Modern Linguistics 现代语言学, 2023, 11(11), 5534-5543
Published Online November 2023 in Hans. https://www.hanspub.org/journal/ml
https://doi.org/10.12677/ml.2023.1111742

Hans汉斯

外语视频直播课堂的模态配置研究

吴玲娟[1]，瞿　桃[2*]

[1]上海第二工业大学外文学院，上海
[2]同济大学外国语学院，上海

收稿日期：2023年10月12日；录用日期：2023年11月15日；发布日期：2023年11月28日

摘　要

在多模态交际时代，可选择的模态呈现多样性，选择哪些模态以及对所选模态进行怎样的配置才能更加有效地表达意义，成为了多模态交际是否成功的一个关键问题。本文以优质少儿英语视频直播课程为研究语料，首先描述了外语视频直播课堂模态配置的制约因素，接着结合视频多模态标注结果阐释了教师模态配置的总体特征，甄别了各教学环节可优先选择的模态配置方式，总结了不同模态的功能及模态间的协同关系，并以教师讲解知识点环节为例进行了具体分析；最后，本文提出了外语视频直播课堂模态配置应遵循的四条原则，以期给切实改善外语视频直播课堂教学效果提供一定的参考。

关键词

外语视频直播课堂，模态配置，制约因素，特征，机制

A Study on the Configuration of Modes in Live Video-Streamed Foreign Language Classes

Lingjuan Wu[1], Tao Qu[2*]

[1]School of Foreign Languages and Cultural Communication, Shanghai Polytechnic University, Shanghai
[2]School of Foreign Languages, Tongji University, Shanghai

Received: Oct. 12th, 2023; accepted: Nov. 15th, 2023; published: Nov. 28th, 2023

Abstract

The choice and configuration of modes are critical to the success of multimodal communication

*通讯作者。

文章引用: 吴玲娟, 瞿桃. 外语视频直播课堂的模态配置研究[J]. 现代语言学, 2023, 11(11): 5534-5543.
DOI: 10.12677/ml.2023.1111742

which involves a wide variety of semiotic resources. This study, taking high-quality live video-streamed English classes for children as an example, first analyzes the constraining factors of the configuration of modes in live video-streamed foreign language classes, then summarizes the features of modal arrangement by teachers, identifies the ideal ways of arranging semiotic resources in major teaching phases and elaborates on the functions and synergetic relations among various modes. After that, the features and the mechanism of modal configuration in the phase of overt instruction are elaborated. Finally, four principles governing the configuration of modes are put forward, with a view to improving the teaching quality of live video-streamed foreign language classes.

Keywords

Live Video-Streamed Foreign Language Class, Modal Configuration, Constraining Factors, Features, Mechanism

1. 引言

在多模态话语设计中，交际者在语境因素的促动下从可及的符号资源系统中选择出需要的模态后，还要在框定的时空内对所选的资源进行组合与安排，即配置模态[1] [2]。从符号学角度看，配置是一种模态布局，选择出来的符号要进行合理安排，要考虑符号表征意义的顺序以及哪种符号安排方式对交际对方效果最好[1]。在多模态交际时代，可选择的模态呈现多样性，选择哪些模态、对所选择的模态进行怎样的配置才能更加有效地表达意义，成为了多模态交际是否成功的一个关键问题。可供选择和配置的模态资源既可以是单模态资源，也可以是多模态资源，当一种模态的供用特征不能很好地体现意义时，交际者便要对多种模态进行搭配组合，让各种模态的供用特征相互补充、相互协同来共同体现交际者需要表达的意义[3]。

近年来，随着网络直播技术的成熟，能提供浸润式全外语学习环境、上课时间和地点灵活、生动有趣且互动性强的视频直播外语课程越来越受到我国外语学习者的青睐。当前极少数研究从不同角度分析了此新型类在线课程的多模态话语[4] [5] [6]，但并未对课堂各种模态的配置展开深入研究。而分析在线外语课堂上教师模态配置的理据、特征及机制有助于我们更加深入地了解此类新型课堂多模态话语意义的体现过程并有助于切实改善课堂教学效果。基于此，本文以 30 节优质少儿英语视频直播课程为例，通过对教学视频的多模态标注和分析来探索外语视频直播课堂多模态话语意义建构中教师模态配置的制约因素，总结不同教学环节模态配置的总体特征、可优先选择的模态配置方式和不同模态的功能及模态间的协同关系，并以教师讲解知识点环节为例进行了具体分析，最后尝试提出外语视频直播课堂模态配置应遵循的原则。

