

From Language to the Semiosphere: A Systematic Review of the Social Semiotic Approach to Multimodal Discourse Analysis

XUAN Junbin, SONG Qijun

Guangdong University of Education, Guangzhou, Guangdong 510800, China

Abstract: This systematic review traces the theoretical and methodological trajectory of the social semiotic approach to multimodal discourse analysis. Originating in M.A.K. Halliday's foundational work on language as a social semiotic, the approach has evolved into a comprehensive framework for analysing how meaning is made through the interplay of diverse semiotic resources in social contexts. We chart its development through the critical-ideological turn of Hodge and Kress, the landmark application to visual communication by Kress and van Leeuwen, and its formalisation as Systemic Functional Multimodal Discourse Analysis (SF-MDA). The review synthesises the core principles of the approach—including the motivated sign, the concept of semiotic resource, and the centrality of the metafunctional principle—and examines its extension to non-linguistic modes such as sound, music, and gesture. We survey its application across key contemporary domains, including digital media, advertising, and education, highlighting its analytical power in deconstructing complex multimodal texts. Finally, the paper critically assesses the methodological challenges, scholarly debates, and future directions of the field, considering the impact of interactive digital technologies and the rise of computational semiotics. It is concluded that the social semiotic approach remains an indispensable critical and analytical tool for understanding the complex, multimodal nature of contemporary communication.

Keywords: Social semiotics; Multimodal discourse analysis; Discourse studies; Systematic review

I. Introduction: The Social Turn in Semiotic Inquiry

Semiotics, broadly defined as the study of signs, has undergone a significant paradigm shift over the past half-century, moving from abstract, structuralist models toward a socially grounded theory of meaning-making. This evolution has culminated in the development of social semiotics, an approach that provides the theoretical and analytical foundation for multimodal discourse analysis. This review charts the trajectory of this approach, demonstrating how its core principles have been developed and adapted to account for

the increasingly complex nature of contemporary communication.

The intellectual context for social semiotics is best understood as a critical response to its structuralist predecessors. While foundational theories from Ferdinand de Saussure and Charles Sanders Peirce established the core components of the sign, their frameworks often lacked a robust engagement with the social world. Saussure's diadic model of the sign as a union of a signifier (form) and a signified (concept), and his distinction between the abstract system of language (*langue*) and its individual use (*parole*), provided a powerful

methodology for linguistics. However, a central tenet of his theory—the principle of the "arbitrary" nature of the sign—posited that the link between signifier and signified is based on convention, lacking any natural connection. This focus on the abstract, self-contained system tended to de-emphasise the role of the individual agent and the immediate social context in the act of making meaning.

Social semiotics initiates a fundamental rupture with this tradition by challenging the principle of arbitrariness. It posits that the relationship between form and meaning is not arbitrary but is instead motivated by the "interest of the sign-maker". This shift is more than a minor theoretical adjustment; it represents a profound epistemological break that redefines the object of semiotic inquiry. Where Saussurean analysis focuses on the stabilising force of convention within a system, social semiotics asserts that sign-makers actively choose a signifier as an "apt means" to express a signified, a choice driven by their interests and purposes within a specific social environment.⁴ The analytical focus thus moves from convention to motivation and from arbitrariness to interest, relocating the site of meaning from the abstract system to the concrete, socially-situated agent—the "meaning-maker". This reframing makes any semiotic analysis an inherently social and political investigation. The primary question is no longer "What does this sign mean within the system?" but rather "Why was this particular sign made, by whom, in this context, and to what effect?" This opens the door to a critical analysis of power, ideology, and social strategy, which has become a hallmark of the social semiotic approach.

The core tenet of social semiotics is that meaning is a social product, inextricably linked to social life, processes, and structures. Signs and

messages cannot be understood in isolation; they must always be situated within their sociocultural context.⁴ From this perspective, semiosis—the process of meaning-making in all its complexity—becomes the primary focus of research, and contexts are understood as being constituted by the communicative processes that take place within them. This approach examines how individuals and groups use semiotic resources—words, images, behaviours, sounds, and more—to enact social life and perpetuate or challenge existing power structures.

