Sentiment Analysis of Literary Texts vs. Reader's Emotional Responses

Tatiana Sherstinova, Anna Moskvina, Margarita Kirina, Asia Karysheva,
Evgenia Kolpashchikova, Polina Maksimenko, Anastasia Seinova, Ruslan Rodionov
National Research University Higher School of Economics, Saint Petersburg
Saint Petersburg, Russia
{tsherstinova, admoskvina, mkirina}@hse.ru
{askarysheva, eokolpaschikova, pimaksimenko, arseynova, rarodionov}@edu.hse.ru

***Abstract*—Sentiment analysis is a relevant task in natural language processing, which is often conducted on Internet texts to analyze reviews of products, services, posts, and comments on social media. Unlike most studies, our work focuses on literary texts in the Russian language written during the emotionally charged period of the early 20th century, when a change in the Russian political system and a fundamental shift in society's way of life occurred as a result of socio-economic upheavals such as wars and revolutions. It is assumed that literary texts from this period should also be rich in emotional vocabulary. The main goal of the research described in this article is to study the correlation between the results of sentiment analysis, performed on the material of Russian short stories using several automatic methods, and the average expert evaluation of the emotions that these same texts evoke in modern readers. The results of the study contribute to understanding the evaluative component of literary works vocabulary and can also be used to build recommendation systems aimed at selecting literary texts for readers.**

I. Introduction

Sentiment analysis refers to a growing field of natural language processing (NLP) and is aimed at detecting emotional valence of text based on its content [1]. Most often, sentiment analysis is applied to Internet texts, namely reviews of goods, services, posts, and comments on social networks, in order to assess consumer opinion [2], [3]. Sociologists and political scientists actively use it to study and evaluate public opinion, as well as the attitudes of specific social groups towards a variety of issues [4]. Analyzing the results of such studies enables the development of more effective strategies for promoting goods and services, shaping public opinion, and conducting election campaigns. Sentiment analysis tools are typically used to characterize the underlying opinions in text in terms of their positivity or negativity. More advanced approaches may also take into account the emotions associated with the text, such as anger, joy, surprise, sadness, and others [5], [6].

Sentiment analysis methods can be broadly classified into two categories: lexicon-based and machine learning-based [7]. Regarding lexicon-based approaches, it should be noted that they typically use not only dictionaries of emotive words but also rules, including grammatical, syntactic, and the lexical ones. The approaches that use dictionaries and linguistic rules are based on the following logic. Each word in the dictionary is associated with a sentiment score. Then, words from the dictionary are compared with words from the analyzed text, and their sentiment scores are calculated according to a proposed formula [1]. When compiling dictionaries, it is worth noting that pre-existing lexicons can be manually expanded to include subject-specific terms. Typically, human assessors are responsible for the markup of these lexical resources (e.g., WordNet, LIWC, ANEW, VADER, and others) [8]. Machine learning approaches can be categorized as either supervised or unsupervised methods [9]. Supervised machine learning approaches typically use classification algorithms.

While the application of sentiment analysis methods to literary texts may not have an obvious utilitarian purpose, it can shed light on the lexical features of such texts. These methods, along with classical stylometric techniques, expand the range of quantitative methods available for philological analysis. Successful examples of sentiment analysis applied to literary texts include the following works: [10], [11], [12], [13], [14].

Our proposed quantitative study aims to explore the relationship between the lexical tonality of Russian stories written approximately 100 years ago, measured by various computer methods, and the emotional responses of readers to these texts, determined through expert evaluation. The results of our study contribute to the understanding of the evaluative component of the vocabulary in literary works. Additionally, these results can be used to enhance the development of recommendation systems that suggest literary texts to readers.

II. Data Description

The study was conducted on a sample of 210 texts from the Russian Short Story Corpus of 1900–1930, created to model the language and style of Russian short stories prose of the period under consideration [15], as well as allowing to conduct digital studies of Russian literature. Unlike most philological resources focused on presenting the legacy of the most famous and popular writers, the Russian Short Story Corpus includes literary texts from a broad range of Russian prose writers, including well-known authors along with less prominent ones, and even virtually forgotten [16], [17]. This resource is actively used for various linguistic, literary, and DH studies [18], [19], [20], [21], [22], [23].

The sample comprises 210 stories that were selected randomly in a manner that accurately represents three historical periods: 1) the pre-war period (1900-1913), 2) the period of acute social cataclysms such as wars and revolutions (1914-1922), and 3) the early Soviet period (1923-1930). To ensure proper representation, only one story per author was included. The sample contains a total of 713,245 words.

Each short story was manually evaluated by 3 experts to obtain an objective reader's assessment. A total of 630 texts were read and annotated. The considerable amount of expert work accounts for a relatively small sample size.

All the stories in the sample, along with their summaries, are available on the Russian Short Story Corpus website at https://russian-short-stories.ru/story [24].

