La Tercera and Wikipedia: the relationship between the news and the editions in the encyclopedia during the Social Outbreak of 2019

Carla Toro Fernández
carlaa.toro@gmail.com

Sebastián Urrejola
seburrejola@gmail.com

Abstract

This ongoing investigation examines the relationship between media discourses in *La Tercera* news and Wikipedia editions during the 2019 Social Outbreak in Chile. By analyzing the content on both platforms, the study aims to explore the impact of the media on constructing a "story" reflected in the encyclopedia and investigate relevant phenomena such as misinformation and news incidence on different platforms. The research also raises questions about media concentration and bias on Wikipedia reporting and their influence on international audiences.

Keywords: Media construction of reality – Wikipedia – Discourse Analysis – Social Outburst – Media discourse

Introduction

New digital media intensifies the speed and reach of news dissemination, but also enables disinformation, which is difficult to detect. Media and information literacy is essential to fight this issue, particularly since filter bubbles create personalized realities for each user (Pariser, 2011). Assessing the quality of information is crucial, particularly during times of crisis when falsehoods rise, and the press verifies information. However, media impartiality is not assured, and those in control may mold reality. In Chile, for example, media ownership is concentrated among elite groups, resulting in an information duopoly between La Tercera and El Mercurio.

The overexposure to news and filter bubbles has led many people to turn to centralized platforms like Wikipedia (Sáez-Trumper, 2019). Wikipedia's use of secondary sources and adherence to policies and rules to maintain a standard of verifiably ensures objectivity in its reporting. However, the predominance of news sources in Wikipedia articles raises questions about whether media concentration and bias affect the platform's reporting. Understanding the interaction between news from media and the content on Wikipedia can provide insight into how stories will be preserved.

The Chilean Social Outbreak (CSO) of 2019 has been the subject of study in various disciplines, particularly in communication, due to the role of the media in the events. The reality reported by the hegemonic media contradicted the collectively imagined reality to such an extent that it sparked societal reactions, such as protests and attacks on media headquarters. Araya & Farías (2022) have identified factors such as the different approaches of the media and the varying discursive constructions used during the events. La Tercera is one of the media sources that is studied and the second most read in the country, so we chose it for this investigation.

The present investigation aims to analyze the media discourses in *La Tercera* news and Wikipedia editions during the 2019 Social Outbreak in Chile. Through the analysis of editions in the platform and media news, we intended to seek the relationship that media had in the construction of the "story" that was reflected in the encyclopedia. This study also allows the exploration of various phenomena that are relevant today, such as misinformation and the incidence of news on different platforms.

Methods

First, we gathered news from La Tercera that contained the word "protest", "government" or "social outbreak", from October 18, 2019, to November 15, 2019. In addition to this, we also obtained data from SeroBOT, an automated bot that reverts vandalism in articles on the Spanish Wikipedia. SeroBOT's contributions provided us with a unique perspective on the CSO.

We constructed a graph with articles that were edited during the social outbreak in Chile. In this graph, each node represents an article from SeroBOT's dataset, while each edge corresponds to a link between articles. Finally for each node we obtained the amount of direct connections to other nodes in the graph, this metric is known as degree centrality. We used this measure to make a list of articles that we could sort to identify the most central nodes in the network.

Then, after manually filtering the sorted list of nodes to identify the articles that were most closely related to Chile or the social outbreak, we proceeded to analyze the

vandalism and other edits made to these articles during the specified period. This allowed us to gain insights into the types of misinformation and propaganda that were being spread through Wikipedia during the CSO, and to identify potential sources of bias or manipulation in the online discourse surrounding these events. Overall, the use of degree centrality provided a valuable tool for identifying the key nodes in our network and focusing our analysis on the most relevant articles for our research question. Ultimately, through the "what links here" feature of the Social Outbreak article, we compiled a corpus of articles pertaining to the aforementioned event. Our focus rested on those articles which contained more than a singular sentence alluding to the CSO, as this approach allowed us to apply the analysis technique that was chosen for this study.

