

# Travelling in Style: phases in the writing of Virginia Woolf

Shaping the story through the narrative use of locations &  
contextualizing the novel in the evolution of modernist writing style

---

Leiden University, Netherlands

Faculty of Humanities

Leiden Center of Digital Humanities

Minor Digital Humanities

Study year 2020 - 2021 semester 2

5170V08

Digital Detectives: Text Mining in Practice

Dr. J. Prokic

Aurora E. L. Turkenburg

Laura Semlič

Sophie Granger

## **Table of contents**

<b>Introduction</b>	<b>2</b>
<b>Literature Review</b>	<b>4</b>
<b>Data</b>	<b>6</b>
Source Files	6
Image Assets	7
Results Reuse	7
<b>Methods</b>	<b>7</b>
Tools	8
Writing Style	9
Narration	12
Maps	13
<b>Results</b>	<b>15</b>
Stylometry	15
Network Analysis	20
Geographical Analysis	23
Sentiments	24
<b>Discussion</b>	<b>28</b>
<b>Conclusion</b>	<b>31</b>
<b>Bibliography</b>	<b>34</b>
<b>Appendix</b>	<b>36</b>

# Introduction

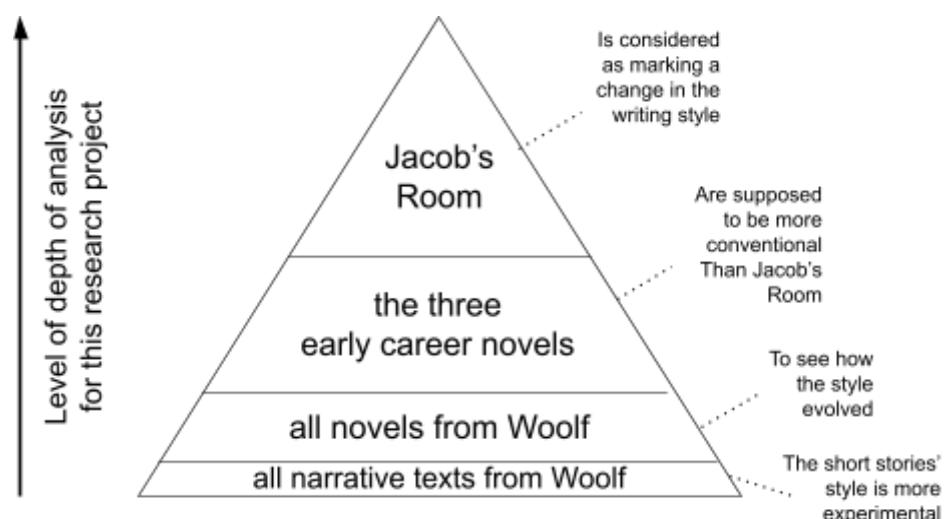
In 1919, Virginia Woolf published an essay titled “Modern Fiction” in which she critiques the longstanding tradition of realist writing, arguing that novels written in this tradition fail to capture life as humans experience it. She calls for abandoning the conventions of realist writing, which, she argues, is usually more plot driven and “symmetrically arranged” (Woolf 2152), and instead suggests that novels should “record the atoms as they fall upon the mind in the order in which they fall [and] trace the pattern, however disconnected and incoherent in appearance” (Woolf 2152). In other words, “Modern Fiction” provides Woolf’s definition and understanding of what were to become characteristic features associated not only with her writing, but with Modernism in general: stream of consciousness and free indirect discourse. Taking a comparative approach, this project examines the changes in Woolf’s style of writing across her oeuvre, with a particular focus on her first three novels: *The Voyage Out* (1913), *Night and Day* (1919), and *Jacob’s Room* (1922).

Since these two terms play a part in our project, it is useful to first briefly explain them. Free indirect discourse is a narratological device which enables authors to mediate characters’ speech in a way which makes it difficult to distinguish between the voice of the narrator and the character. In that sense, “the narrator disappears” (Klarer 30). Closely related to this device, is modernist dialogism which Hammond et al. define as the “tendency of modernist writers to include mutually differentially and often ideologically opposed voices in their works” (Hammond 52), which raises the question of which character is speaking or thinking certain thoughts. Since free indirect discourse, unlike direct discourse, is commonly not enclosed within quotation marks, determining who speaks or thinks can be particularly difficult (Abbott 237). Another frequently employed technique in modernist writing is the stream of consciousness. This device is used to convey the subconscious thought patterns of literary characters (Klarer 192).

Upon completing *Jacob’s Room* in 1922, Woolf remarked that with it she “found out how to begin (at 40) to say something in [her] own voice” (qtd. in Raitt 29). That *Jacob’s Room* represents a turning point in Woolf’s writing, and a transition towards high

modernism, seems to have been largely accepted by scholars (Goldman 49; Raitt 29). In light of recent studies discussed below which have used methods in digital humanities to detect narrative techniques characteristic of modernism, such as free indirect discourse, in Woolf's writing, this project examines the changes in Woolf's style across her oeuvre, with a particular focus on her first three novels: *The Voyage Out* (1913), *Night and Day* (1919), and *Jacob's Room* (1922). While the project offers a global perspective on changes taking places in Woolf's oeuvre, particular attention is paid to first three novels since we are particularly interested in whether, using the methods we have studied during Digital Detectives, we could detect a stylistic change in *Jacob's Room* in comparison to *The Voyage Out* and *Night and Day*. Detecting such a change would confirm Woolf's, as well as scholars' claim that with her third novel she has really found her voice. Additionally, the project looks more closely at the internal structure of these three novels by examining the progression of storytelling with regard to sentiments. Since *The Voyage Out* and *Jacob's Room* are, at their core, stories about a journey, this project also looks at the geographical progression of the stories.

*Fig. 1. Level of depth of analysis for this research project.*



## Literature Review

Considerable scholarly attention, both in traditional literary studies and within digital humanities, has been given to Woolf's writing style. For instance, Violeta Sotirova has studied modernist dialogism in her analysis of four excerpts from *To the Lighthouse*. She has identified various linguistic features which delineate one character's consciousness from another, such as sentence-initial connectives (e.g. and, for; and: so), repetition, and similar lexical choices (12-15). Woolf's modernist dialogism has also been taken up by Hammond et al. who, working within digital humanities, used an annotated version of *To the Lighthouse* to train a computational stylistics model to detect instances of free indirect discourse in an untagged text. While this proved difficult, the authors were more successful in training the model to "generate stylistic profiles for narration, spoken direct discourse, silent (thought) direct discourse, and FID" (Hammond et al. 67), and to determine stylistic profiles of characters. This project resulted in a website, *The Brown Stocking*, which allows users to compare the student annotated version of the novel with the machine learned instances of FID. On a bigger range of texts, the project *Hearing Silent Woolf*, led by Brandon Walsh, worked on a supervised classification model to detect internal and external speech. In addition to this, an analysis on the use of quotation marks was done. Yaxiao Cui focuses on a slightly different aspect of Woolf's style: rather than examining where the shifts in characters' consciousness occur, Cui questions which "linguistic mechanism ... facilitates" (175) this shift. Cui concludes that it is the use of parentheticals in *Mrs Dalloway* which enable Woolf to present "multiple points of view within a short stretch of text" (175), and therefore evoke a sense of simultaneity.

Not only Woolf's experimental style, but her work in general, has been of interest to scholars within digital humanities. In the past fifteen years, two digitisation projects have been carried out: *Woolf Online*, and *Modernist Archives Publishing Project (MAPP)*. While the first focuses specifically on *To the Lighthouse*, MAPP is a digital archive of various textual resources relating to the Woolfs and their publishing firm, The Hogarth Press.

Alongside her writing style, Woolf is also known for her interest in urban environments (see: Evans and Cornish 2017). There are several examples of mapping projects in digital humanities which take Woolf's work as its subject. Google's Art and Culture platform in collaboration with the National Museum of Women in the Arts made, for important locations about her life and novel, a thread of interactive 360 degrees street view photos. In the context of a digital humanities course about Woolf, Amanda Golden blogs about the subtleties in the use of locations in the novel. Their resulting maps focus on the roads characters took or on mapped objects that are present in the story. A project which has certainly been an inspiration for our own geographical analysis of the novels is the Authorial London project which includes roughly 193 works by writers who have lived in and written about London. The purpose of the project is to map the London locations found in these works onto a map of the city, with each location including a reference to all the works in which it is mentioned. While Authorial London does include Woolf's works, it does not allow the user to compare the locations of Woolf's several novels on a map simultaneously. The map from The Londonist contains locations around the whole world from Woolf's ten novels, and while each book has a designated colour, due to the amount of markers, it is difficult to reach conclusions. Our project aims to fill this gap by plotting a map of frequent locations from *The Voyage Out*, *Night and Day*, and *Jacob's Room* onto a single map. Such a map enables the users to observe where the locations of these three novels intersect or deviate.

## Data

The data used in this project is a corpus which consists of 13 texts written by Virginia Woolf. Ten of these are novels, two are collections of short stories and one is a short story. The corpus provides a good overview of Woolf's oeuvre as it extends from her first publication, *The Voyage Out*, first published in 1915, to her last novel, *Between the Acts*, published posthumously in 1941. Such a broad range of work is particularly useful for the examination of stylistic changes in her writing.

All novels were obtained from Project Gutenberg, six of which from its general - and seven from the Australian domain. The latter domain was used as these novels were not available through the general website due to copyright issues, and a larger corpus of novels was desired for this project. *Table 1* shows the titles of all texts, including the information about the encoding, where the text was obtained, and the year of first publication.

While this project focuses on all 13 of these texts, special attention is paid to *The Voyage Out*, *Night and Day*, and *Jacob's Room*. There are several posthumous publications of short stories from Woolf, these weren't taken into account.

### Source Files

*Table 1. Filename, text, encoding, source and year of first publication of the works by Virginia Woolf examined in this paper.*

Filename .txt	text	encoding	source	Year of first publication
0100101	To the Lighthouse	ASCII	gutenberg.net.au	1927
0200331	Orlando	ASCII	gutenberg.net.au	1928
0200991	Mrs. Dalloway	ISO-8859	gutenberg.net.au	1925
0201091	The Waves	ISO-8859	gutenberg.net.au	1931
0301041	Flush A Biography	ISO-8859	gutenberg.net.au	1933
0301171	Between the Acts	ISO-8859	gutenberg.net.au	1941

0301221	The Years	ISO-8859	gutenberg.net.au	1937
1245-0	Night and Day	UTF-8	gutenberg.org	1919
144-0	The Voyage Out	UTF-8	gutenberg.org	1915
5670	Jacob's Room	ASCII	gutenberg.org	1922
pg29220	Monday or Tuesday	UTF-8	gutenberg.org	1921
pg63107	Mrs Dalloway in Bond Street	UTF-8	gutenberg.org	1923
pg63230	Two Stories (The Mark on the Wall)	UTF-8	gutenberg.org	1917

## Image Assets

The Pictorial Plan of London, 1923, Geographia Ltd, 19 x 29 inches, digitized by Geographicus:

[https://commons.wikimedia.org/wiki/File:1923\\_Geographica\\_Pocket\\_Map\\_of\\_London,\\_England\\_-\\_Geographicus\\_-\\_London-geographica-1923.jpg](https://commons.wikimedia.org/wiki/File:1923_Geographica_Pocket_Map_of_London,_England_-_Geographicus_-_London-geographica-1923.jpg)

## Results Reuse

Resulting from the text mining, several csv files with data about locations have been made with data such as coordinates, sentiments and chapters. These were used for map making and network analysis.

## Methods

In order to examine the changes in Woolf's style of writing, and determine whether *Jacob's Room* is the turning point in Woolf's transition to a more modernist style of writing, we made use of a set of methods from the digital humanities. A brief description of the methods, as well as the way in which they were used to answer the research questions is explained in more detail in the remainder of this chapter. Attention will be given to the writing style, locations and named entities contained within Virginia Woolf's novels *Jacob's Room* (1922), *The Voyage Out* (1915), and *Night and Day* (1919).

### Tools

A total of 11 tools were used for this analysis, most of them using the programming language python:

- The *Natural Language Toolkit* or NLTK (explained in more detail below)
- *SpaCy* is a library for natural language processing that can automatically detect named entities in text, such as locations, personal names, and timestamps. This project used *SpaCy* to identify both locations and personal names that were subsequently visualized.
- *VADER* is a sentiment analysis tool. This project used *VADER* for sentiment analysis of whole novels, as well as sentences in which locations are mentioned in *Jacob's Room*.
- *Pandas* is a library for numerical data-analysis. This project used *Pandas* for complex data calculations needed for the purpose of writing-style analysis.
- *Numpy* is a library for mathematical computing. This project used *Numpy* for complex data calculations needed for writing-style analysis.

- *Textstat* is a statistics python library that contains multiple functions to assess the readability of text. The formulas have been developed by and for evaluating children's literature, army instruction booklets and marketing copywriting.
- *Scikit-learn* is a machine learning library created specifically for python. This project used *Scikit* for clustering tri-grams in a writing-style analysis.
- *Matplotlib* is a data visualization tool that uses code to display a myriad of plots, including barplots, scatterplots and wordclouds. This project used *Matplotlib* to visualize part of the data calculated using *Pandas*, *Numpy*, *Textstat* and *Scikit*.
- *Folium* is a data visualization tool that specializes in the design and display of intricate maps. This project used *Folium* to visualize the locations identified by *SpaCy*.
- *NetworkX* is a graph calculation and visualization tool, used here to connect locations between them.
- *Weebly* is a website creation tool, used in this project to visualize the interactive story map and store unused visualizations.

As discussed above, the source data for this project was a corpus of 13 texts by Virginia Woolf. Their origins can be seen in table 1 below. After obtaining all texts from Project Gutenberg in .txt format, individual files were preprocessed since they contained unwanted text, such as introductions and metadata. For this step, we made use of The *Natural Language Toolkit* or NLTK, which is a platform of libraries created for statistical natural language processing. Once the files were preprocessed, they were stored in variables for use by other tools.

## Writing Style

In order to address how Woolf's writing style has changed over time, we compared all thirteen texts across various metrics. We devised a function which calculated and returned the following metrics:

- The date of publication
- Number of sentences
- Percent of quote pairs
- Percent of quote pairs with apostrophes
- Number of sentences starting with quotation marks
- Ratio of sentences starting with quotation marks
- Number of characters (letters and symbols)
- Ratio of characters per sentence
- Average length of sentences in words
- Longest sentence in characters
- Number of special characters
- Ratio of special characters by total number of special characters
- Ratio of special characters by sentences
- Percentage of each special characters by total number of special characters

In order to study the change in Woolf's style from her more conventionally designed plots to those where she experiments with the use of stream of consciousness and free indirect discourse, we wrote a function which would help us study the presence of dialogue. While scholars such as Brandon Walsh, Adam Hammond and Julia Brooke have studied in a computational way Woolf's use of free indirect discourse, their methods appeared too advanced for our level of computational skill. Instead of focusing on replication or developing a function which would detect free indirect discourse, we focused on a function which would help us detect the presence of dialogue. This function detects the presence of quotation marks that indicate dialogue for each sentence. This method, however, has a disadvantage: it detects only clearly indicated direct dialogue while leaving out subtle reported and narrated dialogues. However when the narrator is a diegetic character, the indirect dialogues can be interpreted as being thoughts. Another limitation of this function is that it does not detect direct dialogue placed later on in the sentences. To solve this problem, we measured the frequency of quotation pairs by characters. In a revised version of this function, we added a component which also counts the apostrophes since these are used to indicate dialogues in *Orlando*. This measure is less precise since the apostrophes are also used grammatically, and the result therefore does not reflect the quantity of dialogue in the text. However, assuming the proportion of grammatical apostrophes is similar in all the texts, it can still be used to compare it with the other texts.

To get a better grasp of how these metrics have evolved throughout time, we presented our findings in chronologically ordered bar plots, using Matplotlib. While we visualised these metrics for all data, we also made visualisations for smaller subsets: the three early novels that are often compared (*The Voyage Out*, *Night and Day*, *Jacob's Room*), all the novels and all the short stories. These subsets of our corpus vary a lot from each other, but taking a closer look at each makes sense in a comparative analysis. The short stories vary the most in their format, and can be considered as outliers. However they are also texts where Woolf could experiment the most with the innovative and unorthodox writing styles.

As Woolf's later writings try to imitate the way humans think, sentences get unconventional structures. To try to characterise this, the textstat python library, which contains multiple functions to assess the readability of texts, was used. The metrics mentioned above were all plotted in a graph for each set of texts. Although *Textstat* was initially developed for evaluating children's literature, army instruction booklets and marketing copywriting, and was not designed with literary texts in mind, we assume some changes in Woolf's way of writing will be detected. More specifically, we expect that those sentences which imitate the structure of thoughts will be more difficult to read.

A method more often used to compare literary texts, although for authorship attribution, is the stylometric analysis of n-grams. However, this method enables us to compare the similarity in writing style between texts from the same author and look for the sentence-initial connectives and repetitions that are typical to characterize a character's thinking. With this technique we can look at recurrent patterns of words and syntactic structures by extracting the part of speech labels with *NLTK*. We calculated the trigram distance matrix. With the help of the heatmap function from *Matplotlib*, we visualised the results in a colored graph matrix. The results were clustered in groups determined by the elbow method with a *YellowBrick* visualization and visualized in a dendrogram tree.

## Narration

To study the narration, two aspects were taken into consideration: the use of geographical references and the repartition of strong emotional sentences.

Determining the sentiment of individual sentences enables us to detect the high and low points in the tone of the writing, as well as study any pattern that may emerge. To determine the sentiment of sentences, each sentence was plotted in a stripe colored by the sentiment score provided by the *VADER* module. The visualisation of results was inspired by the famous climate change temperature stripe. However, we made a simpler graph. To make the results more readable, the text was divided in a number of parts and the average sentiment score of the sentences in each part was plotted. For the set of the three early career novels the pattern of sentiments was annotated and an interpretation was given.

To see how the sentiments were reflected in the writing style, we looked at the correlation of sentiment strength and sentence length. Using Matplotlib, these were plotted in graphs and scatter plots.

The sentiments associated with the corresponding ten most recurrent locations were also plotted in a graph, in the order of appearance in the text. Although not ordered chronologically in respect to the timeline of the story, it still gives a rough idea, as the order of events is often something done consciously when crafting the story. Instead of looking at locations in a temporal way, the geographical one may be more suited.

To get an idea when locations are mentioned in the novel, different visualizations have been made with *Rawgraphs.io*. A matrix plot and streamgraph were made. Additionally, an arc diagram was plotted. The parameters were configured in a way to minimise the overlap with the links. Locations are then ordered in a way to be near the chapters with which they are the most related. The resulting order of locations is not based on the appearance of locations in chronological order, but instead depends on the importance of the location in regards to the plot. With this we can extract Jacob's Voyage.

## Maps

Using the OpenStreetMap database, coordinates for locations that appeared more than once were requested. Subsequently, these coordinates were formatted in a csv file containing their frequency and sentiment score that could be read in a Pandas dataframe for fast plotting with the Leaflet based Folium library. With a plugin and a list containing the locations by chapter, an animated and interactive heatmap was made, showing where the story took place and which locations were mentioned in each chapter.

To see how locations were associated with each other and with which chapters, a network analysis was made. Another type of visualisation is a plot with dots on a line for each location in fifteen most recurrent ones, representing when they appear in the text.

Leaving the temporal aspect aside, and looking at the emotional aspect, each location was represented on a map with a transparent circle depending on their average sentiment score: yellow when the sentiment could not be calculated, and two values of green and red according to the strength of positive or negative emotion. Hovering the cursor above a circle provides the name of the location and the score. When clicking on the location, a panel shows up with the sentences where the location appears and the sentiment associated with it. This allows the viewer to see how a location is described in the novels, but also to see how the sentiment was calculated. This also shows sometimes that the location on the map and in the story aren't always the same. For example, the continent 'America' is marked on the map in the Netherlands, where a homonymous village is

To be able to compare the use of locations in different novels, a map using the same type of circles was made, although the color now represents the corresponding novel. A box was added to control which novels are shown and in which order they are superposed. Playing with the fact that the circles are transparent, the overlap is made clear. An additional map was made to show Woolf's London. Locations in London were filtered out with a function looking at the coordinates and if they were in a rectangular

range that represented the city. A map issued by WikiMedia from the same time period was used as background.

These maps allow us to look at the novels in a distant way, in the sense of distant computational reading and satellite perspective on the world map. Wanting to go closer to the story in the novel, an interactive map was made with the StoryMap tool by Knightlab. Images and other sources are presented in an interactive way.

Another way to look at locations is to use graph analysis. This method is often used to study networks between people, but also can be used with location instead to plan routes. For our project, we will look at locations as if they were characters, not taking geographical aspects into account, and using chapters as 'middlemen'. This allows us to look for cultural links between locations that are established in the novel. To do this degree and centrality was calculated to find node communities.

# Results

While authors across the world are renowned for their specific, personal writing styles, it is style that shows a difference across the ages in Virginia Woolf's writing. This chapter will discuss the results of the computational analysis of the works of Virginia Woolf discussed in the previous chapter. Attention will be given to stylometry, network analysis and finally a geographical analysis. In the following chapter, a conclusion will be drawn from the results detailed below.

