

# A Review of Multimodality Research: Origins and Developments<sup>1</sup>

**Yunzhi Yang**

Beijing Language and Culture University, China

## *Abstract*

Multimodality has become a new fast-developing research field due to its interdisciplinary and inclusive properties. This paper is organized into two main sections: one dedicated to exploring the origins and theoretical formation of multimodal discourse analysis, and the other to giving illustrations about multimodality research situations, addressing a series of published developments (including authored books, edited collections, and journal articles ranging from 1994 to 2019). The review is arranged in a chronological and selective way, using both English-language and Chinese sources. Following Sigrid Norris (2016), a simple time-division approach is taken; the evolution of multimodality is generally classified as initial, more established, and new stages.

*Keywords: multimodal discourse analysis, multimodality, semiotics, discourse, literature review*

## **1. Introduction**

The rise of multimodal research apparently marks a significant turn in theorizing and analyzing meaning in discourse and semiotics. Given the large expansion and transmutation of the definition of “discourse”, we might ask whether there has been a

traceable turning point during the theoretical formation, and if it would be possible to speculate some academic effects about the emergence of multimodality. This question arises in the disciplinary landscape between discourse analysis and multimodal discourse analysis. As this paper is just designed to present research trends, it is not my aim to trace or explore that specific boundary point. Rather, it is to establish kinds and points of connection between multimodal research and discourse analysis.

## **2. Origins, Transition of Discourse Analysis & Multimodal Discourse Analysis**

The origin of the term “discourse analysis” can be traced to 1952. This term is proposed by American structural linguist Zellig Sabbettai Harris (October 23, 1909 – May 22, 1992), “however, the term first came into general use following the publication of a series of papers by Zellig Harris from 1952 reporting on work from which he developed transformational grammar in the late 1930s.”<sup>2</sup> Then, in the late 1960s and 1970s, the original period of discourse analysis, a large variety of contributions to cross-discipline studies of DA began to develop, including semiotics, psycholinguistics, sociolinguistics, and pragmatics, each of which was subject to its own conceptions, dimensions of analysis, and methodologies. “In origin, and drawing on its early links with traditional sociolinguistics, discourse analysis concerned the study of language use ‘above the sentence’, and focused predominantly on talk and interaction e.g. Hymes, 1972a, 1972b; Coulthard, 1977; Berry, 1981; Tannen, 1984; Sinclair and Coulthard, 1992; Schiffrin, 1994” (Iedema, 2003, p. 30). Perhaps the most significant development in that phase was the emergence of Michael Halliday’s theory. He moved linguistics from the concentration of sentences towards a focus on “text”. Halliday argues that the grammar of a language is not a code, not a set of rules for producing correct sentences, but a resource for making meanings (Halliday, 1978, p. 192). No matter what form it takes, a discourse is best regarded as a unit not of form but of meaning. Halliday’s insights made the analysis of language one part of a “social semiotic” which was beyond the structure of sentences. Following that, Michael

Halliday developed SFL as an “applied linguistics”, achieving the task of describing the grammatical system of language through its various functions. On the basis of the metafunction hypothesis, language is no longer simply conceived as a set of rules, but a semiotic system with a number of subsystems. Halliday’s remarkable work proved very helpful to our understanding of language as a complex system of meaning-making. Furthermore, it made people reflect on the nature of human communication, and how this kind of thinking pattern may no longer be restricted to language but also could be manifested through a variety of communicative modes. Soon afterwards, a group of researchers named the “Hallidayan School” emerged and published various studies about other communicative modes besides language. They mainly utilized linguistic concepts derived originally from Saussure’s principles, Peirce’s ideas, and Halliday’s general theories of language to study non-linguistic modes in the late 1980s. Rick Iedema (2003) summarized this phase as following: “While the semiologists on the continent had been struggling to make De Saussure relevant to the analysis of film and photography (Metz, 1974, 1977; Barthes, 1977), and in America Peircean semiotics and sign theory was flourishing with the work of Thomas Sebeok (Sebeok, 1994; *Semiotica*), the connection between an explicitly systemic-relational approach to meaning making and the analysis of semioses other than language was only just about to be made” (Iedema, 2003, p. 32).

### **3. The Initial Stage of Multimodal Research**

Under these fruitful connections, discourse analysis is thus undergoing a major shift from analyzing language, or mono-mode, to dealing with multi-mode meaning-making practices such as: music, body language, facial expressions, images, architecture, and a great variety of communicative modes. In the beginning, much of this early cross-semiotic or multi-modal work was published in authored and edited books, and only few were published in journals (Norris, 2016, p. 1).

Two foundational books featured and captured the spirit of semiotic resources with non-verbal communication as of the 1990s, Michael O’Toole’s *The Language*

of *Displayed Art* (1994/2011), and Gunther Kress and Theo van Leeuwen's *Reading Images: The Grammar of Visual Design* (1996/2006). Taken together, these two groundbreaking books evidenced a rapidly growing interest across the fields of non-verbal communication, and both were concerned with semiotic resources and visual images. Generally speaking, by the beginning of multimodal research, the need to deal with aspects of visual information alongside verbal language gained acceptance. At the same time, relative notions and frameworks urgently needed to be developed for extending and applying linguistic description to other forms of communication. Under this circumstance, Gunther Kress and Theo van Leeuwen's *Reading Images: The Grammar of Visual Design* has been the most influential book by far. The approach they proposed still structures the work of many multimodal researchers to this day, particularly those who are concerned about visual forms of communication. In their book, Kress and van Leeuwen utilized Halliday's theory to make foundation for the development of Visual Grammar. They adopted the theoretical notion of "metafunction" and created visual grammar with three dimensions: representational meanings, interactive meanings, and compositional meanings. In 1994, the other pioneer in MDA, Michael O'Toole, published his book *The Language of Displayed Art*, focusing more on art works and including architecture and sculpture in addition to painting. The core of this book is the contention that semiotics can assist in the search for a language, and people's perceptions of a work of art can be shared. And also, this book explores the grammar of the visual arts of painting, sculpture, and architecture, proposing that viewers simultaneously read three different kinds of meanings: what is represented (representational meaning), how it engages us (modal meaning) and how it is composed (compositional meaning).

