

## Reading experience of Fernanda Melchor's work: analysis of reviews on Goodreads

Mariana Reyes-Abundes   
Tecnológico de Monterrey, México  
[A01550689@tec.mx](mailto:A01550689@tec.mx)

Manuel Cebral-Loureda   
Tecnológico de Monterrey, México  
[manuel.cebral@tec.mx](mailto:manuel.cebral@tec.mx)

Received: 23/01/2023

Accepted: 27/09/2023

### Abstract

Digital culture influences the reading experience of social network users. This paper analyses the reading experience of hundreds of users of the social network Goodreads by using Natural Language Processing techniques with R programming applied to the book reviews by the Mexican writer Fernanda Melchor. Directly and inversely frequent words were identified, as were emotions linked to the reviews and the most relevant lexemes and verbs, and a classification of topics was performed as well. The results show how readers reflect, in a personal way, on social aspects, such as violence or poverty, giving continuity to the themes addressed by the author herself.

**Keywords:** Digital humanities; reading promotion; literary reviews; 2.0 technologies; text mining; Melchor, Fernanda (1982-).

**How to cite:** Reyes-Abundes, M., & Cebral-Loureda, M. (2024). Reading experience of Fernanda Melchor's work: analysis of Goodreads reviews. *Ocnos*, 23(1). [https://doi.org/10.18239/ocnos\\_2024.23.1.363](https://doi.org/10.18239/ocnos_2024.23.1.363)



## Experiencia de lectura de la obra de Fernanda Melchor: análisis de las reseñas en Goodreads

Mariana Reyes-Abundes   
Tecnológico de Monterrey, México  
[A01550689@tec.mx](mailto:A01550689@tec.mx)

Manuel Cebral-Loureda   
Tecnológico de Monterrey, México  
[manuel.cebral@tec.mx](mailto:manuel.cebral@tec.mx)

Recibido: 23/01/2023

Aceptado: 27/09/2023

### Resumen

La experiencia de lectura de los usuarios de las redes sociales está influenciada por la cultura digital. En este artículo se analiza la experiencia de lectura de cientos de usuarios de la red social Goodreads, mediante el uso de técnicas de Procesamiento de Lenguaje Natural con programación en R, aplicadas a las reseñas de los libros de la escritora mexicana Fernanda Melchor. Se encontraron las palabras directa e inversamente frecuentes, las emociones vinculadas a las reseñas, los lexemas y verbos más relevantes, y se practicó una clasificación de tópicos. Los resultados muestran cómo los lectores reflexionan, de manera personal, sobre aspectos sociales, como la violencia o la pobreza, dando continuidad a las temáticas que aborda la propia autora.

**Palabras clave:** Humanidades digitales; promoción de la lectura; reseñas literarias; tecnologías 2.0; minería de textos; Melchor, Fernanda (1982-).

**Cómo citar:** Reyes-Abundes, M., & Cebral-Loureda, M. (2024). Experiencia de lectura de la obra de Fernanda Melchor: análisis de las reseñas de Goodreads. *Ocnos*, 23(1). [https://doi.org/10.18239/ocnos\\_2024.23.1.363](https://doi.org/10.18239/ocnos_2024.23.1.363)



## Introduction

Reading is an activity that, although regularly performed alone, occurs in a rich social context due to the influence of several people and processes. Books allow several kinds of social interaction: the selection of a book may be motivated by the recommendation of other readers, the promotion done by publishers and authors, the reading process, and the interaction with other readers. Reading mediation comprises the interaction between books and readers to facilitate the dialogue between them (Cerrillo et al., 2002), so to initiate, reading needs both. All these processes involve several people and entities and take place in both physical and virtual environments; however, digital tools have allowed readers to interact directly and more frequently, sometimes among them or with other agents, by using social networks and reading platforms.

The objective of this paper is to analyse the acceptance of literary pieces discussed in digital media and how they create a community that shares reviews, ratings, opinions, and comments on books, broadening the boundaries of the topics in a book depending on how readers internalize them. To this end, the acceptance of the work by the Mexican writer Fernanda Melchor, one of the most awarded and best-selling authors in recent years, will be studied through an analysis that includes the collection of reviews, ratings, and publications shared on Goodreads, a digital platform aimed at promoting reading. Goodreads provides the opportunity to collect a substantial number of reviews as it is one of the most popular reading Apps in the world, with over 90 million users (Barnett, 2021). The objective of examining this large number of reviews is to learn what information about reading practices is being provided by readers and how, by applying text mining, we can obtain results that help to raise questions about the reading experience of hundreds of users.

## Social Reading and Digital Media

Today, the reading experience is mainly influenced by digital culture and platforms; however, it can occur in both virtual and physical spaces. A recent report on reading practices among Mexican women shows that 60% of female readers find out about novelties on social networks (Quiroga, 2020), displacing the traditional bookstore as a direct communication channel. This type of data shows that readers select their books with the help of other readers and that phenomena such as content rating on the Internet respond to a sociocultural reality that assumes that reading is a social and community-based practice. (Alonso-Arévalo & Cordón-García, 2014; García-Roca, 2020). Digital and physical spaces coexist and enrich reading, so it is not difficult to imagine a reader who interacts with both indiscriminately: the reader can read a book review in an online magazine, search for recommendations on social networks, read in physical format or download an article in PDF and read it on the cell phone, share on social networks the latest reading, etc. All these activities are performed daily and are part of reading as a hybrid social activity.

Similarly, there are other spaces that are not exclusively dedicated to promoting reading, where literary reviews, reading challenges, and opinions about authors and books are also shared in audiovisual format. Among them, booktubers (Youtube users) and booktokers (TikTok users) have gained relevance. Known as book influencers, they review content, give their opinions, promote challenges, and, in general, encourage the habit of reading to an audience made up mostly of young and adolescent users of these platforms. Consequently, they emerge as reading mediators before an audience that has the possibility of interacting asynchronously from the functionalities of the platforms and through other social networks (Paladines-Paredes & Aliagas,

2021). Many of these content creators have become new book marketers, recommending texts and writers depending on the taste of their customers, with a powerful impact on reading practices and book circulation (Saez, 2022). Thus, the legitimization model of the cultural field --in which traditional instances are academia, showrooms, and publishing houses (Altamirano & Sarlo, 1990)-- is being challenged, making room for new digital practices of mediation and literary criticism.

