published: Aug. 29th, 2025

Abstract

Multimodal discourse analysis, a focus in linguistic research in recent years, functions as an effective tool for realizing discourse intentions and objective evidence for examining meaning transmission effects. This study takes the promotional trailer of the 22nd Shanghai International Arts Festival as corpus. Guided by multimodal discourse analysis theory, it analyzes the video-constructed context,

and via the multimodal analysis software Elan, conducts hierarchical annotation and statistics of various modes to illustrate their interconnections and meaning construction. It aims to reveal the new connotations of Shanghai's art and culture, enhance promotional video production quality, and facilitate the international communication of China's artistic image.

Keywords

Multimodal Discourse Theory, Elan, Promotional Videos

Copyright © 2025 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

为提高社会凝聚力与国家文化软实力，强化我国在全球性信息传播中的竞争力与影响力，政府部门与传媒行业密切合作，旨在共同制定依托新媒体塑造并维护良好国家形象的发展策略，扩展国家形象与民族精神的跨文化传播范围与途径。上海国际艺术节，作为文化部主办，上海市承办的国家级国际艺术节，是中国一项重要的对外交流活动。它以吸收世界优秀文化，推动中外文化交流为宗旨，历经 22 届，现已成长为对外文化交流的重要平台。而先导片，本质上作为活动预热性质的“预告片”，通常时长较短，通过紧凑的节奏，精彩的画面，在短时间内吸引观众，让观众产生共鸣。优质的先导片，能够通过用户的自发传播扩大活动的辐射范围，从而吸引潜在受众。作为艺术节重要的宣传物料，先导片结合了图像，颜色，声音，动作，文本等多种符号资源，是一种典型的多模态动态语篇，通过分析各个符号资源之间的相互关联，能帮助提高视频的制作质量，更好的达到宣传目的。

上海国际艺术节先导片作为文化传播载体，符号资源丰富，其宣传力度很大程度依赖于多模态话语的建构，故本研究选取艺术节宣传片为语料，借助多模态分析软件 Elan，基于多模态话语分析理论和视觉语法对该动态语篇进行分析，以期能够为中国对外文化形象的塑造提供借鉴，助力中国艺术文化在对外交流中的输出。

2. 多模态话语分析

近年来，越来越多的多模态研究者在探索社会符号的同时，开始注重不同学科视角的融合，使得多模态逐渐成为一个跨学科的概念。Harris 认为，话语分析，作为研究语言的一种方法，主要聚焦于人们日常生活中所使用的语言，通过对语言特征的观察和描述，进而探究其在实际场景中所承载的功能及具体用法。多模态话语分析里，语言连同绘画、雕塑、音乐、舞蹈等非语言符号，同属于符号资源范畴，都具有意义的表征与传播作用。多模态话语分析领域的代表学者 Kress 与 Leeuwen 认为，“模态”是人间沟通的符号，作为承载意图与交流意义的社会文化资源，其范畴涵盖各类工具及物质形式[1]。国内学者顾曰国曾在《多模态感官系统与语言研究》一书中提到，“模态”就是人类依靠视觉、听觉、触觉、嗅觉等多种感官系统，与外界事物及环境开展交流互动的一种方式手段。他认为，模态指的是人类借助感官系统与外部环境进行交际互动的一种方式[2]。张德禄提出多模态话语分析的理论，他指出多模态是人类交际的普遍现象，多模态话语分析是指人借助感官系统通过语言、图像、声音、动作等多种手段和符号资源跟外部事物或者环境交流互动的途径。多模态话语分析通过整合语言及其他符号资源展开研究，使得符号在社会交际中具有更强的应用价值。

在新闻传播领域，多模态话语分析的对象多为图片或时长较短的广告，对于时长稍长的宣传片研究较少。经分析发现，宣传片因其宣传性质，符号资源会更丰富，符号意义会更显著。在多模态分析领域，语言和非语言符号资源(如图像、音乐、空间安排等)共同构成符号资源，达到表达传递意义的目的。宣传片通过图像、字幕、背景音、独白和旁白等多种符号资源协同表达传递信息，是一种典型的多模态动态语篇[3]。当前对动态语篇分析模式主要有两种：一是基于系统功能语言学的转录模式，用多模态软件标注语篇并转化为图文，结合视觉语法、语域、衔接等理论分析多模态符号的意义建构；二是从文化层、语境层、内容层、表达层等视角分析动态语篇如何构建，强调表达层与内容层均以语境层为基础，并且都受到特定语境的严格制约[4]。

