Global Coherence, Local Uncertainty – Towards a Theoretical Framework for Assessing Literary Quality

Yuri Bizzoni[1,†], Pascale Feldkamp[1,†] and Kristoffer Nielbo[1,†]

[1]Center for Humanities Computing Aarhus, Jens Chr. Skous Vej 4, Building 1483, DK-8000 Aarhus C, Denmark

Abstract

A theoretical framework for evaluating literary quality through analyzing narrative structures using simplified narrative representations in the form of story arcs is presented. This framework proposes two complementary models: the first employs Approximate Entropy to measure local unpredictability, while the second utilizes fractal analysis to assess global coherence. When applied to a substantial corpus of 9,089 novels, the findings indicate that narratives characterized by high literary quality, as indicated by reader ratings, exhibit a balance of local unpredictability and global coherence. This dual approach provides a formal and empirical basis for assessing literary quality and emphasizes the importance of considering intrinsic properties and reader perception in literary studies.

Keywords

literature, information theory, fractal theory, aesthetic theory

1. Introduction

Quality assessment of literature is a highly contested matter. Positions in the debate range from constructivist context dependency ('the success of a work of literature depends entirely on its context') to work internalism ('success depends on work-internal features') [3, 36]. While context dependency is evidenced by the variety and seemingly chaotic dynamics of, e.g., bestseller lists, the constructivist argument ignores the convergence of an empirical 'canon' for many readers over time and space [34, 59]. Moreover, to attribute the longevity or popularity of certain books to purely contextual factors would seem to be at odds with large-scale consensus among readers, which appear far from volatile [1, 60, 39].

Several shifts have played a role in making terms like "literary quality" or "classics" unpopular in the discipline of literary studies, even said to belong to the "precritical era of criticism itself" [23]: Methodological shifts that move the scholarly focus from evaluation to interpretation [9], an expansion of the conceptual boundaries of "literature" to encompass texts that challenge traditional ideas of beauty or enjoyment [62], but also constructivist and postcolonial shifts that bring attention to the context of literary evaluation and tradition, canon representativity [24] and the inequality of cultural production [55, 53].

CHR 2024: Computational Humanities Research Conference, December 4–6, 2024, Aarhus, Denmark

† These authors contributed equally.

yuri.bizzoni@cc.au.dk (Y. Bizzoni); pascale.moreira@cc.au.dk (P. Feldkamp); kln@cas.au.dk (K. Nielbo)
0000-0002-6981-7903 (Y. Bizzoni); 0000-0002-2434-4268 (P. Feldkamp); 0000-0002-5116-5070 (K. Nielbo)

Context-dependent positions on literary appreciation range from simply prioritizing context over work-internal factors [23], to beliefs on the contextual determinacy of aesthetic evaluation [19, 15]. An extreme perspective holds that critical evaluation reflects culturally dominant voices, arguing that canon judgments "are only the instruments of entrenched interests" [24, p. iv]. In this sense, a disparity appears to have arisen between a scholarly "denial of quality" [62, 46], and the multitude of quality judgements that are practiced within the literary culture (literary awards, classics book series, prescriptive creative writing courses, etc.).

Conversely, *work-internalist positions* are closer, at least in their purest form, to universalist claims, where aesthetic judgments are seen as arising solely from the interaction between the individual reader and writer [10]. On this position, aesthetic pleasure tends to be considered as an a-historical or culturally universal experience and sidetracks as unimportant any "historically contingent" aspect of literature [16].

We can simplify the concept of literary quality by modeling *perceived* literary quality, thereby differentiating between a contested intrinsic value and the perceived value of a work, which anchors the quality assessment in the reader's aesthetic experience (i.e., a situated experience).

Empirical aesthetics offers well-formed theoretical work supported by empirical findings on the nature of aesthetic experience[22] and preferences at the psychological level [47, 11], as well as cultural and contextual influences on the aesthetic experience [12]. Yet, empirical aesthetics has predominantly focused on the visual modality, in particular of pictorial art, attempting to find patterns correlating with the appreciation of paintings [22]. Within the domain of paintings, prevalent research topics include aesthetic appreciation and judgment [37], their perceptual and cognitive underpinnings [38, 61], and, finally, emotional response [54, 58, 47]. The appreciation of literary prose has received significantly less attention than pictorial art [22, 11], and studies tend to focus on general aspects of aesthetic appreciation of language [33] and literary response [41, 42, 18].

Approaches in empirical aesthetics that are based on fractal theory are particularly promising since they cross aesthetic modalities, that is, they seem to be valid for pictorial [56, 51] literary [43, 44] and musical [40, 25] arts. Collectively, these approaches state that the appreciation of art depends on the presence of fractal-like scale-invariant properties. Although the explanation for this can vary, it is typically grounded in neural mechanisms of sensory coding [51]: a prototypical example is Jackson Pollock's abstract expressionism, where his drip painting technique was used to compose self-similar patterns repeated at multiple scales, which enables "Pollock authentication" through fractal analysis [56]. The same properties can be found in literature, where linguistic properties tend to display fractal behavior [43], but where positively evaluated literary narratives tend to be characterized by a higher degree of self-similarity [8, 4, 5]. These findings of multiscale self-similar repetitions in a set of linguistic properties (e.g., sentence structure, literary entities, or lines) are generally translated into *global* coherence or consistency of the aesthetic object [5, 7, 50]. In literary prose, for example, it is argued that a high-quality story is characterized by a narrative coherence and multiscale predictability that distinguish it from bland and unpredictable stories [28]. Similarly, high-quality paintings have visual elements that remain consistent at different magnification levels. A recent addition to this global quality property is a *local* property in literary prose that can identify quality [43, 44]. More specifically, canonical works show higher degrees of sequential unpredictability compared to noncanonical fiction [44]. Combining these two quality indicators, we expect that high-quality literature, that is, literature that is appreciated by a majority, should display a tension between global coherence and local unpredictability and, furthermore, that a model of perceived literary quality should account for both.

