**Modern Linguistics** 现代语言学, 2025, 13(8), 608-612
Published Online August 2025 in Hans. https://www.hanspub.org/journal/ml
https://doi.org/10.12677/ml.2025.138877

Hans汉斯

# 基于ELAN的自闭症儿童教学视频多模态互动机制研究

## ——以特教学校一则教学视频为例

**高子涵**

曲阜师范大学翻译学院，山东 日照

收稿日期：2025年7月3日；录用日期：2025年8月5日；发布日期：2025年8月20日

## 摘　要

**随着多媒体技术的发展和普及，多模态互动分析成为语言学研究热议的中心。教学视频是典型的动态多模态语篇，却少有文献针对其多模态资源进行系统的分析，特别是针对自闭症儿童的教学视频。本研究以ELAN多模态分析软件作为模态标注工具，结合多模态互动分析理论和多模态互动分析综合理论框架，对特教学校一则教学视频中的各交流模态在课堂教学中所起的作用以及这些模态之间的互动机制进行分析。研究发现：在自闭症儿童教学视频中，语言模态、视觉模态(含肢体动作、教具展示等)、听觉模态(含语音语调、辅助音效等)呈现“主导－辅助－协同”的层级互动关系，其中教师的肢体动作与教具视觉符号的协同使用对维持自闭症儿童注意力效果显著；模态转换频率与学生回应度呈正相关，适度的模态切换可提升教学互动效果。**

## 关键词

**自闭症儿童教学视频，ELAN，多模态互动分析**

# A Study on the Multimodal Interaction Mechanism in Teaching Videos for Children with Autism Based on ELAN

## —A Case of a Teaching Video from a Special Education School

**Zihan Gao**

School of Translation Studies, Qufu Normal University, Rizhao Shandong

Received: Jul. 3rd, 2025; accepted: Aug. 5th, 2025; published: Aug. 20th, 2025

文章引用: 高子涵. 基于 ELAN 的自闭症儿童教学视频多模态互动机制研究[J]. 现代语言学, 2025, 13(8): 608-612.
DOI: 10.12677/ml.2025.138877

## Abstract

**With the development and popularization of multimedia technology, multimodal interaction analysis has become a focal point in linguistic research. Teaching videos are typical dynamic multimodal discourses, yet there is scant literature systematically analyzing their multimodal resources, especially those targeting teaching videos for children with autism. This study employs ELAN multimodal analysis software as the modality annotation tool, integrating theories of multimodal interaction analysis and a comprehensive theoretical framework, to examine the roles of various communicative modalities within a teaching video from a special education school and the interactive mechanisms among these modalities. The findings reveal that in teaching videos for children with autism, linguistic modalities, visual modalities (including gestures, display of teaching aids, etc.), and auditory modalities (including voice tone, auxiliary sound effects, etc.) exhibit a hierarchical interactive relationship of "dominant-auxiliary-collaborative". Among these, the synergistic use of teachers' gestures and visual symbols of teaching aids shows a significant effect on maintaining the attention of children with autism. Moreover, there is a positive correlation between the frequency of modality switching and student responsiveness, suggesting that moderate modality switching can enhance the efficiency of teaching interactions.**

## Keywords

**Teaching Videos for Children with Autism, ELAN, Multimodal Interaction Analysis**

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

短视频以内容紧凑、传播便捷等优势，迅速融入人们的生活。对于特教学校而言，短视频的这些特性恰好契合特殊教育的教学需求。自闭症儿童由于身心特点，注意力集中时间较短且不同个体间理解能力存在差异，传统课堂教学方式有时难以满足他们的学习需要，而教学短视频能以直观形象的方式呈现教学内容，满足多元化学习的需求。此外，自闭症儿童教学视频作为一种典型且富有特色的多模态语篇，在短时间内，老师需要运用多种模态吸引学生的注意并有效传授知识，从而达到预期的教学效果。随着“多模态”概念的提出，教学视频的多模态互动分析已形成一定的体系。

