

# Affect as a proxy for literary mood

Emily Öhman<sup>1</sup> and Riikka Rossi<sup>2</sup>

<sup>1</sup>Waseda University, Japan

<sup>2</sup>University of Helsinki, Finland

Corresponding author: Emily Öhman , ohman@waseda.jp

## Abstract

We propose to use affect as a proxy for mood in literary texts. In this study, we explore the differences in computationally detecting tone versus detecting mood. Methodologically we utilize affective word embeddings to look at the affective distribution in different text segments. We also present a simple yet efficient and effective method of enhancing emotion lexicons to take both semantic shift and the domain of the text into account producing real-world congruent results closely matching both contemporary and modern qualitative analyses.

## I INTRODUCTION

In this study, we explore how the literary concept of mood can be studied and detected with computational methods. We propose to use affect as a proxy for mood and test our hypothesis first quantitatively on different segmentations of the different texts, and then qualitatively against expert close-readings of the same texts. We use multiple different natural language processing (NLP) approaches to attempt to identify the origins, construction, and location of mood. For this purpose, we have collected a corpus of nearly 1000 literary works published in Finnish around the year 1900. Our paper utilizes many common NLP methods and approaches from computational literary studies and affective computing, but besides our pilot study [Öhman and Rossi, 2022], to our knowledge, it is the first study to combine emotion detection/sentiment analysis with the study of mood in texts.

We focus on *mood* as it is one of the more ephemeral yet pervasive aspects of a literary text. Mood is typically described as the atmosphere that the author creates through their word choices, style, and use of imagery and can sometimes even include *tone*. The line between tone and mood can be difficult to draw, particularly when approaching the topic with computational tools, but succinctly the difference can be explained as *mood* being about how the reader feels about the text, and *tone* about how the implied author feels about it and uses words to convey their own attitude towards a topic or subject<sup>1</sup> [Turco, 2020].

Although many different literary tools contribute to the creation of mood, perhaps the most important one, that influences all the other tools as well, is the choice of words. Typically in literature, the intentionality of word choice is higher than in most other text genres (e.g. social media posts) [Mitchell, 2008, Öhman et al., 2020]. It is therefore a great subject for analyzing the use and intensity of emotion-associated words in text. As there are so many components that help create mood, in a larger sense mood is not reducible to a single aspect, but generated by a set of textual elements, however, we can use the importance of word choice to our advantage. If

---

<sup>1</sup>On the concepts of tone and mood in literary studies, see Richards [1929], Ngai [2005], Flatley [2008].

we look at the affective distribution of the words in texts using different text segments, we can attempt to pinpoint where in a text mood is created and how that links to affect.

We suggest that the computational study of the valence of the lexicon of a literary text can be valuable in providing an accurate picture of the distribution of the positive and negative valence in a text continuum, and thus help us to better understand the relationship between the linguistic qualities of a text and its perceived emotional effects, particularly the mood of a text. Furthermore, we contribute to the discussion of the most suitable tools for interdisciplinary work including the debate about lexicon-based methods versus machine learning as the most "accurate" [van Atteveldt et al., 2021, Öhman, 2021, Teodorescu and Mohammad, 2022].

## II BACKGROUND & PREVIOUS WORK

Around the turn of the millennium the "affective turn" took place [Smith, 2011, Armstrong, 2014]. It was a shift in attitude regarding the importance of affect in humanities and social science research, including literary studies [Kim and Bianco, 2007]. Literature can be considered a domain where the affective functions of language are of principal importance [Hogan, 2011]. The affective turn has led to a significant increase in research that focuses on the affective side of text in literature [Armstrong, 2014].

The affective power of literary texts has been acknowledged since Aristotle's *Poetics*, but in the 20th century, many schools of thought such as formalism, new criticism, structuralism and post-structuralism orientated the attention to formal and structural aspects of texts, whereas the study of emotions was excluded and considered as susceptible to researchers' subjective emotions. Research topics range from the study of literature and empathy [Keen, 2007] to the study of literature and cognition [Hogan, 2011]), negative affects and tone in texts [Ngai, 2005] to empirical perspectives [Sklar, 2013, Van Lissa et al., 2016], and even emotions specific to Finnish literature [Rossi, 2020, Rossi and Lyytikäinen, 2022].

Recent studies on emotions in Finnish literature demonstrate that Finnish literature presents us with a rich body of work for developing the general theory of literature and emotions and to study the ways in which genre-specific emotional effects vary culturally and historically [Rossi and Lyytikäinen, 2022]. While this research has opened up new perspectives it has also demonstrated there are a number of gaps and complex questions to resolve. Along with the affective turn, the question of a text's overall emotional tone or mood has aroused vivid interest (e.g. Ngai Ngai 2005, Lyytikäinen 2017, and Rossi Rossi 2020). However, a systematic theory of how tone and mood are created and triggered in the reader is still in the works.

We suggest that a study of the emotional valence of the lexicon measured quantitatively provides a new approach that can help with understanding the components of a text's mood [Öhman and Rossi, 2021, 2022]. Furthermore, it has been shown that lexicon-based methods can achieve better results in emotion classification tasks than machine learning models, especially when the text segment size is optimized [Öhman, 2021, Teodorescu and Mohammad, 2022].

Parallel to the affective turn, sentiment analysis became an active field of research within natural language processing and computer science [Mäntylä et al., 2018]. Sentiment analysis and emotion detection has been used with literary works with varying levels of success; Kim and Klinger [2018] provide a substantive overview of sentiment analysis and emotion detection as it is used in CLS. Although many of the papers cited in the review are interesting and innovative, virtually none of them deal with topics that are common in more traditional literary studies. Common CLS topics are genre classification by emotion, story-type, sentiment tracking, and

sentiment recognition (see e.g. Sprugnoli et al. 2016; Schmidt and Burghardt 2018; Amano et al. 2023).

