Jurnal Riset Rumpun Ilmu Bahasa
Volume 4, Nomor 3, Desember 2025
e-ISSN: 2829-0143; p-ISSN: 2829-0151, Hal. 283-291
DOI: https://doi.org/10.55606/jurribah.v4i3.7028
Tersedia: https://prin.or.id/index.php/JURRIBAH

Multimodality Analysis in YouTube Educational Content

(Case Study on the Realization of Meaning in *SUARA BERKELAS* Video "Kalau Kamu Mau Punya CRITICAL THINKING, Jangan Skip Obrolan Ini!")

Nisa Silitonga[1*], Selly Sitompul[2], Sonya Pasaribu[3], Sonya Silitonga[4], Muhammad Natsir[5]

[1-5] Sastra Inggris, Fakultas Bahasa dan Seni, Universitas Negeri Medan, Indonesia
*Penulis Korespondensi: nisasilitonga@gmail.com[1]

Abstract. In the digital era, YouTube has become a powerful platform for sharing knowledge, shaping perspectives, and encouraging critical engagement. This study explores how multimodal elements—verbal, visual, and audio—construct meaning in the educational video "Kalau Kamu Mau Punya Critical Thinking, Jangan Skip Obrolan Ini!" from the Suara Berkelas channel. Using a qualitative case study approach, the research interprets how these modes work together to support critical thinking. Data were collected through video observation, transcription, and documentation of multimodal features, then analyzed thematically using Kress and van Leeuwen's multimodal framework and Facione's critical thinking model. The findings show that the verbal mode structures logical arguments and rhetorical appeals, the visual mode emphasizes meaning through gestures, expressions, and text overlays, while the audio mode builds emotional engagement through tone, pauses, and background music. Together, these modes create a coherent and persuasive discourse that not only informs but also invites viewers to reflect critically. The study contributes theoretically to the understanding of multimodality in digital education and practically provides insights for educators and content creators to design more interactive, reflective, and engaging educational content.

Keywords: Critical Thinking; Digital Media Literacy; Educational Discourse; Multimodality; Qualitative Case Study.

1. INTRODUCTION

In today's increasingly complex information era, critical thinking has become one of the essential skills required in both academic and professional contexts. YouTube, as a leading video-sharing platform, has become a primary source of information for many people, offering a wide variety of educational content capable of stimulating reflection and analysis from multiple perspectives. One such video that draws attention is from the channel SUARA BERKELAS, entitled *"Kalau Kamu Mau Punya CRITICAL THINKING, Jangan Skip Obrolan Ini!"*. This video not only introduces the concept of critical thinking but also demonstrates how multimodality in information delivery can enhance understanding and audience engagement.

Naskah Masuk: 15 September 2025; Revisi: 28 September 2025; Diterima: 14 Oktober 2025; Tersedia: 16 Oktober 2025

Recent studies indicate that the use of multimodality in educational content can optimize learning experiences, as it facilitates cognitive processes by effectively combining text, image, and sound (Mayer, 2020). However, despite the growing body of research on multimodality in education, a knowledge gap remains regarding how these multimodal elements contribute to the comprehension of critical thinking, particularly among younger audiences. Previous research has mostly focused on theoretical aspects without examining in depth how the interaction between text, image, and sound in videos realizes meaning and influences viewers' perspectives on the discussed topics.

Through an in-depth analysis of this video, the present article seeks to explore how multimodal elements interact in constructing an understanding of critical thinking. The theoretical contribution of this study is to provide a new perspective on multimodality in learning contexts, while the practical implication lies in offering insights that educators and content creators may use to design more inclusive and engaging learning materials. Accordingly, this research seeks to fill existing gaps and provide a deeper understanding of the importance of multimodality in today's digital age.