In general, social reading platforms offer the possibility of interacting with books directly while providing services focused on promoting reading. On Goodreads, it is possible not only to comment on any book in its database; users can also rate them using a rating mechanism, create their book tags, or view other users' comments and reviews. Consequently, Goodreads offers a reliable and valuable data source to analyse different contemporary reading practices from online reviews, as has been found in a variety of research (Thelwall & Kousha, 2017; Wang, Liu et al., 2019).

## Women's Writing: The Case of Fernanda Melchor

In recent years, books written by women have been published and promoted for reading, challenging the traditionally male literary canon and giving relevance and visibility to women within the publishing circuit (Carroto, 2017; González Arceo, 2017). This publishing phenomenon, which is gradually but significantly striking readers, presents itself as an opportunity for analysis, generating measurable data, especially regarding the acceptance of its books, which is the subject of interest in this paper. However, since talking about women writers is a very broad topic, we decided to choose Fernanda Melchor as a case study. Melchor stands out for the outstanding editorial and literary success of her novel *Temporada de Huracanes* (2017), published in more than twenty languages, with several reprints and a paperback edition (2022). Other books by the author are the chronicle *Aquí no es Miami* (2013), the novels *Falsa Liebre* (2013), and *Páradis* (2021).

This writer and translator from Veracruz also enjoys the support of the media and literary critics and has been awarded prizes such as the *Anna Seghers* in 2019 and the *International Literature Prize* 2019, among others. Her texts have been the subject of research in different academic publications (Lemus, 2022; Godínez- Rivas & Nieto, 2019), and she has a solid base of followers on social networks. Consequently, analysing the acceptance of her work allows us to visualize and study the logic and relationships of a publishing system in transformation (Parnell & Driscoll, 2021), always considering that a literary bestseller is a double-sided cultural product: it quickly reaches a large audience, while gaining literary recognition (Pouly, 2016). Melchor has become one of the Mexican writers with more reviews in Spanish on Goodreads. *Temporada de Huracanes* alone has 2,082 reviews, a higher number than other books by contemporary Mexican women writers such as *La Hija Única* by Guadalupe Nettel with 1,043 reviews, and very similar to *Nuestra Parte de Noche* (2019) by the Argentinian Mariana Enríquez with 2,752.

## Fernanda Melchor on Goodreads

As of 2016, the Goodreads platform does not allow access to its content openly since the API (Application Programming Interface) to the full text of its reviews is not available. Therefore, it was necessary to use the *web scraping* technique to collect the reviews as it allows data to be extracted directly from the web, in particular for the parts of the reviews that can be viewed by the general public. Goodreads enables the visualization of open data, but some is only accessible

when the user is active on the platform. In this case, a Goodreads user account was used to visualize as much data as possible.

While there are more than 3,000 reviews, Goodreads only allows 300 views of each book for each filter -there are three display filters: newest, oldest, and most popular, as well as a language filter-. To access the reviews in a tabular format *Listly* a Chrome extension was used. This helped collect public data to form a database with 2,048 reviews in Spanish -575 correspond to *Páradais*, 1,127 to *Temporada de Huracanes*, 292 to *Aquí no es Miami*, and 54 to *Falsa Liebre*-. The database consists of the following fields: user, date, rating, review, likes, and comments. It is noteworthy that in this study, only the textual analysis was performed using the review field, the users are not mentioned to protect their privacy, and the rating or comments are not taken into account.

**Table 1. Metadata collected on Goodreads as of July 21, 2022 from Fernanda Melchor's books**

	Publication Year	Total Reviews	Total Ratings	Writing %	Rating (1-5)	Average Length of Reviews	Review Languages
<i>Páradais</i>	2021	914	4,738	19.3%	3.75	106	15
<i>Temporada de Huracanes</i>	2017	3,704	17,975	20.6%	4.09	103	20
<i>Aquí no es Miami</i>	2014	312	1,649	18.9%	4.26	84	4
<i>Falsa Liebre</i>	2013	59	271	21.8%	3.89	99.89	1

According to the metadata collected, the author's most popular book is *Temporada de Huracanes*, with more reviews and a significantly higher number of ratings; however, the best-rated of all her books is *Aquí no es Miami*, with an average of 4.26 stars. It is interesting to mention that they all have a very similar percentage of reviews, resulting in approximately 20% of the people who read Melchor's books writing a review on the platform. The average length of the reviews ranges from 83 words for the shortest, to 106 words for the longest, which corresponds to *Páradais*. This very variable data, since there can be reviews of one word and others of more than 1,000 words, is a fact that helps to understand how the writing of reviews per book behaves. Another interesting fact is the languages in which reviews are published: although Spanish has been chosen for the sample collection, it should not be forgotten that some books are translated into more languages, which increases the number of reviews written compared to books that are not translated -the case of *Falsa Liebre*-.

## Methodology

Mixed methods were used in this study. Firstly, some approaches in digital humanities offer methodologies to quantify stories and predict the commercial success of books and authors (Toubia et al., 2021; Wang, Yucesoy et al., 2019), as well as of digital bookstores and electronic literature (Pressman et al., 2015). Complementarily, cultural analytics has been devoted to studying social networks processing large volumes of images, hashtags, and comments, demonstrating that quantitative data-driven methods offer a new framework for describing cultural artifacts, experiences, and dynamics (Manovich, 2020).

This approach to cultural analytics is relevant in this research as a considerable number of comments and reviews collected in reading platforms and networks will be analysed following the logic proposed by Lev Manovich (2020), seeking to represent culture through data and statistical

models. Furthermore, the object of study is also understood from the cultural analytics approach: culture is depicted by how it is produced in the present, from low culture or popular culture, as a source of legitimation rather than on specialized literary criticism or high culture. Finally, unlike traditional methods of literary studies, in this paper, we seek to analyse the entire sample collected online -limited to the case study- without having to choose a few more representative data.