This paper will presents theoretical framework on two models of literary quality that utilize simplified narrative representations in the form of story arcs [29, 49, 45]. The first model uses approximate entropy to capture local unpredictability of the story arc [44], and the second model is based on fractal analysis to characterize the global coherence of the story arc [28]. Finally, we present an application of the models to a large collection of literary texts that when applied to the texts' story arcs, covary in non-trivial ways.

2. Proposal and Methodology

In this section, we present the theoretical foundation for models of literary quality, focusing on the integration of global coherence and local uncertainty within narrative structures. By employing metrics from empirical aesthetics, fractal theory and information theory, we propose two complementary approaches to quantify these aspects: approximate entropy and the Hurst exponent. Specifically, we suggest the for literature to be appreciated, it should manifest a tension, or positive correlation, between global coherence and local unpredictability (Fig. 1).

2.1. Approximate entropy - a measure of local predictability

Approximate entropy ($ApEn$) is a statistical metric that quantifies the complexity and regularity inherent in the time series data. Within the framework of narrative analysis, $ApEn$ provides a means to assess the predictability and structural characteristics of story arcs [44]. A lower $ApEn$ value signifies a more predictable and regular story arc, indicative of a repetitive and 'boring' narrative structure. In contrast, a higher $ApEn$ value denotes a more complex and less predictable arc, suggesting a novel and intricate storyline, 1. This metric facilitates the quantitative analysis of local narrative predictability, enabling a formal evaluation of literary quality through empirical methodologies. For the specific question of evaluating literary quality, we expect that higher local $ApEn$ will be associated with higher literary appreciation, c.f., [44].

In order to estimate $ApEn$ for story arc $X = x(1), \ldots, x(n)$, subsequences of length m, $y_i^m = [x(i), \ldots, x(i+(m-1))]$, and tolerance r, we start by computing the Chebyshev distance between each sub-sequence y_i^m and y_j^m

$$d_{i,j}^m = \max_k \mid y_i^m - y_j^m(k) \mid$$

for each sub-sequence y_i^m to compute the count C

$$C_i^m(r) = \frac{1}{n-m+1} \sum_{j=1}^{n-m+1} H(r - d_{i,j}^m)$$

where $H(.)$ is the Heaviside function

$$H(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

then define

$$\phi^m(r) = \frac{1}{n-m+1} \sum_{i=1}^{n-m+1} \log(C_i^m(r))$$

where log(.) is the natural logarithm. Repeat the above for all sub-sequences of length $m + 1$ to compute $\phi^(m + 1)(r)$, then $ApEn$ is

$$ApEn(m, r) = \phi^m(r) - \phi^{m+1}(r)$$

A note on parameter selections, the choice of m and r can influence the estimation of $ApEn$. Typically, m is set to 2 or 3 (in this study $m = 2$), while r is chosen as a small percentage of the standard deviation of the story arc (in this study $0.2 * SD$). It is important to note that the values of the optimal parameters may vary depending on the specific application and the characteristics of the data.

2.2. The Hurst exponent - a measure of global coherence

Story arcs can be modeled as fractal processes to understand their time-dependent self-similarity. By representing narratives as one-dimensional time series, we can apply fractal analysis to model the underlying pattern of coherence [28, 8, 6]. This approach allows us to quantify the degree of self-similarity in the narrative flow. Using the Hurst exponent, H, which quantifies the persistence of a story arc, we can differentiate between persistent, anti-persistent, and short-memory processes, thus characterizing the temporal dynamics in three qualitative ranges [28], 1. We expect that literary appreciation will be associated with coherence, which translates to a persistent story arc such that narrative fluctuations at shorter time scales are approximate copies of narrative fluctuations at longer time scales.

The Hurst H exponent quantifies persistence or memory in time series, where $0 < H < 0.5$ is an anti-persistent process, $H = 0.5$ is a short-memory process, and $0.5 < H < 1$ is a persistent process [20]. For the present proposal, a persistent process indicates continuity of the story arc (i.e., sentiment states will last for a long time). An anti-persistent indicates rigidity (i.e., sentiment states will rapidly decay to a mean state). Finally, short memory indicates a lack of continuity (sentiment states will only be correlated at short time scales).

