



palgrave▶pivot

A Discursive Perspective on Wikipedia

More than an Encyclopaedia?

Susanne Kopf

palgrave
macmillan

A Discursive Perspective on Wikipedia

Susanne Kopf

A Discursive Perspective on Wikipedia

More than an Encyclopaedia?

palgrave
macmillan

Susanne Kopf
English Business Communication
Vienna University of Economics
and Business
Vienna, Austria

ISBN 978-3-031-11023-8 ISBN 978-3-031-11024-5 (eBook)
<https://doi.org/10.1007/978-3-031-11024-5>

© The Editor(s) (if applicable) and The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Cover illustration: © John Rawsterne/patternhead.com

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG

The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

*Dedicated to baby Elia, ‘mein Spröttchen’ and beloved
co-author—thank you for allowing me to finish this piece :-)*

CONTENTS

1	Introduction—Contextualisation Matters	1
1.1	<i>Research on Wikipedia</i>	4
1.2	<i>Discourse Studies and the Importance of Context</i>	8
1.2.1	<i>Critical Discourse Studies and Contextualisation</i>	8
1.2.2	<i>CMDA and CDS in the Digital Age</i>	11
1.2.3	<i>The Web 2.0, Social Media and Contextualisation</i>	15
1.2.4	<i>Contextualising Wikipedia and Wikipedia Discourse</i>	17
1.3	<i>Chapter Overview</i>	19
	<i>References</i>	20
2	Wikipedia’s Architecture and Modus Operandi	27
2.1	<i>Wikipedia Site Structure</i>	29
2.1.1	<i>The Article Page</i>	30
2.1.2	<i>The Talk Page</i>	34
2.2	<i>Wikipedia’s Modus Operandi</i>	36
2.2.1	<i>Wikipedia Rules</i>	37
2.2.2	<i>Wikipedia Governance</i>	46
2.2.3	<i>Limiting Editing Activity—Banning, Blocking and Protecting</i>	52
	<i>References</i>	53

3	Wikipedia Talk Page Discourse I	59
3.1	<i>Methods of Analysis</i>	62
3.2	<i>Moves Analysis</i>	64
3.2.1	<i>TP Debate Thread-Initial Postings and Threaded Arguments</i>	64
3.2.2	<i>TP Debate Thread Response Postings and Their Components</i>	76
3.2.3	<i>Posting Components</i>	84
	<i>References</i>	89
4	Wikipedia Talk Page Discourse II	91
4.1	<i>Corpus Linguistic Analysis</i>	92
4.1.1	<i>'I', 'You' and 'We': The Editors and the Community</i>	94
4.1.2	<i>Debating Sources and Wikipedia Policy</i>	99
4.1.3	<i>Debating Article Editing</i>	103
4.2	<i>Wikipedia's Modus Operandi and Wikipedia Discourse</i>	109
	<i>References</i>	113
5	Wikipedia's Position and Function in Society	115
5.1	<i>Wikipedia Consumers and Producers</i>	117
5.1.1	<i>Wikipedia's Reach and Readership</i>	117
5.1.2	<i>Wikipedia's Contributors—'Wikipedians'</i>	121
5.2	<i>Wikipedia—More Than an Encyclopaedia?</i>	124
5.2.1	<i>The Public Sphere</i>	125
5.2.2	<i>Wikipedia's Function as a Transnational Public Sphere</i>	128
	<i>References</i>	137
6	Summary and Outlook	143
	References	147
	Index	163

LIST OF FIGURES

Fig. 1.1	CMDA and CDS integrated	14
Fig. 2.1	Sample article	33
Fig. 2.2	Sample TP thread	36
Fig. 2.3	Wikipedia—the Neutral Point of View policy	37
Fig. 4.1	Signing and timestamping TP postings	93
Fig. 5.1	Percentages of individuals using the internet in 2019 (International Telecommunication Union, 2020, p. 7)	117
Fig. 5.2	Leading language versions of Wikipedia (Anderson et al., 2016)	119
Fig. 5.3	Editor motivations (Wikimedia Foundation, 2011b, p. 13)	123