As interesting and innovative as previous CLS work is, it is typically of little use to literary scholars studying affect. These CLS approaches rarely work in harmony with literary analysis in the traditional sense and usually do not even touch upon the topics that interest literary scholars such as tone, mood, and emotion evocation. We hope that our efforts will contribute to merging the talents and knowledge within CLS, digital humanities, NLP, as well as traditional literary studies.

Although there are exceptions (see e.g. Hu et al. 2021 and Herrmann et al. 2019), it is somewhat rare for studies within the field of CLS to have literary experts working on the project, and many such projects rely heavily on the analysis of the quantitative results conducted by experts of NLP rather than experts of literature. This is not a problem in only CLS, but in many other interdisciplinary fields, particularly those with a computational element [Bartlett et al., 2018]. This is why we think it is imperative to conduct CLS (and other interdisciplinary studies) with domain experts and not just NLP knowledge and literary data.

CLS as a field has been criticized for providing either obvious results or ephemeral results that are not robust enough for repeat analysis [Da, 2019]. Furthermore, there seems to be a pervasive belief that state-of-the-art methods from NLP are the most accurate when used on non-standard unstructured data regardless of the research question or the downstream application of such methods. It is commonly suggested that machine learning methods are more accurate than lexicon-based ones (see e.g. van Atteveldt et al. 2021), but several recent papers have suggested this is not the case, especially when dealing with emotional arcs and ideal bin sizes in verbose domains [Öhman, 2021, Teodorescu and Mohammad, 2022]. We have taken the criticism of both camps to heart in an effort to produce robust and reliable but also useful and interpretable results.

### III DATA

Our data collection was simple and straightforward. Although there are R packages and Python libraries in existence for handling Project Gutenberg downloads, Project Gutenberg discourages mass downloads using such methods, thus we used their recommended method for filtering works using http queries and downloading a set smaller amount of books at a time. We downloaded the first<sup>2</sup> 1000 books from Project Gutenberg<sup>3</sup>, with two filter criteria: (1) the language was Finnish, and (2) the text was in utf-8 plain text format. There is currently no method for filtering out translated works so we estimate that the dataset consists of approximately 50% translated texts. The data is publicly available<sup>4</sup>.

We used the simple gutenberg-cleaner<sup>5</sup> to get rid of the preamble and the legal text at the end of the book, then we created a regex to extract key information such as the title, the name of the author, the year of publication, and whether the book was originally written in Finnish. The translation status of the book was extracted based on whether the terms *suomentaja*, *suomennettu*, *suomentanut*, or any version of *kääntäjä/käännös/käännetty* etc.<sup>6</sup> were present within the first ten lines of text after the preamble was removed.

---

<sup>2</sup>Presumably in order of entry to the database.

<sup>3</sup><https://www.gutenberg.org/>

<sup>4</sup><https://github.com/esohman/FinLit-corpus>

<sup>5</sup><https://libraries.io/pypi/gutenberg-cleaner>

<sup>6</sup>All the terms roughly translate to "translator" or "translated".

Although we filtered books based on their encoding, a fairly large number of the books were not actually utf-8 encoded and had to be decoded and re-encoded. We used automatic encoding detection and tried to convert the texts to utf-8, but for some works this failed and in the end due to these encoding issues, our final corpus consists of 975 books instead of 1000. A vast majority (95+%) were written or translated between the years 1850 and 1925 and over 90% after 1880, with only a few instances of older texts, which means that the language used in the texts can be considered Modern Finnish [Forsman Svensson, 2011]. The final data consists of 2,938,032 sentences and 41,417,116 tokens.

### 3.1 The Emotion Intensity Lexicon

We used the Finnish Emotion Intensity Lexicon (FEIL) [Öhman, 2022] as a starting point for detecting affective terms<sup>7</sup>. FEIL is based on the NRC emotion lexicon [Mohammad and Turney, 2013] and emotion intensity lexicon [Mohammad, 2018] and has been adapted for Finnish. It lists words alongside the emotions they are associated with as well as the intensity of the associated emotion as a number between 0 and 1. The emotions roughly correlate with Plutchik's wheel of emotions [Plutchik, 1980] and contains the emotions *anger*, *anticipation*, *disgust*, *fear*, *joy*, *sadness*, and *trust*. We follow the best practices and ethical guidelines as set forth by Mohammad [2022a,b].

## IV METHOD

Finnish is a fairly easy language to work with in terms of NLP. There are numerous high-quality resources that are actively maintained and new tools are constantly being developed for various written standards of Finnish and Finnish-adjacent languages. Nearly all of these resources are also open-source and freely available [Hämäläinen and Alnajjar, 2021]. Therefore, we were privileged enough to have several different lemmatizers and tokenizers at our disposal, some even specifically made for older Finnish texts. We also added to this list by creating a Finnish version of the `chapterize` Python package<sup>8</sup> package.

Once the data was cleaned, we preprocessed it by lemmatizing, splitting it into paragraphs, and tokenizing the texts. We attempted to fine-tune Finnish BERT [Virtanen et al., 2019] to work with our texts (as per Gururangan et al. 2022), but the vocabulary was not improved sufficiently to work with data so different from the original training data for Finnish BERT. In order to find the best tool for our data we tried multiple different lemmatization tools. These tools included the Turku Neural Parser [Kanerva et al., 2019], `murre` [Partanen et al., 2019, Hämäläinen et al., 2021], and both the *experimental* and *news* Finnish spaCy models. In the end we settled for the Turku Neural Parser as the results were the most accurate (see table 1 for an example) and all words were not only parsed but parsed correctly in context as well.

In the example, the Turku Neural Parser was the only one able to correctly parse the nonstandard form *kahvians* – standard form: *kahviansa* – partitive case of 3rd. pers. sing./plur. coffee. Incidentally, in the dissertation of Airio [2009] *kahviansa* is discussed as an example of "parasite words" since it can be mistakenly split into *kahvi* (coffee) and *ansa* (trap), something none of the lemmatizers did. With careful optimism, we take this as a demonstration of how good lemmatizers for morphologically complex languages have become in the past decade.