III. Measuring Reader's Emotional Responses to Literary Texts

A. Methodology for obtaining expert evaluation

To obtain expert evaluation, we conducted an experiment designed to measure readers' perception. Participants were asked to read each story and rate their emotional response by evaluating each of the six basic emotions (happiness, sadness, disgust, surprise, anger, fear) [25] they experienced while reading. This list of emotions, originally proposed by Paul Ekman, was chosen as it is also used in SentiArt, a tool for sentiment analysis of literary texts [26] employed later in this study. Additionally, participants were asked to rate each story on a scale of 1 to 10 based on how much they liked it in general. All participants were philology students at the Higher School of Economics in Saint Petersburg.

To prevent the influence of writers' personalities on the evaluation, the participants were not provided with the information on the author. In anonymous form, the majority of the stories are relatively unrecognizable, even among philology specialists. The only exception might have been Ivan Bunin's “Light Breathing”. Each of the 210 texts in the sample was evaluated independently by three participants in the experiment.

The reader's emotional response was evaluated on a three-point scale, where 0 indicated no emotional response, 1 indicated a weak emotional response, and 2 indicated a strong one. Additionally, the participants were given the opportunity to add some comments and provide explanations for any of their ratings.

The overall impression of each story was evaluated using a rating scale ranging from 0 to 10, where 0 represents the lowest score and 10 represents the highest. The participants were responsible for interpreting the scale, and the ratings have been thoroughly analyzed in a separate study [27].

While it may be argued that the respondent pool for the experiments lacks sufficient socio-demographic diversity, we believe that students of philology are quite a sensible choice for the research objectives. All respondents were instructed to approach the task of reading the stories with a calm and neutral emotional state, without any sense of haste or urgency. They were asked to focus solely on reading the stories and assessing their emotional response.

B. Data processing methodology

In order to summarize the data collected from the three respondents, we employed a cumulative sum approach where the points assigned to each emotion were totaled for each story. The final score ranges from 0, if none of the respondents noted a manifestation of the corresponding emotion, to 6, if the emotion received the maximum score from all three respondents.

C. Results obtained

Fig. 1 shows six histograms reflecting the distribution of cumulative scores for each emotion. Four out of six graphs (happiness, disgust, anger, and fear) display a similar pattern, where the predominant rating is “zero”. This indicates that readers either did not experience these emotions while reading the texts or only experienced them to a very weak extent. However, one can see from the histograms that happiness and disgust are more common than anger and fear. Surprise and sadness are more typical reactions, with sadness manifested to a much greater extent.

Fig. 2 contains visualization of the correlations between different emotions evoked by reading. One can note a weak inverse relationship between happiness and sadness ($r = -0.308$) — high indicators of happiness, as it were, displace sadness. Similar relationship, but expressed to an even lesser extent, can be observed between happiness and negative emotions: disgust ($r = -0{,}272$), fear ($r = -0{,}15$) and anger ($r = -0{,}157$). There is a very weak direct correlation between happiness and surprise ($r = 0{,}140$).

A weak direct correlation is observed between sadness, anger, disgust and fear. It is most pronounced between disgust and anger ($r = 0{,}519$), as well as between fear and anger ($r = 0{,}398$). As for surprise, its correlation with negative emotions is low in absolute values. However, one can notice a greater variety of trends concerning this emotion, which requires a separate study.

Table I presents the average values of the cumulative scores for the sample as a whole and separately for each period.

TABLE I. Emotional Assessment Averages for the Sample and Historical Periods

Emotions	Periods			On average
	I (1900-1913)	II (1914-1922)	III (1923-1930)	
Happiness	1,457	1,662	1,870	1,662
Sadness	3,743	3,099	3,130	3,324
Disgust	1,614	1,394	1,754	1,586
Surprise	1,729	1,789	1,826	1,781
Anger	1,057	0,761	0,971	0,929
Fear	1,357	1,070	1,319	1,248

Table I shows that the most frequent emotional response among readers is sadness. Almost all the texts (around 94% of the sample) were characterized by at least one of the respondents as evoking this emotion. Moreover, a significant part of the stories was evaluated as very sad (with scores from 4 to 6). Surprise scores second in terms of frequency as its average score is 1,78 (in 76% of the stories, at least one respondent noted its presence). Happiness and disgust exhibit close average values (1,66 and 1,59, respectively). Different scores of these emotions are noticed in 70% of the texts. Least of all, the readers experienced fear and anger since different degrees of fear are noted in 59% of the stories, of anger — in 53%.