LIST OF TABLES

Table 1.1	Page et al.'s (2014, p. 11) adaptation of Herring's medium and situation factors (2007)	13
Table 2.1	Conduct policies and guidelines (<i>Wikipedia:Edit Warring</i> , Wikipedia, 2021; <i>Wikipedia:List of Policies and Guidelines</i> , Wikipedia, 2021; <i>Wikipedia:Vandalism</i> , Wikipedia, 2021)	44
Table 2.2	Editor categories and privileges (<i>Wikipedia:Requests for Adminship</i> , Wikipedia, 2021; <i>Wikipedia:Rollback</i> , Wikipedia, 2021; <i>Wikipedia:User Access Levels</i> , Wikipedia, 2021)	50
Table 3.1	The article page (AP) and the talk page (TP) corpus	61
Table 3.2	Three thread-initial posting types	66
Table 3.3	Sample TP thread	71
Table 3.4	Typical posting types—for the sake of completeness, the components, to be discussed in more detail later, are also given (see more in 3.3.3).	77
Table 3.5	The main components	85



Introduction—Contextualisation Matters

Abstract Chapter one sets the scene for this pivot publication. In addition to giving an outline of the monograph’s research objectives and presenting an overview of past research on Wikipedia, this chapter details the theoretical framework motivating the holistic approach to Wikipedia taken in this study. I discuss why it is important to know about Wikipedia’s modus operandi, technological affordances, policies and general site characteristics as well as its position in society and its potential societal functions. In this context, I draw on Herring’s research in Computer-Mediated Discourse Analysis (CMDA) and complement this with a discussion of developments in (Critical) Discourse Studies to underscore the importance of contextualisation and theorisation when examining digital discourse. Finally, this chapter gives a brief outline of the remaining chapters.

Keywords Wikipedia · (Critical) discourse studies (CDS) · Computer-mediated discourse analysis (CMDA) · Social media critical discourse analysis (SM-CDS) · Contextualisation

This pivot publication examines the English version of Wikipedia, the online encyclopaedia, from a discourse-analytical perspective. Overall, it seeks to address the following questions:

1. How does Wikipedia work? That is, specifically, what are key medium-specific features both regarding Wikipedia's overall site architecture and its *modus operandi* (MO)?
2. With a focus on the Wikipedia talk pages, i.e., sites where Wikipedia editors can discuss the encyclopaedic entries, what are key characteristics of Wikipedia discourse? Connected to this, how do the medium-specific features possibly shape this Wikipedia discourse?
3. In terms of reach and readership, what position does Wikipedia hold in a global society and what are its limitations in this regard?
4. What potential function can Wikipedia, especially the Wikipedia discussion sites, fulfil beyond serving as a collaboratively created encyclopaedia designed in accordance with the social media (SM) communicative paradigm? Specifically, and taking into account Wikipedia's technological affordances and *modus operandi*, its discourse material and its position in a global context, can Wikipedia function as a transnational public sphere?

Questions (1) and (2) focus on the encyclopaedia's unique MO, its medium-specific characteristics and how these affect Wikipedia discourse. The former question is addressed in Chapter 2 and the latter in Chapters 3 and 4. In this context, it is worth underscoring that the discourse-analytical examination focuses on Wikipedia *discussion* sites rather than its encyclopaedic articles (see Sect. 2.1 for information on Wikipedia's architecture and Sect. 1.1 for an overview of the existing extensive research dedicated to the encyclopaedic aspect of Wikipedia). Research questions (3) and (4) shed light on Wikipedia's position and possible function in a global society. This is done by, on the one hand, exploring its authors, its reach and readership. On the other hand, I also discuss and assess the site's societal role beyond providing a repository of encyclopaedic information. That is, I focus on if and how Wikipedia may serve as a public sphere that supports transnational debate and collective opinion-formation processes. In answering these four research questions, readers of this pivot publication get a comprehensive overview not only of how Wikipedia functions, its key (linguistic and non-linguistic) characteristics and design, but will also take away a clear understanding of Wikipedia's place in today's online media landscape and its potential to be more than an encyclopaedia. Thus, this publication presents a first foray into a holistic exploration of Wikipedia from a discourse-analytical perspective.

