

Book Reviews



Bateman, John, Janina Wildfeuer, and Tuomo Hiippala (2017). *Multimodality: Foundations, Research and Analysis – A Problem-Oriented Introduction*. Berlin: Mouton de Gruyter. 416 pages.

Multimodality has grown substantially as a method of inquiry and analysis over the last twenty years. Some notable problems with multimodality though are who owns the field and how does one approach a multimodal analysis? These problems are taken up in Bateman, Wildfeuer and Hiippala's (2017) textbook that offers a different approach to multimodality by attempting to remove multimodality from the confines of disciplinary boundaries and paradigms. The book contains eighteen chapters spread over three parts of the text: Part I *Working your way into multimodality* (Chapters 1–4), Part II *Methods and Analysis* (Chapters 5–7), and Part III *Use Cases* (divided into 5 use case areas Chapters 8–18). The idea set forth in this introduction to multimodality is that disciplinary boundaries do not work so well in multimodal research and that good multimodal analysis entails exploring disciplines which have engaged in the analysis of a particular type of text, say film studies, art history, and/or photography, to a name a few. The rhetorical approach taken in the book is scaffolding where one concept is presented in one chapter and built on in the following ones.

In Chapter 1 *Introduction: the challenge of multimodality* the authors set up the problem of multimodality, noting that it is now considered insufficient to focus solely upon individual forms of communicative expression. They introduce the notion of media convergence which illustrates how it is that media themselves are not consumed in isolation but often in the context of other media. They problematize the definition of mode offering a series of quotes (page 18) but do not define mode in this chapter. The chapter also drew attention to problems of defining multimodality, studying multimodal phenomenon, and the problem of identifying objects as multimodal.

The focus of Chapter 2, *Recognizing multimodality: origins and inspirations*, is about defining the problem space of multimodality which offers a short history of a variety of disciplines that have handled multimodality. One of the problems with multimodality, as the authors' see it, is the fact that there are numerous ways that multimodality has been approached for analysis and one must choose, as Jewitt, Bezemer, and O'Halloran (2016) emphasize, which approach to take. The authors draw attention to the scope of what multimodal research is from language to literary studies to human-computer interaction. This leads up to a discussion of the metafunctions derived from systemic-functional linguistics and how some researchers such as Kress and van Leeuwen (1996) and Jewitt et al., (2016) use this as a starting point for research in multimodality. The authors emphasize that this is not the approach they take, noting how it is potentially problematic to use such a distinction on modes other than language. This leads into a discussion of semiotics beginning with Saussure and Peirce discussing the work of both, and arguing that many researchers have gotten Peirce wrong. An important message that the authors mention in this chapter is that multimodality is not a new idea, many authors have noted this (Norris 2012; 2004; Levine et al. 2004), however the authors also discuss how until quite recently there has been an overwhelming preference for monomodality. What is more, while there are numerous approaches to multimodality, the authors believe that no one approach can account for multimodality as a discipline in itself. A graphic is also introduced which shows how speech, sound, image, text and media (the five components the authors identify for multimodality) are interrelated.

Chapter 3 *Where is multimodality? Communicative situations and their media*, attempts to navigate the numerous terms that can be used for multimodality and defining terms used in the approach taken within the book. The problem is that multimodality means different things to different researchers from different paradigms. To solve this problem the authors offer a number of situations which they define as multimodal and here they introduce a term which is used throughout the book, canvas. Canvas is where the communicative situation is taking place, this is not necessarily a physical space but an area or region where semiotic activity is being displayed. Canvases can be 2D or 3D, embodied or not, offer a role, a performance structure, or can be symmetrical or asymmetrical. The last section of the chapter on medium/media introduces a series of terms that make up the typology used to describe communicative situations to enable them to be analyzed multimodally. Canvases can thus be 2D or 3D, static or dynamic, transient or permanent, or offer a role of participant or observer. Other terms of importance are ergodic which refers to whether the reader or participant has work to do within the canvas.