As noted above, the emerging of multimodal discourse analysis in discourse studies moves the focus on the study of language per se to the research of language in combination with other communicative resources. The real traceable marriage of the terms "multimodal" and "discourse" appeared in 2001, though, with the publication of Gunther Kress and Theo van Leeuwen's other work: *Multimodal Discourse: The modes and media of contemporary communication*. (see also Constantinou, 2005,

p. 602). This book contains a new way of thinking and outlines a new theory of communication for the age of interactive multimedia. “In trying to demonstrate the characteristics of these multimodal ensembles we have sketched a multimodal theory of communication which concentrates on two things: (1) the semiotic resources of communication, the modes and the media used, and (2) the communicative practices in which these resources are used” (Kress & van Leeuwen, 2001, p. 111). In this book, the authors provided many inspiring theoretical foundations. Some definitions of theoretical concepts and notions were first explained and illustrated, for instance, they first used the term “multimodal discourse”, defined the term “multimodality”, and outlined a clear demarcation between “modes” and “media”. Kress and van Leeuwen (2001) thus are evidenced as the first to use the term “multimodal discourse” and also the first to come up with “multimodality” in academic publications. (see also Constantinou, 2005, p. 602; Jewitt, Bezemer & O’Halloran, 2016, p. 2) Since then, there has formed an invisible bond tied up within discourse analysis and multimodal research.

Since these groundbreaking works appeared in the 1990s and early 2000s, many theoretical, methodological, and analytical developments across a multitude of disciplines have been put forward by researchers from all over the world. In 2001, van Leeuwen and Carey Jewitt edited the *Handbook of Visual Analysis*. This handbook demonstrates the importance of visual data and mainly provides various methods for visual analysis, including content analysis, historical analysis, structuralist analysis, iconography, psychoanalysis, social semiotic analysis, and so on.

The year 2002 witnessed another burst of newly emergent works. Carey Jewitt published an article which demonstrated the shift from the written page to the multimodal screen. Terry Royce (2002) explored the complementarities of various semiotic resources in multimodal discourse analysis and studied multimodality in the TESOL classroom. In the same year, Kress again teamed up with van Leeuwen (2002) and published the thesis “Colour as a Semiotic Mode: Notes for a Grammar of Colour”. They considered color to be a semiotic mode and discussed the possibility of extending the use of grammar to color as a communicational resource, further

developing their earlier work.

In 2003, Rick Iedema published his pioneering work *Multimodality, Resemiotization: Extending the Analysis of Discourse as Multi-Semiotic Practice*. Since then, the notion of “resemiotization” has been cited by many scholars. In the same year, Ron Scollon and Suzie Wong Scollon published their book *Discourses in Place: Language in the Material World*. This textbook examines the social meanings of the “situatedness” of language and analyzes discourses located in the material world within the framework of “geosemiotics”.

In 2004, three important edited books about multimodal discourse analysis came out. Kay L. O’Halloran’s book *Multimodal Discourse Analysis: Systemic-Functional Perspectives* (2004), Philip LeVine and Ronald Scollon’s *Discourse and Technology: Multimodal Discourse Analysis* (2004) and Eija Ventola, Cassily Charles and Martin Kaltenbacher’s *Perspectives on Multimodality* (2004). Philip LeVine and Ronald Scollon’s (2004) book basically investigates the interrelation between discourse and technology within the domain of multimodality, and mainly concerns the impact of communication technologies on meaning making. Eija Ventola, Cassily Charles, and Martin Kaltenbacher’s (2004) *Perspectives on Multimodality* mainly demonstrates several different research approaches to multimodality. Collected papers are organized into chapters forming two main sections. The first part represents a cross-section of current perspectives on multimodal discourse with a special focus on theoretical and methodological issues. Part Two covers a wide range of applications for multimodal description in fields such as mathematics, education, museum design, entertainment, medicine, and so on. Around the same time, Kay L. O’Halloran (2004) published her representative work: *Multimodal Discourse Analysis: Systemic-Functional Perspectives*. This book mainly contains the theories and practice of the analysis of discourse and the theoretical approach informing these research efforts is Michael Halliday’s systemic functional theory of language, which is extended to other semiotic resources. Kay L. O’Halloran (2004) commented, in the introduction to her book: “The research presented here represents the early stages in a shift of focus in linguistic enquiry where language use is no longer theorized as an isolated phenomenon

(see, for example, Baldry, 2000; Kress, 2003; Kress and van Leeuwen, 1996, 2001; Iedema, 2003; Ventola et al., forthcoming). The analysis and interpretation of language use is contextualized in conjunction with other semiotic resources which are simultaneously used for the construction of meaning” (O’Halloran, 2004, p. 1). In this edited collection, distinguished scholars such as Anthony Baldry and Michael O’Toole both contributed their work as invited papers. Anthony Baldry discussed multimodal transcription and multimodal concordance, while Michael O’Toole (2004) analyzed the Sydney Opera House using Halliday’s framework, and proposed it as a systemic-functional semiotic model of architecture. Similarly, Andrea Hofinger and Eija Ventola (2004) applied multimodal analysis to examples of the Mozart Wohnhaus (Mozart Residence) in Salzburg, Australia. The authors mainly focused on the interaction between pictures and spoken language. Also in this book, Lim Fei (2004) proposed a new theoretical framework (the Integrative Multi-Semiotic Model) to analyze a multimodal text that involved the co-deployment of language and visual images.

