

Multimodal Discourse Analysis of Digital Practices: Theories, Methods, and Future Directions for Museum Website Communication

Qian Li^{1,2,*}

¹ Department of Foreign Language studies, Jinan University, China

² Department of Language Science and Technology, The Hong Kong Polytechnic University, Hung Hom, Hong Kong, China

*Corresponding author: Qian Li

Abstract

This article reviews research on computer-mediated communication, Internet-mediated communication, and multimodal discourse analysis, with a particular focus on museum websites as digital cultural interfaces. It argues that museum websites in multimodal environments in which cultural meaning, institutional identity, and user engagement are constructed through the interaction of text, image, colour, layout, and navigation. Drawing on MDA, MCDA, social semiotics, and Visual Grammar, the article identifies key theoretical foundations and highlights major research gaps. Special attention is given to the representation of Lingnan culture in museum websites and to the balance between cultural symbolism, usability, and intercultural accessibility.

Keywords

digital communication, multimodal discourse analysis, museum websites, Lingnan culture, visual grammar

1. Computer-Mediated Communication and Internet-Mediated Communication

The emergence of new media has generated heightened scholarly interest in the term “Netspeak” [1] to describe the language utilised on the Internet. The Internet exhibits a profusion of media texts, encompassing synchronous (chat groups, Facebook messages, MSN, SMS) and asynchronous (email, webpages) forms. The advent of digital media has led to the development of a new genre termed computer-mediated communication (CMC), which incorporates both verbal and non-verbal components—such as abbreviations, acronyms, and emoticons—to express meaning [2]. Computer-mediated discourse analysis offers insights into sociolinguistics and social semiotics by examining the convergence of linguistic techniques and social activities. This methodology corresponds with Wodak's discourse-historical framework and Fairclough's critical analytical perspective [3].

Additionally, numerous studies investigate Internet-mediated communication (IMC) via cognitive pragmatics, specifically relevance theory and politeness theory [4]. These frameworks clarify Internet users' identities and the interpretation of information in a digital setting. The increasing prevalence of interactive

digital technologies, such as websites, has amplified the need to comprehend how each modality produces symbolic meaning, independently and jointly, within the social framework in which it is utilised.

2. Multimodal Discourse Analysis

The advent of multimodal discourse analysis (MDA) as a theoretical framework has been essential in elucidating the complexities of digital communication. It recognises that meaning-making is a dynamic process influenced by the interaction of various modalities in the digital domain, such as text, images, and videos. MDA is fundamentally interdisciplinary, incorporating multiple fields such as anthropology, philosophy, psychology, cultural studies, art, linguistics, and semiotics. It analyses verbal and non-verbal communication elements and investigates their interaction in generating meaning. The beginnings of MDA are rooted in Halliday's examination of social semiotics, notably articulated in his work *Language as a Social Semiotic*. Multimodality, a relatively recent academic discipline, is strongly linked to the significant idea of *Visual Grammar* presented by Kress and van Leeuwen. MDA has frequently employed this theory, notably in two major studies: Kress and van Leeuwen's *Reading Images* [5] and O'Toole's *The Language of Displayed Art* [6]. Since then, research in Multimodal Discourse Analysis (MDA) has advanced through two principal pathways. The first approach entails analysing the semantic possibilities embedded in mono-semiotic modes, whereas the second explores the interconnections among various semiotic resources within a multimodal framework.

In this setting, Multimodal Critical Discourse Analysis (MCDA) employs a qualitative analytical framework to examine and evaluate the multifaceted creation of meaning. This method clarifies the impact of these constructions on power dynamics and emphasises the significance of language in interactive and sociocultural contexts. Furthermore, it elucidates how linguistic choices contribute to the formation of social meanings, hence influencing and altering power relations. The approaches utilised in this domain include social semiotics, interactive theory, and cognitive linguistics. The multimodal social semiotic approach is based on Halliday's theory of social semiotics. This theory posits that language serves three meta-functions: it represents experiences and logical relations (conceptual meaning), expresses emotions and attitudes while forming social relationships (interpersonal meaning), and integrates both conceptual and interpersonal meanings into a cohesive and coherent text (textual meaning) [7]. Systemic Functional Multimodal Discourse Analysis (SF-MDA) has swiftly evolved, drawing inspiration from Halliday's social semiotics theory. This method emphasises the analysis of discourses utilising various symbolic resources, including verbal and non-verbal language, gestures, architecture, lighting, camera movement, gaze, and angles. A variety of scholars have made contributions to this domain, including O'Halloran [8], Baldry and Thibault [9], Lemke [10], Royce and Bowcher [11], and Feng and O'Halloran [12, 13].