Detrended fluctuation analysis (DFA) [48] is a widely used method for estimating the Hurst exponent for a time series. DFA consists of five steps, 1) initially a random walk process is constructed from the time series:

$$u(n) = \sum_{k=1}^{n} (x_k - \overline{x}), \quad n = 1, 2, \cdots, N,$$

where $\overline{x}$ is the mean of the series $x(k)$, $k = 1, 2, \cdots, N$; 2) dividing the constructed random walk process into non-overlapping segments; 3) determining the local trends of each segment as the best polynomial fit; 4) getting the variance of the differences between the random walk process and the local trends; and 5) determining the average variance over all the segments. DFA may involve discontinuities at the boundaries of adjacent segments. Such discontinuities can be detrimental when the data contain trends [27], non-stationarity [32], or nonlinear oscillatory components [13, 26]. Adaptive fractal analysis (AFA) provides a robust alternative to DFA that solves these issues [20]. The main advantage of AFA over DFA is that it identifies a global smooth trend, which is obtained by optimally combining local linear or polynomial fitting, and thus no longer suffers from DFA's problem of discontinuities of adjacent segments. As a result, AFA can automatically deal with arbitrary, strong nonlinear trends that are not unusual to encounter in story arcs[20, 26, 28].

AFA is based on a multiscale non-linear adaptive decomposition algorithm [20]. The first step involves partitioning the time series under study into overlapping segments of length $w = 2n + 1$, where neighboring segments overlap by $n + 1$ points. In each segment, the time series is fitted with the best polynomial of order M, obtained using the standard least squares regression; the fitted polynomials in overlapped regions are combined to yield a single global smooth trend. Denoting the fitted polynomials for the $i - th$ and $(i + 1) - th$ segments by $y^i(l_1)$ and $y^{(i+1)}(l_2)$, respectively, where $l_1, l_2 = 1, \cdots, 2n + 1$, we define the fitting for the overlapped region as

$$y^{(c)}(l) = w_1 y^{(i)}(l + n) + w_2 y^{(i+1)}(l), \quad l = 1, 2, \cdots, n + 1,$$

where $w_1 = \left(1 - \frac{l-1}{n}\right)$ and $w_2 = \frac{l-1}{n}$ can be written as $(1 - d_j/n)$ for $j = 1, 2$, and where d_j denotes the distances between the point and the centers of $y^{(i)}$ and $y^{(i+1)}$, respectively. Note that the weights decrease linearly with the distance between the point and the center of the segment. Such weighting ensures symmetry and effectively eliminates any jumps or discontinuities around the boundaries of neighboring segments. As a result, the global trend is smooth at the non-boundary points and has the right and left derivatives at the boundary [52]. The parameters of each local fit are determined by maximizing the goodness of fit in each segment. The different polynomials in the overlapping part of each segment are combined so that the global fit will be the best (smoothest) fit of the overall time series. Note that, even if $M = 1$ is selected, i.e., the local fits are linear, the global trend signal will still be nonlinear. For an arbitrary window size w, we determine, for the random walk process $u(i)$, a global trend $v(i), i = 1, 2, \cdots, N$, where N is the length of the walk. The residual of the fit, $u(i) - v(i)$, characterizes fluctuations around the global trend, and its variance yields the Hurst parameter H according to the following scaling equation:

$$F(w) = \left[\frac{1}{N} \sum_{i=1}^{N} (u(i) - v(i))^2\right]^{1/2} \sim w^H.$$

By computing the global fits, the residual, and the variance between original random walk process and the fitted trend for each window size w, we can plot $\log_2 F(w)$ as a function of $\log_2 w$. The presence of fractal scaling amounts to a linear relation in the plot, with the slope of the relation providing an estimate of H.

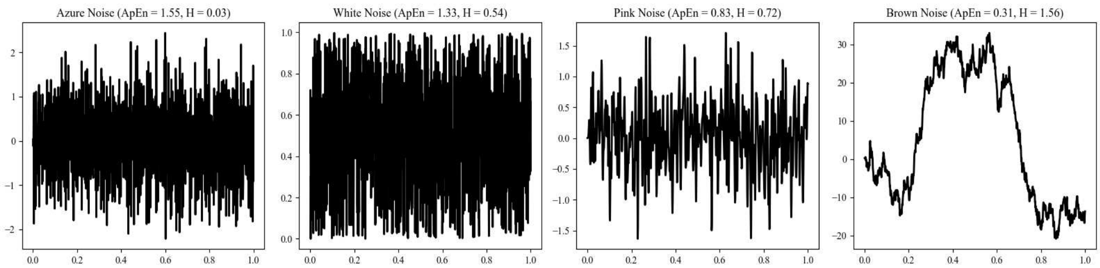

Figure 1: To illustrate the 'boldness' of the proposed tension between global coherence and local unpredictability, we generated four time series characterized by decreasing levels of local unpredictability and increasing levels of global persistence. Starting from left to right, azure noise ($ApEn = 1.55, H = 0.03$), white noise ($ApEn = 1.33, H = 0.54$), 3) pink noise ($ApEn = 0.83, H = 0.72$), 4) brown/black noise ($ApEn = 0.31, H = 1.56$). It should be apparent the both intuitively and theoretically unpredictability and persistence are anti-correlated.

3. Application

To illustrate the practical application of our proposed models, we apply the models to a substantial corpus of literary texts. By analyzing the sentiment arcs of the narratives, we quantify the local unpredictability using $ApEn$ and the global coherence using the Hurst exponent. This application demonstrates how our models align with quality indices, providing empirical support for the notion that a balance of global coherence and local uncertainty enhances perceived literary quality.