In order to develop this methodology, computational analytics was applied using Natural Language Processing (NLP) techniques in RStudio, an integrated environment for command management of the R programming language, performing the following tasks:

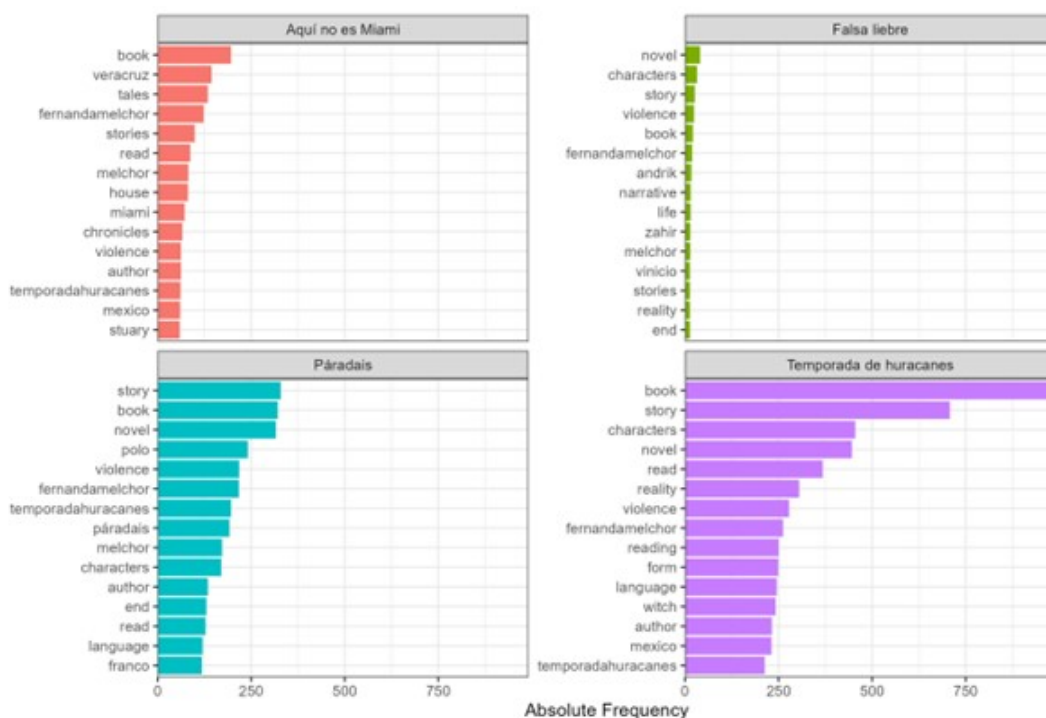
- Direct and inverse frequency analysis of each book: using the *tidytext* package (Silge and Robinson, 2022), the reviews were divided into words, and then *stopwords* were eliminated using the list provided by the *stopwords* package (Benoit et al., 2021). With the remaining words, the most repeated words are counted for each document, in this case, each of the books to which the reviews belong. Complementarily, the term frequency-inverse document frequency, also known as TF-IDF statistic refers to the frequency of a term adjusted to the rarity of its use in the set of documents, measuring the importance of a word in a document (Silge & Robinson, 2017; Sparck-Jones, 1972).
- Sentiment analysis and emotion recognition based on previous lexicons: this is a technique used to automatically classify a text according to a set of words previously designated as positive and negative, as well as others linked to basic emotions. The words have a score for each feeling and emotion so that the frequency of those words in the analysed texts is counted as the increase of those feelings and emotions. In this study, the NRC lexicon (Mohammed et al., 2013) was used. It is noteworthy that being aware of the kind of feelings that are predominant in the reviews does not determine the acceptance or disapproval of the book, as is reflected in the ratings of the books provided by that platform.
- Grammatical tagging: it involves identifying the category (noun, verb, adjective, etc.) to which a word corresponds, its root or lexeme, and its syntactic positioning in a sentence. In this case, the *UDPipe* package (Wijffels, 2022) was used, which has pre-trained artificial intelligence models for several languages -including Spanish- and processes texts delivering the above-mentioned characteristics.
- Bigrams and word networks: by separating words into pairs, the relationship works just like the principle of the Markov chain, that is, it reveals the sequential order in which these words occur in the text, enabling a more detailed understanding of their use. Word networks, presented as nodes that allow the identification of multiple relationships, were created from the bigrams. By observing these networks, topics were formed manually according to how some bigrams come together and form thematic clusters, while others remain more dispersed.
- Topic modelling: this is a *machine-learning* technique that seeks to build topics or themes based on the distribution of words in a set of texts (Becerra & López-Alurralde, 2021). To construct the topics in the sample, the Latent Dirichlet Allocation model (Blei et al., 2003), found in the *topicmodels* package (Grün & Hornik, 2021), was used. This automatic modelling technique was used complementarily to the previous process, which already yielded a first approach to corpus modelling, hoping to obtain a more comprehensive approach. To set up the algorithm, it is necessary to set the value of k, that is, the number of topics to generate. Since this value strongly influences the result, it was calculated using the *perplexity* measure (Blei et al., 2003), obtaining that k=4 is a good configuration value.

The use of digital methods and text mining techniques in Goodreads is not new; there is already research that analyses emotions to identify reactions to books (Driscoll & Rehberg-Sedo, 2019) using tools such as Python and Excel for data extraction and processing (Sánchez-García et al, 2021) or that analyses exploratory and quantitative user participation in Goodreads (García-Roca, 2020). However, this paper proposes a methodology for analysis from different perspectives using different techniques.

## Results

**Figure 1. Absolute Frequency**

Note. Most frequent words in Goodreads reviews of Fernanda Melchor’s novels.

