information

MDPI

Article

Correlations and Fractality in Sentence-Level Sentiment Analysis Based on VADER for Literary Texts

Ricardo Hernández-Pérez [1,†], Pablo Lara-Martínez [2], Bibiana Obregón-Quintana [2], Larry S. Liebovitch [3,4] and Lev Guzmán-Vargas [5,*]

1 Escuela Superior de Física y Matemáticas, Instituto Politécnico Nacional, Edif. 9, 2o. Piso, UP Zacatenco, Mexico City 07738, Mexico
2 Facultad de Ciencias, Universidad Nacional Autónoma de México, Mexico City 04510, Mexico; p4blolara@ciencias.unam.mx (P.L.-M.); bobregon@ciencias.unam.mx (B.O.-Q.)
3 Department of Physics, Queens College, City University of New York, New York, NY 11367, USA; lsl2140@columbia.edu
4 Advanced Consortium on Cooperation, Conflict and Complexity (AC4), Climate School, Columbia University, New York, NY 10027, USA
5 Unidad Interdisciplinaria en Ingeniería y Tecnologías Avanzadas, Instituto Politécnico Nacional, Av. IPN No. 2580, L. Ticomán, Mexico City 07340, Mexico
* Correspondence: lguzmanv@ipn.mx
† Sadly, Ricardo passed away on 11 October 2024. This is our tribute to our dear friend, colleague and exceptional scientist.

Abstract: We perform a sentence-level sentiment analysis study of different literary texts in English language. Each text is converted into a series in which the data points are the sentiment value of each sentence obtained using the sentiment analysis tool (VADER). By applying the Detrended Fluctuation Analysis (DFA) and the Higuchi Fractal Dimension (HFD) methods to these sentiment series, we find that they are monofractal with long-term correlations, which can be explained by the fact that the writing process has memory by construction, with a sentiment evolution that is self-similar. Furthermore, we discretize these series by applying a classification approach which transforms the series into a one on which each data point has only three possible values, corresponding to positive, neutral or negative sentiments. We map these three-states series to a Markov chain and investigate the transitions of sentiment from one sentence to the next, obtaining a state transition matrix for each book that provides information on the probability of transitioning between sentiments from one sentence to the next. This approach shows that there are biases towards increasing the probability of switching to neutral or positive sentences. The two approaches supplement each other, since the long-term correlation approach allows a global assessment of the sentiment of the book, while the state transition matrix approach provides local information about the sentiment evolution along the text.

Keywords: sentiment analysis; opinion mining; social systems

check for updates

Citation: Hernández-Pérez, R.; Lara-Martínez, P.; Obregón-Quintana, B.; Liebovitch, L.S.; Guzmán-Vargas, L. Correlations and Fractality in Sentence-Level Sentiment Analysis Based on VADER for Literary Texts. *Information* **2024**, *15*, 698. https://doi.org/10.3390/info15110698

Academic Editor: Fei Liu

Received: 2 October 2024
Revised: 18 October 2024
Accepted: 21 October 2024
Published: 4 November 2024

1. Introduction

Sentiment analysis, or opinion mining, is an active research area in the field of natural language processing that analyzes people's opinions that are loaded with sentiments and emotions via the computational treatment of subjectivity in text that is produced in different platforms [1–5]. Sentiment analysis mainly considers two aspects: categorical sentiment analysis and dimensional sentiment analysis. Categorical sentiment analysis basically classifies emotions under categories or labels based on the fact that these emotions can be classified using a low number of basic emotions. In contrast, in dimensional analysis, sentiment can be represented by emotions along continuous dimensions rather than discrete categories [6,7]. In general, two basic dimensions are included: valence and intensity. Sentiment analysis makes use of various lexicons to quantify emotional intensity [8–13].

There are several lexicons built from very diverse sources, such as the Social Contextualized Affect Lexicon (SO-CAL), SentiWordNet, AFINN, and ANEW, to name a few [14–17]. SO-CAL aims to analyze sentiment in social network contexts, captures how emotions are expressed in informal language, and evaluates emotions in terms of valence, arousal, and dominance [18–20]. AFINN focuses on evaluating conversations in social network posts and assigns words a score ranging from negative to positive, providing a direct numerical sentiment score [21]. SentiWordNet is an extension of WordNet that provides three scores for each word: positive, negative, and objective, allowing for a nuanced understanding of word sentiment [22,23]. ANEW provides emotional ratings of words, including valence, arousal, and dominance values, and is used in the context of emotional profiling analysis [24–27].

A recent study conducted an extensive literature survey on sentiment analysis to identify existing sentence-level methods covering several different techniques in order to perform a comparison with each other to identify their advantages, disadvantages, and limitations [28]. The authors of that study suggested that there is no single method that always achieves the best prediction performance for all different datasets. However, they found that the method called VADER (for Valence Aware Dictionary for sEntiment Reasoning) was the most consistent method in different experiments that were performed. VADER uses a combination of qualitative and quantitative methods to produce and then empirically validate a sentiment lexicon that is especially attuned to microblog-like contexts. Moreover, after incorporating some lexical rules that embody grammatical and syntactical conventions that humans use when expressing or emphasizing sentiment intensity, the accuracy of the sentiment analysis engine improves across several domain contexts (social media text, NY Times editorials, movie reviews, and product reviews) [29]. Therefore, we select VADER to perform the sentence-level sentiment analysis in the present work.

On the other hand, previous studies performed on different corpora in several languages taken mainly from Web-based content (online forums, blogs, Twitter, etc.) have found that positive words carry less information [30] and that words with a positive emotional content are more frequently used [30,31]. However, these previous studies have been centered in a word-by-word sentiment analysis. Are these conclusions still valid if the analysis is performed at a higher level, i.e., at sentence level?

Our goal is to analyze the evolution of the sentiment along the text by examining the statistics associated to the sentiment value of each sentence and then assess the presence of correlations. We find that, for all the texts analyzed, the correlations follow power laws with exponent values that differ from the randomized case, confirming the existence of long-term memory. We also quantify the dependence of sentiment values on transitions to positive, negative, or neutral, and find that either neutral or positive sentences are more likely to be followed by neutral or positive ones as well. We focus on single-author works for which the writing dynamics differ from those of Web-based content (online forums, blogs, Twitter, etc.) since the authors take more time to conclude the text after reviewing it as part of the writing process and, oftentimes, the texts are also subject to the editor's review process. Therefore, these literary works are subject to a longer production process, implying that the selection of words in the sentences is more thoroughly performed. Thus, our approach supplements the previous studies based on a high number of authors in online opinion forums or Twitter, and then this allows to compare the behavior of an ensemble of authors on shorter texts versus the longer texts produced by single authors under more scrutiny.

This paper is organized as follows: in Section 2, we present the related work. In Section 3, we describe briefly the sample of books that were considered. In Section 4, we describe the methods to obtain series of sentiment values from the texts which are processed to investigate correlations. In Section 5, we discuss the results. Finally, the concluding remarks are given in Section 6.

2. Related Work

Sentiment analysis in literary texts is a developing field that leverages computational techniques to understand and interpret the emotional dimensions that appear in literary texts. As is widely recognized, emotions play a transcendental role in narratives or stories developed by authors, shaping characters' motivations and influencing readers. The study of emotions in literature has been modified to include more quantitative analyses through computational methods, which has allowed for more extensive and systematic investigations of how emotions are displayed through a literary text. VADER offers an important tool for sentiment analysis within literary studies, as it allows the emotional dimensions of texts to be quantified. It has been noted that it also has certain limitations—especially in relation to sarcasm and domain specificity—but its ease of use and effectiveness make it an attractive option for both qualitative and quantitative literary analysis. As the field has progressed, integrating tools such as VADER into literary research can enhance our understanding of emotional dynamics in literature. For example, Bizzoni et al. [32] explored the complexities of analyzing sentiment in narrative texts such as those of E. Hemingway. Yeruva et al. [33] analyzed data from literary texts related to human feeling in classical Greek tragedy. More recently, S. Vinodini [34] conducted a sentiment analysis of selected novels by the celebrated Brazilian author Paulo Coelho.

