

Multimodal Discourse Analysis of Cinematography in “Green Book”

Yesika Natalina Sidabutar¹, Immanuel Prasetya Gintings², Berlin Sibaran³

^{1, 2, 3} *Program Studi Linguistik Terapan Bahasa Inggris, Pascasarjana Universitas Negeri Medan*
Email: yesika.sidabutar@uhn.ac.id

Published: 30/04/2025

How to cite (in APA style):

Sidabutar, Y. N., Gintings, I. P., & Sibaran, Berlin. (2025). Multimodal Discourse Analysis of Cinematography in “Green Book.” *Retorika: Jurnal Ilmu Bahasa*, 11(1), 281-288. DOI: 10.55637/jr.11.1.10075.281-288

Abstract-“Green Book,” directed by Peter Farrelly, depicts the evolving relationship between Dr. Don Shirley, an African-American pianist, and Tony “Lip” Vallelonga, an Italian-American bouncer, during a 1960s concert tour in the American South. This study employs Multimodal Discourse Analysis (MDA) and Systemic Functional Linguistics (SFL) to examine how the film conveys complex social messages about race, identity, and friendship. Utilizing theories by Halliday & Matthiessen (2013), Kress & van Leeuwen (2020), and (O’Halloran, 2011), the research focuses on 23 key scenes set in a Cadillac. The study addresses three main questions: What messages are conveyed in “Green Book”? How are these meanings realized? Why are these meanings realized in their specific social context? Findings reveal significant messages about the struggles and resilience of African Americans, realized through linguistic and visual elements. The confined space of the Cadillac symbolizes broader societal issues, emphasizing personal and cultural conflicts. This research contributes to understanding multimodal discourse analysis in film studies and provides insights for filmmakers, educators, and media analysts. “Green Book” demonstrates the power of film as a medium for social commentary and cultural reflection.

Keywords: Cultural Dynamics, Film Analysis, Multimodal Discourse, Systemic Functional Linguistics, Visual Grammar, Green Book, Cinematography.

I INTRODUCTION

This research explores how the film “Green Book” employs multimodal discourse to convey complex social messages about race, identity, and friendship. Utilizing Systemic Functional Linguistics (SFL), Visual Grammar, and Multimodal Discourse Analysis (MDA), this study aims to analyze key scenes to uncover the interplay between linguistic and visual elements and their contributions to the film’s thematic depth. The study focuses on how these semiotic resources create meaning within the historical and social context of the 1960s American South, aiming to provide new insights into the use of

multimodal elements in film analysis.

“Green Book,” directed by Peter Farrelly, is a film inspired by the true story of Dr. Don Shirley, a well-known African American pianist, and Tony “Lip” Vallelonga, an Italian American bouncer hired to drive Shirley through the Deep South during a concert tour in 1962. The film’s historical setting during the Civil Rights Movement and its use of “The Negro Motorist Green Book” highlight the pervasive racial discrimination of the era. This context makes “Green Book” ideal for detailed multimodal analysis, examining the interplay between visual and textual elements to communicate complex

social messages about race, identity, and friendship.

Movies are a powerful medium that can educate and enhance understanding through visual storytelling, effectively conveying information and messages to audiences. Sharjeel & Dadabhoy (2013) explain that movies present clear images that make their messages easily understandable, although some viewers may find these messages challenging to grasp. Films often convey moral values by combining visual elements such as cinematography and editing with narrative structures, engaging audiences and addressing complex social issues and ethical dilemmas in an accessible manner. This integration creates a more immersive and impactful experience for viewers, enhancing their ability to understand and reflect on the conveyed messages.

Brown (2011) asserts that films reflect societal values and cultural norms, often challenging or reinforcing prevailing ideologies. By presenting characters and situations that embody specific virtues or vices, movies prompt audiences to consider their own moral beliefs and the ethical implications of their actions. The capacity of films to evoke emotional responses strengthens the audience's connection to the characters and the story, further enhancing their effectiveness in conveying messages. This emotional involvement can result in a deeper comprehension of the themes and messages presented in the movie, demonstrating the potential of films as practical educational tools.