The different lemmatizers had different strengths and weaknesses. For example, spaCy is very

---

<sup>7</sup><https://github.com/Helsinki-NLP/SELF-FEIL>

<sup>8</sup>The original: <https://pypi.org/project/chapterize/> and the Finnish version: <https://github.com/esohman/chapterize-fi>

<b>Translation</b>	The provost sits down in his rocking chair, stands his pipe on the floor against the table leg, and starts drinking his coffee
<b>Original</b>	Rovasti istuu keinuoliinsa, panee piippunsa lattialle pöydän jalkaa vasten pystyyn ja rupeaa juomaan kahvians
<b>Murre (hist)</b>	(öljy)movasti istua keinuoli panna piippu latija pöytä jalka vaste pysty ja ruveta juoma kahvis
<b>spaCy (news_lg)</b>	Rovasti istuutua keinuoliinsa panee piippu lattia pöytä jalka vasten pystyyn ja rupeata juoda kahvians
<b>spaCy (exp. /w voikko)</b>	rovasti istuutua keinuoli panna piippu lattia pöytä jalka vasten pystyyn ja ruveta juoda kahvians
<b>UralicNLP</b>	Rovastilrovasti istuutua panna piippu lattia pöytä jalka vasten pystyynpysty ja ruveta juomaljuoda
<b>Turku Neural Parser</b>	rovasti istuutua keinuoli panna piippu lattia pöytä jalka vasten pystyyn ja ruveta juoda kahvi
<b>Ideal lemmatization</b>	rovasti istuutua keinuoli panna piippu lattia pöytä jalka vasten pystyyn ja ruveta juoda kahvi

Table 1: Example of lemmatization using different lemmatizers for Finnish.

easy to use and install, it is reasonably fast for this amount of text, and also produces other information that most of the other lemmatizers leave out. It is therefore great for exploring other aspects of the data such as NER and similarity scores. However, the Turku Neural Parser produced the best results in an easy-to-use pipeline<sup>9</sup>. Omorfi [Pirinen, 2015] and FinPos [Silfverberg et al., 2016] might also have been good candidates for lemmatizers, but these proved difficult to install on the platforms available for this project.

Once we had the texts lemmatized and tokenized, we attempted to identify the first three paragraphs of each text. This was a more complex process than expected as even after removing the preambles/headers and footers, miscellaneous metadata of various shape cluttered the start of the book. There was very little uniformity or even commonly recurring patterns of where the actual text of the book or chapters starts. This led us to create a Finnish version of the `chapterize` package for Python, which enabled use to split texts into chapter and recognize opening paragraphs when used in conjunction with the sentence and paragraph ids provided by the `conllu` metadata generated by Turku Neural Parser. We used two different text sections as targets for overall mood detection: the first three paragraphs of each book, and the first 200 tokens from each chapter in each book.

The size of the bins was chosen based on findings by Teodorescu and Mohammad [2022] that suggest even at a few hundred tokens, lexicon-based methods can excel at estimating emotion arcs beyond current state-of-the-art machine learning capabilities. They suggest that lexicon-based approaches are more suitable "for applications where simple, interpretable, low-cost, and low-carbon-footprint systems are desired" [Teodorescu and Mohammad, 2022, 8].

Previous studies [Öhman and Rossi, 2021] have made it clear that although the results from lexicon-based emotion detection can be very accurate in terms of real-world congruency, certain words can easily obfuscate the results. One of the literary works used as quality assessment and proof-of-concept the novel *Rautatie* (tr. as *The Railroad*) by *Juhani Aho* focuses on the novelty of a railroad track coming to a peripheral rural village and naturally contains many instances of the word *railroad*, the fact that the word itself was associated strongly with *trust* in the lexicon skewed the results and were not representative of the level of overall *trust* in the novel. Therefore we opted to remove the term from the lexicon.

On the other hand, FEIL contains mostly contemporary words and their contemporary emotion associations. It is reasonable to assume that these words and their associations would have been subject to semantic shift. Hence we needed to ensure that first and foremost the most common words in our texts that in our subjective opinion have an emotion association are indeed in the lexicon, and secondly that the most common words that match words in the emotion lexicon are labeled correctly for emotions at reasonable intensities. These two steps are iterative and continuous in that they should be repeated whenever the lexicon or lemmatization is altered.

<sup>9</sup><http://turkunlp.org/Turku-neural-parser-pipeline/>

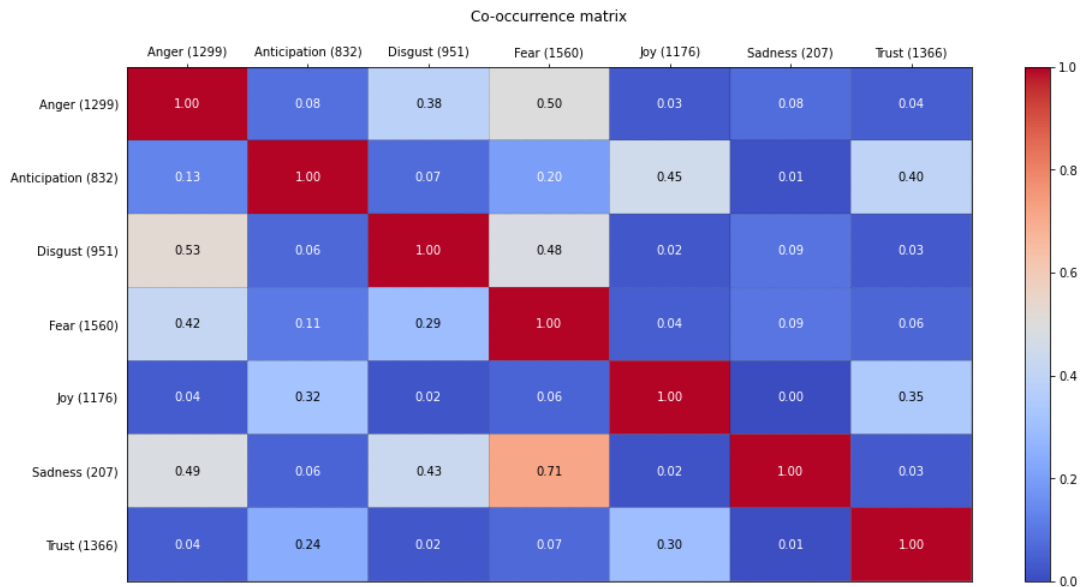


Figure 1: Co-occurrence of emotions in FEIL.

