

# 多模态话语建构中的模态融合模式研究\*

同济大学 张德禄

**提要:**本文主要从多模态话语建构的角度探讨多模态话语如何产生,重点探讨文化语境、情景语境和交际意图等语境因素如何促动多模态话语建构,哪些意义特征参与构建多模态话语意义,多模态符号又是如何相互配置和融合,从而共同建构话语的。结果发现:1)多模态话语涉及多种符号系统的意义特征,包括符号实体的物理特性以及符号媒介在承载符号意义方面的特性等;2)多模态话语中的符号之间存在主次、核心外围、前景背景、对等互补以及伴生等配置关系。上述发现表明,多模态话语建构是一个十分复杂的过程,需要考虑符号的可及性、专业性和智能性、供用特征的互补性、模态之间的相容性以及模态协同和配置的动态性和平衡,还要考虑模态符号选择中应该规避的矛盾和内耗等问题。本研究对社会交际中如何利用多种符号资源构建适切的多模态话语具有一定启示。

**关键词:**多模态话语;符号系统;模态配置;模态融合

[中图分类号] H05 [文献标识码] A [文章编号] 1003-6105(2023)04-0439-13

DOI:10.20071/j.cnki.xdwy.20230420.012

## On the Models of Modal Integration in Multimodal Discourse Construction

ZHANG Delu Tongji University

**Abstract:** This paper is mainly concerned with how multimodal discourse is generated from the perspective of multimodal discourse construction. Specifically, it investigates how the construction of multimodal discourse is motivated by such contextual factors as culture, situation, communicative intention and so on, what semantic features are employed to construct the meaning of multimodal discourse and how modal signs are synergized and integrated to construct multimodal discourse. It is found that multimodal discourse has the semantic features of various semiotic systems involved, including the physical characteristics of entities in question and relevant media features involved in realizing the meaning of the sign, etc. Signs in multimodal discourse are related to each other in such ways as main-dependent, core-peripheral, equivalent-complementary, foregrounding-backgrounding, and accompanied-accompanying. The construction of multimodal discourse is therefore a complex process, which needs the designer to consider the accessibility of the signs, the field of specialty and intelligent nature of the signs, the complementarity of affordances and their characteristic uses, the compatibility between modes, the dynamics and

\*本文为国家社科基金项目“多模态话语建构中模态配置的机制和机理研究”(22BYY002)的阶段性成果。

balance of mode coordination and patterning, and the contradictions and internal friction that should be avoided in the selection of modal signs. Hopefully, this study can shed some light on the construction of appropriate multimodal discourse through multimodal resources in social communication.

**Key words:** multimodal discourse; semiotic system; multimodal synergy; modal integration

## 1. 引言

多模态话语研究从诞生之日起 (O' Toole 1994; Kress & van Leeuwen 1996/2006/2020), 到现在已有近 30 年时间。在这段时间里, 研究重点基本集中在多模态话语的理解和分析上, 而对多模态话语的建构则研究不足。另外, 多模态话语的建构既涉及不同的模态 (modality) 或符号系统 (semiotic systems), 也涉及相应的媒介 (medium) 和物质实体 (physical entities)。也就是说, 话语的意义不仅由模态体现, 同时也由与这些模态相关的媒介和物质实体体现, 因而需要从多个层面来探讨多模态话语意义的建构。这就涉及对符号的供用特征 (affordances) (Gibson 1979) 的研究, 以及对多模态话语建构中模态选择和配置模式的研究。

多模态话语的建构涉及创作或生产多模态话语, 属于多模态话语写作。较早进行这类研究的是 Jewitt (2005), 重点探讨了多模态与 21 世纪的阅读和写作的关系。在写作部分, 她探讨了屏幕语篇的建构。屏幕语篇都是多模态语篇, 涉及图像、声音、动画及其他交际模态。她认为, 在多模态话语建构中, “不同的模态提供不同的具体资源来构建意义, 不同模态对交际意义的贡献不同。” (Jewitt 2005)。她提供了一个运用不同模态的符号资源来建构话语的例子: 游戏平台 (PlayStation) 上的一个游戏 Ico, 说的是 Ico 和他的朋友 Yorda 如何从坟墓中经过艰难险阻、战胜群魔, 胜利逃脱的故事。其中, 对本游戏的核心人物 Yorda 的短瞬性特征的建构是通过多模态设计进行的。Yorda 的短瞬特性表现在其话语和对其话语的书面转写上, 也表现在其不确定、易变、模糊不清的身形上, 同时还通过其平静的声音、既软又慢的鬼魂式的、不确定和不完整的手势以及漂浮的运动体现出来。体现 Yorda 这个人物的每一种模态都表示相同的概念: 她生活于城堡内外交界处的狭小空间内, 所以要想办法逃脱。

