

An Overview of Research on Multimodal Teaching and Learning

Yan Lv *, Wenqian Ni, Yujiao Tan

School of Languages and Literature, University of South China, Hengyang, Hunan, 421001, China

* Corresponding Author: Yan Lv

Abstract: With the development of modern information technology, there emerges an increasing trend of multimodal teaching and learning. Teachers are required to employ a wide range of materials such as textbooks, diagrams, photographs and drawings to facilitate students' learning. Thus, the significance of applying multimodal textbook discourse into classroom setting is also becoming increasingly pronounced. This article takes a retrospective look at the related previous studies on multimodal teaching and learning, which cover a wide variety of perspectives, including the multimodal discourse analysis, the multimodal classroom teaching and the multimodal textbook analysis. Based on the literature review, the deficiencies of current multimodal studies and suggestions for further research have been proposed in the last part of the article. The current study can give important enlightenment to the research and practice of multimodal teaching and learning, especially with the aid of the multimodal textbook discourse analysis.

Keywords: Multimodal Teaching and Learning; Multimodal Discourse Analysis; Multimodal Classroom Teaching; Multimodal Textbook Analysis.

1. Introduction

With the development of modern information technology, there emerges an increasing trend of multimodal communication, challenging people's traditional thinking modes. Consequently, many domestic and foreign scholars have extended the social symbolic characteristics of language to the field of intersemiotics, conducting a series of multimodal studies in a variety of fields, especially in the field of education.

In the procedure of teaching and learning, visual images enable students to relate to learning materials in meaningful ways. In compliance with the needs of the visual world of today, illustrations or images can serve as valuable teaching tools, "bringing to the eye what otherwise can only be imagined". (Evan, Watson & Willows, 1978:86) Therefore, visual images in textbooks can stimulate students' interest and encourage them to learn language in more creative ways.

As early as 1977, Barthes explored the interaction between images and language in his book "Rhetoric of the Image". In 1996, New London Group put forward multiliteracies on the basis of social semiotics. They pointed out that in classroom teaching, teachers should combine people's sense organs, modern technology and surroundings to cultivate students' multiliteracies. Since then, scholars and educators have carried out extensive theoretical and empirical research on this topic.

While the research in China started late, but in the past 20 years, there has also been a research upsurge of "multimodal studies", and the research perspective also shows a diversified trend. In light of this, this article will give a comprehensive overview of the multimodal studies from the following aspects: the multimodal discourse analysis, multimodal classroom teaching, and multimodal textbook analysis.

2. Previous Studies on Multimodal Discourse Analysis

The previous studies of multimodal theory should begin with the review of multimodal discourse analysis (MDA), which can be understood as the foundation for the understanding of the multimodal studies, and it can be examined from both the theoretical and practical development of this theory.

2.1. Overview of the Theoretical Studies of the MDA

The theoretical underpinning of multimodal discourse analysis primarily stems from systemic functional linguistics, as proposed by Halliday(1985), encompassing conceptual function, interpersonal function, and discourse function. Later, in their book "Reading Images: The Grammar of Visual Design", Kress and Leeuwen (1996) have analyzed multimodal discourse from the perspective of social semiotics, establishing the framework of MDA. From then on, many scholars have come to this field and believed that discourse is multimodal in essence(Scollon & Levine, 2004; Kress, 2010; Jewitt, 2016) and that nonverbal resources can also construct potential meanings, new possibilities and challenges emerge for discourse analysts, especially those in the field of multimodal discourse analysis. Li's(2003) introduction of Kress and Leeuwen's "Visual Grammar"(1996) was an endeavor at understanding language as social semiotics. Subsequently, Hu(2007) advocated including various sensory modalities such as visual, auditory, as well as abstract models, structures, and subject consciousness within the scope of research. Meanwhile, he pointed out that at the time when we are entering into a new era marked by multimodal social semiotics, the nurturing of multiliteracy should draw our attention. Zhu(2007) discussed the theoretical basis and research methods of multimodal discourse analysis, specifically, it covers the scopes of the

definition, nature, theoretical basis, content, methodology and significance of multimodal discourse analysis. The latest development is made by Zhang(2023), who explored the theoretical framework of multimodal research while investigating how multimodal discourse is generated from the perspective of multimodal discourse construction.