本文主要基于第一种分析模式，从多模态话语分析的角度出发，借助动态语篇分析软件 Elan6.7，对上海国际艺术节先导片中的符号资源进行统计分析，同时依据视觉语法理论来解读图像中的意义，探究如何在多个模态维度建构国家文化对外形象宣传，同时也可作为宣传片制作方面提供更多的思路和理论依据。

3. 先导片视频构建的语境分析

英国人类学家马林诺夫斯基最早将“语境”运用到语言文化研究中。语境，也称为语言环境，是理解话语的重要方式，分为文化语境和情景语境，语言所处的环境能够反映语言的背景和深层含义，通过分析先导片视频的语境，能够帮助了解先导片所要表达的具体思想、情感和內容。

3.1. 文化语境分析

文化语境是指说话人所属的话语群体所涵盖的历史文化、习俗、思维模式、伦理道德等文化语境，并且该话语群体都能理解其中的内在含义[5]。Baldry 和 Thibault 指出可以从语篇的内容层面或表达层面的变化找出多模态语篇不同阶段之间的过渡[6]。因此，从视频上来看，可以将这个宣传片的视觉主题划分为 5 个阶段，如表 1 所示：

Table 1. Main plots of the discourse

表 1. 语篇主要情节阶段

语篇阶段	主要内容
第一阶段(开头)	由“what is art”引入主题
第二阶段(上海)	上海这座城市的艺术氛围
第三阶段(艺术节)	上海国际艺术节
第四阶段(重要性)	艺术节在对外文化交流中的重要性
第五阶段(结尾)	众人特写镜头结束

本文所选取的宣传片短片汇集了多位海内外知名艺术家以及演绎机构的诚挚祝福，也展现了各界对艺术节的期待以及艺术与城市关系的深刻理解。该宣传片以“what is art”引入主题，分为三个不同的板块介绍上海国际艺术节。随后分阶段介绍上海、艺术节以及艺术节在对外文化交流中的重要作用。同时视频中夹杂着大量的人物采访，凸显大众对于艺术的理解，增强观众对于视频观点的认同，最后在结尾处，画面采用分屏、快速切换等剪辑手法，配合同步声音，视觉模态和语言模态相协同，以众人齐说“盛典回归，艺享全球”结束，将个体话语升华为群体共识，增强信息的说服力，促进观众从“理解”到“认同”，进而引发情感共鸣。

3.2. 情景语境分析

情景语境是文化语境的具体呈现,包含话语范围、话语基调和话语方式三个维度。其中,话语范围(语场)指在交际过程中,说话者所涉及的主题以及相关情境因素;话语基调(语旨)则强调交际参与者的社会角色以及互动意图。说话方式(语式)是语言活动所采用的语言形式或者媒介,比如口头语和书面语就是两种不同的语式。根据情景语境的定义划分,情景语境在该视频中主要表现为:① 话语范围。该先导片中涉及到大量的人物采访,他们分别表达了自己对于艺术的想法以及上海艺术节举办的重要性。② 话语基调。由于受到疫情影响,历经三年上海国际艺术节再次举办,关注度很高。所以视频的创造者根据艺术文化的内涵以及受众的审美水平、接受能力选择多模态符号并合理搭配,从而提高视频的受众欢迎度也在宣传上海国际艺术节上起到至关重要的作用。③ 话语方式是语言指所采用的不同形式或媒介。先导片大多是人物面对面采访同时伴有艺术节的部分节目。后期剪辑中将人物的采访对话与艺术节的主题相联系来宣传国际艺术节。