2. LITERATURE REVIEW

The theoretical foundation of this study is based on the theory of multimodality developed by Kress and van Leeuwen (2006), which asserts that meaning in communication is constructed through the interaction of various modes, such as verbal text and visual imagery. Within this theory, the visual mode is not merely a supplement to verbal text but possesses an independent role in conveying meaning while remaining closely interconnected with verbal elements. The theory divides meaning into three metafunctions: representational (the conveying of ideas), interactive (the communicative relationship), and compositional (the structure and arrangement of communicative elements). In contrast to Barthes, who argued that visual meaning depends heavily on verbal text, Kress and van Leeuwen emphasized that visuals construct their own independent messages aligned with, but not subordinate to, verbal text.

Furthermore, Facione's (2015) theory of critical thinking also forms an important foundation for this study. This theory highlights critical thinking as a process of analyzing, evaluating, and reflecting upon received information, which can be stimulated and developed through multimodal interaction in educational videos. The integration of critical thinking theory and multimodality provides a comprehensive framework for exploring how meaning is realized through the synergy of text, image, and sound in YouTube educational content.

In short, the central distinction between these theoretical foundations lies in Kress and van Leeuwen's emphasis on the independence and interaction of visual and verbal modes in meaning-making, while Facione's theory highlights the analytical capacity that can be enhanced through multimodal approaches in digital education. Together, these two perspectives offer a robust scientific foundation for understanding and analyzing the complex and multidimensional nature of educational content.

Previous studies on multimodality in digital learning media, particularly on YouTube, suggest that the integration of text, image, and sound enhances both the effectiveness of material delivery and audience comprehension. For example, Ramadhan (2025), in his study on Indonesian as a Foreign Language (BIPA) instructional videos, demonstrated that combining visual elements such as images, gestures, and on-screen text with verbal discourse, including narration and intonation, creates strong multimodal interaction. This combination not only increases the attractiveness of the material but also strengthens both linguistic and cultural understanding among foreign learners. Similarly, Gogola (2023), in her study on English vocabulary learning via TikTok videos, revealed that simultaneous use of various modes of communication enables learners to grasp material more easily while eliciting positive audience responses. These studies employed qualitative approaches with content and multimodal discourse analysis techniques, making them methodologically relevant to the present research. However, most prior studies primarily focused on visual and verbal modes, without deeply integrating sound as a key element in fostering critical thinking. Therefore, this study aims to address that gap by adopting a more holistic multimodal approach, particularly in SUARA BERKELAS educational content, to examine how meaning is realized in text, image, and sound to contribute to the development of critical thinking.

3. METHOD

This study employs a qualitative approach. According to Denzin and Lincoln (2011), qualitative research is a naturalistic and interpretive approach that examines phenomena in their natural contexts while seeking to interpret the meanings individuals ascribe to these phenomena. This approach is chosen because the study aims to understand the meaning realized through text, image, and sound in the SUARA BERKELAS video rather than quantify numerical data. Thus, the research emphasizes the interpretive process of meaning-making and its connection to the development of audience critical thinking.

Data Sources

a. **Primary Data**: The educational video from the SUARA BERKELAS YouTube channel that fulfills the criteria of integrating text, visual, and auditory elements relevant to a learning context.
b. **Supporting Data**: Expert evaluation and interviews (e.g., linguistics scholars specializing in multimodality, educational media experts, or digital learning practitioners).

Data Collection Techniques

a. **Online observation**: Watching and recording multimodal elements in the video.
b. **Documentation**: Preparing transcripts of narration, screenshots of visuals, and notes on sound and intonation.
c. **Expert judgement/interviews**: Engaging 2–3 experts to validate the researcher's analysis and provide perspectives on the contribution of multimodality to audience critical thinking.

Data Analysis Techniques

a. Multimodal analysis using Kress & van Leeuwen's (2006) framework, including representational, interactive, and compositional metafunctions.
b. Critical thinking analysis based on Facione's (2015) model.
c. Expert validation by comparing the researcher's findings with expert assessments to generate more objective and comprehensive interpretations.
d. Synthesis of results through integration of observations and expert insights to address the research questions.