The critical discourse analysis (CDA) technique was utilized to comprehend the discourse of the news, articles, and vandalism. This approach employs a multi-level framework shown in Table 1 to deconstruct the text, thus enabling a thorough understanding of its underlying messages and ideologies. The sociopolitical context studied in the investigation had ideological polarization between power structures and citizenship (Atria & Rovira, 2021), and these are the cases where CDA is applied. This technique allows for the inclusion of a relationship between text and context, to explore how the socially produced ideas and objects (in texts) that inhabit the world (reality) were initially created, and how they are maintained and supported in a context. In this way, thanks to the CDA, it will be possible to give a more in-depth meaning to the news and also to the edits on Wikipedia.

In Table 1, the level of a framework refers to the stages of analysis used to understand the underlying discourse. At the first level, we examine the text to identify the themes and topics being presented, as well as any assumptions or biases that may be present. At the second level, the focus shifts to the language itself, including connotations of specific words and any ambiguity present in the text. The aim of this analysis is to uncover how the discourse is being presented, including its goals and underlying power dynamics. The context of the discourse is also important, as it influences how the discourse is constructed and how it is received by different audiences.

Preliminary Results

The results shown here are part of an ongoing investigation and shed light on the narrative being presented in different data gathered in this study. For the discourse delivered by La Tercera we found that it emphasized on the violence carried out by protesters, the impact on the economy, and the measures adopted by the government for the disturbs. In the case of the Wikipedia vandalisms, there was a clear focus on two opposing viewpoints, with one side depicting the protesters as violent and the other showing the police forces as violent. Our analysis uncovered that a significant portion of vandalisms were intended to undermine the opposing person or institution's ideology, rather than to provide unbiased and accurate information.

Finally, we found that Wikipedia articles primarily focused on the violence by the police and the disturbs by the protesters were on a second plane. Many articles provided information about police repression and human rights violations during the CSO, something that was not found in La Tercera discourse.

Conclusion

By analyzing each data source (La Tercera, Wikipedia articles and vandalisms), we have managed to identify patterns and connections between them and the context in which they operate. Our study has important implications for the use of media as a reliable source on Wikipedia. It emphasizes the need for critical evaluation of the quality and bias of media sources, as well as the potential impact of these sources on the content of Wikipedia articles.

Finally, our preliminary results provided a foundation for further research into the complex interplay between media and Wikipedia. Our next steps are to keep exploring how all this data is related to one another, and to answer the question of how is the relationship between La Tercera and the editions on the Spanish Wikipedia. We hope that this study contributes to a greater understanding of the opportunities and challenges of knowledge creation in the digital age.

References

Araya Seguel, C. & Farías Farías, M. (2022). *Construcción discursiva del "Estallido Social" en tres medios de ciberprensa chilena*. Revista de Linguistica, Filosofía y Literatura. 32(2), 344-364

Atria, J., & Rovira Kaltwasser, C. (2021). *Las elites chilenas y su (des) conexión con la sociedad*. Nueva sociedad, (295), 57-71.

Browne-Sartori, R., & Castillo-Hinojosa, A. M. (2013). *Análisis Crítico del Discurso de la representación intercultural en la prensa chilena*. Convergencia. Revista de Ciencias Sociales, 20(62),45-69

Pariser, E. (2011). *The Filter Bubble*. Van Haren Publishing

Sáez Trumper, D. (2019). *Online desinformation and the role of Wikipedia*. arXiv.

Documentation

1.Plane meaning/text	Thematic level: global meanings
	Level of local meanings: Implicit or indirect Explicit or direct
2. Formal plan / Text-Context	Subtle formal structures Contextual level

Table 1. *Multi-level framework used for the Critical Discourse Analysis. Derivate from Browne & Castillo (2013).*