

David West Brown and Danielle Zawodny Wetzel, editors. *Corpora and Rhetorically Informed Text Analysis*. John Benjamins, 2023. Pp. 292. Hardbound 99.00€, ebook 99.00 €. ISBN: 9789027213815 (Hardbound).

Open access review under a [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/).

DOI: <https://doi.org/10.24197/ersjes.45.2024.296-300>

Exploring the field of computational rhetoric, or as the authors describe it, rhetorically oriented corpus analysis, this book delves into how computational techniques can be used to study and generate rhetorical strategies and persuasive texts.

The authors explain various applications of a tool called DocuScope, which is freely available as an online application, and also available to download and use in Python scripts via a dedicated library. It provides powerful text analysis capabilities, making it an invaluable resource for researchers and students. Importantly, DocuScope is free to use for non-commercial purposes, providing an accessible entry point for those wishing to explore computational rhetoric and rhetorically-oriented corpus analysis.

The book is not a manual or tutorial on how to use the tool, but rather explains different areas of research where such a tool can be useful and how it can be used to obtain relevant data for the discipline. The fact that DocuScope is freely available for research and that the book covers several areas makes it of interest to a wide audience. This is an edited book that includes contributions from scholars from rhetoric, writing studies, education, data science, political science, and literary and cultural studies. Thus, the book is divided into six parts with two contributions in each part.

The first part of the book, “DocuScope and Computational Rhetoric,” offers, on the one hand, a history of the project including interesting concepts and theories of rhetoric, and on the other hand an introduction to the book itself. Authors here point out the goal of this volume: “both to reflect past applications of DocuScope in a range of historical, literary, educational, professional, and political contexts, as well as to provide space for new and experimental applications” (25).

In Part 2 of the book titled “Variation across Academic Disciplines and Contexts,” the chapters explore genre similarities and differences,

providing valuable insights for scholars seeking methods to describe rhetorical tasks effectively. In the first chapter, “DocuScope, Multi-Dimensional Analysis, and Student Writing,” DeJeu and Brown compare undergraduate and graduate student writing, highlighting the stability of academic discourse across different contexts. In doing this, authors compare tagging systems and patterns of disciplinary variation. In the author’s words, this chapter serves as a methodological case study showing how placements along dimensions can be illustrated using parallel stick plots and how the constituents of dimensions can be compared using radar plots. This chapter also highlights the distinctions between linguistically oriented taggers, which describe syntactic, morphosyntactic, and lexical structures, and rhetorically oriented taggers, capable of analyzing the topical, organizational, and affective features of texts. Beigman’s work, “Narrative Writing from Users in the Wild,” focuses on tools that measure the distance between narrative and argument, emphasizing similarities that inform effective feedback for revision. The primary objective of the chapter is to evaluate the feasibility of incorporating automated feedback features into a writing support tool using genre analysis. Both chapters contribute to our understanding of academic writing by examining genre variations and emphasizing the importance of critical linguistic awareness in the writing process.

Part 3, “Writing Pedagogy, Access and Equity” discusses writing pedagogy, justice, and genre norms. It highlights how genre boundaries are exploited by expert members of discourse communities for private intentions within socially recognized purposes. Laura Aull, in her chapter “Language Patterns in Secondary and Postsecondary Student Writing”, illustrates significant rhetorical differences between secondary and postsecondary writing, which are tied to the varying assignment tasks at each level. She also illustrates how DocuScope language clusters can be utilized to enhance students’ metacognitive awareness of writing practices across these educational stages. Additionally, Aull asserts that generalized claims like “students can’t write” (94). perpetuate a language regulation paradigm centered around error, rather than fostering a language exploration paradigm that emphasizes knowledge and analysis. Oliveri et al. chapter “Understanding Social Justice Features in Statistics Writing” is a case study of DocuScope in undergraduate writing, paying a special attention to language use from the perspective of social justice. Although the limitations of the case study, the results demonstrate the usefulness of

DocuScope to provide both teachers and students with concrete and meaningful information about the writing practice.