The evolution of social semiotics from a theory of language to a comprehensive framework for multimodal analysis represents a necessary and powerful adaptation to an increasingly complex communicative environment, which Yuri Lotman termed the "semiosphere". This paper will trace this trajectory, beginning with the foundational principles established by M.A.K. Halliday's view of language as a social semiotic. It will then explore the critical expansion of the theory by Robert Hodge and Gunther Kress, who integrated analyses of power and ideology. Subsequently, the review will detail the landmark application of these principles to visual communication by Kress and Theo van Leeuwen, a development that solidified the field of multimodal analysis. Finally, the paper will examine the formalisation of the approach as Systemic Functional Multimodal Discourse Analysis (SF-MDA), its extension to other semiotic modes, and its contemporary applications and future challenges.

II. The Hallidayan Foundation: Language as Social Semiotic

The theoretical bedrock of the entire social semiotic multimodal project is the work of M.A.K. Halliday. His development of Systemic Functional

Linguistics (SFL) provided the conceptual architecture and analytical tools that would later be extended to analyse semiotic systems beyond language.

Halliday's revolutionary contribution was to reconceptualise language not as a set of formal, abstract rules, but as a "social semiotic". In his foundational 1978 book, *Language as Social Semiotic*, he argued that language is a unique system of signs with a social function, designed to express meanings that are created within a social system.¹ This perspective emphasises the indivisible unity of the text, its context, and the broader social structure.¹ For Halliday, the internal organization of language is not arbitrary; rather, it "embodies a positive reflection of the functions that language has evolved to serve in the life of social man". This functional orientation allows analysts to describe and explain how social reality is encoded in language, both as a means of reflecting on the world and as a means of acting symbolically upon it.

A core innovation of Halliday's theory was to place the concept of "choice" at its centre. Language is conceived not as a rigid code but as a flexible "resource for construing meaning," or what he termed a "meaning potential". The linguistic system offers speakers a vast set of semantic options—what they can mean in a given culture. A "text," whether spoken or written, represents the specific choices made from this potential in a particular instance of communication. A simple analogy is a traffic light system: the meaning potential is the set of options {'stop' OR 'caution' OR 'go'}, while the text is the actual selection of one option, such as 'stop', which is then realised by the form (red). This principle applies to language on a vastly more complex scale, where every linguistic act is an act of meaning.

The cornerstone of SFL, and the key that

unlocks its application to multimodality, is the metafunctional principle. Halliday argued that any act of language is not monofunctional but is, in fact, doing several things at once. These functions cluster into three broad "metafunctions" that are simultaneously expressed in every clause. The first of these is the ideational metafunction, which serves to construe human experience by representing the world around and inside us in terms of processes, participants, and circumstances. This is further divided into the experiential function for representing experience and the logical function for managing relations between clauses.¹¹ The second, the interpersonal metafunction, is the function for enacting social relationships and interacting with others. It is used to establish social roles, express attitudes, ask questions, and give commands, realised through systems like Mood and Modality. Finally, the textual metafunction organises the message itself, enabling it to cohere internally and with its context. It structures information flow through systems such as Theme and Rheme to ensure the text is effective as communication.

SFL provides the analytical tools to deconstruct a text into these three simultaneous strands of meaning, much like a prism refracts white light into its constituent colours. This trifunctional architecture is not merely a descriptive convenience; it is the crucial theoretical innovation that allows the entire field of social semiotic multimodal analysis to exist. Because these functions are derived from the fundamental requirements of any comprehensive communication system—to represent the world, to interact with others, and to create coherent messages—they are not necessarily language-specific. The hypothesis that any semiotic mode aiming to function as a full system of communication must also serve these three metafunctions provides a universal,

function-based architecture for meaning. This creates a common theoretical platform and a shared analytical vocabulary for systematically comparing how different modes, such as image, sound, and gesture, realise ideational, interpersonal, and textual meanings. Without this principle, multimodal analysis would lack a coherent, unifying theoretical core, risking a return to the impressionistic "running commentaries" that characterised early work in the field.

III. The Critical Imperative: Ideology and Power in Hodge and Kress's Social Semiotics

While Halliday provided the functional linguistic foundation, it was Robert Hodge and Gunther Kress who infused the social semiotic project with a critical, political dimension. In their seminal works, particularly *Language as Ideology* and *Social Semiotics*, they expanded the framework to explicitly account for the roles of power and ideology in meaning-making.

