

动态多模态话语的模态协同研究

——以城市宣传片为例

何广惠, 李 成

武汉工程大学, 外语学院, 湖北 武汉

收稿日期: 2022年2月26日; 录用日期: 2022年4月8日; 发布日期: 2022年4月15日

摘 要

以武汉2020城市宣传片为语料, 结合ELAN软件的分割、分层、标注功能, 探讨多模态话语中不同模态的特点, 展现视觉模态和听觉模态的相互配合及其在镜头速率与运镜方式层面上的体现, 同时分析了模态之间的协同以及多模态话语的意义建构。研究发现模态之间的关系随时间呈动态变化, 其标注分布规律图揭示了模态的选择取决于最佳搭配, 并非完全自由随意, 模态协同可通过相互协调、相互强化以实现意义构建。

关键词

动态多模态, 话语分析, 模态协同, 城市宣传片

On the Modal Synergy of Dynamic Multi-Modal Discourse

—A Case Study of City Promotion Video

Guanghui He, Cheng Li

School of Foreign Languages, Wuhan Institute of Technology, Wuhan Hubei

Received: Feb. 26th, 2022; accepted: Apr. 8th, 2022; published: Apr. 15th, 2022

Abstract

Wuhan 2020 City Promotion Video has been analyzed with the help of ELAN software. According to the function of segmentation, tier and annotation, the characteristics of different modalities in multi-modal discourse are presented, revealing the mutual cooperation of visual and auditory

modalities and the representation on lens speed and camera moves. The production of the video embodies the synergy between modalities and the meaning construction of multi-modal discourse. The study indicates that the relationship between the modalities changes dynamically over time. The annotation distribution map reveals that the choice of modal depends on the best collocation and is not completely free and arbitrary. Modal coordination can achieve meaning construction through mutual coordination and mutual reinforcement.

Keywords

Dynamic Multi-Modal, Discourse Analysis, Modal Synergy, City Promotion Video

Copyright © 2022 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

随着科技高速发展, 人类的交际活动呈现智能化、可视化、多模态化特征。语言外因素包括空间、手势、凝视、身势、移动、音乐、三维事物、图像、图表等成为交际媒介, 实现表情达意功能[1]。这种运用听觉、视觉、触觉等多种感觉, 运用语言、图像、声音、动作等多种符号资源进行交际的现象就是多模态话语[2]。Van Leeuwen 认为多模态是多种符号的组合[3]。Baldry 和 Thibault 认为多模态强调表达语篇意义时多种模态的相互配合和运用[4]。总之, 多模态指口头或书面交际中语言、图像、颜色, 音乐和其它符号系统在内的多元化的方式或交流频道[5]。

多模态话语的研究对象从静态(漫画、海报等)拓展到动态(广告、电视访谈、宣传片等) [6]。城市宣传片作为一种典型的动态多模态话语, 由多种符号资源如图像、语言、音乐、声响等互动共同产生意义, 其动态性在于时空的变化, 即在一定的时间和空间内持续呈现不同的画面, 保证画面的完整性和统一性 [7]。城市宣传片的研究视角主要涉及传播学、营销学、符号学和叙事学, 多模态视角下的城市宣传片研究近年来逐渐兴起。但是, 多模态话语研究仍以质性研究为主, 着重描述和总结, 缺乏数据的支持。近年来, 国内外学者设计出多模态话语分析软件(Praat, ELAN, Anvil, Transcriber AG), 提供大数据的支持。不同研究者也利用这些软件进行实证研究, 但成果较少。而且相对于其它软件, ELAN 的友好性较高, 实用性强[8]。

武汉 2020 城市宣传片“相见在武汉”由武汉市文化和旅游局发布, 历时 3 分 40 秒, 包含 192 个画面, 主要特点是没有旁白, 没有字幕, 没有景点介绍, 以轻音乐结尾。这是与以往城市宣传片的不同之处。本文借助多模态话语分析软件揭示视频中每个情态的特征, 研究不同情态的意义建构, 探讨了宣传视频主题的呈现方式。

