

Valuation of Emotionality and Osgood Factors in Fiction Prose Using Machine-Generated Dictionaries (on the Material of the Strugatsky Brothers Trilogy About M. Kammerer)

Rinat Albertovich Bakirov¹, Anton Sergeevich Afanasyev², Vladimir Vladimirovich Bochkarev³

Abstract

The relevance of the work lies in the fact that Osgood factors and emotive characteristics of the text can be extrapolated from the Russian fiction prose, as the most adequate type of author's and reader's expression, to the assessment of moods in society in a specifically given period. To analyse them, we need deep-learning neural networks capable of assessing positivity and other factors taking into account the context - with the help of dictionaries based on the survey of informants and machine extrapolation. From a methodological point of view, we used a dictionary with positivity ratings for Russian language words CartaSlovSent, as well as data from multilingual NRC VAD, Hedonometer dictionaries to identify categories of valence, dominance and arousal in texts. To construct a series of average text positivity, all word forms contained in the selected dictionary were found and the values of their positivity ratings were averaged. The results of the study showed that in the Strugatsky brothers' trilogy about Maxim Kammerer, "The Inhabited Island" with the corresponding descriptive negative narrative related to the image of the planet Saraksh is the most indicative for positivity ratings. At the same time, throughout the trilogy, Osgood factors are spread throughout the text according to the development of plot and compositional elements. "The Beetle in the Anthill" and "Waves Extinguish the Wind" are similar in the degree of expression of Osgood factors, including due to a similar dominant chronotope.

Keywords: *Emotionality, Osgood Factors, Fiction Prose, Machine-Generated Dictionaries, Strugatsky Brothers.*

Introduction

The creative work of A. and B. Strugatsky's creative work has been studied by Russian, European and American literary critics of different scientific schools since their lifetime and continues now (Amusin, 1996; Bardasova, 1995; Breeva, 2015; Izotov, 2002; Kuznetsova, 2004; Nadezhkina, 2008; Strugatsky, 2003; Kajtoch, 1993; Potts, 1991; Salvestroni, 1984; Suvin, 1974; Suvin, 1980; Howell, 1990 et al.). Meanwhile, the emotional aspect of their work, connected with the most important, in our opinion, categories identified by Ch. Osgood and characterising the emotive space of texts, including artistic texts, is practically not touched upon in these studies. These are valence, dominance, arousal.

The importance of these concepts in the analysis of literary, historical and cultural phenomena is due to the fact that an individual learns and appropriates them largely in the process of social learning (Maslinsky, 2023). In addition, the analysis of the works of the leading fiction authors of the Russian literary field of the second half of the 20th century allows us to reconstruct the key concepts of the epoch related to the understanding of alternative history and the future in the mass consciousness. For this purpose, it is necessary to search for general regularities in artistic texts, primarily related to the expression of emotions and Osgoodian factors in them, which will allow us to rise in literary studies from the level of private interpretations to large generalisations.

A vivid episode in Russian cultural history, illustrating the dependence of the emotional-psychological on the political and social, is the actualisation of the utopian and anti-utopian in the Brezhnev era (Breeva, 2015; Nadezhkina, 2008). In turn, important in our case is the question of periodisation of the work of the writers themselves. The worldview of the Strugatskys evolved rapidly. Their works of the 1960s, 1970s

¹ Candidate of Philology, researcher at the Laboratory of Multidisciplinary text studies of KFU, researcher at the Laboratory for Digital Studies of Literature and Folklore of the Pushkin House of the Russian Academy of Sciences. r1nt@ya.ru.

² Doctor of Philology, Professor, leading researcher at the Laboratory of Multidisciplinary text studies of KFU, Head of the Department of Russian Literature and Methods of teaching it at KFU.

³ Engineer at the Laboratory of Multidisciplinary text studies of KFU, a researcher at the Laboratory of Linguistics and Artificial Intelligence of KFU.

and 1980s in many respects differ from each other much more than the works of other authors, and the most intensive evolution of the writers' views was in the 1960s.