On the other hand, long-range correlations are characterized by a relationship or memory between subsequences separated by significant distances. In the context of literature, this can be understood as recurrence of ideas or emotional motifs that recur throughout a text, regardless of their physical distance in the narrative. The presence of correlations in literary texts has been reported in various contexts such as word lengths [35–37], sentences [38], and sentiment trajectories [39]. Also, it has been reported that the temporal organization of written texts can be analyzed from a multifractal perspective [40,41].

3. Input Data

The input data are comprised of 50 ebooks in English language downloaded from the websites of the Gutenberg Project (http://www.gutenberg.org/, last accessed date: 27 May 2024) and the Project Gutenberg Australia (http://gutenberg.net.au/, last accessed date: 27 May 2024). The titles of the written texts are described online at https://figshare.com/articles/dataset/A_sentence-level_sentiment_analysis_of_some_literary_texts_inEnglish_language/27092008, last accessed date: 30 September 2024. There was no particular strategy to select the titles. The selected books were first published in different epochs and countries, therefore offering diversity in time and cultures of origin. We remove the preface, table of contents, etc., of the books and only retain the text.

4. Methods

4.1. Vader Sentiment Values

VADER provides four scores for the sentiment and intensity of a sentence: positive (`pos`), neutral (`neu`), negative (`neg`), and compound (`com`). The `pos`, `neu`, and `neg` scores are ratios for proportions of text that fall in each category; therefore, these should all add up to be 1. These are the most useful metrics for multidimensional measures of sentiment for a given sentence. Moreover, the `compound` score is computed by summing the valence scores of each word in the lexicon, adjusted according to certain rules, and then normalized to be a floating number between -1 (most extreme negative) and $+1$ (most extreme positive) [29]. This is the most useful metric for a single unidimensional measure of sentiment for a given sentence.

Using the `compound` score, standardized thresholds can be set for classifying sentences as either positive, neutral, or negative. Typical threshold values are [29]:

- positive sentiment: $\texttt{compound} \geq 0.05$,
- neutral sentiment: $-0.05 < \texttt{compound} < 0.05$,
- negative sentiment: $\texttt{compound} \leq -0.05$.

In the present work, we choose the `compound` score since we are interested in a single unidimensional measure of sentiment for each sentence in order to study correlations of sentiment along the text as well as the evolution of sentiment by studying the changes in sentiment from one sentence to the next.

In order to estimate the sentiment of the sentences in a text, we start by parsing the text to split it into sentences. Each sentence is then evaluated with VADER to obtain its compound sentiment score, and therefore, the text is mapped to a data series of sentiment values. Figure 1 shows an example of these sentiment series for an specific book. The complete resulting corpus containing the sentiment scores for each book can be found online at https://figshare.com/articles/dataset/A_sentence-level_sentiment_analysis_of_some_literary_texts_inEnglish_language/27092008, last accessed date: 30 September 2024. These series are then processed to investigate the kind of correlations that they contain, for which we apply the well-known Higuchi Fractal Dimension (HFD) [42] and the Detrended Fluctuation Analysis (DFA) [43] methods.

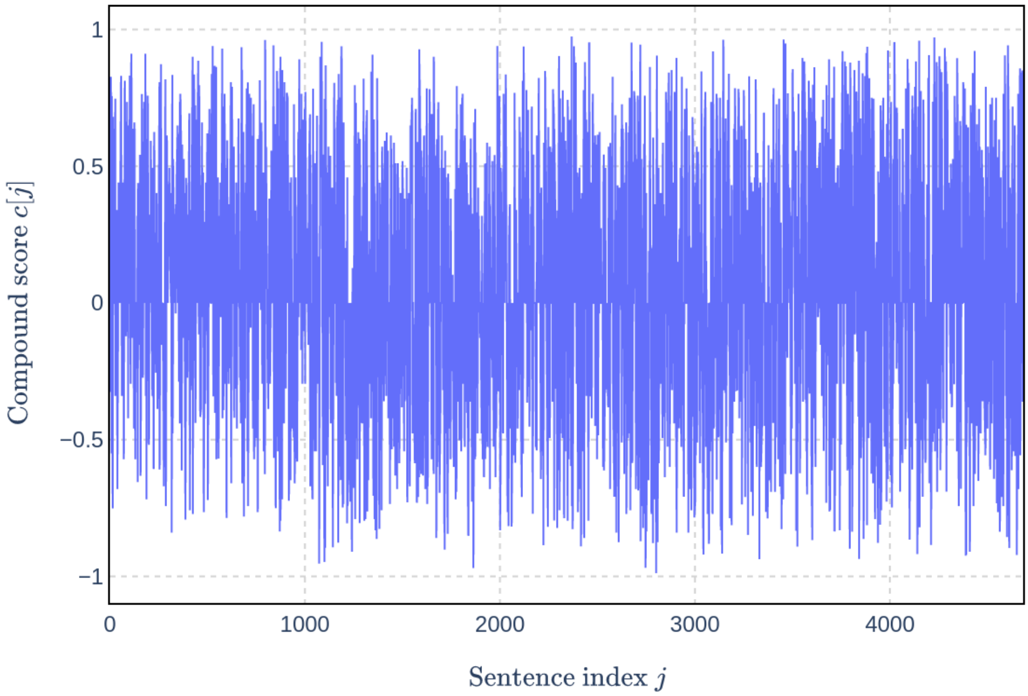

Figure 1. Example of one of the data series of sentiment values estimated with VADER. In this case, for the book The Adventures of Sherlock Holmes by Arthur Conan Doyle.

Moreover, we use the classification threshold mentioned previously to transform each series of compound sentiment into a series of three possible discrete values (states) of sentiment: positive, neutral, or negative (see Figure 2). In order to study the sentiment transitions along the text, we count the number of transitions of each type. For instance, let us suppose that the kth sentence is classified as positive and the $k+1$th is classified as negative, then the number of transitions of the type $\texttt{pos} \to \texttt{neg}$ is increased by one. Once the counting of transitions is finalized, we normalize the matrix to obtain a state transition matrix $\mathbf{p}$ for each one of the books in our sample:

$$\mathbf{p} = \begin{pmatrix} p_{\texttt{pos}\to\texttt{pos}} & p_{\texttt{pos}\to\texttt{neu}} & p_{\texttt{pos}\to\texttt{neg}} \\ p_{\texttt{neu}\to\texttt{pos}} & p_{\texttt{neu}\to\texttt{neu}} & p_{\texttt{neu}\to\texttt{neg}} \\ p_{\texttt{neg}\to\texttt{pos}} & p_{\texttt{neg}\to\texttt{neu}} & p_{\texttt{neg}\to\texttt{neg}} \end{pmatrix} \tag{1}$$

Therefore, we map the evolution of the classified sentiment to a Markov chain for each book, with the state transition coefficients given by Equation (1). Shannon used Markov chains to predict sequences of words, observing that a reasonable approximation to English could be produced if each word was chosen based not just on the previous word but on the last few words, and introduced a mathematical framework for analyzing the information provided by the word [44]. Also, Markov chains have been used to compute how probable each instance of a word is based on the last few words, providing a way to measure the predictability of a word in its context [45]. Interestingly, Markov himself published a report in which he computes the probabilities that a vowel is followed by a vowel or a consonant in a text excerpt containing 20,000 Russian letters of the Russian alphabet from Pushkin's novel [46].