多模态互动分析(Multimodal Interaction Analysis)已经成为语言学、教育学等多个领域的研究焦点。多模态语篇不仅包括文字、声音，还包括图像、动作等多种形式的信息载体，这些不同的信息载体相互作用，共同构建了一个更为丰富和复杂的交流环境(Jewitt, 2009) [1]。尤其是在教育领域，如何有效利用这些多模态资源来促进学生的学习，尤其是对于有特殊教育需要的学生群体，如自闭症儿童，成为一个值得深入探讨的问题。对此，本文拟对一则自闭症教学视频进行多模态互动分析，旨在探讨自闭症儿童教学视频中语言模态和非语言模态的特征，以期为特教老师如何协同多种模态进行互动提供可行建议。

## 2. 研究背景

### 2.1. 多模态互动理论框架的构建与发展

多模态互动分析聚焦于人类在交流中通过语言、手势、表情、图像、声音等多种符号资源传递意义、

协调行动的动态过程，是近年来语言学、教育学、社会学等领域的研究热点。

#### 2.1.1. 国外多模态互动分析研究进展

国外对多模态互动分析的研究起步较早，理论基础扎实，形成了较为系统的研究体系。Kress 与 van Leeuwen (1996/2006) [2]在 Halliday 系统功能语言学基础上，提出“多模态社会符号学”理论，将语言之外的视觉、听觉等符号视为独立的意义系统，强调不同模态通过“模态间互补性”共同构建意义。该理论为多模态互动的“意义构建”提供了核心分析框架，被广泛应用于图像、手势等非语言模态的功能研究。Goodwin (2000)提出“情境化符号”概念，强调互动中模态的选择需依赖具体情境(如场景、参与者关系)，并通过对课堂、医疗等场景的实证研究，揭示了手势、眼神等模态在协调互动节奏、传递隐含意图中的作用[3]。Scollon (2003)结合会话分析理论，提出“多模态互动分析”(Multimodal Interaction Analysis)方法论，关注模态在时间序列中的分布规律，以及参与者如何通过模态转换实现互动目标(如话题转换、行动协作) [4]。

#### 2.1.2. 国内多模态互动分析研究进展

多模态互动分析需要与其他学科学者共同研究，提倡多学科交叉研究，才能使多模态互动分析更全面更有价值。国内有关多模态互动分析的研究还处于起步阶段，在文本语言和语法结构之间进行浅层次探索，初期以理论引介为主，近年来实证研究逐渐增多，聚焦教育、传媒等领域。胡壮麟(2007)较早系统引介多模态理论，提出“多模态话语分析”框架，强调结合汉语语境与文化特点分析模态互动，例如汉语声调(听觉模态)与表情(视觉模态)在情感表达中的协同作用[5]。张德禄(2018)构建了“多模态话语分析综合理论框架”，将模态互动关系划分为“协同、补充、替代、强化”等类型，为国内实证研究提供了可操作的分析工具[6]。

### 2.2. 特殊教育视频多模态互动分析的新发展

国外对特殊教育教学视频多模态互动分析的研究起步较早，在理论层面，依托社会符号学、互动社会语言学等理论为研究奠基。比较著名的是从互动社会语言学视角出发，关注互动中情境对模态使用的影响，指出在特殊教育场景下，教师和学生的互动模态选择与教学情境紧密相关，例如在一对一教学和小组教学中，教师的手势、语言等模态的运用会有所不同。

国内特殊教育教学视频多模态互动分析研究初期主要是对国外理论的引介。胡壮麟(2007)率先将多模态话语分析理论引入国内[5]，张德禄(2018)进一步构建多模态话语分析综合理论框架，将模态互动关系细化为协同、补充、替代、强化等类型，使其更契合国内特殊教育研究需求，为教学视频中多模态互动机制分析提供了可操作的理论工具[6]。在自闭症儿童教学方面，相关研究显示，情绪主题绘本教学能有效提升儿童基本情绪认知、共情、心理理论、情绪归因及情绪管理等能力，能激发自闭症儿童的学习兴趣，提升他们的语言理解和表达能力(Aviram, 2009) [7]。