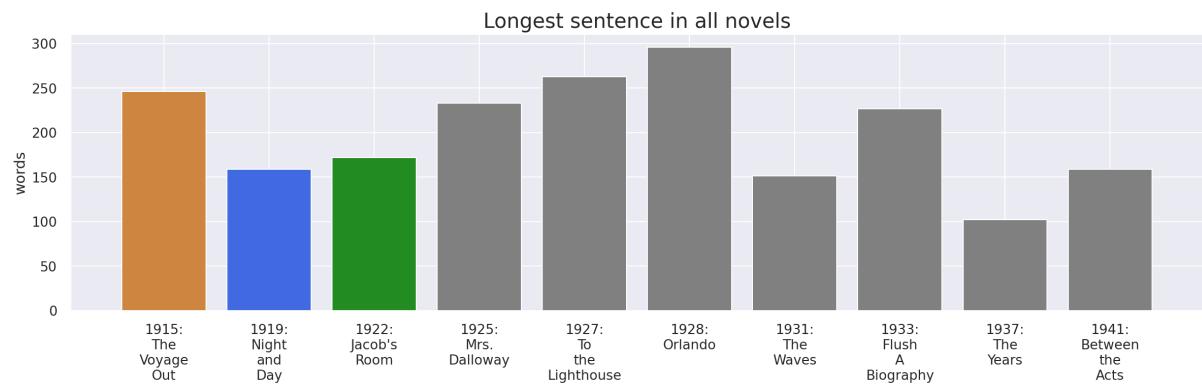
It is important to note that the images used in this chapter present only a smaller part of the full amount of visualizations produced. Unused images for further reference can be found on [thedigitaldetectivesjacobsroomproject.weebly.com](http://thedigitaldetectivesjacobsroomproject.weebly.com) and in the annex. Additional images display analyses of sentence length, special character use, complexity, geographical analysis and sentiment analysis of several of Virginia Woolf's novels.

## Stylometry

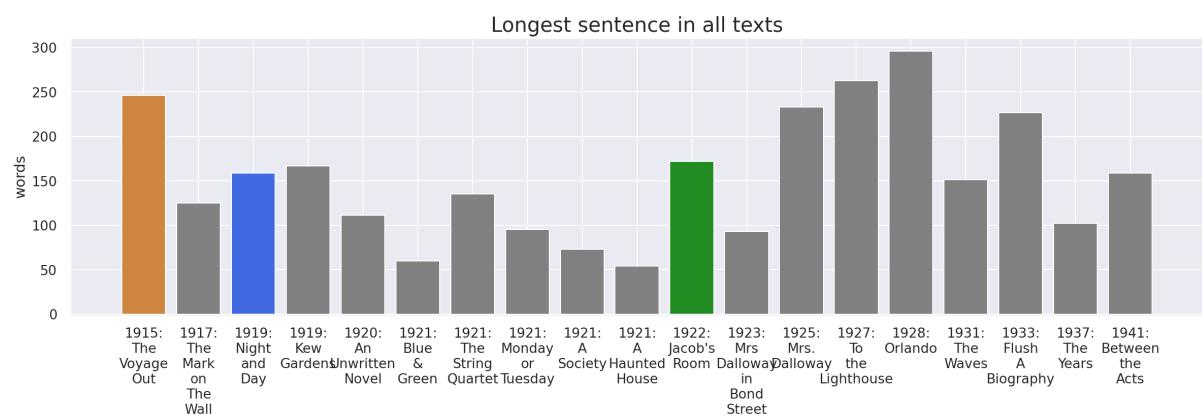
As discussed in the previous chapter, a stylometry analysis of the works of Virginia Woolf was conducted on the basis of sentence length, appearance of special characters, levels of dialogue, readability and complexity, and finally a cluster analysis of all works.

As can be seen in *Figure 2*, the longest sentences in all of Woolf's novels analysed for this project are in *Orlando* (roughly 300), *To the Lighthouse* (just over 250), *The Voyage Out*, and *Mrs Dalloway* respectively. Out of the three early career novels, the longest sentence is in Woolf's first novel, *The Voyage Out*, as shown in *Figure 2*. These findings are the same when taking Woolf's entire corpus in consideration, as shown in *Figure 3*. The average sentence length in novels continued to grow until 1928, culminating in *Orlando* with an average of 26.979 words per sentence. After this, the average sentence length sharply declines, stabilising at around 10-15 words per sentence (see *Figure 4*).

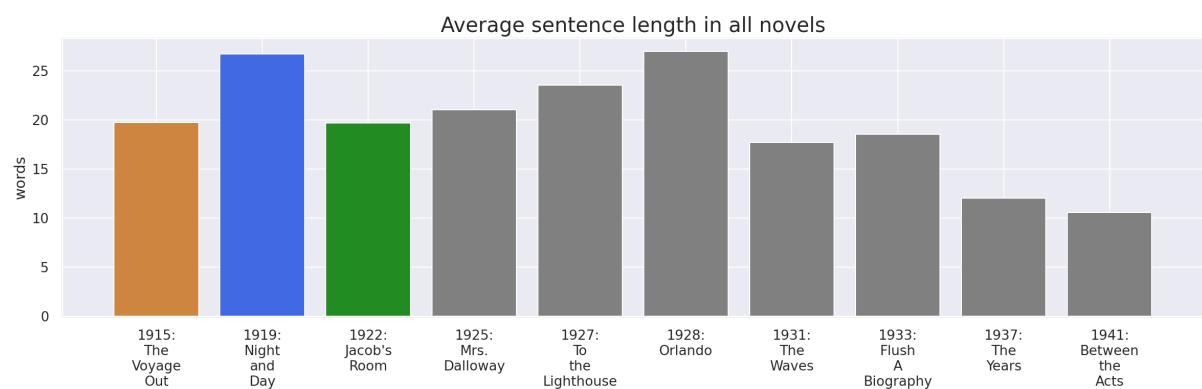
*Fig. 2. Graph showing the length of the longest sentence in all novels written by Virginia Woolf examined in this paper.*



*Fig. 3. Graph showing the length of the longest sentence in all works (novels and short stories) written by Virginia Woolf examined in this paper.*

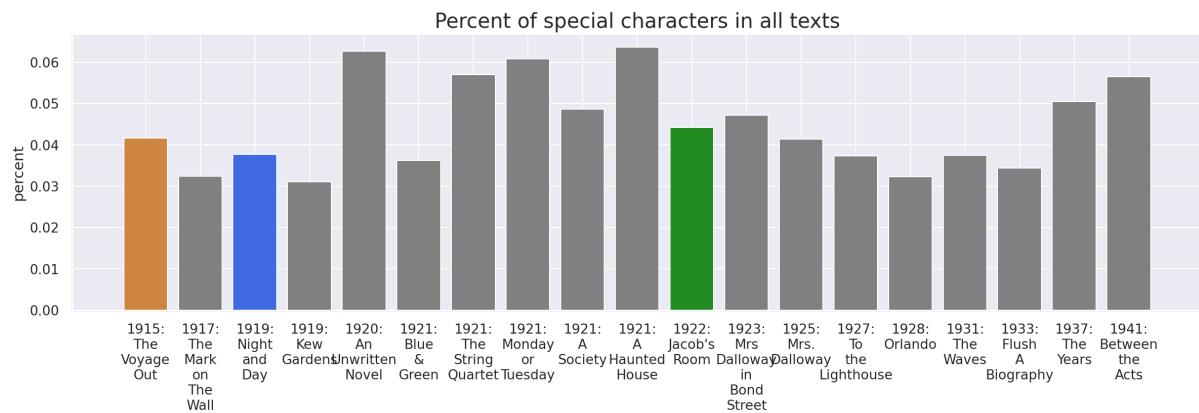


*Fig. 4. Graph showing the average sentence length in all novels written by Virginia Woolf examined in this paper.*



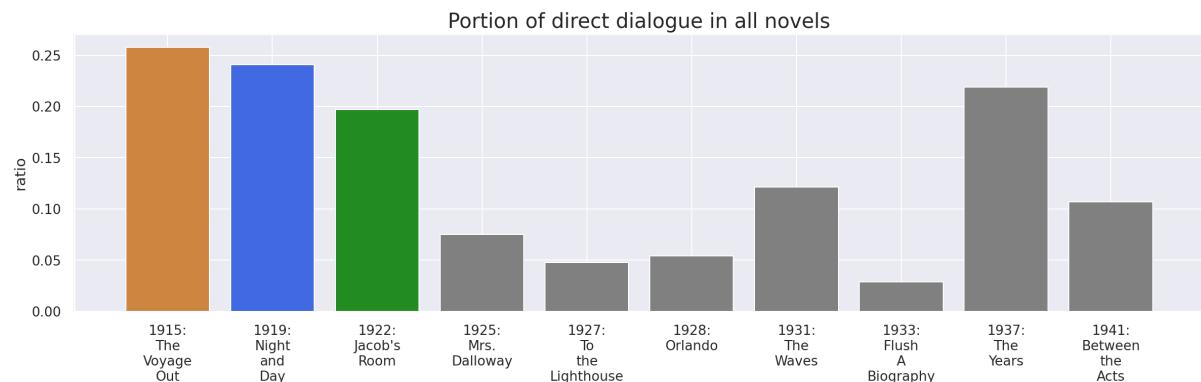
It is perhaps this drop in average sentence length that causes the use of special characters to go up in the last two works by Virginia Woolf. *Figure 5* shows that the average percentage of special characters stays around 0.04 in all works. The last two works, however, have an average that lies above 0.05.

*Fig. 5. Graph showing the percent of special characters used in all works (novels and short stories) written by Virginia Woolf examined in this paper.*



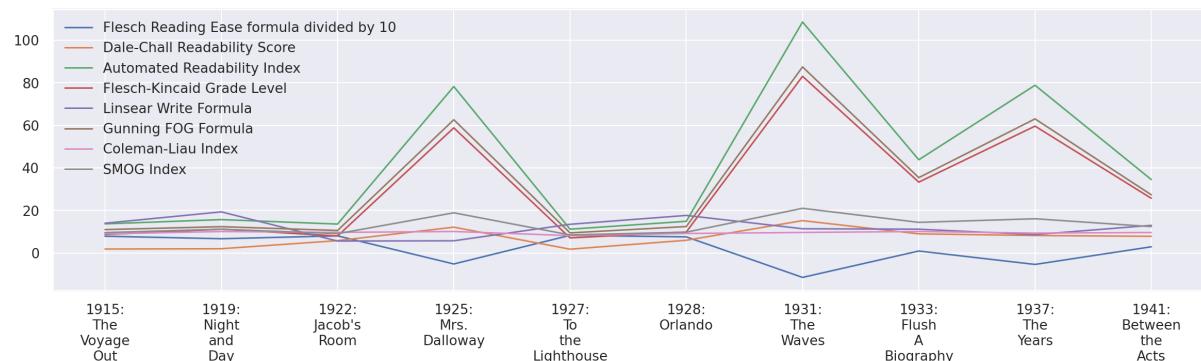
The level of dialogue also differs by work. *Figure 6* shows that whereas *The Voyage Out* and *Night and Day* have similar ratios of direct dialogue at around 0.25, this ratio slightly decreases in *Jacob's Room*. The level of dialogue drastically decreases with the appearance of *Mrs. Dalloway* with the ratio of 0.075. This ratio continues to drop until the last two novels by Virginia Woolf, *The Years* and *Between the Acts*, when the ratio of direct dialogue again increases.

*Fig. 6. Graph showing the portion of direct dialogue as opposed to all text in the novels written by Virginia Woolf examined in this paper.*



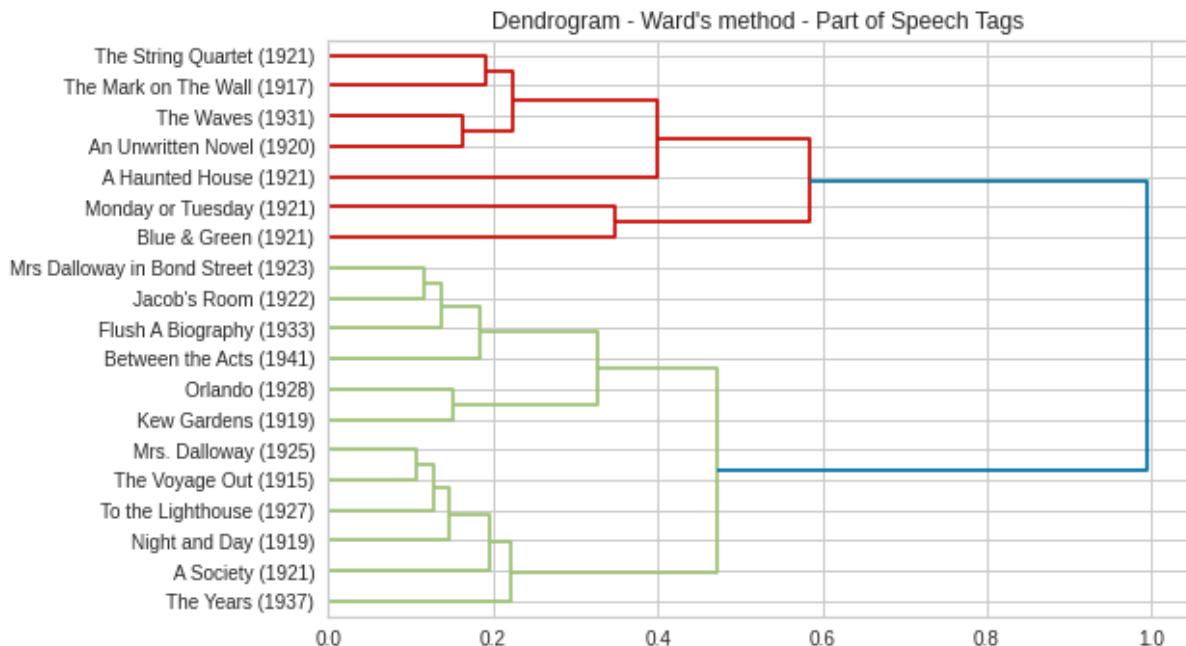
A comparison of all 10 novels in terms of readability indicates that *Mrs Dalloway*, *The Waves* and *The Years* score as being the most difficult. The Automated Readability Index, Gunning FOG Formula and Flesch-Kincaid Grade Levels tend to pick on changes in style. These three look at character / words / sentences ratios. Comparing these results with the bar chart about average sentence length, it is surprising to see that the novels with high sentence lengths (*Night and Day*, *To the Lighthouse* and *Orlando*) don't have high levels. The high score of the three difficult novels could mean that these tend to have longer words. *The Voyage Out*, *Night and Day*, *Jacob's Room*, *To the Lighthouse* and *Orlando* have similar values, inside a range of 0 to 20.

*Fig. 7. Graph showing the readability of all novels, based on the indexes of the Flesch Reading Ease Formula, Dale-Chall Readability score, Automated Readability index, Flesch-Kincaid Grade level, Linsear Write Formula, Gunning Fog formula, Coleman-Liau Index, and SMOG index, written by Virginia Woolf examined in this paper.*

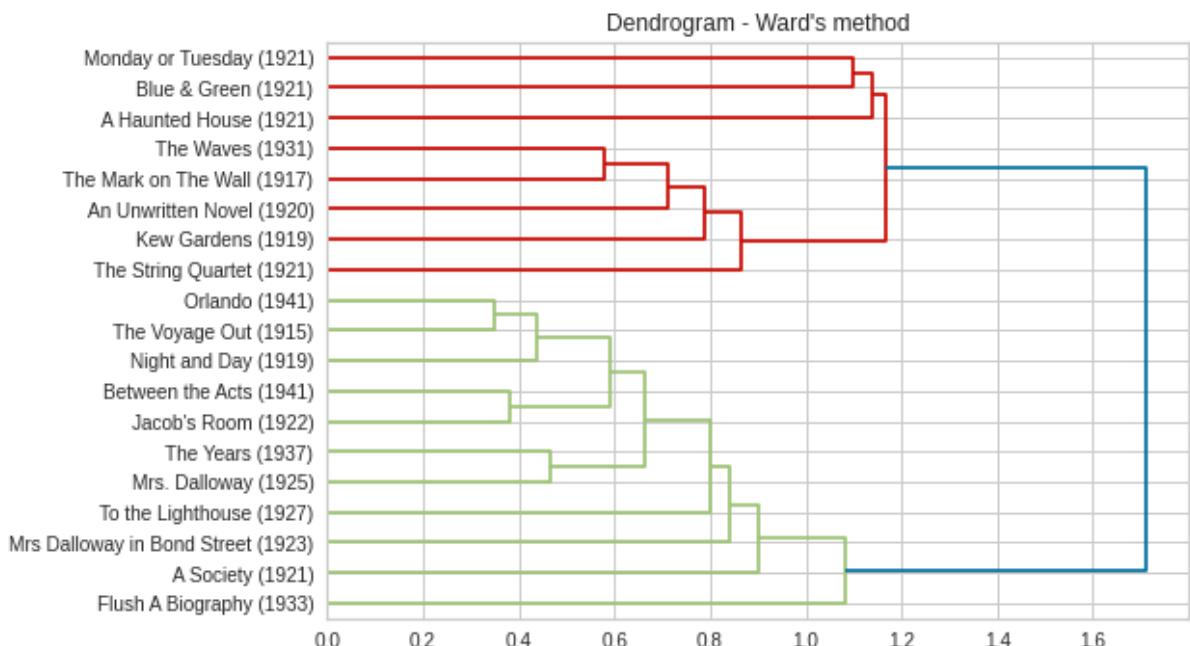


The "Mrs Dalloway in Bond Street" is a short story grouped with the novels too, which is not that surprising, as it was later turned into a novel. In a general way, the texts are placed next to texts that were published close in time to them, often around three years for word clustering and five years for part of speech clustering. This is not true for the 'misclassified' texts. Besides this, there are some exceptions. In terms of word usage, *Orlando* is put next to the first written novel, and *Between the Acts* and *The Years* around the early works. *Flush A Biography* stands out too. When the part of speech tags is the focus, *Flush* and *Orlando* are near each other, probably due to both being fictitious biographies. Clustering allows us to map the similarity between the texts. One was done on raw texts, and the other one on the part of speech tags detected by the *NLTK* module. The novels and short stories are divided quite well in two groups, with the exception of *The Waves*.

*Fig. 8 a. Trigram analysis clustering graph showing the proximity between the novels written by Virginia Woolf.*



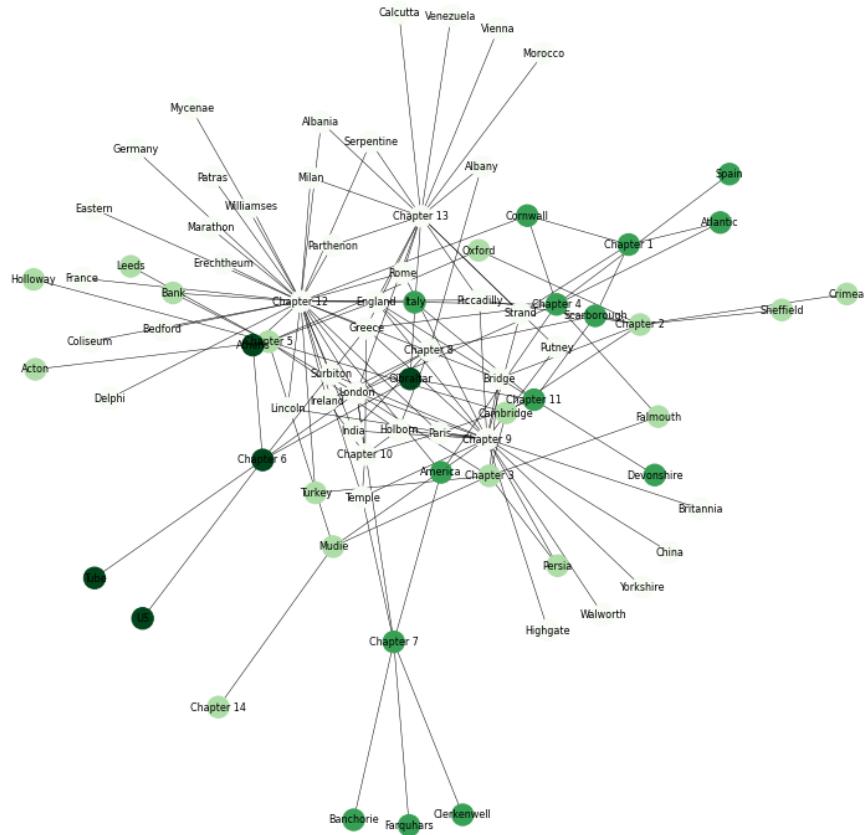
*Fig. 8 b. Part of Speech analysis clustering graph showing the proximity between the novels written by Virginia Woolf.*



## Network Analysis

A network analysis was conducted on the locations mentioned in the individual chapters of *Jacob's Room*. This novel, which has travel at its core, contains a myriad of locations in Great Britain, Europe, and even the Middle East. As shown in the image below, all 14 chapters mention at least one location. Most locations are found in chapter 12, which mentions 19 locations in total (see *Figure 9*). The least locations are mentioned in chapter 2, which mentions only 5 locations.