Data Validity

Data validity is ensured through source triangulation (video, observational notes, expert judgment) and member checking with experts to maintain consistency and accountability of interpretation.

4. RESEARCH FINDING AND DISCUSSION

Research Findings

The multimodal analysis of the Suara Berkelas video entitled "Kalau Kamu Mau Punya Critical Thinking, Jangan Skip Obrolan Ini!" (If You Want to Have Critical Thinking, Don't Skip This Conversation!) demonstrates the realization of meaning through three primary modes: verbal, visual, and audio.

Verbal Mode

The verbal mode is realized through the use of:

a. **Lexical choices**, such as persuasive expressions ("let's think critically," "do not get trapped"), modal verbs ("must," "need"), and rhetorical questions.
b. **Narrative structure**, which begins with the introduction of a problem, followed by argumentation, and ends with reflective invitations.
c. **Function**, which emphasizes the urgency of critical thinking while simultaneously guiding the audience through the process of reflection.

Visual Mode

The visual dimension is conveyed through:

a. **Gestures and expressions**, such as hand movements, eye contact, and serious facial expressions.
b. **On-screen text**, where key terms like *critical thinking* or *do not skip* are highlighted.
c. **Composition**, which employs close-up framing to strengthen intimacy with the audience.
d. **Function**, which reinforces verbal messages and fosters emotional connection with viewers.

Audio Mode

The audio mode is manifested through:

a. **Intonation**, where the speaker employs rising and falling tones to highlight important points.
b. **Strategic pauses**, which are used to provide space for reflection.
c. **Background music**, characterized by minimal, slow-paced tones that create a serious and reflective atmosphere.
d. **Function**, which enhances persuasiveness and underscores key messages.

Multimodal Analysis Table

Time Segment	Transkrip Verbal (ID)	Verbal Transcript (EN)	Visual Mode	Audio Mode	Realized Meaning
00:05 – 00:20	"Kalau kamu mau punya *critical thinking*, jangan skip pembicaraan ini…"	"If you want to have *critical thinking*, don't skip this conversation…"	Close-up of speaker's face, pointing gesture	Firm intonation, no background music	Opens with urgency and captures the audience's attention from the very beginning.
01:10 – 01:45	"Banyak orang salah kaprah, menganggap berpikir kritis itu sekadar pintar berdebat."	"Many people misunderstand, assuming that critical thinking is merely about debating skillfully."	Serious facial expression, on-screen text "misunderstanding"	Falling then rising intonation on the word *"debating"*	Clarifies the misconception of critical thinking.
02:20 – 02:50	"Berpikir kritis bukan soal siapa yang paling pintar bicara, tapi siapa yang bisa melihat masalah dengan jernih."	"Critical thinking is not about who speaks most eloquently, but who can view problems clearly."	Open hand gesture, direct eye contact	Persuasive intonation, pause on the word *"clearly"*	Shifts the meaning from debating toward deeper analytical thinking.
03:00 – 03:40	"Berpikir kritis artinya mampu menganalisis informasi sebelum mempercayainya."	"Critical thinking means the ability to analyze information before believing it."	Simple graphic slide, on-screen text *"analyze information"*	Clear intonation, pauses at each point, soft background music	Provides a systematic definition, guiding the audience's understanding.
04:05 – 04:40	"Kalau kita asal percaya, kita mudah dimanipulasi."	"If we believe blindly, we are easily manipulated."	Serious expression, furrowed brows	Firm intonation, emphasis on the word *"manipulated"*	Highlights risks, stressing the importance of healthy skepticism.
05:15 – 05:50	"Coba pikirkan kembali, apakah kamu sering menerima informasi tanpa mengecek kebenarannya?"	"Think again, do you often accept information without checking its accuracy?"	Close-up framing, reflective hand gesture	Lowered intonation, long pause	Encourages personal reflection and introspection.
07:00 – 07:30	"Critical thinking itu keterampilan, bukan bakat. Artinya semua orang bisa melatihnya."	"Critical thinking is a skill, not a talent. This means everyone can practice it."	Brief smile, sweeping hand gesture	Optimistic intonation, rising background music	Motivates the audience, emphasizing that critical thinking can be trained.
08:20 – 08:50	"Jangan skip pembicaraan ini, karena bisa mengubah cara kamu berpikir."	"Don't skip this conversation, because it can change the way you think."	Close-up shot, intense eye contact	Persuasive intonation, pause before the word *"change"*	Provides a strong closing, leaving a lasting impression.