Removing words that carry little to no emotion association is fairly unproblematic, however, altering emotion labels or intensity scores or adding new words to the lexicon should be avoided. If necessary, such additive alterations require the use of multiple annotators who are not the authors and cross-checking the results using inter-annotator agreement scores [van Atteveldt et al., 2021].

For this reason, we did not want to introduce new biases into the lexicon based on our own interpretations of emotion intensities. Thus we developed an alternative method that would allow us to easily and objectively add words to the lexicon. To enable this we created word embeddings of our texts and used them to look up words in the lexicon with high cosine similarity to the words we wanted to introduce to the lexicon (as per e.g. Maas et al. 2011, Yu et al. 2017, Ye et al. 2018). This led to the association for e.g. the words *kirkas*, *valkoinen*, and *valkea* (clear, white, white/light/bright often referring to fire or morning light) to be identical. The words that needed to be added were relatively few so we were able to manually check that their emotion associations and intensities made sense. For future projects, we intend to employ human annotators in conjunction with word embeddings. Nonetheless, this approach alone showed much promise and was very accurate within the small sample size in both recognizing semantically similar words and in mitigating issues with semantic shift.

Ultimately we ended up removing 128 entries from the lexicon and adding 203 tokens including *rakastaa*, to love. As the FEIL lexicon was based on an English lexicon, this exclusion of such a central emotional term demonstrates the issue with noun and verb distinctions in Finnish compared to English. The noun and verb forms are often the same in English unlike in Finnish where the forms are distinct (cf. to love/a love, to run/a run vs. rakastaa/rakkaus, juosta/juoksu). Many of such instances were addressed in the creation of FEIL with the addition of verb forms copying the intensity scores and emotion associations of the English words, yet many verb forms are still missing from the lexicon.

We used this domain- and period-specific version of FEIL to tabulate normalized (per token count for inter-text comparability), intensity scores for each target text. Other future projects should include checking that both noun and verb forms are found in the lexicon.

title	anger	anticipation	disgust	fear	joy	sadness	trust
Putkinotko	4.22	8.67	5.42	11.75	15.69	1.47	22.07
Kauppa-Lopo	12.45	5.71	7.40	19.11	7.66	0	0
Rautatie	8.56	21.15	0	10.50	13.81	5.52	16.74
Hurskas kurjuus	20.33	9.90	11.25	27.73	9.07	7.29	15.05

Table 2: Normalized emotion scores for each novel’s first three paragraphs.

The co-occurrence matrix shows that emotion associations are not linear. A word associated with *anger* is quite likely (0.5) to also be associated with *fear*, but a word associated with *fear* is slightly less likely to be associated with *anger* (0.42). Words associated with *sadness* are highly likely to also be associated with *fear*, however, due to the very low number of *sadness*-associated words in the lexicon, *fear*-associated words are much less likely to simultaneously also evoke *sadness* (0.09). *Joy* is also the only purely positive emotion in the lexicon, which means it is more likely to be associated with many positive emotions beyond *joy* itself due to annotation process conducted with best-worst scaling [Kiritchenko and Mohammad, 2017] (ranking words associated with a specific emotion in terms of least to most associated).

As mentioned earlier, tone and mood can be difficult to distinguish from each other and can even be intertwined to different degrees, especially if we focus on words alone. However, we argue that the tone of a literary text tends to shift much more within and between chapters, and therefore by focusing on the first paragraphs of each chapter, or even the opening paragraphs of the first chapter only, we can get a fairly accurate idea of the mood of the text due to its more stable nature, with a smaller risk of it being confused with tone.

It is well-established that first impressions matter in literature and beginnings tend to shape the experience of reading and set the mood for the whole text. Tone might vary within one text depending on changing narrative viewpoints or even narrators, or switching from descriptive language to dialogue. Theories of perception (e.g. Perry 1979) argue that the openings play a crucial role in creating a text’s overall emotional disposition. The affective language and emotional effects created at the beginning of a text modify and adjust the reader’s general emotional orientation by shaping up modes of perception and organization of information. For instance, the melancholic mood created in the beginning of Aho’s *Rautatie* (Railroad), or the strong effects of disgust in the beginning of Sillanpää’s *Hurskas kurjuus* (Meek Heritage), are likely to influence the reader experiences later reactions and feelings triggered by narrative events.

## V THE MOOD IN SELECTED TEXTS

As the results mostly consist of a dataframe with emotion intensity scores for each text, the results do not easily lend themselves to visual representations. Therefore, we are focusing the presentation of our results on a small subset of data. We chose four texts based mostly on the second author’s area of expertise and previous intimate analyses of the affective landscapes of these texts. The scores for the chosen texts’ first three paragraphs are presented in table form in table 2 and for both approaches comparatively in figures 2 and 3.

### 5.1 An overview of the selected texts

The first one is Juhani Aho’s breakthrough novel *Rautatie* (tr. as *The Railroad*, 1884). In this text, from the perspective of the implied reader, the novel evokes emotional effects of melancholia and nostalgia, which are characteristic of Aho’s work.