Hodge and Kress were strongly influenced by Halliday's functionalism but also drew heavily on the critical theories of Marx and Freud. Their primary objective was to develop systematic tools for analysing the deep interrelation between power and language.⁴ This work was instrumental in the development of "Critical Linguistics" and later provided a major theoretical input into the field of "Critical Discourse Analysis". They argued that a purely functional description of language was insufficient without a critical examination of the social interests that those functions serve.

For Hodge and Kress, the central assumption is that signs and messages must always be situated within a context of social relations and processes, which are fundamentally structured by power. Consequently, every semiotic act is a social strategy that carries an "ideological content". Meaning

is "produced and reproduced under specific social conditions" and is aimed at positioning individuals or groups in relation to one another. They demonstrated this through analyses of diverse examples, from biblical pronunciation debates to cartoon characters, arguing that every act of meaning-making is a social strategy designed to establish or maintain power and authority.

To operationalise this critical perspective, Hodge and Kress introduced several key concepts. A crucial theoretical contribution is the idea of logonomic systems, defined as a set of rules governing meaning-making that orders social life by stipulating "who may make and receive messages and knowledge under what circumstances and with what behaviours". These systems represent the structures of social control embedded within semiosis; for instance, genres exist only because a social group enforces their constituent rules, and subversive discourses like jokes are analysed as violations of the dominant logonomic code. Drawing from sociology, they also argue that every social message contains two fundamental dimensions: power, reflecting social control, and solidarity, reflecting social cohesion. They posit an interdependence between the powerful and the controlled, where the latter often confirm the power of the former through solidarity with the dominant order. Despite this focus on power, Hodge and Kress did not propose a deterministic model, maintaining that meaning is always negotiated. This principle of the negotiated meaning acknowledges that meaning is "never simply imposed inexorably from above by an omnipotent author through an absolute code," thus recognising the interpreter's agency and the potential for resistance.

This critical turn fundamentally transformed the Hallidayan notion of "context". While Halliday's "context of situation" —comprising Field (what is

happening), Tenor (who is involved), and Mode (the role of language)—is primarily a functional construct for describing the environment of a text, Hodge and Kress's concept of the "logonomic system" politicises this environment. They move beyond describing how the context shapes language to asking why the context is structured in a particular way. This reframes context not as a neutral backdrop for communication but as a politically charged field where meanings are produced, regulated, and contested. This critical edge provides the theoretical bridge between functional linguistics and critical discourse analysis, giving analysts the tools to investigate not just what a text means, but also what interests its meaning serves. This perspective has proven indispensable for the analysis of persuasive discourses like advertising and political propaganda, where uncovering ideology is a primary goal.

IV. The Visual Turn: Kress and van Leeuwen's Grammar of Visual Design

The most significant step in extending social semiotics beyond language was the publication of *Reading Images: The Grammar of Visual Design* by Gunther Kress and Theo van Leeuwen. This landmark work systematically applied the principles of Halliday's social semiotics to visual communication, providing the first comprehensive framework for multimodal analysis and establishing a "tool-kit" for a new generation of scholars.

Kress and van Leeuwen's explicit goal was to move beyond subjective, impressionistic commentary on images and to provide a systematic and comprehensive descriptive framework—a "grammar"—for visual design. Their work is explicitly grounded in social semiotic theory, viewing visual representation as a socially and culturally situated process in which sign-makers

choose the most apt semiotic resources to express meanings guided by their own interests. They argued that just as linguistic grammars govern word combinations, a "grammar of the visual" could explain how depicted elements interact to form visual "statements".

The theoretical architecture of their framework is a direct and systematic adaptation of Halliday's three metafunctions to the visual mode. This metafunction concerns how images represent the world, its events, and its participants. It is divided into two primary process types. Narrative representations depict "unfolding actions and events, processes of change, [and] transitory spatial arrangements". Their key feature is the "vector"—a visual element like a diagonal line, pointing limb, or line of sight that connects participants and signifies an action. In contrast, conceptual representations portray participants "in terms of their class, structure or meaning" in a more stable and generalised way. They lack vectors and instead present entities through their composition, their relationship to other entities, or their symbolic essence.

