

APPROACHES TO MULTIMODAL ANALYSIS IN VISUAL COMMUNICATION

Suku Hyellamada Kenan Ph.D

*Department of Languages and Linguistics,
Gombe State University,
hyellawilbe@gmail.com.*

Aishatu Bello Umar

*School of Secondary Education Languages, Department of English Language,
Federal College of Education (Technical) Gombe.
aishatubu@gmail.com*

Abdullahi Usman Garko

*Department of languages and Linguistics
augarko@gsu.edu.ng*

&

Ibrahim Ruth Ishaku

*Department of General Studies,
Federal College of Horticulture Dadin-Kowa, Gombe State.
Email: ruthishakuibrahim@gmail.com*

Abstract

This paper examines the major approaches to multimodal analysis and their contributions to the study of visual communication. It explores how leading theoretical frameworks explain the processes through which visual texts generate meaning and influence interpretation. The study reviews and compares key perspectives, including the social semiotic model (Kress & van Leeuwen, 2006), systemic functional multimodal discourse analysis (O'Halloran, 2004), genre-based multimodal semiotics (Bateman, 2008), cognitive approaches to visual metaphor (Forceville, 1996, 2017), and multimodal critical discourse analysis (Machin & Mayr, 2012). Attention is given to how these approaches conceptualise representation, interaction, composition, cognition, genre, and ideology in visual meaning-making. Adopting a qualitative multimodal analytical design, the study applies selected frameworks to visual advertisements drawn from MTN and Airtel campaigns to demonstrate their analytical value. The analysis is guided primarily by Kress and van Leeuwen's social semiotic framework and Forceville's cognitive theory of visual metaphor, while insights from systemic functional, genre-based, and critical discourse perspectives support interpretation. Findings reveal that visual persuasion is achieved through the interaction of linguistic and visual resources, including images, layout, colour, and metaphorical structures that shape audience perception and construct social meanings. The study concludes that no single framework fully captures the complexity of visual communication; rather, an integrative approach that combines semiotic, cognitive, functional, and critical

perspectives provides a more comprehensive understanding of contemporary multimodal texts.

Keywords: critical discourse, cognitive visual metaphor, genre-based multimodal semiotics, multimodal analysis, social semiotics, systemic functional.

1.0 Introduction

In the twenty-first century, communication has become increasingly multimodal, with meaning constructed not solely through language but through the dynamic interplay of text, image, sound, gesture, and spatial design. Among these, the visual mode has attained particular prominence in shaping how information, emotion, and ideology are communicated in contemporary culture. From print advertisements and films to social media platforms and digital interfaces, images have evolved into powerful vehicles of persuasion, identity formation, and social positioning. Consequently, understanding how visual messages function demands analytical frameworks that transcend linguistic analysis to encompass the wider semiotic and multimodal resources through which meaning is constructed.

The field of Multimodal Discourse Analysis (MDA) has emerged as a critical response to this need, providing scholars and analysts with systematic tools to examine how different semiotic modes combine to create meaning in communicative events. Over the past three decades, multimodality has developed into a diverse and interdisciplinary domain, drawing from social semiotics, cognitive linguistics, systemic functional linguistics, media studies, and critical discourse analysis. Each of these traditions contributes unique perspectives to the understanding of visual communication as a meaning-making system: social semiotics foregrounds the cultural and communicative functions of signs; systemic approaches map the grammatical and intersemiotic organization of multimodal texts; cognitive approaches explore the mental and metaphorical processes underlying interpretation; and critical frameworks reveal the ideological and power-laden dimensions of visual discourse.

Despite the proliferation of multimodal research, there remains a pressing need for comparative clarity concerning the theoretical assumptions, analytical foci, and methodological orientations of the major approaches to multimodal analysis. While scholars generally agree that meaning is distributed across multiple modes, they differ in emphasis and scope. For instance, Kress and van Leeuwen (2006) emphasize the social and communicative functions of semiotic resources within the social semiotic model. O'Halloran (2004) extends systemic functional linguistics to multimodal contexts, focusing on structural relationships between modes. Forceville (2017) applies a cognitive metaphor framework to examine conceptual mappings in visual media; and Machin and Mayr (2012) integrate multimodal analysis with critical discourse analysis to uncover ideology and power in visual representation.