Furthermore, SF-MDA research has suggested the application of Appraisal Theory [14] to broaden discourse analysis beyond clauses within the Systemic Functional Linguistics (SFL) framework [15]. This addition enables academics to elucidate the formation of social identities within multimodal discourse and to do genre analysis. The second primary approach, Interactive Multimodal Discourse Analysis, derives from anthropology, conversation analysis, ethnography, and sociolinguistics. Prominent scholars like Ron Scollon [16] and Sigrid Norris [17] examine the integration and formation of diverse modes within social interactions, identities, and relationships. Jewitt [18] asserts that their work offers comprehensive descriptions of social interactions, emphasising the links among social actors within intricate institutional practices. Moreover, Cognitive Linguistics theories, such as conceptual metaphor and mental spaces, impact MDA research, primarily through Charles Forceville's contributions. His research centres on developing conceptual metaphors within multimodal representations, encompassing visuals, sounds, and gestures across diverse media, including print, television advertisements, comics, cartoons, and films [19]. By applying a critical cognitive analysis and combining insights from cognitive linguistics, visual semiotics, and critical discourse analysis, this interdisciplinary approach allows for a robust examination of how visual elements can influence and reflect ideological and perceptual perspectives, thus elucidating the strategic use of visual symbolic methods for distinct communication objectives. The ideological implications of visual-spatial orientation have been highlighted in many explorations in the domain of advertisement, such as Feng [20], which provides a comprehensive understanding of how spatial orientations can shape viewer's perceptions and ideologies with the insightful analysis of many visual elements from a variety of empirical examples in advertisements. These examples can help to implicate how spatial orientation can convey messages about

power relations, social status and cultural values. The review also highlights an expanding literature on multimodal translation, especially subtitle translation [21], which seeks to bridge multimodal discourse analysis and translation studies from a multidisciplinary perspective to address the intricacies and complexities of cross-cultural communication.

In China, Zhanzi Li was a groundbreaking academic who used visual grammar in multimodal discourse analysis, as evidenced in his article “A Social Semiotic Analysis of Multimodal Texts”. After this fundamental work, scholars such as Yongsheng Zhu [22], Zhuanglin Hu [23], and Delu Zhang [24] have markedly progressed the examination of subtitle translation from a multimodal viewpoint, resulting in a plethora of productive research in this domain. Much of the study on multimodal translation uses Multimodal Discourse Analysis (MDA) to examine diverse discourses and modalities, particularly within film and television, focusing on subtitle translation.

3. Multimodal Discourse Analysis of Digital Practices

Extensive research has been undertaken in multimodal critical discourse analysis, clearly demonstrating the semiotic roles of corporate culture. This study utilises Visual Grammar to examine commercials, highlighting the interdependence and complementarity of distinct modes within an advertisement. The research encompasses both television advertising and online advertisements. The analysis of television advertisements has yielded advantageous outcomes through applying Systemic Functional Linguistics, Critical Discourse Analysis, Multimodal Discourse Analysis, and inter-modal interactions. The findings indicate that the social semiotic framework can thoroughly elucidate the connection between visual representation and ideological foundations. Concurrently, the analysis of online advertisements has gained prominence in the 2020s, driven by developments in digital technology and the rise of live streaming and short films, including e-advertisements [25-27].