儿童教育是每个家长都特别重视的事情。以前的语言发展教育主要培养儿童的语言能力, 特别是读写能力 (literacy), 但现在的儿童语言发展教育同时包含其他能力, 使其整体素质同步提高。因此, 所涉模态不仅包括听说, 还包括唱歌、跳舞、表演、绘画、照相、看录像等。Switzer (2009) 报道了在美国的巴西移民家庭如何教育子女的调查报告, 调查形式为家访, 孩子是 3-4 岁的儿童。家长及其他家庭成员都参与儿童能力发展的教育。结果表明, 儿童语言和读写能力的发展是通过其与父母的交流进行的。这种交流并非单纯靠语言, 而是一种多模态构建。因此, Grillo (1989: 15) 用“交际实践” (communicative practice) 来指通过语言

进行的社会交际活动,并且把读写能力看作一个大的社会语境中的一种交际活动。这样一来,多模态语篇就指通过多种手段表达意义的语篇,可以包括口语或书面语词汇,也可以不包括口语或书面语词汇 (Walsh 2006)。多模态在此包括图像(如照片、图画、绘画)、声音语篇(如声响和音乐)和身体运动模态(如舞蹈、表演、电影)等。每个模态都是一个交际系统,因此理论上同等重要。Grillo (1989)的研究对象是居住在同一个区域的孩子,有的父母已来美国 10 年,有的则更短。他们的英语水平跟来美国的时间相匹配。教学活动则包括教儿童唱歌或学习音乐,用语言指导他们用手指弹奏音乐,或者唱 ABC 歌。这样,唱歌、音乐、口语、手势等模态同时用来建构多模态教学话语,完成教学活动。在此过程中,孩子的整体素养,包括语言能力、动手能力和感知能力,都得到了发展。

还有一个与多模态话语建构具有天然亲密关系的学科是修辞学,因为修辞学研究的重点是话语生成,从古希腊时期的口头演讲,到现在的写作研究,都是如此。但以前修辞学主要研究语言修辞,不强调其他模态的作用。然而,随着最近几十年来视觉修辞 (visual rhetoric) 的不断发展,修辞学成为了多模态话语建构的主渠道。例如, Murray (2009: 137) 探讨了写作中的非话语修辞,即图像修辞。非话语修辞是用以描述非语言语篇的修辞实践,依赖(充满情感的)图像和非语言意义来劝说、打动观众或创作不可言说的(或非语言)意义的一种修辞理论。Murray (2009: 138-139) 认为,符号化不是通过语言实现的,而是通过图像实现的,据此,人的一切心理和情感活动(如想象、意图、意识、情感等)都要通过图像来实现。最终,话语表达中的问题也可以通过图像来实现,特别是那些模糊的、直觉的、难以说出口的,或难以言说的,都可以通过图像来表达。Murray (2009: 166) 还认为,“照片、字体、颜色、空间、声道都是修辞性的。”他设计了一个“非语言话语写作模式”(non-discursive discursive composing model) (Murray 2009: 151),把形象、意向、情感、感觉、味道等都直接纳入话语建构中,而不是集中于命题和知识。

以上研究对语言话语如何扩展为多模态话语以及多模态话语如何组成做了有益的探索,但还有两个问题需要进一步研究。一是多模态话语建构在不同的层次上有什么不同的特征;二是多模态话语建构中每个模态的潜在意义是如何在话语中现实化的。下面重点探讨这两个问题。

## 2. 话语与多模态话语

研究话语的建构,首先要清楚话语是什么。概括地讲,对话语的定义可以分为两个大类。一类将话语视作社会交际的过程和产物,另一类将话语视作知识类型。前者以 Halliday 的系统功能语言学为代表,后者以 Foucault (1977 1984) 为代表。首先,从前者来看, Halliday 从社会符号的角度研究语言,认为“语言不是由句子组成,而是由语篇,或话语组成——是在某种人际语境中交换意义”

(Halliday 1978: 2)。在此,话语是“在某种人际语境中交换的意义”,且话语和语篇没有区别。接着,他又说,“这涉及既要把注意力同时聚焦于探索实际和潜势,又要从我们称为文化的、无限复杂的意义潜势网络的角度释解话语和它背后的语言系统的难题。”在这里,话语与语言系统相对,因而话语是实例、实际,不是系统。Fairclough (1992: 8) 认为,“语言用途——‘话语’——给社会定型,并被社会定型”……“很常见的情况是,语言运用是由社会决定,且语言要随其运用语境的变化而变化。”在此,话语是语言用途,是社会运用,因而也是语言运用实例,不是语言系统。

从把话语作为知识类型的角度讲,Foucault (1977: 182-183) 认为,“知识是可在话语实践中讲的东西。”“任何话语实践都可通过其所形成的知识来定义。”据此,van Leeuwen (2005: 94) 和 Kress & van Leeuwen (2001: 4) 都把话语定义为“通过社会建构的有关现实的某个方面的知识。”