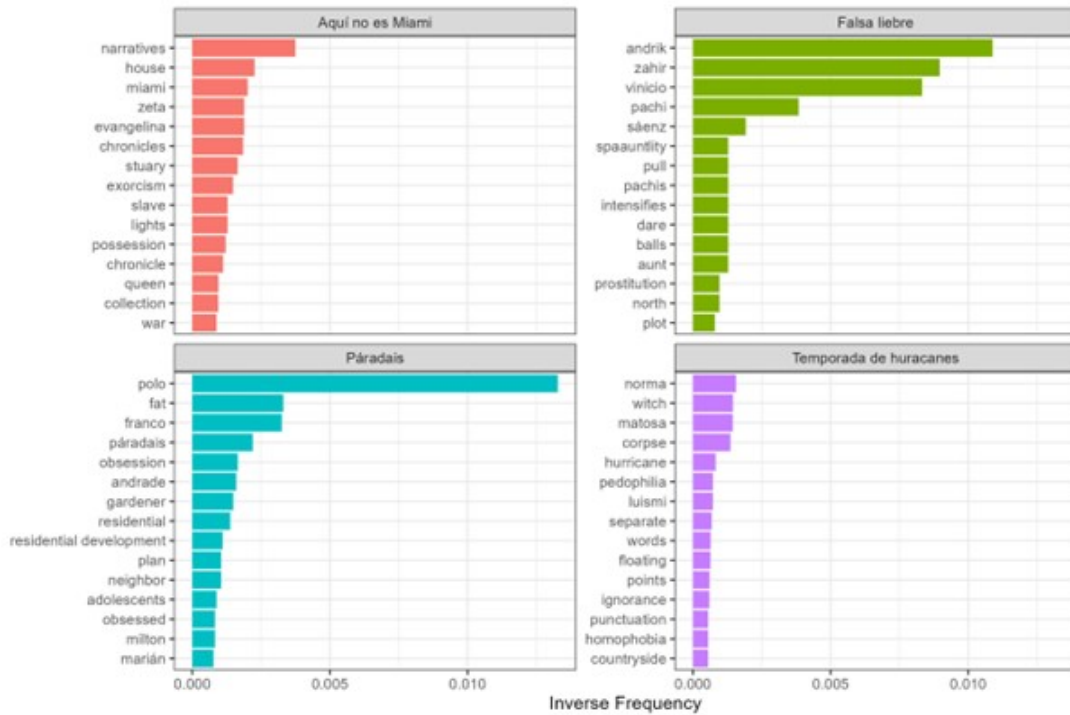


Source: Own development based on R programming.

Concerning the absolute frequency of the most repeated words in the reviews of the four novels, figure 1 shows the comparison with the 15 most frequent words. As we can see, the most repeated words are “book”, “story”, and “novel”, which refer to the description of the book, with a very similar structure in the four sets of reviews. In the case of *Temporada de Huracanes*, “violence” (278), “language” (245), “Mexico” (231), and some characters such as “witch” (242) are also mentioned. In *Páradais*, “violence” (218) and “language” (120) are also cited, which seems to be a constant in Melchor’s books, although “characters” (170) such as “Polo” (241), the “end” (130) of the novel, and references to the “author” (134) also stand out. Similarly, reference to “*Temporada de Huracanes*” is repeated in *Aquí no es Miami* (61) and *Páradais* (196), reflecting that readers have also read *Temporada de Huracanes* and commented on it in the reviews of these books.

**Figure 2. Inverse Frequency (TF-IDF)**

Note. Fifteen most frequent words according to the TF-IDF calculation or inverse document frequency for the selected reviews.

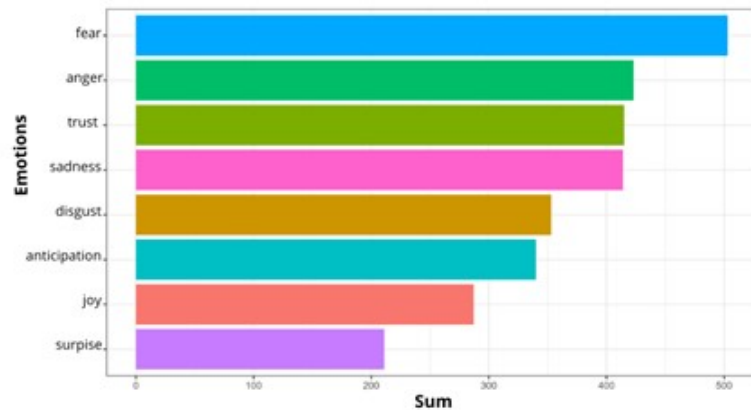


Source: Own development based on R programming.

When making the inverse word frequency comparison, the aim is to measure the importance of a word in the sample by decreasing the commonly used words and increasing those related to the reviews of each book. In this case, the higher the inverse frequency value in the figure, the less that word will appear in the other sets of reviews (per book). Figure 2 reveals that words with higher inverse frequency are associated with characters and places that appear in each book: “norma”, “witch”, “matosa”, “Luismi” in *Temporada de Huracanes*; “Polo”, “Franco”, “residential” in *Páradais*; “Evangelina,” “Miami”, “estuary” in *Aquí no es Miami*; “Andrik” “Zahir”, “Vinicio” in *Falsa Liebre*; and other words such as “narratives” and “chronicles” in *Aquí no es Miami*, which unlike the other three books is not a novel. This is to be expected since the reviews are about the characters, the places, and the elements of the book that are significant for each person; however, it is more interesting to consider how they also mention words that reflect the reader’s interpretation of the not so explicit themes of the novels. In the case of *Temporada de Huracanes*, they talk about “homophobia”, “pedophilia”, and “ignorance”, which denotes a deeper interpretation of the book, referring to important themes that underlie the novel. These readers contribute with their literary and personal reflection and interpretation to the Goodreads community, stimulating what Pierre Lévy (1997) would call a collective intelligence, which helps users to delve deeper into the text in a way that is very different from the individual (García-Roca, 2016).



Figure 3. Emotion Recognition



Source: Own development based on R programming.