In multimodal discourse analysis (MDA) of video games [28, 29], scholars investigate the interaction of visual components, including character design, animation, dialogue, and sound effects, to construct an immersive gaming experience and convey ideological concepts while influencing social reality. Research on websites examines meaning-making processes and the role of semiotic resources in shaping online hypertext to improve user engagement and encourage potential actions. These studies predominantly encompass two domains: education [30], and commerce [31-33]. Nonetheless, there is a significant deficiency in studies concerning the cultural communication of museum websites, especially concerning the equilibrium between intercultural communication and multimodal web design. Compelling and inclusive results are still limited. Researchers have investigated the impact of websites on user experiences and behaviour [34], frequently utilising descriptive methodologies like case studies to evaluate user reactions post-browsing. Nevertheless, corpus-based discourse analysis and empirical research have not been adequately investigated.

Although previous researchers have contributed valuable approaches and theories for evaluating multimodal digital discourse, the domain of multimodal new media is intricate and varied, necessitating an interdisciplinary approach. The swift progression of technology has produced diverse multimodal texts, such as virtual and augmented reality, which require more theoretical investigation and provide new issues for discourse analysis.

4. Theoretical Framework

Van Leeuwen's *Visual Grammar* offers a comprehensive framework for analysing visual pictures by applying linguistic notions to the visual realm. It outlines three essential components of visual meaning: representational meaning (the content shown), interactional meaning (the engagement between the viewer and the participant), and compositional meaning (the arrangement of visual elements). Each dimension aligns with Halliday's metafunctions in systemic functional linguistics.

The framework further analyses the role of colour and other visual indicators, emphasising essential variables such as informational value, salience, and framing in compositional significance. Although it substantially improves our comprehension of visual rhetoric and multimodal communication, the theory possesses limits. It encounters difficulties associated with interpretive subjectivity, intricacy, and an emphasis on static imagery. Furthermore, it may require greater comprehensiveness to grasp the impact of

cultural context and viewers' emotional involvement. Furthermore, as visual communication increasingly occurs in digital media, the framework may require adaptation to suit the unique characteristics of these dynamic modes.

5. Research Gap and Implications for Future Studies

Though empirical studies are fruitful and insightful, the scope of visual data analysed is limited, which still needs to include a broader range of advancing technologies on language usage and communication strategies in Netspeak that affect user interactions at different times and identity formation across various cultural contexts. Building on this observation, the current literature reveals several more specific gaps that jointly constrain the explanatory power and generalisability of existing findings, particularly when the research object shifts from static or semi-static multimodal artefacts to dynamic, interactive, platform-shaped digital discourse.

First, data scope and representativeness remain a persistent issue. Much work in CMC/IMC and MDA still relies on relatively bounded datasets—selected screenshots, single campaigns, or a small number of webpages—partly because multimodal annotation is labour-intensive and because websites and social media interfaces change frequently. However, Netspeak is increasingly multimodal and increasingly embedded in interactive environments. Therefore, dynamic and interactive elements such as website and social media should be incorporated. Future research may expand the data to include more diverse contexts, focusing on comparative analyses across other digital platforms or demographic groups to enhance the generalizability of the findings in comprehending the subtleties of meaning-making in computer-mediated communication. In practical terms, this suggests building broader samples that include different museum (national vs. local), different communication channels (official websites, mobile sites, mini-programs, social media accounts), and different audience segments (domestic/international; general visitors/scholars/students).

Second, the object of analysis needs to be reconceptualised from “page” to “path.” In contemporary IMC, meaning is often created through sequential actions—entering through a search engine, scrolling through a landing page, opening a collections database, using filters, switching languages, and clicking calls-to-action. Yet many multimodal studies still treat webpages as if they were stable visual posters. To address this, a deeper interpretation of temporal and spatial arrangement in hypertext should be conducted through more discourse and cognitive analysis. Specifically, future research could explore how hypertext in these multimodal contexts contribute to the viewer’s perception of cultural construction. Here, “hypertext” is not only a compositional perspective about how menus and links are arranged, but also a cognitive-pragmatic dimension that link structure, information, and interface cues which shape attention allocation and relevance-driven inference, thereby guiding how users construct museum narratives.