In the present work, we are interested not in the word length, but in the sentiment of the sentence. Given that we wrote a sentence with certain sentiment, what is the probability that the next sentence either switches to another sentiment state or stays on the same one? How does the sentiment of a certain sentence relate to the sentiment of the previous few sentences?

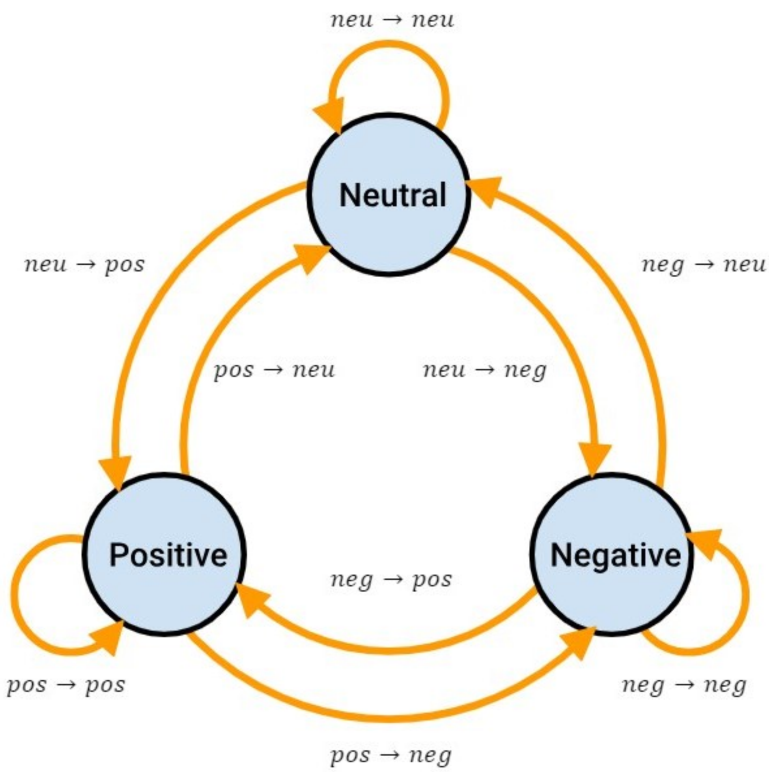

Figure 2. Diagram depicting the possible transitions between sentiment values of sentences in a text.

4.2. Fractal Dimension Method and Detrended Fluctuation Analysis

Fractals are complex patterns that show self-similarity at different scales and are characterized by a dimension that is not necessarily an integer value, which quantifies the complexity of the fractal pattern. In the case of time series, which are considered non-isotropic fractals because they extend over time while the values of the variable remain in a bounded range (auto-affinity), the fractal characterization can be performed using several methodologies. To capture this statistical self-affinity property, it is usually resorted to quantify a measure of the object at different scales. This leads to a power law ($G(x) \sim x^a$), where the scaling exponent a represents the level of spatio-temporal organization displayed by the time series. The power spectrum is the method traditionally used to characterize autocorrelations in time series. For example, we consider a stationary stochastic process with an autocorrelation function following a power law $C(s) \sim s^{\gamma}$, where s is the lag and γ is the correlation exponent, $0 < \gamma < 1$. The presence of long-range correlations is related to the fact that the mean correlation time diverges for infinite time series. According to the Wiener–Khintchin theorem, the power spectrum is the Fourier transform of the autocorrelation function $C(s)$, and for the case described above, we have the scaling relation $S(f) \sim f^{-\beta}$, where f is the frequency and β is called the spectral exponent and is related to the correlation exponent by $\gamma = 1 - \beta$. However, the estimation of β may not be as accurate when we have a nonstationary series. Alternative methods have been proposed for evaluating the correlations and fractal properties of stationary and nonstationary time series [42,47–49]. Here, we briefly describe two robust methods for the estimation of the fractal scaling exponents:

(i) Fractal Dimension Method. We evaluate the time organization in sequences of sentiment values by means of the Fractal Dimension Method (FDM), also known as Higuchi Fractal Dimension (HFD) method [50]. Briefly, we explain the main steps of the FDM [50,51]. Given the time series $x_1, x_2, ..., x_N$, we construct new time series x_m^k defined as $x_m, x_{m+k}, x_{m+2k}, ..., x_{(m+[\frac{N-k}{k}]\cdot k)}$, with $m = 1, 2, 3, ..., k$, [ ] denoting Gauss' notation, that is, the bigger integer and m and k are integers that indicate the initial time and the interval time, respectively. The length of the curve x_m^k, is defined as

$$L_m(k) = \frac{1}{k}\left[\left(\sum_{i=1}^{[\frac{N-m}{k}]} |x(m+ik) - x(m+(i-1)k)|\right)\frac{N-1}{[\frac{N-m}{k}]k}\right] \tag{2}$$

and the term $(N-1)/[(N-m)/k]k$ represents a normalization factor. The length of each sequence x_m^k is considered to construct the length of the original curve for the time interval k, $\langle L(k)\rangle$. Finally, if a scaling behavior of the form $\langle L(k)\rangle \propto k^{-D}$ is observed, then the curve is fractal with dimension D [50]. It is known that the fractal dimension is related to the spectral exponent β by means of $\beta = 5 - 2D$ [50]. We notice that this relationship is valid for $1 < D < 2$ and $1 < \beta < 3$, and thus FDM is not a reliable method for signals with strong anticorrelated behavior, that is, for $-1 < \beta < 0$. To overcome this problem, we first integrate the original time series prior to applying the standard FDM. In this way, for processes which can be described as the first derivative of fluctuations with spectral exponent within the interval $1 < \beta < 3$, the relationship between β and D changes to $\beta = 3 - 2D$. A process with positive long-range correlations leads to $D < 1.5$, whereas for anticorrelated processes, $D > 1.5$. The irregular fluctuations with no memory are represented by $D = 1.5$.

(ii) Detrended Fluctuation Analysis. This method was introduced to quantify long-range correlations in the heartbeat interval time series and DNA sequences [43,48,52]. The DFA method has been used to explore the presence of correlations in different areas of science, from heartbeat [52] and stock markets to earthquakes [53]. Briefly, the DFA is described as follows: First, we integrate the original time series to obtain $y(k) = \sum_{i=1}^{k}[x(i) - x_{ave}]$. The resulting series is divided into boxes of size n. For each box, a straight line is fitted to the points, $y_n(k)$. Next, the line points are subtracted from the integrated series, $y(k)$, in each box. The root mean square fluctuation of the integrated and detrended series is calculated by means of

$$F(n) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}[y(k) - y_n(k)]^2}. \tag{3}$$

This process is taken over several scales (box sizes) to obtain a power law behavior $F(n) \sim n^{\alpha}$, with α being an exponent which reflects self-similar and correlation properties of the signal. An uncorrelated signal leads to $\alpha = 0.5$, $\alpha = 1$ represents a long-range correlated process ($1/f$ noise), and $\alpha = 1.5$ corresponds to a Brownian motion. For $0.5 < \alpha < 1$, the scaling exponent α and the Hurst exponent provide the same information, and therefore both can be used interchangeably [54]. The exponent α is also related to the spectral exponent β (within the interval $0 < \beta < 1$) through relation $\beta = 2\alpha - 1$. Thus, fractal dimension D is related to α by means of $\alpha = 2 - D$, as expected for auto-affine signals [47,55].