Figure 3 shows the counting of emotions in all Melchor’s book reviews. Fear stands out, followed by anger, confidence, sadness, and disgust, followed by expectation, joy, and surprise. The words that most stand out for the emotion of fear are “violence”, “poverty”, “misery”, “doubt”, and “witch”, which refer not so much to the perception of the books but to their subject matter. In the case of anger, the results are very similar since “violence”, “witch”, and “murder” contribute to the score for both emotions. When reviewing the words that relate to trust, “real”, “Franco” (character in a novel), and “strong” are the ones with the highest frequency. Finally, it is noteworthy that the structure of emotions is very similar in the reviews of all the books, with words such as “violence” and “murder” predominating, very well defining the prevailing plot of Melchor’s novels.

Figure 4. Word Cloud of Feelings

Note. Positive feelings -in blue- and negative feelings -in red- in Fernanda Melchor’s books reviews.

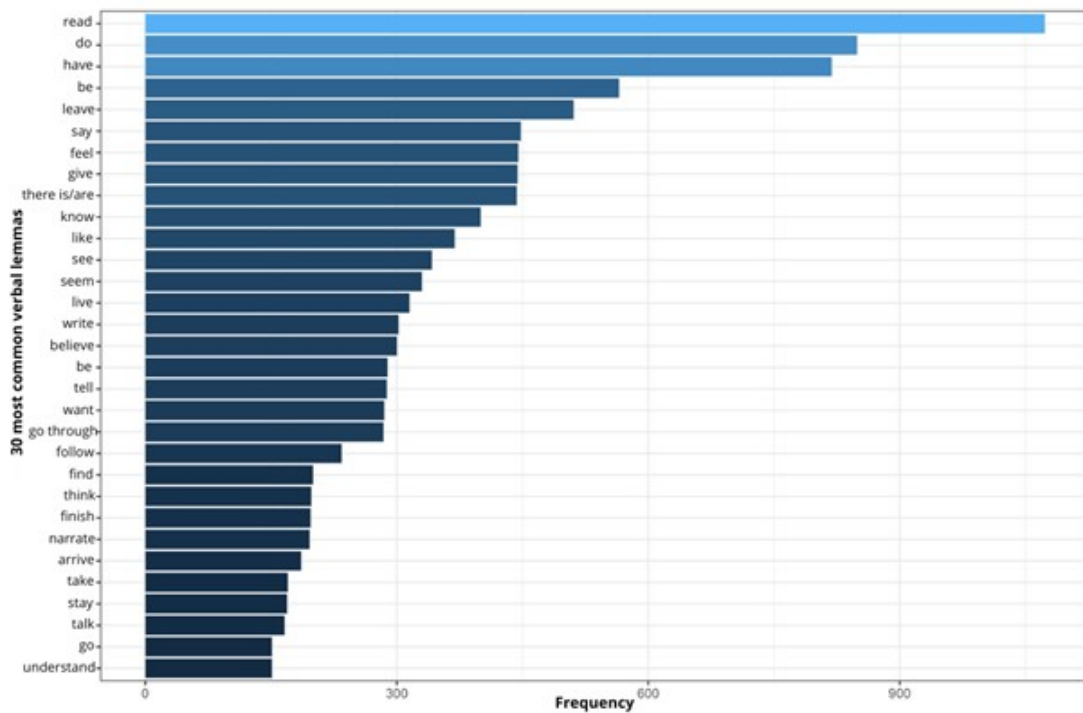


Source: Own development based on R programming.

To visualize the load of feelings, it has been decided to use a word cloud, unlike the previous information shown in tables, because the impact of the words can be better appreciated as a set

of meanings than hierarchically, with the largest ones being the most frequent, the blue ones corresponding to positive feelings and the red ones to negative ones. As can be seen in figure 4, the load of negative feelings is quite strongly motivated by the presence of terms such as “violence”, “poverty”, “misery”, “fear”, “difficult”, “cruel”, and “crude”. These words are linked to content and problematics in the author’s books. For example, when talking about violence and poverty or, more precisely, in *Temporada de Huracanes*, when referring to the character of the witch, there is no doubt that when talking about this topic, the comments will have a negative charge. However, terms with positive connotations also appear in the analysis. Interestingly, they are those that are not so much related to the contents but rather to the reading experience, finding words such as: “form”, “style”, “reader”, “rhythm”, “sensation”, “reading”, “interesting”, “telling”, “perspective”.

Figure 5. Most Frequent Mottos



Source: Own development based on R programming.

In this section, we analyse the main grammatical entities used in the reviews, where nouns stand out - remember that the most frequent word in the sample is “book”, punctuation marks, determiners, prepositions, verbs, pronouns, adjectives, and adverbs. We decided to delve deeper into the verbs used to understand the content of the reviews concerning the readers’ behaviour: what reactions and actions does reading provoke in the readers?

For this purpose, the lexical roots of the most frequent verbs in the reviews were taken and graphed in figure 5 from highest to lowest. The verb “read” is in first place, showing that the reviews are a reflection of the reading experience of each of the users. Following, verbs such as “feel”, “like”, and “know” are found, similarly reflecting a personal perception shared by the users in the reviews. In the case of the verbs .think”, “show”, “understand”, and “find”, it is possible that they refer to what the book means to each reader and how receptivity can transcend to other aspects of their life. The writing of the reviews emphasizes what books make readers feel,

reflecting a personal relationship with the read text, mirroring each reader's own experience. Hence, the most frequent verbs are those commonly used to describe impressions of a book, talk about previous knowledge or knowledge acquired through reading, or readers' expectations in a rich exchange of personal and metacognitive experiences.