### 2.3. 研究问题

本研究将多模态互动分析理论应用于自闭症儿童教学视频，通过构建“模态－互动－效果”的分析框架，为动态语篇的模态互动分析提供新的视角，旨在研究以下问题：

1) 自闭症儿童教学视频中包含哪些主要交流模态？各模态在教学互动中分别承担何种功能？
2) 不同模态之间存在怎样的互动机制(如协同、补充、替代等)？
3) 多模态互动模式与自闭症儿童的学习回应之间是否存在关联？

## 3. 研究方法

### 3.1. 语料

本文语料选取某特教学校录制的一则自闭症儿童“颜色认知”教学视频作为研究样本。该视频时长

5 分 25 秒，参与者为 1 名特教教师(女，5 年教龄)和 2 名自闭症儿童(男，6~7 岁，经评估为中度自闭症，具备基本语言理解能力)。视频内容为教师引导学生识别红、黄、蓝三种颜色，包含教具展示(彩色积木)、肢体示范(指向、拿起)、语言指令等多模态互动场景。

### 3.2. 语料处理

基于上述语料，本研究构建“三维分析框架”：模态识别层：识别视频中的核心模态(语言、视觉、听觉)及子模态(如视觉模态含肢体动作、教具、字幕等)；功能分析层：分析各模态的教学功能(如语言模态承担指令传递功能，视觉模态承担注意力吸引功能)；互动机制层：通过模态出现的时间序列、频率及关联强度，归纳模态间的协同、补充、替代等互动关系。

笔者对视频进行逐帧编码，标注各模态的出现时间、类型及功能。编码体系分为三部分：语言模态(L)：教师语言(L1：指令语；L2：鼓励语)、学生语言(L3：回应语)；视觉模态(V)：教师肢体(V1：指向；V2：演示)、教具(V3：积木展示；V4：颜色卡片)、学生动作(V5：模仿；V6：操作)；听觉模态(A)：教师语音语调(A1：平稳；A2：上扬)、辅助音效(A3：颜色对应音效，如“红色”伴随“叮”声)。随后，将视频转为 MP4 格式导入 ELAN，生成时间轴(精确至 0.1 秒)并对视频中的模态进行标注，标注一致性达 92%，分歧部分通过讨论统一。最后导出标注数据，形成“时间－模态－功能”对应表(见表 1)。

**Table 1.** Summary of video language and non-language modal annotation data
**表 1.** 视频语言模态及非语言模态标注数据汇总表

| 主模态 | 子模态 | 出现次数 | 占比 |
|---|---|---|---|
| 语言 | L1 指令语 | 22 | 24.2% |
| | L2 鼓励语 | 8 | 8.8% |
| | L3 学生回应 | 5 | 5.5% |
| 视觉 | V1 教师指向 | 15 | 16.5% |
| | V2 教师演示 | 9 | 9.9% |
| | V3 积木展示 | 10 | 11.0% |
| | V4 颜色卡片 | 8 | 8.8% |
| | V5 学生模仿 | 4 | 4.4% |
| | V6 学生操作 | 6 | 6.6% |
| 听觉 | A1 平稳语调 | 7 | 7.7% |
| | A2 上扬语调 | 5 | 5.5% |
| | A3 辅助音效 | 2 | 2.2% |

## 4. 语料模态类型与功能分布分析

笔者首先采用定性方法对教师语言模态及非语言模态进行编码分类，重点关注各模态的协同配合方式。同时运用定量手段统计关键指标，包括不同模态的出现频率、持续时间以及模态切换间隔，并计算自闭症儿童对特定模态刺激的响应延迟时间。经标注，视频中共出现 3 类主模态、9 个子模态(见表 1)。其中视觉模态出现频率最高(42 次)，占比 46.2%；语言模态次之(35 次，38.5%)；听觉模态(14 次，15.3%)最低。数据显示，语言模态主要以以指令语(L1)为主，如“拿起红色积木”，承担知识传递核心功能；鼓励语(L2)如“真棒”，主要用于强化学生积极行为。学生回应语(L3)较少，且多为单字(如“红”)，反映