Against this background, the aim of this paper is threefold. First, it seeks to review and compare key approaches to multimodal analysis as they pertain to visual communication. ~~Second~~, It aims to demonstrate the practical application of selected frameworks, specifically, the Social Semiotic Approach and the Cognitive Metaphor Approach, through an illustrative analysis of digital advertisements from MTN and Airtel, two leading Nigerian telecommunication brands. Finally, it discusses the broader implications of these approaches for future research in multimodal and visual communication studies.

By integrating theoretical discussion with applied illustration, this paper contributes to the growing dialogue on multimodal theory and practice. It argues that visual communication research benefits most from methodological pluralism, an openness to combining insights from semiotic, cognitive, and critical paradigms. Such an integrative stance is vital in an era of media convergence and digital hybridity, where the boundaries of communication are constantly reshaped by technological innovation and cultural transformation.

2.0 Theoretical Background

2.1 From Semiotics to Multimodality

The theoretical roots of multimodal analysis lie in semiotics, the study of signs and how they generate meaning. Ferdinand de Saussure (1916) defined the linguistic sign as a dual entity composed of the signifier (sound pattern or form) and the signified (concept). Meaning, for Saussure, emerges from the arbitrary relationship between these two elements and the structural differences within a language system. This view, grounded in structural linguistics, established the principle that meaning is relational rather than inherent.

In contrast, Charles Sanders Peirce (1931–1958) proposed a triadic model of the sign, comprising the representamen (form), the object (referent), and the interpretant (the sense made of the sign). Peirce's framework was more dynamic, allowing for the interpretive process of semiosis – the continuous production of meaning through signs. He also categorized signs into icons, indexes, and symbols, based on their mode of relation to the object they represent. These early models of semiotics established the foundation for later theories that conceptualized images, gestures, and other forms of representation as semiotic resources alongside language.

Semiotics, therefore, laid the groundwork for understanding how visual elements operate as signs in communicative processes. The field expanded in the twentieth century through the works of Eco (1976), Barthes (1977), and others who explored how meaning is constructed and naturalized in visual and cultural texts such as photographs, advertisements, and fashion. Barthes's distinction between denotation and connotation – the literal and associative meanings of an image –

remains particularly relevant to visual communication, offering a framework for interpreting the ideological underpinnings of media images.

2.2 Systemic Functional Linguistics and Meaning Functions

The transition from classical semiotics to multimodal analysis was significantly influenced by Michael Halliday's (1978) Systemic Functional Linguistics (SFL), which views language as a social semiotic system. Halliday proposed that all communication simultaneously realizes three metafunctions: the ideational (representation of experience), the interpersonal (enactment of social relations), and the textual (organization of information into coherent discourse). These metafunctions describe how meaning operates at multiple levels of communication – representational, social, and structural.

Halliday's model provided a functional understanding of meaning that extended beyond syntax or lexis to the purposes and contexts of communication. This orientation toward meaning as social action inspired the development of multimodal theories that apply similar principles to non-linguistic modes. In multimodal contexts, the metafunctions serve as an analytical foundation for exploring how images, gestures, and layouts contribute to meaning in ways analogous to linguistic grammar.

2.3 The Rise of Social Semiotics and Multimodal Theory

Building on Halliday's work, Gunther Kress and Theo van Leeuwen (1996, 2006) pioneered a social semiotic approach to multimodality, which treats all modes as culturally and socially shaped resources for communication. They argue that visual communication has its own grammar – a set of semiotic conventions that determine how meaning is encoded and interpreted. In *Reading images: The grammar of visual design* (2006), they identify three metafunctional categories corresponding to Halliday's linguistic metafunctions: representational meaning (depicting participants and actions), interactive meaning (establishing relations between image and viewer), and compositional meaning (arranging elements into a coherent whole). This framework marked a paradigmatic shift from viewing language as the primary vehicle of meaning to recognizing multimodality as the norm in communication. Kress (2010) further contends that multimodality reflects the broader social turn in communication studies, emphasizing how meaning-making is shaped by the social, cultural, and technological contexts in which signs are produced and interpreted.