*Fig. 9. Network analysis image showing a network of chapters in Jacob's Room and the locations mentioned in these chapters.*



*Table 2. Betweenness centrality of locations mentioned in Jacob's Room, written by Virginia Woolf.*

Name of the node	Betweenness centrality
1	0.01135699504
Scarborough	0.0174118157
Cornwall	0.01746851309
Strand	0.01568609063
Atlantic	0.001841947447
2	0.06915225381
Sheffield	0
Cambridge	0.008530615967
Crimea	0
Piccadilly	0.05017765276
Bridge	0.03983408637
Oxford	0.00394348079

Name of the node	Betweenness centrality
8	0.03094760301
Ireland	0.005440167093
India	0.01005513193
Albany	0.001744453865
9	0.2316799002
Putney	0.003652804697
Walworth	0
Lincoln	0.005154932158
Yorkshire	0
Britannia	0
Rome	0.01399867178
China	0

3	0.03513267338	Highgate	0
Mudie	0.009004417877	Paris	0.02118619672
Persia	0.002181307048	Surbiton	0.005154932158
Falmouth	0.002082532428	10	0.0118293804
London	0.1999899823	11	0.04007683086
Turkey	0.004998408796	Devonshire	0
4	0.06883330251	12	0.4594365795
Italy	0.0155991904	Parthenon	0.005190934921
Greece	0.0587735749	Erechtheum	0
Spain	0	Marathon	0
America	0.02012573724	Williamses	0
5	0.1145845449	Milan	0.005190934921
Bank	0.01217715569	Patras	0
England	0.06972677255	Bedford	0
Leeds	0	Coliseum	0
Holborn	0.01194264135	France	0
Acton	0	Serpentine	0.005190934921
Holloway	0	Mycenae	0
6	0.05945526503	Germany	0
Gibraltar	0.01102033007	Delphi	0
Athens	0.007974364498	Albania	0.005190934921
US	0	Eastern	0
Tube	0	13	0.1746657045
7	0.07846335248	Morocco	0
Temple	0.02966104937	Venezuela	0
Clerkenwell	0	Vienna	0
Farquhars	0	Calcutta	0
Banchorie	0		

The chapter-location network in *Jacob's Room* is dense and not centralized (see *Figure 9*). No location is mentioned by all chapters. The betweenness centralities of the network are recorded in *Table 2*. The highest scores, and thus a bridging position is given to the following locations from highest to lowest scores:

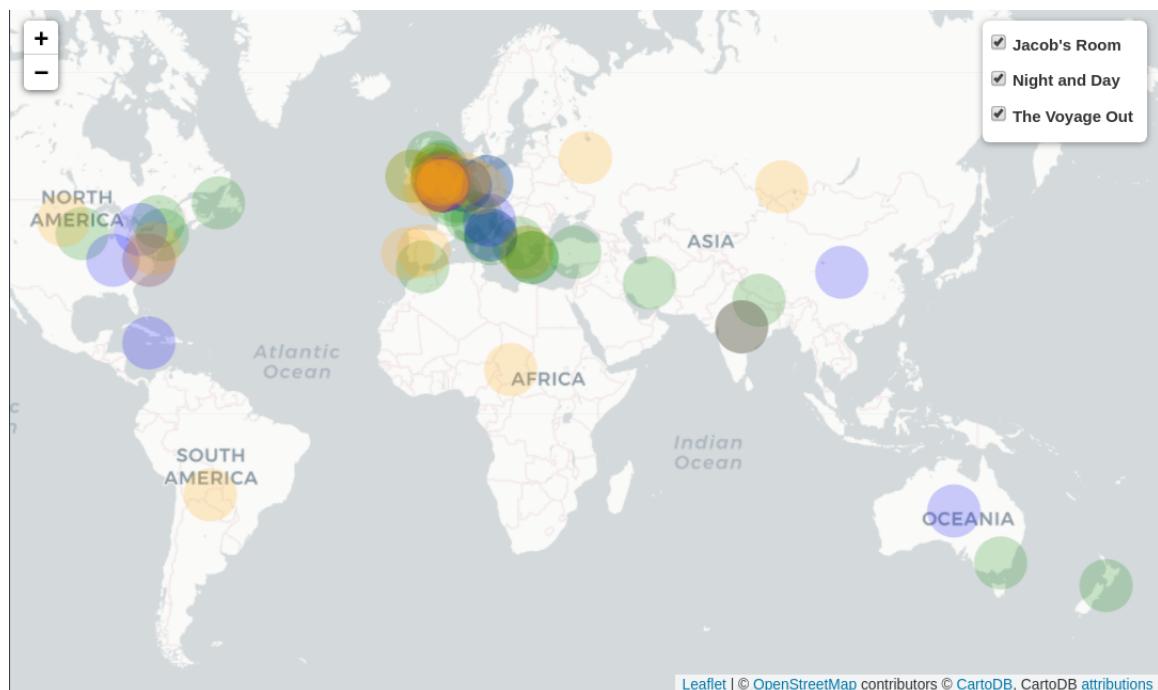
- London
- England
- Greece
- Piccadilly Road

- Paris
- America
- Cornwall

## Geographical Analysis

A geographical analysis was conducted of three of Virginia Woolf's major works. *Jacob's Room*, *The Voyage Out* and *Night and Day* all start in London. As travel stories, however, all three novels feature other locations across the globe. Mapped out first across all three books, prominent locations feature on an interactive map as discussed in the previous chapter, and available via [thedigitaldetectivesjacobsroomproject.weebly.com](http://thedigitaldetectivesjacobsroomproject.weebly.com). An overview of the locations can be viewed in *Figure 10*. Sentiment analysis was furthermore used on the three novels as a whole. Finally, the sentences featuring prominent locations in *Jacob's Room* were subjected to sentiment analysis as well. The results of both sentiment analyses are discussed below.

*Fig. 10. Map showing the locations mentioned in Jacob's Room, The Voyage Out and Night and Day, written by Virginia Woolf.*



## Sentiments

As seen in the image below, the three early career novels by Virginia Woolf score mostly positive on text-sentiment. Many differences in sentiment can be found, however, starting with the endings of the three books, of which only *Jacob's Room* scores significantly high on negative sentiment. This highest negative score is almost fitting in a line of sentiment scores on the novel that has the most negative scores overall.

The highest positive scores overall can be found in *Night and Day* (see *Figure 11*). Looking more closely at the two sentences with the highest scores, sentence 300 with a

'He should have felt that his own sister was more original, and had greater vitality than Miss Hilbery had; but his main impression of Katharine now was of a person of great vitality and composure; and at the moment he could not perceive what poor dear Joan had gained from the fact that she was the granddaughter of a man who kept a shop, and herself earned her own living.'

compound score of 0.975, [OBJ] sentence 1572 with a score of 0.9777

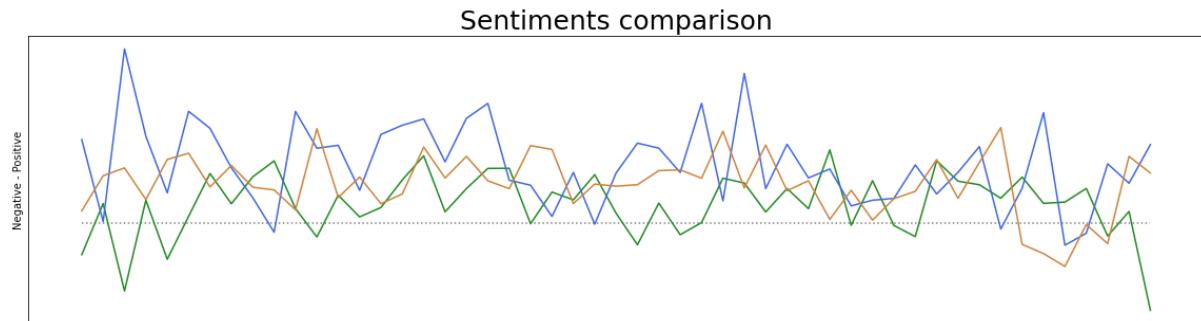
'She wouldn't understand it, but I like her very much as she is.' For some time they discussed what the women had better do; and as Ralph became genuinely interested in the question, Mary unconsciously let her attention wander, and a great desire came over her to talk to Ralph about her own feelings; or, at any rate, about something personal, so that she might see what he felt for her; but she resisted this wish.'

and sentence 3494 with a score of 0.975 are all describing women.

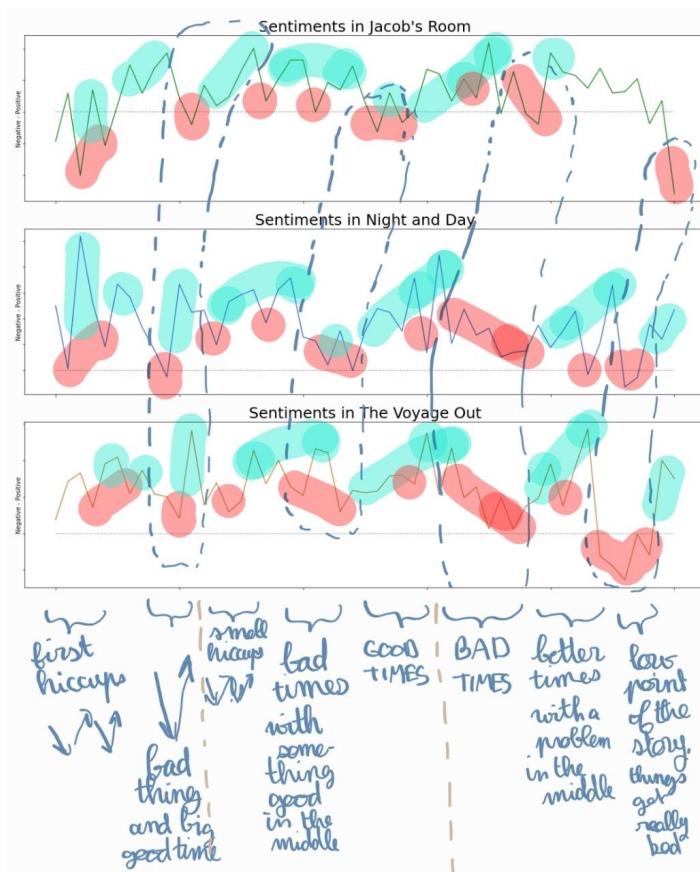
'She had been too lately and too roughly mastered by facts to take an easy pleasure in the relief of renunciation; such satisfaction as she felt came only from the discovery that, having renounced everything that made life happy, easy, splendid, individual, there remained a hard reality, unimpaired by one's personal adventures, remote as the stars, unquenchable as they are.'

While all novels clearly fluctuate in sentiment over the course of the story, the overall shape of the fluctuations overlap considerably between *The Voyage Out* and *Night and Day*. Both novels start on a positive note, end on a positive note, and experience a set of negative scores near the end of the novel, as shown in *Figure 12*. Sentiments and tension are known to have been used to structure a story. Think of the hero's journey monomyth described by Campbell, a story structure that can be found in countless stories in whole history. These patterns can create engagement from the reader for the characters. In the field of cinematography, it is not uncommon to make a graph representing them when planning the scenes in the script.

*Fig. 11. Sentiment comparison graph showing the average scores in textual sentiments for 50 parts of the entire novels Jacob's Room, The Voyage Out, and Night and Day, written by Virginia Woolf.*



*Fig. 12. Annotated sentiment comparison graph showing a comparison of sentiment fluctuations in the novels Jacob's Room (green), The Voyage Out (brown), and Night and Day (blue), written by Virginia Woolf.*



Prominent locations mentioned in the novel *Jacob's Room* feature both Great Britain, France and Greece. The locations' scores on sentiment fluctuate widely, with both negative and positive scores present for all locations.

One location stands out, however, due to its high scores on both sentiments: France. Starting the novel with a highly negative score, this location scores a highly positive score at the end of the novel. This score reaching 0.6249 is only trumped by the ending score of London, which scored a 0.857.

Interestingly, three of the locations mentioned which are part of Great Britain (see 'England', 'London', and 'the British museum') start the novel with relatively positive scores, but all drop to neutral scores or below early on in the novel (see *Figure 49 in the appendix*). These neutral and negative scores then pick up throughout the remainder of the novel, ending on relatively high positive scores overall.

*Fig. 13. Comparative image of the sentiments of locations in Jacob's Room, versus The Voyage Out and Night and Day*

name	Jacob's Room		The Voyage Out		Night and Day	
	frequency	sentiment	frequency	sentiment	frequency	sentiment
America	3	0.4298666667			7	0.2502357143
Cambridge	11	-0.0279	18	-0.047595		
England	8	0.18670625	45	-0.004595744 681	17	0.2131625
Hyde Park	6	-0.067916666 67			2	0
India	3	0.13438	6	-0.06787	9	0.1185933333
India	3	0.13438	6	-0.06787		
Ireland	5	0.22758	2	-0.35015		
Italy	6	-0.063316666 67			2	0.17675
London	21	0.1507125	47	0.1487863636	49	0.1633811321
Strand	3	-0.2425			16	0.1470761905

England, London, and India are the only locations mentioned more than once in all the three early career novels. The sentiments associated with them vary from novel,

except for London (0.15) and Cambridge (- 0,03). Ireland is painted negatively in *The Voyage Out* while positively in *Jacob's Room*. Woolf associates different sentiments with the locations depending on the story, however for London, the overall sentiments are really consistent, around 0,15.

## Discussion

The results of our analysis reveal that the longest sentences in Woolf's novels can be found in *Orlando*, *To the Lighthouse*, *The Voyage Out*, and *Mrs Dalloway* respectively. Excluding *The Voyage Out*, this finding matches scholars' claims that these three novels represent "the culmination of her experimental lyric technique, a tour de force in high modernist poetic fiction" (Goldman 49). The length of these sentences indicates that they are more likely to be written in a stream of consciousness technique or in free indirect discourse. Taking a closer look at Woolf's first three novels reveals that, surprisingly, the longest sentence can be found in her first novel, *The Voyage Out*. This novel also features in the list of all novels with the longest sentence, and is in this sense an outlier. This result does not match our expectations, since we did not expect her first novel to be closer in sentence length to her more experimental novels.

The rise in special characters in Woolf's final two works, which we see in *Figure 5*, could be explained by the decrease in average sentence lengths in these two works. Shorter sentences could indicate a lesser degree of free indirect discourse or stream of consciousness, and therefore more direct dialogue. Since direct dialogue is usually enclosed within quotation marks, this would explain the rise in special characters.

Regarding the presence of dialogue, we hypothesised that if a text was more experimental in nature, the proportion of dialogue would diminish since Woolf would have most likely used devices such as stream of consciousness and free indirect discourse, which as discussed in the literature review, are usually not enclosed within question marks. Indeed, the results confirm our hypothesis: the amount of dialogue drastically decreases in *Mrs Dalloway*, and remains low for the remaining works considered to be the pinnacle of her modernist writing (*To the Lighthouse*, *Orlando*, *The Waves*). The fact that, in comparison to *The Voyage Out* and *Night and Day*, a small decrease in the proportion of dialogue can be detected, suggests that *Jacob's Room* was indeed a slight turning point in Woolf's writing style. This, however, is not particularly visible in the readability scores, where *Mrs Dalloway* and *The Waves* score particularly high in comparison to *Jacob's Room*. The particularly high score for *The Waves* supports

Suzanne Raitt's questioning of whether *Jacob's Room* is "the definitive realisation of [Woolf's] fictional voice" (29) and whether this should not be taken "with a certain degree of reserve" (29). Raitt makes this observation based on Woolf's own diary entry in which she exclaims "I think I am to embody, at last, the exact shapes my brain holds. What a long toil to reach this beginning - if *The Waves* is my first work in my own style" (Diary 4; quoted in Raitt 29).

Network analysis yields interesting results. A difference in general location mentions can be detected among the chapters: whereas chapter 12 mentions locations in Greece almost exclusively, chapters 7 and 9 only mention locations in Great Britain. Six of these locations, furthermore, appear as locations with a high centrality. This could be an indicator of Jacob's actual geographical location. Most locations are mentioned in these 3 chapters alone.

One outlier can be found in the network, which was mistyped as a location by the program used. China, a word used to describe fine powder porcelain, was used in *Jacob's Room* to describe tea drinking supplies, such as teacups. However, as the word features widely in the novel, being used a total of 8 times, of which only once used as a location in chapter 4. Taken out of context, the word doubles as porcelain the other 7 times, but is mistyped as a location. This points to the shortcomings of tools such as *SpaCy* which are used for Named Entity Extraction.

The sentiment analysis of *Jacob's Room*, *The Voyage Out* and *Night and Day* shows a difference in the overall sentiment style between *Jacob's Room* and the other novels. The difference might point to a writing style in which negativity, especially near the end of the novel - or rather a 'bad ending' - is key.

On the basis of locations, two conclusions can be drawn regarding the sentiment in *Jacob's Room*. First, it is important to acknowledge that the first mention of France in the novel is surrounded by negatively scored words. As the location - a word itself - is deemed neutral in sentiment by *VADER*, we must discuss the sentence in which it is found to see if the location is described negatively. Notable is the word 'broken', a negatively scored word, that appears in the sentence in which 'France' is found in (Woolf

1922, chapter 12). While the sentence speaks of France as a subject of ‘imagination’ (*ibid.*), ‘broken’ here refers to the state of a toy (see figure 14). This instance explains why we find a negative score here. It has nothing to do with France as a location, which on a deep reading level can be scored positive as a location that ‘draws the superfluous imagination’ (*ibid.*) and is thus deemed interesting by Jacob.

*Fig. 14. Screenshot of sentence in which ‘France’ is mentioned in chapter 12 of Jacob’s Room, written by Virginia Woolf.*

No doubt we should be, on the whole, much worse off than we are without our astonishing gift for illusion. At the age of twelve or so, having given up dolls and broken our steam engines, France, but much more probably Italy, and India almost for a certainty, draws the superfluous imagination. One's

Second, we can wonder if the sentiments scored on sentences in which Great Britain’s locations appear show us the traveling incentives of Jacob in *Jacob’s Room*. Starting with positive scores, these locations can be seen as the favorite locations Jacob likes to visit. Their neutral and negative scores later on in the novel might betray his travel incentive - he is no longer entranced by the locations, and instead wishes to travel abroad. The locations’ increasing sentiment scores near the end of the novel might betray a wish to visit them again.

## Conclusion

This paper sought to discern the changes in the writing style across Virginia Woolf's oeuvre. In order to do this, 13 texts, including novels and short stories were analysed on text level, with *Jacob's Room*, *The Voyage Out* and *Night and Day* specifically scrutinized. The analysis included, alongside stylometry, a network analysis and a geographical sentiment analysis.

The aim of this paper was to examine the stylistic changes across Woolf's *oeuvre*, with special attention being paid to *Jacob's Room*, which is often regarded as a turning point in Woolf's writing career. The analysis of the corpus confirms that while *Jacob's Room* does indeed mark a change in Woolf's writing style, this change is less pronounced than initially expected. While *Jacob's Room* does have a higher proportion of dialogue when compared to Woolf's high modernist works, this is not the case when compared to two of her earlier works. According to the analysis, *The Voyage Out*, as well as *Night and Day*, both exhibit more dialogue than *Jacob's Room*, indicating a slight shift towards characteristically modernist techniques of writing.

Looking at the entire corpus, this paper strongly suggests that three phases can be found in the writing of Virginia Woolf. The first phase consists of the first three novels: *Jacob's Room*, *The Voyage Out* and *Night and Day*. The second phase roughly starts with the publication of *Mrs. Dalloway* in 1925. *To the Lighthouse*, *Orlando*, and *The Waves* could all be categorised as belonging to the middle period. This phase is known for Woolf's use of narrative techniques characteristic of modernism, such as stream of consciousness and free indirect discourse. Our analysis indicated that a third phase of Woolf's writing consists of her last two novels: *The Years* (1937) and *Between the Acts* (1941). The fact that the proportion of dialogue and special characters in these two novels is higher in comparison to the novels in the middle period seems to suggest that these two novels are closer in style to the earlier novels. This certainly holds true for *The Years*, which, according to Jane Goldman, marks a "return to a more conservative novelistic form" (78). Goldman explains that *Between the Acts*, on the other hand, is again more experimental in form since it's "chapterless and sectionless" (83). The clustering

of these two novels, as explained above, could be due to the fact that *Between the Acts* was published posthumously. A potential reason for the clustering of these two novels into the same period could be the fact that *Between the Acts* was published posthumously, and was therefore completed by Woolf's husband, Leonard Woolf. Clustering confirms the presence of these phases.

Stylometry analysis and geographical sentiment analysis can be used to define the features of the first phase. *Jacob's Room*, *The Voyage Out* and *Night and Day* feature similar amounts of direct dialogue, sentence length and use of special characters. Apart from this, sentiment analysis on these novels shows similarities, with *Jacob's Room* as an outlier with a higher overall negative sentiment.

Network analysis, on the other hand, cannot be used to discern the phases in writing. What it does show, however, is the importance of geographical locations and the place of travel in the works of Virginia Woolf, which were scrutinized in both the maps and sentiment analysis on locations of *Jacob's Room*.

To solve our research question, methods were used which are not often applied in digital literary studies, such as the unconventional network analysis, readability formulas and sentiment arcs. *Narrative Natural Language Processing* is a domain which is still in development. There is work done on making automated extraction of events, storylines, specific emotions and dialogues possible with machine learning. These tools may become more accessible in the coming years and give new approaches to the study of storytelling in novels but also in blog posts and news.