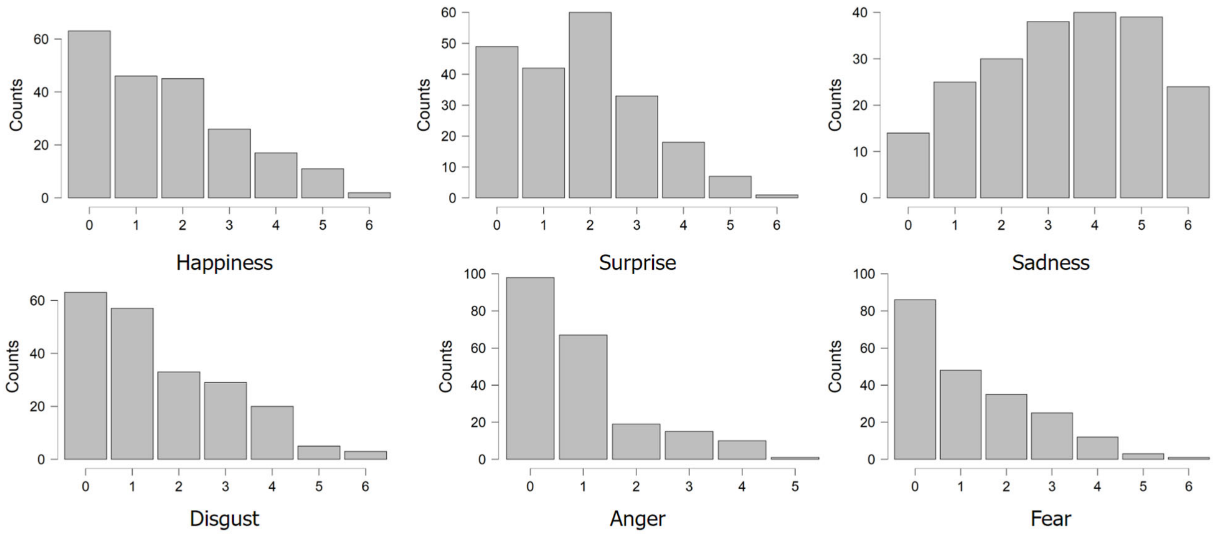

Fig. 1. Distribution of obtained scores across stories for different emotions

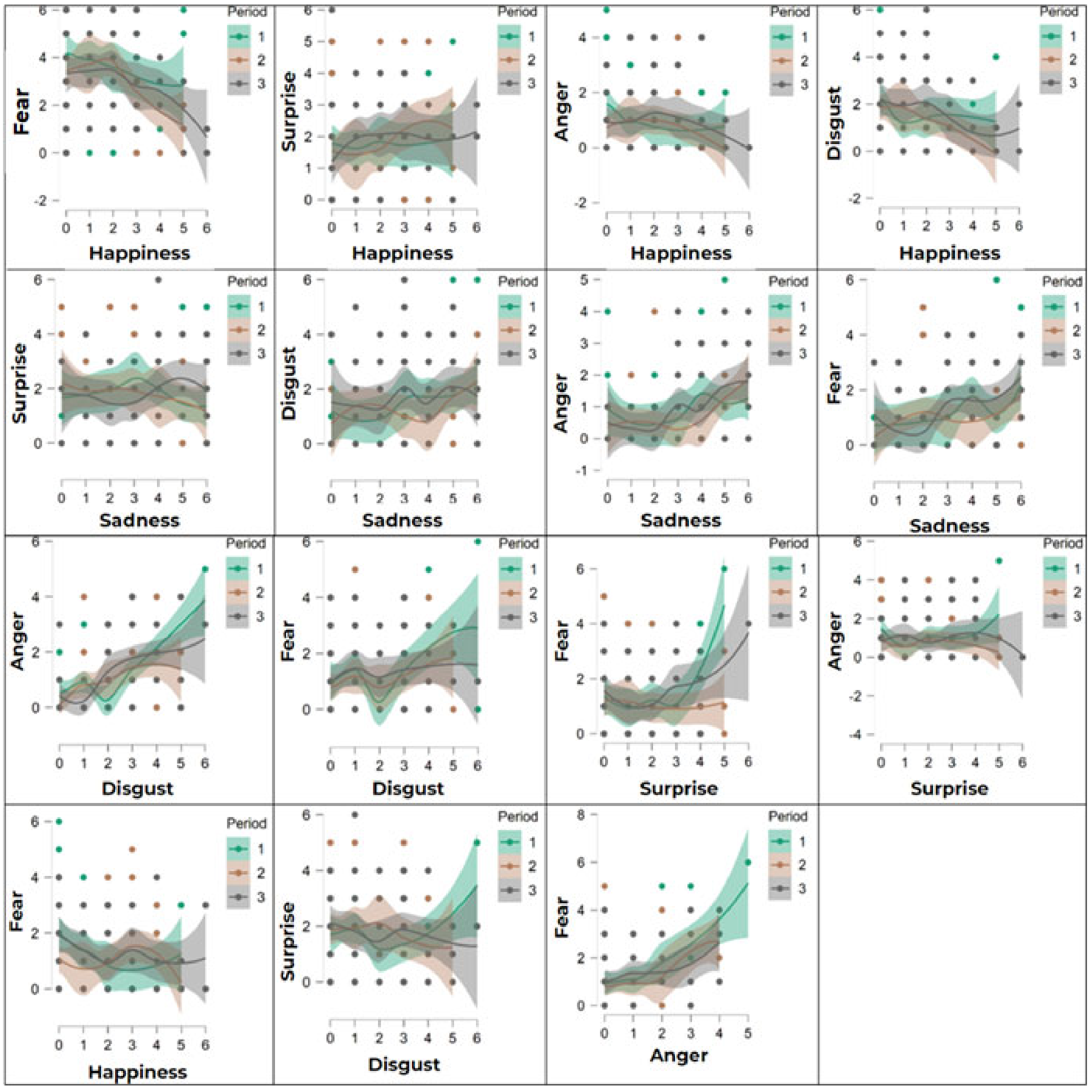

Fig. 2. Visualization of correlation scores for different pairs of emotions, taking into account the period of the story's writing

IV. Sentiment Analysis Methodology

In this research, we use both dictionary-based and machine learning methods.