Discussion

The multimodal analysis of the Suara Berkelas video illustrates how meaning is coherently constructed through the interplay of verbal, visual, and audio modes. The findings from eight video segments can be elaborated as follows:

Verbal Mode.

The speaker frequently employs imperative statements and rhetorical questions, which function persuasively by positioning the audience as active participants. Definitional statements, such as "critical thinking means the ability to analyze information before believing it," provide clear conceptual framing. This supports Facione's (2015) view that critical thinking involves analysis, reflection, and evaluative judgment, with verbal language serving as the foundation for logical argumentation.

Visual Mode.

Visual elements strongly reinforce the verbal message. Close-up shots displaying serious expressions enhance the speaker's authority, while gestures signal urgency and openness. On-screen text such as "misunderstanding" and "analyze information" highlights key concepts for easier retention. As argued by Kress and van Leeuwen (2006), visuals convey representational meaning and contribute symbolic value rather than merely supplementing speech.

Audio Mode.

The audio dimension adds emotional depth. Rising and falling intonation underscores important concepts such as manipulation and clarity. Extended pauses following rhetorical questions provide space for audience reflection, while subtle background music fosters a reflective and motivational atmosphere. These findings support Mayer's (2020) multimedia learning theory, which posits that auditory features guide attention toward key information.

Integration of Modes.

The three modes function synergistically. For example, in the opening (00:05–00:20), verbal invitations are reinforced by pointing gestures (visual) and firm intonation (audio), establishing urgency. In the core explanation (03:00–03:40), verbal definitions are emphasized through visual text and clear intonation, ensuring comprehensibility. In the conclusion (08:20–08:50), verbal exhortation, close-up visual framing, and persuasive intonation combine to deliver a compelling final message. This demonstrates interactional and compositional meaning (Kress & van Leeuwen, 2006), whereby meaning emerges from interrelations across semiotic modes.

Implications for Critical Thinking.

The video not only delivers information but also actively trains the audience in critical thinking through multimodal strategies. Verbal content provides logical definitions and arguments; visuals reinforce salient points; and audio directs focus and emotional engagement. These findings are consistent with Snelson (2018) and Ramadhan (2025), who found YouTube-based multimodal content effective in fostering critical awareness. The analysis affirms that the Suara Berkelas video successfully integrates verbal, visual, and audio modes to construct persuasive, educational, and reflective discourse, reinforcing the realization of "critical thinking" in line with Kress and van Leeuwen's (2006) multimodal framework.

5. CONCLUSION

Based on the multimodal analysis of the Suara Berkelas video "Kalau Kamu Mau Punya Critical Thinking, Jangan Skip Obrolan Ini!", several conclusions can be drawn. The study demonstrates that meaning in the video is not solely conveyed through spoken language but is strengthened by visual and auditory components that work in concert. Verbally, the speaker employs persuasive diction, rhetorical questions, and structured definitions to emphasize the importance of critical thinking. Visually, facial expressions, gestures, and on-screen text reinforce spoken messages and foster emotional engagement with the audience. Auditory elements, including intonation, strategic pauses, and background music, add emotional resonance and enhance persuasiveness. The integration of these three modes creates effective message delivery, constructs strong meaning, and encourages reflective thinking. These findings corroborate Kress and van Leeuwen's (2006) theory of multimodality and align with Facione's (2015) assertion that critical thinking is a skill that can be cultivated.