4. 先导片多模态语篇设计分析

4.1. Elan 分析赋码方案

Elan 是由英国学者卡普兰和特纳共同开发的可用于多模态话语分析的标注软件,能够对语言、图像、文字、声音等多种模态进行多层次标注,多角度分析,能够帮助研究者更好理解视频中各种模态之间的关系,相比其他软件,ELAN 可以循环播放,精确定位到 0.01 秒,用户可以自定义标注编码表,标注与文本、声音、图像可以同步进行[7]。本文选取的视频内容发布于 2023 年 7 月 10 号,在艺术节官网,视频号,“Shanghai let’s meet”海外英文账号进行扩散,全网点赞转发量超 5000 余次。先导片的发布正值艺术节国际化程度创历史新高之际(境外项目占比 60%,首演占比显著提升),也间接强化了海外对于艺术节的关注。例如,美国芭蕾舞剧院时隔 23 年再次重返上海,在上海大剧院首次展演的《尼伯龙根的指环》加座售罄,这些亮点均通过《纽约时报》《卫报》等国际媒体传播,而先导片作为其预热内容,也为后续报道奠定了基础。故本研究借助 Elan6.7 软件从影像、图像和表达方式 3 方面对语料进行标注分析,见表 2;根据表 2 赋码方案对语篇进行分层标注,详见图 1。

Table 2. Coding scheme

表 2. 赋码方案

类型	赋码	含义
表达方式	EN	Expression-Narration 旁白
	EM	Expression-Monologue 独白
	AP	Animation Person (影像人)
影像	AT	Animation Thing (影像物体)
	ATP	Animation Thing & Person (影像物体\人)
	IT	Image Text (文字)
图像	IP	Image Picture (图片)
	ITP	Image Text Picture (文字\图片)

4.2. 先导片模态统计分析

经多模态分析软件 Elan 分析后,各层级的统计数据见图 2 和图 3。该视频作为上海国际艺术节的宣

传物料，时长为 5 分 22 秒，是一种结合了图像、字幕、音乐、旁白等多种符号资源的动态语篇，通过 Elan 对各层进行标注后，见图 2，Expression 的模态标记数量为 43 个，占比最大，在表达层面上，EM 独自总标注时长为 147.83，EN 旁白总标注时长为 45.17，而 AP 总时长仅为 7.1。由数据得知，该语篇表达方式主要是以采访人独白为主，不定时穿插旁白。独白能直接传递受访者的主观体验，使得叙事更加真实，让观点更贴近大众，而旁白的穿插，则承担“公共视角”的引导功能，例如，视频以“**What is art**”引入，插入不同视角对于艺术看法的旁白，从而引导观众关注艺术话题。Animation 影像也是媒体进行宣传的重要手段，标记数量为 17 个。图 3，在影像中，ATP 居于首位，总标注时长为 66.25，说明视频中影像、物体和人共同出现在画面中的占比较大。而图像模态也是该宣传片的主要叙事模态，视频中有大量的镜头移动和切换来表现时间变化和空间转移，能够直接触发观众的情感反应，协同重构语言的意义，使语言更加具象化。在宣传类视频中，音乐也是必不可少的一部分，经分析发现，该先导片的背景音乐全部为主观音乐，且音乐的时长和视频总时长一样，所以并不具备参考意义。从图表分析可以看出，各种模态的使用不仅会增强动态语篇的可观赏性，同时也会增加受众的获取信息量。