1) Dictionary-based sentiment analysis

The open-source program used for calculations was Orange [28], which facilitates machine learning and data visualization and also supports sentiment analysis.

Orange provides a dictionary-based approach for sentiment analysis, which relies on lists of evaluative vocabulary. Users have the option to use the default list, which is the Multilingual Sentiment Lexicons (MSL) collection developed by the Data Science Lab [29], or create their own custom dictionaries. The input requires two lists: one with positive words and the other with negative words, each on a new line. Weighted scores are not used here; each word is considered either positive, negative, or neutral. We compare the results obtained by using the Multilingual Sentiment Lexicons (MSL) and the RusSentiLex (RSL) evaluative vocabulary, which is a well-known resource for the Russian language [30]. To ensure comparability, we removed the neutral vocabulary (marked as "neutral") and context-dependent word forms (marked as "positive/negative") from the RusSentiLex dictionary.

To measure the sentiment, Orange employs the technique originally proposed in [31]. The overall sentiment of the text is determined by subtracting the number of distinct negative words from the number of distinct positive words, then dividing the resulting value by the length of the document in words and multiplying it by 100. Therefore, the frequency of words does not impact the sentiment score, and the overall sentiment of the text is primarily determined by the diversity of positive and negative words as well as the length of the document.

When working with Orange, we conducted two sets of calculations: the first set without any preprocessing of the texts, and the second set involved lemmatization of the short stories. Lemmatization was performed using the ru_core_news_sm model of the spaCy library [32].

2) Dostoevsky library

Dostoevsky is a sentiment analysis Python library for the Russian language. It is based on the FastText model which was trained on the RuSentiment dataset introduced in [33]. RuSentiment is an extended database built on a sample of 6950 posts from social network VKontakte, written in Russian. The authors claim that due to the choice of the platform, the RuSentiment data sample differs in the variability of the vocabulary used and the length of publications [ibid.].

The sample was compiled as follows: initially, 31,185 posts were annotated, of which 21,268 were selected randomly. The posts included in the sample "were 10-800 characters in length, at least 50% of which were alphabetical, and at least 30% used the Russian Cyrillic alphabet" [ibid.]. All links, postcards and posts with a large number of hashtags were excluded. The selected data (6,950 posts) was annotated manually by experts. The test dataset for training consisted of 2,967 posts, which were rated on a three-point scale ("positive", "negative", "neutral"). The "positive" and "negative" labels were given to the texts which contained an explicit or implicit expression of an internal emotional state ("mood") or attitude to an object ("evaluation") — positive or negative, respectively. Texts in the "neutral" category were defined as not containing expressed sentiment (e.g., factual questions, descriptions, commercial information). Furthermore, the labels "speech act" and "skip" were implemented. The "speech act" category included posts containing speech acts frequently found in the data (expressions of gratitude, greetings, congratulations), which were not included in the "positive" category because they are "very formulaic". The "skip" label was given to non-Russian, "noisy" or "unclear" texts [ibid.].

As a result of processing the sample using the Dostoevsky library, all texts obtained the probability values for five labels ("positive", "negative", "neutral", "speech", "skip"), ranging from 0 to 1. The scores were calculated in three variants: 1) for words, 2) for sentences, and 3) for a whole text.

3) SentiArt

SentiArt is a tool for sentiment analysis of literary texts based on a vector space model, which was introduced in [26]. Unlike dictionary or word-list based sentiment analysis methods, the methodology employed in SentiArt does not rely on data that has been manually marked up according to its valence. Instead, one starts with a list of labels, e.g., 'good' and 'bad' for positive and negative sentiment, respectively, and gets the embeddings for the labels using a vector space model. Words in a text are also vectorized via the same model. The association degree between a word and every label is then computed. When it comes to interpreting the values obtained, as authors explain, "if a given test word is — on average — more similar to a set of positive labels like GOOD than to the opposite set, it will be classified as having a positive valence and vice versa" [ibid.]. It should be noted that SentiArt proved to achieve adequate results in predicting the emotional potential of literary texts and outperformed other machine learning based classifiers in sentiment analysis [ibid.].

We followed the logic proposed by SentiArt developers with a few modifications. We have applied a vector space model for measuring texts' emotionality in the same fashion the experts did, not getting to classifying data as positive or negative. First, the labels were chosen as the words clearly representing six basic emotions. One word represented one emotion: "*ispugats'a*" ("to take fright"), "*zloj*" ("angry"), "*otvratitelnyj*" ("disgusting"), "*schastlivyj*" ("happy"), "*grustny*" ("sad"), and "*udivit'sa*" ("to be surprised"). The labels were chosen intuitively as having strong correspondence to the given emotions in their literal sense, being frequent and as stylistically neutral as emotions could be.