5. Results and Discussion

5.1. Correlations Analysis

Figure 3 shows the log–log plots for the HFD and DFA analyses performed on the data series of compound sentiment scores for all the books in our sample. As can be seen, these data series exhibit fractal behavior with different scaling exponents for both analyses. In order to further study these correlations, we shuffle the sentences in each book prior to estimating the VADER sentiment and the resulting data series are analyzed with the HFD and DFA methods.

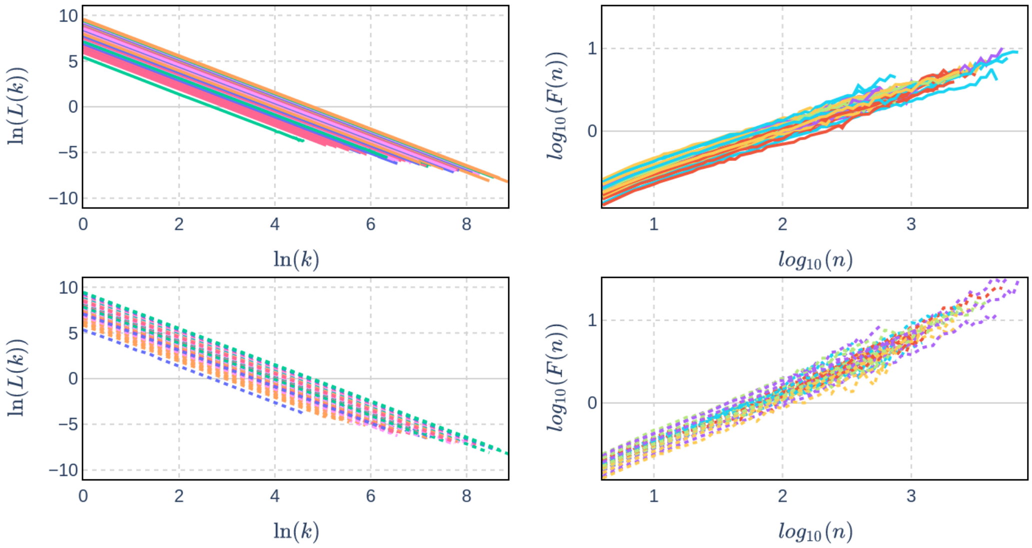

Figure 3. Top row: Log–log plots of $\langle L(k) \rangle$ vs. k for the HFD (**left**) and $F(n)$ vs n for the DFA (**right**). Each trace with different color on both plots corresponds to one book. As can be seen, the composite sentiment data series are monofractal and can be characterized by a single scaling exponent, either D or α. **Buttom row**: Similar log–log plots as in the top row but for the series resulting after shuffling the sentences in the books.

Figure 4 shows a scatter plot of α vs. D for the original texts and the shuffled ones. The mean values for the scaling exponents for the original texts are $D = 1.38 \pm 0.04$ and $\alpha = 0.62 \pm 0.04$. Even though there is scattering in the values for the scaling exponents, we can see from the values of α that the data series of sentiment are slightly long-term correlated. On the other hand, the mean values of the scaling exponents for the shuffled texts are $D = 1.49 \pm 0.03$ and $\alpha = 0.50 \pm 0.02$. We perform a two-sample t-test for each scaling exponent between the original sequences and the shuffled ones, with the null hypothesis that the two populations have no statistically significant difference. In other words, we take the population of one scaling exponent computed for the original sequence (for instance, α) and perform the t-test to assess statistical difference to the population of the values of α obtained for the shuffled sequences, and similar approach is followed for the fractal dimension D. The p-values obtained for the scaling exponents are: $p_D = 7.4 \times 10^{-29}$ and $p_\alpha = 4.5 \times 10^{-22}$. They allow us to reject the null hypothesis, i.e., we find statistically significant differences in α and D between the original sequences and the randomized ones.

We also observe that there is a good concordance between the median values obtained from both methods. By using the relationship $\alpha = 2 - D$, it is easy to show that $\bar{D} = 1.38$ yields to $\bar{\alpha} = 0.62$, which matches the value obtained by means of the DFA. This also applies for the case of the shuffled sequences.

We compute the fraction of the number of sentences on each sentiment state to the total number of sentences for each book. Figure 5 shows box plots for each fraction for all the books considered in our study, with the median values of each population shown within the box. As can be seen, the majority of sentences are valued as neutral, followed by the positive and then the negative ones. The observation that there are more positive than negative sentences in the books under study is in agreement with recent studies reporting a positive bias in human expression [30,31], although in those works, the positivity in the language is based on a word-by-word analysis, and in our case, the analysis is made on a sentence-by-sentence basis, which supplements and extends those previously reported results.

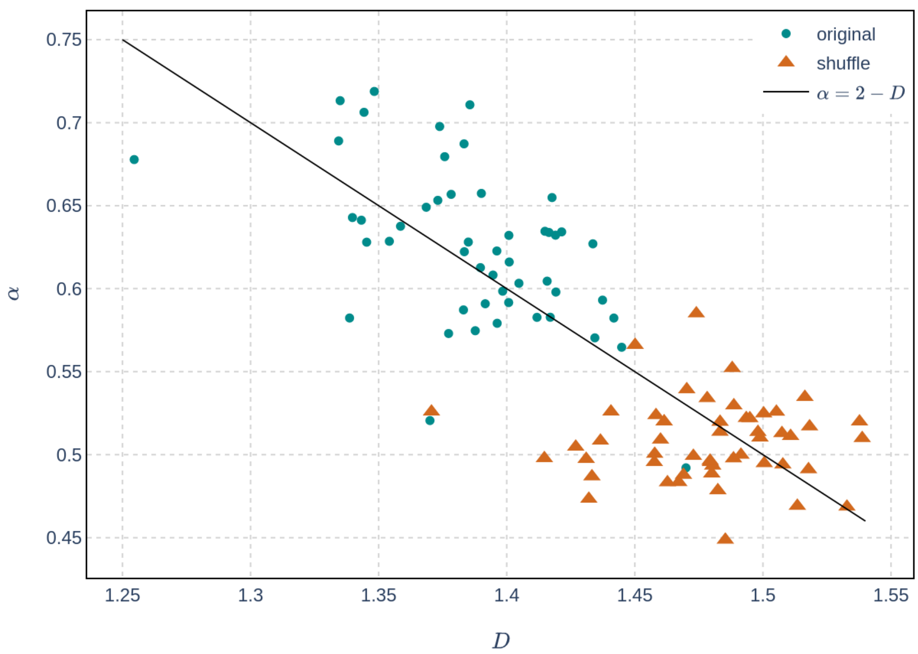

Figure 4. Scaling exponents for the HFD and DFA for the original texts and the shuffled ones. The straight line corresponds to the relationship $\alpha = 2 - D$.