Third, there is an under-examined layer of platform and infrastructural mediation. Digital discourse is influenced by factors that are semiotically consequential but not always visible in the text: responsive layouts (mobile vs. desktop), algorithmic ranking (search results within a site), multiple choice tested designs, and analytics-driven content. For museum websites, this can affect what collections appear central, which stories are foregrounded during major events, and whether multilingual content is genuinely equivalent. Future studies should therefore document interface across devices and time, and consider how they contribute to ideological meaning-making.

Fourth, the field needs stronger comparative and intercultural designs. The earlier literature often advances convincing interpretations within one cultural or institutional setting, but without systematic comparison. In that case, it is difficult to claim that observed patterns reflect broader tendencies in multimodal digital discourse. Comparative analyses across platforms (e.g., museum website vs. WeChat/Weibo short posts; long-form exhibition pages vs. short-video explainers) and across cultures (Chinese museums vs. overseas museums) would clarify what is culturally specific, what is institutionally conventional, and what is technologically constrained.

Fifth, methodological development is required to balance interpretive depth with rigour. Visual Grammar and SF-MDA provide rich descriptive tools, but they are sometimes criticized for interpretive subjectivity and a tendency to prioritize static imagery. Future studies can respond by adopting mixed methods: (a) multimodal corpora and transparent coding schemes for interface and visual features; (b) user-centred

evidence such as interviews, questionnaires, eye-tracking or clickstream logs to validate claims about engagement and comprehension; and (c) MCDA triangulation with institutional documents (mission statements, curatorial notes, design guidelines) to link micro semiotic choices to macro ideological positions. Such approaches would allow scholars to retain the strengths of qualitative analysis while increasing replicability and explanatory robustness.

In short, while prior research has established valuable theoretical foundations for analyzing multi-modal digital discourse, the accelerating evolution of technologies and genres (including interactive websites, short-video ecosystems, and emerging VR/AR museum experiences) calls for broader datasets, path-based analysis of hypertext, and integrated discourse–cognitive interaction frameworks. These steps can substantially enhance how we understand meaning-making in Netspeak and in digital communication.

6. Contribution and Significance

Museum websites are essential tools for promoting cross-cultural communication and enhancing user involvement. These websites are necessary to preserve cultural history, develop cultural awareness, and encourage Chinese identity and national image by making digital content available to a broader audience. Museum websites facilitate cultural exchange and promote the global appreciation of China's rich cultures and historical narratives. Building on this core significance, multimodal discourse analysis of museum websites offers contributions at three interconnected levels: cultural–ideological representation, intercultural accessibility, and theoretical–methodological advancement.

6.1 Cultural Representation, National Image, and Multimodal Identity

From an MDA/MCDA perspective, museum websites are not neutral information portals; they are multimodal sites where cultural memory and institutional authority are designed and performed. Semiotic choices in colour, typography, imagery, layout, and navigation contribute to how users perceive legitimacy, continuity, and cultural value. In the context of Lingnan culture, these design elements can be mobilized to communicate a distinct regional identity rooted in southern China's historical, artistic, and social traditions. For example, colour schemes may draw on the visual sensibility of Lingnan architecture, folk art, and seasonal landscape, using warm, vivid, and nature-related tones to evoke the region's climatic environment and cultural atmosphere. Typography may incorporate stylistic features associated with Lingnan calligraphy, traditional inscriptions, or the visual elegance found in local artistic heritage, thereby suggesting cultural depth and historical continuity. Imagery can further strengthen this semiotic framing through references to emblematic elements of Lingnan culture, such as arcade buildings, ancestral halls, Cantonese opera, lion dance, Lingnan gardens, embroidery, ceramics, and motifs associated with the Pearl River Delta. In addition, website layout and spatial organization may reflect aesthetic principles found in Lingnan architecture and garden design, including openness, layered space, adaptability to climate, and a balance between decorative richness and practical functionality.