For the nexts stage, we used Word2vec Continuous Skipgram model trained on full Russian National Corpus. The model under consideration contains 185K words and is implemented in the gensim library [34]. During preprocessing, the texts were lemmatized using stanza library [35]. Each word in a given story received six scores based on how close it was to label words in a vector space (from 0 to 1). Semantic relatedness was understood as cosine similarity between the embeddings obtained for the labels and the embeddings for the words in the texts. The final scores for 210 short stories were then calculated as mean values for each emotion.

The emotional scores of short stories were calculated in two different ways. The first way was to consider all the words. The second way included a threshold, i.e. the mean was computed only using cosine similarity values that are higher than 0,5. The hypothesis was that leaving only words referring to emotions more strongly would increase the differentiation in final scores.

V. Results of Automated Sentiment Analysis of Literary Texts

A. Analyzing Sentiment with Emotive Lexicons

Calculations performed in Orange according to the algorithm [31] can generate sentiment values that fall within the range of -100 to 100. However, in our case, the final values turned out to be very small in absolute value and mostly negative, which indicates the predominance of words with negative sentiment. Table II shows the main statistics for the distribution of sentiment scores for the two used dictionaries — Multilingual Sentiment Lexicons (MSL) and RusSentiLex (RSL).

TABLE II. Comparison of Sentiment Scores for 2 Sentiment Dictionaries

Statistics	MSL	RSL
Min (negative)	-5,848	-7,992
Max (positive)	3,992	4,403
Range	9,840	12,395
Mean	-0,416	-1,453
SD	1,3556	1,718
Median	-0,574	-1,778

The information presented in Table II suggests that using the RusSentiLex dictionary results in a wider range of sentiment values. In the context of sentiment rating, this broader spread of values can be regarded as an advantage. Moreover, the average sentiment values are shifted towards more negative ratings, which can be explained by the larger size of the negative dictionary itself [36]. However, the correlation coefficient between the sentiment values obtained on lemmatized texts processed by these two dictionaries turned out to be quite high, namely 0.818. Spearman's rank correlation coefficient, obtained to compare the ranks of stories by sentiment, turned out to be slightly less than 0.740.

The most negative short story, according to both lexicons, was the story of Artyom Vesyoly “Pod krasnymi znamenami” (“Under the Red Flags”) (1923), which tells about the civil war. The most positive stories are different, however, depending on the lexicon, but the story by Alexander Belenson-Lugin “Egipetskaya predskazatel'nica” (“The Egyptian Fortune Teller”) (1921) and the well-known story “Legkoe dykhanie” (“Light Breathing”) by Ivan Bunin (1916), that can hardly be called positive by content, received the highest values.

B. Results of Sentiment Analysis using Dostoevsky

Breaking down each story into words and sentences and calculating the average sentiment score resulted in a significant number of neutral values, which did not provide us with useful insights. Therefore, we have chosen to exclude these results from analysis.

It was decided to focus on the data obtained for the entire text of the stories. Table III presents statistics on positive and negative sentiment scores, each measured on a scale from 0 to 1. These data are summarized in the parameter Pos/Neg, which represents the ratio of positive word sentiment to negative word sentiment.

TABLE III. Statistics for Positive and Negative Sentiment Scores

Statistics	Positive	Negative	Pos/Neg
Minimum	0,052	0,207	0,155
Maximum	0,228	0,446	0,909
Range	0,176	0,239	0,754
Mean	0,125	0,305	0,416
SD	0,034	0,043	0,126
Median	0,119	0,301	0,398

Table III demonstrates that, similar to the results obtained using the vocabulary method, the sentiment values produced have very small absolute values, with a predominance of negative lexicon.

Sergey Semenov’s “Sumerki” (“Twilight”) (1909) appeared to be the most negative text, while Artyom Vesyoly’s leading story in the vocabulary approach ranked only 25. The most positive texts according to the algorithm are Ivan Kataev’s “Avtobus” (“Bus”) (1929) and Ivan Bunin’s “Legkoe dyhanie” (“Light Breathing”) (1916), while “Egipetskaya predskazatel'nica” (“The Egyptian Fortune Teller”) was pushed to a distant 41st place. The sentiment scores obtained for the "positive" and "negative" scales were compared with the expert values assigned to the texts by respondents.

C. Results of Sentiment Analysis using SentiArt

In contrast to the previous approaches, the SentiArt method does not operate within a positive-negative dichotomy. Instead, it measures the degree of expression of six different emotions: happiness, surprise, sadness, disgust, anger, and fear. As mentioned above, the calculation was carried out using two methods: 1) cosine similarity between the words and the labels, obtaining the mean value for each of the labels, 2) the mean was computed with threshold — using cosine similarity values that are higher than 0.5.

Table IV shows the statistics for all 6 major emotions calculated using both methods. The statistics obtained by the first algorithm give surprisingly similar values for all six emotions. It can be assumed that such a picture is explained by the fact that, from the point of view of distributive semantics, the words denoting the studied emotions behave in a similar way compared to all other words in the vocabulary. The range of values is more widely spread with the counting option that uses a threshold of 0.5.