2.4 Defining Mode and Semiotic Resource

A central concept in multimodal theory is the distinction between mode and semiotic resource. According to Kress (2010), a mode is an organized set of semiotic resources, such as language, image, gesture, or sound, through which people communicate. Each mode has its own materiality, affordances, and limitations. For example, visual modes can represent spatial relations more effectively than language, while linguistic modes can express temporal and abstract concepts with greater precision. A semiotic resource, on the other hand, refers to

the specific means by which meaning is realized within a mode, such as colour, typography, camera angle, or gesture. These resources are not neutral; they are culturally shaped and historically developed, reflecting social practices and communicative needs. As Kress and van Leeuwen (2001) note, “signs are motivated” because their forms and meanings arise from social interests rather than arbitrary convention.

2.5 Meaning as Multimodal Integration

Multimodal discourse analysis (MDA) extends these theoretical principles by examining how multiple modes interact within a communicative event. Meaning is understood as intersemiotic, emerging from the interplay among modes rather than residing in any single one. For instance, in a television advertisement, linguistic text, voice-over narration, camera movement, music, and visual imagery all contribute to the overall meaning. The analytical challenge is to understand how these elements are coordinated to produce coherent and persuasive messages. O’Halloran (2011) describes this process as one of intersemiotic complementarity, where each mode fulfills a specific role while supporting or enhancing others. Similarly, Baldry and Thibault (2006) conceptualize multimodal texts as integrated meaning systems that require analytical tools capable of mapping cross-modal relations. This principle of multimodal integration has become foundational in contemporary visual communication studies, especially in the analysis of advertisements, films, and digital media, where modes operate simultaneously to engage viewers cognitively and emotionally.

In summary, the theoretical background of multimodal analysis emerges from the convergence of semiotics, systemic functional linguistics, and social semiotics. These traditions collectively emphasize that meaning is socially constructed, functionally motivated, and distributed across multiple semiotic modes. The rise of multimodal discourse analysis has transformed visual communication research by providing a systematic means to investigate how images, layouts, and other visual resources work in tandem with language to communicate social, cultural, and ideological meanings.

3.0 Major Approaches To Multimodal Analysis In Visual Communication

3.1 The Social Semiotic Approach (Kress & van Leeuwen)

The social semiotic approach developed by Gunther Kress and Theo van Leeuwen (1996, 2006) represents one of the most influential frameworks in multimodal discourse studies. Rooted in Halliday’s systemic functional linguistics and Peircean semiotics, this approach views communication as a social practice where signs are motivated by the interests and intentions of their creators. Meaning is thus not fixed but dynamically produced within cultural and situational contexts. Kress and van Leeuwen (2006) propose that visual communication, like language, operates through three metafunctions that shape its grammar of meaning:

- i. Representational meaning refers to how images depict people, places, and things, either as narrative structures (involving action and vectors)

- or conceptual structures (depicting stable relationships, classifications, or attributes).
- ii. Interactive meaning concerns the relationship between the image and its viewers, realized through gaze (demand or offer), social distance (intimacy or detachment), and perspective (power relations between viewer and participant).
- iii. Compositional meaning deals with how visual elements are arranged to form a coherent whole, governed by information value (left/right, top/bottom positioning), salience (visual prominence), and framing (boundaries and separations between elements).

Through these metafunctions, the social semiotic model reveals how visual design communicates ideational content, constructs viewer relationships, and organizes meaning compositionally. For instance, in an advertisement, a direct gaze from a model may create a sense of engagement (interactive meaning), while the colour red might draw attention to the brand logo (compositional meaning).

The strength of this approach lies in its ability to integrate semiotic theory with practical analysis of real-world visuals such as advertisements, magazine covers, and digital interfaces. However, it has been critiqued for being descriptive rather than explanatory, offering tools for analysis but not always providing insights into the cognitive or ideological processes underlying visual meaning (Jewitt, 2016). Nevertheless, its accessibility and systematicity have made it foundational to multimodal research.