经过近 30 年的发展, 动态话语的研究多聚焦于语言和图片, 且偏向于定性分析。本文采用定量统计分析方法, 研究武汉城市宣传视频中不同模态之间的协同作用, 展示视觉模态和听觉模态的协同工作模式, 准确展现意义的建构过程。观看者能更深刻、更准确地掌握动态话语中的意义, 从而提高对城市宣传视频的欣赏能力。

2. 理论背景

(一) 视觉语法理论概述

Halliday 开创的系统功能语言学认为, 语言具有社会符号性, 语言性质决定语言功能, 可以将语言功能抽象概括为“元功能”。该理论整合了语言、语法、语境等内容, 构建了一个多视角、多层面的话语分析模型, 实现了语言学和社会学的有机结合, 为多模态话语分析奠定了理论基础。视觉语法理论是 Kress & Van Leeuwen 以韩礼德的系统功能理论为基础提出的。该理论认为图像亦具有社会符号性, 主张用系统功能语言学分析图像。“再现意义”“互动意义”和“构图意义”是视觉语法分析的核心。再现意义指图像模态可再现客观事物及其与外部世界的关联, 并基于图像的特征细分为叙事再现和概念再现; 互动意义是指图像通过接触、社会距离和态度等要素, 实现客观事物与外部环境的互动; 构图意义指图像的再现意义与互动意义有机结合, 通过信息值、显著性和框架实现意义构建[9]。

(二) 多模态话语的相关研究

自李战子 2003 发表“多模式话语的社会符号学分析”以来, 众多学者对多模态话语的研究纷至沓来。研究类型主要集中在: 介绍性的理论引入与阐述、多模态话语与意义构建(静态和动态研究对象)、理论框架及研究路径探析三大类。首先在理论引入方面, 胡壮麟探讨了多模态符号学和多媒体符号学的区别, 指出人类进入社会符号学多模态化的新世纪[10]; 朱永生论述了多模态话语的产生和定义、多模态话语分析的理论基础和研究方法[11]; 李战子、陆丹云讨论了多模态话语分析的理论基础、研究途径和发展前景[12]; 此外冯德正就多模态话语分析的基本问题进行了探讨[13]。

其次在多模态话语与意义构建方面, 研究对象不断拓展, 从漫画、海报、网页等静态的多模态话语延伸到广告、宣传片、课堂教学等动态的多模态话语[6]。张德禄、王璐探讨了大学英语课堂教学中不同模态的相互协同, 并指出教学的主模态是口头话语, 其它模态起补充、强化、指引、阐释等作用[14]; 姚银燕、陈晓燕基于一则企业形象电视广告, 分析了视觉模态和听觉模态在企业形象构建中的作用[7]。吴安萍、钟守满对西方报道的政治漫画进行多模态话语分析, 揭示了视觉元素与语言元素之间的互动以构建意义, 向世界传达文化内涵和交际意图, 展示国家形象[15]。

另外在研究框架和路径方面, 王立非、文艳介绍多模态分析软件 ELAN 及其主要功能, 以例证说明 ELAN 用于应用语言学研究, 探讨了多模态研究方法应用的价值和意义[8]。张德禄在已有多模态话语分析框架的基础上提出改进的新模式, 突出其系统性和实例化过程, 显化了符号系统的特性、符号间关系和再符号化过程, 建立了新的多模态话语分析的系统功能综合框架[16]。基于隐喻对意义的建构功能, 孙小孟等以大型纪录片《一带一路》第一集《共同命运》为例, 在认知分析框架下, 结合空间映射与概念整合理论, 从思想维度、价值维度和文化维度分析多模态隐喻对构建“一带一路”话语体系的积极意义[17]。