The happiest short story, according to this distributional semantics approach, was the story by Sergei Auslender “Zanyatye lyudi” (“Busy people”) (1912), the saddest and at the same time the most surprising was the story by Yevgeny Zamyatin “Iks” (“X”) (1919), the greatest disgust was manifested in the story by Lev Gumilevsky "Obnazhennye dushi" ("Naked Souls") (1915), anger is shown to the maximum extent in Alexander Lazarev-Gruzinsky's story "Forget-Me-Nots" ("Nezabudki") (1913), and fear — in the text by Mikhail Sandomirsky "Verochka" ("Verochka") (1915).

TABLE IV. STATISTICS FOR EMOTION SCORES FROM SENTIART

Statistics	Happy	Surprise	Sad	Disgust	Anger	Fear
Without threshold						
Minimum	0,205	0,194	0,218	0,179	0,197	0,224
Maximum	0,248	0,232	0,252	0,207	0,225	0,264
Range	0,043	0,038	0,034	0,028	0,028	0,040
Mean	0,224	0,212	0,233	0,190	0,210	0,243
SD	0,008	0,007	0,007	0,004	0,006	0,007
Median	0,224	0,212	0,233	0,189	0,210	0,243
With 0,5 threshold						
Minimum	0,503	0,501	0,503	0,518	0,501	0,506
Maximum	1,000	1,000	0,829	0,776	1,000	0,896
Range	0,497	0,499	0,326	0,257	0,499	0,390
Mean	0,587	0,771	0,571	0,570	0,613	0,582
SD	0,098	0,218	0,052	0,035	0,110	0,050
Median	0,540	0,750	0,557	0,561	0,587	0,576

The results obtained by the second method with 0.5 threshold give a different distribution of data, while the final correlation between the emotionality values is close to zero. The maximum correlation coefficient was found for surprise (0.17) and happiness (0.15), but these indicators are very small, the rank correlation coefficient does not exceed 0.145 in absolute value.

VI. COMPARING AUTOMATED SENTIMENT ANALYSIS AND READER'S EMOTIONAL FEEDBACK

Table V displays the correlation coefficients for the results of each of the automatic sentiment analysis experiments with the values of readers' emotional responses, which were obtained as a result of the experiment described in Section III of the paper. The correlation coefficients that were statistically significant at the level of $p < 0.05$ are indicated in bold font, while the cells in which the correlation coefficient was significant at the level of $p < 0.001$ are highlighted in gray. The data for SentiArt were calculated separately for each of the six emotions (happiness, surprise, sadness, disgust, anger, and fear).

TABLE V. CORRELATIONS BETWEEN TEXT SENTIMENT SCORES AND READER'S EMOTIONAL RESPONSES

Method	Reader's Emotional Responses					
	Happy	Surprise	Sad	Disgust	Anger	Fear
Dictionary approach						
RusSentiLex	**0,247**	0,102	**-0,255**	**-0,205**	**-0,238**	**-0,289**
MultiSentiLex	**0,241**	0,033	**-0,222**	**-0,149**	**-0,237**	**-0,355**
Dostoevsky						
Positive	-0,012	0,048	**0,229**	-0,038	0,032	0,008
Negative	**-0,281**	-0,075	**0,296**	**0,149**	**0,188**	**0,279**
Pos/Neg	**0,139**	0,084	0,054	-0,099	-0,055	-0,127
SentiArt						
SentiArt	0,089	-0,063	**0,209**	0,114	-0,004	-0,025
SentiArt (0,5)	0,037	0,022	-0,104	**-0,148**	-0,040	0,006

All of the correlations obtained are relatively weak. The strongest correlations are observed for the emotions of fear, sadness, and happiness, and the sentiment scores obtained using the dictionary approach. The sentiment analysis method based on the Dostoevsky library, although trained on social media texts, produced similar results to the dictionary-based approach. It also showed that fear, sadness, and happiness are the emotions that can be identified most effectively. However, the SentiArt method did not achieve the objectives of this research. The correlations obtained using the SentiArt method are both small in absolute value and not statistically significant. This outcome is particularly disappointing as it undermines the relevance of the method for our research purposes.

The findings suggest that the link between the sentiment of a literary text and the emotional response of reader is rather weak. In other words, the presence of negative vocabulary does not necessarily provoke negative emotions in readers, and the presence of positive vocabulary does not guarantee a positive emotional response. While we cannot entirely discount the possibility of a relationship between these factors, it appears that other factors may play a more significant role in shaping the emotional response of readers to literature. For instance, in the case of the dictionary approach results, positive emotions such as happiness and surprise have a positive correlation coefficient, while negative emotions like sadness, disgust, anger, and fear are negatively correlated.

The highest values of fear, sadness and happiness can be explained by the fact that these are the most frequent and easily recognizable emotions, which were more commonly observed in the readers' responses [27].

The results obtained with the SentiArt method appeared to be the least convincing. Perhaps, this is an indication that the method proposed by [26] for English literary texts requires more significant adaptation for Russian.