4.3. 模态具象化分析

基于 Halliday 的系统功能学理论，Kress 和 Leeuwen 将语言元功能理论扩展到视觉图像领域，认为

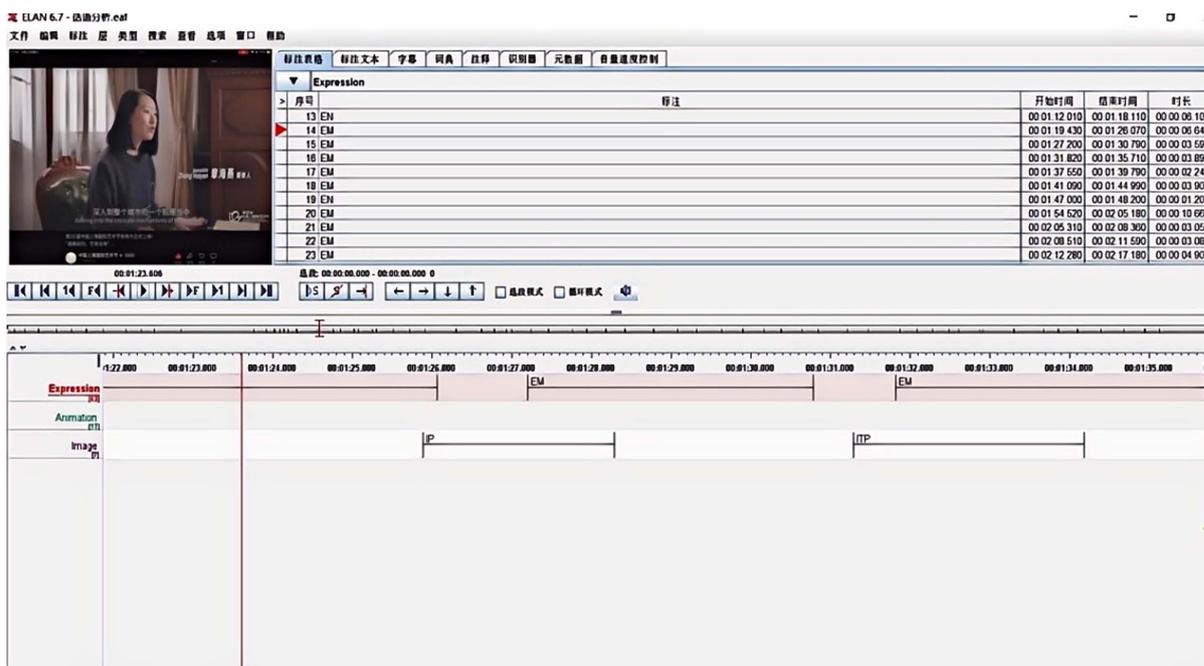


Figure 1. Elan annotation information

图 1. Elan 标注信息

统计参数									
层	标注数量	最小时长	最大时长	平均时长	中间时长	总标注时长	标注时长...	恢复时间	
Expression	43	0.96	13.17	4.935814	3.7	212.24	65.184	1.51	
Animation	17	1.51	8.39	4.526471	3.5	76.95	23.633	64.51	
Image	7	0.76	7.93	2.698571	2.15	18.89	5.802	70.84	

Figure 2. Statistics on layered information of the trailer

图 2. 先导片分层信息统计

选择层:	Expression	<input checked="" type="checkbox"/>	只显示根层						
统计参数									
标注	命中	最小时长	最大时长	平均时长	中间时长	总标注时长	标注时长...	恢复时间	
	2	1.72	10.42	6.07	6.07	12.14	3.729	238.0	
AP	2	3.4	3.7	3.55	3.55	7.1	2.181	3.05	
EM	26	2.12	13.17	5.686	4.6	147.83	45.402	67.83	
EN	13	0.96	10.69	3.475	2.92	45.17	13.873	1.51	

选择层:	Animation	<input checked="" type="checkbox"/>	只显示根层						
统计参数									
标注	命中	最小时长	最大时长	平均时长	中间时长	总标注时长	标注时长...	恢复时间	
AT	2	3.93	6.77	5.35	5.35	10.7	3.286	273.27	
ATP	15	1.51	8.39	4.417	3.33	66.25	20.347	64.51	

选择层:	Image	<input checked="" type="checkbox"/>	只显示根层						
统计参数									
标注	命中	最小时长	最大时长	平均时长	中间时长	总标注时长	标注时长...	恢复时间	
IP	2	0.76	2.4	1.58	1.58	3.16	0.971	85.89	
IT	2	2.15	7.93	5.04	5.04	10.08	3.096	294.38	
ITP	3	0.92	2.9	1.883	1.83	5.65	1.735	70.84	