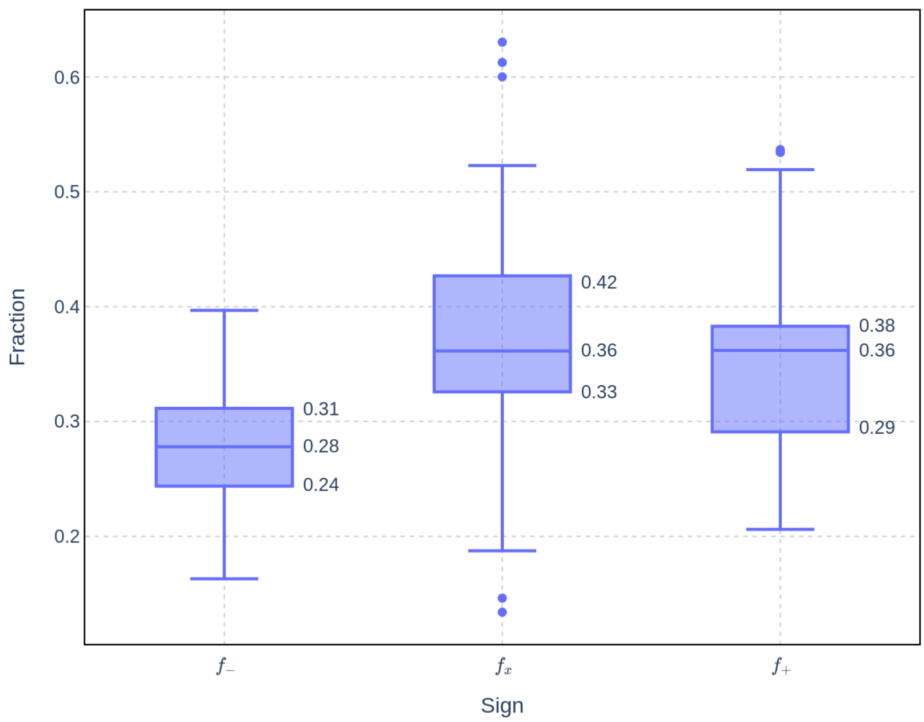

Figure 5. Box plots for the fraction of sentences on each sentiment state (f_+ for positive, $f_\times$ for neutral and f_- for negative sentiments) with respect to the total number of sentences in the book. The values shown are the median (inside the box) and the lower and upper quartile. The whiskers extend from the box to show the range of the data, while the flier points are those past the end of the whiskers and are considered as outliers.

In order to investigate further the bias towards positive sentiment in language, for each book, we obtain the normalized histograms for the compound sentiment series. Figure 6 shows the overlay of the histograms. As can be seen, there is a peak for the sentiments around the neutral values of sentiment. Moreover, we compute the area under the positive (denoted by A_+) and the negative (A_-) sides, which offer the probability of finding positive or negative sentences in the text, respectively. The scatter plot of A_+ vs A_- in the inset of Figure 6 shows that for most of the cases, the probability of finding positive sentences is larger than the one for the negative ones. Moreover, we compute the empirical cumulative distribution function for the series of compound series and compare the positive and negative tails (see Figure 7). As can be seen, the probability of finding sentences with a compound value lower than a certain value c_0 decreases faster for the negative tail.

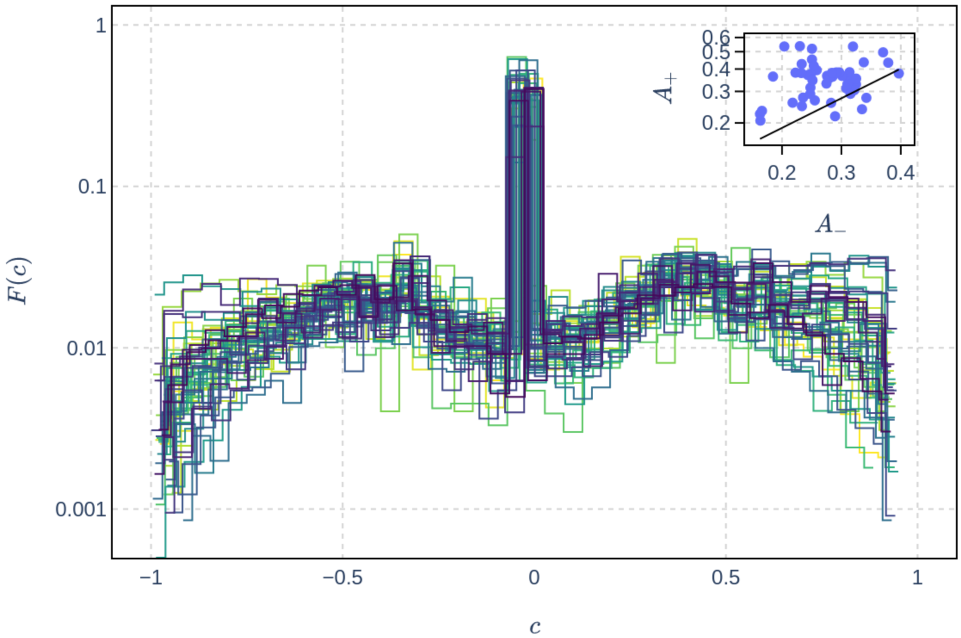

Figure 6. Normalized histogram ($F(c)$) for the compound sentiment c, where each trace with different color corresponds to a book. The inset is a scatter plot of the areas under the histograms for the positive and negative sides, which shows that the probability of finding sentences with positive sentiment is higher than with negative sentiment (the identity line is provided as a visual guide).

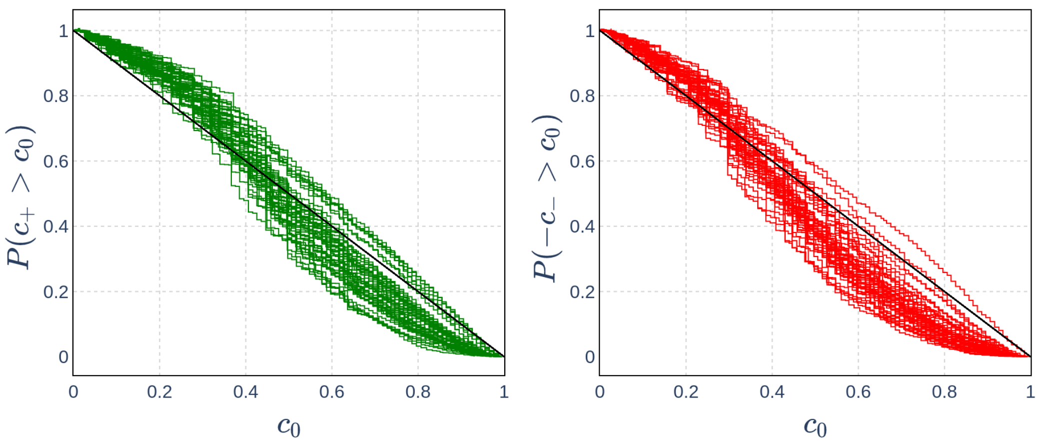

Figure 7. Cumulative distribution function for the compound values for the positive sentiment (**left**) and the negative one (**right**). Each trace corresponds to a book, green color for positive values and red for negative ones. The straight line is provided as a visual guide.

5.2. Observations from the Transition Coefficients

As mentioned in the Methods section, we discretize the compound sentiment series by applying a classification approach. After counting the number of sentiment states transition, we come up with a state transition matrix for each book. Figure 8 shows box plots for the coefficients of each possible transition, for both populations: original text and shuffled text. As can be seen, the most likely transition is from a neutral sentence to another neutral one, followed by the transitions from a positive (negative) sentences to a positive (negative) one.