The typology they develop is used like a calculus to determine the multimodal tools that one may use in order to analyze a specific canvas. First one identifies whether a canvas is linear and unchangeable, this would include such canvases as face-to-face interaction recorded on video or just audio. Second is to determine page-flow – readers/viewers must determine how to process the information on the text or depiction. Third has to do with the amount of work that the consumer has to do in order to engage with the text, this is referred to as immutable ergodic, and fourth is mutable ergodic where the user/viewer/reader/player can alter the text that they are engaging. They suggest four orienting questions for multimodal analysis: who are the sign makers? Who are the intended sign consumers? What is the canvas the sign makers are working? What is the time profile?

It is not until Chapter 4, *What is multimodality? Semiotic modes and a new 'textuality'*, where the authors begin to define what a mode is. They cover several terms and concepts, too many to cover completely here. Their point is that many of the concepts that are used to discuss multimodally were derived in linguistics and/or semiotics, which poses a problem for multimodal analysts who work with texts that can be analyzed with concepts outside of linguistics, or work with texts that do not contain language. They then discuss modes and media, where, "a medium is best seen as a historically stabilized site for the deployment and distribution of some selection semiotic modes for the achievement of varied communicative purposes" (page 123). Noting that the semiotic mode has often been handled by multimodal researchers in ways that are both vague and programmatic. The mode then is taken to have both a material form in terms of the canvases that are used for its articulation, and a semiotic form with regards to the meanings that it takes up. Mode is thus rather vaguely defined. The chapter then discusses modes and genre as well as text and textuality, noting that the latter terms have been problematic and are difficult to use in multimodal texts, hence their preference for terms like canvas.

Chapter 5, *The scope and diversity of empirical research methods for multimodality*, covers broadly research methods that can be used to conduct multimodal analysis. The authors discuss everything from differences between quantitative and qualitative research to sampling and sampling methods in addition to protocols and questionnaires. A section is devoted to annotation and the problem of transcription in multimodal data as well as coding, transcoding and content analysis. Another section is devoted to corpus driven multimodal analysis, the issue being that in order to develop a corpus of multimodal data, for example images of signs of linguistic landscape, a lot of work has to be done by the researcher(s) in order to annotate and code the corpus to make it searchable. Another section is devoted to tools/applications such as

Elan which is an open source software used to code and analyze multimodal data. There is also a great discussion on the use of eye-tracking software for multimodal research, noting how Kress and Van Leeuwen's (1996) application of given/new to images has been taken as fact and there is little empirical research to back this up. The final section has to do with computer driven multimodal research including artificial intelligence and machine learning. One of things emphasized has to do with the fact that within multimodal analysis there are many ways to analyze a data set, and a researcher certainly does not have to be familiar with all of them. Rather what the authors advocate is collaboration with researchers from other disciplines who are intimate with the tools for analysis, for example it would be a lot of work to learn how to use eye tracking software and hardware, not to mention the cost of obtaining it. More can be gained from collaborating with researchers who are already intimate with such methods. The chapter concludes with a section on responsible citation of multimodal artifacts in all contexts from giving a lecture to writing a paper.

Chapter 6, *Are your results saying anything? Some basics*, is effectively a crash course in conducting statistical analysis from applying sampling measures to conducting statistical tests such as correlation. The authors do a very good job at introducing statistics and I really like how they emphasize learning how to read statistical analysis. The authors also discuss the fact that one need not know how to conduct statistical analysis by hand, there are many tools out there to assist from excel to Python and researchers should take advantage of them.

Chapter 7 *Multimodal navigator: how to plan your multimodal research*, offers multimodal researchers with one method for approaching a multimodal research project. They offer eight steps (page 230) for implementing an effective multimodal research project: 1) selecting communicative situations, 2) decomposing the communicative situation in terms of canvases and sub canvases, 3) mapping out genre space, 4) selecting modes to analyze from the data, 5) triangulating the research in terms of other work and genres, 6) performing the analysis, 7) finding patterns, and 8) writing up the results. Bateman, Wildfeuer and Hiippala also discuss traps that researchers should try to avoid. The description trap where researchers describe the situation instead of applying analysis to it. The pseudotechnicality trap where researchers overuse technical terms yet do not have the data to back up their use. The circularity (or 20/20 hindsight) trap where the technical description relies too heavily on the analytical results. Lastly, they discuss the importance of triangulating data and results.

