Anouck Kuypers & Xinran Li

**Emotional Patterns in Horror Literature: A Sentiment Analysis Across Time and Text Lengths**

**Abstract**

In this paper, we present a computational analysis of emotional patterns in horror stories spanning across different time periods and text lengths; specifically novel, novella, short story and novelette. A corpus of 118 horror stories was assembled from freely available texts provided by Project Gutenberg, which subsequently underwent the necessary preprocessing. Using the NRC Emotion Lexicon, matching tokens of each story could be linked to positive and negative sentiments, as well as eight specific, basic emotions. This was done to allow for both a sentiment-focused and emotion-focused analysis. The results revealed a range of emotional patterns within the genre. Surprisingly, the horror corpus consists of more positive than negative expressions. However, as stories become more recent, they consistently become less emotionally intensive. Regarding specific emotions, the corpus is characterised heavily by words associated with trust, followed by fear and anticipation, while surprise and disgust is uncommon. Additionally, multiple factors can influence the general emotional patterns of the genre to a certain extent. Works by the same author exhibited similar patterns, but text length also plays an important role. Among others, it influenced both sentiment and emotional clustering. While some patterns emerged regarding the publication date of the stories, both within the sentiment analyses and emotional analyses, most were not clear enough to lead to definitive conclusion, which could in part be due to poor distribution in the corpus.

**1 Introduction**

In the literary world, the genre of horror occupies a distinctive place and remains compelling from its emergence in the 18th century to the present day. Characterized by themes such as the supernatural, the irrational and the grotesque, horror is often defined by literary scholars in relation to all sorts of negative emotions it provokes: the horror story "shocks, or even frightens the reader, or perhaps induces a feeling of repulsion or loathing" (Cuddon, 1984); the horror genre is "focused on evoking emotions of dread, fear and tension" (Prohászková, 2012), etc. How closely the genre is tied to its emotional effects also makes it interesting to examine how emotions unfold within the texts. Horror narratives often involve emotional fluctuations, such as the alternation of intense fear and relief, or the gradual building of suspense followed by a release; these emotional dynamics have been noted in traditional literary studies, and interpreted as a powerful device for engaging readers and creating emotional impacts (Kersh, 2024).

While traditional literary criticism has explored more of the psychological dimensions of emotions in horror, the growing field of computational literary studies has provided us with new means to analyze these emotions. Sentiment analysis techniques have made it possible to quantify and visualize emotional beats, and enabled distant reading of a large corpus, allowing us to identify and compare emotional trends at scale. Building on these approaches, previous studies have identified a set of common, basic emotional arc shapes across diverse works of fiction, based on sentiment scores (Reagan et al., 2016), and have also found associations between specific kinds of emotions and particular literary genres (Kim et al., 2017; Samothrakis & Fasli, 2015). Using the NRC Emotional Lexicon which maps words to both binary sentiment categories and eight basic kinds of emotions (Mohammad & Turney,

2013), we seek to investigate the emotional patterns in horror fiction, asking how these patterns change over time or vary by work lengths.

The remainder of this paper is structured as follows: Section 2 outlines the state of the art by reviewing related research works. Section 3 introduces the corpus we constructed and necessary preprocessing steps. Section 4 explains our methodology. Finally, we present our results in section 5 and discuss the conclusions in section 6.

## 2 Related research

The practice of sentiment analysis is becoming increasingly popular since the turn of the century as it encompasses a variety of tasks such as sentiment classification or emotion detection and can be employed in multiple fields. The technique allows for the detection and analysis of emotional patterns and their intensity within text-based research (Cui et al., 2023, pp. 1-3), which is the main aim of this study's investigation into text-based horror stories. Within the field of computational literary studies, similar sentiment analyses have been performed on an author's oeuvre or on corpora containing stories from select genres. Both the applied methods and promising results of these studies form an important precedent that influenced the approach for the current research.

First and foremost, a study by Samothrakis and Fasli (2015) also takes a genre-oriented approach, but on a broader scale, and questions if it is possible to predict a fictional text's genre based on the emotion it contains. An important distinction is made between time, "which portion of a work of fiction portrays what [emotion]", and type, "what emotions are portrayed in what works of fiction" (p. 2). They analyzed both dimensions for eight literary genres by first calculating the presence of six emotions (anger, disgust, fear, joy, sadness and surprise) at the sentence level using WordNet-Affect, a lexicon that contains words with their associated emotions and sentiments. These emotional representations were subsequently smoothed for each story and extremely random forests were trained to distinguish between the eight genres (pp. 2-3). Samothrakis and Fasli (2015) observed that stories from the horror genre exhibited high levels of emotionality and an especially strong correlation to fear. Nonetheless, it was by far the most misclassified genre; often being labelled as fantasy or science fiction (pp. 4-5). However, this does not entail that horror stories do not contain distinct emotional patterns, especially since the predictions were fairly accurate for the other seven genres which implies that a connection between literary genre and emotional content is present. These results call for a more detailed and thorough sentiment analysis of horror stories. To achieve this, the current study will be applying a different approach; partially influenced by a study by Arroyo-Barrigüete.

Arroyo-Barrigüete (2023) delved into the works of H.P. Lovecraft, a prominent horror author of the early 20th century. While he mainly aims to divide Lovecrafts' work into distinct categories based on emotional content, his research reveals some interesting patterns that could be relevant for the genre as a whole. Instead of detecting six emotions like Samothrakis and Fasli (2015), Arroyo-Barrigüete (2023) employs the eight basic emotions from Plutchik's Wheel of Emotions (joy, sadness, fear, anger, anticipation, surprise, disgust and trust). This allows for a more fine-grained analysis and aligns with the emotions in the NRC Emotion Lexicon, a high-quality and widely used lexicon (p. 2; p. 4). Based on the presence of words associated with these emotions, Arroyo-Barrigüete created clusters of

Lovecraft's stories using Euclidean distances and Ward's linkage method[1]. His cluster analysis revealed that fear and sadness are closely connected as they formed a compact group, which means stories with similar patterns for these emotions will likely form a cluster (pp. 6-7). Furthermore, fear, sadness and trust were by far the most dominant emotions in his corpus, while surprise was the least common (pp. 5-6; p. 7-8). Arroyo-Barrigüete displays that author-specific analyses can successfully identify emotional patterns across an oeuvre. Yet, would a corpus representing the horror genre as a whole potentially display similar patterns or is it Lovecraft's personal style that causes these distinct emotional patterns?

While the aforementioned studies solely investigate the presence of specific emotions, Bizzoni and Feldkamp (2024) explore different tools for uncovering positive and negative sentiment patterns in literary texts by testing their performance on a text known for implicitly expressing emotions (pp. 3-4). Of the dictionary-based approaches, the Syuzhet package performed exceptionally very well (pp. 9-11). Syuzhet was developed by Jockers (2015) can be utilised to plot diachronic sentiment or emotional arcs of a narrative by plotting valence (positive minus negative scores) at the sentence-level. This allows for the analysis of emotional trajectory of a text, *not* the actual content of the plot itself (Jockers, 2015; Reagan et. al., 2016, p. 2).

The current study aims to combine the approaches of the aforementioned studies by analysing both the presence and narrative arcs of positive/negative sentiments as well as distinct emotions to develop a more thorough understanding of the emotional patterns within the horror genre and to see if they potentially correlate in any way.

## 3 Data
### 3.1 Corpus
Our corpus consists of horror stories which are freely provided online by Project Gutenberg under the 'Horror tales' category. Whenever different editions of the same story were available, it was ensured that only the oldest version would be included in the corpus. Additionally, stories that span across multiple books—although very rare—were not included, as they likely do not follow the same structure as a single, self-contained tale. For the same reason, books that form a collection of horror stories were also excluded.