综上所述, 多模态话语的研究类型涵盖从理论探索到实证研究, 研究成果大多集中在外语教学领域[18]。多模态话语分析以定性研究为主, 质性研究和量化研究的个案比较有限。近年来, 网络多模态话语的研究方兴未艾, 大体从发展现状、研究主题和研究方法三个层面进行梳理。诸如以 QQ 群话语为例, 考察了这类动态交际中多种模态的融合[19][20]。

3. 研究设计

史兴松利用可视化软件 Citespace 对国内外多模态话语分析做综述研究时提出, 未来研究中应注重研究方法的完善, 利用分析软件以及语料库等增强实证研究的可信度[21]。自王立非推介 Elan 软件之后, 其使用度并不高, 研究成果也不多。因此, 本文借用 Elan 软件, 以武汉 2020 城市宣传片为研究对象, 探讨三个问题: 1) 武汉城市宣传片中不同模态各有什么特点? 2) 武汉城市宣传片中各个模态是如何构建意义的? 3) 武汉城市宣传片中各个模态是如何凸显主题的?

对动态视频的多模态话语分析, 需要结合两种路径, 即视觉语法以及系统功能语法理论框架, 以多

模态语料库技术手段为辅助。以音视频流切分与标注工具 ELAN6.0 软件作为研究工具, 对武汉城市宣传片进行多层次切分与标注并进行数据分析。ELAN 软件(EUDCO Linguistic Annotator)是由荷兰马普心理语言学研究所开发的一款便利实用的多模态语料分析工具, 为语言分析者提供技术支持[21]。ELAN 能实现对视频文件的层级同步标注, 包括话语内容、语音语调、面部表情、手部动作等。王立非、文艳总结了 ELAN 软件的以下优点: 视频播放精确到 0.01 秒; 标注与文本、声音、图像同步; 标注层无限, 标注编码表自行定义; 标注与标注间互相链接, 标注结果可根据需求进行多种格式的输出[8]。基于 ELAN 软件, 从音韵特征(T-/P-/St-)、图像(I-)、影像(A-)、表达方式(ES-)、音乐(M-)、镜头速率(LS-)、运镜方式(C-)等方面对武汉城市宣传片进行赋码编写。

第一类标注特征为音韵特征 P-[Prosody], 包括重音[PS_t]; 语速快[PFT]、语速慢[PST]; 长停顿[PPL]、短停顿[PPS]。第二类标注特征为图像特征 I-[Image], 包括文字[IT]、图片[IP]和文字—图片[ITP]。第三类标注特征是影像 A-[Animation], 分别为人影像[AP]、物影像[TP]、人—物影像[ATP]。第四类特征为表达方式 ES-[Expression], 包括旁白[ESN]、独白[ESM]、对话[ESD]。第五类标注特征为音乐 M-[Music], 包括背景音乐[BGM]、独奏音乐[MS]。第六类标注特征为镜头速率 LS-[Lens Speed], 包括慢镜头速率[SLS]、快镜头速率[FLS]。第七类标注特征为运镜方式 C-[Camerawork], 包括横摇[CP]、直摇[CT]、拉近/推远[CD]、平移[CTr]、鸟瞰[CB]以及焦点转移[CRF]。

根据叙事特点对视频进行时间段的分割, 对分段的视频内容进行标注, 将标注后的文件分为若干层级, 还可根据不同的需求对视频内容进行自主分层。本研究的分层主要涉及以下几个方面: 运镜方式、镜头速率、语速、重音、停顿、图像、影像、表达方式、音乐。对分割分层标注后的视频反复查看对比, 导出标注的数据, 结合视觉语法以及系统功能语法理论框架, 做进一步的量化分析。

4. 研究结果

(一) 视觉模态分析

根据视觉语法, 图像意义由三部分组成: 再现意义、互动意义和构图意义[3]。动态视频是由一帧帧的画面组成, 而且视觉模态的主体是影像以及图像。利用 ELAN 软件对宣传片进行视觉模态中的图像以及影像标注得出下表。