More specifically about Woolf, in this paper the method was based on extracting information about style and narration from the text, however external sources of information can be taken into account. Authorial London has a dataset of manually assigned locations, and could be useful to have a more precise insight in the way Woolf uses locations in her stories. The downside is that it is limited to London. Using more complex geolocalisation tools or human verification would make it possible to have more accurate data on all the locations. Another interesting source of data is the text.study database, which contains narrative tropes. While it doesn't allow to see their

evolution through the novel, to study their presence in Woolf's whole oeuvre can give a better grasp of the change in topics. Regarding the style, a lot of Woolf's drafts have been digitized. A computational study of those, and their final version shows conscious changes made by Woolf in her manner of writing.

In conclusion, the analysis of Woolf's writing style reveals that her first three novels, including *Jacob's Room*, can be clustered together. The first phase is followed by the phase in the middle which includes Woolf's most famous works, such as *To The Lighthouse*, *Orlando*, and *Mrs Dalloway*, which are also the most experimental in form. Woolf's final two novels, *The Years* and *Between the Acts*, form a third phase of her writing. The transition between the three periods is not always smooth in the statistics, showing a conscious choice to write in a different way.

## Bibliography

Abbott, H. Porter. *The Cambridge Introduction to Narrative*. 2nd. ed., Cambridge University Press, 2008.

Brown, Matt. *Mapped: All Virginia Woolf's Novels*. The Londonist, 2017.  
<https://londonist.com/london/books-and-poetry/mapped-all-virginia-woolf-s-novels>. Accessed 19 June 2021.

Caughie, Pamela L., Nick Hayward, Mark Hussey, Peter Shillingsburg, and George K. Thiruvathukal, eds. *Woolf Online*. <http://www.woolfonline.com>. Accessed 17 June 2021.

Cui, Yaxiao. "Parentheticals and the presentation of multipersonal consciousness: A stylistic analysis of *Mrs Dalloway*." *Language and Literature*, vol. 23, no. 2, 2014, pp. 175-187. Sage, doi: 10.1177/0963947014528112.

Evans, Elizabeth F., and Sarah E. Cornish. *Woolf and the City*. Liverpool University Press, 2010.

Evans, Martin, et al. *Authorial London*. Stanford University, 2011,  
<https://authorial.stanford.edu/>. Accessed 17 June 2021.

Google Arts and Culture, and National Museum of Women in the Arts. *Take a Tour of Virginia Woolf's Life in London: Get an insight into what inspired the author with Street View*.  
<https://artsandculture.google.com/story/1wXhSk5yw18liw>. Accessed 19 June 2021.

Golgen, Amanda. *Mapping Jacob's Room*, TECHStyle, 2013.  
<http://techstyle.lmc.gatech.edu/mapping-jacobs-room/>. Accessed 19 June 2021.

Goldman, Jane. "From *Mrs Dalloway* to *The Waves*: New elegy and lyric experimentalism." *The Cambridge Companion to Virginia Woolf*, edited by Susan Sellers, Cambridge University Press, 2010, pp. 49-69. doi:  
<https://doi-org.ezproxy.leidenuniv.nl/10.1017/CCOL9780521896948.003>

---. *The Cambridge Introduction to Virginia Woolf*. Cambridge University Press, 2006. doi:  
<https://doi-org.ezproxy.leidenuniv.nl/10.1017/CBO9780511607295>

Hammond, Adam, et al. "Using models of lexical style to quantify free indirect discourse in modernist fiction." *Digital Scholarship in the Humanities*, vol. 32, no. 2, 2017, pp. 234-250. doi: <https://doi-org.ezproxy.leidenuniv.nl/10.1093/llc/fqv072>.

---. "Modeling Modernist Dialogism: Close Reading with Big Data." *Reading Modernism with Machines*, edited by Ross, S. and J. O' Sullivan, Palgrave Macmillan, 2016, pp. 49-77. [https://doi-org.ezproxy.leidenuniv.nl/10.1057/978-1-37-59569-0\\_3](https://doi-org.ezproxy.leidenuniv.nl/10.1057/978-1-37-59569-0_3).

Hammond, Adam, and Julian Brooke. *The Brown Stocking*, 2013, University of Toronto,  
<http://brownstocking.org/about/fid.html>.

Klarer, Mario. *An Introduction to Literary Studies*. 3rd. ed., Routledge, 2011.

Raitt, Suzanne. "Virginia Woolf's early novels: Finding a voice." *The Cambridge Companion to Virginia Woolf*, edited by Susan Sellers, Cambridge University Press, 2010, pp. 29-48. doi:  
<https://doi-org.ezproxy.leidenuniv.nl/10.1017/CCOL9780521896948.002>

Sotirova, Violeta. "Woolf's Experiments with Consciousness in Fiction." *Contemporary Stylistics*, edited by Marina Lambrou and Peter Stockwell, Continuum, 2010, pp. 7-18. *EBSCOhost*,  
<search.ebscohost.com.ezproxy.leidenuniv.nl:2048/login.aspx?direct=true&db=e000xww&AN=377557&site=ehost-live>.

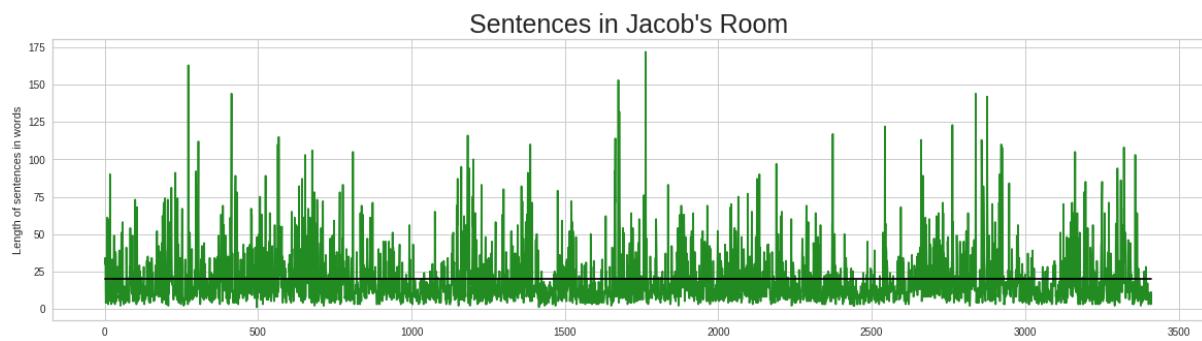
Walsh, Brandon. "Reading Speech: Virginia Woolf, Machine Learning, and the Quotation Mark." *Scholars's Lab*, University of Virginia Library, 17 May 2016,  
<https://scholarslab.lib.virginia.edu/blog/reading-speech-virginia-woolf-machine-learning-and-the-quotation-mark/>. Accessed on June 20, 2021.

Woolf, Virginia. "Modern Fiction." The Norton Anthology English Literature, edited by Stephen Greenblatt et al., 9th ed., W. W. Norton & Company, 2012, pp. 2150-2155.

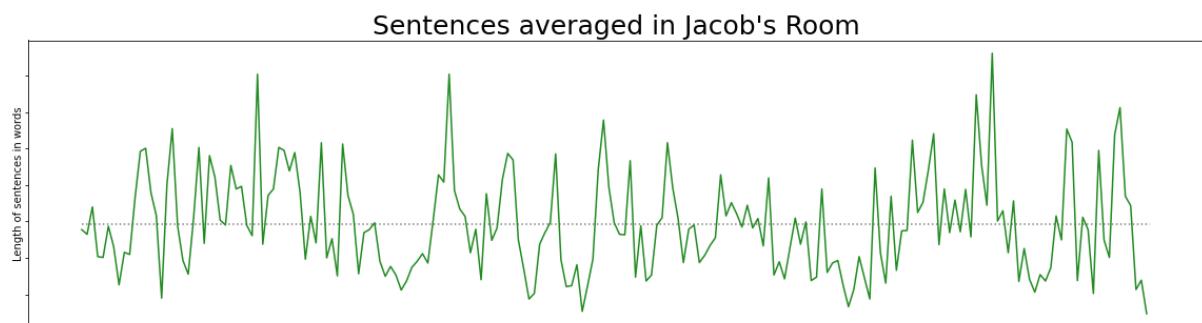
## Appendix

This appendix contains all graphs created for this paper. These graphs can also be found on the website created for this paper:  
[thedigitaldetectivesjacobsroomproject.weebly.com](http://thedigitaldetectivesjacobsroomproject.weebly.com).

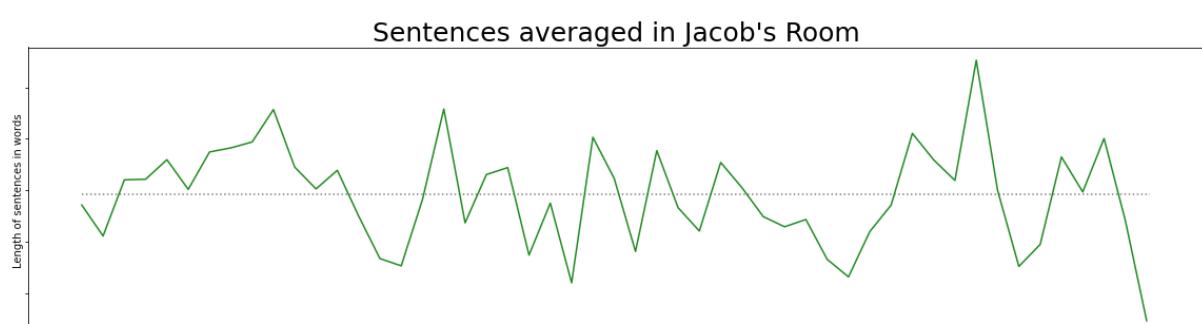
*Fig. 1. Sentence length in words in Jacob's Room.*



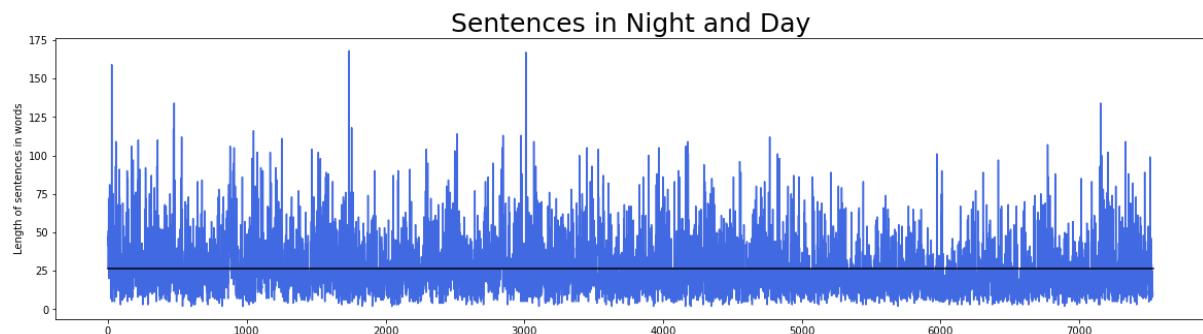
*Fig. 2. Sentence length in words in Jacob's Room, averaged.*



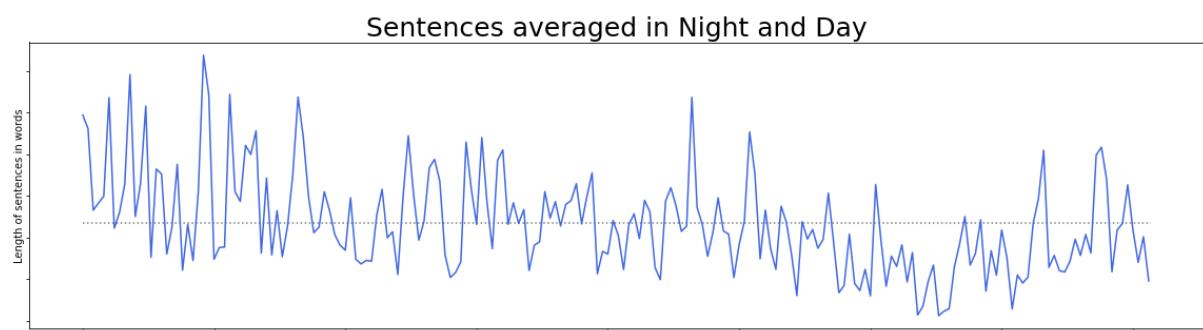
*Fig. 3. Sentence length in words in Jacob's Room, most averaged.*



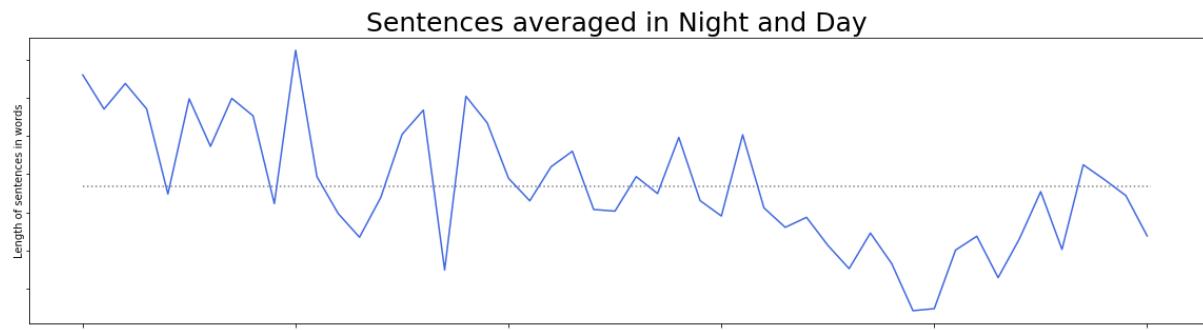
*Fig. 4. Sentence length in words in Night and Day.*



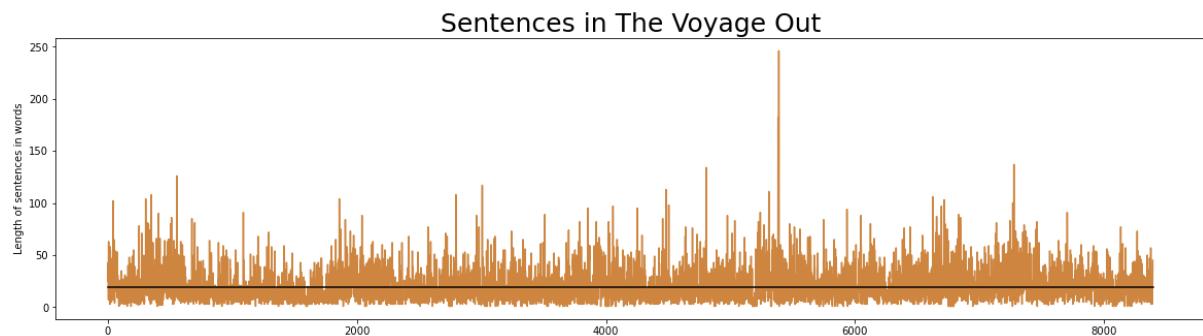
*Fig. 5. Sentence length in words in Night and Day, averaged.*



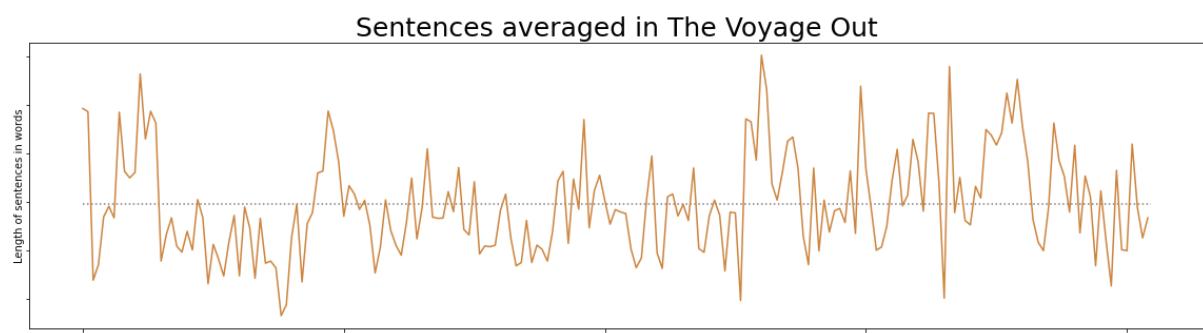
*Fig. 6. Sentence length in words in Night and Day, most averaged.*



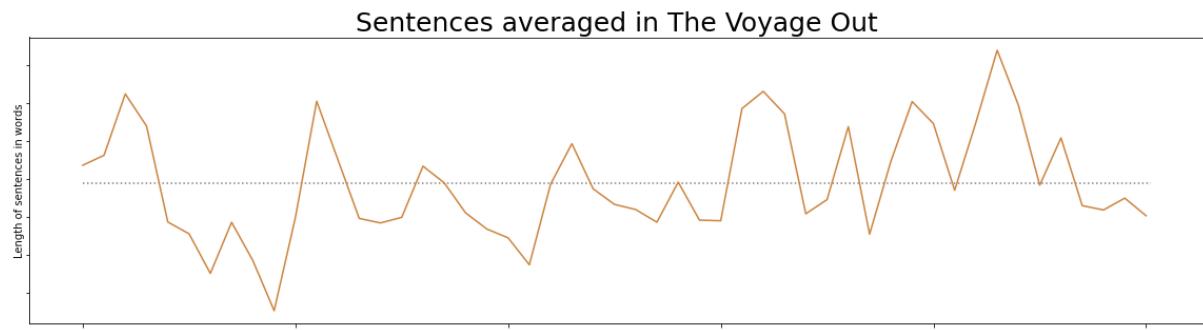
*Fig. 7. Sentence length in words in The Voyage Out.*



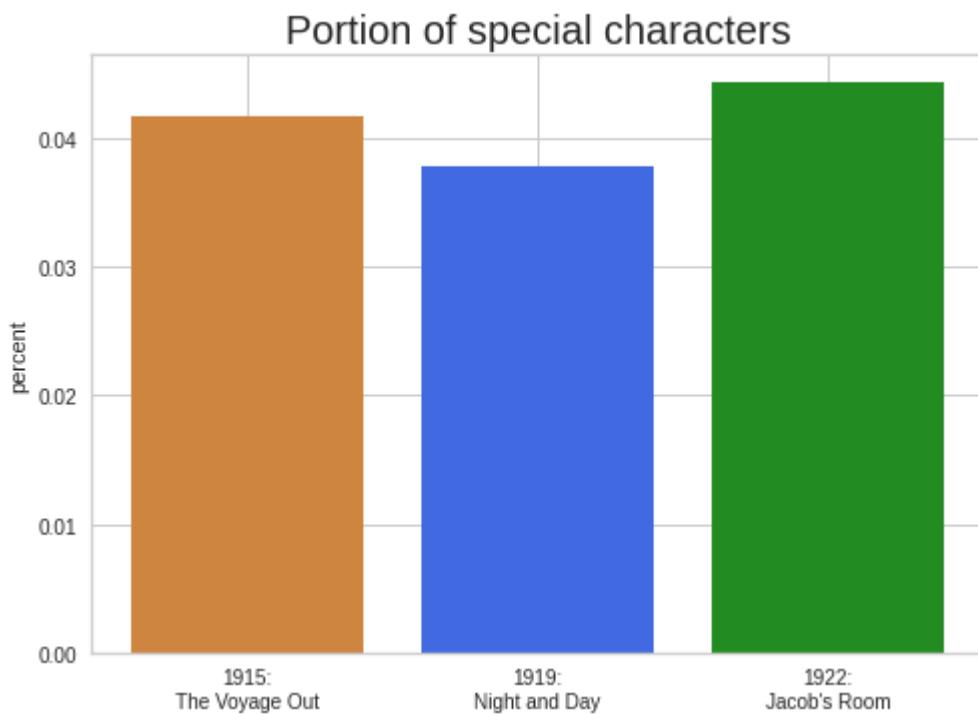
*Fig. 8. Sentence length in words in The Voyage Out, averaged.*



*Fig. 9. Sentence length in words in The Voyage Out, most averaged.*

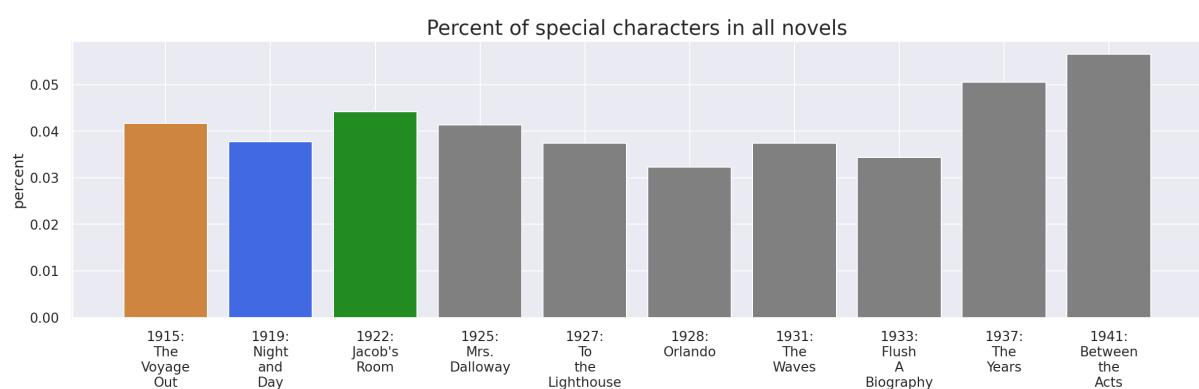


*Fig. 10. Portion of special characters in The Voyage Out, Night and Day, and Jacob's Room.*



*Fig. 11. Portion of direct dialogue in *The Voyage Out*, *Night and Day* and *Jacob's Room*.*

*Fig. 12. Percent of special characters in all novels by Virginia Woolf.*



*Fig 13. Percent of quotation pairs in all novels by Virginia Woolf.*

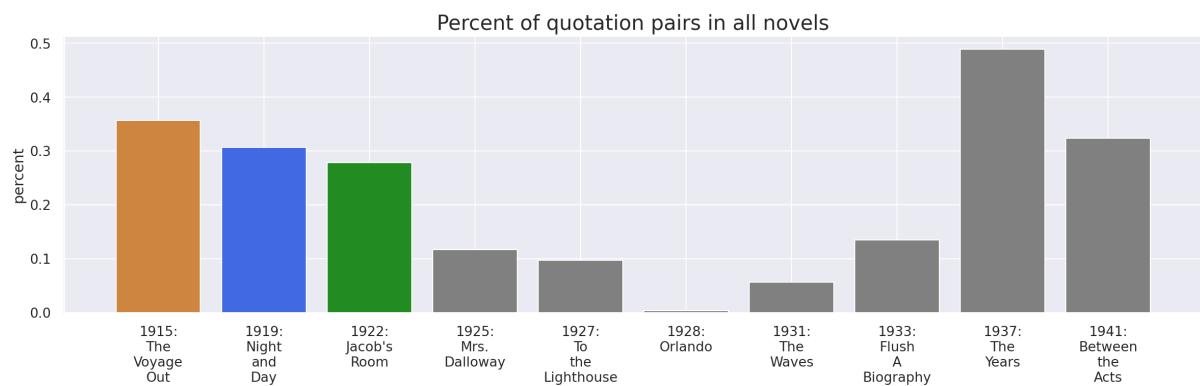


Fig. 14. Percent of quotation pairs and apostrophes in all novels by Virginia Woolf.