The second one is Minna Canth's *Kauppa-Lopo* (no translation, the title refers to the protagonist's nickname, 1889), a tragic story of poverty and illness. The beginning of the novella, set in prison, underlines the anti-hero's ugly appearance, but the narrative contrasts the physical ugliness with an inner goodness: she is described as good-hearted and compassionate towards other people. The story was met with anger in contemporary audience, and the critics did not value her social criticism. Canth's naturalism was considered "poor art" by her contemporaries and she was accused of being an admirer of disgust, "destroying the laws of beauty, unfolding ugliness in every sense" [Rossi, 2007, 52].

The third one is Frans Emil Sillanpää's *Hurskas kurjuus* (tr. as *Meek Heritage*, literally "Sacred Misery", 1919), which begins with a shocking prologue which anticipates the death of the protagonist: it describes the execution of a poor tenant farmer who had ended up as a Red Guard soldier in the Finnish Civil War (1918). Despite the negative emotions and the tragic events the narrator also expresses trust and comfort in the future of the Finnish nation. When Sillanpää's novel appeared it was met with confusion.

The fourth and last one is *Putkinotko* (no translation, 1919-20) by Joel Lehtonen. Like Sillanpää's *Sacred Misery*, this novel tracks the tensions that escalated in the Finnish Civil War in 1918. The novel's protagonist, a good-hearted yet self-willed tenant farmer resigns to obey the landlord and instead resorts to illegal distillery to support the family. The novel is emotionally ambivalent: the idyllic descriptions of Finnish summer nature and the comic elements are likely to arouse positive emotions, while the unembellished description of poverty intends to evoke moral anger and sadness over social inequality.

## 5.2 Results

The results for the texts are presented in figures 2 and 3.

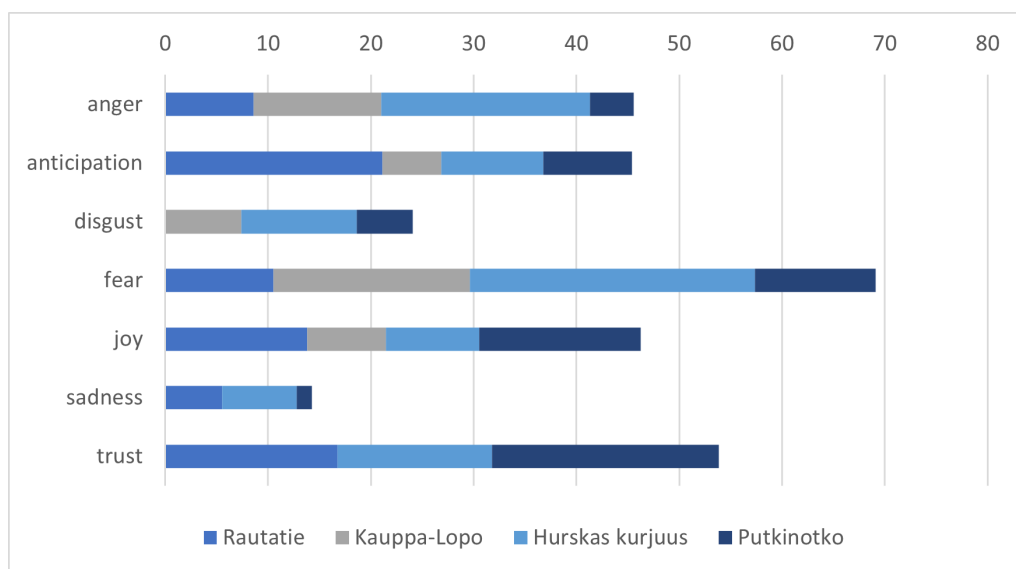


Figure 2: Emotion word distribution in first three paragraphs per 1000 words

In figure 2 we can see some patterns emerge. In particular, the laconism of *Rautatie* is evident when compared to the other authors and the strong early emotional impact of *Hurskas Kurjuus* becomes very apparent, with *fear* and *anger*, but also *sadness* being particularly notable. *Fear* and to some extent *anger* are also highly present in *Kauppa-Lopo*, likely due to the descriptions of the prison environment and the appearance of the protagonist. *Trust* and *joy* are the most



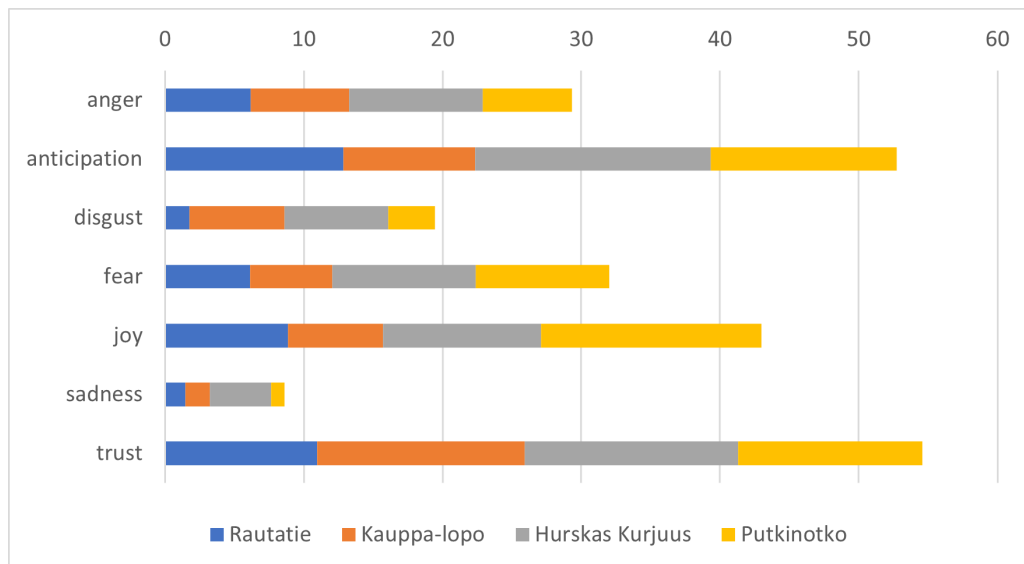


Figure 3: Emotion word distribution in the first 200 tokens of each chapter per 1000 words

notable emotions in *Putkinotko*, perhaps due to the detailed descriptions of the idyllic landscape that dominate the opening chapter.