Table 1. Statistical table of visual modal annotation

表 1. 视觉模态标注统计表

模态	标注	命中	总标注时长	标注时长百分比
Animation	AP	20	73.81	33.573
	AT	21	51.91	23.612
	ATP	26	92.7	42.165
Image	IP	1	0.75	0.341
	IT	2	9.987	4.543

宣传片是经历新冠肺炎疫情背景下由武汉市文化和旅游局制作推广。根据表 1, 动态视频中图像的占比明显低于影像。在 Animation 标注统计表中, 人影像[AP]、人—物影像[ATP]的标注时长百分比都在 30% 以上。人影像[P]的出现标注时长百分比为 75.738, 表明该宣传片凸显以人为本的理念。

关于图像(Image), 该宣传片的标题贯穿始终。根据图 1 和图 2, 其布局方式起初位于左上角, 结尾转为中心。根据表 1, 图像文字[IT]、图片[IP]的标注数量偏少。



Figure 1. Location of key word (top left corner)

图 1. 主题字位置图(左上角)



Figure 2. Location of key word (center)

图 2. 主题字位置图(中心)

Kress & Leeuwen 认为, 图像中的参与者和环境对再现意义有着重要影响。宣传片中的人与景相融合无不彰显人与自然和谐共存的主题[3]。“相见在武汉”多使用叙事过程, 以武汉人的特色早点、一家三口欢笑画面、五次握手特写来表征城市的文化习俗、健康的生活方式、和谐的生活氛围; 视频低频率使用概念过程, 主要表征如黄鹤楼、武汉关、晴川阁、昙华林, 叙事意义和概念意义共同构建再现意义。互动意义强调图像制造者、图像中的事物和观看者的关系[3]。长镜头下的长江大桥、武汉电视塔和野营用的帐篷使观看者领略到城市的整体风貌。宣传片中的舞者、扬帆的男子和荡舟的女子等画面采用正面和水平视角, 有利于拉近与观众的距离。人与景物的协同使景物不再是独立的个体, 而是与人紧密相联, 无形之中彰显了人与自然的和谐统一。信息值、显著性与取景三者共同建构构图意义, 宣传片中以中心-边缘结构最多, 上下、左右结构为辅助, 其中人物的服装颜色以白色、红色和天空蓝为主, 增强显著性, 揭示主题意义。宣传片中包含大量灯光秀, 照亮整个夜空, 璀璨绚丽、赏心悦目, 充分展示繁华都市中的静谧。该宣传片不仅有武汉人眼中的武汉, 也有游客镜头中的武汉。不同视角以及运镜的运用都优化了意义的表达。

(二) 听觉模态分析

宣传片“相见在武汉”听觉模态的特点在于没有旁白、对话以及独白, 其中音乐模态对意义构建发挥了重要作用。根据表 2, 音乐模态标注时长占比为 96.247, 主要为背景音乐[BGM]和独奏音乐[MS]。背景音乐为钢琴曲以及和声, 独奏音乐为同名音乐《相见在武汉》。该曲目是陈文非(演唱者)专门为后疫情时代的武汉而创作, 其歌词包含武汉的地标景观。因此该曲目也为表达视频主题、构建意义发挥了积极作用。

Table 3. Statistical table of visual and auditory modal
表 3. 听觉模态和视觉模态数据统计

模态类型	层	标注数量	总标注时长	标注时长百分比
听觉模态	Tempo	10	210.7	95.838
	Stress	5	2.82	1.283
	Pause	3	7.86	3.575
	Music	2	211.6	96.247
视觉模态	Animation	67	218.42	99.35
	Image	3	10.737	4.884
	Lens Speed	10	215.859	98.185
	Cameraswork	29	85.784	39.019