3.2 Systemic Functional Multimodal Discourse Analysis (O'Halloran)

Building on Halliday's linguistic theory, Kay O'Halloran (2004, 2011) developed Systemic Functional Multimodal Discourse Analysis (SF-MDA), which applies the metafunctional principles of SFL to multimodal data. O'Halloran's model examines how different semiotic modes, such as language, image, and mathematical symbolism, interact to produce meaning in complex communicative events.

This approach is characterized by analytical precision and methodological rigour. It employs detailed annotation systems to map intersemiotic relationships, focusing on how modes complement or extend each other's meaning potential. For instance, in educational materials or scientific texts, language may express abstract concepts while diagrams provide spatial and relational clarity. O'Halloran (2008) identifies three analytical dimensions within SF-MDA:

- i. The ideational metafunction, concerned with how multimodal ensembles represent experiences and processes.
- ii. The interpersonal metafunction, which examines evaluative and attitudinal meanings across modes.
- iii. The textual metafunction, which explores how multimodal elements achieve cohesion and coherence within a text.

A notable contribution of this framework is its computational and corpus-based dimension, which allows for the digital annotation and visualization of multimodal data (O'Halloran, et al, 2017). This has made SF-MDA particularly influential in educational, scientific, and digital communication research where multimodal complexity demands formalized methods.

The limitation of SF-MDA, however, lies in its high level of technicality, which can make it less accessible for analysts working with interpretive or critical approaches. Nonetheless, its systematic approach provides a robust foundation for empirical multimodal analysis.

3.3 Multimodal Semiotics and Genre-based Models (Bateman and Wildfeuer)

Another major framework is John Bateman's Genre and Multimodality (GeM) model (Bateman, 2008), which combines insights from semiotics, text linguistics, and genre theory to analyse the structure and organization of multimodal documents. Bateman's model conceptualizes multimodal artefacts as hierarchically organized systems that follow genre conventions and rhetorical structures. The GeM framework identifies multiple layers of organization:

- i. The document structure, which concerns physical layout and segmentation;
- ii. The layout structure, which governs spatial arrangement and design; and
- iii. The rhetorical structure, which captures relations between multimodal elements (e.g., elaboration, contrast, cause–effect).

This multi-level approach enables analysts to describe how images, text blocks, and layout devices work together to produce coherent multimodal texts. Wildfeuer (2014) extends this perspective to film analysis, emphasizing narrative coherence and inter-modal cohesion.

The genre-based approach is particularly effective for analysing digital media, web interfaces, comics, and print advertisements, where meaning depends on how visual and textual elements are organized according to genre expectations. Its major strength lies in its attention to formal structure and multimodal cohesion, but its limitation is that it may underemphasize ideological or cognitive aspects of meaning (Bateman, 2014).

3.4 Cognitive and Multimodal Metaphor Approaches (Forceville)

The cognitive approach to multimodality, developed primarily by Charles Forceville (1996, 2009, 2017), integrates insights from Conceptual Metaphor Theory (CMT) (Lakoff & Johnson, 1980) into the study of visual and multimodal texts. This framework assumes that much of human thought is metaphorical, and that metaphors structure how we understand abstract concepts through more concrete experiences. In multimodal communication, metaphors are not limited to language but can also be expressed visually, aurally, or gesturally. For example, in advertising, an image of a car with wings may visually instantiate the conceptual

metaphor “A FAST CAR IS A BIRD.” Forceville distinguishes between monomodal metaphors, where the metaphor is realized within a single mode (e.g., visual), and multimodal metaphors, where different modes jointly realize source and target domains (e.g., a visual image and accompanying text).

This approach provides valuable insights into how viewers interpret meaning through cross-domain mappings between conceptual and perceptual structures. It has been applied widely in the analysis of advertising, political cartoons, and film, offering a psychologically grounded perspective on multimodal communication.

Its key strength lies in its attention to cognitive processes and audience interpretation, highlighting how multimodal messages evoke emotion and inference. However, critics argue that it sometimes neglects the socio-cultural and ideological contexts that influence metaphor interpretation (Machin, 2016). Despite this, the cognitive framework remains a crucial complement to social semiotic and critical approaches.

