

## BOOK REVIEW



### Review of *Research methods for digital discourse analysis*

*Hanwool Choe, University of Hong Kong*

---

#### Research methods for digital discourse analysis

Camilla Vásquez (Ed.)

2022

ISBN: 9781350166837

US \$ 39.95

352 pp.

Bloomsbury Publishing

London, United Kingdom



Digital technology has become a crucial component of social interaction and many discourse analysts specifically focus on how people communicate in online contexts. Up to now, however, far too little attention has been paid to the methodological considerations of digital discourse analysis. Vásquez's (2022) timely edited volume offers a comprehensive account of research methods for analyzing digital discourse on various platforms (e.g., YouTube, Instagram, online dating applications, GoFundMe, and Snapchat) and in several languages (e.g., English, Cantonese, Russian, and Ukrainian). The diverse methodological perspectives explained by discourse analysts in different academic disciplines (e.g., linguistics, communication, and writing studies) and from different regional backgrounds (e.g., Hong Kong, Germany, and the United States) make for a well-rounded collection. Another strength of the book lies in each contributor's extensive reflection on methodological issues, challenges, and concerns that they encountered at the time of their research. The contributors offer hands-on advice for dealing with those challenges as well as specific recommendations and suggestions. By exploring a range of methods as well as trending topics in digital discourse analysis, this book provides excellent guidelines on conducting research from different methodological perspectives.

The book is organized into 13 chapters, with an opening chapter written by Vásquez, the editor of the volume. The first half of the text (Chapters 2 to 7) highlights important methodological challenges that may arise at different phases of the research process, from the application of relevant theory to data sampling and collection. The second half of the book (Chapters 8 to 14) is focused on popular methodological approaches, such as corpus linguistics and ethnography, and it addresses the analysis of multilingual, multimodal, and cross-platform data.

In Chapter 1, Vásquez situates the book in light of increasing interest in digital discourse analysis, as evidenced by its use in recent literature and its selection as a theme of scholarly conferences and journals in the field of linguistics. In doing so, the editor emphasizes the book's niche, which lies in providing a space for discussing discourse analytic research methods, which, Vásquez notes, are not usually included in publications due to space limitations. By pinpointing issues that are unique to digital discourse analysis, this useful introduction sets the stage for the way the book tackles those issues.

Chapters 2 and 3 focus on the use of theory in studying digital discourse analysis. In Chapter 2, Tovaes

reminds us of the role of theory in analyzing digital discourse data. The author explains the formulation of research questions based on personal ‘noticing’ and the way theory can provide insights into the analysis of that noticing. Tovares demonstrates this process by describing her research on the carnivalesque strategy of trolling, or online provocation, wherein she links online reader comments mocking Russian President, Vladimir Putin, to Bakhtin’s (1984a, b) theory of Carnival. Tovares also discusses common mistakes that researchers make when applying theories to the analysis of online data, and ways to avoid them. In describing the application of theory as engaging in conversation with other scholars, the chapter emphasizes that digital discourse analysis should be an integrated work of connecting description and interpretation of data to theory.

In Chapter 3, Kiesling discusses ways to operationalize qualitative theoretical constructs, primarily based on spoken discourse, in order to uncover language patterns in digital discourse data quantitatively. Specifically, the chapter illustrates how the notion of stance (see Du Bois 2007; Jaffe 2009) can be used for computational and corpus-based approaches to analyzing data, thus highlighting the manner in which specific digital platforms can affect the way a certain theoretical concept is operationalized and analyzed. Using a study conducted in collaboration with a computer scientist on Reddit discussion threads, Kiesling walks readers through theory building. Specifically, Kiesling describes how the research team refined previous definitions of stance to make them suitable for annotation, coding each user comment based on the annotation and then characterizing patterns. According to Kiesling, in doing this kind of digital discourse research, it is important to identify and to crystallize aspects of a concept that can be observed and analyzed in online data through iterative testing and adjustments. The chapter also suggests that interdisciplinary collaboration can contribute to the validity of decisions around operationalizing theoretical constructs and the reliability of coding.