For the remaining 118 qualifying stories, a large datafile was assembled which contained both the raw texts and any informative metadata. The HTML files available on Project Gutenberg provided a limited amount of metadata which could be directly extracted, specifically titles, author names and any applicable subgenres. However, the original publication year[2] of each story was manually annotated, as well as the type of story: 'novel', 'novella', 'novelette' or 'short story'. These four categories are a great indicator of a story's length, which allows us to analyse if they affect the detected emotional patterns in any specific manner. The table below shows the distribution of these categories across the corpus. Naturally, entries in the novelette category are outliers which means that any patterns they may present cannot be safely regarded as characteristics of the overall story

---

[1] Ward's linkage method is an agglomerative clustering approach that groups data by merging clusters in a way that keeps the total within-cluster variation as small as possible at each step. Due to its tendency of generating compact groups, it appears advantageous for classification (Arroyo-Barriguëte 2023).

[2] We opted for the year of a story's first publication since the original year that a story was written could only be found for a small subset of the corpus.

type. Nonetheless, the other categories, especially the short story and novella categories, are more adequately represented.

| Short story | 61 |
|---|---|
| novel | 40 |
| novella | 14 |
| novelette | 3 |

**Table 1**: distribution of short stories, novels, novellas and novelettes in the corpus

Lastly, the corpus contains works by a vast array of different authors; namely 72 in total. This guarantees that the following analyses will reflect the horror genre itself, instead of the works of a few specific authors. The only authors with a notable amount of entries in the corpus are H.P. Lovecraft, who single-handedly wrote 14 stories, and Robert E. Howard, who wrote 11 stories. The finalised corpus, as described here, will be made available as a JSON file alongside this research project.[3]

**3.2 Preprocessing**
Before the analyses can proceed, some basic preprocessing must first be performed on the raw text data. First, any leading and trailing whitespaces have been removed and every text was converted to lowercase, so that capitalized words would not be regarded as different from their non-capitalized counterparts. Additionally, any characters that are not alphabetic characters or whitespaces were removed. Lastly, two different methods were applied to create a representation of each horror story. Each text in the corpus received its own bag-of-words model which contained the unique words within that text and their absolute frequencies. Furthermore, after tokenizing each story at the sentence level, a vocabulary of unique words was created for each sentence. As will be described in section 4, both representations will be utilized to perform different analyses.

**4 Methodology**
**4.1 Sentiment Analysis Framework**
We based our sentiment analysis on the NRC Emotion Lexicon (Mohammad & Turney, 2013), a widely used resource in multiple fields including Natural Language Processing, Digital Humanities and Computational Literary Studies. This lexicon contains a total of 14,182 English words, each annotated with ten score labels of 0 or 1, indicating their association to positive/negative sentiments as well as eight basic emotions: anger, anticipation, disgust, fear, joy, sadness, surprise and trust. As the first word-emotion association lexicon, it has facilitated many influential studies and inspired the development of relevant research tools and projects.

Using such a lexicon, our analysis operated at two levels: one focusing on sentiment polarity (positive vs. negative), and the other targeting specific emotions. Having applied the Bag-of-Words algorithm to the texts in preprocessing, we were able to detect the presence

---

[3] The JSON file can be downloaded at: https://github.com/Anouck-K/CLS_Project_Kuypers_Li

and distribution of sentiment- or emotion-related words according to the NRC Lexicon. For each text in the corpus, we calculated both raw counts of sentiment- or emotion-associated words and their normalized frequencies (raw counts divided by total word count of the text) to enable a fair comparison between texts of different lengths.

These first steps serve as the basis for three subsequent stages of analysis: First, corpus-level analysis which examines general trends in the emotional content of horror works; second, clustering analysis which identifies texts with similar emotional profiles; and finally, emotional arc modelling which tracks emotional trajectories in individual narratives.

### 4.2 Corpus-Level Analysis

To identify broad patterns within the horror genre, we first analyzed the overall distribution and intensity of sentiment and emotion across the dataset. We started by calculating the total number of sentiment- or emotion-associated words in the entire corpus to illustrate the prevalence of each sentiment or emotion category. Then, we proceeded to individual works and ranked all texts by emotional intensity, namely the sum of the normalized frequencies of sentiment- or emotion-related words, and also by normalized frequencies of specific sentiment or emotion categories. This allowed us to locate the most emotional works and inspect their metadata, offering preliminary insights.

Specifically, in the binary sentiment analysis, we studied the relationship between emotional intensity, year of publication and text length by visually assessing pairwise correlations through a scatterplot, which offers an overview of how emotional expression shifts across time periods for each length category.

### 4.3 Clustering

To uncover patterns of similarity among texts, we applied hierarchical clustering to our corpus. This approach groups together texts with similar emotional profiles, in the present study, according to the normalized frequencies of sentiment- or emotion-associated words in each work. Following Arroyo-Barrigüete (2023), we also used Ward's linkage method in combination with Euclidean distance for clustering. Once all the groups were identified, the results were visualized in dendrograms. To explore the potential influences of publication time, text length and personal styles of the authors on emotionality, we enriched the dendrograms with color-coding by publication periods or text length, and added author information alongside each title. Furthermore, we incorporated heatmaps into the dendrograms to display the intensity of each sentiment or emotion category in individual works. This combined visualization simultaneously shows the clustering structure and emotional profiles of individual texts, revealing possible reasons why particular works are grouped together.

### 4.4 Emotional Arc Modelling

To visualize and analyze emotional fluctuations in individual works, we employed a Python reimplementation of Jockers' (2015) *syuzhet* with several adaptations to better fit our own dataset and research objectives.

Modelling emotional trajectories of our horror stories starts with text segmentation. We experimented with two possible approaches: sentence-level segmentation as in *syuzhet*, where texts were tokenized into sentences; and chunk-level segmentation, where each text is divided into 100 equal-sized chunks based on total word count. For each segment, we extracted a vocabulary of unique words, and identified sentiment- or emotion- associated words in the NRC Lexicon.

In our binary sentiment analysis, we modeled emotional valence as in *syuzhet*. For each segment, we calculated separate positive and negative sentiment scores, and computed valence by subtracting the positive score with the negative. To smooth the trajectory, we applied a rolling mean window with a size of 10% of total segments. Moreover, inspired by the emotional content graphs in Samothrakis and Fasli's study (2015), we also tracked the development of prevalent specific emotions by modelling a curve for each of their intensities. Such intensity was calculated as the frequency of the emotion-associated words normalized by segment length, and similarly, smoothing via rolling means was applied. Through these, we hope to look into both overall sentiment progression and more nuanced emotional signals.

## 5 Results
### 5.1 Binary sentiment analysis
### 5.1.1 Distribution analysis
The results of our binary sentiment analysis begin with a somewhat surprising finding: across the entire corpus, positive words (199,767 occurrences) actually outnumber negative ones (171,531). This pattern contradicts the intuitive assumption that horror fiction would be dominated by negative language, as it is a genre known to evoke unpleasant feelings. One possible explanation is that when creating its emotional impact, horror may rely less on pervasive negative expressions, and more on rhetorical techniques such as contrast, irony or sudden shifts of tones: it might be the juxtaposition of peaceful and pleasant moments with disturbing happenings that generates the unsettling effect. Another explanation lies in the general "positivity bias" of the English language: as Kloumann et al. (2012) noted through their analysis of four corpora which are vast and encompassing in terms of size, date range and register, the most frequently used words in English are generally more regarded as positive. This broader linguistic tendency may also partly account for the numerical advantage of positive words, even in the horror genre where the overall emotional tone is expected to be darker.