3.1. Data

To extract the sentiment arcs, we use a corpus of 9,089 novels published in the US from 1880 to 2000, making it one of the largest corpora of manually cleaned contemporary literature [57, 14, 2]. It was assembled by Hoyt Long and Richard Jean So and was based on the number of libraries worldwide that held a copy of each title, favoring titles with a higher number of holdings. It comprises high-brow fiction (e.g., Joyce Carol Oates, Philip Roth) as well as highly known "genre fiction" (e.g., J.R.R Tolkien, Philip K. Dick), bestsellers (e.g., Agatha Christie, George R.R. Martin) and prestigious experimental literature (e.g., James Joyce). The works span from $341 - 714,444$ words, but more than 97% of the novels are longer than 35,000 words. Overall the corpus totals over one billion words. The corpus has an Anglophone bias: most are US or UK authors writing in English. While this allowed us some form of control on cultural variability within the corpus, it also presents a narrower perspective on literary dynamics.

From each novel, we extracted the average valence sentence by sentence using the *Syuzhet* library [30].[1] *Syuzhet* was chosen based on recent research showing that it performs particularly well on *detrended* narrative arcs, returning values close to human annotations [63].

For this corpus various quality proxies were collected: GoodReads average rating and rating

[1]The custom Syuzhet dictionary is extracted from 165,000 human coded sentences from contemporary literary novels [31].

count,[2] library holding numbers,[3] and translation numbers.[4]

3.2. Findings

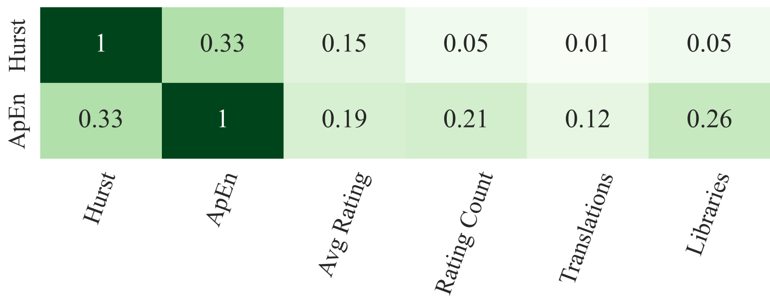

Figure 2: Spearman correlation between the Hurst exponent and approximate entropy (ApEn) of novels in the corpus and four quality proxies: GoodReads average rating and rating count, library holding numbers and translation counts. For all correlations, $p < 0.01$

We estimated $ApEn$ and Hurst exponent for each story arc. We correlated these measures with available metadata quality proxies: GoodReads average rating and rating count, library holdings, and translation counts (see Figure 2). As predicted, we find empirical support that global coherence and local predictability are moderately correlated ($\rho = .33,\ p < .0001$). Furthermore, average ratings are positively associated with both measures ($\rho_{ApEn} = .15$ and $\rho_H = .19,\ p < .0001$). Interestingly, we observe that $ApEn$ is more strongly associated with the additional quality proxies.

One of the most interesting findings is, in our opinion, the positive correlation of Hurst exponent and $ApEn$ in our corpus in general, and with at least one of the proxies of reception. As we show in 1, the Hurst exponent and $ApEn$ tend to be divergent measures, and in fact they have negative correlations for artificial or simpler time series [21]. In other words, a completely random series will have both a higher $ApEn$ and a lower Hurst exponent than a completely trending series. However, Hurst exponent and $ApEn$ do not measure the same thing, and in series with complex behaviors it is not the case that the highest Hurst corresponds to the lowest entropy (see, e.g., Kristoufek and Vosvrda [35]).

Overall, $ApEn$ (as we have computed it) here compares each embedding to all the other embeddings in the series. We can imagine a case in which a time series that has some degree of persistence in its overall trend, but which consists, at the sentence-level, of relatively unpredictable shorter sequences. Such a case would not result in a negative correlation between the measures. If the level of "noise" (unpredictability) is both relatively low and constant throughout a trending series, it is likely to result in a high $ApEn$ and a high Hurst exponent.

This effect may be due to the ability of $ApEn$ to capture shorter-range variations in the sentimental fluctuations of a novel, closer to style than to overall narrative structure. For example,

[2] I.e., the average score assigned to a book by users, and how many GoodReads users assigned a score.
[3] The number of libraries which hold a copy of the book as indexed on WorldCat.
[4] The number of translations of a book in the *Index Translationum* database.

a relatively low $ApEn$ could represent a text where the same small-scale sentiment patterns are repeated in a structured and predictable fashion. Reversely, a high Hurst with high $ApEn$ might represent a text where the overall narrative arc is highly coherent (and predictable), but where the small-scale succession of positive and negative sentences is more chaotic or unpredictable. In this case, readers may find short-term uncertainty but an overall smoother and foreseeable experience: a trend is builds up through a very diverse set of sentimental "zig-zags", rather than through a simple linear succession going from the saddest to the happiest point. Narratives that manage to strike a balance between short-term surprising patterns and long-term coherent structures might be one of the reasons for the observed correlations. Finally, while $ApEn$ slightly correlates with all our metrics of reception, a linear positive relation with the Hurst exponent of sentiment arcs is present only in the case of average GoodReads ratings. Interestingly, this is the only metric that is not related to spread or popularity, while the other three more or less indirectly measure how popular, known or disseminated a text is [17].