2. 外语视频直播课堂模态配置的制约因素

2.1. 多模态话语意义建构中模态配置的制约因素

多模态话语意义建构中模态的配置不是随意进行的，而是受到一系列因素的制约，包括语境因素、可及的符号资源系统以及不同符号资源的供用特征。首先是语境因素，因为人类所有的语言活动都是在一定的文化语境和情景语境中发生的。文化语境由作为文化的主要存在形式的意识形态和作为话语模式

选择潜势的体裁(体裁结构潜势)组成。文化语境从宏观上限定和支配人类的交流模式，但是具体采用哪种交际模式，还需要在具体的情景语境中确定[7]。情境语境是指语言活动的直接环境，它是文化语境的一个实例，是一个产生语篇的语境[8]。情景语境涉及三个变项：话语范围指发生了什么事、从事什么活动等；话语基调表示参与者之间的关系，包括交际者的地位、态度、情感等；话语方式指的是语言活动所采用的媒介或渠道如口头模式、书面模式、电子模式等[8] [9]。这三种情境因素共同作用，为具体的语言活动提供语境构型，还可以生成不同类型的语篇结构。

其次，在模态配置过程中，交际者也要考虑有哪些模态可以用来建构意义，因为不同的语境中可供选择的模态会有差别。如面对面即时口头交际中可供选择的模态通常包括有声语言、语调、手势、面部表情、身体姿势、空间距离等；电子邮件交流语境中可供交际者选择的符号资源通常包括文字、字体、字号、图像、版面布局等；传统电话交际语境中交际者可选择的符号资源为口语和相关的副语言资源如音调、音高、音色、口音、停顿等；而如今，人们经常通过视频电话来交流，在这种数字化交际语境中，交际双方也可从口语、图像、面部表情、身体姿势、手势、与声音有关的副语言资源甚至是背景环境等符号资源系统中进行选择。因此，在多模态话语设计中，交际者也需要考虑特定情景中有哪些可及的模态可以用来建构意义。

此外，虽然在多模态话语的建构中可能有很多可以利用的模态系统，但不同模态都有它独特的供用特征和局限性，也就是说，每个符号系统都具有它们自己合适的意义类型和范围，某些模态适合某些交际任务，却不适合其它意义类型和范围；有些意义由特定的模态来体现更为经济、有效，而在某些非专业领域，某些模态可能无法有效地表征意义。因此，了解不同模态的供用特征对于多模态话语意义的建构非常重要，只有熟悉可选模态的供用特征后才能在模态与意义之间进行准确匹配。有效的模态选择要求交际者充分了解相关的符号系统中具体的符号资源的供用特征以及它在话语中可能的语义潜势及局限性，对符号系统及其供用特征的了解越多，对系统中的特征的选择就越准确[10]。

2.2. 少儿英语视频直播课堂模态配置制约因素

根据上述多模态话语意义建构模态配置的制约因素，少儿英语视频直播课堂多模态话语意义建构中模态的选择和配置需要教师除了掌握课程涉及的意识形态、课堂体裁结构成分以及话语范围、话语基调和话语方式等语境因素外，还需要教师考虑教学过程中有哪些模态(符号资源)可供选择以及这些符号具有哪些供用特征。

首先是文化语境，包括意识形态和体裁系统。在意识形态方面，少儿英语视频直播课程性质为主要针对中国中小学生的校外课程，得到我国相关部门的支持，同时也要遵守相关的法规。体裁系统是文化语境的一个重要组成部分，教学中模态的选择一方面受整体体裁类型的支配，另一方面也受到课堂教学体裁不同结构成分的支配[3]。少儿英语视频直播课程的体裁结构成分包括问候、主题介绍、提问、发布指令、知识讲解、拓展反馈、引导反馈、纠错反馈、积极评价反馈以及道别这十个必要的教师教学体裁结构成分以及教师介绍课堂规则、教师引导学生复习、教师课堂总结、教师布置作业这四个可选的教师教学体裁结构成分。

就情境语境而言，其三个变项即话语范围、话语基调和话语方式在教学中分别表现为教学内容、师生特点和教学条件。Bourne & Jewitt [11]指出，课堂上教师对模态的选择受制于教学目的和教学重点。小学阶段在线英语课程的总体教学目标是通过师生口头交流和系列符合少年儿童思维水平和认识规律的一系列趣味游戏来培养学生的英语学习兴趣以及英语进行基本的日常口头交际的能力。就具体教学主题内容而言，本文涉及的 30 节课程涵盖了物主代词、时间、动作、家具、介词、兴趣爱好、星期、月份、季