When we ranked the texts by sentiment scores, a clear contrast emerged between the most positive and negative works. Among the top 10 most positive texts, the majority (8) were the earlier works from before the 20th century, and most of them (7) were full-length novels. In contrast, the most negative texts were highly associated with a single author: 7 out of 10 were written by Robert E. Howard. All 10 were published in the 20th century and were short in length (9 short stories and 1 novelette), suggesting that authorial style may play a key role in shaping the overall emotional profile of a text, and in general, publication era and text length are also likely to have a significant impact. Moreover, the 10 works with the highest

overall emotional intensity show a striking overlap (8) with the most positive texts, while only 1 overlaps with the most negative. This suggests that in our corpus, high emotional intensity may be more commonly driven by an abundance of positive rather than negative words. Our following clustering analysis will further clarify such sentiment distribution through a visual heatmap accompanying clustering results.

Figure 1 further illustrates how the emotional intensity of individual works varies by year of publication, with text length categories indicated through color-coded data points. Such a visualization reveals a noticeable downward trend in emotional intensity over time, particularly for novels and novellas, which are relatively evenly distributed across the timespan. Short stories, which are more concentrated in the first half of the 20th century, exhibit greater variability and do not follow a clear trajectory; but overall, the decline is still apparent. This pattern may reflect the historical development within the genre, such as the shift from Gothic supernatural to psychological terror (Sadiqzade 2025). While the earlier Gothic mode depicts external threats which come with emotionally charged language, the psychological type focuses more on exploring internal states of mind, likely resulting in more emotionally-restrained expression. The trend may also resonate with findings in a broader scope from Morin and Acerbi (2016), who observed a general decrease in emotional expression in Anglophone fiction from the 19th century to the turn of the millenium.
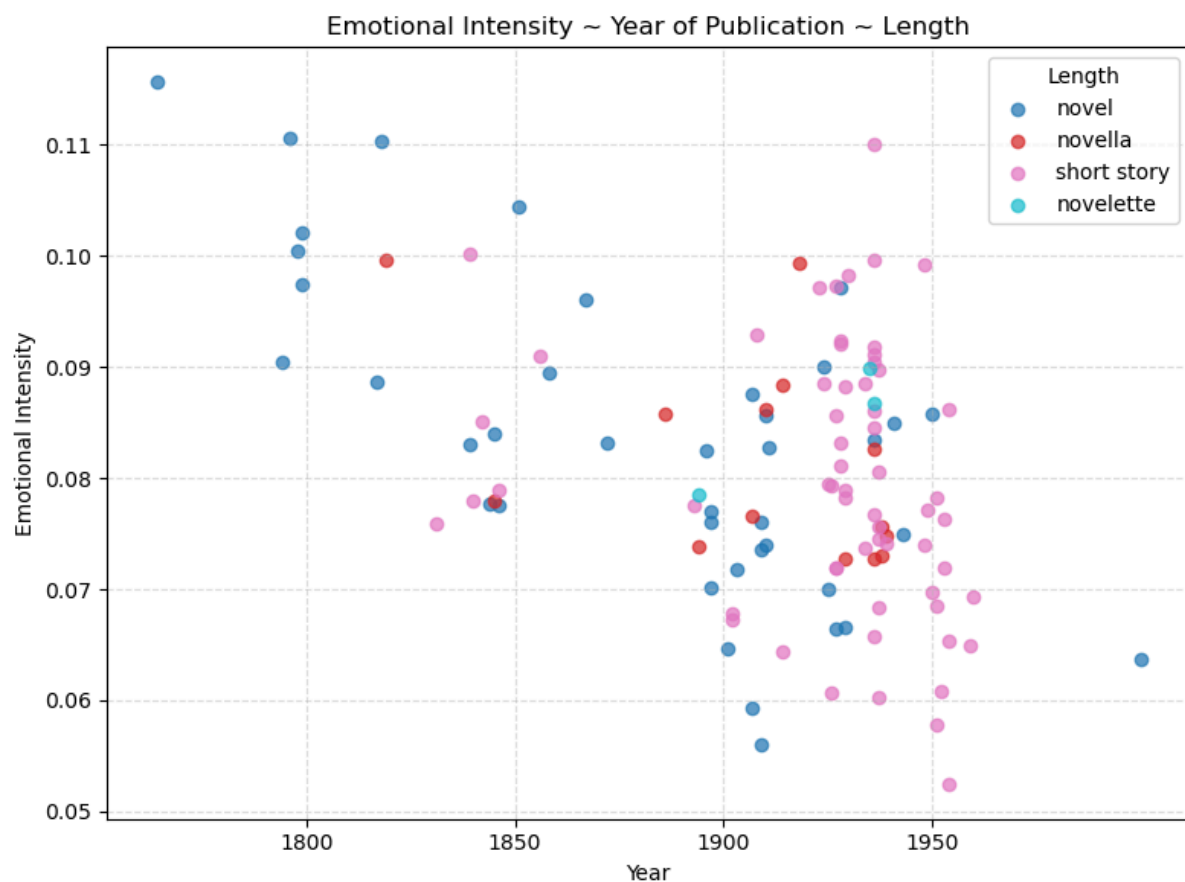


**Figure 1**: Scatter plot of emotional intensity by publication year and text length category

### 5.1.2 Clustering

The hierarchical clustering analysis reveals several patterns in how horror texts group together based on their positive and negative sentiment profiles.

First, text length emerges as a relatively strong predictor of clustering. As shown in the dendrogram colored according to length categories (Appendix A), novels and short stories each form rather distinct clusters, suggesting similar sentiment patterns within these major length categories in our corpus. This likely reflects the influence of substantial difference in narrative scope on sentiment profiles. Notably, even though the novels in our corpus are distributed across a wide timespan, they still cluster more closely than the short stories, which are largely from the early 20th century. This suggests that for a work's sentiment profile, text length may be a more significant determinant than publication era.

The dendrogram colored by time periods (Appendix B) further confirms that temporal factors are less relevant: only the earliest works (pre-1800) fall into a relatively compact cluster, likely due to their higher positivity. Texts from all later periods are to some extent dispersed, reflecting that the positive and negative sentiment patterns may have become more varied over time. While the overall emotional intensity has been steadily declining over time, the balance between positive and negative expressions seems to have developed in more complex ways.

Finally, the author specific pattern we previously observed continues to stand out in the clustering. Towards the bottom of the dendrogram, Robert E. Howard's works cluster tightly together and are also the farthest apart from all other branches. The heatmap made it clear that the distinctive feature of this cluster is exceptionally high negativity and low positivity, reflecting how consistent and extreme Howard's personal style is.

### 5.1.3 Valence arcs

As a preliminary exploration of emotional progression within individual texts in our corpus, we examined their valence arcs, which reflects the sentiment movement throughout the narrative rather than directly representing plot structure (Reagan et al., 2016). Here, we chose to use the original sentence-based arcs to preserve finer-grained local sentiment variations and observe distinctive features of each narrative, instead of chunked arcs which may overly neutralize co-occurring positive and negative signals in larger segments (Jockers, 2015), or super-smoothed normalized arcs via DCT or Fourier transformation as we inspected and compared the arc shapes separately, not overlapped and point-by-point.