## Manual and Automatic Topic Modelling

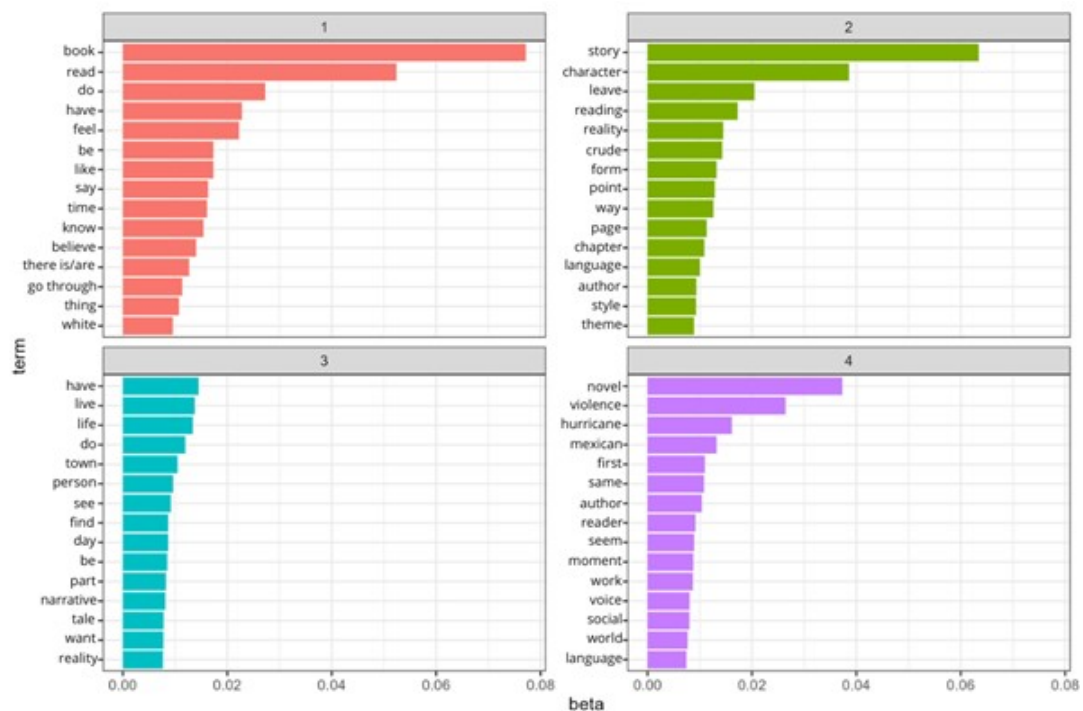
**Table 2. Most Relevant Features Formed by Word Pairs**

Features	Bigrams or Pairs Of Words that Form It
<i>Literary and Book-writing</i>	"narrative structure", "literary style", "narrative voices", "endless paragraphs", "long chapters", "main character", "language: vulgar, popular, coarse, colloquial, crude, strong", "rhythm: fast-paced, vertiginous", "long sentences", "masterful way", "Latin American Literature".
<i>Book Description</i>	"book: unpleasant, hard, short, crude, easy, written, excellent", "bad book", "novel: brutal, short", "story: revolves, told, narrated, real", "author: describes, uses, accomplishes", "suffocating atmosphere", "first pages", "Mexican idioms".
<i>Reading Experience</i>	"bitter taste", "bad vibes", "mixed feelings", "will continue reading", "wanted to continue", "want to read", "read season", "recommend, enjoyed, liked: very much", "could stop", "can feel", "could quit", "impossible to drop", "read fast", "let breathe", "3, 4, 5: stars", "would have liked", "favourite story".
<i>Social Topics</i>	"violence, abuse: sexual", "so much: misfortune, violence", "sexual abuse", "misery, condition, decadence: human", "organized crime", "real facts", "classes, problems, differences: social", "conservative, classist", "class: high, low", "inequality, context, criticism: social", "drug prostitution", "violent Mexico", "modern Veracruz".
<i>External References</i>	"Gabriel García Márquez", "Daniela Aedo", "Random House", "Pedro Páramo", "Vargas Llosa", "Mariana Enríquez", "Mel Gibson"

Note: This table is the result of the breakdown of Fernanda Melchor's book reviews into bigrams.

The most frequent word pairs or bigrams help better understand the main topics discussed in the reviews, providing broader lexical chains than isolated terms. In table 2, these bigrams were grouped in order to identify, manually, these general aspects that make up the reviews. First, there are the literary and writing aspects of the books, which refer to the form, voice, style, narrative, language, and rhythm of the books, as well as to the characters, paragraphs, and chapters, and in general to Latin American literature. The description of the reading experience is another frequent topic, expressed through words such as "unpleasant", "brutal", "raw", and "easy", as well as through direct reference to the author and the way she writes, the atmosphere she creates, and the use of idioms. The reading experience encompasses all those elements that have to do with the reader's acceptance and interpretation of the books: what the reader thought ("I enjoyed it a lot"); how the reader felt ("bitter taste"); how the reader read it ("impossible to put it down"); the reader's expectations ("I would have liked"); or the reader's conclusions ("I highly recommend it"). Another aspect that stands out in Melchor's books, according to the reviews, is the social issues addressed: violence, abuse, crime, poverty, and inequality. These elements are perceived by readers as highly relevant, reflecting part of "everyday life" and "violent Mexico". Finally, we were able to identify some external references to writers and novels that readers relate to Melchor's books, such as "García Márquez" and "Pedro Páramo".

Figure 6. Topic Modelling



Source: Own development based on R programming.

The result of applying topic modelling automatically using the LDA algorithm was a set of topics to which a number has been assigned and does not reflect any particular meaning or priority (figure 6). Nevertheless, these results were reviewed to assign a suitable label to each topic. For example, in the case of topic 1, words such as “read”, “feel”, “like”, “know”, and “believe”, refer to the reading experience that each user describes when reading a particular book. In topic 2, “story”, “crude”, “form”, “chapter”, “language”, and “author” stand out, all of which refer to the description and literary structure of each work. The third topic is more varied, as the words “town”, “life”, “narrative”, “reality”, and “find” may refer to the readers’ perception of the subject matter of the books. The fourth topic, with words such as “violence”, “Mexican”, “moment”, “voice”, “social”, “world”, and “language”, makes direct reference to the themes and social aspects that each book deals with. Up to this point, the results from the topic modelling are similar to those found manually in the bigrams because, although they are not the same, they highlight the same points: reading experience, literary aspects, and themes of each book (social, violence, poverty). The only significant difference would be that in the topic modelling, there is no mention of external references to other books or writers, which can be considered an important element for qualitative analysis, complementing automatic and computational analysis.