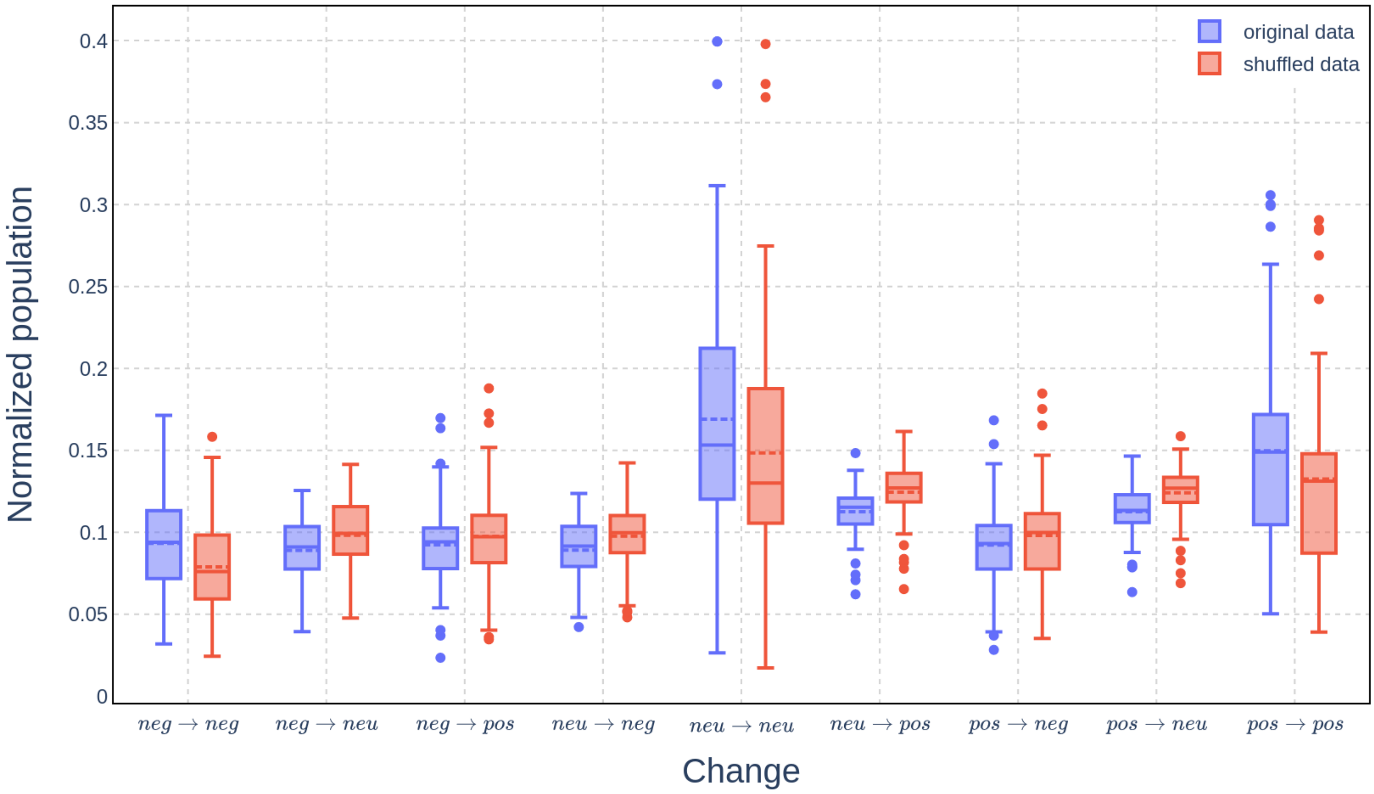

Figure 8. Box plots for the normalized populations of each state transition coefficient. The values for the original text and its shuffled version are given side by side (box plots in red color represent the shuffled text). The median values are marked within each box plot with a dotted line, while the mean values are marked with a solid line. The whiskers extend from the box to show the range of the data, while the flier points are those past the end of the whiskers and are considered as outliers.

We see that for the three states, the transition to a neutral sentence is the one with higher probability, which suggests that the neutral sentence is an inflection point for the sentiment when changing emotion (positive to negative or vice versa) or that these sentences are providing general and contextual information in the narrative. Starting from a positive sentence, it is more likely to move to a positive sentence, and from a negative sentence to a negative one, except that the latter transition is less differentiated with respect to the transitions to neutral or positive sentences, while moving from a neutral to a neutral sentence is markedly more likely compared to the other two transitions. Moreover, the medians of the transition coefficients show that the transition matrices are not symmetric, since there is a certain bias for transitions to positive and to neutral sentences. In summary, Table 1 shows the list of state transitions in descending order on the probabilities of transitioning to a certain state given a start on a particular one.

Table 1. List of state transitions in descending order of transition probabilities.

Initial State	First Most Likely Transition	Second	Third
positive	positive	neutral	negative
neutral	neutral	positive	negative
negative	negative	positive	neutral

On the other hand, Figure 8 also shows the state transition coefficients for the shuffled sentences. As can be seen, the transition matrix between states of the same type (*pos*–*pos*, *neg*–*neg* and *neu*–*neu*) exhibit mean values that are smaller than the values corresponding to the original data, while for the other transitions, the differences between the mean values is less noticeable. We see that for the shuffled sentences, the transition coefficient to a certain state is the same independently of on which state we started with, while there is a bias towards neutral sentences followed by positive ones. Again, the transition to negative sentences has the lower probability. Therefore, even in this case of shuffled sentences, there is a bias towards positive sentiment in human language, as previously reported in other studies [56]. One interesting thing to note here is that the median values of the transition coefficients for the shuffled texts follow the same proportion that the fraction of text that it is in a certain sentiment state (see Figure 5).

The long-term correlations in the sentiment data series that are revealed by the DFA analysis are somehow expected since, by construction, the writing process has memory given that the narrative is about concatenating sentences that are related. Our results on the presence of long-term correlations with values in fractal exponents, which reveal slight persistence, are in agreement with previous results reported for sentiment value sequences and those based on word-length [35,37,38], sentence-length [57], and sentiment arcs [58]. In addition, this approach could eventually be used to evaluate certain individual characteristics of the texts, which would provide an alternative way of identifying them through their organization across different scales. On the other hand, by introducing a Markov chain approach, we find that the sentiment state of a sentence depends on the immediately preceding one. This approach allows us to reveal the biases on the transitions to neutral and positive sentences in the evolution of the sentiment along the text. The narrative has sentiment scores that are correlated, but the story being told can vary in emotion and intensity from one sentence to the next; therefore, the sentiment evolution along the text could be mapped to a Markov chain. In other words, the long-term correlation approach allows to make a global assessment of the book sentiment, while the state transition matrix approach provides local information on the evolution of sentiment along the text. These local and global properties have been recognized as features present in systems that exhibit, at the same time, the presence of event "clustering" and long-term correlations [59–61]. The main limitation of this study lies in the sentiment values returned by VADER, where the compound sentiment construction may exhibit biases when analyzing content with sarcasm or other representations where the tool is not robust. Although these evaluations can be made more robust by considering other sentiment analysis tools, it remains a complex task. Further studies are needed to explore the feasibility of obtaining specific features of sentiment evolution in literary texts, and whether the levels of correlations are representative of the type of text being analyzed to determine properties such as literary genres, authors, etc.

6. Conclusions

We perform a sentence-level sentiment analysis of several literary works in English. This study has two approaches: (i) a global one through the assessment of the correlations in the sentiment along the text and (ii) a local one through the analysis of the transitions of sentiment from a sentence to the next. The global assessment, performed by applying the Higuchi Fractal Dimension and the Detrended Fluctuation Analysis methods, shows that the data series of sentiment scores is monofractal with long-term correlations, which can be explained by the fact that the writing process has memory by construction with a sentiment evolution that is self-similar. The local assessment, performed by counting the sentiment state transitions, shows that there are biases that make more likely the transitions to neutral and positive sentences, implying a bias towards positive sentiment in the texts. Previous studies based on word-level sentiment analysis have found a similar bias to positive sentiment [56]. These two approaches supplement each other because the long-term correlation approach allows a global assessment of the sentiment of the book, while the state transition matrix approach provides local information about the sentiment evolution along the text.