While the specific patterns of individual valence arcs vary, some broad tendencies can be observed across the corpus, potentially representing general features of the genre. The majority of the arcs show a clear direction, such as generally rising or falling curves, or structures consisting of rises and falls with a few crucial turning points. These arcs often include a few distinguishable peaks or troughs that stand out from the surrounding sentiment context. The following figures 2 and 3 presents two relatively representative examples: *Edgar Huntly; or, Memoirs of a Sleep-Walker* shows an arc with clear peaks and valleys,

while "Black Canaan" owns an arc with an apparent downward trend[4]. In contrast, only a small subset of texts exhibit relatively flat sentiment trajectories or constant fluctuations around a certain valence level, with more obscure directional movements. The features of the arcs may serve as a reflection of narrative strategies: major shifts in sentiment tones across the narrative likely contribute to creating the horror effect. Although we do not attempt to systematically classify or quantify the various characteristics of the story shapes further here, they can serve as a starting point for relevant future studies which aim to examine story shapes in more detail.
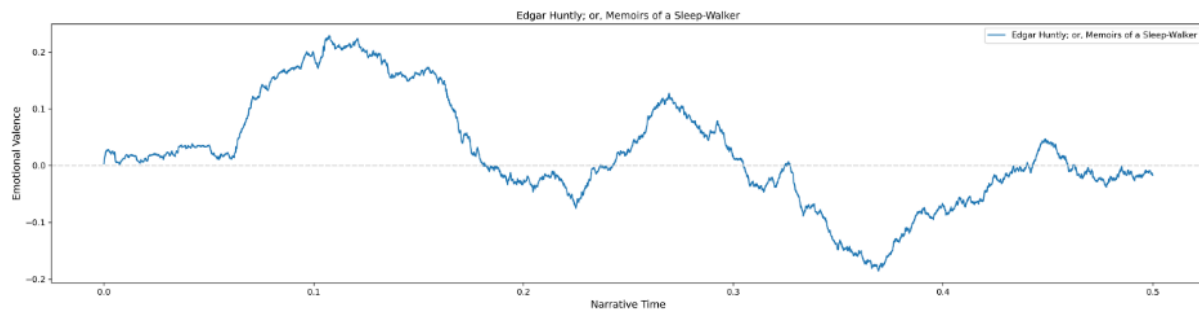


**Figure 2**: Valence arc of *Edgar Huntly; or, Memoirs of a Sleep-Walker*, a novel by Charles Brockden Brown, published in 1799.
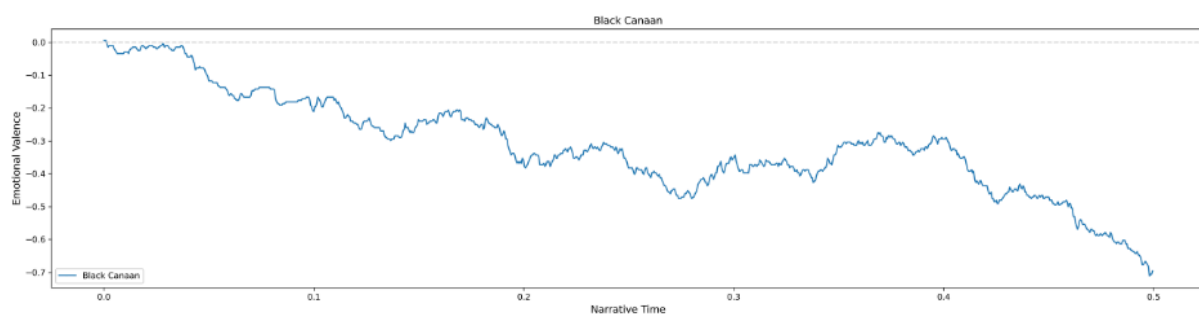


**Figure 3**: Valence arc of *Black Canaan*, a short story by Robert E. Howard, published in 1936.

While it remains challenging to generalize valence arc features across large groups of texts, identifying patterns from an author with a particularly distinctive style can still be relatively straightforward. Once again, the valence arcs of Robert E. Howard's works (like "Black Canaan" above) appear notable for a general decline throughout, and these arcs typically remain below the neutral baseline (the horizontal dashed grey line representing a valence value of 0.0) over the course of the story. This pattern aligns with our previous findings regarding the significant negativity in Howard's works, and presents one more component to his negative sentiment profile: overall decline in valence within individual narratives.

**5.2 Distinct emotions**
**5.2.1 distribution analysis**
By utilising the filtered NRC Lexicon, where only eight basic emotions remain (anger, anticipation, disgust, fear, joy, sadness, surprise and trust), a matrix could be constructed

---

[4] The complete set of figures for the valence arcs, together with the code used for generating them, can also be accessed at https://github.com/Anouck-K/CLS_Project_Kuypers_Li.

where each vector represents one horror story and which contains the absolute frequencies of words linked to each emotion. Table 2 displays the first few rows of this matrix.

|   | trust | fear | sadness | anger | surprise | disgust | joy | anticipation |
|---|---|---|---|---|---|---|---|---|
| **0** | 1552 | 1308 | 1300 | 1030 | 689 | 781 | 859 | 1144 |
| **1** | 922 | 659 | 543 | 477 | 397 | 326 | 304 | 800 |
| **2** | 698 | 815 | 702 | 591 | 343 | 477 | 516 | 670 |
| … | | | | | | | | |

**Table 2**: A matrix containing the absolute frequencies of all emotion-associated words in each horror story of the corpus.

As seen below in Figure 4, the emotional distributions reveal that trust is the most dominant emotion in the corpus by a fairly large margin. It is followed by fear and anticipation, with similar frequencies, and sadness is the fourth most common emotion. On the other hand, words that indicate surprise or disgust seem to be uncharacteristic of the corpus. This aligns in part with the findings of Arroyo-Barrigüete (2023), who observed that fear, sadness and trust are the most common and surprise is the least common in H.P. Lovecraft's work (p. 6). Additionally, despite Samothrakis and Fasli (2015) inspecting only six emotions, he also observed that fear defines horror, while surprise is the least relevant emotion (p. 8). While the observations in Figure 4 are not an exact match, it does display noticeable similarities to both studies which indicate that trust and fear can confidently be seen as staples of the genre's emotional content, while the opposite rings true for surprise. Furthermore, it is interesting that words linked to disgust appear the least in our horror corpus. While Arroyo-Barrigüete (2023) makes no specific observations regarding the emotion, it does rank second to last or last in his story-specific analyses (p. 5). Despite him analysing a specific author, the observed emotional patterns seem to more or less align with those observed in the genre as a whole.

It must be noted that the combined appearance of these emotions is not irrelevant. According to Plutchik's Wheel of Emotions, fear combined with sadness indicates despair, yet trust and fear stands for submission (Dennison, 2024, pp. 4, 6). These potentially relevant emotional combinations will be further explored in section 5.2.2.
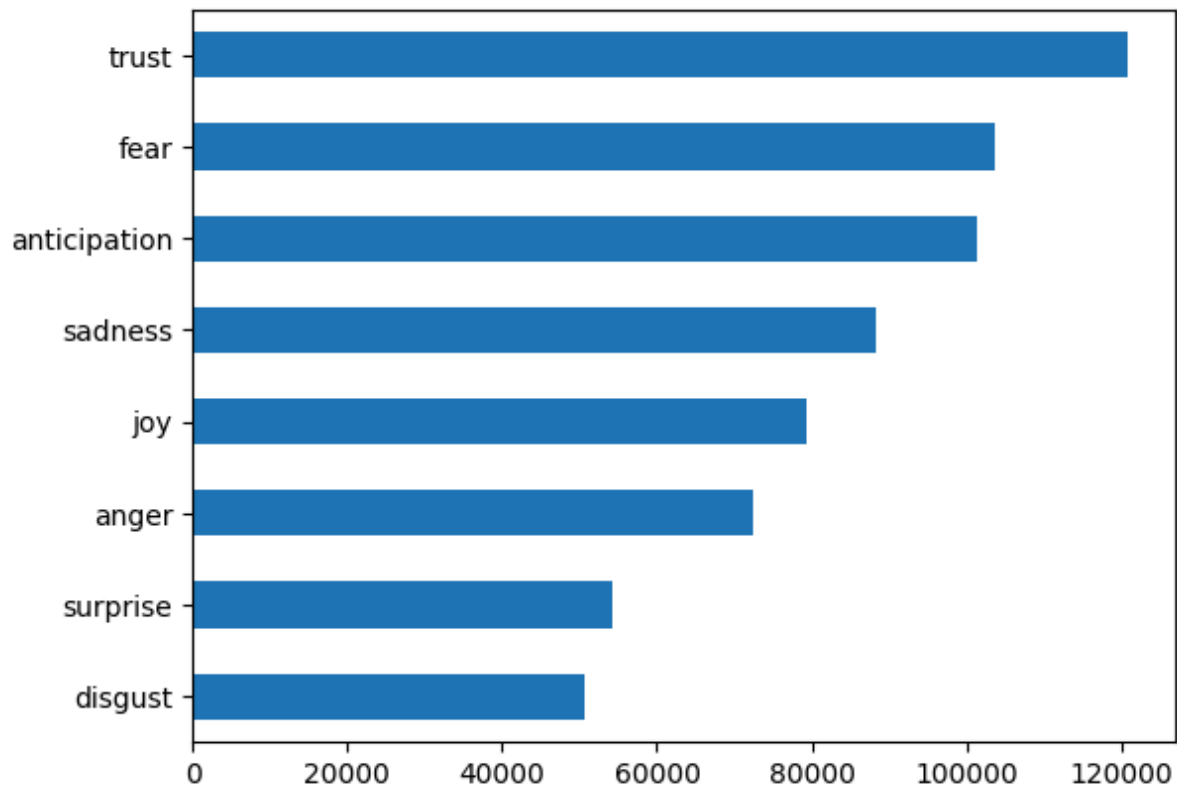
**Figure 4:** Distribution of words associated with trust, fear, anticipation, sadness, joy, anger, surprise and disgust in the horror corpus.