4. Concluding Remarks

Our theoretical proposal aims to avoid the contentious issue of universal features by focusing on the reader's perception of literary quality. By redefining literary quality as perceived literary quality, we emphasize the reader's experience and the subjective nature of aesthetic appreciation. Despite this focus on perception, our approach still leans towards intrinsic properties, as our models of approximate entropy and the Hurst exponent are grounded in the structural characteristics of the narratives themselves.

Our findings provide initial empirical support for the hypothesis that tension between global coherence and local unpredictability contributes to perceived literary quality, as reflected in readers' average ratings. However, there is ample room for further research to explain more of the variance in literary quality assessments. Context-sensitive variables such as genre and reader demographics could significantly shape literary appreciation. Less indirect quality proxies with a higher temporal resolution, e.g., reading time, can also provide important correctives to the current proposal.

By combining the strengths of empirical aesthetics with fractal theory and information theory, our proposal offers a robust framework for evaluating literary quality. For future research, our aim is to empirically investigate how the appreciation of literary language is influenced by or drawn to specific forms of temporal organization of language.

Acknowledgments

This research was supported by the Velux Foundation (Grant Name *The Fabula-Net*). Nielbo's work was supported by the Carlsberg Foundation (Grant Number CF23-1583) and Aarhus University Research Foundation (Grant Name *The Golden Imprint*). Bizzoni and Nielbo's work was supported by the Innovation Fund Denmark (Grant Number 4298-00018B).

References

[1] J. Archer and M. L. Jockers. *The Bestseller Code*. London: Penguin books, 2017.

[2] Y. Bizzoni, I. M. S. Lassen, T. Peura, M. R. Thomsen, and K. L. Nielbo. "Predicting Literary Quality - How Perspectivist Should We Be?" In: *1st Workshop on Perspectivist Approaches to Disagreement in NLP, NLPerspectives 2022 as part of Language Resources and Evaluation Conference, LREC 2022 Workshop*. European Language Resources Association, 2022, pp. 20–25.

[3] Y. Bizzoni, P. Moreira, N. Dwenger, I. Lassen, M. Thomsen, and K. Nielbo. "Good Reads and Easy Novels: Readability and Literary Quality in a Corpus of US-published Fiction". In: *Proceedings of the 24th Nordic Conference on Computational Linguistics (NoDaLiDa)*. Tórshavn, Faroe Islands: University of Tartu Library, 2023, pp. 42–51. URL: https://aclanthology.org/2023.nodalida-1.5.

[4] Y. Bizzoni, P. Moreira, M. R. Thomsen, and K. Nielbo. "Sentimental Matters - Predicting Literary Quality by Sentiment Analysis and Stylometric Features". In: *Proceedings of the 13th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis*. Ed. by J. Barnes, O. De Clercq, and R. Klinger. Toronto, Canada: Association for Computational Linguistics, 2023, pp. 11–18. DOI: 10.18653/v1/2023.wassa-1.2. URL: https://aclanthology.org/2023.wassa-1.2.

[5] Y. Bizzoni, P. F. Moreira, M. R. Thomsen, and K. L. Nielbo. "The Fractality of Sentiment Arcs for Literary Quality Assessment: the Case of Nobel Laureates". In: *Journal of Data Mining & Digital Humanities* Nlp4dh (2023), p. 11406. DOI: 10.46298/jdmdh.11406. URL: https://jdmdh.episciences.org/11406.

[6] Y. Bizzoni, P. F. Moreira, M. R. Thomsen, and K. L. Nielbo. "The Fractality of Sentiment Arcs for Literary Quality Assessment: the Case of Nobel Laureates". In: *Journal of Data Mining & Digital Humanities* Nlp4dh (2023), p. 11406. DOI: 10.46298/jdmdh.11406. URL: https://jdmdh.episciences.org/11406.

[7] Y. Bizzoni, T. Peura, M. R. Thomsen, and K. Nielbo. "Sentiment Dynamics of Success: Fractal Scaling of Story Arcs Predicts Reader Preferences". In: *Proceedings of the Workshop on Natural Language Processing for Digital Humanities*. Ed. by M. Hämäläinen, K. Alnajjar, N. Partanen, and J. Rueter. NIT Silchar, India: NLP Association of India (NLPAI), 2021, pp. 1–6. URL: https://aclanthology.org/2021.nlp4dh-1.1.

[8] Y. Bizzoni, T. Peura, M. R. Thomsen, and K. L. Nielbo. "Fractal Sentiments and Fairy Tales: Fractal scaling of narrative arcs as predictor of the perceived quality of Andersen's fairy tales". In: *Journal of Data Mining & Digital Humanities* Nlp4dh (2022). DOI: 10.46298/jdmdh.9154.

[9] E. Bjerck Hagen, C. Hamm, F. Helmich Pedersen, J. M. Sejersted, and E. Vassenden. "Literary Quality: Historical Perspectives". In: *Contested Qualities*. Ed. by K. O. Eliassen, J. Hovden, and O. Prytz. Fagbokforlaget, 2018, pp. 47–74. URL: https://interaccio.diba.cat/sites/interaccio.diba.cat/files/contested%5C%5Fqualities%5C%5F0.pdf%5C#page=48.