I have been fascinated with and have wanted to explore Wikipedia in the described manner for years and for several reasons. First, Wikipedia is a staple of today's internet and I daresay anybody who is reading these lines has used Wikipedia before. A comparatively 'old' example of the participatory internet—Wikipedia was founded in 2001—its prominence in the online landscape is undeniable and enduring. Wikipedia's clout is reflected in the fact that it consistently ranks among the top 15 most visited websites globally (Alexa, 2021) and it boasts an average of 7.5 billion page visits per month to the English Wikipedia alone (between October 2020 and November 2021) (Wikipedia, 2021). Moreover, the project's popularity seems unlikely to diminish as search engines, such as Google, further boost Wikipedia content and data. For example, on the one hand, Google draws on Wikipedia content, e.g., to fill the info boxes Google displays for many searches. On the other hand, Google has been known to guide its audience to Wikipedia entries in that it prominently features Wikipedia articles in its search results (McMahon et al., 2017). In short, it is hard to imagine the internet without Wikipedia.

Even more than the site's prominence, Wikipedia's structure and its modus operandi are fascinating and thus motivate this pivot publication. Wikipedia is a collaboratively created encyclopaedia and it is unsurprising that it is not only different from other encyclopaedias but also different from other user-created content platforms. Apart from the fact that Wikipedia comes in numerous language versions whose study could (and does) fill books, Wikipedia's collaborative writing processes, the editor hierarchy, the thread structure of Wikipedia discussion sites and the policy construct guiding the encyclopaedic production are but a few of Wikipedia's fascinating aspects. What is more, Wikipedia's potential to fulfil a particular societal function beyond the one it aims for—being a free and comprehensive encyclopaedia—is exciting. As an encyclopaedia, the website obviously functions as a repository of knowledge but since Wikipedia also provides space for discussions among editors, I would argue that it could possibly also serve as a space for transnational debate and opinion-formation processes. For all of these reasons, it is not surprising that the site has inspired a substantial amount of research already.

1.1 RESEARCH ON WIKIPEDIA

Wikipedia has been studied in various fields and from various angles and the following review of existing scholarly engagement with the site can only be very selective and cursory. A body of research, for example, has focused on theoretical examinations of Wikipedia. In this context, the encyclopaedia's governance and its policy-creation processes have yielded particular attention (e.g., Jemielniak, 2016; Kostakis, 2010). Here, descriptions of Wikipedia have reached from Konieczny's 'ad hoc-racy', a term inspired by the Wikipedia community's ability to develop policies ad hoc (2010), to van Dijck's 'democratic bureaucracy', stemming from Wikipedia's sizeable and ever-growing body of rules and set decision-making procedures (2013). In addition, a body of research has centred on Wikipedia editors and Wikipedia's creation process. Here, various studies have homed in on Wikipedia contributors—the so-called 'Wikipedians'—as individuals without confirmed expert status, queried the motivations behind Wikipedians' voluntary work and considered the societal implications of their willingness to put in substantial but free labour (Bruns, 2008; Lund, 2017; Messner & DiStaso, 2013; Rafaeli et al., 2009; Stegbauer, 2009).

Apart from these more theoretically oriented studies, a number of researchers have addressed the collaborative writing and content-creation process and have drawn on Wikipedia data to shed light on, e.g., how contributors establish credibility, negotiate which content to include or exclude from Wikipedia articles and how collective memory is built on the website (Ferron & Massa, 2011; Oxley, 2010). Some work in this context has focused on automated mapping of controversies and processes of consensus building on Wikipedia, for instance, Borra et al.'s (2015) 'Contropedia', a tool that aims to allow the visualisation of edits to articles to identify controversial aspects in Wikipedia articles. Another aspect of Wikipedia's content-creation process that has received research attention focuses on the inclusiveness of the website's peer collaboration. For example, Kittur et al. and Wilson examine whether Wikipedia can harness the 'wisdom of the crowd' but arrive at the conclusion that it is an elite user group that builds the majority of content (Kittur & Kraut, 2008; Kittur et al., 2007; Wilson, 2014). Also examining peer collaboration and decision-making processes through a content analysis of English Wikipedia data, Schneider et al.'s (2012) study is particularly notable as it highlights the prominent role Wikipedia's policies play and

it finds that novice contributors are disadvantaged due to their lack of knowledge of policy (see Sect. 2.2.2 for more on Wikipedia’s editor hierarchy). Studying the Wikipedia community ethnographically, Jemielniak and Reagle cast light on how Wikipedia’s culture of cooperation may alleviate tendencies towards incivility often found in other anonymous online contexts (Reagle, 2010) but also how conflict and authority building are key aspects of Wikipedia’s *modus operandi* (Jemielniak, 2014).