VII. CONCLUSION

The article presents the findings of a sentiment analysis conducted on a sample of short Russian prose texts that were written approximately 100 years ago. The sentiment analysis was performed using three different methods, including both dictionary-based and machine learning-based approaches. The results of the sentiment analysis were also compared with those of a reader evaluation experiment in which the same stories were rated based on the emotions they elicited.

The study's findings suggest that the presence of "positive" or "negative" vocabulary in a text has only a weak association with the reader's overall emotional response. However, it is also important to note that the correlation coefficients obtained are statistically significant and align with logical expectations. Therefore, while sentiment analysis is a useful tool, it is not sufficient on its own to create effective book recommendation systems that consider the emotional impact on readers. Other factors, such as plot, style, and narrative dynamism, should also be considered.

The relatively low correlation rates observed between the analyzed phenomena could also be attributed to the limitations of the dictionaries and training samples used in the study. Both are based on contemporary linguistic material and may not be suitable for analyzing the vocabulary of literary texts written a century ago. Therefore, it is essential to use appropriate dictionaries and text datasets from the corresponding time period when studying literary texts.

The comparison of results obtained by different methods is challenging due to the variation in approaches used to analyze the material, particularly when transitioning from specific emotions to the "negative-positive" scale and vice versa. To overcome this challenge, a special technique proposed by [26] may require further adaptation for use in analyzing Russian literary texts.

The preliminary assessment of the results obtained from different methods suggests that a dictionary approach using the RusSentiLex dictionary is preferable as it shows the maximum correlation value for all six emotions. However, for processing literary texts, it is recommended to expand the dictionary with bookish, poetic, and "obsolete" vocabulary.

ACKNOWLEDGMENT

This article is an output of a research project "Text as Big Data: Modeling Convergent Processes in Language and Speech using Digital Methods" implemented as part of the Basic Research Program at the National Research University Higher School of Economics (HSE University).

REFERENCES

[1] E. I. Bol'shakova, K. V. Voroncov, N. E. Efremova, E. S. Klyshinskij, N. V. Lukashevich, A. S. Sapin, *Avtomaticheskaya obrabotka tekstov na estestvennom yazyke i analiz dannyh*. HSE University Press, 2017.

[2] S. H. Lye and P. L. Teh, "Customer Intent Prediction using Sentiment Analysis Techniques", *in 2021 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Cracow, Poland*, Sep. 2021, pp. 185–190. doi: 10.1109/IDAACS53288.2021.9660391.

[3] B. Pang and L. Lee, "Opinion mining and sentiment analysis", *Foundations and Trends in information retrieval 2.1*, vol. 2, 2008, pp. 1–135.

[4] D. Vilares, M. Thelwall, M. A. Alonso, "The megaphone of the people? Spanish SentiStrength for real-time analysis of political tweets", *Journal of Information Science*, vol. 41, no. 6, Dec. 2015, pp. 799–813, doi: 10.1177/0165551515598926.

[5] B. Liu, *Sentiment analysis: mining opinions, sentiments, and emotions*. New York, NY: Cambridge University Press, 2015.

[6] P. Nandwani and R. Verma, "A review on sentiment analysis and emotion detection from text", *Social Network Analysis and Mining*, vol. 11, no. 1, p. 81, Aug. 2021, doi: 10.1007/s13278-021-00776-6.

[7] A. G. Pazelskaya and A. N. Solovyov, "Method for determining emotions in texts in Russian", *Computational linguistics and intelligent technologies: Based on the materials of the annual International Conference "Dialogue"*, no. 10, 2011.

[8] C. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text", *in Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8, no. 1, May 2014.

[9] S. Smetanin, "The Applications of Sentiment Analysis for Russian Language Texts: Current Challenges and Future Perspectives", *IEEE Access*, vol. 8, 2020, pp. 110693–110719.

[10] E. Kim, R. Klinger, "A Survey on Sentiment and Emotion Analysis for Computational Literary Studies", *Zeitschrift für digitale Geisteswissenschaften*, 2019. https://arxiv.org/ftp/arxiv/papers/1808/1808.03137.pdf

[11] N. Eric and S. Henry, "Character-to-Character Sentiment Analysis in Shakespeare's Plays", *in Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, 2013, pp. 479–483. Association for Computational Linguistics, https://aclanthology.org/P13-2085.pdf

[12] E. Kim et al., "Investigating the Relationship between Literary Genres and Emotional Plot Development", *in Proceedings of the ACL Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, 2017, pp. 17–26. Association for Computational Linguistics, https://aclanthology.org/W17-2203.pdf

[13] S. Min and J. Park, "Modelling Narrative Structure and Dynamics with Networks, Sentiment Analysis, and Topic Modeling." *PLoS ONE, edited by Thilo Gross*, vol.14, Dec. 2019.