The use cases are organized in terms of how a researcher might employ the multimodal navigator to analyze multimodal data. So, the first use case area 1: temporal, unscripted; Temporal refers to the importance of the linearity of time and unscripted refers to the fact that there is no script, things emerge in real-time. Chapter 8 *Gesture and face-to-face interaction*, gives an overview of such research and analysis. Beginning with a discussion of conversation analysis (Sidnell 2010) and turn-taking leading up to a discussion of Scollon's mediated discourse analysis (Scollon 2001), nexus analysis (Scollon and Scollon 2004) and geosemiotics (Scollon and Scollon 2003), which the authors commend for its emphasis on real-time interaction and focusing upon language as just one aspect of the material world. In addition the authors discuss Norris's approach to multimodal discourse analysis (Norris 2004). The authors then discuss a case analysis of interaction, more specifically classroom interaction (Jocuns 2012). Drawing on the work of Kendon (2009; 2004) and McNeill (1992; 2007) and others, they set about developing an analysis of gestures from a classroom interaction. One of the takeaways here again is how often such analyses stick to language and gesture only, when the objects that make up the classroom are equally important. The blackboard and maps that the teacher is gesturing toward as material objects need to be considered as well.

Use case area 2: temporal, scripted, Chapter 9 *Performances and the performing arts*, begins with a discussion of scripted behavior and then launches in to two analyses of performance in the arts, theatre and music. In terms of the former they analyze a performance at The Globe theatre. One of the interesting things here is that the actual text of the play does not always play a part in such multimodal analyses. This took me a bit to figure out, but in effect such analyses focus upon the performance itself not the rehearsals leading up to it, including the blocking and other aspects of staging. The latter part of the chapter focuses upon a "live" performance of the Berlin philharmonic. The scare quotes denote the fact that the live recording analyzed was uploaded to the Berlin Philharmonic's webpage. Such recordings they note are video artefacts which are a new medium and as the authors' note such performances are descriptive mediums.

Use case area 3: spatial, static includes three chapters: Chapter 10 *Layout space*, Chapter 11 *Diagrams and infographics*, and Chapter 12 *Comics and graphic novels*. All three chapters were very interesting and good reads in terms of their analysis. Chapter 10 analyzes a tourist brochure and the key take away here was how layout designers design their products using modular grids that divide up a page. In the analysis they show how the brochure is constructed through various modular grids in the design process arguing

that this structure should be the unit of analysis for this type of document. This preference is based on design as a discipline, which uses such structures in construing documents. In Chapter 11 the diagrammatic mode is the unit of analysis and instructions from Ikea based on Bezemer and Kress' (2016) analysis are used to present a case to analyze diagrams in the first part. Their focus here is in part critique of Bezemer and Kress (2016) who, the authors believe, made too many pre-packaged judgements based on a particular theory. The second part of the chapter analyzes an infographic of the solar powered airplane that navigated the world using Bateman's (2008) rhetorical structure theory which was developed to show how some texts, like infographics, are perceived as coherent whole texts as opposed to segmented sequences. What is emphasized in both analyses is how diagrams and infographics through the diagrammatic mode, support a variety of rhetorical relations through images and texts. Chapter 12 focuses upon comics and graphic novels and the analysis focuses upon a page from a graphic novel. This chapter is very thorough and filled with a lot of interesting information and analysis that a neophyte to comics as a text would find interesting and helpful. They show how comic techniques such as staggering (Cohn 2013), where a border is forced around other panels, has both a narrative and cohesive function. The latter portion of the chapter focuses upon the meta-comic, which is a self-referential and/or reflexive comic. For example McCloud's (1994) discussion of comics is presented as a comic, they also discuss dissertations and academic articles which have been presented as comics.