Given the vast amount of data available online and the ease of accessing that data, Chapters 4 and 5 hone in on the sampling and collection of relevant digital discourse. In Chapter 4, along with an introduction to several approaches of sampling online discourse (spanning from a corpus of millions of words to a small set of interactions among users), Pihlaja delves into ways to identify, define, and limit an appropriate sample. Using his case study on the interactions of Christians, Muslims, and atheists on YouTube and on podcasts, the author discusses how and why data sampling and collection require attention not only to the framing of a particular online community, but also to the selection of representative platforms and text types. In Chapter 5, Kreis focuses on right-wing discourses and counter discourses on Twitter during the European migrant crisis, elucidating important aspects to consider in collecting, storing, and managing (meta)data. These aspects include whether or not data from a particular platform may be used for research purposes, issues of privacy and anonymity in light of the ephemeral nature of some social media content, and boundaries between public and private contexts (i.e., content that is private and sensitive but publicly available).

In Chapter 6, Tagg and Spilioti turn attention to the ethics of digital discourse analysis, which, according to the authors, needs to be approached case by case. Through the presentation of two case studies—Twitter data with mass media images and WhatsApp messages among close friends—the authors reflect on the ethical challenges that emerged at different stages of their research, from the initial design to the dissemination of textual and visual data. The chapter also includes interviews with other scholars who have dealt with potentially distressing data, such as digital tutorials created by pick-up artists and video interviews with patients and healthcare staff, to share diverse perspectives on ethical practices in doing digital discourse analysis.

Paulus, in Chapter 7, discusses more technical aspects of research methods for digital discourse analysis. The author shares how narrative analysis can be carried out with qualitative data analysis software (QDAS) tools, specifically [ATLAS.ti](#) (Windows version 7). Employing as an example the researcher’s narrative study on GoFundMe, a crowdfunding platform used to raise money, Paulus presents how the ATLAS.ti supports and facilitates the process of extracting, annotating, coding, visualizing, and reporting data. The author also gives practical recommendations and tips regarding how to account for the use of QDAS in the

methods section of a scholarly paper and how to select a proper QDAS package for data analysis.

Chapters 8, 9, and 10 illustrate how technological affordances of social media platforms center around selecting a research method and overcoming methodological challenges. In Chapter 8, Lee investigates multilingual hashtags, or searchable key phrases with a pound sign (#), on Instagram during the 2014 Umbrella Movement in Hong Kong. The author uses a collaborative study to illustrate the sampling and coding of hashtag data. In addition, Lee explains several benefits that researchers can gain from interviewing users (i.e., hashtag posters) to enrich the analysis. In Chapter 9, Page presents ways for dealing with issues arising during the collection of Snapchat data, which consists of pictures, videos, text captions, emoji, and stickers that are only available for twenty-four hours. The author also discusses methodological choices that she and her collaborators make in order to transcribe, annotate, code, and analyze multimodal data. In particular, Page shares the annotation manual that they developed to assist with categorizing, assessing, and rating a set of multimodal data. Then, in Chapter 10, Bahtia defines video blogs (vlogs) as “visual texts” (p.183) that combine several communicative modes. Through a multi-methodological approach to socio-cognition, genre, and language, Bahtia investigates “interdiscursive identity performance” (p.183) wherein a YouTube personality ‘curates’ an authentic- and a professional-self while providing make-up tutorials.

Chapter 11 shifts focus to how digital content can be (re)contextualized by different platforms. Brunner and Diemer theorize on the nature of cross-platform analysis (CPA) through their case study of how companies respond to customer communication depending on the social media platform, thus (re)shaping customer involvement and brand identities. The authors’ discussion instructs readers on how to set up CPA research, which involves both qualitative and quantitative data analyses as well as platform-based interpretations, and explains how the CPA approach optimizes digital marketing strategies.