This metafunction concerns how images establish a relationship between the producer, the represented participants, and the viewer, thereby constructing the social interaction enacted by the image. Key analytical systems address this interaction. One is contact, established through gaze, which makes a fundamental distinction between images where a figure looks directly at the viewer to create a 'demand' and those where the figure looks away, 'offering' themselves for contemplation. Another system is social distance, where framing, analogous to camera shots, manipulates the perceived distance; a close-up suggests intimacy while a long shot creates impersonal distance. Perspective, or angle, is

also crucial for constructing power relations. A low angle looking up at a subject invests them with power, while a high angle looking down can suggest vulnerability; similarly, a frontal horizontal angle suggests engagement, while an oblique one implies detachment. Finally, modality refers to the representation's truth-value or credibility, which exists on a gradient from low modality, like a cartoon, to high modality, like a high-resolution photograph that appears true to life. This metafunction concerns how the representational and interactive elements are arranged into a cohesive and meaningful whole—the "text" of the image. Key systems govern this arrangement. Information value dictates that the placement of elements carries specific meanings derived from Western reading practices: the left is the 'Given' (known information) and the right is the 'New'; the top is the 'Ideal' (generalized essence) and the bottom is the 'Real' (practical information), with a central position indicating the 'Nucleus'.²⁴ Salience refers to how elements are made to attract the viewer's attention through features like size, focus, colour contrast, or foreground placement. Lastly, framing involves the use of lines or compositional divisions to either connect or disconnect elements. Strongly framed elements are presented as separate units of meaning, while unframed elements are shown as belonging together.

The primary achievement of Reading Images was not in proving that images possess a grammar identical to that of language, but in demonstrating the immense heuristic power of using the analogy of grammar to systematise the analysis of visual meaning. By seeking functional equivalents in the visual domain for robust SFL concepts—for example, mapping linguistic "transitivity" onto visual "narrative processes" with vectors, or linguistic "mood" onto visual "gaze"—Kress and

van Leeuwen created a "tool-kit" that gave analysts a structured vocabulary and a systematic method for deconstructing an image. This transformed the field by making visual analysis teachable, replicable, and more rigorous. Whilst valid critiques have been raised regarding the tightness of the analogy, the subjectivity of interpretation, and a potential over-emphasis on producer intent, the framework's enduring influence demonstrates the value of this cross-modal heuristic in providing the first comprehensive, function-based language for talking about how images are structured to make meaning.

V. Systematising Multimodality: The SF-MDA Framework

Following the successful application of social semiotic principles to visual design, the field moved toward a more formalised and integrated methodology capable of handling the full complexity of multimodal phenomena. This led to the development of Systemic Functional Multimodal Discourse Analysis (SF-MDA), a framework that extends SFL theory to analyse meaning arising from the simultaneous use of multiple semiotic resources.

SF-MDA is an approach concerned with the theory and practice of analysing meaning in discourses ranging from printed texts to dynamic and interactive digital environments. It is explicitly a social semiotic theory, founded on the principle that all meaning is context-dependent. It aims to develop theoretical and practical tools for analysing how semiotic resources—such as spoken and written language, visual imagery, mathematical symbolism, gesture, music, and architecture—combine to make meaning.

The major strength of SF-MDA, and its core organising principle, is Halliday's metafunctional principle. This principle provides a powerful

integrating platform for theorising how different semiotic resources interact to create meaning. The framework posits that in any multimodal text, choices from different semiotic systems work together to simultaneously construct ideational meaning (representing the world), enact interpersonal meaning (establishing social relations), and create textual meaning (organising the discourse into a coherent whole). This allows analysts to examine the unique functionalities of each resource and to analyse how choices across these resources are orchestrated to achieve specific social objectives, such as persuading a consumer or constructing knowledge in a textbook.

The primary concern of SF-MDA is the phenomenon of "intersemiosis," which refers to the study of how meaning is created, integrated, and expanded through the combination of choices from different semiotic systems.¹⁵ The analysis moves beyond examining each mode in isolation to mapping the "integrative meaning" that arises from their interplay. For example, an SF-MDA of a print advertisement would not only analyse the language and the image separately but would focus on how linguistic and visual choices combine to construct a reality and a persuasive message that would be impossible using either resource alone. This includes studying how meaning can be metaphorically transferred between modes, creating what are known as "semiotic metaphors".