Use case area 4: spatial, dynamic Chapter 13 *Film and the moving (audio-) visual image*, Chapter 14 *Audiovisual presentations*. Films in whatever form (motion pictures, short videos on YouTube) are very difficult to digest analytically. Here the authors emphasize that researchers who are interested in the multimodal analysis of film should consider what has been done in film analysis already, as opposed to constructing a new analytical tool based on a theory. Page 331 contains a very nice table indicating the variety of areas that one may account for in the analysis of film from camera movement, camera angles, shot scales and others. Using a short excerpt from one of the Diehard films, they show how combining discourse relations between shots and events is one way that films can be analyzed multimodally. Chapter 14 on Audiovisual presentations discusses how one can analyze PowerPoint presentations and the amount of work that has addressed them. One very good point that they make here is that many of these analyses are very decontextualized as not all PowerPoint presentations were meant to be consumed as isolated texts. Rather they were designed to be consumed as a part of lecture or talking presentation of some form. To that end the authors argue that researchers should focus

upon the whole communicative situation within which the PowerPoint is located, not just the PowerPoint text itself.

Use case area 5: spatiotemporal, interactive: ‘media that bite back,’ Chapter 15 *Webpages and dynamic visualizations*, Chapter 16 *Social media*, and Chapter 17 *Computer and video games*. In Chapter 15 the authors introduce the term dynamic data visualization to discuss and analyze webpages. Webpages are dynamic data visualizations to the degree that they change over-time, have varied content and allow for the user to manipulate visualizations in an interactive manner. The example analysis is from a webpage called Connected China, a news article, which is presented in a networked diagram where one clicks on different nodes to read and see how the connection was made. They emphasize that the webpage being analyzed is thus a dynamic diagrammatic mode as different parts of the page are dynamically activated by the user. Chapter 16 on Social media uses Instagram as an example analysis. One of the interesting things analytically for social media is how different apps focus on the users and the different roles that they can take up (passive consumer, active media creator). One of the features of social media that is important to consider is how things such as “likes” and hashtags create ambient affiliation (Zappavigna 2012; 2018) among users. Users feel aligned to other users in virtual communities. Their analysis of Instagram focuses upon the different canvases on the page that can be analyzed from the content image itself, the like, and the comments. The authors also discuss how because all social media have their own Application Programming Interfaces (APIs) that allow researchers to search content, e.g. Twitter hashtags, thus making social media analysis quantitatively friendly as well as searchable via programming languages like Python. Chapter 17’s focus is on computer and video games and the authors offer two analyses of different types of video games: the strategy game Civilization V and a first-person shooter Armed Assault 3. In terms of the strategy game we learn that most of the content is 2D but also hosts a 3D environment. Additionally, it has a variety of canvases that offer analysis of interactivity as well as the manipulation of time. One of the interesting things with such games has to do with user choice in deciding different courses of action. The first person shooter Armed Assault 3 also contains both 2D and 3D environments. One of the points emphasized is how the 3D canvases of such games have a lot of semiotic potential. There are different types of ergodic work done in game within both games. In the strategy game the ergodic work is accomplished in a 2D environment in linear time whereas the ergodic work in Armed Assault 3 is accomplished in real-time.

In Chapter 18 Final words: ready, steady, analyse! The authors summarize their lengthy discussion. One huge take away for me from this chapter was

about collaboration. Multimodal analysis is a paradigm that warrants collaboration and Evaluation.

One criticism of this work is that there is a lot of rhetoric about dangers and weaknesses in much multimodal analyses without offering specifics of the work that is weak or considered weak. The authors go to great lengths throughout the book mentioning such dangers and suggest that pre-packaged multimodal analysis is dangerous and misleading. I found this rhetoric a bit off putting to read. What the authors do in this book is set out to design an alternative approach which they view as the way that multimodal analysis should be conducted. I don't really believe as strongly as they do that one needs to consult what designers or art historians have done in order to conduct a proper multimodal analysis. One can still be theory or method driven using a grounded theory, which is widely accepted in social science. Some other criticisms have to do with the amount of new terminology introduced, which I found frustrating and convoluted at times. One other criticism has to do with how the authors handled the social, or did not fully engage it. For example, in the use case of the first-person shooter they tended to focus on the lone user interacting with the game. Yet many first-person shooters are on-line in multiplayer environments (massively multiplayer on-line games MMOs) where one can also interact and play with friends and strangers; YouTube is filled with numerous video clips from such games. The use cases were the part of the book that I really looked forward too but at times I found the authors hard to follow because of the terminology and constant technical rhetoric. Some of the interesting things they discussed was previous research on the specific cases. For example, the majority of the discussion in Use case area 3 was new to me and the manner in which previous work was articulated as well as the analysis was quite engaging. The audience for this work is quite varied. I believe graduate students who are wishing to embark on a study of multimodality would find many of the chapters presented here helpful. Current multimodal researchers may find some of the rhetoric and terminology frustrating as I did, but I also believe they cover quite a vast array of objects that can be handled multimodally. In contrast many such textbooks focus on face-to-face interaction or images, but here the authors discuss quite a broad range of multimodal research and they should be commended for such a daunting task.