In 2005, Martinec and Salway presented their classic framework for image-text relations. Their research investigates the relationship between image and writing with detailed grammar-derived text-image classification. They draw on systemic functional grammar to build their classification system around two main dimensions: status and logicosemantic relations. Relations between texts and images are considered to assign their elements either equal or unequal status, and the nature of the logical semantic linkage between elements are considered as elaboration, extension, and enhancement. In the same year, Kay L. O’Halloran (2005) published another work *Mathematical Discourse: Language, Symbolism and Visual Images*. This book examines mathematical discourse from the perspective of Michael Halliday’s social semiotic theory and explores intersemiotic relations between language, images, and mathematic symbolism.

The literature review presented here represents the early stage of multimodal research. As Carey Jewitt commented: “The starting point for multimodality is to extend the social interpretation of language and its meanings to the whole range

of representational and communicational modes or semiotic resources for making meaning with employed in a culture – such as images, writing, gesture, gaze, speech, posture” (Jewitt, 2014, p. 1).

#### **4. More Established Stage of Multimodal Research**

After the year 2005, the rapid evolution and advancement of the field of multimodality started. Scholars with different research backgrounds begin to use the term “multimodal” and “multimodality”. A variety of distinct theoretical concepts and frameworks continue to emerge in multimodal studies. Meanwhile, it has recently begun to attract a lot of attention from Chinese academic researchers and many scholars in China have noticed the necessity to fill the gaps in multimodal analysis.

It is in this period that previously established directions in multimodal research are being adopted by new scholars, and those who laid the foundations for the field of multimodality continue to offer new ideas and perspectives to further improve their work. For instance, Kress and van Leeuwen (2006) delivered the second version of *Reading Images: The Grammar of Visual Design*. This work, first published in 1996, is a seminal work in the field of multimodality and one of the very few to be entirely dedicated to the building of a new theory. In the second version (2006), the authors refined sections in the first version (1996) about the differences between language and visual communication and brought out some broader semiotic principles that connect, not just language and image, but all the multiple modes in multimodal communication.

The increasing popularity of multimodal research can also be evidenced by various publications (e.g. Baldry and Thibault, 2006; Pahl, 2007; Bateman, 2008; Jewitt, 2008; Norris, 2009; Stenglin, 2009). Baldry and Thibault’s (2006) *Multimodal Transcription and Text Analysis* provided concrete solutions for multimodal text analysis and transcription of printed texts. They came up with two significant ideas: the meaning compression principle and the resource integration principle. Kate Pahl (2007), drawing on a 2-year ethnographic study in an infants’ school in England,

argued that it could be possible to look at children's texts in relation to the lens of literacy events and practices from New Literacy studies. One year later, John A. Bateman (2008) published his book *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents*. This book proposes a new framework for the multimodal genre, which opens up the systematic, corpus-based and theoretically rigorous approach to the description and analysis of multimodal documents. In the same year, Carey Jewitt (2008) investigated multimodality and literacy in the school classroom. Her thesis introduced New Literacies, multimodality, and multiliteracies. Sigrid Norris (2009) discussed the concepts of "modal density" and "modal configuration" and illustrated these notions by investigating social actions. Maree Kristen Stenglin (2009) demonstrated that three-dimensional spaces could also be organized as a multifunctional semiotic resource (a mode). The author proposed a grammar for three-dimensional space and articulated a set of principles to analyze and design three-dimensional spaces with Halliday's notions.

With the increasing emergence of various directions in multimodal research, handbooks and introductions to this field appeared. Examples include *Introduction to Multimodal Analysis* conducted by David Machin (2007) as well as *The Routledge Handbook of Multimodal Analysis* (2009) edited by Carey Jewitt. In David Machin's textbook, the author acknowledges his indebtedness to the scholarly work of Gunther Kress and Theo van Leeuwen in the study of visual representation, and clearly explains this groundbreaking approach to visual analysis in a very comprehensive and systematic way. In 2009, Routledge published his first handbook about multimodality, *The Routledge Handbook of Multimodal Analysis* (2009) edited by Carey Jewitt. "This handbook accomplishes two things: one, the book brings together and showcases the thoughts of many scholars working in and towards multimodality at the time; and two, the handbook offers a moment of reflection on just where multimodality sees itself in 2009" (Norris, 2016, p. 1). Even Gunther Kress noted the importance of the publication of this handbook in his 2010 book, *Multimodality: A Social Semiotic Approach to Contemporary Communication*. "Carey Jewitt's *Handbook of Multimodal Analysis* (2009) gives a very good sense of the range and diversity of that work; and

of the interconnections” (Kress, 2010, p. xiii).