The effectiveness of "Green Book" in addressing its central themes is deeply rooted in its masterful use of cinematography, editing, and narrative techniques. The film immerses the audience in the historical and social contexts of the 1960s American South, using the backdrop of the Civil Rights Movement and significant racial tensions. The meticulous visual storytelling in "Green Book" is instrumental in achieving this immersion, with every shot carefully composed to convey the nuances of the era, from segregated environments to intimate moments between characters. The use of color palettes, such as muted tones reflecting the subdued atmosphere of the South and vibrant hues symbolizing hope and change, further enhances the narrative's depth.

Strategic editing choices, such as the pacing of scenes and the juxtaposition of contrasting images, heighten the emotional impact on the viewer. This attention to detail in visual composition and editing brings the

historical context to life, allowing the audience to intimately experience the characters' challenges and triumphs. By engaging viewers on both an intellectual and emotional level, "Green Book" effectively conveys its messages about race, identity, and friendship, making the film a powerful tool for reflecting on and discussing these critical social issues.

A comprehensive review of existing literature reveals significant contributions to understanding multimodal discourse analysis, visual grammar, and the interplay between text and images in films. Halliday & Matthiessen (2013) provide an essential foundation for understanding how language functions as a semiotic system, offering insights into how different modes of communication, including visual elements, work together to create meaning. This aligns with the multimodal approach, where visual and textual elements are integrated to convey messages, which is crucial for analyzing films by emphasizing the interplay between various semiotic resources in constructing a cohesive narrative.

Kress & van Leeuwen's (2020) framework for visual grammar provides a robust methodology for analyzing film features. They argue that understanding visual elements such as composition, color, and perspective is crucial for interpreting a film's overall meaning and conveying complex social and cultural messages. Their framework includes representational, interactive, and compositional meanings, which are essential for interpreting semiotic resources in visual media.

O'Halloran (2011) extends language study to include resources such as images, scientific symbolism, gesture, action, sound, and music. This comprehensive approach has been applied to various forms of media, including advertisements and movie posters, showcasing how multimodal discourse analysis can uncover complex interactions between different semiotic modes and provide a deeper understanding of how meaning is constructed in media.

Building on these theoretical perspectives, Chen & Gao (2013) analyzed movie posters to illustrate the role of visual and textual interplay in creating impactful and memorable promotional materials, demonstrating the broad applicability of multimodal discourse analysis. Similarly, Saputra & Rosa (2020) highlighted how advertisements use multimodal elements to convey persuasive messages, showcasing the practical applications of multimodal discourse

analysis in commercial media.

Bo's (2018) analysis of "Argo" demonstrates how films construct multimodal discourse through context, language, and imagery interplay to propagate ideological messages. This study elaborates on the representational, interactive, and compositional meanings constructed through visual elements, such as depicting actions, reactions, and events that align with the film's ideological stance and using camera angles and shot distances to influence the viewer's emotional and psychological engagement.

Zhu & Huang's (2023) study of "The Tea House" underscores the importance of analyzing film images, music, and text to understand how these elements collaborate to create a cohesive narrative and evoke specific emotional responses. It highlights the effectiveness of multimodal discourse analysis in revealing deeper meanings within films.

Synthesizing these theoretical perspectives allows for a deeper appreciation of how films use multimodal elements to communicate complex social, cultural, and ideological messages. This approach provides a framework for understanding how visual and textual elements work together to address themes such as race, identity, and friendship in historical and social contexts.

To illustrate the research approach, consider the scene where Dr. Shirley and Tony Lip travel through the South. Dr. Shirley requests a Steinway piano and a bottle of Cutty Sark each night. This scene highlights the interplay of visual and textual elements, such as the red cashmere blanket that Lip places over Shirley's legs. In stark contrast to the tension that Lip's frantic driving and smoking suggest, this blanket represents warmth and care. This dynamic interplay of visual cues and dialogue provides a rich tapestry of meaning, portraying the evolving relationship between the characters.