At the same time, these factors may shape the broader impression of Lingnan culture, which requires attention to the complex interplay of multiple semiotic systems embedded in digital communication. In other words, museum websites function simultaneously as educational media, instruments for the articulation and dissemination of regional culture, and institutional brands. Analysing them through multimodal frameworks makes it possible to uncover how particular ideologies—such as those related to heritage preservation, regional identity, tradition, modernity, and cultural continuity—are encoded not only in curatorial texts, but also in interface hierarchies that determine what is foregrounded, in image modality that influences how the Lingnan past is rendered as authentic, living, or symbolically stylized, and in interaction design that positions the visitor in specific roles, such as learner, cultural participant, tourist, or global audience.

6.2 Intercultural Communication and User Engagement

Museum websites are often the first point of contact for international audiences. Their success depends not only on aesthetic authenticity but also on communicative clarity and cultural intelligibility. A multimodal approach highlights that intercultural communication is achieved through coordinated resources: explanatory captions, bilingual micro-copy, icon design, layout conventions, and narrative sequencing. This is where user involvement becomes measurable and meaningful: engagement is not simply time-on-site, but the extent to

which users can form coherent understandings of artefacts, follow narrative pathways, and feel invited rather than excluded.

Research in this area can therefore provide practical guidance for museums seeking to reconcile the symbolic expression of Lingnan cultural identity with usability and inclusiveness. For example, highly stylised calligraphic or seal-inspired fonts may strengthen cultural resonance and regional distinctiveness, yet they may also reduce readability for non-native readers or broader international audiences. Similarly, colour schemes derived from Lingnan architectural decoration, folk art, or festive traditions may effectively evoke local cultural associations, but they can also create accessibility challenges if contrast, hierarchy, or visual clarity is not adequately considered. MDA can make these tensions more explicit and offer a principled vocabulary for redesign by coordinating semiotic resources in ways that allow Lingnan cultural symbolism and communicative accessibility to support one another.

References

- [1] David, C. (2001). *Language and the Internet*. Cambridge, CUP. Baldry, A., & Thibault, P. J. (2006). *Multimodal Transcription and Text Analysis: A Multimodal Toolkit and Coursebook with Associated On-line Course* (1st ed.). Equinox Publishing Ltd.
- [2] Bodomo, A. B. (Ed.). (2009). *Computer-mediated communication for linguistics and literacy: Technology and natural language education*. IGI Global
- [3] Tannen, D., & Trester, A. M. (Eds.). (2013). *Discourse 2.0: Language and new media*. Georgetown University Press
- [4] Yus, F. (2011). *Cyberpragmatics: Internet-mediated communication in context*. John Benjamins Publishing Company
- [5] Kress, G., & Van Leeuwen, T. (2020). *Reading images: The grammar of visual design*. Routledge.
- [6] O'Toole, M. (1994). *The language of displayed art*. Leicester University Press.
- [7] Halliday, M.A.K. (1978). *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. Edward Arnold, London.
- [8] O'halloran, K. (2008). *Mathematical discourse: Language, symbolism and visual images*. A&C Black.
- [9] Baldry, A., & Thibault, P. J. (2008). Applications of multimodal concordances. *HERMES-Journal of Language and Communication in Business*, (41), 11-41.
- [10] Lemke, J. L. (2002). Travels in hypermobility. *Visual communication*, 1(3), 299–325.
- [11] Royce, T. D., & Bowcher, W. (2013). *New directions in the analysis of multimodal discourse*. Routledge.
- [12] Feng, D., & O'Halloran, K. L. (2012). Representing emotive meaning in visual images: A social semiotic approach. *Journal of Pragmatics*, 44(14), 2067-2084.
- [13] Feng, D., & O'Halloran, K. L. (2013). The visual representation of metaphor: A social semiotic approach. *Review of Cognitive Linguistics*. Published under the auspices of the Spanish Cognitive Linguistics Association, 11(2), 320-335.
- [14] Martin, J. R., & White, P. R. (2003). *The language of evaluation* (Vol. 2). Basingstoke: Palgrave Macmillan.
- [15] Martin, J.R. & Rose, D. (2003). *Working with Discourse: meaning beyond the clause*. London: Continuum
- [16] Scollon, R. (2001). Action and text: towards an integrated understanding of the place of text in social (inter) action, mediated discourse analysis and the problem of social action. *Methods of critical discourse analysis*, pp. 113, 139–183.
- [17] Norris, S. (2004). *Analysing multimodal interaction: A methodological framework*.