2.2. Overview of the Practical Studies of the MDA

As for the empirical and practical study of the MDA, O'Halloran and Smith(2011) proposed that MDA can extend the study of language itself to the study of language in combination with other resources, such as images, scientific symbolism, gesture, action, music and sound. Bateman(2014) studied the MDA with a particular focus on the utilization of multimodal corpora in empirical research, and emphasized their role in capturing the intricacies of communication beyond linguistic elements alone. By providing a comprehensive overview of the use of multimodal corpora in empirical research, his research served as a valuable resource for scholars and practitioners interested in exploring the complexities of multimodal communication. Ravelli & McMurtrie(2016) expanded the systemic functional linguistics metafunctional analysis into architectural space analysis, proposing the framework of spatial discourse analysis. Subsequently, McMurtrie(2022) further refined this framework, delving into the relationship between spatial constructs and human behavioral patterns. Sinda(2022) employed a descriptive analysis method to investigate the interpersonal significance of the song "LATHI", qualitatively examining how different symbols and patterns such as music, sound, language, color and facial expressions collectively construct interpersonal meaning.

3. Previous Studies on Multimodal Classroom Teaching

From the perspective of comprehensive construction of multimodal classroom teaching(MCT), scholars have conducted numerous studies, and generally it can be classified into two categories: the theoretical and empirical approaches.

3.1. Overview of the Theoretical Studies of MCT

The most commonly seen approach is to combine with different disciplines to form interdisciplinary theoretical perspectives, such as the application of multimodal approach into classroom teaching. Stein(2000:333) believed that multimodal teaching can facilitate synergistic communication among one's brain, body, and other sense organs. Teachers should explore conceptual and practical issues in relation to representation and the activity of meaning making in classrooms. Royce(2002) carried out analysis of the interrelatedness of semiotic resources and analyzed how teachers could apply teaching materials and various modalities to enhancing students' ability of multimodal communication. His study has the potential to conceptualize pedagogical approaches and can provide a foundation for research into how learners interact with multimodal material in the classroom. Zhang and Ding(2013:39) applied the theory of multimodal genre to the procedure study of foreign language teaching. They claimed design does not involve a direct choice of teaching modes and their combinations, but a

choice based on teaching methods, which makes the choice of modes more practical. Chen, Yang and Cao(2014) explored how to utilize the different functions of classrooms, multimedia, and the Internet to optimize the balance between teachers and students, teachers and teaching materials, and students and teaching materials in the multimodal and multimedia environment of college English classrooms.

3.2. Overview of the Empirical Studies of MCT

Another perspective is to study from the empirical perspective, namely, confirming a certain theory through experimental researches such as corpus studies or classroom-based experiments.

Eliane et al.(2008) investigated the cognitive processes of children during multimodal text learning. The study involved 128 fifth-grade students who received four different modes of instruction. They concluded that the teaching mode of oral presentation with accompanying pictures was most beneficial for students' reading learning, leading to improved reading abilities under such instructional conditions. Song(2013) carried out an empirical research, comparing the effectiveness of a multi-dimensional reading and writing teaching method with traditional teaching methods. The results indicated that the former is significantly more effective than the latter, representing a teaching approach that aligns with the characteristics of college English instruction. Abbas(2014) combined quantitative and qualitative methods to explore the impact of multimodal texts on the reading comprehension abilities of secondary school students. The study results indicated that multimodal reading texts had a positive impact on students' reading abilities. Alejandra et al.(2018) investigated the scaffolding function of multimodal teaching on the reading comprehension abilities of fifth-grade students in Chile with different proficiency levels. The study found that after controlling for variables such as scientific knowledge and vocabulary, multimodal scaffolding did have a significant impact on students' understanding of scientific texts. However, this impact was limited to students with lower reading levels. The scaffolding function of multimodal teaching could assist low-level students in understanding abstract scientific knowledge.

Bateman (2014:238) pointed out that with the development of multimodal research, researchers need to validate and evaluate multimodal theories through a large amount of authentic language data. In light of this, Tuomo (2015) collected 89 tourism brochures produced by the University of Helsinki from 1967 to 2008, totaling 1400 pages, to establish a multimodal corpus. He analyzed the structure of each promotional page and its temporal variations, which can provide multimodal discourse analysts with methodological tools to draw empirically-based conclusions about multimodal documents.

4. Previous Studies on Multimodal Textbook Analysis

A social semiotic approach to text places multimodality and agency in meaning making at the centre of attention (Bezemer & Kress 2008). It ascribes meaning to all modes of communication, including image, writing, typography and layout. The previous studies of textbook analysis can be reviewed from the textbook discourse analysis (TDA) and the image-text relation analysis.

4.1. Overview of the Textbook Discourse Analysis

In terms of the TDA, Machin (2007) discussed several changes that had occurred in textbook discourse over the years, drawing attention to the lower status images had in comparison to written texts. Torres (2015) drew upon Kress and van Leeuwen's (2006) visual grammar to examine an EFL textbook taught in South Korean University context. Her findings showed some instances where the visual message was in contradiction with the verbal message, reflecting the embedded ideologies in the texts and images. Bezemer and Kress(2015) have demonstrated that in textbooks, typography and image are used to construct and differentiate between different imagined abilities as much as writing does. This has important implications for researching and evaluating textbooks and text more widely, since they believed that text designed for readers to engage with aspects of the world cannot be fully understood without due attention to all modes operating in that text. Germán(2022) offered a new framework for analyzing textbook discourse, bridging the gap between contemporary ethnographic approaches and multimodality for a contextually sensitive approach which considers the multiplicity of multimodal resources involved in the production and use of textbooks.