节、颜色、动物、身体部位、职业、数字等教学主题。就师生特点而言，本文涉及的在线英语课程教师为母语为英语且拥有 TEFL 证书的外籍英语教师。他们进行全英文授课，发音准确且语调自然。学生方面，本文研究的在线英语课程授课对象为生理年龄介于 6~9 岁的低小阶段(小学 1~3 年级)英语初学者，他们活泼好动、好奇心强、喜爱游戏和模仿、自我意识较强且喜欢被称赞，思维能力处于从具象思维向抽象思维过渡的阶段，更能接受具体、直观的教学信息，对教师抱有一种崇敬和依赖的特殊情感，同时也存在注意力不集中、自控能力弱和认知范围有限等问题。在教学条件方面，少儿英语在课程依托 Classin、Air-class、学点云、拓课云等在线教学直播互动平台等进行授课，师生通过教学课件、人像屏幕和互动聊天区进行多维度实时交流和互动。

视频直播课堂环境中可调用的符号资源(如图 1 所示)包括语言模态和非语言模态两个大类。语言模态主要包括口头语言和书面文字；课堂的非语言符号不仅包括课件上的图像、文字样色和版面布局等符号资源，而且由于师生双方都打开摄像头，师生的手势、表情、上身动作等肢体动作以及教师使用的虚拟贴纸、特效音频、视频和直观教具等教学工具都参与意义建构。此外，教师的物理背景、服饰、发饰、妆容等非语言符号也不仅仅是环境因素，他们也成为了体现意义的模态。

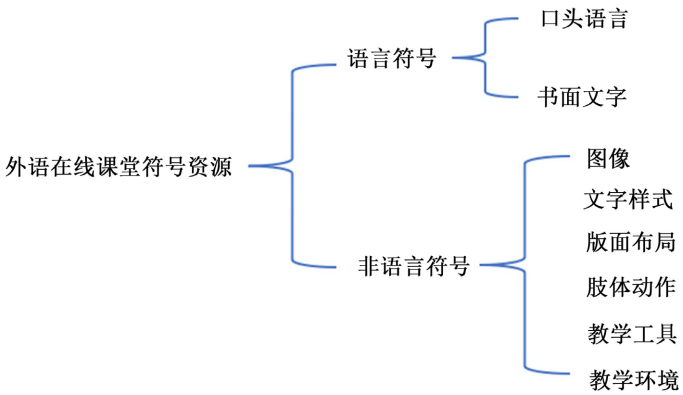

Figure 1. Semiotic resources in live video-streamed foreign language classes
图 1. 外语视频直播课堂符号资源系统

口语、文字、图像、肢体动作、教学工具和教学环境等符号资源各自具有鲜明的供用特征。口语模态具有极强的表意功能，能清楚、直白地传递几乎所有的信息，特别是抽象、笼统的信息；书面语具有长时保留信息、延时反馈的特点，特别是当人们在无法迅速理解转瞬即逝的口语信息的时候，书面语言让人们有足够的时间通过视觉通道来弥补未能理解的口语的含义；文字的呈现和组织方式如字体形状、大小、颜色、空间布局等伴语言的媒体特征对意义整体表达起着重要作用，有时甚至起到关键作用，能引发整体意义的变化[3]；图像具有表达形象逼真、精细、具体和吸引注意力等供用特征，人的一切心理和情感活动，如想象、意图、意识、情感以及模糊的、直觉的或难以言说的信息，都可以通过图像来表达[7]；手势既可以伴随语言使用，又可以脱离语言来表达意义，具有生动性、直接性、指示性、补充性、模拟性的特点；面部表情能有效反映情感和内心状态的变化；头部动作具有语用功能，能示意话轮的转换；身体姿态能有效表达情绪、态度、注意力和投入度等，具有侧重性、控制性、吸引注意力和模拟性的特点；虚拟贴纸生动有趣、简便易用；实物教具能生动、形象、直观地辅助语言表达；音视频能有效吸引注意力；交际物理空间、服装、身体装饰等符号资源也能体现意义，能反映交际双方的个性、权势关系或交际关系的正式程度等[3]。