3.5 Multimodal Critical Discourse Analysis (Machin & Mayr)

The Multimodal Critical Discourse Analysis (MCDA) approach, developed by David Machin (2007) and later refined with Andrea Mayr (2012), integrates multimodal analysis with Critical Discourse Analysis (CDA) to examine how power, ideology, and social relations are encoded in visual and multimodal texts.

MCDA extends the principles of CDA, such as the analysis of dominance, hegemony, and social inequality, to include non-verbal semiotic modes. It examines how semiotic choices in visual design (e.g., gaze, modality, setting, colour, and typography) serve ideological purposes. For example, in political or corporate advertising, the use of low angles may signify authority, while certain colour schemes may connote national identity or cultural values.

Machin and Mayr (2012) argue that multimodal texts are never neutral: their design choices reflect institutional and cultural power relations. The analytical task of MCDA is thus to uncover the hidden ideologies embedded in visual representation and to link semiotic structures with broader socio-political discourses. A distinctive feature of this approach is its focus on semiotic resources as ideological strategies rather than mere aesthetic choices. MCDA has been widely applied to the analysis of news imagery, political campaigns, and corporate advertising, offering critical insights into how visual communication legitimizes authority and shapes public perception.

The strength of MCDA lies in its sociocultural depth and critical orientation, though it can be less precise in its methodological procedures compared to systemic or computational frameworks. Nevertheless, its critical lens ensures that multimodal research engages not only with form and function but also with issues of power, ideology, and representation. In summary, these five frameworks represent complementary yet distinct traditions in multimodal analysis. The social semiotic approach foregrounds communicative practice and cultural meaning; the

systemic functional model emphasizes structural mapping and metafunctional correspondence; the genre-based framework focuses on document structure and multimodal cohesion; the cognitive approach highlights conceptual and interpretive processes; and the critical approach exposes ideological and sociopolitical meanings. Together, they provide a comprehensive methodological toolkit for the study of visual communication.

4.0 Comparative Discussion of Multimodal Approaches

The five major approaches (social semiotic, systemic functional, genre-based, cognitive, and critical discourse), reflect the diverse epistemological orientations and analytical traditions that shape the field of multimodal studies. Each offers unique theoretical insights and methodological advantages, but they also differ significantly in scope, assumptions, and analytical focus. A comparative discussion therefore helps to clarify how these approaches can be positioned relative to one another, and how they might be integrated to produce a more comprehensive understanding of visual communication.

4.1 Epistemological Foundations and Theoretical Orientations

The first major point of divergence among these frameworks lies in their epistemological orientation, that is, the nature of knowledge they seek to produce and the assumptions they make about meaning-making. The social semiotic and systemic functional approaches are grounded in constructivist and functionalist paradigms. Both assume that meaning arises from the social use of semiotic resources and that communicative acts can be analyzed in terms of their functional organization (Halliday, 1978; Kress & van Leeuwen, 2006). These frameworks share a concern with how semiotic choices realize social and communicative functions rather than how they represent cognitive or mental processes.

In contrast, the cognitive approach is rooted in cognitive linguistics and psychology, focusing on the mental representations and conceptual mappings that underlie multimodal meaning (Forceville, 2009; Lakoff & Johnson, 1980). Here, meaning is viewed as the outcome of mental construal rather than social interaction. The genre-based model, on the other hand, adopts a structuralist orientation, seeking to describe how multimodal texts are organized according to genre conventions and textual hierarchies (Bateman, 2008). Meanwhile, the critical discourse approach is distinctly ideological and interpretivist, prioritizing power relations, ideology, and sociopolitical critique (Machin & Mayr, 2012).

4.2 Analytical Focus and Units of Analysis

Another key point of comparison concerns the unit of analysis and analytical focus. The social semiotic approach focuses on the semiotic resources available in a mode, such as, colour, gaze, framing, distance, and how they are deployed in context. It is a qualitative, interpretive model suitable for analysing individual images or short sequences. Systemic functional multimodal discourse analysis (SF-MDA), by contrast, employs a more technical, clause-like segmentation of multimodal data, allowing for detailed intermodal mapping and computational

annotation. It seeks analytical consistency and replicability, which makes it especially valuable in large-scale or corpus-based multimodal research.