The formalisation of SF-MDA and its increasing reliance on digital tools is a direct methodological response to the escalating semiotic density of contemporary communication. Early social semiotic work could effectively analyse relatively simple combinations of text and image in print media. However, the proliferation of digital media has produced texts—such as websites, videos, and interactive applications—where

multiple semiotic resources are densely interwoven and unfold dynamically over time. The manual analysis of such complex phenomena is often intractable, as it is difficult for the human eye to process and identify all the intricate patterns of inter-semiosis. Consequently, the development of SF-MDA as a formal framework, coupled with computational tools for annotation, time-stamping, and pattern visualisation, was not merely a theoretical refinement but a practical necessity. These software applications function as "meta-semiotic tools," enabling more rigorous, empirically grounded analysis of complex data and managing the multidimensionality of multimodal phenomena. The future of rigorous multimodal analysis is therefore increasingly tied to the co-evolution of social semiotic theory and the computational tools capable of managing the dynamic, data-rich texts that now dominate the semiosphere.

VI. Expanding the Modal Repertoire: Analyses of Sound, Music, and Gesture

A key indicator of the robustness of the social semiotic framework is its successful extension to semiotic modes beyond the traditionally dominant domains of language and image. Scholars have adapted the core principles of meaning potential, motivated choice, and the metafunctional architecture to develop systematic analyses of sound, music, and gesture, demonstrating the theory's adaptability.

Building on the social semiotic project, scholars such as Theo van Leeuwen and David Machin have developed frameworks for analysing sound and music as communicative resources rather than purely aesthetic objects. This approach seeks to identify the "semiotic resources" available to composers, musicians, and sound designers. These resources include melody, pitch range, rhythm,

tempo, instrumentation, and vocal quality (e.g., breathiness, articulation). Meaning is understood to arise from socially established conventions and associations that are built up over time. For instance, a particular sound may acquire meaning through "provenance," where cultural associations are imported from one context to another; the sound of pan pipes, for example, may be used to connote 'nature' or 'ancient Latin American cultures' for a Northern European listener. The metafunctional principle is applied to ask how sound is used to construe ideas (ideational), communicate attitudes and emotions (interpersonal), and create coherence and structure within a text, such as a film soundtrack (textual).

The analysis of gesture within a social semiotic framework treats bodily movements as motivated signs that are an integral part of a multimodal communicative act, often tightly synchronised with speech. Research in this area draws on established typologies, such as those developed by David McNeill, which classify gestures based on their semiotic function in relation to the accompanying

speech. These include iconic gestures, which bear a physical resemblance to the objects or actions they represent; metaphoric gestures, which give physical form to abstract ideas; deictic gestures, which are pointing movements that direct attention; and beats, which are small, rhythmic movements that mark the tempo and emphasis of speech without carrying representational content.

A central debate in the field concerns the degree to which gestures form a systematic, code-like system versus a more fluid, context-dependent resource. The social semiotic perspective emphasises their role within specific "usage events," where their form and meaning are shaped by the immediate communicative needs of the participants.

The consistent application of the metafunctional principle across these disparate modes reveals its power as a unifying analytical framework. Table 1 provides a comparative summary of how the core metafunctions are realised in language, image, sound, and gesture, illustrating the "cross-modal translation" that is central to SF-MDA.

Table 1 Comparative of Metafunctional Realisation across Modes

Metafunction	Language (SFL)	Image (Kress & van Leeuwen)	Sound/Music (van Leeuwen/Machin)	Gesture (McNeill/Kendon)
Ideational	Transitivity (Processes, Participants, Circumstances)	Representational: Narrative (vectors) & Conceptual (classification, symbolism)	Provenance, Experiential Associations (e.g., thunder=danger), Melody	Iconic (representing objects/actions), Metaphoric (representing ideas)
Interpersonal	Mood (Statement, Question), Modality, Appraisal	Interactive: Gaze (demand/offer), Angle, Social Distance, Modality	Perspective (subjectivity), Social Distance (e.g., loud vs. soft), Emotional Valence (e.g., major/minor key)	Affect Displays, Regulators, Deictic (establishing joint attention with viewer)
Textual	Theme/Rheme, Cohesion (Reference, Conjunction)	Compositional: Information Value (Given/New), Salience, Framing	Rhythm, Tempo, Cohesion (leitmotifs), Information Structuring (e.g., musical phrasing)	Beats (rhythmic connection to speech), Cohesives (linking parts of discourse)