Figure 3. Summary of layered annotation information by Elan

图 3. Elan 分层标注信息汇总

图像等视觉模态同样具有语言的元功能，都以其独特的形式独立发挥其功能。他们以系统功能学理论为基础，将语言学中的元功能，即概念功能、人际功能和语篇功能拓展至视觉符号层面，提出视觉语法理论，依次对应为再现意义、互动意义和构图意义。宣传片是集字幕、图像、独白、音乐等多种模态符号于一体，符号资源丰富，具有重要宣传意义，因此视觉语法理论为多模态话语分析提供了一个很好的研究视角。

(1) 再现意义。在多模态语篇中，再现意义相当于系统功能语法的概念意义，细分为叙事再现和概念再现，叙事再现是对时间序列的语言建构，侧重于动态过程，概念再现是对抽象观念的符号表征，侧重静态语义，该宣传片属于动态语篇，叙事再现占比较大，主要是通过图像反映客观世界，通过不同场景的转换构建语篇的主题意义。该宣传片中的场景构成主要是：上海这座城市(高楼大厦、公园、酒吧、商圈等)，第二个主要是剧场(演员、道具和舞台)。既通过高楼大厦，商圈来强化上海的“文化名片”，当观众将“上海”与“多元艺术”“活力生活”深度绑定，这座城市的文化软实力会被具象化，进而吸引更多国内外艺术资源；据悉，在艺术节举办期间，来自 54 个国家和地区的上万名艺术家会聚上海，举办了近千场高品质的演出和展览活动，吸引观众超 600 万人次，向世界构建一个真实立体的中国文化形象。图 4 涉及到剧场相关元素，聚焦艺术本体，将剧场的片段真实展现给观众，通过演员状态与舞台细节彰显艺术节的专业性与多元形态。

(2) 互动意义。根据视觉图像分析法，接触、社会距离、态度和情态这四个方面共同构建视觉图像中的互动意义。宣传类视频非常注重互动意义，互动意义的实现依赖于图像参与者的主观能动性，这种主观能动性的发挥，既体现在受众对于图像符号的个性化解读，同时也影响互动关系的构建与递进。接下来本文将围绕这四个维度，详细分析该动态语篇如何通过多模态符号的构建，来实现图像与受众之间的互动。

接触。接触指的是图像参与者和图像互动者间建立的一种想象性接触互动关系，在接触这一环节，图像互动者对于图像互动意义的解读依赖于其主观能动性[8]。接触细分为：索取和提供。二者的差异主要体现在图像参与者是否与受众有眼神交流。索取类图像中，参与者直接与观众进行眼神交流，产生一种理想的互动关系，而在提供类图像中，参与者与观众没有眼神交流，主要提供信息。该先导片既有索取类图像也有提供类图像，但是以提供类图像居多，视频中多为人物采访，并未与观众直接进行延伸接触，无法构成理想的互动关系。图5是宣传片2分33秒的截图，一位国外歌唱家向观众介绍展出内容，



Figure 4. Video screenshot (representational significance in Image Provider Class)

图4. 视频截图(提供类图像再现意义)



Figure 5. Video screenshot (contact in Image Requester Class)

图5. 视频截图(索取类图像中的接触)

视频采用平视的角度，歌唱家目光指向观众，营造出平等的视觉关系，缩短两者的社交距离，仿佛在与观众对话。图6在视频的1分18秒，参与者并未与观众进行眼神交流，整个身子是侧着的，这张图就属于提供类信息，观众需要注意参与者所说的话。通常宣传类语篇提供类图像占比较大，但是该动态语篇提供类图像和需求类图像紧密结合，占比相似，避免视频主体单向输出，强化共同参与感，让受众觉得艺术不是属于少数人的，而是每个人可参与的“共同事业”，拉近与大众的心理距离；其次也能提升传播实效，既以艺术节的实力赢得认可，又能激发大众参与的动力。

社会距离指的是参与者和观众之间的距离，它通过图像的取景来影响社会距离的亲疏远近，分为近景、中景和远景。一般近景能够拉近与观众的距离。该语篇是远景与近景相结合，这样能让观众有身临其境之感。如图7，远景能让观众快速感知场景的规模，了解艺术节的宏大氛围，给观众带来视觉冲击；



Figure 6. Video screenshot (contact in Image Provider Class)

图6. 视频截图(提供类图像中的接触)