When comparing the two different segmentations, first paragraph-only vs. opening paragraphs of each chapter, the latter shows a stronger prevalence of positive emotions. We take this to indicate that the intended opening mood for these novels is constructed on negative emotions intended to evoke strong feelings in the reader. In the former segmentation, the emotion distribution is more varied and seemingly converging on the distribution of emotion words in the lexicon.

## VI DISCUSSION

Our computationally derived results correspond highly with qualitative evaluations of the same target texts in terms of established interpretations of mood specifically when compared to the emotion word distribution of the first three paragraphs of a literary text. The valency and intensity of emotions in the opening paragraphs of the book and the opening paragraphs of all chapters the emotions approach the distribution of emotions in the lexicon in the former segmentation. That is, they become muddled because they are not focused enough. Only minor effects of the authors' style can be discerned when applying this method across all chapters. Not only are the differences between chapters reduced with the former, all-chapters, approach, but the differences between the texts also become less clear. This could also be in part because the focus and narrative choices become more varied and therefore the results average out and start to converge on the distribution of emotions in the lexicon. We recommend that the quest for mood should begin with the opening paragraphs of a text.

The qualitative analysis of the selected works, summarized above, has paid attention to various aspects of depicting and triggering emotions in a literary texts:

1. the characters' and the narrators' emotions
2. the emotional effects targeted at the implied reader
3. the empirical readers' reactions in contemporary reception
4. the texts' tone, the organizing feeling of a literary work, which is never reducible to a reader's emotional response to a text, nor to a text's internal representations of feeling (on the concept of the emotional effect, see Lyytikäinen 2017; on the notion of tone, see Ngai

2005, 28.

It should be noted that evoking emotional effects in literature is not restricted to emotion-associated words or to direct descriptions of the character's emotions. All facets of the narrative, from the description of objects to the narrative point of view and style, including tropes and even the rhythm of the text are important aspects in triggering emotional effects in the reader. As an example, the melancholic tone of Juhani Aho's text is not generated by themes of separation and loss alone but also by Aho's style, which favors fragmentation and loosening of syntax, with a recurring mannerism of three points "...", as a sign of hesitation and withdrawal, even evoking a depressive loss of contact.

The qualitative analysis demonstrates that the selected texts depict and trigger negative emotions in particular: feelings of deception, fear, anxiety, disgust and hatred, anger, moral indignation and melancholia. This can be partly explained by genre-specific emotional effects: a critical naturalist novel tends to shock and challenge its reader by representing and inciting strong negative emotions, which confirm the effect of the reality of a text and direct the reader's attention to the social defects described. For instance, the emotion of *disgust*, a genre-specific emotion of the naturalist novel, is a named emotion and salient in Sillanpää's and Canth's novels in particular [Rossi, 2007, 2017, 2020].

The salience of negative emotions can be explained by the importance of negative emotions in literature and art in general. As discussed by Menninghaus et al. [2017] negative emotions are an important resource for the arts, since negative emotions have been shown to be particularly powerful in securing attention, intense emotional involvement, and high memorability, and hence is precisely what artworks strive for. Narrative plots routinely involve social conflicts and both represent and elicit negative emotions in response to such conflicts: failing marriages, unhappy love, long separations, adultery, betrayed friendship, and the like. In narratives, happiness is generally not described in great detail but rather evoked as a peak moment to be challenged, or as a goal to strive for: Canth and Jotuni depict the characters' desire for happiness and love, which is not possible in today's society; Sillanpää's narrator expresses trust in and hope for the future of Finland, and Aho concludes *The Railroad* with the normative happy ending of a fairy tale - yet this happiness is only evoked, but not described in detail.

## VII FUTURE WORK

This dataset will be used for more robust detection of tone and mood in Finnish literature using many novel approaches. The approaches developed for this dataset will also be used with other literary datasets in many other languages. Our preliminary studies show that the "big data" results support qualitative analyses and further justify the use of purely lexicon-based methods and affect as a proxy for mood when dealing with larger collections of text where word choice is an important factor in evoking affective states in the reader. Explicitly, we can see that the choice of emotion-associated words in the first three paragraphs correlates highly with established analyses of mood in the selected texts. We hope to add established emotion categories from literary affect studies (see e.g. Hogan 2011) to the lexicon as a measure to further improve the usability of the FEIL lexicon for the literary domain [Öhman, 2020]. Additionally, we would like to expand on the methodologies used in this exploratory study and hopefully create more and more robust approaches to tone and mood detection in literature and perhaps also incorporate intentionality detection [Guo et al., 2009] to further separate tone from mood.

## ACKNOWLEDGEMENTS

This work was supported by JSPS KAKENHI Grant Number 22K18154.