4.2. Overview of the Image-Text Relations

The interactive communication power achieved by different semiotic resources in multimodal texts are key aspects of the multimodal discourse study (O'Halloran, 2011; Jewitt, 2016). In this regard, Barthes (1977) was the first to put forward three kinds of image-text relations: illustration, anchorage, and relay, which were applied to the analysis of newspaper photographs, moving images, and dialogue in films. Halliday (1985) showed how logico-semantic relations recur throughout the lexicogrammar. By integrating Barthes' categorization of image-text relations with Halliday's logico-semantic relations, Martinec and Salway(2005) proposed a dual image-text relation framework, namely, the status and logico-semantic relations, which proved useful for distinguishing between image-text relations for (genuinely) new and old media. Zeng(2010) made comments on three key theories concerning image-text relation, i. e. the theory of image-text status, the theory of image-text zone, and the theory of image-text system, based on their theoretic backgrounds and their contents. Xie (2014) took Martinec and Salway's image-textual logical semantic relations system as the theoretical basis, and proposed a framework for studying image-textual relations in "Grimm's Fairy Tales". Wang and Jiang (2015) took the PEP primary school English textbooks as the research object, applied the logical semantic framework of image-text relations to analyzing 283 images and texts, and found that the images and language in the textbooks had reasonable and effective integration of symbolic resources in form and semantics. If fully utilized, they will have profounding promoting effects on the cultivation of students' multimodal reading ability.

5. Conclusion

Previous reviews on multimodal studies have greatly expanded the scope of research in the field of education, which can give important enlightenment to the multimodal teaching practice in the world. After a careful review of the previous related studies, it can be noticed that most of the

current studies have focused on the fields of multimodal discourse analysis, the multimodal classroom teaching and the multimodal textbook analysis, in terms of both theoretical and empirical studies.

In spite of the fruitful research, we can see that there are still many deficiencies in multimodal teaching research. For one thing, the interaction and coordination of various semiotic resources are still not fully investigated, especially regarding the designing of the multimodal textbooks. Besides, the interdisciplinary studies of applying the multimodal discourse theory to other domains, for instance, the advertising design, creative culture industry, and the like, still need more exploration. Furthermore, we still need to innovate the multimodal research methods by combining the qualitative and the quantitative approaches, especially through effective utilization of multimodal corpora and Open AI technology, which can provide researchers and teachers with a large amount of authentic language data. All in all, it is hoped that through the current study, more scholars and educators will keep focusing on this field, to expand the depth and breadth of the multimodal teaching and learning.

Acknowledgments

This work was supported by grants from the Project of Social Science Foundation of Hunan Province (No. 17YBA 347).

References

- [1] Abbsa, M. B. The efficacy of multimodal vs. print-based texts for teaching reading comprehension skills to Iranian school third graders[J]. *International Journal of Language Learning and Applied Linguistics World*, 2014(3):365-380.
- [2] Barthes, R. Rhetoric of the image [A]. in R. Barthes (ed.) *Image Music Text*[C]. London: Fontana,1977: 32–51.
- [3] Bateman, J. Using multimodal copora for empirical research [A]. in C. Jewitt (ed). *Routledge Handbook of Multimodal Analysis*[C]. London: Routledge, 2014: 238-252.
- [4] Bezemer, J. & Kress G. Writing in multimodal texts: A social semiotic account of designs for learning [J]. *Written Communication*, 2008, 25 (2):166–195.
- [5] Bezemer, J., & Kress, G. The textbook in a changing multimodal landscape[A]. in N.M. Klug & H. Stöckl (eds). *Language in Multimodal Contexts* [C]. New York: De Gruyter, 2015: 21–28.
- [6] Chen Z.F.& Yang Y.& Cao Z.H. The construction of the dynamic balance teaching mode in college English class on the basis of multimodality and multimedia [J]. *Journal of Xidian University (Social Science Edition)*, 2014(04):110-115.
- [7] Evan, M. A. & Watson C. & Willows D.M. A naturalistic inquiry into illustrations in instructional textbooks [J]. *The Psychology of Illustration. Instructional Issues*. New York: Springer US, 1978(02): 86-115.
- [8] Germán C. A Multimodal and ethnographic approach to textbook discourse [A]. in A. Kay & L. O'Halloran (eds.), *Routledge Studies in Multimodality* [C]. New York: Routledge, 2022: 1-142.
- [9] Halliday, M.A.K. & Hasan, R. *Language, context and text: Aspects of language in a social-semiotic perspective* [M]. Victoria: Deakin University Press,1985.
- [10] Hu Z.L. Multimodalization in social semiotics [J]. *Foreign Language Teaching and Research*, 2007(01):1-10.