3. 外语视频直播课堂模态配置研究方法

本研究使用免费开源、功能强大的多模态标注工具 ELAN 来进行教学视频的多模态标注。首先对在线课程的师生有声话语进行转写。接着设定师生话语、手势、教学工具、教学环节模态配置等共 11 种标注类型和层级。其中教学环节模态配置层级的标注侧重教师为实现该环节教学目标所采用的模态或模态组合方式，如“口语 + 文字 + 图像”、“口语 + 图像”、“口语 + 文字 + 手势”、“口语 + 手势”、“纯口语”等模态配置类型。最后将各个层级的标注结果如时长、频率等信息导出到 Microsoft Excel 和 SPSS (20.0)进行统计、分析。

由于目前缺乏科学、统一的测试或教学质量监测和评估标准来直接检验少儿英语视频直播课堂教师多模态话语的教学效果，同时考虑到此类课程以口语能力的提升为主要教学目标，绝大部分的课堂活动都是依赖师生言语互动来实现的，而且学生的学习行为是由语言行为、身体动作与信息技术行为协同的多模态化动态信息交互行为[6]，因此将学生在课堂上的有声话语输出率、互动话语输出率、正确答题率、主动发言频率以及师生语言互动率和学生多模态活动率作为客观反映教师多模态话语教学效果的 6 个维度。利用 SPSS 对各个教学环节教师不同模态配置方式的频率与这 6 个教学效果综合评价维度进行相关性分析，同时结合不同模态的功能及相互关系来甄别各教学环节可优先选择的模态配置方式。

4. 少儿英语视频直播课堂模态配置特征及机制

以下分析少儿英语视频直播课堂各个教学环节模态配置的总体特征及机制(不同模态的功能及模态间的协同关系)并以教师知识点讲解环节的模态配置方式为例来进行具体分析。

4.1. 少儿英语视频直播课堂各教学环节模态配置总体特征及机制

少儿英语视频直播课堂的问候、主题介绍、提问、发布指令、知识点讲解、拓展反馈、引导反馈和纠错反馈、积极评价反馈和道别这些教学环节中口语、文字、图像、手势、实物教具、虚拟贴纸、音频、视频等模态的配置方式如表 1 所示，可以发现此类课堂上教师的模态配置具有以下特征：

Table 1. Modal configurations in various teaching phases of live video-streamed foreign language classes
表 1. 少儿英语视频直播课堂各教学环节的模态配置方式

环节	模态配置方式的数量	优先选择的模态配置方式
问候	3	“口语 + 身体动作”
主题介绍	1	“口语 + 文字 + 图像”
提问	9	“口语 + 文字 + 图像”、纯口语模态及“口语 + 图像”
发布指令	8	纯口语、纯手势、“口语 + 身体动作”
知识讲解	9	“口语 + 文字 + 图像”、“口语 + 文字”、“口语 + 文字 + 图像 + 身体动作”、“口语 + 图像”
拓展反馈	7	“口语 + 文字 + 图像”、“口语 + 图像 + 身体动作”、“口语 + 文字 + 图像 + 身体动作”
引导反馈	7	“口语 + 图像”、“口语 + 文字”、“口语 + 图像 + 身体动作”
纠错反馈	5	“口语 + 文字”、“口语 + 图像”、“口语 + 文字 + 身体动作”
积极评价反馈	12	纯口语、“口语 + 身体动作”、“口语 + 身体动作 + 虚拟贴纸”
道别	3	“口语 + 身体动作”

首先，除了标题呈现环节教师只通过 1 种模态配置方式来建构多模态话语意义外，其它各个教学环节的模态配置方式都出现多样化的特点，涉及 3~12 种不同的模态配置方式，特别是在积极评价反馈环节教师采用高达 12 种不同的模态或模态组合来在建构意义。总体来说，“口语 + 文字 + 图像”模态组合和“口语 + 图像”模态组合是整体课堂上教师选择最为频繁的模态配置方式。

其次，不同体裁结构成分的模态配置方式绝大多数是多模态的，多种模态协同配合共同参与意义的建构，但总体而言，值得优先选择的模态配置方式通常涉及 2~4 种模态的搭配组合。具体来说，从表 1 可以看出，“口语 + 身体动作”模态组合为问候和道别环节意义建构的首选模态配置方式；“口语 + 图像 + 文字”模态组合为建构课程标题意义时可以优先考虑的模态配置方式；“口语 + 文字 + 图像”模态组合、纯口语模态及“口语 + 图像”模态组合是建构提问意义时值得优先采用的模态配置方式；在建构指令意义时，纯口语模态、纯身体动作和“口语 + 身体动作”模态组合可以优先采用；“口语 + 文字 + 图像”、“口语 + 文字”、“口语 + 文字 + 图像 + 身体动作”和“口语 + 图像”模态组合是教师讲解环节意义建构时值得优先选择的模态配置方式；“口语 + 文字 + 图像”、“口语 + 图像 + 身体动作”和“口语 + 文字 + 图像 + 身体动作”模态组合是拓展反馈意义建构时可以优先使用的模态配置方式；“口语 + 文字”、“口语 + 图像”和“口语 + 图像 + 身体动作”模态组合是引导反馈意义建构时优先选择的模态配置方式；在建构纠错反馈意义时，“口语 + 图像”、“口语 + 文字”和“口语 + 文字 + 身体动作”模态组合是值得优先考虑的模态配置方式；纯口语、“口语 + 身体动作”组合和“口语 + 身体动作 + 虚拟贴纸”组合是建构积极评价反馈意义时可以优先选择的模态配置方式。