## References

- Eija Airio. *Morphological Problems in IR and CLIR. Applying linguistic methods and approximate string matching tools*. Tampere University Press, 2009.
- Miki Amano, Kenshin Tsumuraya, Minoru Uehara, and Yoshihiro Adachi. An analysis of representative works of Japanese literature based on emotions and topics. In *Advanced Information Networking and Applications: Proceedings of the 37th International Conference on Advanced Information Networking and Applications (AINA-2023), Volume 2*, pages 99–112. Springer, 2023.
- Nancy Armstrong. The affective turn in contemporary fiction. *Contemporary Literature*, 55(3):441–465, 2014.
- Andrew Bartlett, Jamie Lewis, Luis Reyes-Galindo, and Neil Stephens. The locus of legitimate interpretation in Big Data sciences: Lessons for computational social science from-omic biology and high-energy physics. *Big Data & Society*, 5(1):2053951718768831, 2018.
- Nan Z Da. The computational case against computational literary studies. *Critical inquiry*, 45(3):601–639, 2019.
- Jonathan Flatley. *Affective mapping: melancholia and the politics of modernism*. Harvard University Press Cambridge, MA, 2008.
- Pirkko Forsman Svensson. Virtuaalinen vanha kirjaisuomi. [http://www.vvks.info/aanne-ja\\_muotooppi/heittyminen/](http://www.vvks.info/aanne-ja_muotooppi/heittyminen/), 2011.
- Yi Guo, Nan Hua, and Zhiqing Shao. An algorithm for detection of cognitive intentionality in text analysis. In *2009 International Conference on Information Technology and Computer Science*, volume 2, pages 176–179. IEEE, 2009.
- Suchin Gururangan, Mike Lewis, Ari Holtzman, Noah Smith, and Luke Zettlemoyer. DEMix layers: Disentangling domains for modular language modeling. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 5557–5576, Seattle, United States, July 2022. Association for Computational Linguistics. URL <https://aclanthology.org/2022.naacl-main.407>.
- Mika Hämäläinen and Khalid Alnajjar. The current state of Finnish NLP. In *Proceedings of the Seventh International Workshop on Computational Linguistics of Uralic Languages*, pages 65–72, Syktyvkar, Russia (Online), September 2021. Association for Computational Linguistics. URL <https://aclanthology.org/2021.iwclul-1.8>.
- Mika Hämäläinen, Niko Partanen, and Khalid Alnajjar. Lemmatization of historical old literary Finnish texts in modern orthography. In *Actes de la 28e Conférence sur le Traitement Automatique des Langues Naturelles. Volume 1 : conférence principale*, pages 189–198, Lille, France, 6 2021. ATALA. URL <https://aclanthology.org/2021.jeptalnrecital-taln.18>.
- J Berenike Herrmann, Karola Woll, and Aletta G Dorst. Linguistic metaphor identification in German. In *Metaphor Identification in Multiple Languages. MIPVU around the world*, pages 114–135. John Benjamins, 2019.
- Patrick Colm Hogan. *What literature teaches us about emotion*. Cambridge University Press, 2011.
- Qiyue Hu, Bin Liu, Mads Rosendahl Thomsen, Jianbo Gao, and Kristoffer L Nielbo. Dynamic evolution of sentiments in never let me go: Insights from multifractal theory and its implications for literary analysis. *Digital Scholarship in the Humanities*, 36(2):322–332, 2021.
- Jenna Kanerva, Filip Ginter, and Tapio Salakoski. Universal Lemmatizer: A Sequence to Sequence Model for Lemmatizing Universal Dependencies Treebanks. *arXiv preprint arXiv:1902.00972*, 2019.
- Suzanne Keen. *Empathy and the Novel*. Oxford University Press on Demand, 2007.
- Evgeny Kim and Roman Klinger. A survey on sentiment and emotion analysis for computational literary studies. *arXiv preprint arXiv:1808.03137*, 2018.
- Hosu Kim and Jamie Bianco. *The affective turn: Theorizing the social*. Duke University Press, 2007.
- Svetlana Kiritchenko and Saif Mohammad. Best-worst scaling more reliable than rating scales: A case study on sentiment intensity annotation. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 465–470, Vancouver, Canada, July 2017. Association for Computational Linguistics. doi: 10.18653/v1/P17-2074. URL <https://aclanthology.org/P17-2074>.
- Pirjo Lyytikäinen. How to Study Emotion Effects in Literature: Written Emotions in Edgar Allan Poe’s “The Fall of the House of Usher”. In *Writing Emotions: Theoretical Concepts and Selected Case Studies in Literature*, pages 247–64. Bielefeld: Transcript Verlag, 2017.
- Andrew L. Maas, Raymond E. Daly, Peter T. Pham, Dan Huang, Andrew Y. Ng, and Christopher Potts. Learning word vectors for sentiment analysis. In *Proceedings of the 49th Annual Meeting of the Association for Com-*