In linguistics and discourse analysis, Wikipedia, Wikipedia data, the encyclopaedia’s genres and different styles have also inspired research. Myers, among a more general discussion of blogs and wikis, discusses Wikipedia, how articles evolve and what shape talk page discussions take. However, while his discussion gives an overview of how Wikipedia functions, Myers’ work merely aims to provide an introduction to the topic (2010, pp. 129–159). Comparably, Schmied (2012), in his discussion of various new media, also touches on Wikipedia as a source of data for linguistic inquiry and addresses the site’s central features briefly. In contrast to this, Mederake (2016) presents an in-depth examination of the features of Wikipedia, but focuses on its German incarnation. Generally, the German Wikipedia has enjoyed ample attention as researchers such as Pentzold (2009a, 2013), Beyersdorff (2011), Storrer and Kallass have examined the German version of Wikipedia in various regards. Storrer, for instance, focuses on hyperlinks, the connection between Wikipedia article pages (AP) and talk pages (TP) and the usage of particular linguistic elements in both AP and TP discourse (2019). Similarly, Kallass (2015) and Pentzold (2009b) examine the writing processes and article-editing processes in the different spaces with a focus on the German Wikipedia. Regarding the English Wikipedia, Emigh and Herring take a corpus-based approach to provide an examination of formality and informality in Wikipedia. Comparing Wikipedia with traditional encyclopaedias, they conclude—similar to Elia (2006)—that Wikipedia articles are as formal in style as traditional encyclopaedias but that TPs are markedly less so (Emigh & Herring, 2005). Focusing on Wikipedia TPs, Gallus and Bhatia observe differences in how male versus female Wikipedians converse (Gallus & Bhatia, 2020) and Sheypak’s examination of several different-language Wikipedias identifies the different politeness strategies employed in different linguistic contexts (Sheypak, 2015).

A substantial body of research on Wikipedia has focused on Wikipedia articles in particular. Aside from a number of scholars’ examinations of the characteristics and genre of Wikipedia articles (see more on this

in Sect. 2.1.1), Wikipedia content, especially its accuracy and quality, the presence and detection of bias and point of view (POV) has inspired research (e.g., Góngora-Goloubintseff, 2020; Swarts, 2009). Already in Wikipedia's earlier days, in 2005, Wikipedia and the Encyclopaedia Britannica are found comparable in accuracy (Giles, 2005). In 2008, another comparison between Wikipedia and other encyclopaedias concludes that then Wikipedia was about 80% accurate whereas other encyclopaedias rated about 95–96% on accuracy (Holman Rector, 2008). Focusing on the representation of political issues, Brown finds that Wikipedia provides accurate information for the most part but does not present comprehensive information when it comes to more obscure or dated issues (Brown, 2011). Similarly, a later study arrives at the conclusion that Wikipedia articles and Encyclopaedia Britannica articles on political issues are comparable for their lack of bias especially when the Wikipedia article has been subject to a high number of revisions (Greenstein & Zhu, 2018). In contrast, studies on medical information on Wikipedia have attested less accuracy to the online encyclopaedia (Matheson & Matheson-Monnet, 2017; Yacob et al., 2020). All in all, however, “it is noteworthy that reliability studies in almost all other domains (including a couple in health) concluded that Wikipedia is generally a reliable source of information” (Mesgari et al., 2015, p. 228).

While some projects have attempted to devise automated methods of bias detection in Wikipedia articles (Al Khatib et al., 2012; Recasens et al., 2013), others have attempted to shed light on bias across different Wikipedia language versions (Greenstein & Zhu, 2014; Halavais & Lackaff, 2008). Callahan and Herring's project (2011), for example, focuses on cultural bias across different-language Wikipedias. Their findings suggest a systemic cultural bias (also see Jemielniak & Wilamowski, 2017). Page (2014) also examines different-language Wikipedias—she takes a diachronic view on one article in the English and the Italian Wikipedia to highlight the existence and trace the creation of counter narratives, i.e., narratives that present alternatives to the dominant version of certain events. By comparison, Gredel (2017) takes into account both talk pages and article pages of different-language Wikipedias when discussing the site's treatment of the annexation of Crimea to Russia. While these examples already present steps towards addressing (the English) Wikipedia and Wikipedia data—especially article page data—in discourse studies, these studies do not fully take into account both

Wikipedia's body of policies and the wider societal context and potential societal relevance of Wikipedia beyond its functioning as an encyclopaedia.