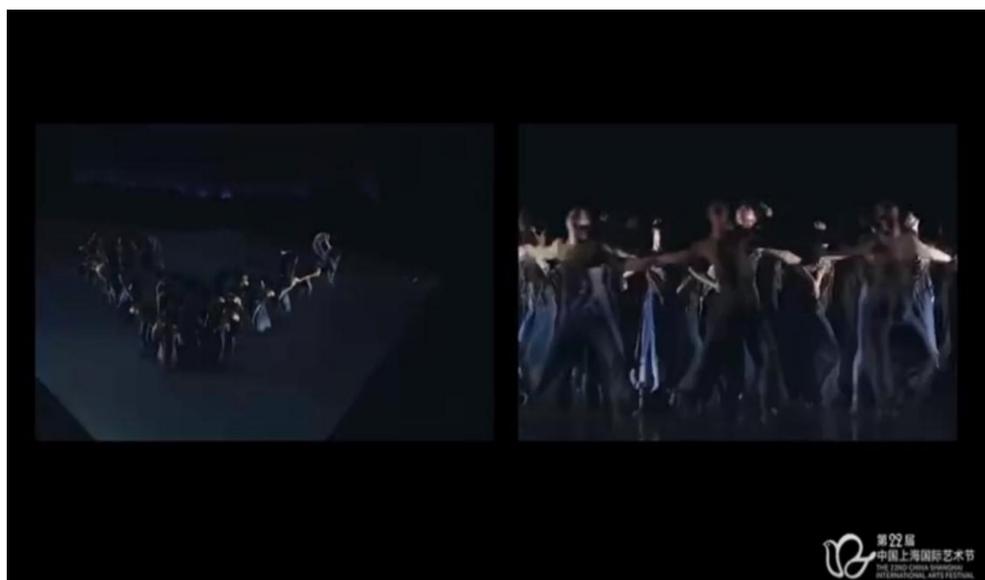


Figure 7. Video screenshot (close and long shots)

图7. 视频截图(近景和远景)

而近景会拉近观众与表演的距离，让他们从旁观者变成仿佛在场的参与者，进而更愿意参与其中。该语篇通过视觉模态，让观众感受到艺术节的演出既高水平(远景展现其规模与专业度)，又具有大众化(近景的细腻表达)，达到扩大文化传播的目的。

态度。态度互动，主要是由图像拍摄视角呈现，视角可以清楚表达参与者态度。拍摄视角主要分为两种：水平和垂直视角。正面水平视角更容易拉近与观众的距离，而倾斜的视角会让观众有一种疏离陌生感。图 8 可以看出图片采用正面的拍摄角度，镜头以与观众平视的角度捕捉场景，能够让观众清晰的捕捉到参与者的表情、眼神以及手势动作，避免因侧面拍摄而带来的疏离感，但是参与者身体和视角是倾斜的，这主要是因为参与者是面对采访者，参与者的表达(如对艺术节的感受、对城市的理解)更易被接受为“真实的个体经验”。正面拍摄与参与者倾斜姿态相结合，观众既不是被动的接受镜头信息，也不是完全游离在外的旁观者，而是通过镜头获得“平等观察”的位置——这种视角暗示“每个人都可以像这样参与对话”，间接强化了艺术节是开放的，大众可参与的。



Figure 8. Video screenshot (attitudinal interaction)

图 8. 视频截图(态度互动)

情态。情态值有高、中、低情态，强调图像的真实度和可信度。多模态语篇多通过元素的清晰性和饱和度等传递情态强度，低饱和度的图像情态较低，反之，色彩鲜艳高饱和度的图像情态值较高。情态程度的高低影响图像所传达信息的真实度和可信度。如图 9，构图整体颜色饱和度较低，高饱和度的颜色通常具有较强的视觉刺激，容易与大众产生距离感。通过弱化视觉冲击，选择低饱和度的颜色构图，能够向大众传递“艺术不是商品，而是生活体验”。柔和的色调会弱化观众对于艺术节高雅、难懂的刻板印象，潜移默化的告诉观众，参与艺术等同于享受日常，是人类精神世界的升华，从而吸引观众主动走进艺术节，真正地实现艺术走向大众。