此外，口语、文字、图像、身体动作、教学工具在多模态话语意义建构中发挥着不同的意义潜势，相互配合来共同表达意义。口语是绝大多数教学环节主要且必要的模态，直截了当地通过听觉通道传递信息；文字、图像、手势、直观教具、虚拟贴纸等模态通过视觉通道对口语主模态起强化、互补等作用，通常处于辅助地位。文字与口语所表征的信息部分重合，强化了口语中的核心信息并在实义层提供口语信息的正确拼写形式，口语则在实义层体现了文字的正确发音，在语义层是对文字的确认和强化；大部分课件图像对口语或文字信息进行强化或补充，如限定语言中人称代词的语义范围、为语言提供具体例证或时间、地点、方式等环境因素，同时图像还能通过卡通表征形式、明艳的色彩以及背景语境对比来吸引学生注意力和激发他们的学习兴趣；文字的字号、颜色、呈现方式也能发挥吸引学生兴趣的人际功能和引导学生注意力分配的语篇功能；身体动作尤其是手势能对口语或文字信息或与核心语言相关的主体图像进行强化、互补或拓展，还可以脱离语言而独立传递指令或积极反馈信息，不仅能增加课堂乐趣，还能引导学生的注意力，提高学生对知识点的理解力和记忆效果；各类教具不仅能生动、形象地体现知识点的语义信息，还能吸引学生的学习兴趣和注意力。此外，值得一提的是，在主题介绍环节，口语、文字和文字均为主要且必要的模态，在知识讲解环节，口语和文字则都是主要且必要的模态。

4.2. 教师知识讲解环节的模态配置特征及机制举隅

以下以少儿英语视频直播课程教师讲解知识点这一环节为例来具体分析教师模态配置的特色及效果。多模态标注结果显示，在少儿英语课堂中，平均每节课(25.8 分钟)教师讲授单词、词组、句子以及语法结构等知识点 58.7 次。教师在讲解知识点时的模态配置方式非常丰富，共选择了 9 种模态或模态组合，如表 2 所示。可以看出，“口语 + 图像 + 文字”模态组合、“口语 + 文字 + 图像 + 身体动作”模态组合、“口语 + 文字”模态组合和“口语 + 图像”模态组合是中外教师在呈现知识点时选择最为频繁的模态配置方式(共占知识呈现总频次的 88.8%)。其它 5 种模态配置方式则很少被选择。

在“口语 + 图像 + 文字”模态组合这一教师首选的模态配置方式中，口语在实体层示范了字母、单词、词组或句子的正确发音，在语义层是对文字的确认和强化，而文字则在实体层提供了字母、单词、

Table 2. The frequencies and percentages of various modal configurations in the phase of overt instruction
表 2. 讲解意义建构不同模态配置方式的频次及占比

	口语 + 图像 + 文字	口语 + 文字 + 图画 + 身体动作	口语 + 文字	口语 + 图像	口语 + 文字 + 身体动作	口语 + 文字 + 直观教具	音频 + 文字 + 图像	口语 + 图像 + 文字 + 直观教具	口语 + 文字 + 图像 + 虚拟教具
频次	20.60	13.80	10.60	7.20	3.80	1.10	0.80	0.70	0.20
占比	35.0%	23.5%	18.0%	12.2%	6.5%	1.9%	1.4%	1.2%	0.3%

词组或句子的正确书写形式，图像与口头语言或书面语言形成拓展或投射的逻辑语义关系，口语、图像和文字模态协同配合表达知识点的概念意义、人际意义和语篇意义。如在例 1 中(见图 2)，课件上浅蓝色纯色背景图上的四幅概念图像——一坨蓝色颜料、一只蓝色蝴蝶、一只蓝色小鸟和一粒蓝莓是非衬线粗体大号蓝色文字 blue 的具体例证，文字和图像构成抽象－具体的逻辑语义关系；虽然文字 blue 直接覆盖在纯蓝色背景图上且其尺寸够大且与四幅概念图像视觉分割明显，方便学生辨认重点文字信息；教师读出课件上的文字，在实体层示范了该词的正确发音，并在语义层对其进行确认和强化，而文字则在实体层提供了单词的正确书写形式。