从把话语作为交流的意义和作为语言运用的实践的角度讲,话语就是意义,属于 Martin (1992) 所称的“话语意义”(discourse semantics)。而从把话语作为知识的角度讲,话语是在社会文化中能够被语境选择的“话语意义”,所以,“作为知识的话语”需要经过从系统到实例,再被符号化的过程,才能成为要交流的意义。

这两种意义的话语也都是多模态话语中的话语,分别由两种多模态话语分析的理论模式使用:系统功能语言学模式把话语作为意义,在话语意义层上与语篇同义。在此,参照 Stockl (2004: 10) 对多模态的定义,多模态话语可以定义为:“通过联合不同的符号系统(模态),并由交际者把所有出现的符号在意义和形式上相互联系而产出(或输入)的意义模式”。社会符号学模式则把话语视作在社会中建构的知识。在此,参照 Kress & van Leeuwen (2001: 20) 对多模态的定义,多模态话语可定义为:“在设计符号产品或事件中运用多种符号模态并将这些符号模态相互结合而产出的知识构型”。两者的区别在于:如果把多模态话语视为意义,则其可直接由模态及其集合来体现;如果把多模态话语视为知识,知识又在语境或交际者的大脑中,则其需要转换为意义才能被符号模态直接体现。

### 3. 多模态话语选择的动因

如果话语是人类进行社会交际的产物,那么任何话语都产生于一定的语境中。根据系统功能语言学理论,语境包括文化语境和情景语境(Halliday 1978; Halliday & Hasan 1989)。文化语境为本言语社团的成员提供了从事社会交际的总体环境。从多模态话语的角度讲,文化语境为本言语社团的成员提供了生存、生活、发展、交际、创新等所有涉及人类互动的环境,并且大体确定了他们的行为模式,称为“行为潜势”(behaviour potential) (Halliday 1973: 49),也确定了他们进行交流的意义范围,即意义潜势(meaning potential)。这样,多模态话语的意义就来源于文化语境,即文化语境所限定的意义潜势。

文化语境限定和支配人类交际的思维模式,包括通常观察到什么、做出何种反应、思考什么等;行为模式,包括经常做什么事,遵循什么交际常规等。人的思维模式和行为模式可共同决定人类在交际中为了完成某个交际任务,达到某种交际目的而选择的交际模式:体裁(genre),即某个交际事件的程序,包括开始、发展、解决问题、结束等阶段。Martin (1984: 25)把体裁定义为,“某个文化中的社团成员按阶段、有目标、有目的的交际活动”。在语言学研究中,“文化语境”一般局限于语言的文化语境。但在多模态话语建构中,体现话语意义的所有模态都包括在内,所以,“文化语境”成为使用所有的符号系统进行交际的文化语境(见张德禄 2016),实际上指某个社团的整个文化。

文化语境只是从宏观上限定和支配人类的交流模式。具体采用哪个模式,还需要在具体的情景语境中确定。根据系统功能语言学理论(Halliday 1978; Halliday & Hasan 1989),情景语境包括三个主要变项:1)语场:发生了什么事件,包括题材;2)语旨:谁和谁讲话,包括交际者之间的关系;3)语式:通过什么交际渠道和媒介进行。情景语境促动了对话语意义的选择,包括表现客观现实、逻辑关系、人际关系、态度、情感、观点、主题、信息等意义。

系统功能语言学对情景语境的定义也适合于多模态话语研究,但要扩展其涵盖范围:从语场的角度讲,它包括由所有的符号系统体现的事件和交际活动(见张德禄 2018)。从语旨的角度讲,多模态话语的交际者更加广泛,包括使用语言和不使用语言进行交际的所有交际者及其关系。从语式的角度讲,在多模态话语建构中,社会交际的范围由语言扩展到通过所有符号系统进行的交际。Martin (1984)在探讨语式时,提出要通过两个连续体来具体说明话语类别,分别是:1)空间/人际连续体,即从空间距离近、人际关系密切到空间距离无限远、人际关系距离无限远之间形成的连续体,包括面对面有视觉、听觉,甚至触觉接触的口头对话、打电话、写微信信息、写短信、大会口头报告、写信、多模态语篇制作、写小说等。2)经验距离,指语言在交际中按作用大小形成的连续体,从主要活动是行为、物理活动,语言只是起辅助作用,到纯粹的语言活动,如写作等。这两个连续体是直接相互联系的:空间/人际距离近,使视觉、听觉、触觉等直接接触成为可能,同时,语言只起辅助作用的社会活动通常是面对面的交流,如聊天、玩游戏、合作修车等。

这样,多模态话语建构的动因可以描述为:交际者生活在一定的文化语境中,其思维模式和行为模式要受到这个文化中交际模式的制约。在一定的情景语境促动下,他需要根据自己的环境和交际目的选择合适的体裁,并且根据语场、语旨和语式选择具体要交流的意义和行为,从而形成多模态话语建构的基本模式。