构图意义。构图意义主要体现在构图元素之间的空间顺序上，主要由空间位置构建，指的是受众对于图像的整体印象，可以从信息值、取景和显著性三方面分析动态语篇的构图意义。图 10 中出现的东方明珠，作为上海最有名的商业地标之一，传递着城市品牌文化的独特性，也是外界感知理解城市的重要窗口。这部先导片借助东方明珠这一具象化符号，意在彰显上海具备承办国际艺术节的实力。它不仅让观众能清晰辨识城市文化所承载的形象特质，也进一步助力城市品牌文化的有效传播，使观众能在潜移默化中接受并认同这一文化符号所传递的文化价值。图 11 中的上海大剧院作为图中的主体部分，其建筑

本身就具有标志性，坐落于城市核心区域，古典与现代融合的外观，很多艺术节的剧目都在该剧院展演，使其天然成为一张靓丽的文化名片，都在明确向观众提供大量信息，向观众展示这座城市的非凡魅力，加深观众对于上海的印象；显著性则指构图在不同程度上吸引观众，比如图 10 中的东方明珠是上海的标志性建筑，为观众所熟知，而图 11 中的上海大剧院五个字也非常显眼，适配传播场景，提高宣传的力度，同时在一定程度上也能够让外国友人快速建立城市认知，这就是上海，展示城市的时代特征和发展实力。从构图框架来看，图中的主要信息都占据了整个图像的一大部分，突出了主要信息，能够让受众跳出复杂的背景介绍，直接建立与城市的连接，加深对于这座城市的印象。



Figure 9. Video screenshot (low-saturation color composition)

图 9. 视频截图(低饱和度颜色构图)



Figure 10. Video screenshot (Framing in Compositional Meaning)

图 10. 视频截图(构图意义中的取景)



Figure 11. Video screenshot (Salience in Compositional Meaning)

图 11. 视频截图(构图意义中的显著性)

5. 结论

本文借助多模态语篇分析软件 Elan 上海国际艺术节先导片中多模态符号资源进行分层标注和统计,同时从语境层面分析了文化语境和情景语境;并且参照视觉图像分析法,分析了该先导片的再现意义、互动意义和构图意义,论述了不同模态之间如何相互协作,共同构建语篇的整体意义。研究发现,该先导片通过旁白、独白、图像、主观音乐等多种符号资源的整合,塑造了一个充满艺术活力的城市,宣传了上海国际艺术节;视觉与听觉模态的相互协作,使得该动态语篇的再现、互动以及构图意义得到更好地呈现,增强受众对于城市以及其文化内涵的认可。本文借助软件分析解读了宣传片多模态语篇,了解多种模态协同作用,以期能够对于此类宣传视频传播中国文化,促进对外文化交流提供借鉴。

参考文献

- [1] Kress, G. and van Leeuwen, T. (2001) *Multimodal Discourse: The Modes and Media of Contemporary Communication*. Arnold.
- [2] 顾曰国. 多模态感官系统与语言研究[J]. 当代语言学, 2015, 17(4): 448-469.
- [3] 翟海霞. 安徽省旅游形象广告宣传片的多元模态话语分析[J]. 海外英语, 2020(7): 238-239, 241.
- [4] 耿敬北, 陈子娟. 文献记录片的多模态动态语篇分析——以《指点江山: 毛泽东诗词故事》之《沁园春·长沙》为例[J]. 西安外国语大学学报, 2014, 22(4): 24-28.
- [5] 黄国文. 功能语篇分析纵横谈[J]. 外语与外语教学, 2001(12): 1-4, 19.
- [6] Baldry, A. and Thibault, P.J. (2006) *Multimodal Transcription and Text Analysis*. Equinox.
- [7] 王立非, 文艳. 应用语言学研究的多元模态分析方法[J]. 外语电化教学, 2008(3): 8-12.
- [8] 陈俊. 城市旅游宣传片的多元模态语篇互动意义解读——以淮安旅游宣传片为例[J]. 现代交际, 2015(9): 51-53.