- putational Linguistics: Human Language Technologies*, pages 142–150, Portland, Oregon, USA, June 2011. Association for Computational Linguistics. URL <https://aclanthology.org/P11-1015>.
- Winfried Menninghaus, Valentin Wagner, Julian Hanich, Eugen Wassiliwizky, Thomas Jacobsen, and Stefan Koelsch. The distancing-embracing model of the enjoyment of negative emotions in art reception. *Behavioral and Brain Sciences*, 40, 2017.
- Kaye Mitchell. *Intention and Text: Towards an Intentionality of Literary Form*. Bloomsbury Publishing, 2008.
- Saif Mohammad. Ethics sheets for AI tasks. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 8368–8379, Dublin, Ireland, May 2022a. Association for Computational Linguistics. doi: 10.18653/v1/2022.acl-long.573. URL <https://aclanthology.org/2022.acl-long.573>.
- Saif M. Mohammad. Word affect intensities. In *Proceedings of the 11th Edition of the Language Resources and Evaluation Conference (LREC-2018)*, Miyazaki, Japan, 2018.
- Saif M Mohammad. Best practices in the creation and use of emotion lexicons. *arXiv preprint arXiv:2210.07206*, 2022b.
- Saif M. Mohammad and Peter D. Turney. Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29(3):436–465, 2013.
- Mika V. Mäntylä, Daniel Graziotin, and Miikka Kuuttila. The evolution of sentiment analysis—a review of research topics, venues, and top cited papers. *Computer Science Review*, 27:16 – 32, 2018. ISSN 1574-0137. doi: <https://doi.org/10.1016/j.cosrev.2017.10.002>. URL <http://www.sciencedirect.com/science/article/pii/S1574013717300606>.
- Sianne Ngai. *Ugly feelings*, volume 6. Harvard University Press Cambridge, MA, 2005.
- Emily Öhman. Emotion Annotation: Rethinking Emotion Categorization. In *DHN Post-Proceedings*, pages 134–144, 2020.
- Emily Öhman. The validity of lexicon-based emotion analysis in interdisciplinary research. In *Proceedings of the ICON 2021 workshop NLP4DH*. The Association for Computational Linguistics, 2021.
- Emily Öhman. SELF & FEIL: Emotion Lexicons for Finnish. In *Proceedings of the 6th Digital Humanities in the Nordic and Baltic Countries conference*. CEUR Workshop Proceedings, 2022.
- Emily Öhman and Riikka Rossi. Affect and Emotions in Finnish Literature: Combining Qualitative and Quantitative Approaches. In *The Language of Emotions: Building and Applying Computational Methods for Emotion Detection for English and Beyond*, 2021.
- Emily Öhman and Riikka Rossi. Computational Exploration of the Origin of Mood in Literary Texts. *NLP4DH 2022@Asian Association for Computational Linguistics*, page 8, 2022.
- Emily Öhman, Marc Pàmies, Kaisla Kajava, and Jörg Tiedemann. Xed: A multilingual dataset for sentiment analysis and emotion detection. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 6542–6552, 2020.
- Niko Partanen, Mika Hämäläinen, and Khalid Alnajjar. Dialect text normalization to normative standard Finnish. In *Proceedings of the 5th Workshop on Noisy User-generated Text (W-NUT 2019)*, pages 141–146, Hong Kong, China, November 2019. Association for Computational Linguistics. doi: 10.18653/v1/D19-5519. URL <https://aclanthology.org/D19-5519>.
- Menakhem Perry. Literary dynamics: how the order of a text creates its meanings [with an analysis of Faulkner's "A Rose for Emily"]. *Poetics today*, 1(1/2):35–361, 1979.
- Tommi A Pirinen. Omorfi—free and open source morphological lexical database for Finnish. In *Proceedings of the 20th Nordic conference of computational linguistics (NODALIDA 2015)*, pages 313–315, 2015.
- Robert Plutchik. A general psychoevolutionary theory of emotion. *Theories of emotion*, 1:3–31, 1980.
- I.A. Richards. *Practical Criticism*, volume 1964. Routledge: London, 1929.
- Riikka Rossi. *Le naturalisme finlandais. Une conception entropique du quotidien*. Suomalaisen Kirjallisuuden Seura, 2007.
- Riikka Rossi. Writing Disgust, Writing Realities. *Sabine Schönfellner, Gudrun Tockner (eds.) Writing Emotions*, page 277, 2017.
- Riikka Rossi. *Alkukantaisuus ja tunteet — Primitivismi 1900-luvun alun suomalaisessa kirjallisuudessa*. Number 1456 in Suomalaisen Kirjallisuuden Seuran toimituksia. Suomalaisen Kirjallisuuden Seura, 2020.
- Riikka Rossi and Pirjo Lyytikäinen. Pohjoisia tunteita. *AVAIN-Kirjallisuudentutkimuksen aikakauslehti*, 19(1):3–9, 2022.
- Thomas Schmidt and Manuel Burghardt. An Evaluation of Lexicon-based Sentiment Analysis Techniques for the Plays of Gotthold Ephraim Lessing. In *COLING 2018*, page 139. Association for Computational Linguistics, 2018.
- Miikka Silfverberg, Teemu Ruokolainen, Krister Lindén, and Mikko Kurimo. Finnpos: an open-source morphological tagging and lemmatization toolkit for Finnish. *Language Resources and Evaluation*, 50(4):863–878,

2016.

- Howard Sklar. *The Art of Sympathy in Fiction : Forms of Ethical and Emotional Persuasion*. Linguistic Approaches to Literature. John Benjamins Publishing Company, 2013. ISBN 9789027233509. URL <http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=547591&site=ehost-live&scope=site>.
- Rachel Greenwald Smith. Postmodernism and the affective turn. *Twentieth Century Literature*, 57(3/4):423–446, 2011.
- Rachele Sprugnoli, Sara Tonelli, Alessandro Marchetti, and Giovanni Moretti. Towards sentiment analysis for historical texts. *Digital Scholarship in the Humanities*, 31(4):762–772, 2016.
- Daniela Teodorescu and Saif M Mohammad. Frustratingly easy sentiment analysis of text streams: Generating high-quality emotion arcs using emotion lexicons. *arXiv preprint arXiv:2210.07381*, 2022.
- Lewis Turco. *The book of literary terms: the genres of fiction, drama, nonfiction, literary criticism, and scholarship*. University of New Mexico Press, 2020.
- Wouter van Atteveldt, Mariken ACG van der Velden, and Mark Boukes. The validity of sentiment analysis: Comparing manual annotation, crowd-coding, dictionary approaches, and machine learning algorithms. *Communication Methods and Measures*, 15(2):121–140, 2021.
- Caspar J Van Lissa, Marco Caracciolo, Thom Van Duuren, and Bram Van Leuveren. Difficult Empathy-The Effect of Narrative Perspective on Readers’ Engagement with a First-Person Narrator. *Diegesis*, 5(1), 2016.
- Antti Virtanen, Jenna Kanerva, Rami Ilo, Jouni Luoma, Juhani Luotolahti, Tapio Salakoski, Filip Ginter, and Sampo Pyysalo. Multilingual is not enough: BERT for Finnish. *arXiv preprint arXiv:1912.07076*, 2019.
- Zhe Ye, Fang Li, and Timothy Baldwin. Encoding sentiment information into word vectors for sentiment analysis. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 997–1007, Santa Fe, New Mexico, USA, August 2018. Association for Computational Linguistics. URL <https://aclanthology.org/C18-1085>.
- Liang-Chih Yu, Jin Wang, K Robert Lai, and Xuejie Zhang. Refining word embeddings using intensity scores for sentiment analysis. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 26(3):671–681, 2017.