II. METHODS

This study employs a descriptive qualitative research methodology, specifically focusing on multimodal discourse analysis (MDA) to investigate the cinematographic elements in the film "Green Book." The qualitative approach is particularly suitable for this research as it allows for an in-depth interpretative analysis of complex visual and linguistic data, essential for understanding the film's multifaceted nature. Descriptive qualitative

research aims to provide a comprehensive and detailed description of the phenomenon under study, differing from other qualitative approaches by offering straightforward descriptions of events, drawing from the general principles of naturalistic inquiry where phenomena are studied in natural settings without manipulation or control (Sandelowski, 2000). This method is appropriate for exploring subjects that require a deep understanding of the context and the meanings that participants assign to their experiences.

Multimodal discourse analysis (MDA) integrates the study of language with other semiotic resources such as images, sounds, and gestures. O'Halloran (2011) explains that MDA examines how different modes of communication combine to create meaning. This methodology is particularly suited for film analysis, where visual, auditory, and textual elements work together to produce a cohesive narrative.

Systemic Functional Linguistics (SFL), developed by Halliday & Matthiessen (2013), provides the foundational framework for analyzing the film's narrative and dialogue. SFL focuses on language's ideational, interpersonal, and textual metafunctions, which are crucial for understanding how language constructs meaning. This approach allows for a detailed examination of the primary messages and moral lessons in "Green Book" by analyzing its dialogue and narrative structure.

Kress & van Leeuwen (2020) state that visual grammar framework is used to analyze the film's visual elements, such as composition, color, and perspective. This framework examines how visual elements contribute to meaning-making, depicting the narrative world (representational meaning), engaging the viewer (interactive meaning), and arranging elements within the frame (compositional meaning).

The primary data source for this study is the film "Green Book," directed by Peter Farrelly. The analysis focuses on specific scenes within the film, particularly those set inside the car, as they are pivotal to the narrative and rich in dialogue and visual cues. All scenes inside the car were identified and cataloged, ensuring no relevant scenes were overlooked. For each identified scene, the corresponding sections of the screenplay were extracted to analyze the written dialogue and stage directions. Each scene was closely examined for its visual elements, including shot composition, camera angles, lighting, and color use. Detailed notes were made

on how these visual aspects contribute to the overall narrative and thematic expression. The script and the film's visual representation were compared side-by-side to highlight how language and visuals interact to create meaning.

Data analysis in this study integrates the frameworks of Systemic Functional Linguistics (SFL), Kress and Van Leeuwen's visual grammar, and Multimodal Discourse Analysis (MDA) to systematically break down the cinematographic elements of "Green Book." The ideational, interpersonal, and textual metafunctions of language in the film's dialogue were analyzed to uncover the primary messages and moral lessons, coding the dialogues to identify processes, social roles, and discourse structures. The film's visual elements, such as composition, color, and perspective, were examined to understand how they contribute to the narrative and thematic depth, identifying different shot types (long shots, medium shots, close-ups) and their roles in storytelling. The findings from the linguistic and visual analyses were synthesized using MDA, integrating findings to understand how multimodal elements interact to convey the film's messages, providing a holistic view of how different semiotic modes work together. The socio-cultural and historical context of the film was analyzed to provide a deeper understanding of the narrative, including researching the historical background of the 1960s American South and its influence on the film.

To illustrate the research approach, the process began with scene selection and segmentation, where the film was reviewed to identify and segment relevant car scenes, cataloging each for further analysis. The corresponding sections of the screenplay were then extracted for each identified scene, providing a textual basis for comparison with the visual representation. Each scene was examined closely for its visual elements, analyzing shot composition, camera angles, lighting, and color use, with detailed notes on how these aspects contribute to the narrative and thematic expression. The dialogue within the scenes was analyzed using the SFL framework to understand how language constructs meaning. The visual elements were analyzed using Kress and Van Leeuwen's visual grammar framework to identify and interpret representational, interactive, and compositional meanings. The results from the linguistic and visual analyses were synthesized using MDA to understand how multimodal elements interact to convey the film's messages.

The socio-cultural and historical context of the film was analyzed to provide a deeper understanding of the narrative, including researching the historical background of the 1960s American South and its influence on the film's portrayal of race and identity. Detailed transcriptions of each scene were created, capturing verbal and visual elements, and semiotic analysis was applied to interpret the significance of these multimodal interactions, creating a comprehensive record for analysis.