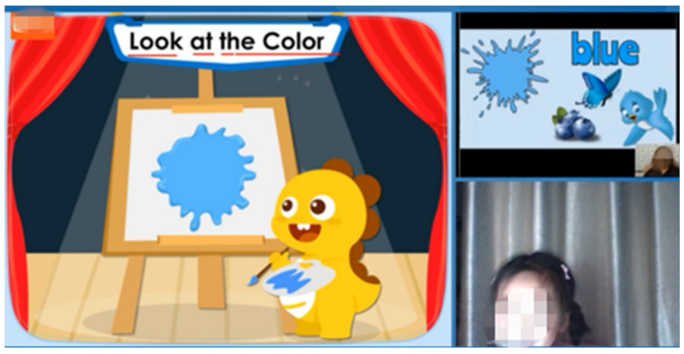

Figure 2. The multimodal ensemble of “Speech + Written Text + Image”
图 2. “口语 + 文字 + 图像”模态组合

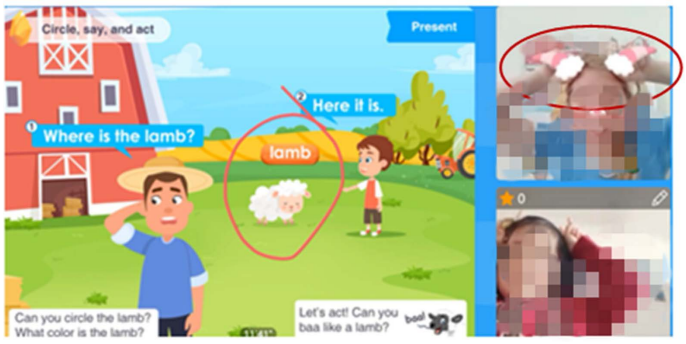

Figure 3. The multimodal ensemble of “Speech + Written Text + Image + Body Language”
图 3. “口语 + 文字 + 图像 + 身体动作”模态组合

在“口语 + 文字 + 图像 + 身体动作”模态组合这种教师选择第二频繁的模态配置方式中，口语在实体层示范了字母、单词、词组或句子的正确发音，在语义层对文字概念意义进行确认和强化，而文字

则在实体层提供了字母、单词、词组或句子的正确书写形式，图像与文字形成拓展或投射关系，身体动作特别是手势则强化了口语、文字和图像中的核心信息，口语、图像、文字和身体动作模态协同配合表达知识点的概念意义、人际意义和语篇意义。如在例 2 (见图 3)中，课件上高语境化农场背景图中有多处文字和多个表征参与者，红色圆形线框引导学习者注意力聚焦在洁白卷毛小绵羊卡通表征参与者和上方的橘色文本框内的白色非衬线中号文字 lamb；教师读出单词 lamb，口语和文字虽然在语义层重复，但口语不仅在实体层示范了单词的正确发音，同时也是对文字的确认和强化，而文字则在实体层提供了口语信息的正确书写形式；课件上的蓝天白云、青草红房以及两个卡通人物表征参与者让学生仿佛置身于农场中，而且还拓展了语言信息，生动、形象地展示了绵羊的形象、生活环境等额外信息；课件上可爱的小绵羊表征参与者以及教师头上的虚拟绵羊头则如为单词 lamb 提供了具体的例证；此外，教师双手放在头顶，做出羊角的手势，生动地描述了绵羊的明显外部特征，不仅增强了语言的表达效果，也能提高学生的课堂参与度，学生一边跟读一边也模仿教师做出羊角手势。口语、文字、图像和手势有机结合，形成了一个紧凑的信息整体，栩栩如生地解释了 lamb 这个词的含义，同时也能增强学生的理解和记忆效果。

“口语 + 文字”是教师呈现知识点时选择第三频繁的模态配置方式。教师读出课件上的文字信息，口头语言和书面语言这两种模态分别在听觉通道和视觉通道传递语义信息，虽然两者在语义上部分重复，但口语是对文字内容的确认和强化，同时也提供了句子的正确发音，而文字则在实体层提供了口头表达的正确书写形式。如在例 3 (见图 4)中，大面积纯深蓝色背景图上方黄色文本框内的黑色文字 They are sitting on the couch.与文本框及课件背景图色彩构成鲜明对比，在视觉上易于学生辨认；教师直接读出这个句子，强化了文字语义信息并在实体层示范了该句子的正确发音；文字本身则在实体层提供了教师口语的正确书写形式，同时也对口语内容进行了确认和强化。