The genre-based approach uses the text or document as its unit of analysis, focusing on how visual and verbal components are arranged hierarchically to create coherent multimodal ensembles. This model is particularly suited to structured media, such as websites, manuals, and print layouts where genre conventions determine design choices. In cognitive approaches, the unit of analysis is the conceptual structure or mapping between source and target domains across modes. This allows analysts to explain how audiences infer meaning and experience affective or emotional responses through multimodal metaphor. Similarly, MCDA operates on the discourse or ideological formation as its primary unit, connecting visual signs to broader sociopolitical narratives. It moves beyond the immediate text to examine institutional authorship, power asymmetries, and representational politics.

Taken together, these distinctions show that while some frameworks privilege micro-level semiotic detail (e.g., Kress & van Leeuwen, 2006), others focus on macro-level discursive structures (e.g., Machin & Mayr, 2012). An integrative approach thus requires analysts to choose their level of granularity according to the research aim.

4.3 Methodological Complementarities and Integration

Despite their theoretical differences, these frameworks can be seen as complementary rather than mutually exclusive. Each captures a different dimension of multimodal meaning: the structural, functional, cognitive, and ideological. For instance, an advertisement could be analyzed using social semiotic principles to identify its compositional and interactive meanings, while cognitive metaphor theory could explain how the visual rhetoric engages conceptual associations. MCDA, in turn, could reveal the ideological work performed by those semiotic choices. Such integration allows for a multi-layered understanding of how images persuade, position, and naturalize social meanings. This multi-perspectival synthesis aligns with Jewitt's (2016) argument that multimodal research should be "methodologically plural," combining descriptive, interpretive, and critical lenses. It also echoes O'Halloran's (2011) call for intersemiotic complementarity, where different analytical tools capture distinct but interconnected facets of meaning.

However, integration requires theoretical coherence and methodological transparency. Analysts must clarify how frameworks are combined and ensure that the underlying assumptions are compatible. For example, cognitive and critical approaches may conflict in epistemology, one focusing on mental construal, the other on social power, but they can be linked through the shared concept of representation, which mediates between cognition and ideology (Machin, 2016).

4.4 Ideology, Representation, and Agency in Multimodal Communication

A crucial theme across these approaches is the question of representation and agency, who creates meaning, and to what effect. The social semiotic model views sign-makers as socially situated agents who select modes and signs according to

their interests (Kress, 2010). This introduces agency into the design process: producers choose resources to achieve communicative goals within cultural constraints. By contrast, the critical discourse approach shifts attention from individual agency to structural power. It argues that the available semiotic resources themselves are ideologically loaded – certain ways of seeing are privileged while others are marginalized (Machin & Mayr, 2012). The cognitive perspective offers a third angle, emphasizing how individuals construct meaning through embodied experience and metaphorical reasoning. Here, representation reflects cognitive patterns rather than explicit ideology.

Taken together, these perspectives suggest that multimodal communication operates through a dynamic interplay between individual agency, cultural convention, and institutional power. A comprehensive multimodal analysis must therefore account for both the design choices of producers and the interpretive frameworks of audiences, as well as the discursive structures of power within which communication takes place.

4.5 Strengths, Limitations, and Research Implications

Each analytical framework in multimodal discourse studies brings distinctive strengths and limitations that shape the scope and depth of interpretation. The social semiotic approach, grounded in the work of Kress and van Leeuwen (2006), provides a flexible and accessible framework for examining how meaning is constructed through semiotic resources such as image, colour, composition, and typography. Its major strength lies in its ability to uncover the socially and culturally situated nature of meaning-making practices. However, the approach has been criticized for being overly descriptive, with limited empirical validation and occasional neglect of systematic methodological procedures (Bateman, 2014).

In contrast, the Systemic Functional–Multimodal Discourse Analysis (SF-MDA) model offers greater rigour and systematicity, drawing upon Halliday's metafunctional principles to account for ideational, interpersonal, and textual meanings across modes. While this model enhances analytical precision and reproducibility, it can become excessively technical, limiting its accessibility and interpretive depth, especially in analyses that demand cultural or ideological critique.