由表 3 得知,“相见在武汉”以影像为传递信息的主要模态。在标注数量上,影像 A-的统计数最多,达到 67 个,包括人影像、物影像以及人-物影像。听觉模态中音乐 M-的标注时长百分比为 96.247,说明音乐模态也发挥了语篇意义的构建作用。

以宣传片开头为例,首先展示主题以及 logo 形象,以影像加文字的组合呈现,影像作为主要模态出现,通过视觉的直观感受,传达视频的主题,其它模态包括音乐和文字,音乐模态对影像起互补的作用,主要体现在非强化协调层面,音乐旋律缓慢,伴着钢琴的节奏,呼应整个语篇优美的前景化特征。在初始阶段,文字与图像呈现互补关系,通过文字和影像的融合,文字随着缓慢的音乐出现,以明示主题。

1. 模态内部协同

在视觉模态内,各分模态相互协作,促进视觉模态的发展。首先,使用运镜[C]与影像[A],增添气氛和情感。除了用来追踪主体之外,适当的横摇[CP]和直摇[CT]可以揭晓原本不在画面里的事物,让观看者了解环境和距离,甚至凝造惊喜的感觉。拍摄长江上的渡轮时,用拉近/推远[CD]手法让观看者有亲历其境的感觉,自然投入感情。采用平移[CTr]展示东湖上划船的红衣女子,周围的美景尽收眼底。鸟瞰[CB]常用于高角度拍摄,拍摄磨山时便是用该手法,让观众俯瞰全景,产生壮观和旷达的感觉。焦点转移[CRF]可让移动中的主体保持清晰,或者引导观众转移焦点。根据图 3,物影像[AT]大多运用鸟瞰[CB],如野营的帐篷;人影像[AP]大多运用平移[CTr]以及焦点转换[CRF],人-物影像[ATP]运用横摇[CP]和直摇[CT]以及拉近/推远[CD]。这些基于不同影像的运镜方式和影像的协同作用,达到宣传城市的目的,特别是拉近/推远[CD]运镜的使用增加了亲和力,体现“相见在武汉”的主题。

听觉模态内部间相互合作,辅助音乐和背景音乐共同构建语篇意义。根据图 3,重音[PS_t]与其他模态的关系十分密切。首先,每次重音都标志着新的旋律的出现;其次,三次重音的出现意为音乐节奏快慢的转换;第三,最后一次重音的出现标志着音乐类型的转换([BGM]-[MS])。此外,停顿[P]也与音乐相关:第一次停顿标志着音乐类型的转换([BGM]-[MS]),第二次停顿标志着音乐的停止。各分模态的协同作用促进听觉模态的意义构建。

2. 模态间协同

在多模态话语构建的语篇中,模态间的协作是必须的,也是必然的[22]。因此,充分利用动态视频的立体传播优势,使视觉模态和听觉模态两个系统相互协作、相互补充,便于观众从不同视角欣赏宣传片。

视觉模态与听觉模态之间的相互作用主要有两种:强化和补充。第一:强化,听觉模态的协同作用可强化视觉模态中获得的信息;第二:补充,听觉模态的协同作用可弥补视觉模态所缺失的或者观众没有完全接受的信息[22]。宣传片中,运镜速率[LS]与音乐节奏[T]的标注区间大致相同,影像[A]各标注类

型的转换速度与音乐节奏[T]联系密切。音乐节奏快的标注区间影像的标注类型转换频率较高, 相反音乐节奏慢的标注区间影像的标注类型转换频率较慢。而且, 影像标注转换频率较高的区间运镜方式呈多变趋势。此外, 音乐类型不同, 影像的转换频率也不同。在背景音乐标注区间内, 其影像的转换频率相对于独奏音乐标注区间较高。重音[PS_t]与运镜方式[C]之间也存在联系, 第一、三、四、五次重音标注之后运镜方式的转变频率开始变快。第一次停顿[P]前后运镜方式的转变频率随之发生变化。