Figure 4. The multimodal ensemble of “Speech + Written Text”
图 4. “口语 + 文字”模态组合

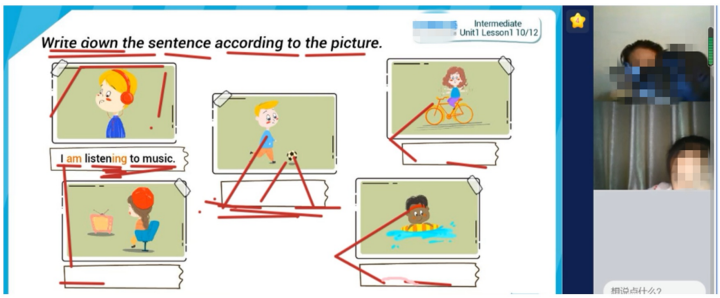

Figure 5. The multimodal ensemble of “Speech + Image”
图 5. “口语 + 图像”模态组合

“口语 + 图像”模态组合也是教师在讲解知识点时选择的一种模态配置方式。如在例 4 (见图 5)中，教师在读出第一个示范句子 I am listening to music.后，由于时间的限制没有让学生回答后面四幅独立黑边浅绿色底文本框中的叙事图像所对应的句子，也没有将正确答案写出，而是直接口头示范了正确句子如 I am watching TV、I am playing football 等并要求学生跟读。教师直接通过听觉通道对学生进行明确指导，即提供了图像对应句子的正确发音，图像则为教师口头语言知识点提供了具体的例证，口语和图像相辅相成，在时间有限的情况下也能帮助学生理解并熟悉 I am doing sth.这一句型。

虽然 SPSS 相关性分析结果(见表 3)显示教师知识讲解环节这 4 种主要的模态配置方式与教学效果各维度没有显著正向或负向相关关系，但从表 2 的教师知识点讲解环节不同模态配置方式的频次及占比可以看出，几乎所有(98.6%)的知识讲解方式都涉及口语模态，绝大部分(87.8%)知识讲解方式涉及文字模态，近 3/4 (73.2%)涉及图像模态，近 1/3 (29.3%)涉及身体动作模态尤其是手势，涉及的音频、实物、虚拟贴纸模态非常少。由此可见，口语和文字是知识讲解环节的必要模态，口语在实体层示范了字母、单词、词组或句子的正确发音，在语义层对文字信息进行确认和强化；文字则在实体层提供了字母、单词、词组或句子的正确书写形式。图像、身体动作和实物教具为可选模态，图像与语言形成拓展或投射的逻辑语义关系，为语言信息提供具体例证或时间、地点、方式等环境因素，同时通过卡通表征形式、色彩、背景语境对比来吸引学生注意力。此外，适量的手势等身体动作不仅能对核心的语言信息进行强化，促进学生对关键知识点的理解和记忆，能拓展表达无法通过语言表达的信息，还能引导学生的注意力；实物教具能直观、形象地体现知识点的语义信息，吸引学生的注意力；音频则能独立或与教师口语联合体现听觉意义。口语和文字这两种主要模态可以与图像、手势这两种辅助模态进行搭配组合，形成的“口语 + 文字 + 图像”、“口语 + 文字 + 图像 + 身体动作”、“口语 + 文字”模态组合能比较有效地传递教学信息。

Table 3. The result of correlation analysis between the frequencies of major multimodal ensembles and various pedagogical effectiveness indicators
表 3. 讲解环节主要模态配置方式的频率与教学效果各维度相关性分析结果

		师生口语互动率	学生多模态活动率	学生话语率	学生互动话语率	学生回答正确率	学生主动发言频次
口语 + 图像 + 文字	r	0.619	0.041	−0.036	0.544	0.324	−0.242
	p	0.062	0.831	0.850	0.052	0.081	0.198
口语 + 文字 + 图像 + 身体动作	r	0.610	−0.002	−0.104	0.590	0.165	−0.271
	p	0.051	0.990	0.586	0.057	0.383	0.148
口语 + 文字	r	0.200	−0.055	−0.001	0.369	0.009	−0.142
	p	0.289	0.771	0.996	0.065	0.963	0.455
口语 + 图像	r	0.483	0.001	−0.041	0.546	0.176	−0.291
	p	0.053	0.997	0.829	0.062	0.353	0.119