- [11] Jewitt, C. Multimodal analysis [A]. in Georgakopoulou A. & T. Spilioti (eds.), *The Routledge Handbook of Language and Digital Communication* [C]. London and New York: Routledge, 2016:69–84.
- [12] Kress, G., & Van Leeuwen, T. *Reading image: The grammar of visual design*. [M]. New York: Routledge, 1996.
- [13] Kress, G. *Multimodality: A sociosemiotic approach to contemporary communication* [M]. London: Routledge, 2010.
- [14] Li Z.Z. Social semiotic approach to multimodal discourse [J]. *Foreign Language Research*, 2003,81(05):1-8.
- [15] Machin, D. *Introduction to multimodal analysis* [M]. London and New York: Bloomsbury, 2007.
- [16] Martine R. & Salway A. A system for image-text relations in new (and old) media [J]. *Visual Communication*, 2005(03): 337-371.
- [17] McMurtrie, R.J. Observing, recording, visualising and interpreting visitors' movement patterns in art museums: A mixed method approach [J]. *Multimodality & Society*, 2022, 2(2): 93-113.
- [18] Meneses, A. & Escobar, J. & Veliz, S. The effects of multimodal texts on science reading comprehension in Chilean fifth-graders: text scaffolding and comprehension [J]. *International Journal of Science Education*, 2018(18):2226-2244.
- [19] New London Group. *A pedagogy of multiliteracies: Designing social futures*. *Harvard Educational Review* [J]. 1996, 66(1): 60-93.
- [20] O'Halloran, K. L., & Smith, B. A. Multimodal studies [A]. in K. L. O'Halloran & B. A. Smith (eds.), *Multimodal Studies: Exploring Issues and Domains* [C]. London: Routledge, 2011: 1-13.
- [21] Ravelli, L. J. & McMurtrie R. J. *Multimodality in the built environment: Spatial discourse analysis* [M]. New York: Routledge, 2016.
- [22] Wang, R. H. & Jiang, G.Y. Multimodal foreign language teaching: Blending of image and language resources [J]. *Journal of Schooling Studies*, 2015,12(03):84-90.
- [23] Royce, T. Multimodality in the TESOL classroom: Exploring visual-verbal synergy [J]. *TESOL Quarterly*, 2002, 36(2): 191-205.
- [24] Scollon, R., & Levine, P. Multimodal discourse analysis as the confluence of discourse and technology [A]. in P. Levine & R. Scollon (eds.), *Discourse & Technology: Multimodal Discourse Analysis* [C]. Washington: Georgetown University Press, 2004: 1–6.
- [25] Segers, E. & Verhoeven, L. & Hulstijn-Hendrikse, N. Cognitive processes in children's multimedia text learning [J]. *Applied Cognitive Psychology*, 2008,22(3):375-387.
- [26] Sinda M. & Ahmad J. A Multimodal discourse analysis of "Lathi" song [J]. *International Journal of Linguistics, Literature and Translation*, 2022, 5(1): 66-72.
- [27] Song Q.W. The empirical research on multimodality and the cultivation of multi-dimensional reading and writing abilities in college English [J]. *Foreign Languages Research*, 2013,138 (2): 55–59.
- [28] Stein, P. Rethinking resources multimodal pedagogies in the ESL Classroom [J]. *TESOL Quarterly*, 2000,34 (2): 333-336.
- [29] Tuomo H. *The Structure of multimodal documents: An empirical approach* [M]. New York: Routledge, 2015.
- [30] Torres, G. P. 'Reading' world link: A visual social semiotic analysis of an EFL textbook [J]. *International Journal of English Language Education*, 2015, 3(1):239-253.
- [31] Xie N.N. Analysis of logico-semantic expansion of image-text relations in Grimm's Tales [J]. *Foreign Language Teaching*, 2014, 35(01): 21-25.
- [32] Zhang D.L. & Ding Z.F. Exploration of the multimodal selective framework [J]. *Foreign Language World*, 2013, 156 (03): 39-56.
- [33] Zeng F.B. A study on the construction of image-text relations and its models in multimodal discourse: Comments and analysis on the three image-text relations theories [J]. *Foreign Language and Literature*, 2010, 26(04):60-64.
- [34] Zhang D.L. On the models of modal integration in multimodal discourse construction [J]. *Modern Foreign Languages*, 2023, 46(04):439-451.
- [35] Zhu Y.S. Theory and methodology of multimodal discourse analysis [J]. *Foreign Language Research*, 2007,138(05):82-86.