As noted earlier, generally speaking, multimodal analysis in China is definitely still an emerging, new field of research which has to address disciplinary borders and linguistic domains. However, the growing interest in conferences on multimodality as well as many publications that have arisen from many perspectives show that Chinese linguistics finally seems to be overcoming its traditional restrictions and limitations.

According to Zeng Fangben (2009), Chinese studies on multimodal discourse analysis started from the first introduction of visual grammar made by Li Zhanzi in 2003. Then Hu Zhuanglin (2007), Zhu Yongsheng (2007), and Gu Yueguo (2007) went further to start the study of multimodal discourse. Hu Zhuanglin (2007) analyzed the differences between multimodal semiotics and multimedia semiotics. He expressed his ideas about some relevant problems such as the replacement of books with screens, the designing of multimodal signs, and the realization of coherence. Zhu Yongsheng (2007) came up with four issues in multimodal discourse analysis: the first one concerned the origin of multimodal discourse; the second one was the definition; the third one involved the nature and theoretical basis of the subject; the fourth one was the content, research methods, and significance of multimodal discourse analysis. Gu Yueguo (2007) made a distinction between multimedia learning and multimodal learning. He also proposed an analytic model named agent-based modeling language (AML developed by the author) and used it to datamine the multimedia and multimodal interactions for computer simulation.

It is noteworthy that Wei Qinhong (2009) has made a valuable contribution to multimodal discourse analysis by her book entitled *Multimodality and Multimodal Discourse in Visual Surroundings*. This is the first published reference book about multimodal discourse in China. In this book, the author first discussed globalization and “new writing”, and provided some new research methods. Wei Qinhong is also the first multimodal discourse analyst in China to give a detailed explanation and illustration about the differences between mode and modality. Around the same time, another Chinese pioneer in multimodality, Zhang Delu (2009), taking a systemic functional linguistic perspective, provided a theoretical framework for multimodal

discourse analysis. The theoretical framework he outlined mainly consists of five levels, including semantics, form, expression, context, and culture. Then in 2010, his focus moved to investigating the concept of design and the selection of modalities in multimodal foreign language teaching.

As presented earlier, while many early writings about multimodal research were collected in books, the rest of the “multimodality” research came together in the journal *Visual Communication*, whose first issue appeared in 2002. *Visual Communication* thus became the rallying point for scholars who were interested in multimodality.

Since then, a myriad of edited volumes and papers that are dedicated to multimodality have been produced. With more and more platforms emerging for scholars working in different disciplines, various directions in multimodality are beginning to blossom, including media studies, semiotics, new literacy studies, education, sociology, and psychology, addressing a vital and broadening range of different research questions.

## **5. The New Stage of Multimodal Research**

As stated by Sigrid Norris (2016), most early writings about multimodal research were published in authored or edited books and only a few published in journals. However, the year 2012 marks a shift in that trend. First, most works concerning multimodality since then have been published in the form of journal articles. Second, scholars from across the world have begun to contribute to the growth of multimodal research.

In 2012, Wendy L. Bowcher (Professor at Sun Yat-sen University, China) edited the book *Multimodal Texts from Around the World*. It is a collection of twelve original articles by scholars from around the world with fascinating and absorbing insights about multimodal texts. This book is a first in multimodal discourse studies, as it focuses primarily on texts from non-English speaking countries, and the chapters are arranged according to the regions from which the texts originate, including Africa, the Middle East, Europe, Russia, Asia, Oceania, and Brazil.

In New Zealand, Paul White (2012) shared his points about the challenges to traditional marketing theory from a multimodal analytical approach. White's work argues that the convergence of traditional media with new communication technologies is challenging the precepts of conventional marketing theory regarding the design of posters and billboards and the reception of the messages they mediate. Also in New Zealand, Sigrid Norris (2013) redefined the important term, "mode". She explicated the definition of "mode" as a system of mediated action with regularities and illustrated this concept by taking a perfume blog entry as her starting point.

In Singapore, a team of researchers in a multimodal analysis lab (Interactive Digital Media Institute at the National University of Singapore), led by Kay L. O'Halloran, developed an interactive software for multimodal analysis. This software permits multimodal data to be generated in a systematic manner and facilitates empirically grounded results. Then, Kay L. O'Halloran (2012) coauthored an article with Alexey Podlasov, Alvin Chua, and Marissa K. L. E to present the design and functionalities of their newly developed interactive software for multimodal analysis.

In 2013, Jennifer Rowsell, Gunther Kress, and Brain Street wrote an article together in England. They investigated body art (tattoos on a subject's body) and approached the task from three different perspectives (they termed as three "optics"): social semiotics, sedimented identity in texts, and new literacy studies in relation to multimodality, drawing on ethnographic perspectives. In this article, each optic illustrated how a woman constructed her identity through her body art. Emilia Djonov and Theo van Leeuwen (2013) presented an article on layout and PowerPoint. Drawing on developments in the visual arts, graphic design, and social semiotics, the authors demonstrated two central and complementary principles for layout design and layout analysis: the grid and composition.