#### 4. 多模态话语的意义建构

如上所述,话语来源于文化语境,即文化语境给话语的产生和存在提供了环

境,产生于情景语境,即根据情景语境和交际目的选择要交流的意义,从而产生话语。传统上一般认为,话语是对客观现实的表达,但实际上,由于受语境因素和交际目的的制约,话语不是对客观现实完整真实的记录和表现,而是要对事实进行修改、删减、突出、淡化等处理,从而使其成为讲话者要表现自己的意义的话语。在此,话语的两个定义可融为一体,用于描述话语的生成过程。

把话语作为在社会中建构的知识的观点实际上是把话语作为意义的观点的话语建构过程,即话语实践(Kress & van Leeuwen 2001)。在这个过程中,交际者要选择合适的话语来实现自己的交际目的,就需要汇聚其要表达的意义所涉及的相关知识,从中选择出有利于实现自己的交际目的的知识类型,并把它们聚集在一起,成为自己的意义。这些知识来源于他要讲述的交际事件。这样,他可以对原交际事件进行客观的描述。但在大多数情况下,对事件的直接描述并不符合他的目的,因此,要对其进行加工处理。

首先,交际者可以对原事件进行改编、删减等,如在新闻报道中,选择自己感兴趣的部分进行报道,舍去自己不感兴趣的事件;删减自己不喜欢的参与者,突出个别对自己有利的参与者及其事迹等。例如,在俄乌战争报道中,俄方报道(2022年11月13日):“俄罗斯国防部称,俄军在顿涅茨克地区和卢甘斯克地区击退了乌军的军事行动并继续发动进攻,占领了马奥尔斯克村,同时打击了乌克兰空军位于切尔卡瑟的一座军火库。”而乌方则报道(2022年11月13日):“乌军13日在顿涅茨克地区和卢甘斯克地区击退了俄军的进攻,并继续在赫尔松地区第聂伯河右岸进行排雷行动”(人民网+,2022年11月22日)。双方的报道存在矛盾,因此这里需要理清的问题是:是否双方报道的都是事实?如果是,则他们各自的“胜利”可能发生在不同的时间段。如果不是,则可能一方的报道是颠倒是非,也可能双方报道的都不是事实。

第二,交际者还可以在表现客观实际的同时,附加上讲述者的政治观点,以及情感、态度、好恶等,或者根据交际者自己的观点、态度、好恶来对客观事实进行取舍或者改编。例如,美国常驻联合国代表格林菲尔德在联合国安理会就乌克兰局势召开的紧急会议上表示,“俄罗斯今日的决定是为战争制造借口,其行动将带来严重后果。如果俄罗斯选择进一步“侵略”乌克兰,世界将会看到大量的生命损失和流离失所乃至难民危机”(人民网+,2022年2月21日21:00)。在此,格林菲尔德把俄罗斯的决定定性为“制造借口”,认为这种行动将“带来严重后果”,并且把俄乌战争定义为俄罗斯“侵略”乌克兰,认为将会造成“大量的生命损失和流离失所乃至难民危机”。这里不仅涉及对俄罗斯的负面态度和观点,而且还在对事件性质和结果的判定上持负面观点。而普京则认为,“乌克兰已“丧失主权”:沦为美国外交工具和生化试验场”(人民网+,青蜂侠Bee 2022年10月27日06:51)。在此,“丧失”和“沦为”不仅表现了普京对乌克兰及西方世界的负面态度,同时,也把这两个事件描述为负面行为。

最后,同一个事件,或称“话语域”(discourse domain),还可以通过不同的渠道和方式表现。如其可以根据不同的交际目的,通过不同的体裁表现。例如,海湾战争可以通过新闻进行报道,也可以通过闲聊讲述、编成电影、电视剧等播放。除了讲述实际事件以外,还可以增加评论、判断、假设、评价、批评、褒奖等;在电影和小说中,还可以增加虚构成分等。

然而,意义不能独立存在,必须由特定的符号或符号组合加以体现,如由语言体现,或者由语言与其他模态配合体现,还可以由语言之外的符号,如图像、动画等体现。这就意味着,选择了要表达的意义就需要选择合适的模态或模态组合来体现它。每个已经确立的模态或符号系统都由一组相互联系的符号组成,最终形成一个系统,具有相互对立、意义互补、随时供选择等特性。语言系统是最成熟的意义系统。但也有的符号系统会慢慢过时,会被淘汰;也有的正在成长进化,没有形成完整的符号系统,也就是说,这个系统所涵盖的意义范围并不完整,也不完善,而是处于完善和发展中,如某些网络话语符号系统。