Author Contributions: Conceptualization, R.H.-P. and L.G.-V.; methodology, R.H.-P., P.L.-M., B.O.-Q. and L.S.L.; software, R.H.-P. and P.L.-M.; validation, R.H.-P., P.L.-M., B.O.-Q. and L.S.L.; formal analysis, R.H.-P., B.O.-Q., L.S.L. and L.G.-V.; investigation, R.H.-P., P.L.-M., B.O.-Q., L.S.L. and L.G.-V.; resources, B.O.-Q. and L.G.-V.; data curation, R.H.-P. and P.L.-M.; writing—original draft preparation, R.H.-P.; writing—review and editing, R.H.-P., P.L.-M., B.O.-Q., L.S.L. and L.G.-V.; visualization, R.H.-P. and P.L.-M.; supervision, R.H.-P. and L.G.-V.; project administration, R.H.-P. and L.G.-V.; funding acquisition, B.O.-Q. and L.G.-V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All corpora used in this study are available from the repository at https://figshare.com/articles/dataset/A_sentence-level_sentiment_analysis_of_some_literary_texts_inEnglish_language/27092008, last accessed day: 30 September 2024.

Acknowledgments: The plots in this paper were produced with `matplotlib (3.9.2)` [62]. This work was partially supported by COFAA-IPN, EDI-IPN, UNAM-PAPIIT (IN111822) and CONAHCYT (México).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Dodds, P.S.; Danforth, C.M. Measuring the Happiness of Large-Scale Written Expression: Songs, Blogs, and Presidents. *J. Happiness Stud.* **2010**, *11*, 441–456. [CrossRef]
2. Dodds, P.S.; Harris, K.D.; Kloumann, I.M.; Bliss, C.A.; Danforth, C.M. Temporal Patterns of Happiness and Information in a Global Social Network: Hedonometrics and Twitter. *PLoS ONE* **2011**, *6*, e26752. [CrossRef] [PubMed]
3. Chmiel, A.; Sobkowicz, P.; Sienkiewicz, J.; Paltoglou, G.; Buckley, K.; Thelwall, M.; Hołyst, J.A. Negative emotions boost user activity at BBC forum. *Physica A Stat. Mech. Its Appl.* **2011**, *390*, 2936–2944. [CrossRef]
4. Chmiel, A.; Sienkiewicz, J.; Thelwall, M.; Paltoglou, G.; Buckley, K.; Kappas, A.; Hołyst, J.A. Collective Emotions Online and Their Influence on Community Life. *PLoS ONE* **2011**, *6*, e22207. [CrossRef] [PubMed]
5. Wankhade, M.; Rao, A.C.S.; Kulkarni, C. A survey on sentiment analysis methods, applications, and challenges. *Artif. Intell. Rev.* **2022**, *55*, 5731–5780. [CrossRef]
6. Calvo, R.A.; Mac Kim, S. Emotions in text: Dimensional and categorical models. *Comput. Intell.* **2013**, *29*, 527–543. [CrossRef]
7. Jain, M.; Jindal, R.; Jain, A. Building Domain-Specific Sentiment Lexicon Using Random Walk-Based Model on Common-Sense Semantic Network. In *International Conference on Innovative Computing and Communication*; Springer: Singapore, 2023; pp. 193–204.
8. Cambria, E.; Poria, S.; Hazarika, D.; Kwok, K. SenticNet 5: Discovering conceptual primitives for sentiment analysis by means of context embeddings. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018; Volume 32.
9. Akhtar, M.S.; Ekbal, A.; Cambria, E. How intense are you? Predicting intensities of emotions and sentiments using stacked ensemble [application notes]. *IEEE Comput. Intell. Mag.* **2020**, *15*, 64–75. [CrossRef]
10. Nandwani, P.; Verma, R. A review on sentiment analysis and emotion detection from text. *Soc. Netw. Anal. Min.* **2021**, *11*, 81. [CrossRef]
11. Xie, H.; Lin, W.; Lin, S.; Wang, J.; Yu, L.C. A multi-dimensional relation model for dimensional sentiment analysis. *Inf. Sci.* **2021**, *579*, 832–844. [CrossRef]
12. Roy, B.; Das, S. Perceptible sentiment analysis of students' WhatsApp group chats in valence, arousal, and dominance space. *Soc. Netw. Anal. Min.* **2022**, *13*, 9. [CrossRef]
13. Rizhinashvili, D.; Sham, A.H.; Anbarjafari, G. Enhanced speech emotion recognition using averaged valence arousal dominance mapping and deep neural networks. *Signal Image Video Process.* **2024**, *18*, 7445–7454. [CrossRef]

14. Sharma, S.S.; Dutta, G. SentiDraw: Using star ratings of reviews to develop domain specific sentiment lexicon for polarity determination. *Inf. Process. Manag.* **2021**, *58*, 102412. [CrossRef]
15. Baldwin, J.; Brunsdon, T.; Gaudoin, J.; Hirsch, L. Comparative Analysis of Lexicon-Based Sentiment Analysis Methods. 2023. Available online: https://ssrn.com/abstract=4531226 (accessed on 27 May 2024).
16. Oh, N. sentibank: A Unified Resource of Sentiment Lexicons and Dictionaries. *Proc. Int. AAAI Conf. Web Soc. Media* **2024**, *18*, 2003–2013. [CrossRef]
17. Ophir, Y.; Walter, D. Computational Sentiment Analysis. In *Emotions in the Digital World: Exploring Affective Experience and Expression in Online Interactions*; Oxford University Press: Oxford, UK, 2023; p. 114.
18. Taboada, M.; Brooke, J.; Tofiloski, M.; Voll, K.; Stede, M. Lexicon-based methods for sentiment analysis. *Comput. Linguist.* **2011**, *37*, 267–307. [CrossRef]
19. Deng, S.; Sinha, A.P.; Zhao, H. Adapting sentiment lexicons to domain-specific social media texts. *Decis. Support Syst.* **2017**, *94*, 65–76. [CrossRef]
20. Itani, M.; Roast, C.; Al-Khayatt, S. Developing resources for sentiment analysis of informal Arabic text in social media. *Procedia Comput. Sci.* **2017**, *117*, 129–136. [CrossRef]
21. Thelwall, M. The Heart and Soul of the Web? Sentiment Strength Detection in the Social Web with SentiStrength. In *Cyberemotions. Understanding Complex Systems*; Springer: Cham, Switzerland, 2017. [CrossRef]
22. Sebastiani, F.; Esuli, A. Sentiwordnet: A publicly available lexical resource for opinion mining. In Proceedings of the 5th International Conference on Language Resources and Evaluation, Genoa, Italy, 22–28 May 2006; European Language Resources Association (ELRA): Genoa, Italy, 2006; pp. 417–422.
23. Khan, F.H.; Qamar, U.; Bashir, S. SentiMI: Introducing point-wise mutual information with SentiWordNet to improve sentiment polarity detection. *Appl. Soft Comput.* **2016**, *39*, 140–153. [CrossRef]
24. Warriner, A.B.; Kuperman, V.; Brysbaert, M. Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behav. Res. Methods* **2013**, *45*, 1191–1207. [CrossRef]
25. Lee, L.H.; Li, J.H.; Yu, L.C. Chinese EmoBank: Building Valence-Arousal Resources for Dimensional Sentiment Analysis. *ACM Trans. Asian Low-Resour. Lang. Inf. Process.* **2022**, *21*, 65. [CrossRef]
26. Bradley, M.; Lang, P.J. *Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings*; Technical Report C-1; The Center for Research in Psychophysiology, University of Florida: Gainesville, FL, USA, 1999. Available online: https://pdodds.w3.uvm.edu/teaching/courses/2009-08UVM-300/docs/others/everything/bradley1999a.pdf (accessed on 21 October 2024).
27. Nielsen, F.Å. A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. *arXiv* **2011**, arXiv:1103.2903.
28. Ribeiro, F.N.; Araújo, M.; Gonçalves, P.; Gonçalves, M.A.; Benevenuto, F. SentiBench—A benchmark comparison of state-of-the-practice sentiment analysis methods. *EPJ Data Sci.* **2016**, *5*, 23. [CrossRef]
29. Hutto, C.J.; Gilbert, E.E. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proc. Int. AAAI Conf. Web Soc. Media* **2014**, *8*, 216–225. [CrossRef]
30. Garcia, D.; Garas, A.; Schweitzer, F. Positive words carry less information than negative words. *EPJ Data Sci.* **2012**, *1*, 3. [CrossRef]
31. Kloumann, I.M.; Danforth, C.M.; Harris, K.D.; Bliss, C.A.; Dodds, P.S. Positivity of the English Language. *PLoS ONE* **2012**, *7*, e29484. [CrossRef]
32. Yuri, B.; Pascale, F. Sentiment Analysis for Literary Texts: Hemingway as a Case-study. *J. Data Min. Digit. Humanit.* **2024**, *NLP4DH*. [CrossRef]
33. Yeruva, V.K.; Chandrashekar, M.; Lee, Y.; Rydberg-Cox, J.; Blanton, V.; Oyler, N.A. Interpretation of sentiment analysis in aeschylus's Greek tragedy. In Proceedings of the 4th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature, Barcelona, Spain, 12 December 2020; pp. 138–146.
34. Vinodini, S. Analyzing Sentiments in Paulo Coelho's Literary Works Using VADER Sentiment Analysis. In Proceedings of the 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), Chennai, India, 14–15 December 2023; pp. 1–6.
35. Ausloos, M. Generalized Hurst exponent and multifractal function of original and translated texts mapped into frequency and length time series. *Phys. Rev. E* **2012**, *86*, 031108. [CrossRef]
36. Rodriguez, E.; Aguilar-Cornejo, M.; Femat, R.; Alvarez-Ramirez, J. Scale and time dependence of serial correlations in word-length time series of written texts. *Physica A Stat. Mech. Its Appl.* **2014**, *414*, 378–386. [CrossRef]
37. Guzmán-Vargas, L.; Obregón-Quintana, B.; Aguilar-Velázquez, D.; Hernández-Pérez, R.; Liebovitch, L.S. Word-Length Correlations and Memory in Large Texts: A Visibility Network Analysis. *Entropy* **2015**, *17*, 7798–7810. [CrossRef]
38. Drożdż, S.; Oświęcimka, P.; Kulig, A.; Kwapień, J.; Bazarnik, K.; Grabska-Gradzińska, I.; Rybicki, J.; Stanuszek, M. Quantifying origin and character of long-range correlations in narrative texts. *Inf. Sci.* **2016**, *331*, 32–44. [CrossRef]
39. Gromov, V.A.; Dang, Q.N. Semantic and sentiment trajectories of literary masterpieces. *Chaos Solitons Fractals* **2023**, *175*, 113934. [CrossRef]
40. Ausloos, M. Measuring complexity with multifractals in texts. Translation effects. *Chaos Solitons Fractals* **2012**, *45*, 1349–1357. [CrossRef]
41. Carrizales-Velazquez, C.; Donner, R.V.; Guzmán-Vargas, L. Generalization of Higuchi's fractal dimension for multifractal analysis of time series with limited length. *Nonlinear Dyn.* **2022**, *108*, 417–431. [CrossRef]