## Conclusions

The different techniques used for the computational analysis of Fernanda Melchor’s book reviews proved helpful when applied to a corpus extracted from a socio-digital platform such as Goodreads using natural language processing techniques. Due to the closeness of the reviews to the discussed books, they contain many literary elements, such as places, adjectives, or characters, combined with the readers’ experiences, feelings, and emotions.

The frequency analysis has been the first approach for the quantitative study of the corpus, providing insight into the main elements of the reviews. Even if the researcher has not read it, he or she can get a general idea of what it is about. In this case, the most repeated words are: “book”, “novel”, “read”, or “story”, which refer to general aspects of reading. Complementarily, the inverse frequency highlights components specific to each book reviewed, emphasizing their particularity concerning the corpus of which they are part, highlighting themes and discussions specific to each story. It also shows the dynamics of this social network, which generates different readings due to the multiplicity of users that integrate it, so that “the reception process does not end with the hypotext, but expands as much as the reader and the community wish” (García-Roca, 2016, p. 49).

From the recognition of feelings and emotions, it is possible to identify the main words readers associate with each book. It is important to note that, although the reviews of Melchor’s books are categorized as “negative” or associated with “fear”, they are not related to the reader’s perception or liking for that book, but rather to its themes, which describe aspects of violence, poverty, and social inequality. This could be deduced by analysing the prominent words of each emotion and being familiar with the content of the novels. It could be said that these books do not end after the reading, but transcend in the readers’ understanding of issues such as violence, poverty, and inequality, and they communicate this in their reviews. Therefore, one of the functions of literature is manifested to awaken and enliven social consciousness: “There where the reader’s literary experience enters the horizon of expectations of their life practice, it shapes their understanding of the world and thus impacts their forms of social behavior” (Jauss, 2013, p. 200).

Grammatical tagging reveals how reviews are written and what their syntax is. In this case, nouns and adjectives are quite important, since most users provide a descriptive review of the books and different literary and reading aspects. The analysis of the most used verbs served to understand the reading experience of each user, their acceptance of the book, and what it contributes to them. It can be concluded that the writing of reviews is a performative act, which assumes that the review will be read, thus constituting a display of social identity. (Driscoll & Rehberg-Sedo, 2019). Verbs also show the metacognitive richness present in these spaces for experience exchange since “the act of reading multiplies in the personalized or anonymous collaboration of individual opinions and queries” (García-Canclini, 2015, p. 21). Reading is social and collaborative in many senses since it involves several actors, influencing the acceptance of books.

Two methods were used for topic labelling: separation by bigrams to find clusters (manual) and modelling by topics (automatic). In both, similar results were obtained, sharing three main topics: reading experience, literary aspects, and the social theme of each book. The questioning of these topics shows how the virtual space in which the reviews are shared is, at the same time, a place of social communication and personal expression, where reading is disseminated and commented on in an intimate but also critical fashion so that “The act of reading is transformed into social conversation and common exchange” (Lluch, 2014). That is particularly true in the case of Fernanda Melchor, given the markedly social character of her books.

This study was conducted on an author that we find of great interest, nevertheless, there are other authors to whom the same study could be applied, allowing us to verify how the results differ while maintaining the methodology. Furthermore, an in-depth analysis of the topics of the reviews and the readers’ perception could be complemented by more qualitative and ethnographic techniques, such as interviews or focus groups, which can be done in future research, deepening the scope of the quantitative analysis presented here.

## Funding

Invitation *Challenge Based Research Funding Program*. 2022\_Project ID: E061-EHE-GI02-D-T3-E.

## References

- Alonso-Arévalo, J., & Córdón-García, J. A. (2014). Lectura social, metadatos y visibilidad de la información. *XLV Jornadas Mexicanas de Bibliotecología*. <http://hdl.handle.net/10760/23095>
- Altamirano, C., & Sarlo, B. (1990). *Conceptos de sociología literaria*. Centro Editor de América Latina.
- Barnett, C. (2021, noviembre 9). Who owns Goodreads? *Book Riot*. <https://bookriot.com/who-owns-goodreads/>
- Becerra, G., & López-Alurralde, J. P. (2021). *Introducción a la ciencia de datos con R. Una guía rápida para científicos sociales*. Bookdown. [https://bookdown.org/gaston\\_becerra/curso-intro-r/](https://bookdown.org/gaston_becerra/curso-intro-r/)
- Benoit, K., Muhr, D., & Watanabe, K. (2021). *Stopwords: Multilingual stopword lists (2.3)*. <https://cran.r-project.org/web/packages/stopwords/stopwords.pdf>
- Blei, D. M., Ng, A., & Michael I, J. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Carroto, P. (2017, agosto 14). El otro 'boom' latinoamericano es femenino. *El País*. [https://elpais.com/cultura/2017/08/13/actualidad/1502641791\\_807871.html](https://elpais.com/cultura/2017/08/13/actualidad/1502641791_807871.html)
- Cerrillo, P. C., Larrañaga, E., & Yubero, S. (2002). *Libros, lectores y mediadores: La formación de los hábitos lectores como proceso de aprendizaje*. Ediciones de la Universidad de Castilla-La Mancha.
- Driscoll, B., & Rehberg-Sedo, D. (2019). Faraway, So Close: Seeing the intimacy in Goodreads reviews. *Qualitative Inquiry*, 25(3), 248-259. <https://doi.org/10.1177/1077800418801375>
- García-Canclini, N. (2015). Leer en papel y en pantallas: El giro Antropológico. In N. García-Canclini, V. Gerber-Bicecci, A. López-Ojeda, E. Nivón-Bolán, C. Pérez-Camacho, C. Pinochet-Cobos, & R. Winocur-Iparraguirre, *Hacia una antropología de los lectores* (pp. 1-37). D - Ediciones Culturales Paidós.
- García-Roca, A. (2016). Reading Practices Affinity Spaces: Participatory Forms of Digital Culture. *Ocnos*, 15(1), 42-51. [https://doi.org/10.18239/ocnos\\_2016.15.1.979](https://doi.org/10.18239/ocnos_2016.15.1.979)
- García-Roca, A. (2020). Lectura virtualmente digital: El reto colectivo de interpretación textual. *Cinta de Moebio*, 67, 65-74. <https://doi.org/10.4067/s0717-554x2020000100065>
- Godínez-Rivas, G. L., & Nieto, L. R. (2019). Queers and bewitched: Temporada de huracanes of Fernanda Melchor. *Anclajes*, 23(3), 59-70. <https://doi.org/10.19137/anclajes-2019-2335>
- González Arceo, L. (2017, agosto 19). Libros B4 Tipos: Las chicas que están revolucionando la forma de leer autoras. *Malvestida*. <https://malvestida.com/2017/08/libros-b4-tipos-las-chicas-estan-revolucionando-la-forma-leer-autoras/>
- Grün, B., & Hornik, K. (2021). *Package «Topicmodels». Topicmodels: Topic Models*. <https://cran.r-project.org/web/packages/topicmodels/index.html>
- Jauss, H. R. (2013). *La historia de la literatura como provocación*. Editorial Gredos.
- Lemus, R. (2022). En los márgenes de la nación: Geografías imaginarias en Temporada de huracanes (2017) de Fernanda Melchor. *Bulletin of Hispanic Studies*, 99(2), 163-170. <https://doi.org/10.3828/bhs.2022.12>
- Lévy, P. (1997). *L'intelligence collective: Pour une anthropologie du cyberspace*. La Découverte.