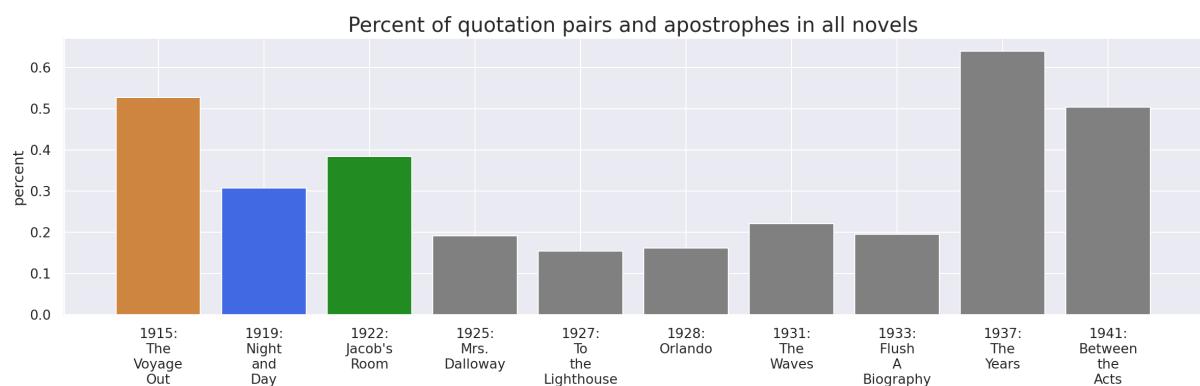


Fig. 15. Longest sentence in all short stories.

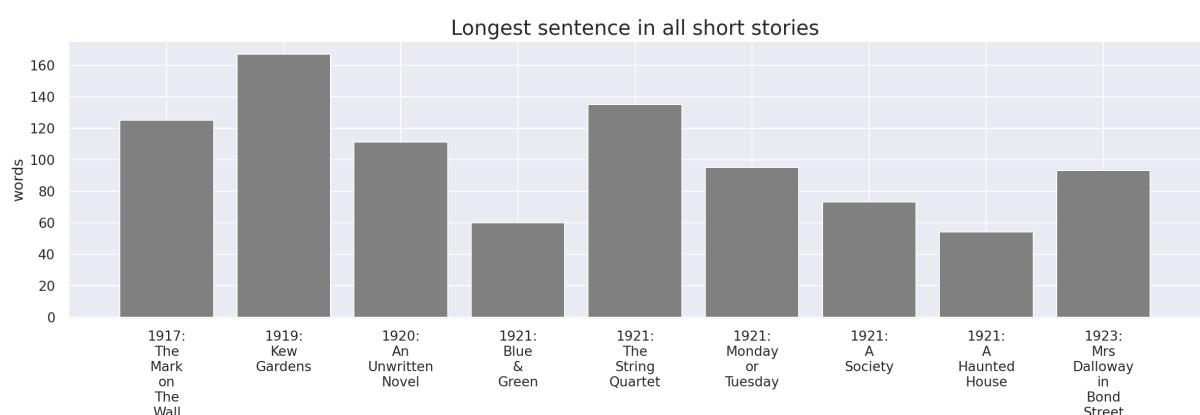
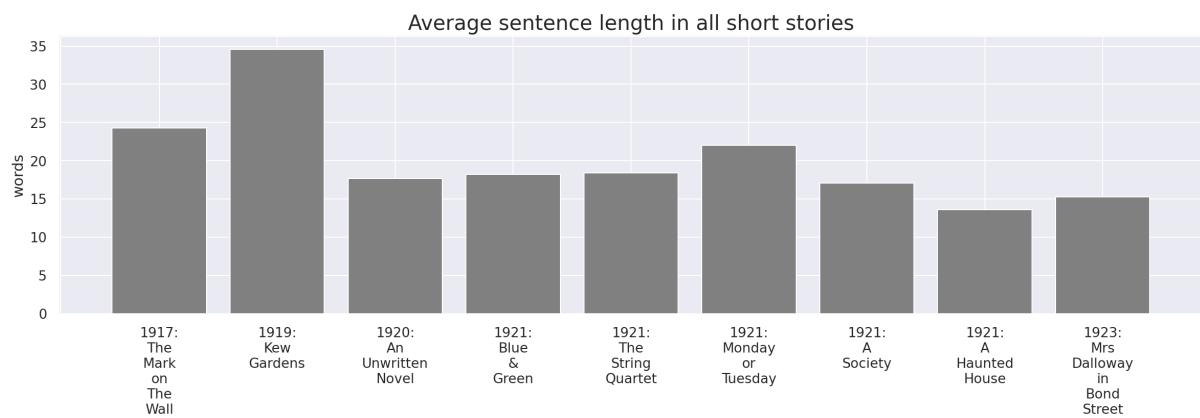
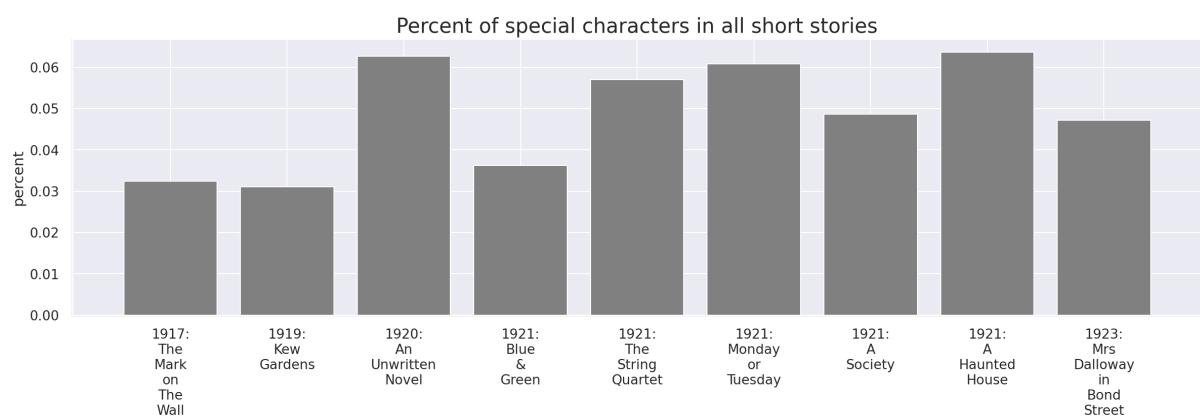


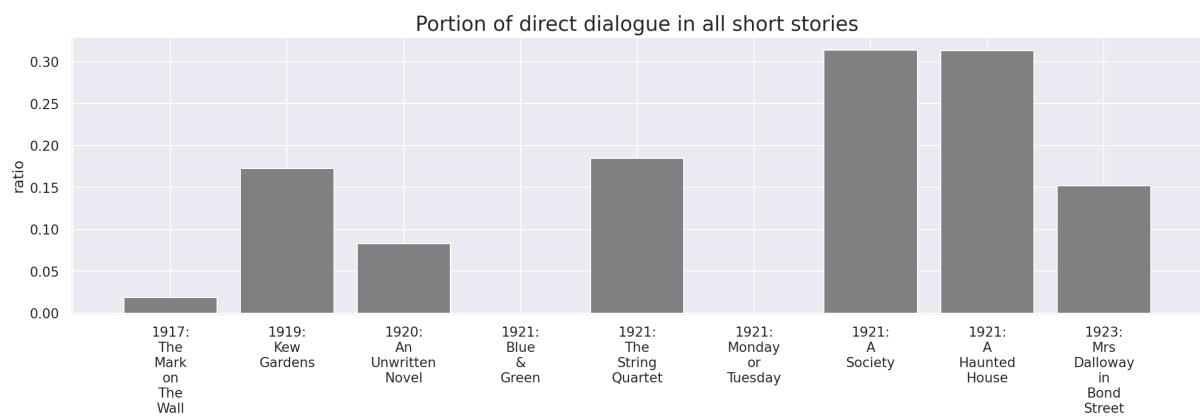
Fig. 16. Average sentence length in all short stories.



*Fig. 17. Percent of special characters in all short stories.*



*Fig. 18. Portion of direct dialogue in all short stories.*



*Fig. 19. Percent of quotation pairs used in all short stories.*

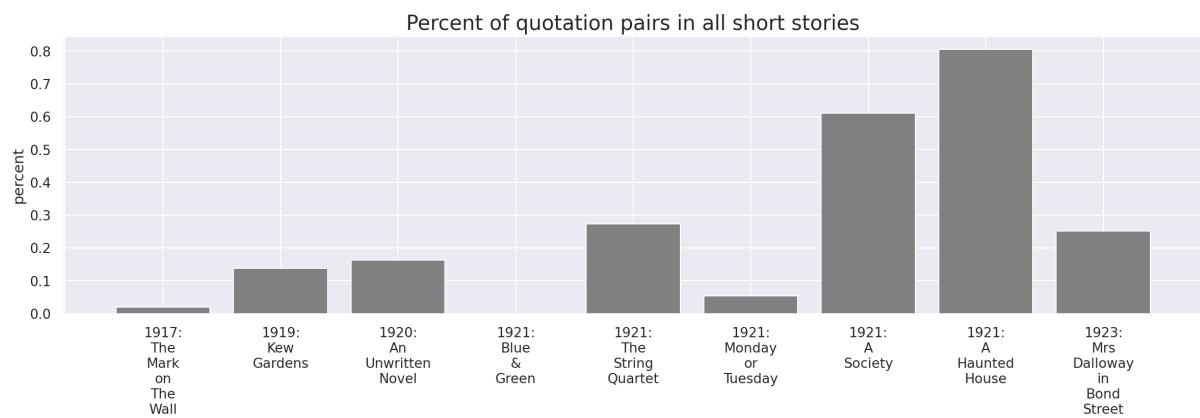


Fig. 20. Percent of quotation pairs and apostrophes in all short stories.

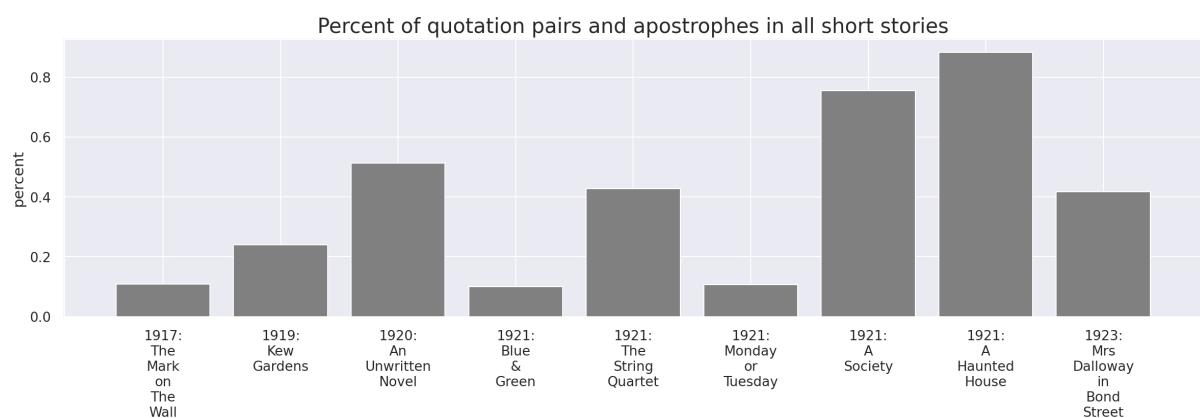


Fig. 21. Average sentence length in all texts.

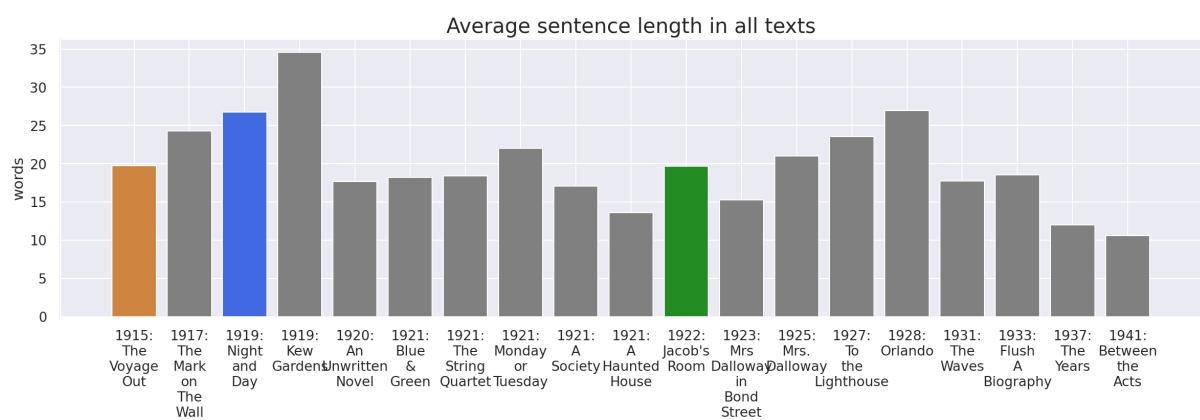


Fig. 22. Portion of direct dialogue in all texts.

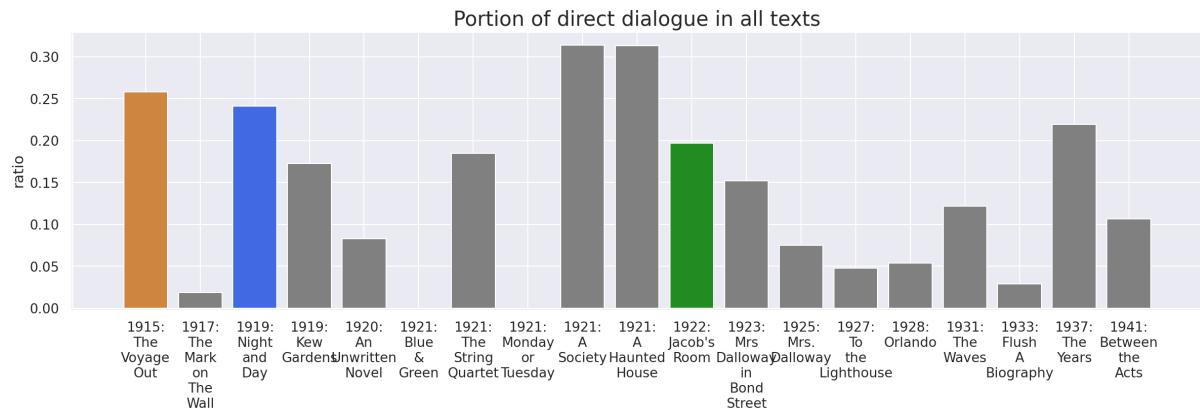


Fig. 23. Percent of quotation pairs in all texts.

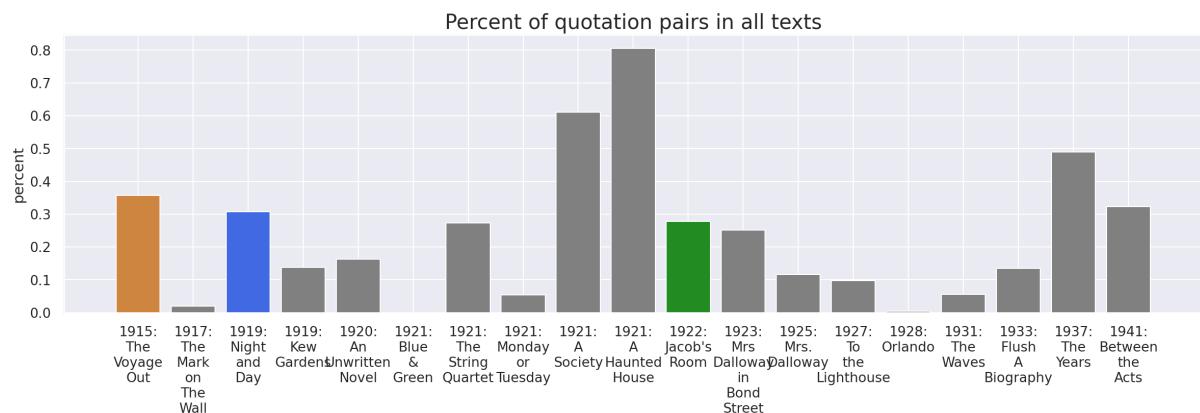


Fig. 24. Percent of quotation pairs and apostrophes in all texts.

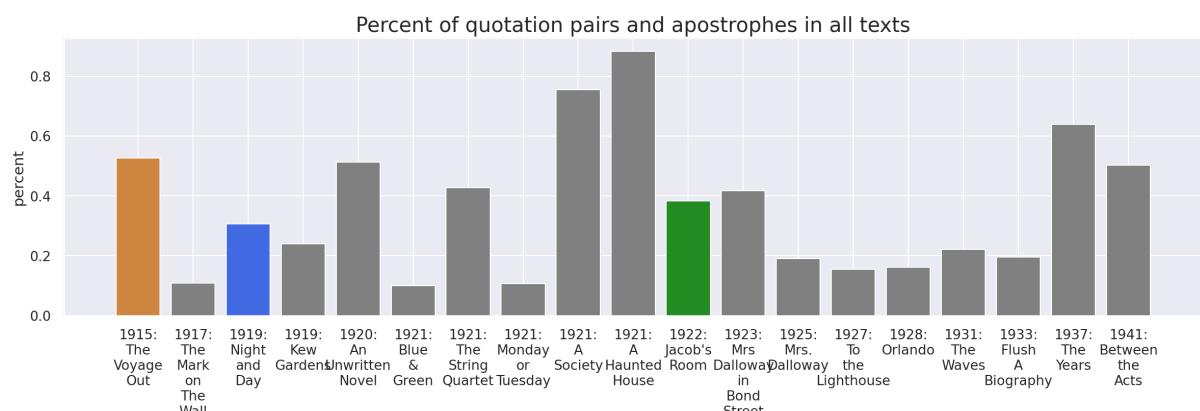


Fig. 25. Readability chart.