In the same year, another significant paper about multimodality was published in *Cognitive Science*. This paper, titled *Visual Narrative Structure*, was conducted by American cognitive scientist and comics theorist Neil Cohn (2013a). Around the same time, he summarized his work in a book named *The Visual Language of Comics: Introduction to the Structure and Cognition of Sequential Images* which was

published by Bloomsbury. His hierarchic interpretation of images challenged many of the existing conceptions of linguistics and filled a gap by presenting the theory of narrative grammar. Generally speaking, visual grammar (proposed by Gunther Kress and Theo van Leeuwen in 1996) describes how compositional factors integrate for understanding individual images, while visual narrative grammar (proposed by Neil Cohn in 2013) deals with several pictures arranged in a time sequence. Cohn's work argues that the structure and comprehension of graphic images, particularly sequential images, should parallel the study of verbal and signed languages. The essence of his theories is that visual narrative grammar organizes the structure of sequential images in the visual language, similar to the way that syntax organizes words into coherent sentences. He describes the basic narrative categories and their relationship to a canonical narrative arc, which demands that the canonical arc be reconsidered as a generative schema whereby any narrative category can be expanded into a node in a tree structure. Narrative "pacing" is interpreted as a reflection of various patterns of this embedding: conjunction, left branching trees, center-embedded constituencies, and others. Following this, the basic parameters of this narrative grammar could be outlined as: peaks, initials and releases (core categories), establishers and prolongations (expendable categories). By the way, there exists a lot of variation in the naming of Neil Cohn's grammar: visual narrative structure/visual narrative grammar/visual language grammar.

In Germany, John A. Bateman (2014) came out with another important work *Text and Image: A Critical Introduction to the Visual-Verbal Divide*. This introductory textbook gave a broad and interdisciplinary explanation of text-image relations and explored various approaches to multimodality (including socio-semiotics, visual communication, psycholinguistic approaches to discourse, rhetorical approaches to advertising and visual persuasion, and cognitive metaphor theory).

In 2015, Belinda Crawford Camiciottoli (at the University of Pisa, Italy) and Inmaculada Fortanet-Gómez (at Universitat Jaume I, Spain) co-edited the book *Multimodal Analysis in Academic Settings: From Research to Teaching*. The authors presented 10 innovative articles on the multimodal dimension of discourse specific

to academic settings. This book is mainly composed of two parts: the first is research communications (focused on the multimodal features of conference presentations), and the second is classroom applications (exploring the role of multimodality in the classroom through five different classroom events, including lectures, student presentations, and listening activities).

In 2016, Sigrid Norris (Professor at Auckland University of Technology in New Zealand) edited *Multimodality (Critical Concepts in Linguistics), Volumes I-IV*. One important feature of these volumes is the breadth of contributions, resulting in a well-rounded view of multimodal studies. The four volumes collected the major achievements in the field of multimodality over the period from 1956 to 2014, and demonstrated the developments and research trends. This collection is a rare and valuable reference to offer extensive background reading for anyone who is interested in multimodal research. In the same year, Tuomo Hiippala (2016) in Finland published his book *The Structure of Multimodal Documents*. Hiippala provided a new framework for describing and interpreting the textual, visual, and spatial aspects of page-based multimodal documents.

In Madrid, Paula Pérez Sobrino (Universidad Politécnica of Madrid) published *Multimodal Metaphor and Metonymy in Advertising* in 2017. The author, taking a cognitive linguistic perspective, studied meaning (re)construction in advertising and argued that ordinary creativity can be described and assessed by a set of identifiable analytical tools: metaphor, metonymy, and their combinations in multimodal use. This book offers a tentative proposal to identify multimodal metaphor and also this is the first in-depth research monograph to study multimodal metaphor and metonymy in a large corpus of real advertisements (Pérez Sobrino, 2017, p. 206).

In 2018, Arnulf Deppermann (Deutsche Sprache) teamed up with Jürgen Streeck (The University of Texas at Austin) to edit the book *Time in Embodied Interaction: Synchronicity and Sequentiality of Multimodal Resources*. As is generally known, the interactional approach descends from conversation analysis and works concerned with interactional analysis are now increasingly described as “multimodal”. This book contains in-depth illustrations about how people synchronize and sequence modal

resources such as their talk, gaze, gestures, and object-manipulation to accomplish social actions, and this book is the first one to dedicate to the study of the complexities that arise in embodied interaction from the multiplicity of timescales (Deppermann & Streeck, 2018). Arnulf Deppermann is also the scholar who used to comment that “‘multimodality’ is a label which is already worn out and has become most fuzzy by its use in various strands of semiotics, discourse, and media analysis” (Deppermann, 2013, p. 2). (see also Bateman, Wildfeuer & Hiippala, 2017, p. 23).

In Denmark, Nina Nørgaard (2019), taking a new analytical approach, explained semiotic resources involved in the multimodal meaning-making of the novel, covering wording, typography, layout, images, and book-cover design. Nørgaard’s work *Multimodal Stylistics of the Novel: More than Words* got published as one of the Routledge Studies in Multimodality series. In this book, the author advocates that all novels are inherently multimodal and combines traditional stylistic text analysis with social semiotic multimodal theory to propose a new framework for multimodal stylistic analysis of the novel.

Similarly, interest in multimodality has also swept across China. In 2014, Chinese scholars Feng Dezheng and Zhang Delu co-authored a paper with Kay L. O’Halloran from Singapore. This paper was published in *Contemporary Linguistics*. The three authors discussed and shared the present achievements and new advances in multimodal research. Their work argues that the focus on multimodal analysis has gradually transitioned from simply analyzing non-linguistic modes to the interdisciplinary systemic analysis of multimodal meaning making resources, including the research demand for the development of new research tools and the new frameworks which are supposed to integrate traditional linguistics with inter-disciplinary studies. In 2016, Gu Yueguo published his thesis about multimodal experiencing and situated cognition. He illustrated the 3E Model (Experiencer-Experiencing-Experience) and proposed a new framework to investigate the lifespan development of situated cognition. Zhang Yiqiong (2017) published a book entitled *A Multimodal Discourse Study of Online Science News: Synchronic and Diachronic Perspectives* which was grown out of her PhD dissertation, instructed

by Kay L. O'Halloran.