By following this detailed and systematic approach, this study aims to uncover how "Green Book" employs multimodal elements to convey its complex social messages. This comprehensive analysis contributes to a deeper understanding of the film's narrative and thematic depth, demonstrating the power of multimodal discourse in film studies.

Table 1. Preliminary Data

Scene Element	Description	Symbolism/ Meaning
Textual Analysis	Dialogue between characters	Depicts social dynamics and evolving relationships
Visual Analysis	Composition, colour, perspective	Highlights contrasts and reinforces themes of racial tension and friendship
Multimodal Analysis	Integration of visual and textual elements	Conveys complex social messages about race and identity

III. RESULT AND DISCUSSION

This section presents the findings from the multimodal discourse analysis of "Green Book," focusing on how various semiotic resources are employed to convey complex social messages. The analysis highlights the film's use of cinematography, narrative structures, and character interactions to communicate themes of race, identity, and friendship within the historical and social context of the 1960s American South.

Multimodal Analysis of "Green Book"

The analysis of "Green Book" is structured around the three metafunctions of systemic functional linguistics (SFL), ideational,

interpersonal, and textual, and the principles of visual grammar. This comprehensive approach helps to reveal how the film employs multimodal elements to convey its complex social messages.

Ideational Metafunction

The ideational metafunction analysis reveals that the film represents the complex social experiences of its characters through both dialogue and visual elements. For example, Dr. Shirley's dialogue often reflects his intellectual and cultural struggles as an African American pianist in a racially segregated society. The visual representation of Dr. Shirley's isolation in elegant but confined spaces further emphasizes his alienation.

Interpersonal Metafunction

The interpersonal metafunction examines how the film enacts social interactions and power dynamics. The dialogue between Dr. Shirley and Tony Lip illustrates their evolving relationship from employer-employee to friend. The changing use of language, from formal to informal, and the growing frequency of direct eye contact signify their developing mutual respect and understanding.

Textual Metafunction

The textual metafunction focuses on the organization of the film's narrative. The confined space of the Cadillac serves as a microcosm for broader societal issues. The recurring scenes inside the car provide a cohesive narrative structure, reinforcing the themes of personal and social journeys. The thematic progression from conflict to cooperation between the characters mirrors the broader social changes of the era.

Visual Grammar

The visual grammar analysis demonstrates how the film's cinematographic choices convey meaning. The use of color palettes, such as muted tones for scenes depicting racial tension and vibrant hues for moments of hope and friendship, enhances the narrative's emotional impact. Camera angles and shot compositions also play a critical role, with low-angle shots emphasizing moments of power and high-angle shots depicting vulnerability.

Multimodal Integration

The integration of linguistic and visual elements creates a rich, multimodal narrative. For instance, the scene where Dr. Shirley and Tony Lip share a meal in the car uses close-up shots and warm lighting to signify intimacy and trust. The dialogue in this scene reinforces the visual cues, as both characters express vulnerability and empathy, highlighting their deepening bond.

The findings reveal that "Green Book" effectively uses multimodal elements to address complex social themes. The interplay between dialogue, visual composition, and narrative structure allows the film to convey powerful messages about race, identity, and friendship. The use of systemic functional linguistics and visual grammar frameworks provides a comprehensive understanding of how these elements work together to create a cohesive and impactful narrative.

By situating the film within the historical and social backdrop of the 1960s American South, the analysis reveals how "Green Book" reflects and critiques the era's racial dynamics. The film's ability to engage audiences on both intellectual and emotional levels demonstrates the power of multimodal storytelling in film.

Connecting the findings to broader theories and concepts in film studies, semiotics, and social sciences enriches the discussion. The work of Halliday & Matthiessen (2013) on systemic functional linguistics provides a foundation for understanding how language constructs meaning. Kress & van Leeuwen's (2020) visual grammar framework helps to decode the film's visual elements, while O'Halloran (2011) multimodal discourse analysis integrates these perspectives, offering a holistic view of the film's narrative.

The detailed multimodal analysis of "Green Book" shows how the film employs various semiotic resources to create a rich, layered narrative that engages audiences on multiple levels. This approach not only enhances our understanding of individual films but also contributes to broader discussions about the role of media in society, illustrating how films can serve as powerful tools for communication and social commentary.