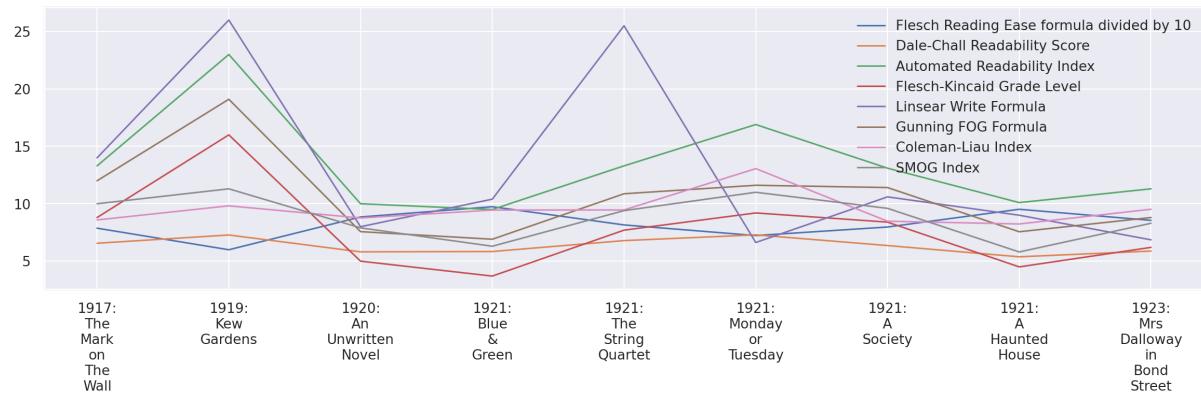
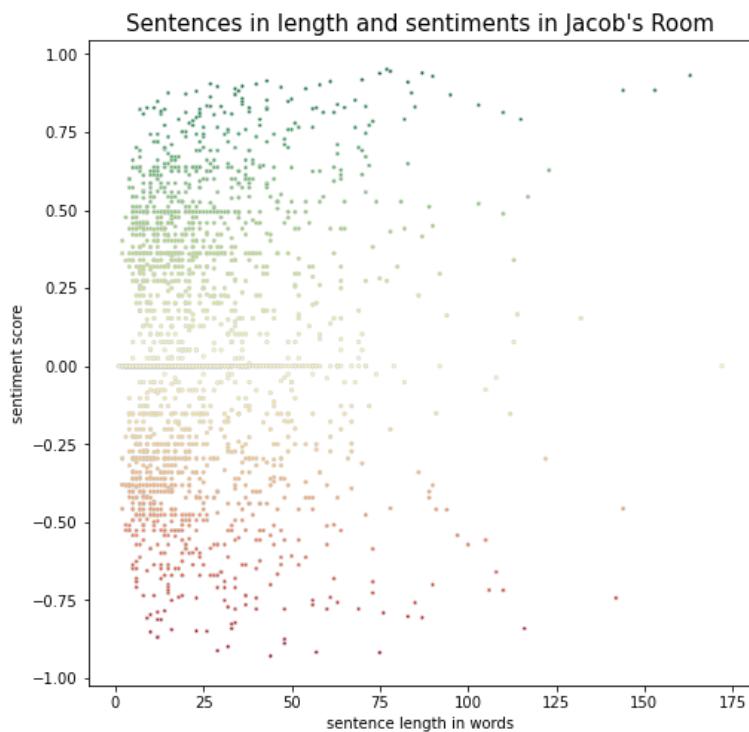
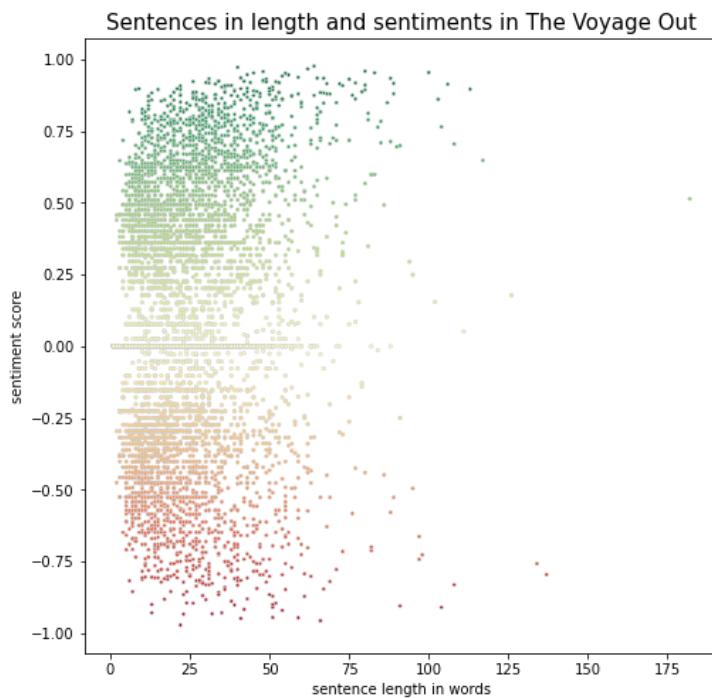


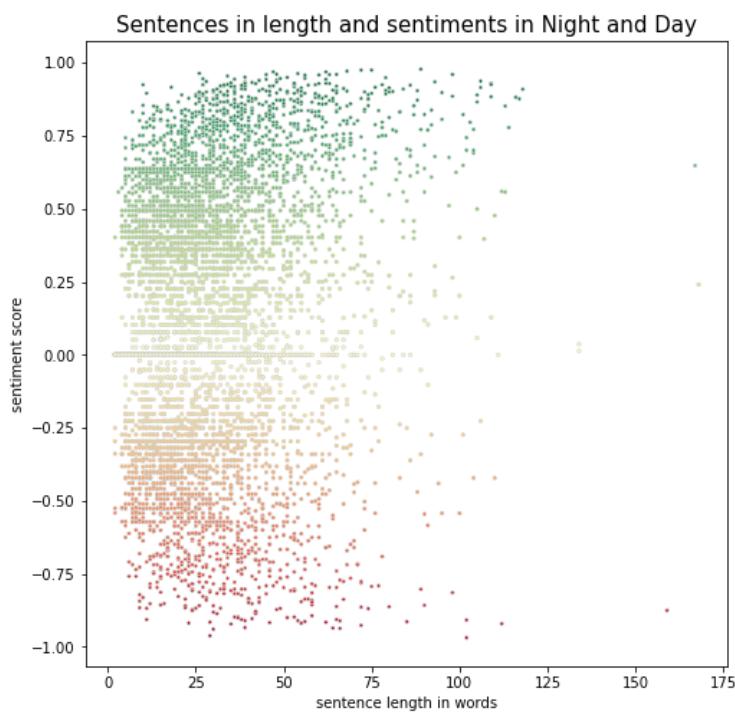
Fig. 26. Sentences in length and sentiments in Jacob's Room.



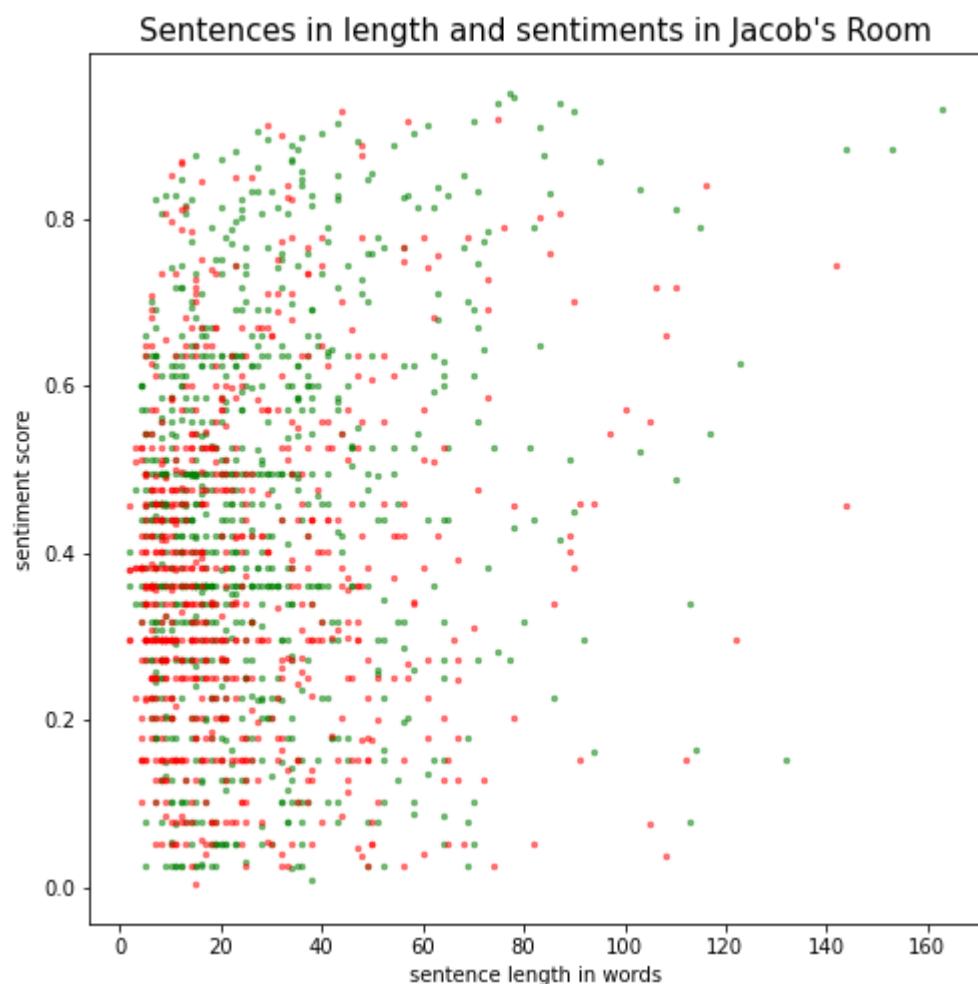
*Fig. 27. Sentences in length and sentiments in The Voyage Out.*



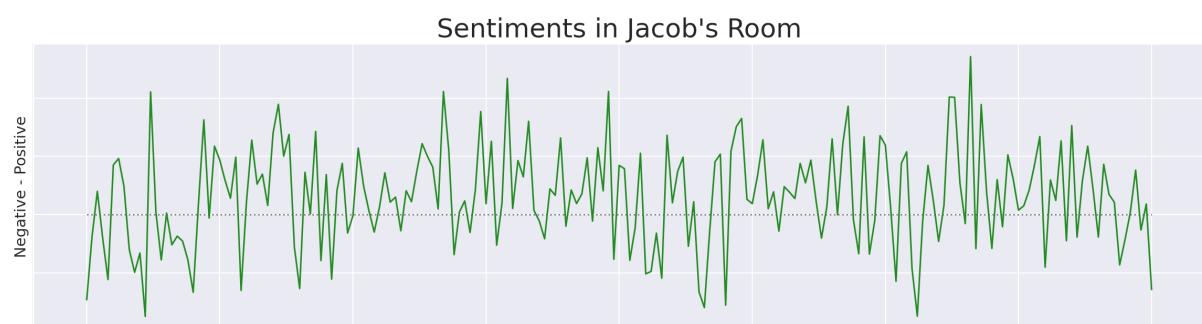
*Fig. 28. Sentences in length and sentiments in Night and Day.*



*Fig. 29. Sentences in length and sentiments Jacob's Room, version 2.*



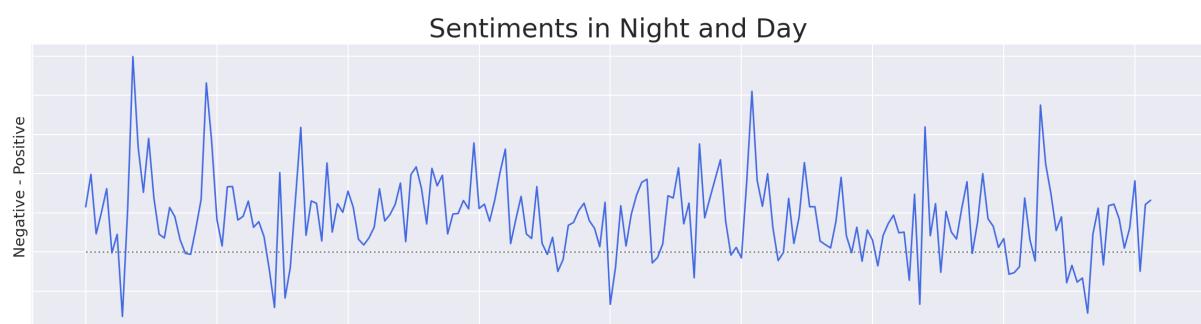
*Fig. 30. Sentiments in Jacob's Room.*



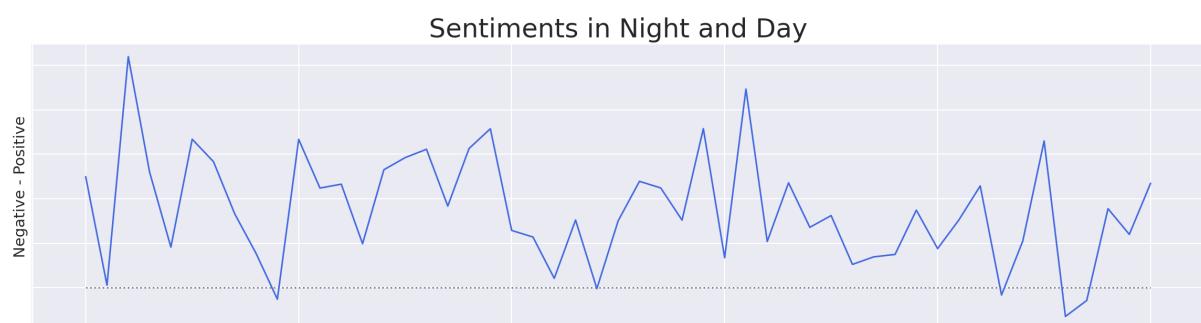
*Fig. 31. Sentiments in Jacob's Room, averaged.*



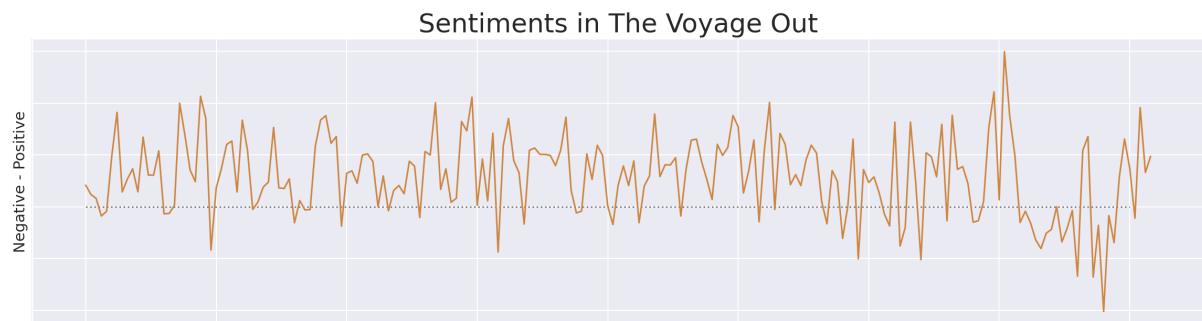
*Fig. 32. Sentiments in Night and Day.*



*Fig. 33. Sentiments in Night and Day, averaged.*



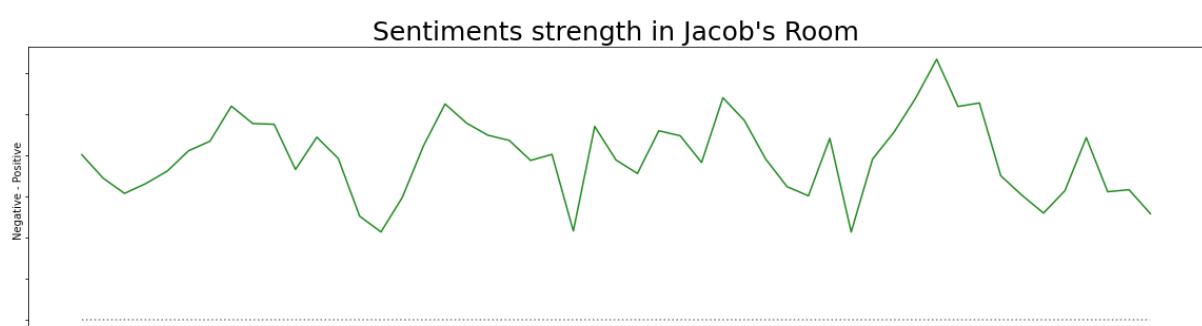
*Fig. 34. Sentiments in The Voyage Out.*



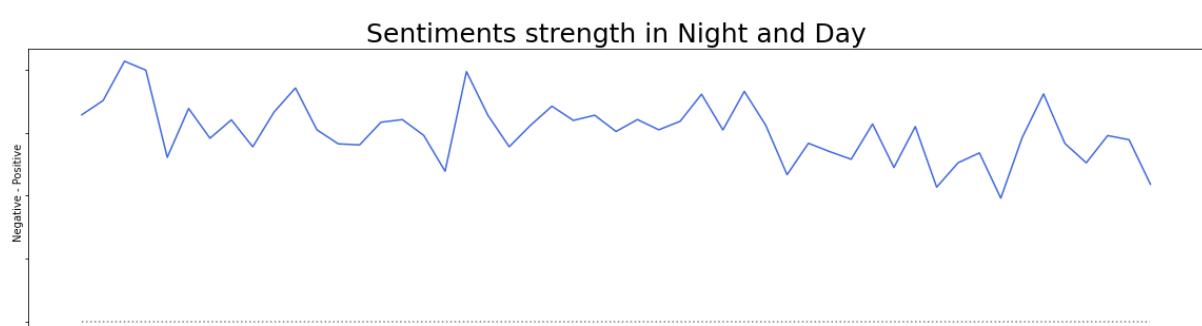
*Fig. 35. Sentiments in The Voyage Out, averaged.*



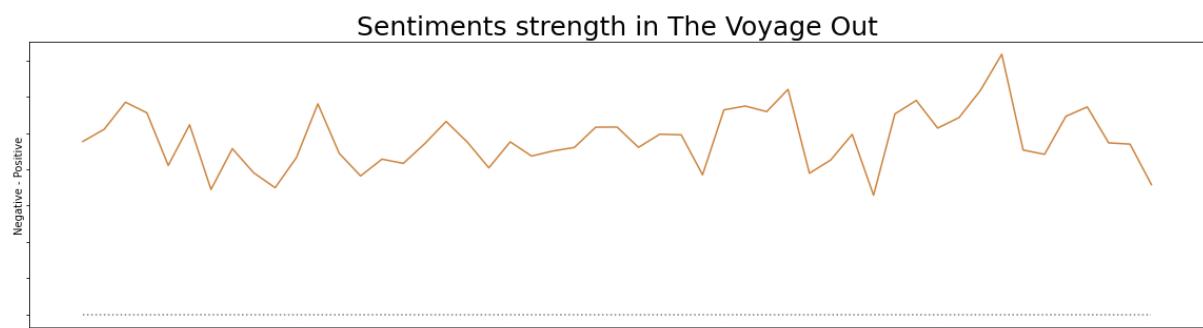
*Fig. 36. Sentiments Strength in Jacob's room.*



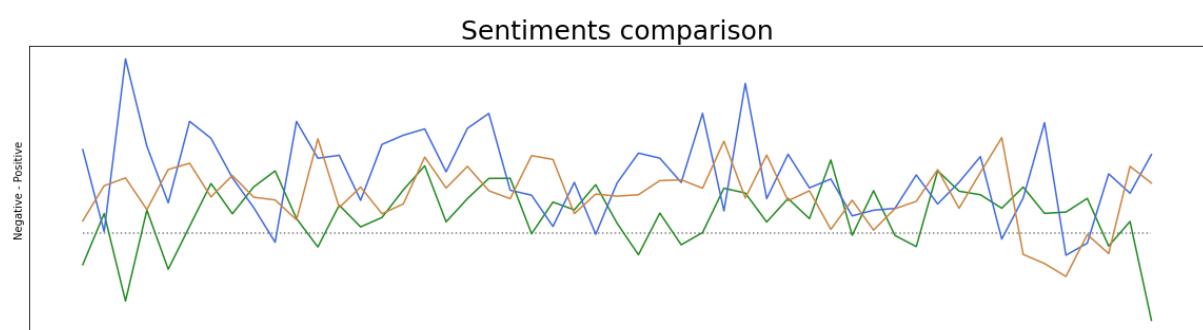
*Fig. 37. Sentiments Strength in Night and Day.*



*Fig. 38. Sentiments Strength in The Voyage Out.*



*Fig. 39. Sentiments Comparison.*



*Fig. 40. Sentences colored by sentiments in Jacob's Room.*



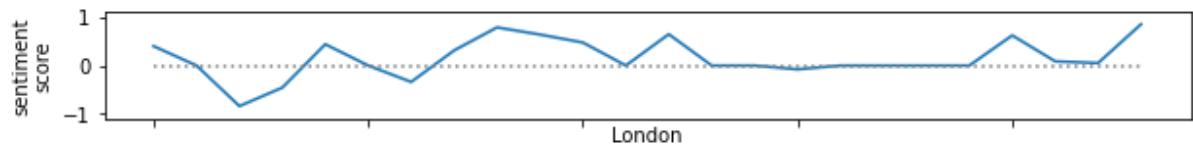
*Fig. 41. Sentences colored by sentiments in Night and Day.*



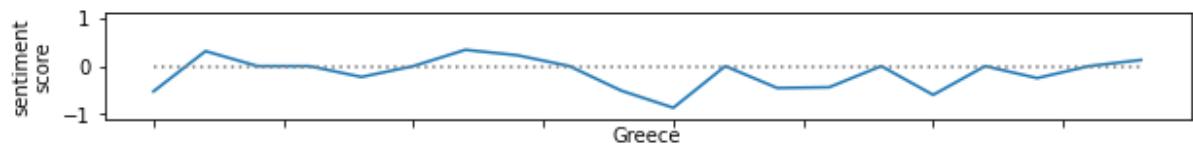
*Fig. 42. Sentences colored by sentiments in The Voyage out.*



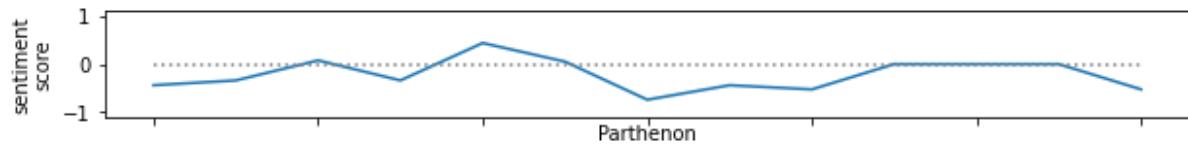
*Fig. 43. Sentiment score of sentences mentioning 'London' in Jacob's Room.*



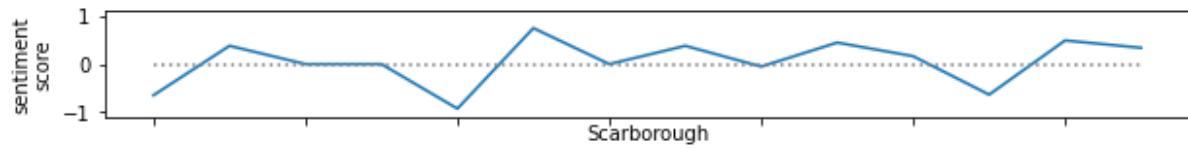
*Fig. 44. Sentiment score of sentences mentioning 'Greece' in Jacob's Room.*



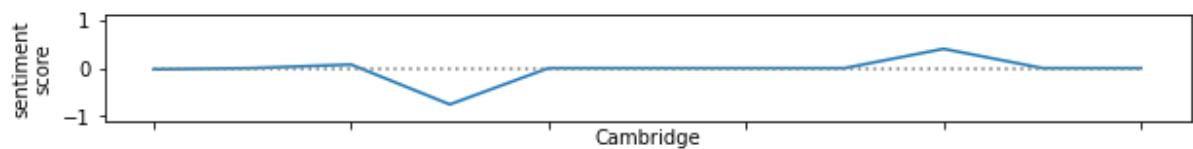
*Fig. 45. Sentiment score of sentences mentioning 'Parthenon' in Jacob's Room.*



*Fig. 46. Sentiment score of sentences mentioning 'Scarborough' in Jacob's Room.*



*Fig. 47. Sentiment score of sentences mentioning 'Cambridge' in Jacob's Room.*



*Fig. 48. Sentiment score of sentences mentioning 'Athens' in Jacob's Room.*

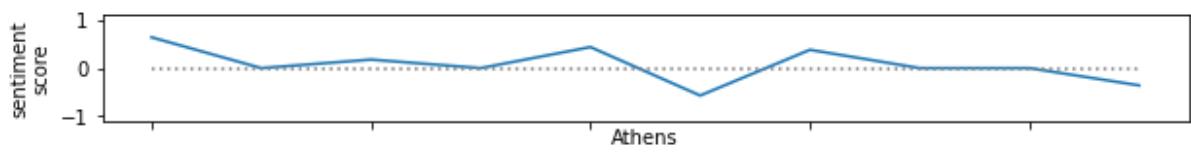


Fig. 49. Sentiment score of sentences mentioning 'the British Museum' in Jacob's Room.