This comparative organisation serves as a powerful pedagogical and analytical heuristic. It encapsulates the central theoretical move of SF-MDA by focusing on the function of semiotic choices rather than their material form. The table

makes the framework's logic transparent, showing, for instance, that the linguistic system of Mood (which establishes the relationship between speaker and listener) is functionally analogous to the visual system of Gaze (which establishes the relationship

between the image and the viewer). This provides a practical guide for analysts seeking to apply the theory to new or combined modes and substantiates the claim that social semiotics offers a "grand theory of multimodality".

VII. Contemporary Applications: Mapping the Social World

The true strength of the social semiotic multimodal framework lies in its adaptability and its wide-ranging application to empirical research across diverse domains of contemporary life. Its function-based core allows it to serve as a flexible and powerful analytical lens for deconstructing how meaning is made in virtually any sphere of human communication.

The framework is increasingly applied to analyse the complex multimodal ensembles of digital media. This includes the analysis of websites, where interactivity itself—the affordance of a text to be acted upon through hyperlinks and other user interface elements—is treated as a semiotic resource that can be analysed metafunctionally. A major focus is social media platforms like Instagram, Facebook, and Twitter/X, where users combine text, images, videos, emojis, hashtags, and platform-specific features to construct identities, build communities, and communicate messages. A key theoretical development in this area is the concept of "semiotic technology." This extends the analysis beyond the user-generated content to include the affordances and constraints of the platform itself—such as its interface design, filters, and underlying algorithms—as an active part of the meaning-making environment.

Advertising has long been a classic domain for social semiotic analysis due to its explicit goal of persuasion. Research in this area deconstructs how verbal and visual elements are strategically

combined to create persuasive messages. A central technique is the transfer of "mythic meanings"—such as prestige, beauty, or power—from culturally resonant signs (e.g., celebrities, luxurious settings, national symbols) to the products being advertised. The analysis treats every choice of colour, font, image composition, and wording as a motivated, ideological act designed to embed the product within a specific cultural value system. Similarly, the framework is used to analyse political communication, from official campaign materials to the social media posts of political figures, to uncover how power, authority, and ideology are represented and enacted multimodally.

In the field of education, the social semiotic approach is used to analyse textbooks, educational software, and classroom interactions to understand how knowledge is represented and how pedagogic relationships are established. This research highlights how different modes have different affordances for learning; for example, an image may be more apt for showing spatial relationships, while writing is better suited for constructing a logical argument. In museum studies, the framework is used to analyse exhibition design as a "multimodal design for learning". The spatial layout, lighting, placement of artefacts, and accompanying texts are analysed as a curated "reading path" that guides the visitor's experience and communicates the institution's particular narrative and ideological stance.

The wide variety of these applications—from static print ads to interactive digital environments, and from commercial persuasion to institutional pedagogy—illustrates that the framework's ongoing relevance is due to its ability to evolve and adapt. Its abstract, high-level principles concerning the social agent, their interests, the resources they utilise, and the fundamental metafunctions of

communication offer a robust yet flexible analytical perspective for exploring new technologies and emerging social interactions, making it a lasting and productive research paradigm.

VIII. Critical Reflections and Future Directions

Despite its explanatory power and wide applicability, the social semiotic approach to multimodal discourse analysis is not without its challenges. The field is characterised by ongoing theoretical debates and is continually adapting to a rapidly changing communicative landscape. This section synthesises the major critiques and maps out key directions for future research.

Several persistent challenges confront scholars working within this paradigm. A significant theoretical issue is the lack of clear, stable definitions for foundational concepts like "semiotic mode" and "semiotic resource," as the intrinsic vagueness of these terms and the difficulty in establishing boundaries between them can complicate empirical analysis. Another critique concerns the risk of subjectivity; the interpretive nature of the analysis, especially in visual communication, has been criticised for being overly subjective and lacking empirical grounding, with the potential for the framework to be used simply to illustrate pre-existing concepts rather than being data-driven. Furthermore, there is a tension with cognitive approaches. Scholars from cognitive paradigms have critiqued the social semiotic focus on social structures, arguing that it can under-theorise the cognitive processes involved in meaning-making and that integrating concepts like conceptual metaphor theory could provide a stronger epistemological basis for its claims.