Table 2. Aspect Analysis

Aspect	Description	Example in "Green Book"
Ideational Metafunction	Analyses how language is used to represent the real world, ideas, and experiences.	In the scene where Dr. Shirley explains the racial discrimination he faces to Tony Lip: "I am not accepted by my own people because I'm not like them." This dialogue represents Dr. Shirley's struggle with his identity as a highly educated African American in a predominantly white society.
Interpersonal Metafunction	Analyses how language reflects social relationships and power dynamics between characters.	In the confrontation scene between Dr. Shirley and Tony Lip regarding the use of racial slurs, Dr. Shirley's authoritative tone and direct eye contact assert his moral stance, while Tony's defensive body language and avoidance of eye contact indicate his initial resistance to change.
Textual Metafunction	Analyses how language and visual elements are organized to form a cohesive narrative.	The recurring scenes inside the Cadillac serve as a microcosm for broader societal issues, with the progression from conflict to cooperation between the characters mirroring broader social changes of the era. The thematic progression from conflict to cooperation mirrors broader social changes of the period.
Visual Grammar	Analyses visual elements such as composition, color, and perspective to understand how they convey meaning.	In a key scene where Dr. Shirley and Tony Lip perform together at a concert, the use of low-angle shots makes Dr. Shirley appear dominant and confident, signifying his mastery and talent. The warm tones of the concert hall contrast with the cold, muted tones of the segregated hotels, highlighting the disparity in their treatment based on race.
Multimodal Integration	Analyses how linguistic and visual elements work together to create a rich and meaningful narrative.	In the scene where Dr. Shirley and Tony Lip share a meal in the car, the use of close-up shots and warm lighting signifies intimacy and trust. The dialogue in this scene reinforces the visual cues, as both characters express vulnerability and empathy, highlighting their deepening bond.

The analysis of "Green Book" through the lenses of systemic functional linguistics, visual grammar, and multimodal discourse analysis has provided a comprehensive understanding of how the film's linguistic and visual elements work

together to convey complex social messages. By examining the ideational, interpersonal, and textual metafunctions alongside the film's visual grammar, this study highlights the intricate ways in which the film addresses themes of race,

identity, and friendship. The integration of these frameworks not only elucidates the film's narrative depth but also demonstrates the power of multimodal storytelling in engaging audiences on both intellectual and emotional levels.

This research underscores the significance of contextualizing film analysis within its historical and social milieu, offering valuable insights into the broader implications of media as a tool for social commentary and cultural reflection. As such, "Green Book" serves as a compelling case study for the potential of multimodal discourse analysis in film studies, paving the way for future research in this dynamic and interdisciplinary field.

IV. CONCLUSION

This study explored how "Green Book" employs multimodal discourse to convey complex social messages about race, identity, and friendship. By integrating systemic functional linguistics (SFL) and visual grammar, the analysis revealed that the film effectively uses linguistic and visual elements to construct its narrative within the historical and social context of the 1960s American South.

"Green Book" conveys messages about the struggles and resilience of African Americans, the complexities of identity, and the transformative power of friendship. These themes are articulated through both the characters' experiences and the film's visual and linguistic elements.

The meanings in "Green Book" are realized through the interplay of dialogue, visual composition, and narrative structure. Systemic functional linguistics (SFL) is used to analyze how language constructs meaning, revealing how Dr. Shirley's dialogue reflects his intellectual and cultural struggles. Visual grammar analysis demonstrates how cinematographic choices, such as color palettes and camera angles, enhance the emotional and thematic depth of the film. For instance, muted tones depict racial tension, while vibrant hues highlight moments of hope and friendship. Low-angle shots emphasize power, and high-angle shots depict vulnerability, contributing to the film's narrative cohesion and impact.

The film's historical and social backdrop of the 1960s American South grounds its narrative in reality, reflecting and critiquing the era's racial dynamics. This context makes the conveyed messages both relevant and powerful, allowing the film to engage audiences on both

intellectual and emotional levels. By situating the film within its historical and social milieu, the study reveals how "Green Book" effectively addresses and critiques the racial dynamics of its era.