In 2018, Hu Zhuanglin (2018), taking the semiotic perspective, provided insights into multimodality and fragmentation. The author illustrated the key concepts and categories of fragmentation as well as its advantages and disadvantages. In 2019, inspired by the definition of the scientific paradigm by Thomas Kuhn (1996), Huang Lihe and Zhang Delu (2019) argued that multimodal study as a research paradigm integrating different approaches could be compared to the “Multi-core Parallel System” in computer science.

Finally, more and more scholars around the world are taking a multimodal approach in their writings. A variety of disciplines and theoretical perspectives should be used to explore different aspects of multimodal research. The landscape of international conferences and events in this field is thriving. Multimodality has become an increasingly popular topic of many seminars and academic conferences, for instance, the ACM International Conference on Multimodal Interaction (the last conference: 21th ACM in Suzhou, Jiangsu, China, 2019) as well as the International Conference on Multimodality (the last conference: 9th ICOM in 2018, at the University of Southern Denmark in Odense, Denmark). Workshops and other events around this topic also provide platforms for researchers to share knowledge and experience, discuss emerging issues, and promote the development of multimodality, in events such as Multimodality and Cultural Change (2015, University of Agder, Norway), Multimodality in Social Media and Digital Environments (2016, Queen Mary University of London, England), the first China Multimodality Forum (2017, Tongji University, Shanghai, China), the International Conference on Multimodal Communication (2018, Hunan Normal University, Hunan, China), and the International Conference on Applicable Linguistics and Multimodality (2019, Shanghai Jiao Tong University, Shanghai, China).

More and more conferences in the field of communication studies, media studies, cultural studies, journalism studies, language, and linguistics have been dedicated exclusively to multimodality, which is showing the interdisciplinary and comprehensive research need of the field. Additionally, multiple edited collections are

evolved from these conferences and forums, for instance: Arlene Archer and Esther Breuer's (2015) *Multimodality in Writing. The State of the Art in Theory, Methodology and Pedagogy* (a collection of thirteen seminal papers); Ognyan Seizov and Janina Wildfeuer's (2017) *New Studies in Multimodality: Conceptual and Methodological Elaborations* (this book is the result of the Second Bremen Conference on Multimodality in September 2015); Evripides Zantides's (2017) *Semiotics and Visual Communication II: Culture of Seduction* (Chapters in this book consist of selected papers that were presented at the 2nd International Conference & Poster Exhibition on Semiotics and Visual Communication at the Cyprus University of Technology in October 2015); Huang Lihe and Zhang Delu's (2018) *Multimodal and Foreign Language Education Research* (this book is the result of the 1st China Multimodality Forum in November 24-26, 2017, Tongji University, Shanghai, China); Elise Seip Tønnessen and Frida Forsgren's (2019) *Multimodality and Aesthetics* (some of the contributions in this volume were presented at the Multimodality and Cultural Change closing conference at the University of Agder in 2015).

To cater for the lively research area, apart from authored books and edited collections, more and more handbooks and introductions to multimodality emerged. For instance, *The Routledge Handbook of Multimodal Analysis 2nd Edition* (C. Jewitt, 2014), the De Gruyter *Handbook of Language in Multimodal Contexts* (N.-M. Klug & H. Stöckl, 2016), *Introducing Multimodality* (C. Jewitt, J. Bezemer, & K. O'Halloran, 2016), and *Multimodality-Foundations, Research and Analysis: A Problem-Oriented Introduction* (J. Bateman, J. Wildfeuer, & T. Hiippala, 2017).

In order to advance knowledge of multimodal resources, the Taylor & Francis Group launched the book series "The Routledge Studies in Multimodality" and De Gruyter Mouton founded the *Journal of Multimodal Communication* as a new venue for exploring diverse and multimodal ways to conduct research and illustrate findings.

## **6. Final Words**

Multimodality is one of the most influential fields for human communication analysis.

It enjoys growing global popularity and is increasingly taken up by scholars with new ideas and unique insights. The number of academic publications on multimodality has grown remarkably over the last twenty years, and demonstrates that the need for exploring multimodality has been recognized. Thus, the era of multimodal communication has arrived and has brought new connections and challenges, which demands openness and motivation for largely interdisciplinary research. Under this circumstance, it has become an irresistible trend for scholars to deal with multimodal research and to redefine the territory of communication in modern times.