The contemporary digital environment presents profound challenges to traditional analytical methods. One major issue involves

dynamic and interactive texts. Page-based methods developed for static media are often insufficient for analysing texts like videos, websites, and games, where meaning unfolds over time and is co-constructed through user interaction. This requires new methodologies capable of capturing these complex, temporally unfolding processes. Another significant challenge is the role of algorithms. In social media environments, the "text" that a user sees is not solely the product of a human sign-maker but is also shaped by non-human agents like recommendation algorithms. Future research must find ways to incorporate the semiotic role of these systems into the analysis of meaning-making.

The field is actively evolving to meet these challenges, with several promising directions for future research. One key trajectory is toward computational semiotics, which involves integrating social semiotics with computational linguistics and Natural Semiotic Processing. This approach formalises analytical methods into algorithms to process large-scale multimodal datasets, enabling more empirically robust findings and the detection of complex patterns invisible to manual analysis. Another urgent frontier is the analysis of AI-generated content. The social semiotic framework is well-suited to deconstruct the ideological assumptions in AI models and to analyse the semiotics of AI-generated texts and images, including phenomena like "deepfakes" and "slopaganda". There is also a need to develop cross-cultural frameworks, as much of the foundational work was developed in a Western context; refining the framework's applicability in non-Western contexts is crucial for building more globally robust theories. Finally, future theoretical work will likely deepen the engagement with agency and the social agent, exploring the complex dialogue between a sign-maker's choices and the institutional

and technological structures that constrain their semiotic action.

The social semiotic approach, born from the analysis of language and static images, is now at a methodological crossroads. It confronts a communicative landscape that is interactive, algorithmic, and increasingly AI-driven, challenging its core assumptions about authorship, textuality, and agency. The response from within the field is a clear trend toward evolution and integration, particularly through the development of new theoretical concepts like "semiotic technology" and a growing engagement with computational methods. The future viability of social semiotic multimodal analysis depends on its ability to successfully navigate this transition and evolve its conceptual toolkit to account for these new forms of semiosis.

IX. Conclusion

This review has traced the intellectual and methodological journey of the social semiotic approach to multimodal discourse analysis, from its origins in M.A.K. Halliday's functional linguistics to its current status as a sophisticated and adaptable framework for analysing contemporary communication. The trajectory reveals a consistent process of theoretical expansion and adaptation: from language to all semiotic resources, from functional description to critical analysis, and from static texts to dynamic digital environments.

The core contribution of this approach has been to provide a systematic, function-based, and critically aware methodology for understanding meaning-making as a social practice. By rejecting the notion of the arbitrary sign in favour of the motivated sign, it places the social agent at the centre of the semiotic process. Through the versatile architecture of the metafunctional principle, it offers a common platform for analysing the intricate

interplay of diverse modes like image, sound, and gesture. And through its integration of critical theory, it equips analysts with the tools to uncover the ideological interests and power relations that underpin all forms of communication.

Despite facing significant methodological challenges in an era of algorithmic curation and artificial intelligence, the social semiotic approach is more relevant than ever. In a world saturated with complex, persuasive, and often ideologically-charged multimodal messages, its capacity to systematically deconstruct how these messages work and what social interests they serve makes it an indispensable tool. For scholars in discourse analysis, communication studies, and education, it provides a powerful lens for fostering critical literacy and for understanding the profoundly social nature of meaning in our multimodal world.

References

- [1] Halliday M A K. Language as social semiotic: The social interpretation of language and meaning[M]. University Park Press, 1978.
- [2] Saussure F de. Course in general linguistics[M]. McGraw-Hill Book Company, 1966.
- [3] Peirce C S. Collected papers of Charles Sanders Peirce (Vols. 1-8)[M]. Belknap Press of Harvard University Press, 1960-1966.
- [4] Hodge R, Kress G. Social semiotics[M]. Polity, 1988.
- [5] Jewitt C. The Routledge handbook of multimodal analysis[M]. Routledge, 2009.
- [6] Lotman J. On the semiosphere[J]. Sign Systems Studies, 2005, 33(1): 205-229.
- [7] Stöckl H. In between modes: Language and image in printed media[M]//Ventola E, Charles C, Kaltenbacher M. Perspectives on multimodality. John Benjamins, 2004: 9-30.
- [8] Hodge R, Kress G. Language as ideology[M].