In Part 4, “Rhetorically Informed Models of Social Interaction,” the authors provide examples of the contribution of DocuScope to our understanding of how people can position themselves as insiders or outsiders based on their communication styles. William Marcelino, the author of the chapter titled “Public Policy Research Applications of DocuScope’s Linguistic Taxonomy,” emphasizes the value of humanistic knowledge as a necessary complement to technical advancements in Natural Language Processing (NLP) and other data-centric disciplines. In his chapter, Marcelino discusses how DocuScope has been employed at the RAND Corporation to address various public policy challenges. The chapter is divided into two sections. The first section demonstrates how the DocuScope taxonomy serves as a statistical tool for identifying patterns in text corpora. For instance, it has been used to analyze open-text responses in large surveys. By applying this taxonomy, researchers can uncover hidden insights and trends within textual data. In the second section, Marcelino highlights how the same taxonomy has enhanced machine learning algorithms. Specifically, it has improved the detection of conspiracy discourse on social media platforms. This application showcases the practical impact of combining linguistic analysis with computational methods. Importantly, Marcelino suggests that DocuScope’s applicability extends beyond rhetoric theory. He also proposes that insights from argument theory in rhetoric and philosophy could help address semantic distance limitations in Large Language Models (LLMs). Ryan M. Omizo, in his chapter titled “Be Positive,” proposes an approach that combines non-negative matrix factorization (NMF) with DocuScope Language Action Type analysis (LATs) for topic discovery. This methodology aims to extract meaningful topic models from texts while also considering their rhetorical organization. Omizo tags texts with DocuScope LATs and spaCy named entities to discover both the topicality and the formal construction of texts. Although this method may not be ideal for classification purposes, it stands out for its ease of use and interpretability.

In Part 5 of the book, “Professional Writing/Professional Genres,” the authors discuss ongoing DocuScope research and its impact on the development of feedback tools for writers. Geisler’s chapter, “DocuScope Write & Audit as an Early Feedback Machine in Genre-Base Writing,” focuses on proposal writing, a genre of written texts designed to secure a

writer the resources necessary to carry out the proposed project. She uses a prototype of a DocuScope-inspired tool called Write & Audit designed to help students revise their own text by providing early feedback. This tool helps students understand their topical themes within the proposal genre. Dreher, on the other hand, in her chapter “From Technical Reporter to Personal Guide,” explores the differences between academic and plain language abstracts using DocuScope, highlighting how writers transition from technical reporting to a more personal guide in plain language abstracts. The goal of these plain language abstracts is to make scientific research accessible to a wider audience, but most authors know little about plain language, making the resulting genre unstable. The analysis has identified six distinct areas of variation that collectively illustrate the transition in authors’ rhetorical approach from technical reporting to serving as personal inquiry guides. This shift may provide a basis for authors, editors, and publishers to develop more effective guidelines based on practical strategies, as well as areas to explore in future user testing. The chapter highlights two recommendations: First, authors should concentrate on guiding readers through essential big-picture questions, key takeaways, and impacts. Second, there should be a shift in the way information is presented in abstracts—from a straightforward report to a captivating story of inquiry and intrigue in the summaries. Prioritizing the use of first-person constructions, authors can emphasize their own voice and invite readers to join them on their research journey.

In Part 6, “Mining History”, authors explore how specific genres wield influence and motivate action. Genres play a crucial role in facilitating conversation; without recurring structures, comprehending the world becomes challenging. In the chapter by Parry-Giles et.al., titled “Strategic Language as a Family of Identity-Based Discourse Registers,” the authors delve into the intersection of language and political campaigns. They argue that language in political campaigns encompasses a variety of discourse registers, each reflecting distinct strategic roles and outcomes. Specifically, the authors identify four key registers: architect, tactician, advisor, and coach. These registers play a crucial role in shaping campaign strategies. Successful campaigns invariably involve a harmonious integration of these discourse registers. To investigate this phenomenon, the authors employed a combination of methods: DocuScope as a text mining tool and a close reading of texts. Fitzsimmons and Wu, in their chapter “Youngs for the Young,” use DocuScope to explain how in 1882, the librarian Hewins decided what texts should be assigned to girls and

boys, setting a precedent into the next century for gendered classification systems. Authors used the DocuScope dictionary, along with other techniques like topic modelling and AHC (Agglomerative Hierarchical Clustering) to examine the rhetorical rich texts of a corpus of early children's literature.

In summary, this book is an invaluable resource for researchers and students across various fields, as it provides examples of using DocuScope in disciplines such as rhetoric, writing studies, education, data science, and literary and cultural studies. Moreover, the application of DocuScope extends beyond these areas, proving useful in fields related to digital humanities, linguistics, sociolinguistics, psychology, marketing, and business. Its potential uses, combined with its free availability for research and comprehensive user guides and technical documentation, make DocuScope a versatile tool for a wide spectrum of users.

ANTONI OLIVER GONZÁLEZ  
Universitat Oberta de Catalunya (UOC)  
aoliverg@uoc.edu