[14] M. Saif, "From Once Upon a Time to Happily Ever After: Tracking Emotions in Novels and Fairy Tales", *in Proceedings of the ACL Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, 2011, pp. 105–114. Association for Computational Linguistics, https://dl.acm.org/doi/pdf/10.5555/2107636.2107650

[15] T. Sherstinova and G. Martynenko, "Linguistic and Stylistic Parameters for the Study of Literary Language in the Corpus of Russian Short Stories of the First Third of the 20th Century", *in R. Piotrowski's Readings in Language Engineering and Applied Linguistics, Proc. of the III International Conference on Language Engineering and Applied Linguistics (PRLEAL-2019), Saint Petersburg, Russia, November 27, 2019, CEUR Workshop Proceedings*, vol. 2552, 2020, pp. 105–120.

[16] G. Ya. Martynenko, T. Yu. Sherstinova, T. I. Popova, A. G. Melnik, E. V. Zamirajlova, "On the Principles of Creation of the Russian Short Stories Corpus of the First Third of the XX Century [O printsipakh sozdaniya korpusa russkogo rasskaza pervoy treti XX veka]", *in Proc. of the XV Int. Conf. on Computer and Cognitive Linguistics 'TEL 2018'*, 2018, pp. 180–197.

[17] G. Ya. Martynenko et al., "Methodological problems of creating a Computer anthology of the Russian story as a language resource for the study of the language and style of Russian fiction in the era of revolutionary changes (the first third of the 20th century)", *Computational Linguistics and Computational Ontologies*, no. 2, 2018, pp. 97–102.

[18] T. Sherstinova, E. Ushakova, A. Melnik, "Measures of syntactic complexity and their change over time (the case of Russian)", *in 2020 27th Conference of Open Innovations Association (FRUCT)*, 2020, pp. 221–229.

[19] A. Grebennikov, N. Marusenko, "The Early XX-century Russian Short Stories Corpora. An Example of Lingvo-statistical analysis", *IMS (CLCO)*, 2020, pp. 21–28.

[20] T. G. Skrebtsova, "Dynamics of Russian short stories themes at the beginning of the 20th century", *Philosophy and the humanities in the information society*, no. 3, 2020, pp. 45–60.

[21] E. Kazartsev, A. Davydova, T. Sherstinova, "Rhythmic Structures of Russian Prose and Occasional Iambs (a Diachronic Case Study)", *in Speech and Computer: 22nd International Conference, SPECOM 2020*, Oct. 2020, pp. 194–203.

[22] T. Yu. Sherstinova, "The World of the Russian Story through the Prism of Modern Digital Technologies", Socio-and psycholinguistic research, vol. 10, 2022, pp. 22–28.

[23] A. Lavrentiev et al., "Using TXM Platform for Research on Language Changes over Time: the Dynamics of Vocabulary and Punctuation in Russian Literary Texts", *Tomsk State University Journal of Philology,* vol. 70, 2021, pp. 69–89.

[24] Corpus of Russian short stories 1900-1930, Web: https://russian-short-stories.ru/.

[25] P. Ekman, "Facial expressions", *in Dalgleish, T., Power, M. (eds.) Handbook of Cognition and Emotion*. Chichester: Wiley, 1999, pp. 301-320.

[26] A. M. Jacobs, "Sentiment Analysis for Words and Fiction Characters from the Perspective of Computational (Neuro-)Poetics", *Front. Robot. AI*, vol. 6, Jul. 2019, doi: 10.3389/frobt.2019.00053.

[27] T. Yu. Sherstinova, E. O. Kolpashchikova, A. R. Seinova, P. I. Maksimenko, R. A. Rodionov, "Russian short story from 1900-1930 and its perception by the reader: Experience in Quantitative Analysis of Literary Text Evaluation" (in print).

[28] Orange, Sentiment Analysis, Web: https://orangedatamining.com.

[29] Multilingual Sentiment Lexicons, Web: https://sites.google.com/site/datascienceslab/projects/multilingualsentiment.

[30] N. V. Lukashevich, A. V. Levchik, Dictionary of evaluative words of the Russian language, Web: labinform.ru/pub/rusentilex/rusentilex_2017.txt.

[31] M. Hu, B. Liu, "Mining opinion features in customer reviews", *Proceedings of AAAI Conference on Artificial Intelligence*, vol. 4, 2004, pp. 755–760.

[32] SpaCy models for Russian, Web: https://spacy.io/models/ru.

[33] A. Rogers et al., "RuSentiment: An Enriched Sentiment Analysis Dataset for Social Media in Russian", *Proc. 27th International Conference on Computational Linguistics*, Aug. 2018, pp. 755–763.

[34] R. Rehurek and P. Sojka, "Gensim — Python Framework for Vector Space Modelling", *NLP Centre, Faculty of Informatics, Masaryk University, Brno, Czech Republic*, vol. 3, 2011.

[35] P. Qi, Y. Zhang, Y. Zhang, J. Bolton and C. D. Manning, "Stanza: A Python Natural Language Processing Toolkit for Many Human Languages", *Association for Computational Linguistics (ACL) System Demonstrations*, 2020.

[36] T. Yu. Sherstinova et al. Sentiment of Literary Texts in the Context of Theme and Reader Preferences (based on Russian short stories from 1900-1930s) (under review).