Similarly, genre-based multimodal models, influenced by Martin and Rose (2008), provide a coherent account of how different modes contribute to the overall cohesion and communicative purpose of a text. These frameworks are particularly valuable for studying pedagogical and institutional genres; however, they tend to overlook the ideological and socio-political dimensions that underpin meaning-making in media and advertising discourse.

Cognitive metaphor approaches, exemplified by Forceville's (1996, 2008) work, offer significant insights into how visual and multimodal metaphors reveal conceptual mappings between source and target domains. Such models highlight the cognitive processes underlying interpretation but may risk individualizing

meaning-making by downplaying the influence of cultural, social, and ideological contexts in shaping audience perception. In contrast, Critical Multimodal Discourse Analysis (CMDA), as advanced by Machin and Mayr (2012), places emphasis on the examination of power, ideology, and social inequality as manifested in multimodal texts. Its critical orientation allows researchers to interrogate how multimodal choices reproduce or challenge dominant ideologies. Yet, without sufficient semiotic grounding, CMDA can sometimes appear impressionistic, privileging ideological readings over detailed semiotic description.

Recognizing the complementarity of these approaches enables researchers to adopt an integrative framework that capitalizes on their respective strengths while mitigating their weaknesses. For instance, an analyst investigating corporate advertisements might employ Kress and van Leeuwen's social semiotic visual grammar to describe the semiotic organization and visual structure of the ads; incorporate Forceville's multimodal metaphor theory to interpret the underlying conceptual mappings that convey brand identity and emotional appeal; and finally, apply Machin and Mayr's critical multimodal discourse framework to uncover how visual and linguistic resources encode ideological meanings and reinforce corporate power relations. Such a triangulated analytical strategy not only enhances the interpretive richness of the study but also aligns descriptive precision with critical depth, yielding a more holistic understanding of multimodal communication.

The implication for multimodal research, therefore, is a move toward hybrid and integrative methodologies that balance analytical depth with socio-cognitive sensitivity. Such synthesis can bridge the gap between form and function, structure and ideology, and sign and cognition – a direction increasingly reflected in contemporary multimodal scholarship (Jewitt, Bezemer, & O'Halloran, 2016). This comparative overview demonstrates that the study of multimodality in visual communication benefits from a multi-dimensional analytical lens. While each approach foregrounds a particular aspect of meaning, none alone can capture the full complexity of multimodal semiosis. The social semiotic and systemic models explain how meaning is structured and realized; the cognitive and genre-based models account for perception and organization; and critical models interrogate power and ideology.

An integrated approach that acknowledges these dimensions can therefore provide a more holistic understanding of how visual communication functions within cultural, cognitive, and ideological frameworks. Such synthesis is particularly valuable in analysing advertising discourse, where persuasion relies simultaneously on semiotic design, conceptual metaphor, emotional appeal, and ideological positioning.

4.6 Integrated Analysis of Social Semiotic and Cognitive Metaphor Frameworks in MTN and Airtel Campaigns:



The analysis of MTN's *Go M.A.D* (Make a Difference) and Airtel's *Data Is Life* campaigns demonstrates how social semiotic and cognitive metaphor frameworks can be integrated to uncover multilayered meanings in visual communication. Drawing on Kress and van Leeuwen's (2006) visual grammar, the social semiotic analysis reveals how each brand mobilizes semiotic resources, such as, colour, gaze, composition, and typography, to construct distinct ideological positions. MTN's campaign emphasizes empowerment and progress through dynamic colour contrast, low-angle shots, and direct gaze, visually encoding action and aspiration. Airtel, conversely, adopts a relational aesthetic, using eye-level shots, domestic settings, and soft lighting to evoke emotional proximity and everyday realism. These semiotic choices show how visual design functions as a mode of social representation, articulating both brand identity and cultural ideology.

Integrating Forceville's (2017) cognitive metaphor framework expands this interpretation from representational to conceptual levels. MTN's metaphors – DATA IS POWER and CONNECTIVITY IS TRANSFORMATION – translate abstract technological functions into embodied experiences of strength, mobility, and progress, visually reinforced through light motifs and upward vectorial composition. Airtel's metaphors – DATA IS LIFE and DATA IS LOVE – construct connectivity as an existential and emotional necessity, symbolized through red palettes, heartbeat motifs, and close-up shots of affection. Through these multimodal metaphors, both brands naturalize digital technology as an essential part of human experience: MTN aligning it with socio-economic development, and Airtel with affective intimacy.