Analysing high, or low, frequencies of specific emotions reveals some interesting patterns in the corpus. In the top 10 stories most associated with fear, seven were short stories, while two were novelettes. Additionally, they tend to originate from the first half of the 20th century. A similar trend emerges for the ten texts most correlated to disgust, consisting of nine short stories and one novelette, and the ten texts with the most words linked to sadness, consisting of seven short stories and one novelette. Thus it seems that, similarly to the sentiment analysis, stories associated with negative emotions like 'fear', 'disgust' or 'sadness' are usually shorter in length. In contrast to this, stories most associated with positive emotions such as 'trust' or 'joy' are full-length novels.

An important side-note is that, Robert E. Howard, whose work is demonstrably negative (see section 5.1.1), fills the majority of the top 10 rankings for 'fear' and 'disgust'; specifically 6 entries of both the 'fear' and 'disgust'' ranking. This once again suggests that authorial style can also affect emotional patterns within a general genre-influenced emotional trend.

### 5.2.2 Clustering
By applying hierarchical clustering to the corpus based on the presence of eight distinct emotions, the dendrograms shown in Appendix C and D were plotted. Firstly, the dendrogram above the heatmaps indicates which emotions behave similarly across the entire corpus. The lines that converge the soonest form the pairs surprise/disgust, trust/anticipation and sadness/anger. These combined emotions respectively stand for unbelief, hope and envy according to Plutchik's Wheel of Emotions. The clustering of surprise and disgust is unsurprising as we observed that they are both extremely

uncharacteristic of the horror genre. However, no clear deduction can be made as of yet regarding the other two pairs. Their respective emotional arcs are likely similar across most stories, which will be investigated further in section 5.2.3.

As can be seen in Appendix C, a clear, large cluster of almost solely novels has formed at the top of the dendrogram which seems to be characterised by a high presence of trust. This indicates that horror novels specifically share similar emotional profiles, which could also be observed in the sentiment analysis in section 5.1.2. Directly below this, a cluster that consists mostly of short stories has formed, which ends at *In the Dark* by Dale Clark. Another short story cluster is visible at the bottom of the dendrogram from *The fearsome touch of death* until *Sea curse*. Most noticeable is their correlation to fear; the first cluster shows moderate to slightly higher levels of trust on the heatmap, while the latter cluster contains a very intense presence of trust. Once again, short stories seem to contain similar emotional profiles. The texts from *He* until *Portrait of a Man with Red Hair A Romantic Macabre* form a mixed cluster of varying text types, seemingly only characterised by low overall emotionality, as indicated by this extremely blue segment on the heatmap.

Appendix D has been colour-coded according to time periods of 50 years. However, it seems that no definite conclusions can be made. Stories from 1900-1949 seem to cluster together, which matches patterns observed in section 5.2.1, but this could also be coincidental due to this time period being highly represented in the corpus.

### 5.2.3 Emotional arcs

As opposed to the sentimental arcs, the emotional arcs can be plotted by dividing the stories into 100 chunks to eliminate noise. The issue of valence scores being neutralised does not occur since the diachronic emotional intensity of specific emotions, which was normalised based on chunk length, is being plotted. For better readability, plots were created for every story for the four most common emotions (trust, fear, anticipation, sadness) in the corpus, and the two least common emotions (surprise, disgust) since, as established, their frequency order slightly differs from prior research.

Seeing as all previous analyses displayed specific patterns based on story length, the plotted emotional arcs were also grouped according to the type of story (novel, novella, novelette, short story). This revealed an extremely interesting pattern. For nearly every single novel, trust was consistently plotted above fear, anticipation and sadness throughout the duration of the story. However, this same pattern does not extend to short stories, where these four emotions display more similar emotional patterns. Figure 5 and 6 are respectively fairly representative of the general patterns observed for novels and short stories. Moreover, the majority of novels displayed the same order of trust appearing consistently at the top of the below, with sadness slightly below it, while anticipation and fear generally displayed lower and similar levels of emotional intensity.
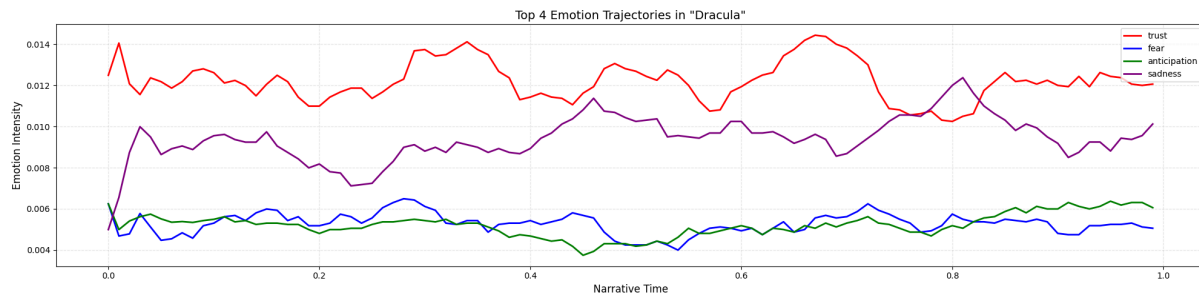
**Figure 5**: Emotional arcs for trust, fear, anticipation and sadness in the novel *Dracula* by Bram Stoker.
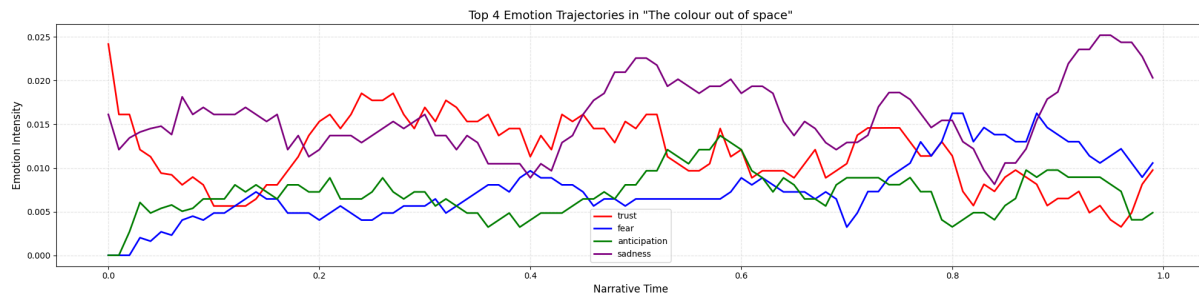


**Figure 6**: Emotional arcs for trust, fear, anticipation and sadness in the short story *The colour out of space* by H.P. Lovecraft.