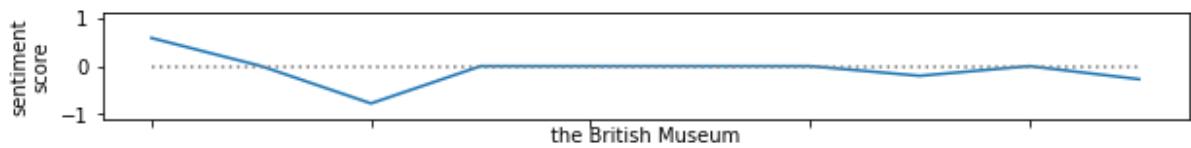


Fig. 50. Sentiment score of sentences mentioning 'England' in Jacob's Room.

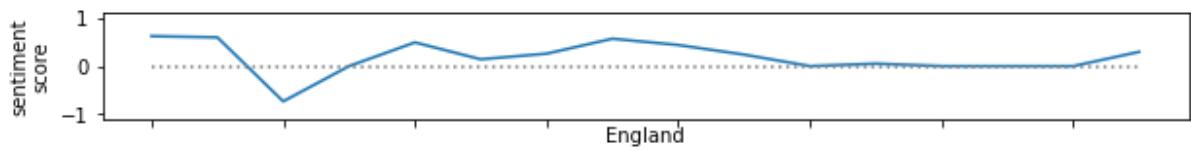


Fig. 51. Sentiment score of sentences mentioning 'Paris' in Jacob's Room.

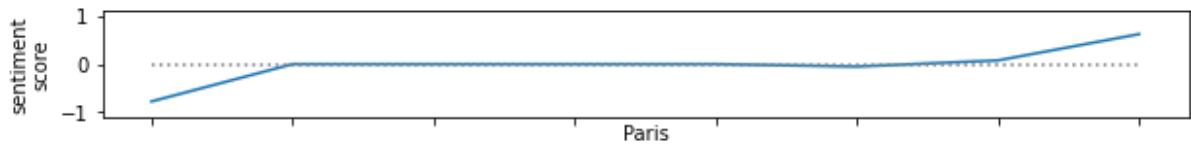


Fig. 52. Map based on sentiment scores of Jacob's Room.

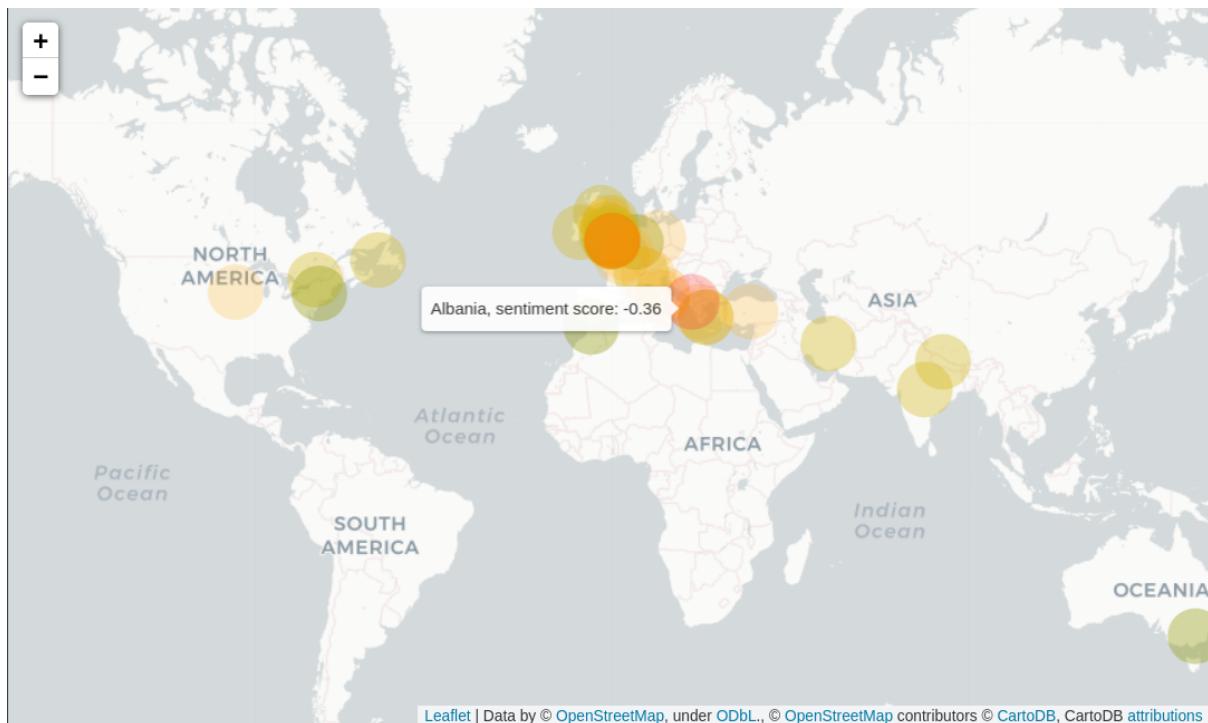


Fig. 53. Example of 'Albania' location in map.

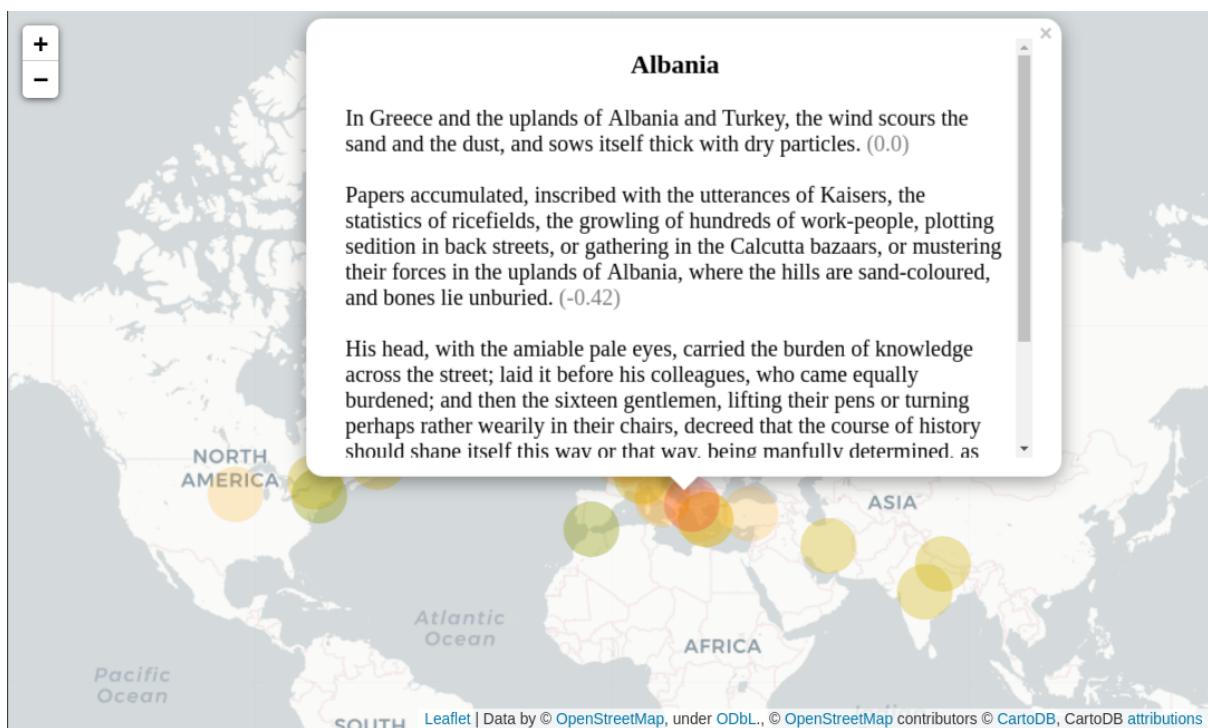


Fig. 54. Map 2.

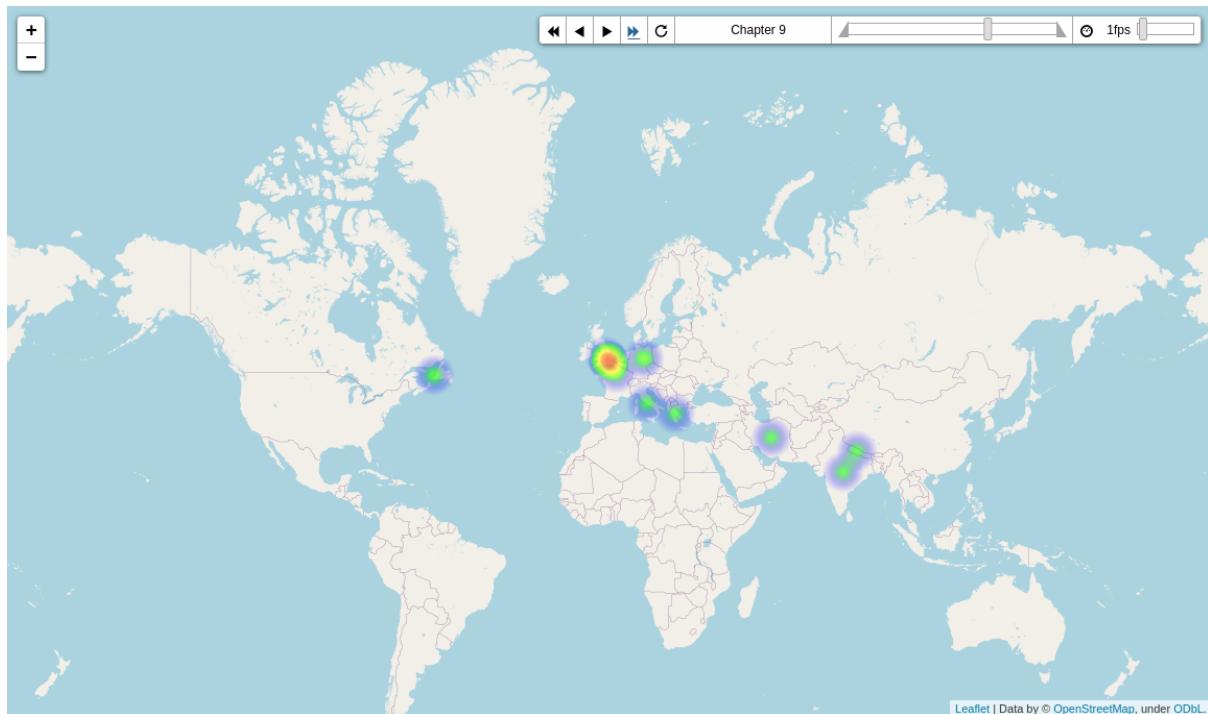
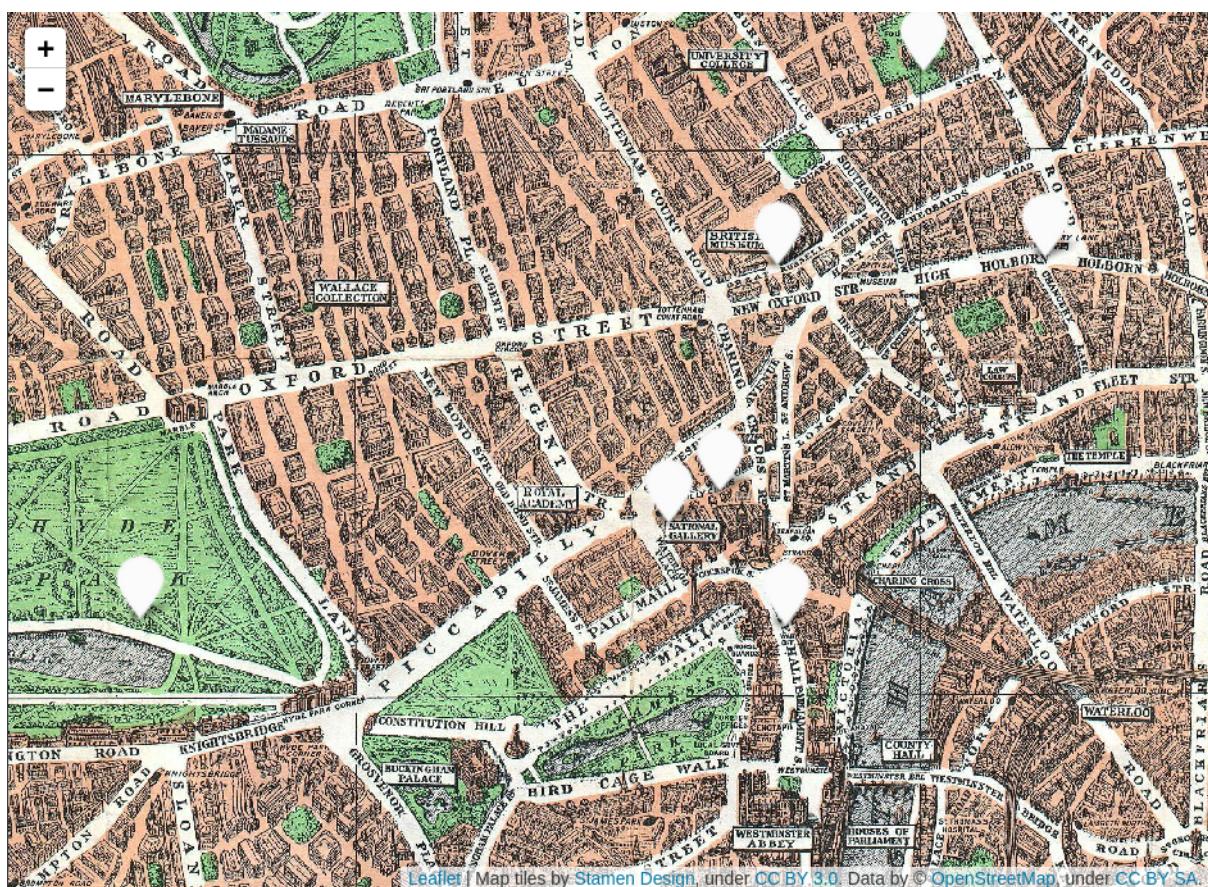
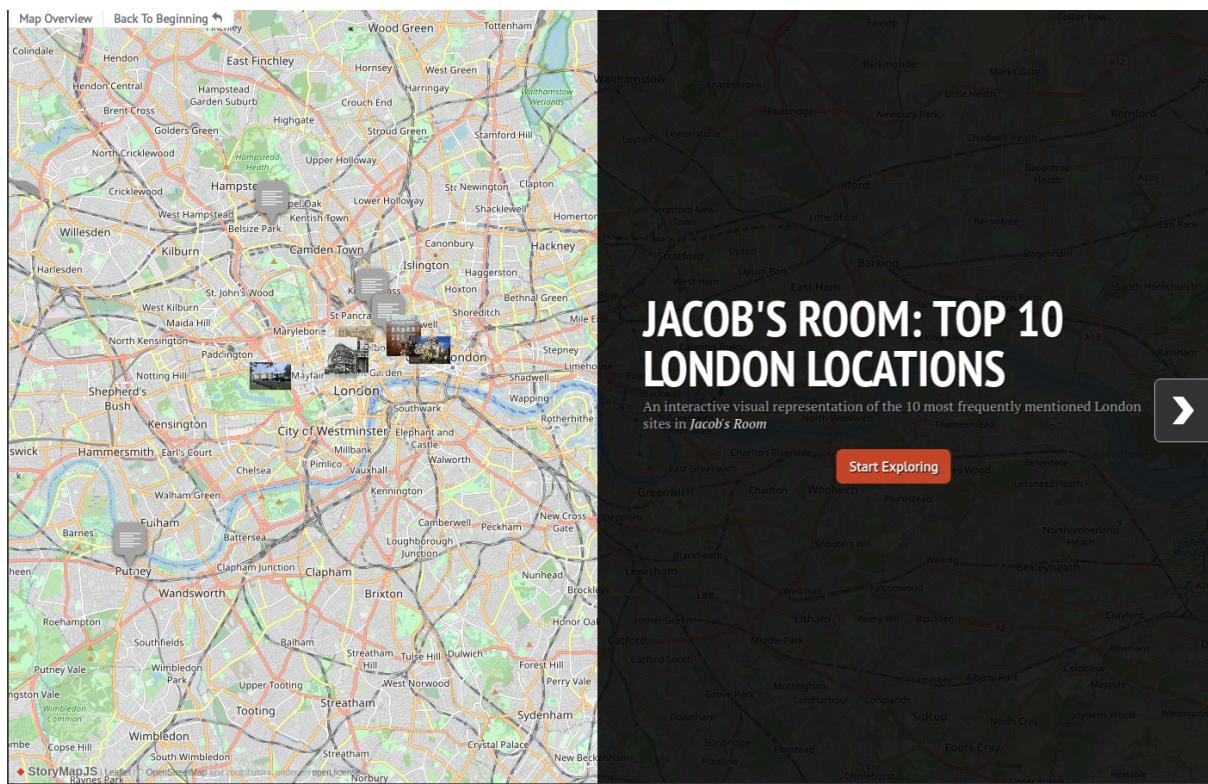


Fig. 55. Example of a 20th century map of London, showing locations in Jacob's Room.