Several recommendations emerge from this study. For educators, multimodality can be employed as a pedagogical strategy to design digital learning materials that are more engaging and capable of fostering critical thinking skills among students. For content creators, balancing verbal, visual, and audio elements is essential to ensure that educational messages are communicated persuasively and meaningfully. For future researchers, the scope of inquiry could be expanded by analyzing a broader selection of educational YouTube videos and combining multimodal analysis with audience surveys to assess actual impacts on critical thinking development. Finally, for audiences, it is important to approach educational content critically, not only understanding the message but also recognizing how multimodal features shape thought processes.

REFERENCES

Arianto, B. and Handayani, B. (2022). Studi fenomenologi youtube sebagai saluran pembelajaran kewargaan desa sukoharjo kabupaten sleman. Jurnal Dialogika Manajemen Dan Administrasi, 3(2), 92-106. https://doi.org/10.31949/dialogika.v3i2.8717

Baldry, A., & Thibault, P. J. (2006). *Multimodal transcription and text analysis: A multimedia toolkit and coursebook*. Equinox.

Bateman, J. A., Wildfeuer, J., & Hiippala, T. (2017). *Multimodality: Foundations, research and analysis – A problem-oriented introduction*. De Gruyter Mouton.

Bina, N. (2021). Pengaruh platform pembelajaran daring youtube terhadap kemampuan matematis di masa pandemik covid-19. Laplace Jurnal Pendidikan Matematika, 4(1), 32-39. https://doi.org/10.31537/laplace.v4i1.461

Choy, M., Hilliard, A., & Dickason, C. (2025). First-year seminars to enhance engineering education. Proceedings of the 2025 IEEE International Conference on Engineering Education. IEEE. https://doi.org/10.1109/ICEE.2025.11115012

Denzin, N. K., & Lincoln, Y. S. (2011). *The Sage handbook of qualitative research* (4th ed.). SAGE Publications.

Djonov, E., & van Leeuwen, T. (2018). Multimodal Discourse Analysis of Social Media: An Introduction to the Semiotics of Social Media. Routledge.

Facione, P. A. (2015). *Critical thinking: What it is and why it counts*. Insight Assessment.

Gogola, G. (2023). The use of multimodal elements in TikTok videos for English vocabulary learning. *Journal of Language and Digital Media*, 7(2), 56–70.

Jewitt, C. (2014). *The Routledge handbook of multimodal analysis* (2nd ed.). Routledge.

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.

Mayer, R. E. (2020). *Multimedia learning* (3rd ed.). Cambridge University Press.

Putri, A., Astagini, N., & Ramli, M. (2023). Strategi produksi televisi palang merah indonesia tentang edukasi kesehatan pada masa pandemi covid-19 melalui media sosial youtube. Jurnal Studi Jurnalistik, 5(1), 40-49. https://doi.org/10.15408/jsj.v5i1.31707

Ramadhan, A. (2025). Analisis multimodalitas dalam video pembelajaran Bahasa Indonesia bagi penutur asing (BIPA) di YouTube. *Jurnal Pendidikan Bahasa dan Sastra*, 15(1), 45–58.

Rybina, N. V., Koshil, N. Y., & Hyryla, O. S. (2025). Artificial intelligence and translation in English language teaching: Opportunities and challenges. Medical Education, 3(2), 45–56. https://ojs.tdmu.edu.ua/index.php/med_osvita/article/view/15494

Siwi, S. R. (2025). Multimodal Discourse Analysis of YouTube Thumbnails in Good Mythical Morning (Skripsi, Universitas Jenderal Soedirman).http://repository.unsoed.ac.id/id/eprint/31715

Vuong, T. T. L., & Kwak, J. T. (2025). VideoPath-LLaVA: Pathology diagnostic reasoning through video instruction tuning. arXiv preprint arXiv:2505.04192. https://arxiv.org/abs/2505.04192