[10] H. Bloom. *The Western canon : the books and school of the ages*. Harcourt Brace, 1994.

[11] M. Burke. “The neuroaesthetics of prose fiction: pitfalls, parameters and prospects”. In: *Frontiers in Human Neuroscience* 9 (2015). DOI: 10.3389/fnhum.2015.00442. URL: http://journal.frontiersin.org/Article/10.3389/fnhum.2015.00442/abstract.

[12] J. Che, X. Sun, V. Gallardo, and M. Nadal. “Cross-cultural empirical aesthetics”. In: *Progress in Brain Research*. Vol. 237. Elsevier, 2018, pp. 77–103. DOI: 10.1016/bs.pbr.2018.03.002. URL: https://linkinghub.elsevier.com/retrieve/pii/S0079612318300025.

[13] Z. Chen, K. Hu, P. Carpena, P. Bernaola-Galvan, H. E. Stanley, and P. C. Ivanov. “Effect of nonlinear filters on detrended fluctuation analysis”. In: *Phys. Rev. E* 71.1 (2005), p. 011104. DOI: 10.1103/PhysRevE.71.011104. URL: https://link.aps.org/doi/10.1103/PhysRevE.71.011104.

[14] J. Cheng. “Fleshing out models of gender in English-language novels (1850–2000)”. In: *Journal of Cultural Analytics* 5.1 (2020), p. 11652. DOI: https://doi.org/10.22148/001c.11652.

[15] J. Culler. *On Deconstruction: Theory and Criticism After Structuralism*. Cornell University Press, 2007.

[16] J. Derrida and D. Attridge. “”This Strange Institution Called Literature: An Interview with Jaques Derrida””. In: *Acts of literature*. London: Routledge, 2017, pp. 33–75.

[17] P. Feldkamp, Y. Bizzoni, M. R. Thomsen, and K. L. Nielbo. *Measuring Literary Quality. Proxies and Perspectives*. Report. Darmstadt, 2024. DOI: 10.26083/tuprints-00027391. URL: https://tuprints.ulb.tu-darmstadt.de/27391/.

[18] O. d. C. Fialho. “Foregrounding and refamiliarization: understanding readers’ response to literary texts”. In: *Language and Literature: International Journal of Stylistics* 16.2 (2007), pp. 105–123. DOI: 10.1177/0963947007075979. URL: http://journals.sagepub.com/doi/10.1177/0963947007075979.

[19] S. E. Fish. *Is there a text in this class? the authority of interpretive communities*. 12. print. Cambridge, Mass.: Harvard Univ. Press, 2003.

[20] J. Gao, J. Hu, and W.-w. Tung. “Facilitating Joint Chaos and Fractal Analysis of Biosignals through Nonlinear Adaptive Filtering”. In: *PLoS ONE* 6.9 (2011). Ed. by M. Perc, e24331. DOI: 10.1371/journal.pone.0024331. URL: http://dx.plos.org/10.1371/journal.pone.0024331.

[21] E. Gospodinova. “Fractal time series analysis by using entropy and hurst exponent”. In: *Proceedings of the 23rd International Conference on Computer Systems and Technologies*. 2022, pp. 69–75.

[22] F. Greb, P. Elvers, and T. Fischinger. “Trends in Empirical Aesthetics: A Review of the Journal *Empirical Studies of the Arts* from 1983 to 2014”. In: *Empirical Studies of the Arts* 35.1 (2017), pp. 3–26. DOI: 10.1177/0276237415625258. URL: http://journals.sagepub.com/doi/10.1177/0276237415625258.

[23] J. Guillory. *Cultural Capital: The Problem of Literary Canon Formation*. University of Chicago Press, 1995.

[24] R. von Hallberg. “Editor’s Introduction ”Canons””. In: *Critical Inquiry, Special Issue ”Canons”* 10.1 (1983), pp. iii–vi. URL: https://www.jstor.org/stable/1343403.

[25] K. J. Hsü and A. Hsü. “Self-similarity of the ”1/f noise” called music.” In: *Proceedings of the National Academy of Sciences* 88.8 (1991), pp. 3507–3509. DOI: 10.1073/pnas.88.8.3507. eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.88.8.3507. URL: https://www.pnas.org/doi/abs/10.1073/pnas.88.8.3507.

[26] J. Hu, J. Gao, and X. Wang. “Multifractal analysis of sunspot time series: the effects of the 11-year cycle and Fourier truncation”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2009.02 (2009), P02066. DOI: 10.1088/1742-5468/2009/02/p02066. URL: http://stacks.iop.org/1742-5468/2009/i=02/a=P02066?key=crossref.879d2c42ec8804831202df82da8d7a1a.

[27] K. Hu, P. C. Ivanov, Z. Chen, P. Carpena, and H. Eugene Stanley. “Effect of trends on detrended fluctuation analysis”. In: *Physical Review E* 64.1 (2001). DOI: 10.1103/PhysRevE.64.011114. URL: https://link.aps.org/doi/10.1103/PhysRevE.64.011114.

[28] Q. Hu, B. Liu, M. R. Thomsen, J. Gao, and K. L. Nielbo. “Dynamic evolution of sentiments in Never Let Me Go: Insights from multifractal theory and its implications for literary analysis”. In: *Digital Scholarship in the Humanities* 36.2 (2020), pp. 322–332. DOI: 10.1093/llc/fqz092. URL: https://doi.org/10.1093/llc/fqz092.