自闭症儿童语言表达的局限性。视觉模态中教师指向(V1)和积木展示(V3)是核心子模态。教师通过指向特定颜色积木(V1 + V3)，将抽象颜色概念与具象物体关联，符合自闭症儿童“具象思维优先”的认知特点；学生操作(V6)如拿起积木，是学习效果的直接体现。听觉模态以上扬语调(A2)为主，多出现在提问句末尾(如“这是黄色吗？↑”)，通过语调变化吸引学生注意；辅助音效(A3)虽出现少，但学生回应率达 100% (如播放“蓝色”音效时，学生均能正确指向蓝色积木)。

再者，采用对比分析法，将特教课堂的多模态互动模式与普通教学视频进行跨案例对比，结果显示当单一模态无法完整传递意义时，其他模态进行补充。例如，教师说“把积木放这里(L1)”(语言模态)，因学生未理解，随即用手比划放置位置(V2) (视觉模态补充)，学生最终完成操作。此类互动占比 23%，多发生于学生初始回应失败后，体现模态的“容错性”。当语言模态效果不佳时，视觉模态替代传递意义。如教师多次说“放下积木(L1)”，学生无回应，教师随即做出放下动作(V2)，学生模仿完成指令。此类互动占比 9%，反映自闭症儿童对动作符号的敏感度高于语言符号。

综上，统计显示，模态转换频率与学生回应率呈正相关($r = 0.68$, $p < 0.05$)。当每分钟模态转换次数为 4~6 次时，学生回应率达 66.7%；而转换次数 < 3 次时，回应率降至 33.3%。此外，“语言－视觉－听觉”三种模态同时出现时(共 3 次)，学生回应率为 100%，表明多模态融合可最大化互动效果。同时，自闭症儿童教学视频的多模态互动呈现“视觉主导、语言辅助、听觉强化”的特征，这与普通儿童教学视频“语言主导”的模式存在显著差异[2]，这是因为自闭症儿童对视觉信息的加工能力优于语言信息，教师需通过视觉模态构建“意义锚点”，再辅以语言和听觉模态深化理解。

## 5. 结语

本文结合多模态互动分析理论和多模态互动分析综合理论框架，对特教学校一则教学视频中的各交流模态在课堂教学中所起的作用以及这些模态之间的互动机制进行分析，发现自闭症儿童教学视频包含语言、视觉、听觉三类主模态，其中视觉模态(尤其是教师肢体与教具)在互动中起主导作用，模态转换频率与学生学习回应呈正相关，适度的多模态融合能适配自闭症儿童的认知特点，提升教学效果。根据数据分析，教师设计教学视频时，应增加视觉模态的比重，优先选用色彩鲜明、形态简单的教具，并确保教师动作与教学内容同步。同时，教师需要根据学生回应调整模态策略，对语言理解较弱的学生，可多采用“动作演示 + 音效”的互动模式。但笔者仅选取一则教学视频作为样本，结论的普适性有待验证。此外，未考虑自闭症儿童个体差异，希望未来可扩大研究对象数量并尽可能进行个体变量分析。

## 参考文献

[1] Jewitt, C. (2009) The Routledge Handbook of Multimodal Analysis. Routledge, 28-39.

[2] Kress, G. and van Leeuwen, T. (2006) Reading Images: The Grammar of Visual Design. 2nd Edition, Routledge.

[3] Gu, Y. (2006) Multimodal Text Analysis: A Corpus Linguistic Approach to Situated Discourse. *Text & Talk*, **26**, 127-167. https://doi.org/10.1515/text.2006.007

[4] Scollon, R. (1998) Mediated Discourse as Social Interaction. Longman.

[5] 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.

[6] 张德禄, 王正. 多模态互动分析框架探索[J]. 中国外语, 2016, 13(2): 54-61, 66.

[7] Aram, D. and Aviram, S. (2009) Mothers' Storybook Reading and Kindergartners' Socioemotional and Literacy Development. *Reading Psychology*, **30**, 175-194. https://doi.org/10.1080/02702710802275348