## 5. 多模态话语建构中模态的选择和配置

多模态话语的意义要通过选择多种模态来体现,涉及多种模态的协同和配置。多模态话语的建构和语言话语的建构是有区别的。在语言话语中,由于参与交际的符号资源已经确定,因而就没有必要审视已有的交际资源,只需考虑表达什么意义,然后用语言(词汇语法)体现即可。但在多模态话语中,话语的构建不仅要考虑什么样的意义模式更加适用,同时还要考虑有哪些模态(符号资源)可供即时选用,从而实现交际目的。这一方面扩展了多模态话语构建意义选择的范围,另一方面也增加了选择合适的已有符号资源的任务。

这样,与语言话语的建构相比,多模态话语建构的模态选择和配置所遇到的新问题包括:1) 对话语意义的选择过程也会受到已有符号资源的制约。这个过程可以是隐性的,也可以是显性的,还可以是两者的结合。例如,交际者由于受到文化语境的熏陶,会自然地意识到在本文化语境中存在哪些现成的符号系统,因而会自然地或典型地选择相关意义模式,并可以通过选择某个符号系统的符号或者几个符号系统的符号,使它们融合来体现话语意义。交际者也可能意识到,“我应该选择这个意义模式,但似乎找不到合适的模态来体现它”,从而出现支支吾吾,或不知所措的现象,或者另取其次,找到尽可能具有相似效果的话语意义模式。从这个角度讲,已有符号资源作为交际话语构建的条件在一定程度上也调节了话语意义模式的选择。2) 模态选择既可以在无意识状态下进行,也可以经过精心设计,主要取决于语式和体裁模式。在所谓“语境依赖性强”的语境中,如面对面会话、执行紧急命令、遇到紧急情况等形势下,交际者需要迅速做出反应,对模态的选择以触手可及的符号系统为主,甚至可以产生无效模态,如面对面对话随机产生的焦虑行为、手势和身势状态等;或者产出负向模态,如虽

然交际者的交际目的是要尽量表现对对方的尊重和敬意,但发话者的面部表情却使受话人感觉到其愤怒或不满情绪;抑或无意识模态,如讲话者随机产生的无意识手势;甚或是模态的交叠和重复等现象,如讲话时随机产生的有一定效果的手势和面部表情。而在连续体的另一端,发话者能够自己设定交际对象,自主选择要交流的话语模式。在这种情况下,发话者可以静下心来精心设计自己要选择的话语模式,还可设计可以体现这个话语模式的模态及其组合模式。这时候,模态之间的配合就会更加密切,不会出现模态选择的无效、负向和交叠重复等现象。

多模态话语的模态选择和配置过程需要话语建构者理清和认识模态选择中的符号及其系统所表现出的几个特征:

1) 理清和认识多模态话语建构所涉及的不同类型的符号系统及其关系。例如,面对面口头对话可能需要选择的模态包括:口头话语、手势、面部表情、口气、语调、身体姿势、空间距离。其中,口头话语是实现发话者交流意义的主要模态,手势和语调一般用以强化、澄清、明晰化要表达的意义,而口气和面部表情则表现发话者对于所表达的意义附加的态度和情感,身体姿势和空间距离则体现交际者之间关系的密切程度以及社会地位和角色等。

2) 要熟悉并深刻认识所使用的模态的系统特征,能够在交际过程中熟练地选择使用。也就是说,发话者在话语构建中选择模态时,需要熟悉所涉及模态系统所具有的特点,如口语是即时的、现场的、听觉的等,也要熟悉这个模态的系统是什么样的,如口语是一个具有三个层次的、线性的符号系统,还要熟悉这一模态系统有哪些次级系统及其选项,每个选项又可体现什么意义等。在此,除了口头话语是每个人必备的能力之外,交际者还需熟练掌握手势系统有几个选项,各自可体现什么意义,有什么作用和效应,是否为这个场合中可用的模态。如果用错,则一方面可能产生误解,另一方面也可能产生负效应,甚至使交际中断,例如不知不觉运用了禁忌手势。其他模态的选择也是如此。例如,不合适的面部表情可以使受话人产生厌恶情绪而导致交际失败。再如,在社会地位差距较大,相互又不熟悉的情况下,体距太近会产生不良效果。

3) 要熟悉并深刻认识所选符号,特别是核心符号的供用特征。交际者需要审视整个符号派生于各个层次的意义。根据社会符号学理论(van Leeuwen 2005: 3-4),社会符号被看作资源,而资源只是能指(signifiers),是进入社会交际领域的可见的行动和物体,具有理论和符号潜势,由其所有以前的实际使用和将来的用途组成。这样,它就不是简单的能指和所指的组合,而是要到社会交际实践中去发现其所指,根据语境来给其匹配合适的所指。从这个角度讲,作为资源的符号实际上是几种潜在意义的集合:

一是物理特性:由实现符号的物质实体表现出来的物理特性。这个特性需要符号依附于其物质实体来体现语境中的意义。

二是媒介特性:符号媒介在承载符号意义方面所表现出的特性,例如,玫瑰花不仅是植物生命体的花朵,而且有美丽的特性。

三是典型的潜在意义:符号潜在意义的核心部分,是这个符号在社会语境中惯常体现的意义,如V形手势表示“胜利、成功”等。

最后是潜存于符号媒体中,没有被发现,有待于将来去发现的意义(van Leeuwen 2005: 5)。

这四种意义都是符号潜在的意义,即在符号被选择后,易于被前景化、纳入符号所体现的整体意义中的意义。但符号实际的所指是根据语境选择的意义,因而可以包括这四种意义,或者其中的部分意义,但更有可能选择语境所激发出来的新的意义,即实指意义,从而成为转换的符号资源(New London Group 1996/2000)(见表1)。

表1 社会符号的能指和所指

社会符号性质	物质实体	媒介特性	典型潜在意义	供用特征	实指意义
社会符号学对意义的研究	研究它对供用特征的影响	能指(资源)			所指
		研究它对意义的承载作用	供用特征的核心部分	供用特征的主要部分	实指意义
	潜在符号意义				实际符号意义

4) 要熟悉并深刻认识符号意义在话语建构中的作用:在多模态话语中,每个模态在整个话语建构中的作用是不同的。例如,在面对面的对话中,口语模式是主模态,承担体现话语整体意义的核心部分,而手势、面部表情等只起辅助作用,身势、体距、空间布局则补充口语模态体现的意义。根据模态间在不同话语模式中的作用及其相互关系,模态间关系可以包括:1) 主-次关系;2) 核心-外围关系;3) 对等互补关系;4) 前景-背景关系;5) 伴生关系等(参见张德禄 2009)。

首先来看主-次关系。主-次关系表示,在多模态话语意义模式建构中,一个或几个模态在体现话语意义模式上起主要作用,同时出现的其他模态则对其起辅助作用,包括修饰作用、完善作用、补充作用、强化作用等,如上例中的口语与其他模态在面对面交流中的作用。

其次是核心-外围关系。核心-外围关系表示某个模态在话语意义建构中居于核心位置,其他模态则从边缘上完善修饰它。例如,有一则公益广告,宣传“讲文明、树新风”。该广告用一个黄色发光的灯泡内嵌上“讲文明、树新风”,同时闪闪发光,而灯泡的周边是文字,由23个四字句组成,都是讲文明、树新风的具体表现。这样,图像“灯泡”为核心模态,而文字则是外围模态(见图1(1))。

再次是对等互补关系。对等互补关系表示两种或多种模态共同体现话语的意义模式,而且每个模态所体现的意义不同,模态相互补充,共同形成一个整

体。例如,图 1 (2) 是一个宣传不要在高空抛物的公益广告,主要包括两个部分:一是文字:“它的杀伤力远大于想象”;二是图像:一只高空抛下的香蕉,顶端变成针形尖头往下落。在此,图像演示物体如何从高空抛下,文字表示这种行为的危害性。最后一句话把它们融合为一体:“严谨高空抛物 创造和谐社区”。

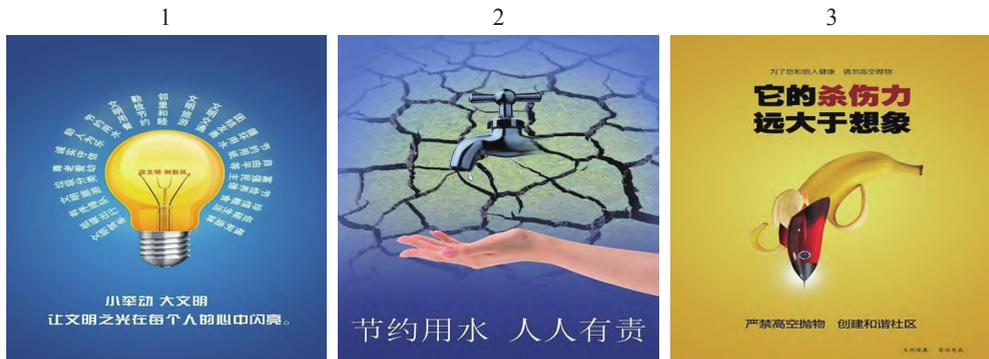


图 1 多模态话语建构中模态之间的关系

再来看前景-背景关系。前景-背景关系表示,在某个多模态话语意义模式建构中,某个模态居于前景中心突出位置,表示该话语的核心意义,而其他模态居于背景之中,对其进行注释或修饰。例如,图 1 (3) 宣传“节约用水,人人有责”,居于前景的是图像,水管只流出一滴水,伸手接不到水,而背景也是图像,表现干枯的土地,下边的文字则提示,“节约用水,人人有责”,背景中的文字模态则有画龙点睛的作用。