[29] M. Jockers. *More Syuzhet Validation.* 2016. URL: http://www.matthewjockers.net/2016/08/11/more-syuzhet-validation/.

[30] M. Jockers. “Package ‘Syuzhet’”. In: *URL: https://cran. r-project. org/web/packages/syuzhet* (2017).

[31] M. L. Jockers. *Syuzhet: Extract Sentiment and Plot Arcs from Text.* 2015. URL: https://github.com/mjockers/syuzhet.

[32] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. E. Stanley. “Multifractal detrended fluctuation analysis of nonstationary time series”. In: *Physica A: Statistical Mechanics and its Applications* 316.1-4 (2002), pp. 87–114.

[33] C. A. Knoop, V. Wagner, T. Jacobsen, and W. Menninghaus. “Mapping the aesthetic space of literature “from below””. In: *Poetics* 56 (2016), pp. 35–49. DOI: 10.1016/j.poetic.2016.02.001. URL: https://linkinghub.elsevier.com/retrieve/pii/S0304422X16000127.

[34] C. Koolen, K. van Dalen-Oskam, A. v. Cranenburgh, and E. Nagelhout. “Literary Quality in the Eye of the Dutch Reader: The National Reader Survey”. In: *Poetics* 79 (2020), pp. 1–13. DOI: https://doi.org/10.1016/j.poetic.2020.101439.

[35] L. Kristoufek and M. Vosvrda. “Measuring capital market efficiency: long-term memory, fractal dimension and approximate entropy”. In: *The European Physical Journal B* 87 (2014), pp. 1–9.

[36] I. M. S. Lassen, Y. Bizzoni, T. Peura, M. R. Thomsen, and K. L. Nielbo. “Reviewer Preferences and Gender Disparities in Aesthetic Judgments”. In: *CEUR Workshop Proceedings.* Antwerp, Belgium, 2022, pp. 280–290. URL: https://ceur-ws.org/Vol-3290/short%5C%5Fpaper1885.pdf.

[37] H. Leder, B. Belke, A. Oeberst, and D. Augustin. "A model of aesthetic appreciation and aesthetic judgments". In: *British Journal of Psychology* 95.4 (2004), pp. 489–508. DOI: 10.1348/0007126042369811. URL: https://bpspsychub.onlinelibrary.wiley.com/doi/10.1348/0007126042369811.

[38] P. J. Locher, P. J. Stappers, and K. Overbeeke. "An empirical evaluation of the visual tightness theory of pictorial composition". In: *Acta Psychologica* 103.3 (1999), pp. 261–280. DOI: 10.1016/s0001-6918(99)00044-x. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-0033253633%5C&doi=10.1016%5C%2fs0001-6918%5C%2899%5C%290004 4-x%5C&partnerID=40%5C&md5=74a7e83ba2126c877f0d0409cba48d96.

[39] S. Maharjan, J. Arevalo, M. Montes, F. A. González, and T. Solorio. "A Multi-task Approach to Predict Likability of Books". In: *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*. Valencia, Spain: Association for Computational Linguistics, 2017, pp. 1217–1227. URL: https://aclanthology.org/E17-1114.

[40] J. McDonough and A. Herczyński. "Fractal patterns in music". In: *Chaos, Solitons & Fractals* 170 (2023), p. 113315. DOI: 10.1016/j.chaos.2023.113315. URL: https://linkinghub.elsevier.com/retrieve/pii/S0960077923002163.

[41] D. S. Miall and D. Kuiken. "Foregrounding, defamiliarization, and affect: Response to literary stories". In: *Poetics* 22.5 (1994), pp. 389–407. DOI: https://doi.org/10.1016/0304-422X(94)00011-5. URL: https://www.sciencedirect.com/science/article/pii/0304422X94000115.

[42] D. S. Miall and D. Kuiken. "The form of reading: Empirical studies of literariness". In: *Poetics* 25.6 (1998), pp. 327–341. DOI: https://doi.org/10.1016/S0304-422X(98)90003-1. URL: https://www.sciencedirect.com/science/article/pii/S0304422X98900031.

[43] M. Mohseni, V. Gast, and C. Redies. "Fractality and Variability in Canonical and Non-Canonical English Fiction and in Non-Fictional Texts". In: *Frontiers in Psychology* 12 (2021). URL: https://www.frontiersin.org/article/10.3389/fpsyg.2021.599063.

[44] M. Mohseni, C. Redies, and V. Gast. "Approximate Entropy in Canonical and Non-Canonical Fiction". In: *Entropy* 24.2 (2022), p. 278. DOI: 10.3390/e24020278. URL: https://www.mdpi.com/1099-4300/24/2/278.

[45] E. Öhman, Y. Bizzoni, P. Feldkamp Moreira, and K. Nielbo. "EmotionArcs: Emotion Arcs for 9,000 Literary Texts". In: *Proceedings of the 8th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature (LaTeCH-CLfL 2024)*. Ed. by Y. Bizzoni, S. Degaetano-Ortlieb, A. Kazantseva, and S. Szpakowicz. St. Julians, Malta: Association for Computational Linguistics, 2024, pp. 51–66. URL: https://aclanthology.org/2024.latechclfl-1.7.