- [18] Jewitt, C. (Ed.). (2009). *The Routledge Handbook of multimodal analysis* (p. 340). London: Routledge
- [19] Forceville, C. (2009). Multimodal metaphor. *Mouton de Gruyter*.
- [20] Feng, D. (2011). Visual space and ideology: A critical cognitive analysis of spatial orientations in advertising. In K. L. O'Halloran and B. Smith. (Eds). *Multimodal Studies: Exploring Issues and Domains*. New York & London: Routledge, pp. 55–75.
- [21] Zhang, M., & Feng, D. (Eds.). (2021). *Multimodal approaches to Chinese-English translation and interpreting*. Routledge.
- [22] Zhu, Yongsheng. (2007). *Duomotai huayu fenxi de lilun jichu yu yanjiu fangfa* [The theoretical foundations and research methods of multimodal discourse analysis]. *Waiyu Xuekan*, (5), 82–86
- [23] Hu, Zhuanglin. (2007). *Shehui fuhaoxue yanjiu zhong de duomotaihua* [Multimodalization in social semiotic studies]. *Yuyan Jiaoxue yu Yanjiu*, (1), 1–10.
- [24] Zhang, Delu. (2009). *Duomotai huayu fenxi zonghe lilun kuangjia tansuo* [Exploring an integrated theoretical framework for multimodal discourse analysis]. *Zhongguo Waiyu*, 6(1), 24–30.
- [25] Rizvi, S. Z., Khan, N. W., & Farooq, M. A. H. W. I. S. H. (2020). A Multimodal Discourse Analysis of E-Advertisement Visuals. *International Journal of Psychosocial Rehabilitation*, 24(10), 4473-4492.
- [26] Saladri, R., Dash, M., & Dash, S. (2020). Multimodal Discourse Analysis of Two Digital Advertising Images from Facebook. *Psychology and Education*, 57(9), 935-939.
- [27] Koteyko, N., & Nerlich, B. (2007). Multimodal discourse analysis of probiotic web advertising. *The International Journal of Language, Society and Culture*, 23(23), 20-31.
- [28] Gee, J. P. (2014). *Unified discourse analysis: Language, reality, virtual worlds and video games*. Routledge.
- [29] Pérez Latorre, Ó. (2015). The social discourse of video games analysis model and case study: GTA IV. *Games and Culture*, 10(5), 415-437.
- [30] Michelson, K., & Álvarez Valencia, J. A. (2016). Study Abroad: Tourism or education? A multimodal social semiotic analysis of institutional discourses of a promotional website. *Discourse & Communication*, 10(3), 235–256.
- [31] Ikhlef, A., & Awad, Z. M. (2023). Gender in fashion advertising on social media: a multimodal discourse analysis approach. *Theory and Practice in Language Studies*, 13(7), 1801-1809.
- [32] Moran, C., & Lee, C. (2013). Selling genital cosmetic surgery to healthy women: a multimodal discourse analysis of Australian surgical websites. *Critical Discourse Studies*, 10(4), 373–391.
- [33] Harvey, K. (2013). Medicalisation, pharmaceutical promotion, and the Internet: a critical multimodal discourse analysis of hair loss websites. *Social Semiotics.*, 23(5), 691–714.
- [34] Suphaborwornrat, W., & Punksarikul, P. (2022). A Multimodal Critical Discourse Analysis of Online Soft Drink Advertisements. *LEARN Journal: Language Education and Acquisition Research Network*, 15(1), 627-653.

Funding

The research received support from Social Science fund Project of Guangdong Province (grant number GD23XWY21).

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).