42. Higuchi, T. Approach to an irregular time series on the basis of the fractal theory. *Physica D Nonlinear Phenom.* **1988**, *31*, 277–283. [CrossRef]
43. Peng, C.K.; Buldyrev, S.V.; Havlin, S.; Simons, M.; Stanley, H.E.; Goldberger, A.L. Mosaic organization of DNA nucleotides. *Phys. Rev. E* **1994**, *49*, 1685–1689. [CrossRef]
44. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 623–656. [CrossRef]
45. Griffths, T.L. Rethinking language: How probabilities shape the words we use. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 3825–3826. [CrossRef]
46. Markov, A.A. An Example of Statistical Investigation of the Text Eugene Onegin Concerning the Connection of Samples in Chains. *Sci. Context* **2006**, *19*, 591–600. [CrossRef]
47. Malamud, B.D.; Turcotte, D.L. Self-affine time series: Measures of weak and strong persistence. *J. Stat. Plan. Inference* **1999**, *80*, 173–196. [CrossRef]
48. Peng, C.K.; Havlin, S.; Stanley, H.E.; Goldberger, A.L. Long-range anti-correlations and non- Gaussian Behavior of the heartbeat. *Phys. Rev. Lett.* **1995**, *70*, 1343–1346. [CrossRef] [PubMed]
49. Rangarajan, G.; Ding, M. Integrated approach to the assessment of long range correlation in time series data. *Phys. Rev. E* **2000**, *61*, 4991–5001. [CrossRef] [PubMed]
50. Higuchi, T. Relationship between the fractal dimension and the power law index for a time series: A numerical investigation. *Physica D Nonlinear Phenom.* **1990**, *46*, 254–264. [CrossRef]
51. Guzmán-Vargas, L.; Angulo-Brown, F. Simple model of the aging effect in heart interbeat time series. *Phys. Rev. E* **2003**, *67*, 052901. [CrossRef] [PubMed]
52. Peng, C.K.; Mietus, J.; Hausdorff, J.; Havlin, S.; Stanley, H.E.; Goldberger, A.L. Quantification of Scaling Exponents and Crossover Phenomena in Nonstationary Heartbeat Time Series. *Chaos* **1995**, *5*, 82–87. [CrossRef]
53. Telesca, L.; Lapenna, V.; Macchiato, M. Multifractal fluctuations in seismic interspike series. *Physica A Stat. Mech. Its Appl.* **2005**, *354*, 629–640. [CrossRef]
54. Bashan, A.; Bartsch, R.; Kantelhardt, J.W.; Havlin, S. Comparison of detrending methods for fluctuation analysis. *Physica A Stat. Mech. Its Appl.* **2008**, *387*, 5080–5090. [CrossRef]
55. Feder, J. *Fractals*; Springer: New York, USA, 1988.
56. Dodds, P.S.; Clark, E.M.; Desu, S.; Frank, M.R.; Reagan, A.J.; Williams, J.R.; Mitchell, L.; Harris, K.D.; Kloumann, I.M.; Bagrow, J.P.; et al. Human language reveals a universal positivity bias. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 2389–2394. [CrossRef]
57. Yang, T.; Gu, C.; Yang, H. Long-range correlations in sentence series from a story of the stone. *PLoS ONE* **2016**, *11*, e0162423. [CrossRef]
58. Bizzoni, Y.; Moreira, P.; Thomsen, M.R.; Nielbo, K.L. The fractality of sentiment arcs for literary quality assessment: The case of Nobel laureates. In Proceedings of the 2nd International Workshop on Natural Language Processing for Digital Humanities, Taipei, Taiwan, 20 November 2022; pp. 31–41.
59. Yamasaki, K.; Muchnik, L.; Havlin, S.; Bunde, A.; Stanley, H.E. Scaling and memory in volatility return intervals in financial markets. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 9424–9428. [CrossRef]
60. Hernández-Pérez, R.; Guzmán-Vargas, L.; Reyes-Ramirez, I.; Angulo-Brown, F. Evolution in time and scales of the stability of heart interbeat rate. *Europhys. Lett.* **2011**, *92*, 68006. [CrossRef]
61. Reyes-Ramírez, I.; Guzmán-Vargas, L. Scaling properties of excursions in heartbeat dynamics. *Europhys. Lett.* **2010**, *89*, 38008. [CrossRef]
62. Hunter, J.D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [CrossRef]