[46] W. van Peer. "Ideology or aesthetic quality?" In: *The quality of literature: linguistic studies in literary evaluation*. Ed. by W. van Peer. Amsterdam ; Philadelphia: John Benjamins Publishing, 2008, pp. 17–29.

[47] M. Pelowski, G. Gerger, Y. Chetouani, P. S. Markey, and H. Leder. “But Is It really Art? The classification of images as ”Art”/”Not Art” and correlation with appraisal and viewer interpersonal differences”. In: *Frontiers in Psychology* 8.Oct (2017). DOI: 10.3389/fpsyg.2017.01729. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85030772360%5C&doi=10.3389%5C%2ffpsyg.2017.01729%5C&partnerID=40%5C&md5=ef6d2c7f622be87d9ec2639df7fab9e1.

[48] C. K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger. “Mosaic organization of DNA nucleotides”. In: *Physical review E* 49.2 (1994), p. 1685.

[49] A. J. Reagan, L. Mitchell, D. Kiley, C. M. Danforth, and P. S. Dodds. “The emotional arcs of stories are dominated by six basic shapes”. In: *EPJ Data Science* 5.1 (2016), pp. 1–12. DOI: https://doi.org/10.1140/epjds/s13688-016-0093-1.

[50] C. Redies. “A universal model of esthetic perception based on the sensory coding of natural stimuli”. In: *Spatial Vision* 21.1-2 (2007), pp. 97–117. DOI: 10.1163/156856807782753886.

[51] C. Redies, J. Hasenstein, and J. Denzler. “Fractal-like image statistics in visual art: similarity to natural scenes”. In: *Spatial Vision* 21.1-2 (2008), pp. 137–148. DOI: 10.1163/156856807782753921. URL: https://brill.com/view/journals/sv/21/1-2/article-p137%5C%5F9.xml.

[52] M. A. Riley, S. Bonnette, N. Kuznetsov, S. Wallot, and J. Gao. “A tutorial introduction to adaptive fractal analysis”. In: *Frontiers in Physiology* 3 (2012). DOI: 10.3389/fphys.2012.00371. URL: http://journal.frontiersin.org/article/10.3389/fphys.2012.00371/abstract.

[53] E. W. Said. *Culture and Imperialism.* 1st Vintage Books ed. New York: Vintage Books, 1994.

[54] I. Schindler, G. Hosoya, W. Menninghaus, U. Beermann, V. Wagner, M. Eid, and K. R. Scherer. “Measuring aesthetic emotions: A review of the literature and a new assessment tool”. In: *PLoS ONE* 12.6 (2017). DOI: 10.1371/journal.pone.0178899. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85020459006%5C&doi=10.1371%5C%2fjournal.pone.0178899%5C&partnerID=40%5C&md5=4da91e23d4877bafb6bb5297c7d15432.

[55] G. C. Spivak. “Can the Subaltern Speak?” In: *Die Philosophin* 14.27 (1988), pp. 42–58. DOI: 10.5840/philosophin200314275.

[56] R. P. Taylor. “Order in Pollock’s Chaos”. In: *Scientific American* 287.6 (2002), pp. 116–121. DOI: 10.1038/scientificamerican1202-116. URL: https://www.scientificamerican.com/article/order-in-pollocks-chaos.

[57] T. Underwood, D. Bamman, and S. Lee. “The transformation of gender in English-language fiction”. In: *Journal of Cultural Analytics* 3.2 (2018), p. 11035. DOI: https://doi.org/10.22148/16.019.

[58] O. Vartanian. “The Contributions of Emotion and Reward to Aesthetic Judgment of Visual Art”. In: *Brain, Beauty, and Art.* Ed. by A. Chatterjee and E. Cardilo. Oxford University PressNew York, 2022, pp. 83–87. DOI: 10.1093/oso/9780197513620.003.0017. URL: https://academic.oup.com/book/39122/chapter/338534586.

[59] M. Walsh and M. Antoniak. “The Goodreads “Classics”: A Computational Study of Readers, Amazon, and Crowdsourced Amateur Criticism”. In: *Journal of Cultural Analytics* 6.2 (2021). DOI: 10.22148/001c.22221.

[60] X. Wang, B. Yucesoy, O. Varol, T. Eliassi-Rad, and A.-L. Barabási. “Success in Books: Predicting Book Sales Before Publication”. In: *EPJ Data Science* 8.1 (2019), p. 31. DOI: 10.1140/epjds/s13688-019-0208-6.

[61] C. Weir and E. Mandes. *Interpreting visual art: A survey of cognitive research about pictures.* 2017, pp. 1–262. DOI: 10.4324/9781351295444. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85045869972%5C&doi=10.4324%5C%2f9781351295444%5C&partnerID=40%5C&md5=1caeacab0150b346ff11abfca21996a1.

[62] R. Wellek. “The Attack on Literature”. In: *The American Scholar* 42.1 (1972), pp. 27–42. URL: https://www.jstor.org/stable/41207073.

[63] B. Yuri and F. Pascale. “Sentiment Analysis for Literary Texts: Hemingway as a Case-study”. In: *Journal of Data Mining & Digital Humanities* Nlp4dh (2024), p. 13155. DOI: 10.46298/jdmdh.13155. URL: https://jdmdh.episciences.org/13155.