最后来看伴生关系。伴生关系表示,在多模态话语建构中,主模态或核心模态会顺便携带其他模态进入话语意义模式,虽然它们仅仅起到辅助作用,甚或是无效模态。例如,在口头对话交流中,有些情况下,手势是可有可无的,因为话语已经把整体意义体现出来,手势是无意识情况下产生的。和话语伴生的还有面部表情、声响、音高等。

## 6. 讨论

多模态话语的建构是一个十分复杂的问题。有的是即时构建,有的则是经过较长时间的准备和计划才能完成。有的主要是语言或抽象的话语,有的则主要由行动、行为体现。有的以情感态度为主,有的则主要表现客观现实。这样,交际者在选择实现话语构建的模态时,需要根据即时的情景语境选择合适的模态来构建话语。交际者选择构建话语的模态涉及以下几个方面的因素:

1) 符号的可及性:在不同话语方式中,话语构建时符号的可及性差别很大。在面对面、无时间间隔的交际中,交际者没有时间和精力去选择现场不可及的模

态,只能利用触手可及的模态来进行交际。如在日常生活会话中,交际者相互偶遇,需要马上进行交际活动。交际者只能选择最可及的模态,如口语、手势、面部表情、身体动作等进行交际,难以利用书写、有计划的行动、不在现场的模态来进行交际,尽管这些模态可能更加有效。例如,路遇同事问一个专业方面的问题,你需要用工具、仪器、设备来展示,但这些模态不可及,你只能选择用口语、手势等来加以说明。在社会交际中,可及性最高的是交际者本身产生的符号,如口语、手势、身势、行动、行为、面部表情、头部运动等。其他的都要根据现场语境来确定。

2) 专业性与智能性:有些符号主要适用于某些领域,如法律、医疗、商务、科技等,在空间上具有一定局限性。这类符号的选择需要具备该专业的合适条件,如在法院、公安局构建法律话语等。而有些符号则具有普通符号不具备的能力。这些模态符号可以打破前面提到的时间、空间、专业等的局限,具有更广泛的普适性,如现在流行的微信、网络、网络电话、动画、机器人等实现的话语。这类模态可以在很大程度上提高交际者的话语构建能力(交际能力)。

3) 供用特征的互补性:在多模态话语研究领域公认的一点是,不同类型的模态具有不同的作用和意义特征,即供用特征。在多模态话语建构中,当一种模态不能确切地实现交际者的目的,表达他的意义时,就需要另一种,或多种模态来补充、完善、强化它。这种情况下,模态之间在多模态话语构建中是互补的。这也是多模态话语能够产生的真谛所在。在不同模态互补的多模态话语中,每个模态的作用并不相同,而是根据各自的意义范围和语境的需要各自占据不同的地位,如核心模态、主模态、平衡模态、辅助性模态、伴生模态等(见上节)。

4) 模态之间的相容性:多模态话语建构需要多种模态的相互协同和融合,但不是任何模态都可置于同一语篇而组成多模态话语。它们必须是相容的。例如,大会发言需要通过口语、面部表情、手势、稿件等共同体现,也可以扩展到PPT上的动画、文字、口语、图像等模态来辅助体现,还可以进一步扩展为放录音、录像进行。但不能以手势和书面语融合体现(除非讲话者是聋哑人)。这类例子还很多。知道哪些模态是相容的、互补的、相互强化增效的,知道哪些模态是不相容的,也是社会交际能力的一种表现。

5) 模态协同和配置的动态性和平衡:以上四个要素在每次多模态话语建构中都可能遇到,而它们之间也需要以一定的关系达到动态平衡,基本标准是所建构的多模态话语是否能使交际者达到交际目的,取得最佳效应。例如,在可及性和专业性之间可能存在矛盾,可及的模态达不到专业的要求,这样,交际者可以有多种选择,如以后补救、通过其他手段弥补或采用新的智能模态。

6) 规避矛盾和内耗:研究每个模态符号的供用特征,一方面是为了使它们更加贴切地相互协同和配合,完成好社会交际任务,另一方面也是为了避免在多模态配置中出现重叠、矛盾、无效等现象。首先,重复是常用的手段,即两

种或多种模态体现几乎相同的意义,如说、画、看相同的内容,一般用以强化、提高效应。但过多不必要的重复则产生冗余、累赘现象。两个模态相互抵消可以使多模态话语失去效应,如上课时大声放音乐,使音乐声音淹没老师讲课的声音。选择理解难度很高的模态来体现比较简单的意义也会产生相似的效果。

由此可见,探讨多模态话语建构是一项十分复杂的工程,需要研究者根据语境和交际意图,综合所选择模态的供用特征,使它们合理配置,以产生最佳搭配效应,产出最适切的多模态话语为准(张德禄 2010)。