## Notes

- 1 This paper is funded by the Fundamental Research Funds for the Central Universities and the Research Funds of Beijing Language and Culture University (19YCX065).
- 2 Retrieved October 26, 2018 from [https://en.wikipedia.org/wiki/Discourse\\_analysis](https://en.wikipedia.org/wiki/Discourse_analysis)

## References

- Archer, A., & Breuer, E. (Eds.). (2015). *Multimodality in writing: The state of the art in theory, methodology and pedagogy*. Amsterdam: Brill.
- Baldry, A. (2004). Phase and transition, type and instance: Patterns in media texts as seen through a multimodal concordancer. In K. L. O'Halloran (Ed.), *Multimodal discourse analysis: Systemic-functional perspectives* (pp. 83-108). London: Continuum.
- Baldry, A., & Thibault, P. J. (2006). *Multimodal transcription and text analysis*. London: Equinox Publishing.
- Bateman, J. A. (2008). *Multimodality and genre: A foundation for the systematic analysis of multimodal documents*. Houndmills, Basingstoke, Hampshire and New York: Palgrave Macmillan.
- Bateman, J. A. (2014). *Text and image: A critical introduction to the visual-verbal divide*. London and New York: Routledge.
- Bateman, J. A., Wildfeuer, J., & Hiippala, T. (2017). *Multimodality: Foundations, research and analysis—A problem-oriented introduction*. Berlin/Boston: Walter de Gruyter.
- Bowcher, W. L. (Ed.). (2012). *Multimodal texts from around the world*. Houndmills, Basingstoke, Hampshire and New York: Palgrave Macmillan.

- Cohn, N. (2013a). Visual narrative structure. *Cognitive Science*, 37(3), 413-452.
- Cohn, N. (2013b). *The visual language of comics: Introduction to the structure and cognition of sequential images*. London and New York: Bloomsbury.
- Constantinou, O. (2005). Multimodal discourse analysis: Media, modes and technologies. *Journal of Sociolinguistics*, 9(4), 602-618.
- Crawford Camiciottoli, B., & Fortanet-Gómez, I. (Eds.). (2015). *Multimodal analysis in academic settings: From research to teaching*. London and New York: Routledge.
- Deppermann, A. (2013). Multimodal interaction from a conversation analytic perspective. *Pragmatics*, 46(1), 1-7.
- Deppermann, A., & Streeck, J. (Eds.). (2018). *Time in embodied interaction: Synchronicity and sequentiality of multimodal resources*. Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Djonov, E., & Van Leeuwen, T. (2013). Between the grid and composition: Layout in PowerPoint's design and use. *Semiotica*, (197), 1-34.
- Fei, L. (2004). Developing an integrative multi-semiotic model. In K. L. O'Halloran (Ed.), *Multimodal discourse analysis: Systemic-Functional perspectives* (pp. 220-246). London: Continuum.
- Feng, D.-Z., Zhang, D.-L., & O'Halloran, K. L. (2014). 多模态语篇分析的进展与前沿 [Advances and frontiers of multimodal discourse analysis]. 《当代语言学》 [Contemporary Linguistics], (1), 88-99.
- Gu, Y.-G. (2007). 多媒体、多模态学习剖析 [On multimedia learning and multimodal learning]. 《外语电化教学》 [Computer-Assisted Foreign Language Education], (2), 3-12.
- Gu, Y.-G. (2016). 当下亲历与认知、多模态感官系统与大数据研究模型——以新生婴儿亲历为例 [Multimodal experiencing, situated cognition and big data with a demonstrative analysis of a newborn baby]. 《当代语言学》 [Contemporary Linguistics], (4), 475-513.
- Halliday, M. A. K. (1978). *Language as social semiotic*. London: Edward Arnold.
- Hiippala, T. (2016). *The structure of multimodal documents*. London and New York: Routledge.
- Hofinger, A., & Ventola, E. (2004). Multimodality in operation: Language and picture in a museum. In E. Ventola, C. Charles, & M. Kaltenbacher (Eds.), *Perspectives on*

- multimodality* (pp. 193-209). Amsterdam: John Benjamins.
- Hu, Z.-L. (2007). 社会符号学研究中的多模态化 [Multimodalization in social semiotics]. 《语言教学与研究》 [*Language Teaching and Linguistic Studies*], (1), 1-10.
- Hu, Z.-L. (2018). Fragmentation from the semiotic perspective. *Language and Semiotic Studies*, 4(4), 1-12.
- Huang, L.-H., & Zhang, D.-L. (2019). 多核并行架构:多模态研究的范式、路径及领域问题之辨 [A multi-core parallel system: Paradigm, approaches and fields in multimodal study]. 《外语教学》 [*Foreign Language Education*], (1), 21-25.
- Iedema, R. (2003). Multimodality, resemiotization: Extending the analysis of discourse as multi-semiotic practice. *Visual Communication*, 2(1), 29-57.
- Jewitt, C. (2002). The move from page to screen: The multimodal reshaping of school English. *Visual Communication*, 1(2), 171-196.
- Jewitt, C. (2008). Multimodality and literacy in school classrooms. *Review of Research in Education*, 32(1), 241-267.
- Jewitt, C. (Ed.). (2009). *The Routledge handbook of multimodal analysis* (1st ed.). London and New York: Routledge.
- Jewitt, C. (Ed.). (2014). *The Routledge handbook of multimodal analysis* (2nd ed.). London and New York: Routledge.
- Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). *Introducing multimodality*. London and New York: Routledge.
- Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. London and New York: Routledge.
- Kress, G. (2012). Multimodal discourse analysis. In J. P. Gee & M. Handford (Eds.), *The Routledge handbook of discourse analysis* (pp. 35-50). London and New York: Routledge.
- Kress, G., & Van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold.
- Kress, G., & Van Leeuwen, T. (2002). Colour as a semiotic mode: Notes for a grammar of colour. *Visual Communication*, 1(3), 343-368.
- Kress, G., & Van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). London and New York: Routledge.
- LeVine, P., & Scollon, R. (Eds.). (2004). *Discourse and technology: Multimodal discourse*