Chapter 12 introduces the use of corpus linguistics for the analysis of Twitter discourse. Lutzky and Kehoe describe their analysis of tweets addressed to and posted by Ryanair, an Irish airline company. To build their corpus of tweets about Ryanair, the researchers used TAGS (Twitter Archiving Google Sheet), a scraping tool used to export tweets into Google Sheets. Then, by sharing their keyword and collocations analysis (i.e., specific words and collocates characterizing the customer service interactions between the airline and its customers), the authors address important points to keep in mind in terms of how to specify, extract, and compare words and collocates from a corpus. According to Lutzky and Kehoe, researchers need to construct and grow a corpus based on a particular list of ‘seed words’ of interest. Seeing as a tweet is a relatively short text, researchers should also consider using whole tweets, rather than setting fixed word boundaries, in the analysis of right/left collocations.

Both Chapters 13 and 14 highlight ethnographic approaches to digital discourse analysis. First, Bolander (Chapter 13) considers where an ethnography of online practices can take place, from what data it can draw, and which techniques are most amenable to this qualitative method. Through the author’s study on personal blogs, Bolander presents fascinating examples of how to use blended data (online data, such as blog posts and comments, and offline data acquired through interviews) and mixed methodologies (online observations, coding, and questionnaire-based interviews). The chapter emphasizes the ‘polyfocal’ nature of online and offline worlds—which are integrated, rather than separate—and how this should be taken into consideration in digital discourse analysis. In Chapter 14, Thompson shares the experience of interviewing participants for research on online dating. Specifically, the author demonstrates researcher reflexivity through reflective interviewing (spontaneously engaging, depending on the participants’ answers, rather than sticking to script-based interviewing) and dialogic action interviews (observing participants in action while interacting with them). The reflective research process, according to Thompson, allows the author to go beyond a mere interpretation of what is seen in the collected data, while also leading to the increased engagement of both parties. Moreover, this reflective approach makes it possible for a researcher to give back to the community studied, as the ethnographic approach heavily relies on interaction with participants and feedback on their online application usage.

In both the breath of information contained therein and the backgrounds of its contributors, this edited

volume captures the diversity and richness of research methods for digital discourse analysis. Collectively, the chapters explain how to carry out digital discourse analysis and which research methods are available to cover a diverse range of digital platforms, discourse types, and languages. The book provides wide-ranging perspectives on research methods and methodological implications in digital discourse analysis. Another notable feature of the volume is that it does not have to be read in order, as each chapter presents a different case study and focuses on unique methodological decisions and practices. For example, if a reader is interested in mixed methods research, Chapters 3, 7, 9, and 13 would serve as practical guidelines for engaging in similar research. The reader could also zero in on platform-specific research methods (e.g., Chapters 5 and 12 focus on Twitter, and Chapters 8 and 11 on Instagram). In addition, an explanatory picture of how to analyze digital discourse is offered through detailed behind-the-scenes accounts of the contributors' research experiences. These research narratives are well-aligned with the book's goal of "pull[ing] back the curtain" on research (p. 3). An especially unique feature of the book is that each chapter ends by providing practical recommendations, guidelines, suggestions, and tips (sometimes with a glossary) based on the contributors' research activities and experiences. However, methodological considerations in other important and rapidly growing areas of digital discourse analysis, such as human-computer interactions, health communication, and virtual human discourse, are relatively underrepresented. I hope to see the book release future editions to document other research methods for digital discourse analysis. The book should be of interest to novice and experienced scholars, students, and individuals who seek to explore methodological challenges, key issues, and cutting-edge topics in contemporary digital discourse analysis.

## References

- Bakhtin, M. (1984a). *Rabelais and his world*. Indiana University Press.
- Bakhtin, M. (1984b). *Problems of Dostoevsky's poetics*. University of Minnesota Press.
- Du Bois, J. W. (2007). The stance triangle. In R. Englebretson (Ed.), *Stancetaking in discourse* (pp. 139-182). John Benjamins Publishing Company.
- Jaffe, A. (Ed.). (2009). *Stance: Sociolinguistic perspectives*. Oxford University Press.

## About the Author

Hanwool Choe (Ph.D., Georgetown University) is an Assistant Professor in the School of English at the University of Hong Kong. Her research interests include digital discourse, language and food, multimodal interaction, and life stories through the frameworks of interactional sociolinguistics and discourse analysis.

**E-mail:** [hanwool@hku.hk](mailto:hanwool@hku.hk)