- Routledge & Kegan Paul, 1979.
- [9] Machin D, Mayr A. How to do critical discourse analysis: A multimodal introduction[M]. Sage Publications Ltd, 2012.
- [10] Kress G, van Leeuwen T. Reading images: The grammar of visual design[M]. Routledge, 1996.
- [11] Kress G, van Leeuwen T. Reading images: The grammar of visual design[M]. 2nd ed. Routledge, 2006.
- [12] Iedema R. Multimodality, resemiotization: A social semiotic view[J]. Visual Communication, 2003, 2(1): 29-57.
- [13] Vygotsky L S. Thought and language[M]. MIT Press, 1962.
- [14] van Leeuwen T. Speech, music, sound[M]. Macmillan, 1999.
- [15] Machin D. Analysing popular music: Image, sound, text[M]. Sage, 2010.
- [16] Bateman J A, Schmidt K H. Multimodal film analysis: How films mean[M]. Routledge, 2012.
- [17] McNeill D. Hand and mind: What gestures reveal about thought[M]. The University of Chicago Press, 1992.
- [18] Goldin-Meadow S. Hearing gesture: How our hands help us think[M]. Harvard University Press, 2003.
- [19] Bateman J A, Wildfeuer J, Hiippala T. Multimodality: Foundations, research and analysis — A problem-oriented introduction[M]. De Gruyter Mouton, 2017.
- [20] Adami E. A social semiotic framework for the multimodal analysis of website interactivity[J]. Visual Communication, 2015, 14(2): 133-153.
- [21] Poulsen S V, Kvåle G. Studying social media as semiotic technology: A social semiotic multimodal framework[J]. Social Semiotics, 2018, 28(5): 700-717.
- [22] Ledin P, Machin D. Doing visual analysis: From theory to practice[M]. SAGE, 2018.
- [23] Lemke J L. Multiplying meaning: Visual and verbal semiotics in scientific text[M]//Martin J R, Veel R. Reading science: Functional perspectives on discourses of science. Routledge, 1998: 87-113.
- [24] O'Halloran K L. Multimodal discourse analysis[M]//The Bloomsbury handbook of discourse analysis, 2011: 249-282.
- [25] Aiello G, Parry K. Visual communication: Understanding images in media culture[M]. SAGE Publications, 2020.
- [26] Feng D, Espindola E. Integrating systemic functional and cognitive approaches to multimodal discourse analysis[J]. Ilha do Desterro, 2013, 64: 85-110.
- [27] Engebretsen M, Kennedy H. Data visualisation in society[M]. Amsterdam University Press, 2020.
- [28] Shackell C. Natural semiotic processing: A new frontier for multimodal computing?[C]//Proceedings of the 2019 5th international conference on computer and technology applications. ACM, 2019: 1-5.
- [29] Diamantopoulou S. An interdisciplinary engagement with Kress's social semiotic concept of agency: Future directions in multimodal research in museums and schools[C]//5th International Webinar on Multimodal Discourse Analysis, Castellón de la Plana, Spain, 2023.
- [30] Dicks B, Soyinka B, Coffey A. Multimodal ethnography[J]. Qualitative Research, 2006, 6(1): 77-96.
- [31] Forceville C. Evaluating the toolkit of Kress and Van Leeuwen's Reading Images. A review article[J]. Journal of Pragmatics, 1999, 31(3): 435-446.
- Fund:** Research on the Coordinated Development of the Greater Bay Area for the 15th National Games

and the National Paralympic Games: Multilingual Communication Mechanisms of the Symbols in the 15th National Games from the Perspective of the Chinese National Community (2025GBA-043)

©2025. This article is copyrighted by the author and Hong Kong Science and Technology Publishing

Group. This work is licensed under a Creative Commons Attribution 4.0 International License.

<http://creativecommons.org/licenses/by/4.0/>



Open Access