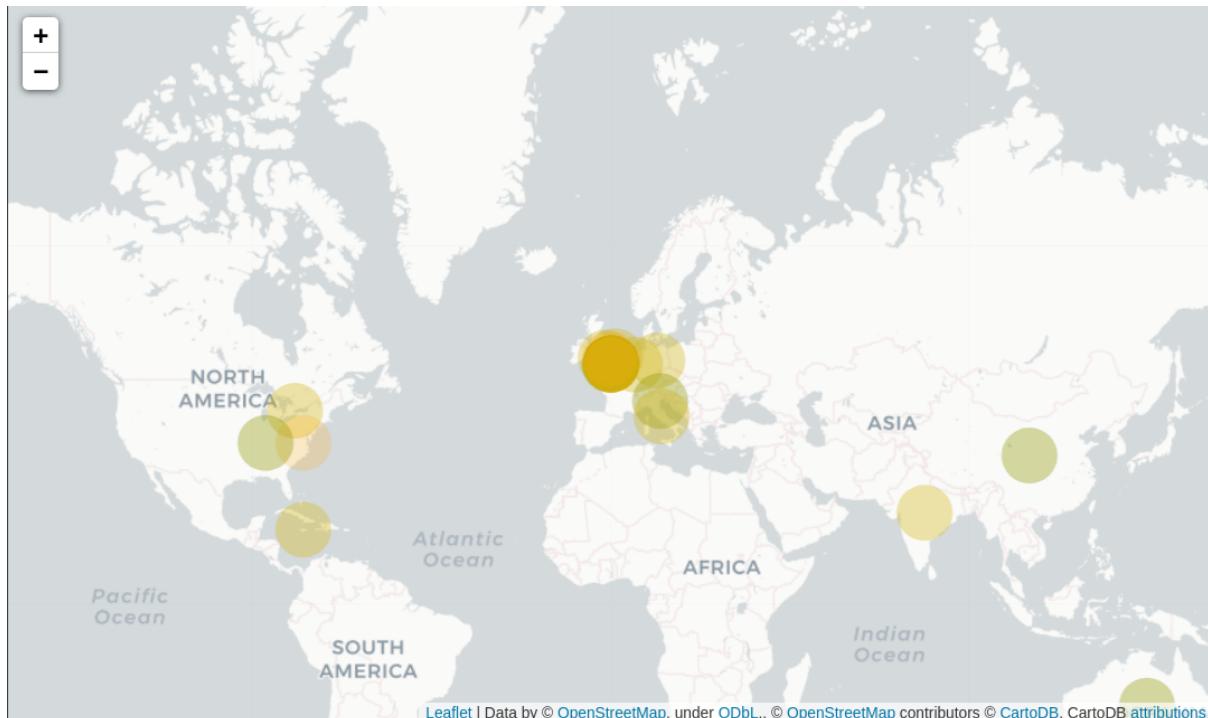


*Fig. 56. Example of a 20th century map of London on the Jacob's Room project website.*



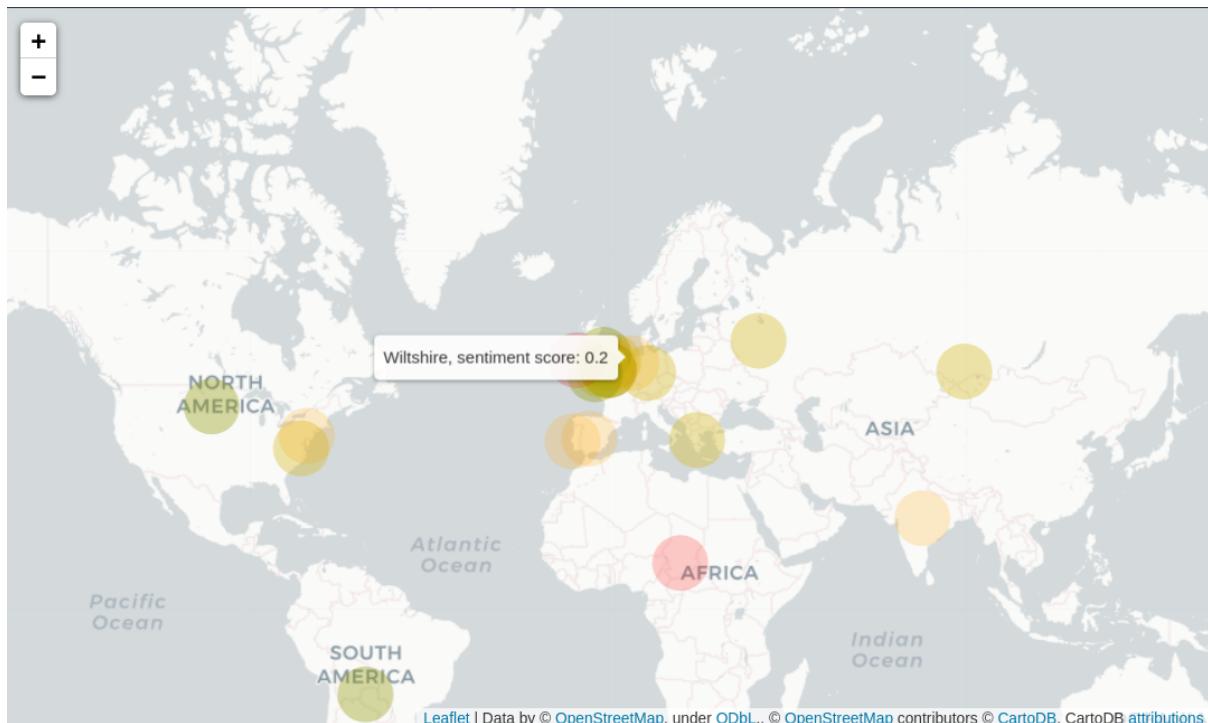
*Fig. 57. Map based on sentiment scores in Night and Day.*

## Night and Day:

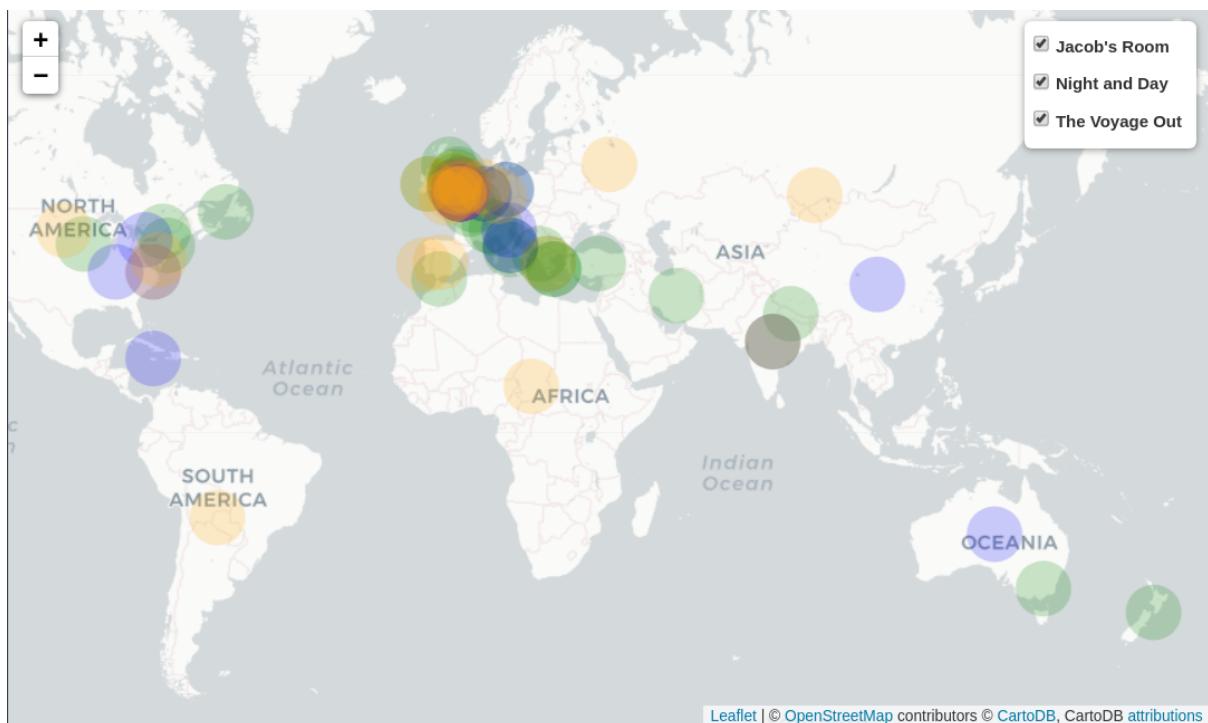


*Fig. 58. Map based on sentiment scores of The Voyage Out.*

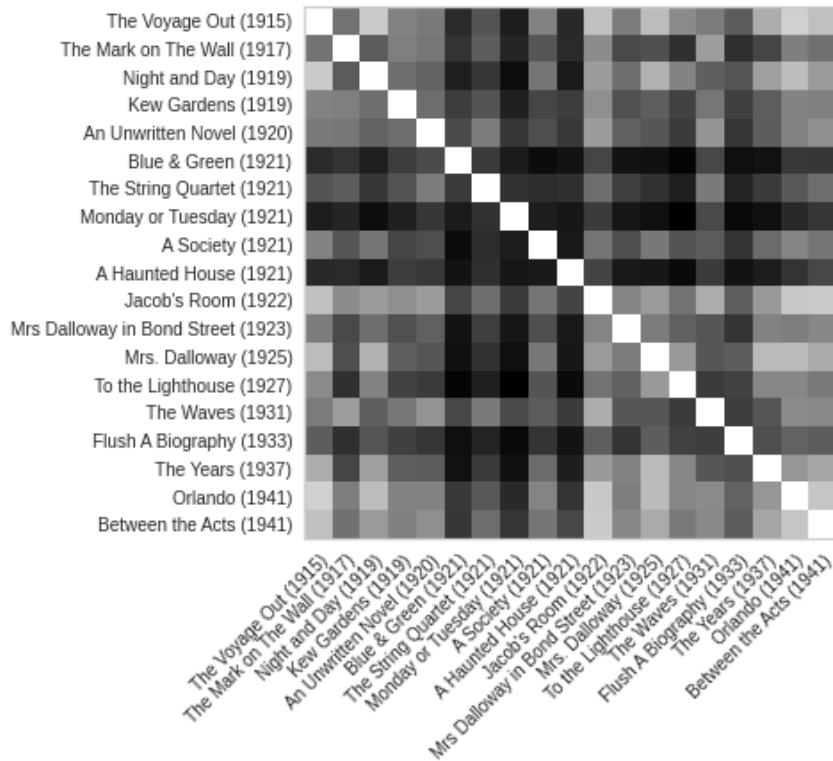
The Voyage Out:



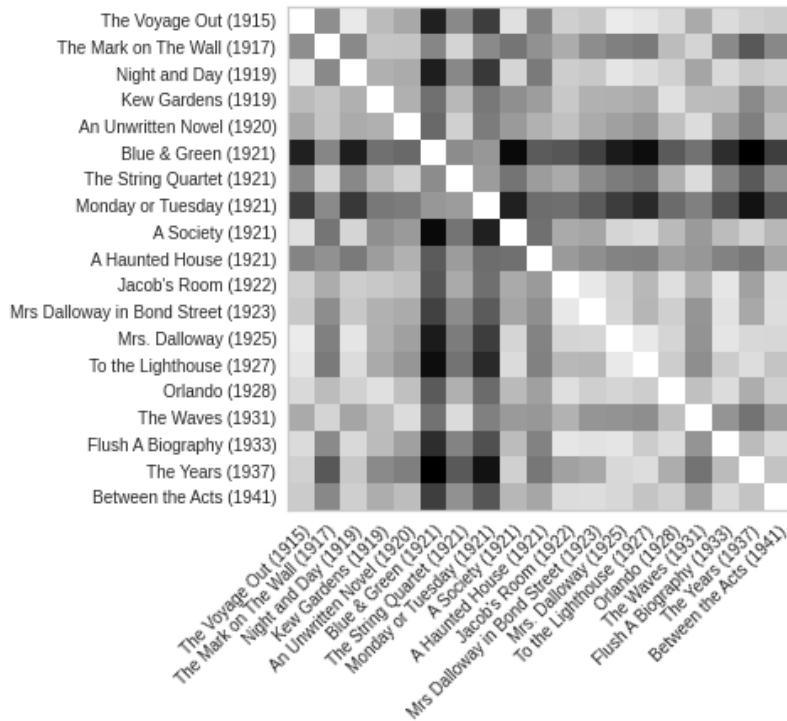
*Fig. 59. Overlap map of sentiment scores in Jacob's Room, Night and Day, and The Voyage out.*



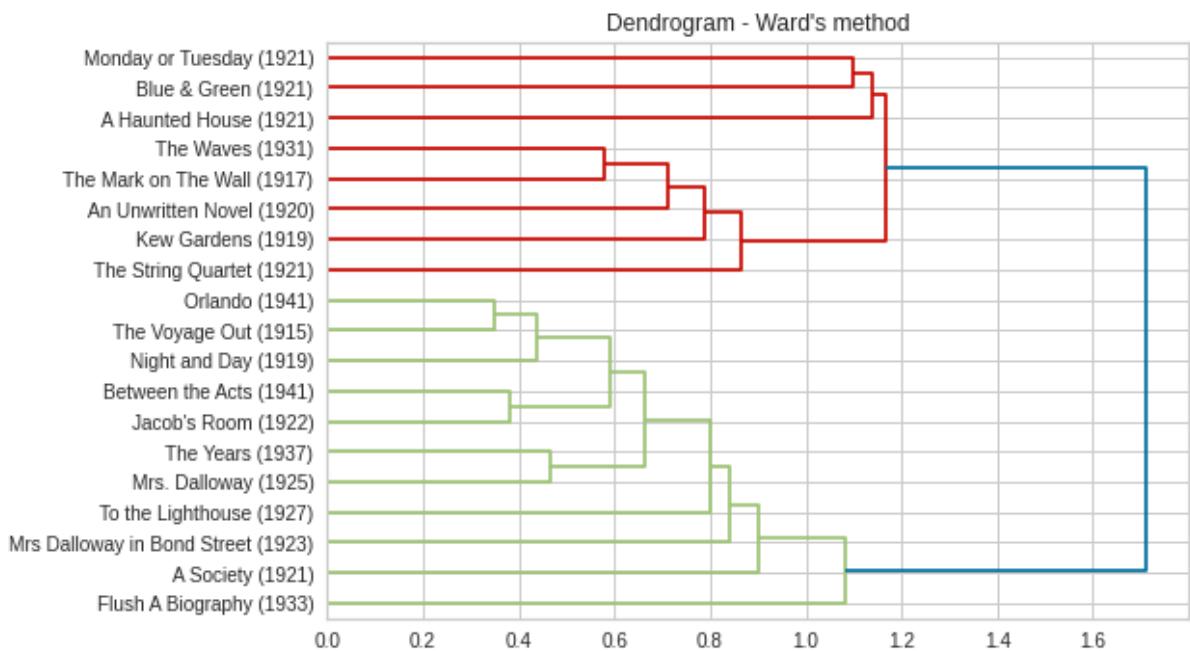
*Fig. 60. Comparison of the words between novels*



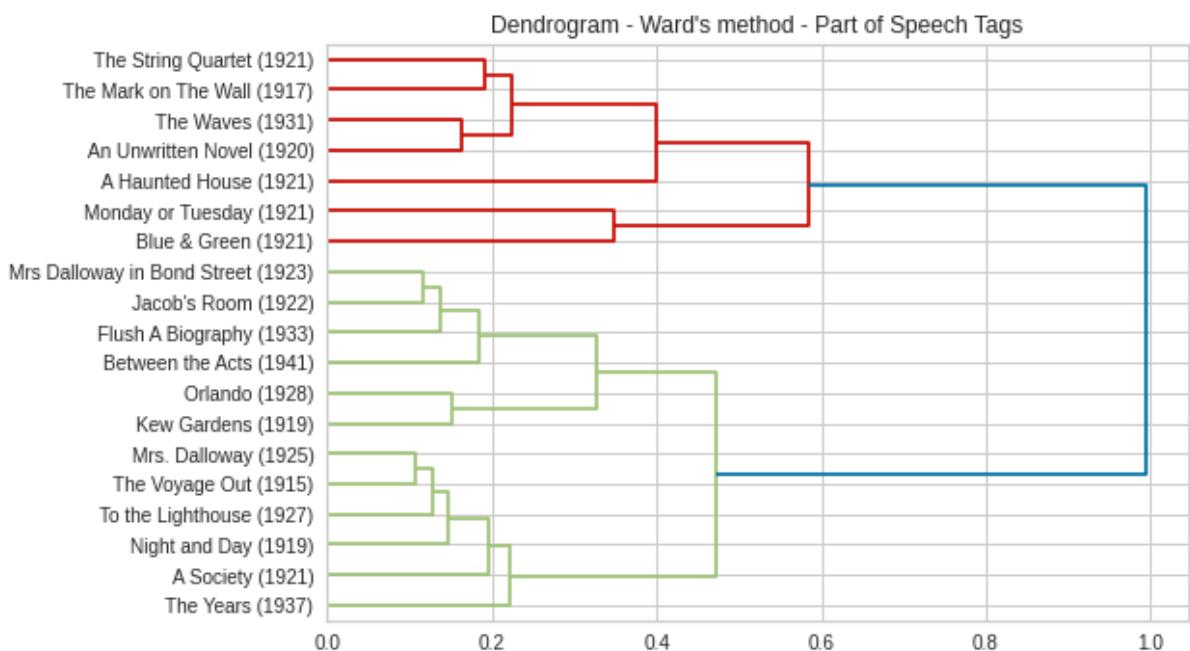
*Fig. 61. Comparison of the trigrams between novels*



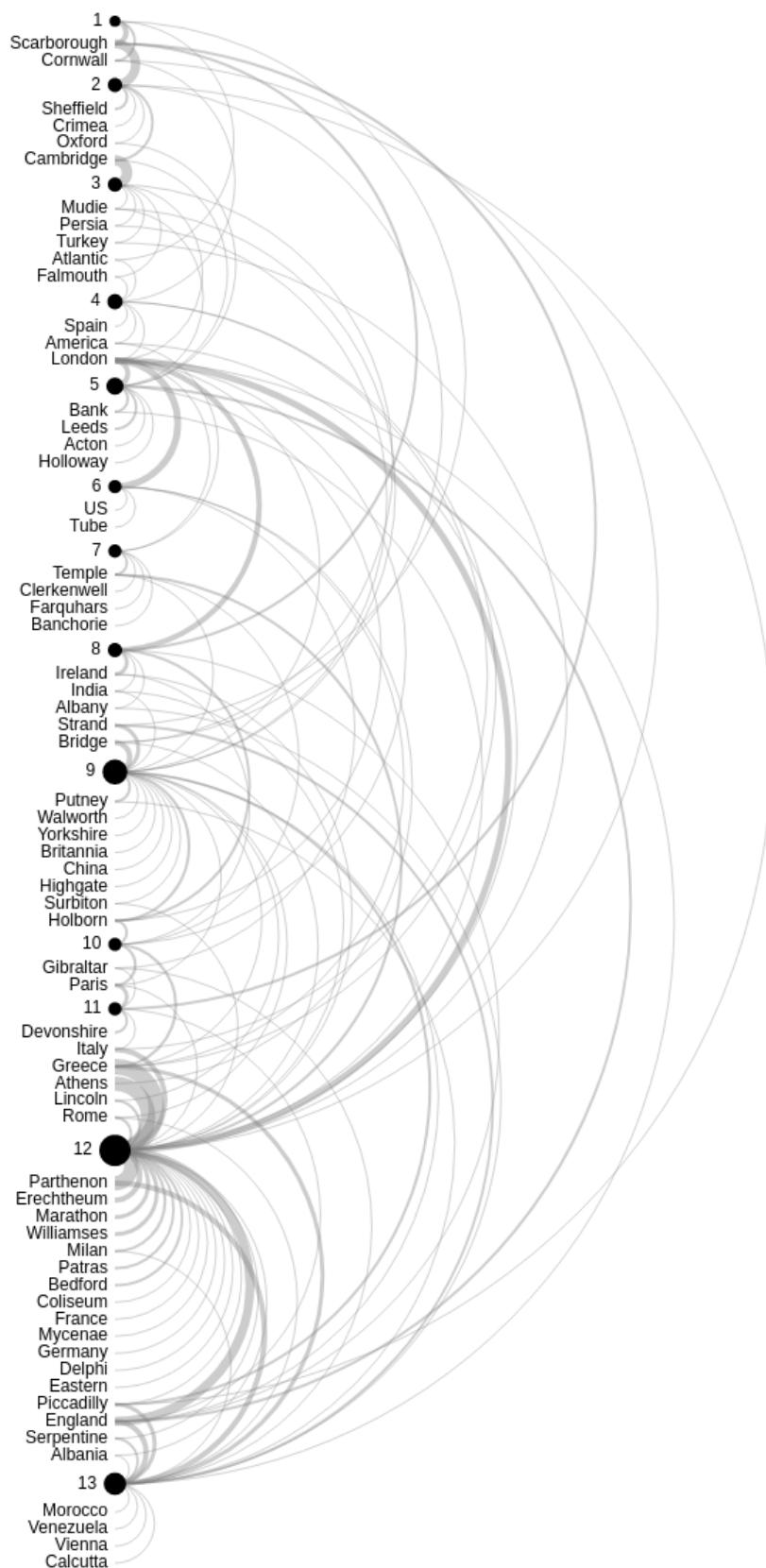
*Fig. 62. Dendrogram using Ward's Method based on trigrams.*



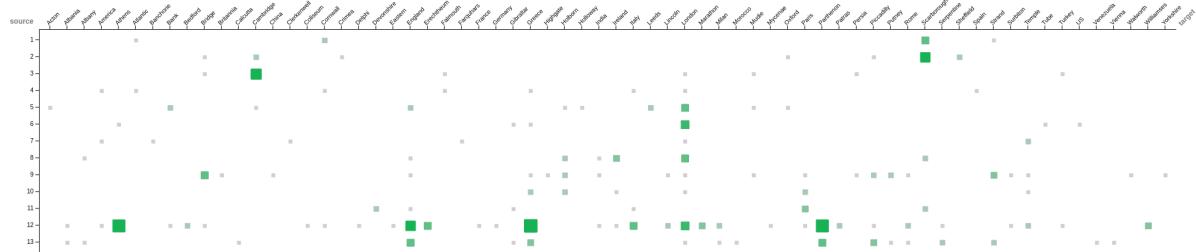
*Fig. 63. Dendrogram using Ward's Method, based on part of speech tags.*



*Fig. 64. Alternative visualization for where locations are mentioned in the chapters of Jacob's Room, with the weight of each line representing the amount of mentions.*



*Fig. 65. Alternative visualization for where locations are mentioned in the chapters of Jacob's Room, with the weight of each block representing the amount of mentions.*



*Fig. 66. Alternative visualization for where locations are mentioned in the chapters of Jacob's Room, with the weight of each location representing the amount of mentions.*