This integrated multimodal approach demonstrates the analytical complementarity of social semiotics and cognitive metaphor theory. While social semiotics elucidates how meaning is structured through visual form and compositional grammar, cognitive metaphor theory reveals how those meanings resonate conceptually and emotionally with viewers. The combined framework enables a holistic interpretation of how Nigerian telecom advertisements transform technical services into cultural narratives of empowerment, belonging, and progress. Methodologically, this synthesis advances multimodal discourse analysis as an interdisciplinary model – capable of decoding both the how (semiotic construction) and the why (cognitive and ideological resonance) of visual meaning in contemporary media.

5.1 Conclusion

This study examined the major theoretical and methodological traditions that underpin multimodal discourse research, highlighting their distinctive contributions and potential intersections. The comparative analysis of the social semiotic, systemic functional, genre-based, cognitive, and critical discourse approaches demonstrates that visual communication is a complex, layered process in which meaning emerges through the dynamic interaction of semiotic, cognitive, and ideological systems. Each framework, whether emphasizing the social use of semiotic resources, the metafunctional organization of modes, the structural conventions of genre, the conceptual mappings of metaphor, or the ideological operations of discourse, offers a particular lens for understanding how images communicate.

However, the findings underscore that no single approach can fully account for the multidimensional nature of multimodal meaning-making. The social semiotic and systemic functional perspectives explain how visual forms structure and organize meaning; cognitive approaches reveal how those meanings resonate at conceptual and affective levels; and critical frameworks interrogate why such meanings matter within broader socio-political contexts. Together, these approaches reveal that visual communication operates simultaneously as a semiotic, cognitive, and ideological process.

Accordingly, the study advocates an integrative analytical paradigm, one that is both methodologically plural and theoretically coherent. By synthesizing the descriptive strengths of social semiotics, the systematic rigor of systemic functional analysis, the cognitive depth of metaphor theory, and the ideological sensitivity of critical discourse analysis, researchers can achieve a more comprehensive understanding of visual communication. This hybrid framework bridges the gap between form and function, cognition and culture, and design and ideology, providing a richer account of how visual texts construct, naturalize, and circulate meaning in contemporary media.

This paper contributes to advancing multimodal discourse analysis as an interdisciplinary field that not only explains how multimodal meanings are made, but also interrogates why they matter in shaping perception, identity, and power. It thus reaffirms the necessity of methodological integration for decoding the complex semiotic work of images in the digital and visual age.

References

- Bateman, J. A. (2008). *Multimodality and genre: A foundation for the systematic analysis of multimodal documents*. Springer.
- Bateman, J. A., Wildfeuer, J., & Hiippala, T. (2017). *Multimodality: Foundations, research and analysis – A problem-oriented introduction*. Walter de Gruyter.
- Forceville, C. (2002). *Pictorial metaphor in advertising*. Routledge.
- Forceville, C. (2017). Visual and multimodal metaphor in advertising: Cultural perspectives. *Styles of Communication*, 9, 24–36.

- Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. Edward Arnold.
- Jewitt, C. (Ed.). (2016). *The Routledge handbook of multimodal analysis* (2nd ed.). Routledge.
- Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.
- Kress, G., & van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. Routledge.
- Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.
- Machin, D. (2007). *Introduction to multimodal analysis*. Bloomsbury Academic.
- Machin, D., & Mayr, A. (2023). *How to do critical discourse analysis: A multimodal introduction*. Routledge.
- Machin, D., & van Leeuwen, T. (2020). *Global media discourse: A critical introduction*. Routledge.
- O'Halloran, K. L. (2004). *Multimodal discourse analysis: Systemic-functional perspectives*. Continuum.
- O'Halloran, K. L., Tan, S., Smith, B. A., & Podlasov, A. (2011). Multimodal analysis within an interactive software environment: Critical discourse perspectives. *Critical Discourse Studies*, 8(2), 109–125. <https://doi.org/10.1080/17405904.2011.558687>
- van Leeuwen, T. (2005). *Introducing social semiotics*. Routledge.