Regarding the latter, O'Sullivan's work (2016) is notable as it already touches on Wikipedia editors as a Community of Practice and how Wikipedia's discussion sites in particular may serve as spaces for transnational debate. Connected to this, several scholars, such as Firer-Blaess and Klemp and Forcehimes, consider Wikipedia's potential to contribute to democratic decision-making (Firer-Blaess, 2011; Klemp & Forcehimes, 2010). However, these studies do not present as comprehensive and holistic an examination of Wikipedia's affordances, the associated discursive practices and their implications for Wikipedia's potential to serve opinion formation and democratic decision-making as presented in this publication. While Lih, already in 2004, presents a discussion of Wikipedia as a site that provides space for citizen journalism (Lih, 2004), Moe's more current empirical analysis of Wikipedia articles is particularly worth mentioning in this context. He examines the Wikipedia entries on 'Islam' across three Scandinavian language versions of Wikipedia in order to illustrate how Wikipedia entries may be interpreted as the end result of different (national) public sphere's treatment of Islam (2019). Rather than viewing Wikipedia as merely allowing an insight into the result of how different public spheres treat certain issues, Hansen et al. home in on Wikipedia's potential to be or become such a public sphere. That is, focusing on one key element associated with functioning public spheres, rational discourse, Hansen et al. study Wikipedia data empirically in order to demonstrate the site's potential to support rational discourse as envisioned by Habermas (2009). Finally and also arguably motivated by considerations regarding democracy and democratic decision-making, Bentzen, on behalf of the European Union's parliament, assesses several online encyclopaedias including Wikipedia, for their allowing citizens easy and free access to relatively reliable information in an era marred by online misinformation (Bentzen, 2018).¹

This brief and by no means comprehensive review of literature on Wikipedia indicates that different academic fields have already begun to study Wikipedia from a variety of angles. The present pivot publication complements existing research by providing a discourse-analytical

¹ See more on Wikipedia's accuracy above.

perspective on Wikipedia connecting the micro, meso and macro level of examination. It aims to forge a bridge from detailed textual analysis to Wikipedia's medium-specific aspects, including its *modus operandi*, structural and technological affordances, to the potential broader societal function possibly associated with Wikipedia.

1.2 DISCOURSE STUDIES AND THE IMPORTANCE OF CONTEXT

The following subsections give an overview of the theoretical background and considerations that this study builds on. Specifically, in addition to defining key concepts and touching on the evolution of (Critical) Discourse Studies in connection with digital discourse, I discuss the importance of context and contextualisation in Discourse Studies. I also outline how and to what degree this study intends to contextualise Wikipedia.

1.2.1 *Critical Discourse Studies and Contextualisation*

In this pivot publication, 'discourse' is understood as language use (together with other forms of semiosis but limited to multimodal elements that can be deployed in digital meaning-making) as a form of social practice. Put in other words, discourse as the manifestation of attitudes, values and norms is understood as a means of creating and shaping the world we live in with all its organisations, power dynamics and social stratifications. At the same time, this discursive social practice and other social practice are viewed as mutually constitutive. That is, non-discursive social practices and realities also shape discourse (Fairclough & Wodak, 1997, p. 259), e.g., Wikipedia's technological design and the site's policies impact Wikipedia discourse (Kopf, 2019b). Therefore, it is important to approach any analysis of discourse material with an understanding of how it is shaped by various factors, such as a digital platform's design.

The definition of discourse underlying this study has notable implications regarding the epistemological value associated with analysing discourse. Among other elements, the 'critical' in Critical Discourse Studies (CDS) refers to the idea that analysing discourse is not an end in itself. Rather, critical discourse analysts examine discourse to shed light on the workings of society—drawing on a variety of analytical methods, they view society through the lens of discourse and try to glean an insight