5. 结语

综上所述, 武汉 2020 城市宣传片的多模态话语特点是视觉模态为主, 与听觉模态有效融合, 通过符号的互动构成和谐的整体, 表达意义。在视觉模态中, 人-物影像模态为主要模态, 而音乐模态是听觉模态的主体。模态之间的关系随时间的变化呈动态变化: 有时以影像为主, 音乐模态为辅; 有时以音乐模态为主, 影像为辅。根据图 3, 在独奏音乐标注区间, 影像的变化频率较低, 且随音乐而发生变化。标题文字、片尾独奏音乐歌词与图像形成互补关系, 两者之间关系的变化与交际事件的进程密切相关, 呈现视觉模态与听觉模态的动态关系。

张德禄指出“模态的选择涉及最佳搭配问题, 不是完全自由、随意的”, 可见不同模态的协同有助于整体意义的搭建, 弥补单一模态的局限性, 强化交际效果。从模态配合的角度, 不同模态不能自相矛盾、相互抵消、彼此无关、互不衔接等; 模态配合以增加正效应为原则, 如相互协调、相互强化等[22]。此外, 本文通过研究发现镜头速率以及运镜方式会强化视觉模态及听觉模态的意义表达。

参考文献

- [1] O'Halloran, K. (2004) *Multimodal Discourse Analysis: Systemic Functional Perspectives*. Continuum, London.
- [2] 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.
- [3] Kress, G. and van Leeuwen, T. (1996) *Reading Images: The Grammar of Visual Design*. Routledge, New York.
- [4] Baldry, A. and Thibault, P. (2006) *Multimodal Transcription and Text Analysis: A Multimedia Toolkit and Course book*. Equinox, London.
- [5] 代树兰. 多模态话语研究的缘起与进展[J]. 外语学刊, 2013(2): 17-23.
- [6] 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.
- [7] 姚银燕, 陈晓燕. 对视频语篇的多模态话语分析——以一则企业形象电视广告为例[J]. 外国语文, 2013, 29(1): 86-91.
- [8] 王立非, 文艳. 应用语言学研究的多模态分析方法[J]. 外语电化教学, 2008(5): 8-12.
- [9] 李德志. 广告类超文本多模态的视觉语法分析[J]. 外语学刊, 2013(2): 7-11.
- [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] 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.
- [12] 李战子, 陆丹云. 多模态符号学: 理论基础, 研究途径与发展前景[J]. 外语研究, 2012(2): 1-8.
- [13] 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.
- [14] 张德禄, 王璐. 多模态话语模态的协同及在外语教学中的体现[J]. 外语学刊, 2010(2): 97-102.
- [15] 吴安萍, 钟守满. 视觉语法与隐喻机制的多模态话语研究[J]. 外语与外语教学, 2014(3): 23-28.
- [16] 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.
- [17] 孙小孟, 贺川, 廖巧云. 多模态隐喻与“一带一路”话语体系建构——评析大型纪录片《一带一路》第一集《共同命运》[J]. 外国语文, 2020, 36(3): 68-74.
- [18] 程瑞兰, 张德禄. 多模态话语分析在中国研究的现状、特点和发展趋势——以期刊成果为例[J]. 中国外语, 2017, 14(3): 36-44.

- [19] 耿敬北, 陈子娟. 网络社区多模态话语分析——以 QQ 群话语为例[J]. 外语教学, 2016, 37(3): 35-39.
- [20] 史兴松, 徐文娟. 近十五年 SSCI 期刊网络多模态话语研究现状及发展趋势分析[J]. 外国语, 2020, 43(3): 55-66.
- [21] Van Uytvanck, D. (2007) EUDICO Linguistic Annotator (ELAN). Version 3.2.0, Manualun.
- [22] 张德禄, 袁艳艳. 动态多模态话语的模态协同研究——以电视天气预报多模态语篇为例[J]. 山东外语教学, 2011, 32(5): 9-16.