Andrew Jocuns

Department of English & Linguistics, Thammasat University, Rangsit &
Bangkok, Thailand

jocunsa@gmail.com

References

- Bateman, J. 2008. *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents*. 2008 edition. Basingstoke, Hampshire: Palgrave Macmillan.
- Bezemer, Jeff and Gunther Kress. 2016. *Multimodality, Learning and Communication: A Social Semiotic Frame*. 1st edition. London; New York: Routledge.
- Cohn, Neil. 2013. *The Visual Language of Comics: Introduction to the Structure and Cognition of Sequential Images*. London; New York: Bloomsbury Academic.
- Jewitt, Carey, Jeff Bezemer, and O'Halloran Kay L. 2016. *Introducing Multimodality*.
- Jocuns, Andrew. 2012. "Classroom Discourse." In *The Encyclopedia of Applied Linguistics*, edited by Carol A. Chapelle. Oxford, UK: Blackwell Publishing Ltd. <http://doi.wiley.com/10.1002/9781405198431.wbeal0134>.
- Kendon, Adam. 2004. *Gesture: Visible Action as Utterance*. Cambridge; New York: Cambridge University Press.
- Kendon, Adam. 2009. *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Cambridge: Cambridge University Press.
- Kress, Gunther and Theo van Leeuwen. 1996. *Reading Images: The Grammar of Visual Design*. London: Routledge.
- LeVine, Philip and Scollon Ronald 2004. *Discourse and Technology: Multimodal Discourse Analysis*. Washington, D.C: Georgetown Univ. Press.
- McCloud, Scott. 1994. *Understanding Comics: The Invisible Art*. Reprint edition. New York: William Morrow Paperbacks.
- McNeill, David. 1992. *Hand and Mind: What Gestures Reveal about Thought*. 1 edition. Chicago: University of Chicago Press.
- McNeill, David. 2007. *Gesture and Thought*. 1 edition. Chicago, Ill.: University of Chicago Press.
- Norris, Sigrid. 2004. *Analyzing Multimodal Interaction: A Methodological Framework*. New York, NY: Routledge.
- Norris, Sigrid, ed. 2012. *Multimodality in Practice: Investigating Theory-in-Practice-through-Methodology*. Routledge Studies in Multimodality 4. New York: Routledge.
- Scollon, Ron. 2001. *Mediated Discourse: The Nexus of Practice*. 1st edition. London; New York: Routledge.
- Scollon, Ron and Suzanne B. K. Scollon. 2003. *Discourses in Place: Language in the Material World*. London: Routledge.
- Scollon, Ron and Suzanne B. K. Scollon. 2004. *Nexus Analysis: Discourse and the Emerging Internet*. New York: Routledge.
- Sidnell, Jack. 2010. *Conversation Analysis: An Introduction*. 1st edition. Chichester, U.K.; Malden, MA: Wiley-Blackwell.

- Zappavigna, Michele. 2012. *Discourse of Twitter and Social Media: How We Use Language to Create Affiliation on the Web*. New York, N.Y: Continuum International Pub. Group.
- Zappavigna, Michele. 2018. *Searchable Talk: Hashtags and Social Media Metadiscourse*. HPOD edition. London ; New York: Bloomsbury Academic.