These patterns could explain why trust ranked as the most common emotion and fear as the second most common in our corpus, despite both prior research (Arroyo-Barrigüete, 2023; Samothrakis & Fasli, 2015) and human intuition pointing to fear being characteristic of horror. It indicates, along with the prior results of this study, that distinctions within patterns of the general horror genre can be made based on story length. Additionally, the arcs for the short stories are often much more angular, while the transitions between chunks are much smoother for novels. Since short stories are naturally shorter, authors likely have to reach each story beat more quickly, resulting in more sudden emotional changes, as shown in the figure below.
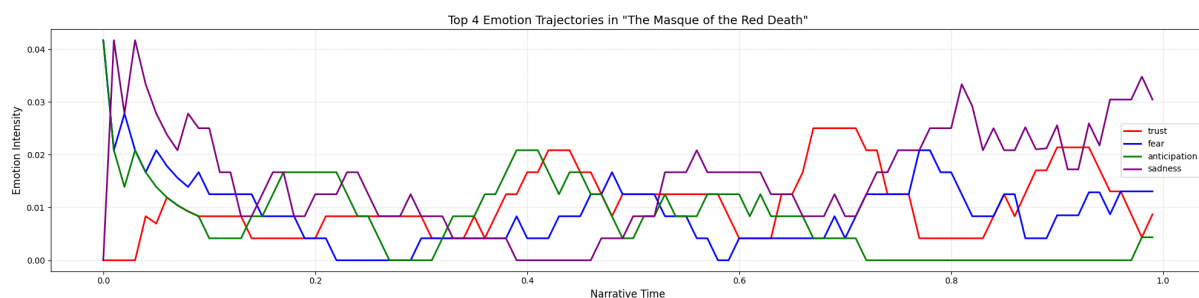


**Figure 6**: Example of angular patterns in the emotional intensities for trust, fear, anticipation and sadness for short stories from *The Masque of the Red Death* by Edgar Allen Poe.

Naturally, the same angular and smooth patterns respectively appear for short stories and novels when comparing their emotional arcs for surprise and disgust. Yet, no clear, general dissimilarities or similarities could be found at the corpus level, or by comparing stories of different lengths. However, there is one pattern that noticeably reappears in a number of plots across the corpus. Quite often, drastically different levels of emotional intensity could

be spotted for surprise and disgust at the beginning of a story. The figures below display what this could look like.
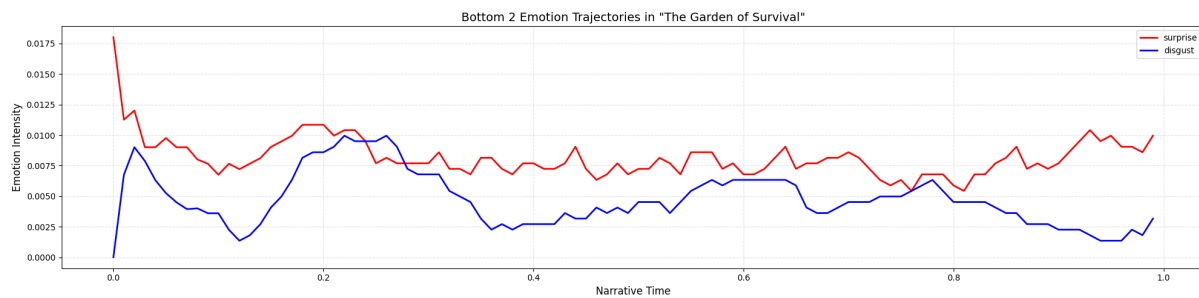


**Figure 7:** Example of emotional arcs for surprise and disgust with unusual intensity levels at the beginning of *The Garden of Survival* by Algernon Blackwood.
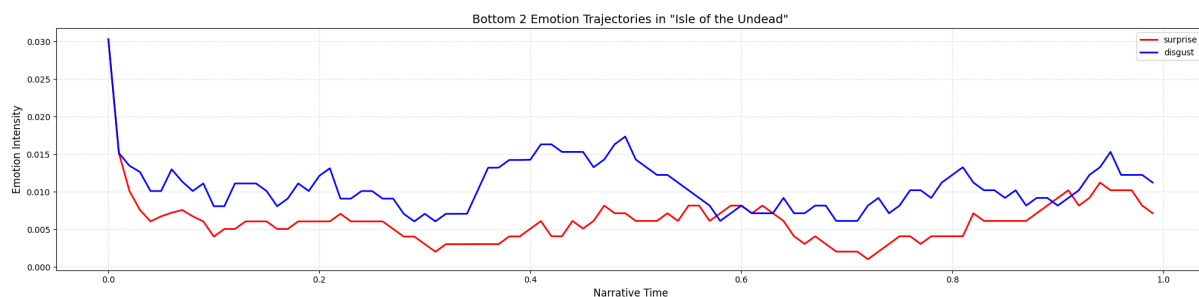


**Figure 8:** Example of emotional arcs for surprise and disgust with unusual intensity levels at the beginning of *Isle of the Undead* by Lloyd Arthur Eshbach.


## 6 Conclusion

This paper analyzes a corpus of 118 English horror stories spanning four length categories and three centuries by employing a range of sentiment analysis methods, which revealed many notable patterns and a series of intriguing findings.

In the binary sentiment analysis, the overall frequency of positive words is surprisingly higher than that of negative ones; longer works are likely to be more positive, while a clear decline in emotional intensity over time is observed. Subsequent hierarchical clustering revealed that text length is a stronger predictor of sentiment profile similarity as opposed to publication era. Distinct novel clusters or short story clusters would form. Additionally, sentence-based valence arcs offered insights into the sentiment progression within individual narratives, with several genre-specific patterns such as distinct rises and falls emerging across the corpus. Throughout this part of the analysis, the works of Robert E. Howard stand out as a compelling case of how distinctive authorial style can significantly influence the sentiment profile of a text even compared to other stories of the same genre.

At the level of eight specific emotions, trust, also generally regarded as a positive emotion, emerges as the most dominant within the genre, followed by fear, anticipation, and sadness, while surprise and disgust are the least relevant. Longer texts also tend to relate to more positive emotions like trust, and this trend is further confirmed by the clustering results, where novels fall in groups characterized by high trust presence. Emotion-specific intensity arcs showed further nuances: as for the top four emotions, most novels have trust prevailing throughout the narrative, whereas for short stories the differences between emotional intensities are far less clear. Interestingly, arcs for surprise and disgust—despite being

14

infrequent overall—often show drastically high frequencies at the beginnings of stories, suggesting that they possibly play a narrative role in story openings.

Overall, our findings challenge certain stereotypical, intuitive and simplistic understandings of the horror genre, such as the impression that horror literature should be completely dominated by negative emotions like fear, and reveal more complex emotional patterns and narrative strategies in representative works. While occasionally counterintuitive, our findings find support in previous computational and traditional literary studies, and also expand their conclusions with new insights. For example, our study provides a possible explanation for the difficulty Samothrakis and Fasli (2015) experienced when trying to predict the horror genre through emotional content: while the genre does display general emotional patterns, factors like text length could also significantly influence the emotional profiles of individual works, possibly complicating genre-level predictions.