- Lluch, G. (2014). Jóvenes y adolescentes hablan de lectura en la red. *Ocnos*, 11, 7-20. [https://doi.org/10.18239/ocnos\\_2014.11.01](https://doi.org/10.18239/ocnos_2014.11.01)
- Manovich, L. (2020). *Cultural Analytics*. MIT Press.
- Mohammad, S. M., Kiritchenko, S., & Zhu, X. (2013). NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets (arXiv:1308.6242). In *Proceedings of the seventh international workshop on Semantic Evaluation Exercises*, June 2013. arXiv. <http://arxiv.org/abs/1308.6242>
- Paladines-Paredes, L., & Aliagas, C. (2021). Booktubers' video reviews as literary mediation spaces. *Ocnos*, 20(1), 38-49. [https://doi.org/10.18239/ocnos\\_2021.20.1.2489](https://doi.org/10.18239/ocnos_2021.20.1.2489)
- Parnell, C., & Driscoll, B. (2021). Institutions, platforms and the production of debut success in contemporary book culture. *Media International Australia*, 1329878X2110361. <https://doi.org/10.1177/1329878X211036192>
- Pouly, M.-P. (2016). Playing both sides of the field: The anatomy of a 'quality' bestseller. *Poetics*, 59, 20-34. <https://doi.org/10.1016/j.poetic.2016.02.002>
- Pressman, J., Marino, M. C., & Douglass, J. (2015). *Reading project: A collaborative analysis of William Poundstone's Project for Tachistoscope (Bottomless Pit)*. University of Iowa Press. <https://doi.org/10.2307/j.ctt20p598m>
- Quiroga, R. (2020). Lectoras mexicanas, el mercado más grande en español. *El Economista*. <https://www.economista.com.mx/arteseideas/Lectoras-mexicanas-el-mercado-mas-grande-en-espanol-20200722-0054.html>
- Saez, V. (2022). De lectores a "influencers". Booktubers, bookstagrammers y booktokers y la circulación de la literatura en redes sociales en Argentina. *Revista Pilquen – sección Ciencias Sociales*, 25(2), 020-046. <https://www.redalyc.org/journal/3475/347572703002/html/>
- Sánchez-García, P., Hernández-Ortega, J., & Rovira-Collado, J. (2021). Reading the social reader: evolution of Spanish children's and young adult literature in Goodreads. *Ocnos*, 20(1), 7-22. [https://doi.org/10.18239/ocnos\\_2021.20.1.2446](https://doi.org/10.18239/ocnos_2021.20.1.2446)
- Silge, J., & Robinson, D. (2017). *Text mining with R: A tidy approach*. O'Reilly.
- Silge, J., & Robinson, D. (2022). *Package 'tidytext'*. *Tidytext: Text Mining using 'dplyr', 'ggplot2', and Other Tidy Tools*. <https://cran.r-project.org/web/packages/tidytext/index.html>
- Sparck-Jones, K. (1972). A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, 28(1), 11-21. <https://doi.org/10.1108/eb026526>
- Thelwall, M., & Kousha, K. (2017). Goodreads: A social network site for book readers. *Journal of the Association for Information Science and Technology*, 68(4), 972-983. <https://doi.org/10.1002/asi.23733>
- Toubia, O., Berger, J., & Eliashberg, J. (2021). How quantifying the shape of stories predicts their success. *Proceedings of the National Academy of Sciences of the United States of America*, 118(26). <https://doi.org/10.1073/pnas.2011695118>
- Wang, K., Liu, X., & Han, Y. (2019). Exploring Goodreads reviews for book impact assessment. *Journal of Informetrics*, 13(3), 874-886. <https://doi.org/10.1016/j.joi.2019.07.003>
- Wang, X., Yucesoy, B., Varol, O., Eliassi-Rad, T., & Barabási, A.-L. (2019). Success in books: Predicting book sales before publication. *EPJ Data Science*, 8(1), 31. <https://doi.org/10.1140/epjds/s13688-019-0208-6>
- Wijffels, J. (2022). *Package 'udpipe'*. *Udpipe: Tokenization, Parts of Speech Tagging, Lemmatization and Dependency Parsing with the 'UDPipe' «NLP» Toolkit*. <https://cran.r-project.org/web/packages/udpipe/udpipe.pdf>