5. 外语视频直播课堂模态配置原则

本研究对外语课堂教学多模态话语意义建构具有一定的启示。在外语视频直播课堂的模态配置过程中，除了需要考虑文化语境中的意识形态、情境语境中的教学条件这几个比较固定的因素以及修辞目的、教学目标和内容、师生特点以及体裁结构成分这几个交互影响因素外，交际者也需要遵循适配原则、主次原则、经济原则和动态调整原则，以增强模态配置的有效性。

首先，适配原则是指教师要根据课程的宏观意识形态、体裁结构、教学目标、教学内容、师生特点、教学条件、修辞目的、不同模态各自的供用特征和词汇语法系统以及相互之间的协同关系，对有声语言、文字、图像、身体动作、教学工具、物理背景等模态进行合理搭配组合，充分刺激学习者的听觉通道和视觉通道，最大限度地表达意义。其次，主次原则是指在线外语课堂的多模态意义建构中，口语是主要的、必要的模态，而文字、图像、手势、直观教具、虚拟贴纸、音频等其他模态，对口语主模态起强化、补充等作用，通常处于辅助地位。教师需要高度重视作为课堂交际主要和重点模态的有声话语的质量，有意识地对其进行精心设计，当然，也不能忽略图像、文字样式、版面布局、手势、虚拟贴纸、特效、音频、视频、以及备玩具、工具、卡片等实物教具等非语言模态资源的表义功能。此外，根据模态配置的经济性原则，教师要做到在意义得到充分表达的情况下，模态资源的配置要尽可能简化，避免因为模态过多分散学生的注意力，产生负面的教学效应。最后，动态原则要求教师根据不同的教学内容和学习者的认知、思维和外语水平并灵活结合学生的课堂即时反应来增加或减少模态来更为有效地表达意义。

6. 小结

本文以优质少儿英语视频直播课程为例分析了外语视频直播课堂模态配置的文化语境、情境语境、可及模态和不同符号资源的供用特征等制约因素；结合视频多模态标注结果分析了教师在不同教学环节模态配置的总体特征、甄别了各教学环节可优先选择的模态配置方式、总结了不同模态的功能及模态间的协同关系，并以教师讲解知识点环节为例进行了具体分析；最后，本文提出了外语视频直播课堂模态配置应遵循的四条原则，以期对切实改善外语视频直播课堂教学效果提供一定的参考。

基金项目

上海市教育科学研究项目“少儿在线外教英语课堂多模态话语协同及有效性研究”(项目编号：C2021003)。

参考文献

[1] Bezemer, J. and Kress, G. (2016) Multimodality, Learning and Communication: A Social Semiotic Frame. Routledge, London. https://doi.org/10.4324/9781315687537

[2] 瞿桃, 王振华. 冲突性磋商话语的多模态设计研究[J]. 现代外语, 2022, 45(6): 780-793.

[3] 张德禄. 多模态话语分析理论与外语教学[M]. 北京: 高等教育出版社, 2015.

[4] 赵淑芬. 儿童在线英语教育的多模态化研究[J]. 兰州教育学院学报, 2017, 33(12): 69-70.

[5] Wu, L.J. (2020) Exploring the Multimodal Features of Courseware for Children’s Live Online English Lessons: A Multimodal Discourse Analysis. *Proceedings of* 2020 *International Conference on Education and E-Learning*, Yamanashi, 6-8 November 2020, 56-61.

[6] 吴玲娟. 外语视频直播课堂的多模态教学行为研究及启示[J]. 现代教育技术, 2022, 32(10): 42-49.

[7] 张德禄. 多模态话语建构中的模态融合模式研究[J]. 现代外语, 2023(4): 439-451.

[8] Halliday, M.A.K. (1978) Language as Social Semiotic: The Social Interpretation of Language and Meaning. Edward Arnold, London.

[9] Halliday, M.A.K. and Matthiessen, C. (2008) An Introduction to Functional Grammar. Edward Arnold, London.

[10] 张德禄, 胡瑞云. 多模态话语建构中的系统、选择与供用特征[J]. 当代修辞学, 2019(5): 68-79.

[11] Bourne, J. and Jewitt, C. (2003) Orchestrating Debate: A Multimodal Analysis of Classroom Interaction. *Reading*, **37**, 64-72. https://doi.org/10.1111/1467-9345.3702004