While making its own contributions, the present study has certain limitations that open up possibilities for future studies. For instance, the corpus could still benefit from including more recent texts, which would further reflect ongoing developments within the horror genre; the currently exploratory analysis of emotional arcs calls for more systematic classification and generalization of story shapes, etc. In conclusion, by performing sentiment analysis on multiple levels, this study demonstrates the potential of computational methods to understanding genre-level emotional patterns, and suggest directions for further research.

**References**

Arroyo-Barrigüete, J. L. (2023). *Sentiment analysis of Lovecraft's fiction writings. Heliyon, 9*(1), e12673. https://doi.org/10.1016/j.heliyon.2022.e12673

Cuddon, J. A. (Ed.). (1984). *The Penguin book of horror stories*. Penguin.

Cui, J., Wang, Z., Ho, S.-B., & Cambria, E. (2023). Survey on sentiment analysis: Evolution of research methods and topics. *Artificial Intelligence Review, 56*(8), 8469–8510. https://doi.org/10.1007/s10462-022-10386-z

Jockers, M. L. (2015, February 2). *Revealing sentiment and plot arcs with the Syuzhet package*. Matthew Jockers. https://www.matthewjockers.net/2015/02/02/syuzhet/

Kersh, S. E. (2024, September 24). *Horror as an enigma: Emotional or cognitive engagement?* Dickinson College. https://blogs.dickinson.edu/403lit/2024/09/24/horror-as-an-enigma-emotional-or-cognitive-engagement/

Kim, E., Padó, S., & Klinger, R. (2017). Investigating the relationship between literary genres and emotional plot development. In *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature* (pp. 17–26). Association for Computational Linguistics.

Kloumann, I. M., Danforth, C. M., Harris, K. D., Bliss, C. A., & Dodds, P. S. (2012). Positivity of the English language. *PLOS ONE, 7*(1), e29484. https://doi.org/10.1371/journal.pone.0029484

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational Intelligence, 29*(3), 436–465. https://doi.org/10.1111/j.1467-8640.2012.00460.x

Morin, O., & Acerbi, A. (2016). Birth of the cool: A two-centuries decline in emotional expression in Anglophone fiction. *Cognition and Emotion, 31*(8), 1663–1675. https://doi.org/10.1080/02699931.2016.1260528

Prohászková, V. (2012). The genre of horror. *American International Journal of Contemporary Research, 2*(4), 132–142.

Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M., & Dodds, P. S. (2016). The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science, 5*, 31. https://doi.org/10.1140/epjds/s13688-016-0093-1

Sadiqzade, Z. (2025). The Gothic legacy: How 19th-century English writers shaped the modern horror genre. *Global Spectrum of Research and Humanities, 2*(3), 35–41.

Samothrakis, S., & Fasli, M. (2015). Emotional sentence annotation helps predict fiction genre. *PLOS ONE, 10*(11), e0141922. https://doi.org/10.1371/journal.pone.0141922

Semeraro, A., Vilella, S., & Ruffo, G. (2021). PyPlutchik: Visualising and comparing emotion-annotated corpora. *PLOS ONE, 16*(9), e0256503. https://doi.org/10.1371/journal.pone.0256503

Yuri, B., & Pascale, F. (2024). Sentiment analysis for literary texts: Hemingway as a case-study. *Journal of Data Mining & Digital Humanities (NLP4DH)*. https://doi.org/10.46298/jdmdh.13155
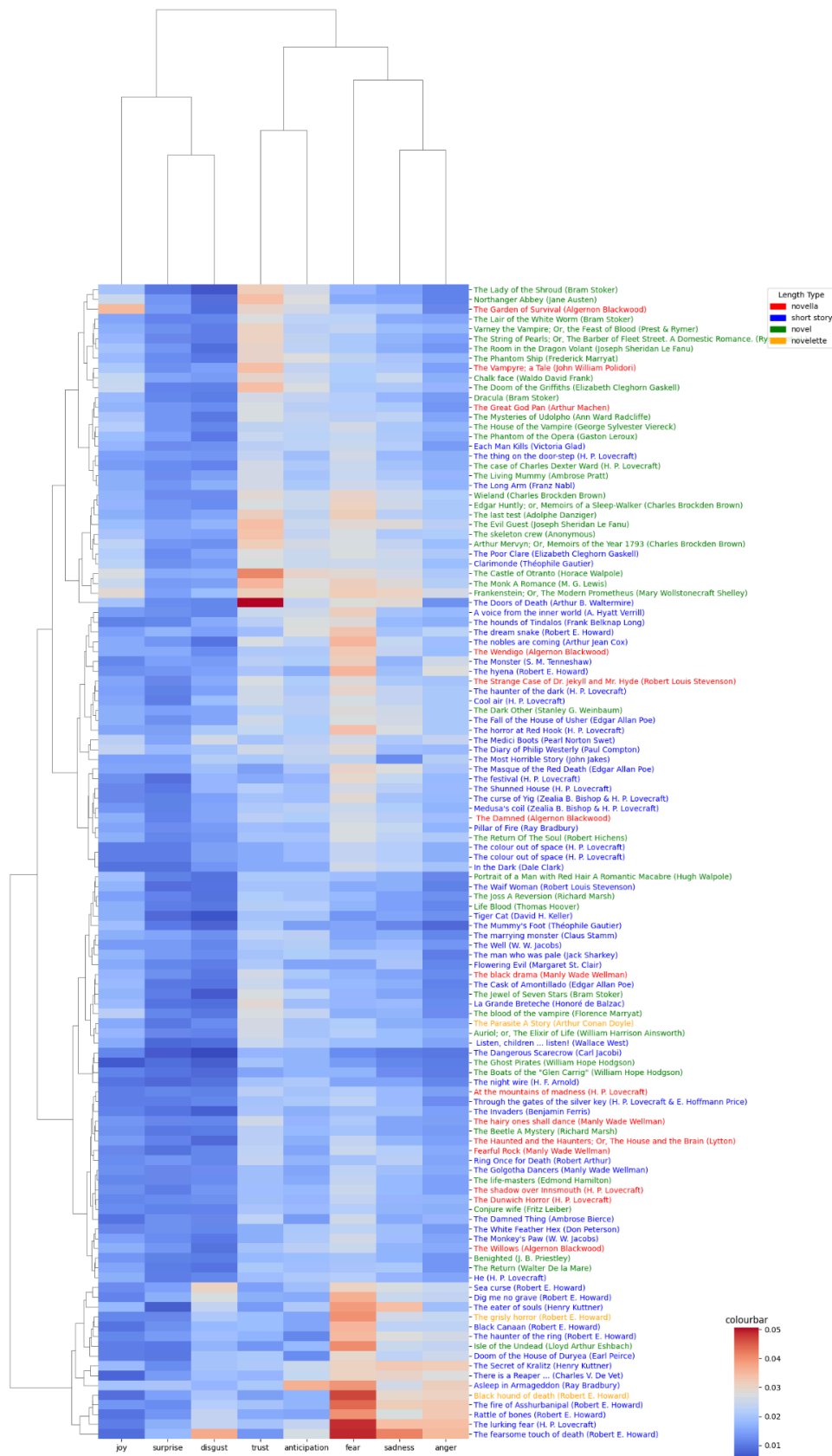
**Appendix**