## 7. 结语

综上所述,多模态话语是人类进行社会交际的产物,是在语境和交际意图的驱动下产出的意义模式或知识构型。这个意义模式是由从多种符号系统中选择的符号相互配置和融合体现的,因此,它就具有所涉及的多种符号系统的意义特征,包括符号实体所具有的物理特性、符号媒介在承载符号意义方面所表现出的特性、符号典型的潜在意义以及潜存于符号媒体中有待于将来去发现的意义。由于符号的不同意义特征在语境中的作用的不同,每个符号在多模态话语建构中的作用也不相同,因而符号之间在多模态话语中就有了主-次、核心-外围、对等互补、前景-背景、伴生等配置关系。总之,多模态话语建构是一个十分复杂的过程,需要考虑符号的可及性、符号的专业性和智能性、供用特征的互补性、模态之间的相容性、模态协同和配置的动态性和平衡,还要考虑模态符号选择中应该规避的矛盾和内耗等。希望本研究能对社会交际中如何通过多种符号资源产生适切的多模态话语具有一定启示。

## References [参考文献]

- Fairclough, N. 2003. *Analyzing Discourse*. London: Routledge.
- Foucault, M. 1977. *The Archeology of Knowledge*. London: Tavistock.
- Foucault, M. 1984. The order of discourse. In M. Shapiro (ed.). *Language and Politics*. Blackwell, 48-77.
- Gibson, J. J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Grillo, R. 1989. Anthropology, language, politics. In R. Grillo (ed.). *Social Anthropology and the Politics of Language*. Cambridge: Cambridge University Press, 1-24.
- Halliday, M. A. K. 1973. *Explorations in the Function of Language*. London: Edward Arnold.
- Halliday, M. A. K. 1978. *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. London: Edward Arnold.
- Halliday, M. A. K. & R. Hasan. 1989. *Language, Text and Context: Aspects of Language in a Social-Semiotic Perspective*. Oxford: Oxford University Press.
- Jewitt, C. 2005. Multimodality, “reading”, and “writing” for the 21<sup>st</sup> century. *Discourse: Studies in the Cultural Politics of Education* (26)3: 315-331.

- Kress, G. & T. van Leeuwen. 1996/2006/2020. *Reading Images: The Grammar of Visual Design*. London: Routledge.
- Kress, G. & T. van Leeuwen. 2001. *Multimodal Discourse: The Modes and Media of Contemporary Communication*. London: Arnold.
- Martin, J. R. 1984. Language, register and genre. In F. Christie (ed.). *Children Writing: A Reader*. Geelong, Vic.: Deakin University Press, 21-9.
- Murray, J. 2009. *Non-Discursive Image: Image and Affect in multimodal composition*. Albany: State University of New York Press.
- New London Group. 1996. A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review* 66(1): 60-92.
- O’Toole, M. 1994. *The Language of Displayed Art*. London: Leicester University Press.
- Stockl, H. 2004. In between modes: Language and image in printed media. In E. Ventola, C. Charles & M. Kaltenbacher (eds.). *Perspectives on Multimodality*. Amsterdam: John Benjamins, 9-30.
- Switzer, S. C. 2009. Multiple modes of communication of young Brazilian children: Singing, drawing, and English language learning. In M. Narey (ed.). *Making Meaning: Constructing Multimodal Perspectives of Language, Literacy, and Learning through Arts-based Early Childhood Education*. Pittsburgh: Springer, 133-142.
- van Leeuwen, T. 2005. *Introducing Social Semiotics*. London: Routledge.
- Walsh, M. 2006. The ‘textual shift’: Examining the reading process with print, visual and multimodal texts. *Australian Journal of Language and Literacy* 29(1): 24-37.
- Zhang, Delu (张德禄). 2009. On a synthetic theoretical framework for multimodal discourse analysis. *Foreign Languages in China* (1): 24-30. [2009, 多模态话语分析综合理论框架探索.《中国外语》第1期.]
- Zhang, Delu (张德禄). 2010. Preliminary investigation into the concept of design and the selection of modalities in multimodal foreign language teaching. *Foreign Languages in China* (3): 48-53. [2010, 多模态外语教学的设计与模态调用初探.《中国外语》第3期.]
- Zhang, Delu (张德禄). 2016. Context of culture in multimodal discourse: A social-semiotic perspective. *Journal of Tianjin Foreign Studies University* (6): 1-7. [2016, 多模态话语的文化语境——社会符号学视角.《天津外国语大学学报》第6期.]
- Zhang, Delu (张德禄). 2018. Context of situation in multimodal text. *Journal of PLA University of Foreign Languages* (3): 1-9. [2018, 多模态话语中的情景语境.《解放军外国语学院学报》第3期.]

收稿日期: 2022-12-08; 作者修改稿, 2023-02-15; 本刊修订, 2023-03-03

通讯作者: 张德禄 <dlzhcn@163.com>

200092 上海市杨浦区 同济大学外语学院

**Corresponding author:** Zhang Delu, School of Foreign Languages, Tongji University, Yangpu District, Shanghai 200092, P. R. China