This analysis demonstrates that "Green Book" successfully employs multimodal elements to address and critique social themes relevant to its historical context. The film's ability to engage audiences both intellectually and emotionally highlights the power of multimodal storytelling as a tool for social commentary and cultural reflection.

Contribution to Film Studies and Multimodal Discourse Analysis

The findings from this study significantly contribute to the broader field of film studies and multimodal discourse analysis. By integrating the theoretical frameworks of SFL, visual grammar, and MDA, this research offers a comprehensive approach to analyzing films. This methodological synergy provides a deeper understanding of how films employ various semiotic resources to create rich, layered narratives that engage audiences on multiple levels.

In film studies, this research underscores the importance of considering both linguistic and visual elements in film analysis. It highlights how films can serve as powerful tools for social commentary, reflecting and critiquing societal issues through a combination of dialogue, cinematography, and narrative structures. The detailed analysis of "Green Book" demonstrates the effectiveness of multimodal storytelling in conveying complex social messages, providing a model for future research in film analysis.

In the realm of multimodal discourse analysis, this study illustrates the applicability of MDA to film analysis. It showcases how different semiotic modes-verbal, visual, and auditory-interact to create cohesive and impactful narratives. By examining the interplay between these modes, the research contributes to a better understanding of how meaning is constructed in multimodal texts, offering insights that can be applied to various forms of media.

Broader Context and Implications

The research situates its findings within the broader field of film studies and multimodal discourse analysis, underscoring its contributions to existing scholarship. The integration of SFL, visual grammar, and MDA not only provides a

robust framework for analyzing "Green Book" but also establishes a precedent for future studies in similar domains. This interdisciplinary approach highlights the film's role in engaging with contemporary issues of race and identity, offering new perspectives on how historical and social contexts influence cinematic storytelling.

By connecting the study's findings to broader theoretical and practical contexts, this research enhances our understanding of the potential of films as educational and reflective tools. The insights gained from this analysis can inform filmmakers, educators, and media analysts in their efforts to create and interpret films that address significant social issues. Furthermore, this study paves the way for future research to explore the evolving techniques of multimodal discourse analysis in various genres and media forms, contributing to the dynamic field of film studies and multimodal analysis.

Future research could expand on these findings by examining multimodal discourse in other films, comparing different genres, or investigating the evolution of these techniques over time. Such studies would further illuminate how multimodal elements enrich film narratives, reinforcing the importance of this interdisciplinary approach in film studies.

Kress, G., & van Leeuwen, T. (2020). *Reading Images*. Routledge.
<https://doi.org/10.4324/9781003099857>

O'Halloran, K. (2011). *Multimodal Discourse Analysis* (K. Hyland & B. Paltridge, Eds.). Continuum International Publishing Group.

Sandelowski, M. (2000). Focus on Research Methods: Whatever Happened to Qualitative Description? *Research in Nursing & Health*, 23, 334–340.

Saputra, M. A., & Rosa, R. N. (2020). Multimodal Analysis of Action Movie Posters. *E-Journal of English Language & Literature*, 9(2), 164–174.

Sharjeel, M. Y., & Dadabhoy, K. (2013). Use of Films for Teaching Social Values in English Classes at Elementary Level. *Journal of Elementary Education*, 23, 41–52.

Zhu, J., & Huang, L. (2023). Multi-modal Discourse Analysis of the Movie the Tea House. *Frontiers in Humanities and Social Sciences*, 3(8), 42–50.
<https://doi.org/10.54691/fhss.v3i8.5536>

REFERENCES

Bo, X. (2018). Multimodal Discourse Analysis of the Movie Argo. *English Language Teaching*, 11(4), 132.
<https://doi.org/10.5539/elt.v11n4p132>

Brown, T. (2011). Using Film in Teaching and Learning About Changing Societies. *International Journal of Lifelong Education*, 30(2), 233–247.
<https://doi.org/10.1080/02601370.2010.547615>

Chen, Y., & Gao, X. (2013, June 17). A Multimodal Discourse Analysis of Movie Posters. *2nd Annual International Conference on Language, Literature and Linguistics (L3 2013)*.
https://doi.org/10.5176/2251-3566_L313.106

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2013). *Halliday's Introduction to Functional Grammar*. Routledge.
<https://doi.org/10.4324/9780203431269>