Appendix A. Hierarchical clustering of all texts based on positive and negative sentiment profiles, color-coded by length type.
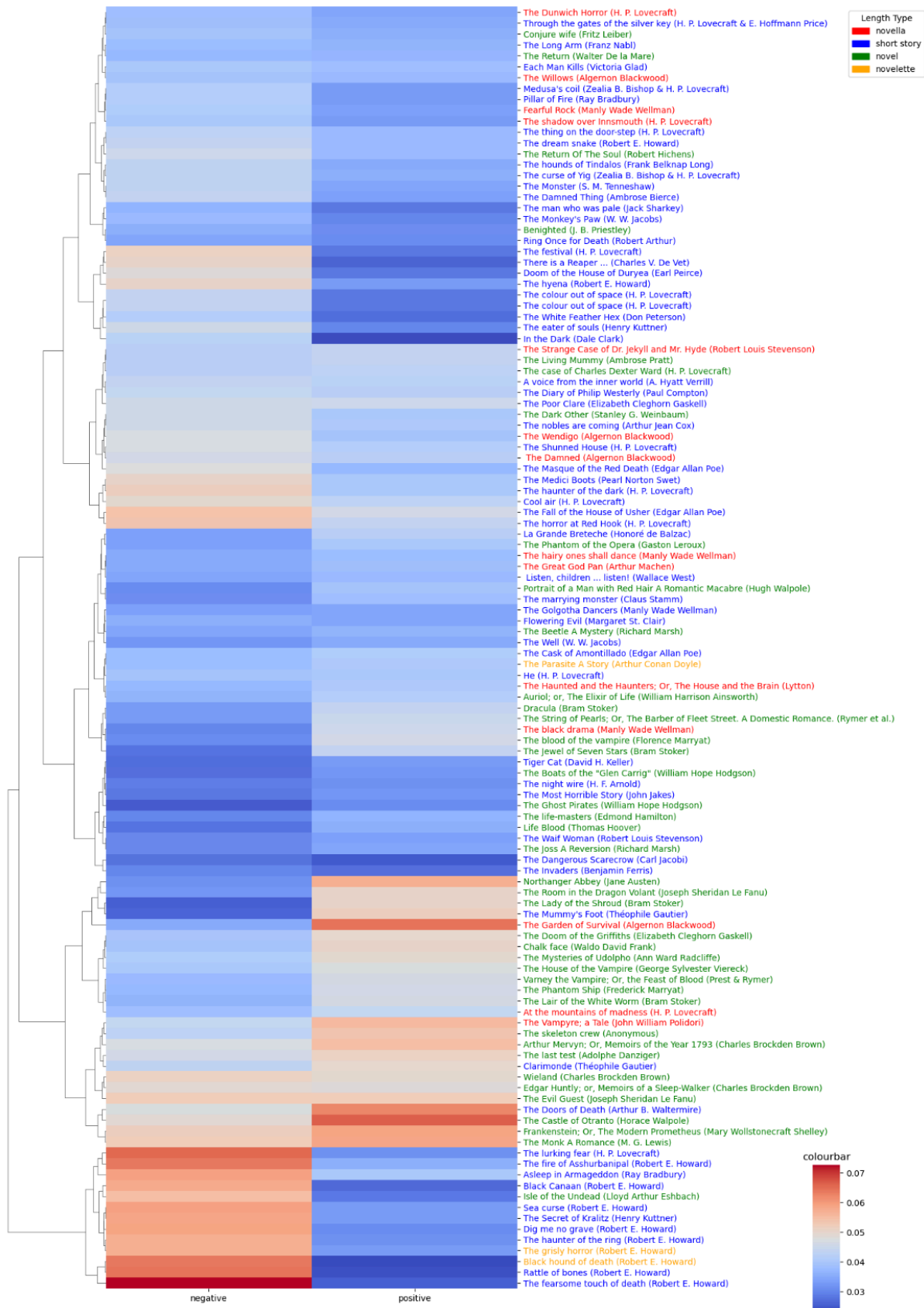
Appendix B. Hierarchical clustering of all texts based on positive and negative sentiment profiles, color-coded by publication era.
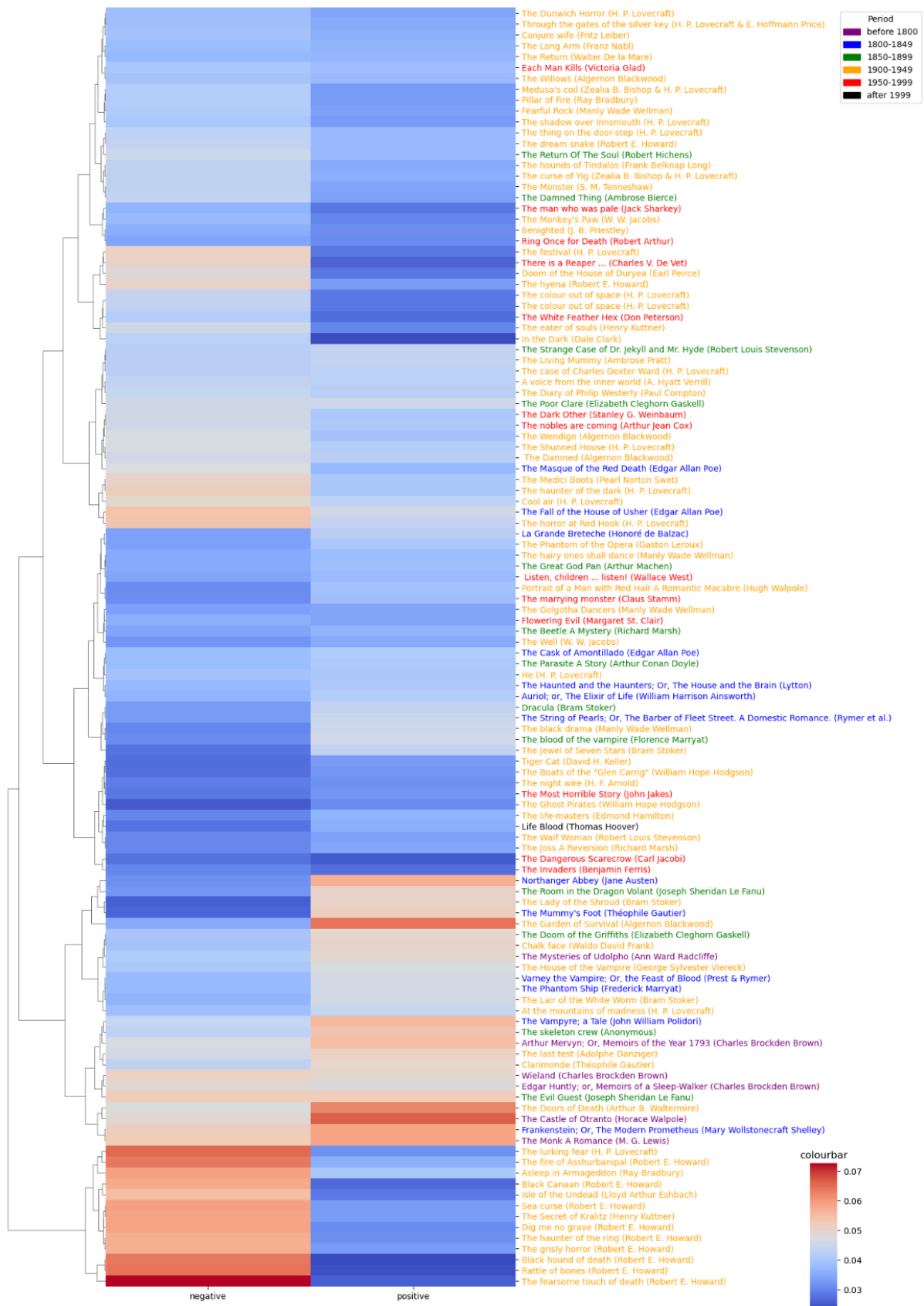
Appendix C. Hierarchical clustering of all texts based on emotional profiles containing 8 distinct emotions, color-coded by length type.

Anouck Kuypers & Xinran Li

Appendix D. Hierarchical clustering of all texts based on emotional profiles containing 8 distinct emotions color-coded by publication period