- analysis*. Washington: Georgetown University Press.
- Li, Z.-Z. (2003). 多模式话语的社会符号学分析 [Social semiotic approach to multimodal discourse]. 《外语研究》 [*Foreign Languages Research*], (5), 1-8.
- Machin, D. (2007). *Introduction to multimodal analysis*. London: Hodder Education.
- Martinec, R., & Salway, A. (2005). A system for image-text relations in new (and old) media. *Visual Communication*, 4(3), 337-371.
- Norris, S. (2009). Modal density and modal configurations: Multimodal actions. In C. Jewitt (Ed.), *The Routledge Handbook of Multimodal Analysis* (1st ed.). (pp. 78-90). London and New York: Routledge.
- Norris, S. (2013). What is a mode? Smell, olfactory perception, and the notion of mode in multimodal mediated theory. *Multimodal Communication*, 2(2), 155-169.
- Norris, S. (2016a). *Multimodality III*. London: Routledge.
- Norris, S. (2016b). *Multimodality IV*. London: Routledge.
- Nørgaard, N. (2019). *Multimodal stylistics of the novel: More than words*. London and New York: Routledge.
- O'Halloran, K. L. (Ed.). (2004). *Multimodal discourse analysis: Systemic-Functional perspectives*. London: Continuum.
- O'Halloran, K. L. (2005) *Mathematical discourse: Language, symbolism and visual images*. London: Continuum.
- O'Halloran, K. L., Podlasov, A., Chua, A., & Marissa, K. L. E. (2012). Interactive software for multimodal analysis. *Visual Communication*, 11(3), 363-381.
- O'Toole, M. (2004). Opera Ludentes: The sydney opera house at work and play. In K. L. O'Halloran (Ed.), *Multimodal discourse analysis: Systemic-Functional perspectives* (pp. 11-27). London: Continuum.
- O'Toole, M. (2011). *The Language of displayed art* (2nd ed.). London: Routledge.
- Pahl, K. (2007). Creativity in events and practices: A lens for understanding children's multimodal texts. *Literacy*, 41(2), 86-92.
- Pérez Sobrino, P. (2017). *Multimodal metaphor and metonymy in advertising*. Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Rowse, J., Kress, G., & Street, B. (2013). Visual optics: Interpreting body art, three ways. *Visual Communication*, 12(1), 97-122.

- Royce, T. (2002). Multimodality in the TESOL classroom: Exploring visual-verbal synergy. *TESOL Quarterly*, 36(2), 191-205.
- Scollon, R., & Scollon, S. W. (2003). *Discourses in place: Language in the material world*. London and New York: Routledge.
- Seizov, O., & Wildfeuer, J. (Eds.). (2017). *New studies in multimodality: Conceptual and methodological elaborations*. London and New York: Bloomsbury.
- Stenglin, M. K. (2009). Space odyssey: Towards a social semiotic model of three-dimensional space. *Visual Communication*, 8(1), 35-64.
- Tønnessen, E. S., & Forsgren, F. (Eds.). (2019). *Multimodality and aesthetics*. London and New York: Routledge.
- Van Leeuwen, T., & Jewitt, C. (Eds.). (2001). *Handbook of visual analysis*. London: Sage Publications.
- Ventola, E., Charles, C., & Kaltenbacher, M. (Eds.). (2004). *Perspectives on multimodality*. Amsterdam: John Benjamins.
- Wei, Q.-H. (2009). 《视觉环境下的多模态化与多模态话语研究》 [*Studies on multimodality and multimodal discourse in visual surroundings*]. Beijing: Science Press.
- White, P. (2012). Multimodality's challenge to marketing theory: A discussion. *Multimodal Communication*, 1(3), 305-323.
- Zantides, E. (Ed.). (2017). *Semiotics and visual communication II: Culture of seduction*. UK: Cambridge Scholars.
- Zeng, F.-B. (2009). 多模态符号整合后语篇意义的嬗变与调控 [Transmutation and control of discourse meaning after the integration of multimodalities]. 《外语教学》 [*Foreign Language Education*], (6), 28-32.
- Zhang, D.-L. (2009). 多模态话语分析综合理论框架探索 [On a synthetic theoretical framework for multimodal discourse analysis]. 《外语教学》 [*Foreign Language Education*], (1), 24-30.
- Zhang, D.-L. (2010). 多模态外语教学的设计与模态调用初探 [Preliminary investigation into the concept of design and the selection of modalities in multimodal foreign language teaching]. 《中国外语》 [*Foreign Languages in China*], 7(3), 48-53.
- Zhang, D.-L., & Huang, L.-H. (Eds.). (2018). 《多模态与外语教育研究》 [*Multimodal and foreign language education research*]. Shanghai: Tongji University Press.

Zhang, Y.-Q. (2017). 《网络科学新闻的多模态语篇研究——共时和历时视角》  
[*A multimodal discourse study of online science news: Synchronic and diachronic perspectives*]. Beijing: Science Press.

Zhu, Y.-S. (2007). 多模态话语分析的理论基础与研究方法 [Theory and methodology of multimodal discourse analysis]. 《外语学刊》 [*Foreign Language Research*], (5), 82-86.

**(Copy editing: Curtis Harrison)**

### **About the author**

Yunzhi Yang (305165131@qq.com), born in 1990, is a Ph.D. candidate at Beijing Language and Culture University. Her main research direction is Foreign Languages and Applied Linguistics.