It seems to us an important observation that in Strugatsky's work the 1960s are characterised by a movement from acceptance of the attitudes proposed by society to doubt in them and a sharp critical attitude, which reached its peak by 1967 - the time of the creation of the novel "The Inhabited Island", for the 1970s, which resulted in the creation of the novel "The Beetle in the Anthill" (1979), is more characterised by reflection on their cause, and the novel "Waves Extinguish the Wind" (1984) sums up these reflections (Nadezhkina, 2008).

Research Method

We used a dictionary with positivity ratings of Russian language words MapSlovSent (Kulagin, 2021), as well as data from multilingual NRC VAD Hedonometer dictionaries (Mohammad, 2018; Dodds, 2015). An overview of available tone dictionaries for Russian can be found in (Kotelnikov et al., 2020). In (Bochkarev et al., 2023; Bochkarev et al., 2024), 6 general vocabulary dictionaries are compared, and it is concluded that they can be divided into 2 groups based on the correlation of word positivity scores, as well as in terms of their application to estimating the average tone of a text. From each group, we selected one dictionary of the largest volume, MapWordsCent and Hedonometer, containing positivity ratings for 46 thousand and 9.9 thousand words, respectively. We also used for the experiments the well-known multilingual NRC VAD dictionary containing data for 20 thousand words of the Russian language.

In (Bochkarev et al., 2024), human ratings from 6 Russian tone dictionaries were used to train neural network models of word positivity. These models use pre-trained fasttext vectors as input (Grave et al., 2018). The dataset presented in 2018 includes vectors for 2 million Russian words, which allowed machine learning of positivity ratings for the widest possible range of words. Table 1 shows the percentage of unique words from the considered works for which positivity ratings are available in the dictionaries.

Table 1. Unique Words from the Considered Works for Which Positivity Ratings Are Available in the Dictionaries.

work	MapWords	Hedonometer	NRC VAD	MICAI
The Inhabited Island	16.0	19.8	11.8	95.5
The Beetle in the Anthill	16.5	24.8	13.0	95.2
Waves Extinguish the Wind	16.1	26.5	13.3	93.0

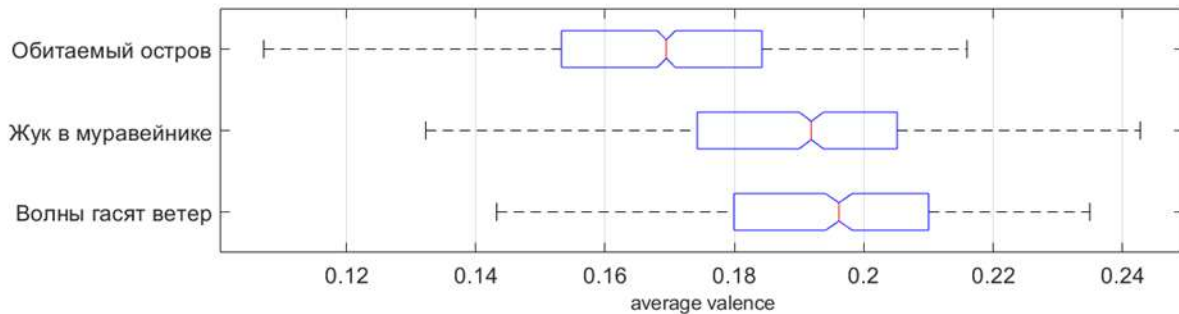
Result

The average positivity series of a text is constructed as follows. In a fixed-width sliding window, we find all word forms contained in the selected dictionary and average their positivity ratings. Experiments were conducted with windows of lengths 250, 500, 1000, 2000 and 4000 words. With small window lengths we obtain a highly noisy graph; with excessively long window lengths, information about changes in average positivity may be lost. As a compromise between the requirement of sufficient resolution and the degree of noise in the obtained curves, we chose a window length of 2000 for the analysis. In this case, we obtain

reasonably smooth curves, and the observed changes in mean positivity are consistent with the division of the text into chapters.

In general, the average positivity values of the texts under consideration are shown in Diagram 1. In this case, we have taken three curves from the graphs of individual works, and made “mustache boxes” on them. Thus, we can see within what limits the average (2000 words) positivity of the text in each of the three works typically fluctuates.

Diagram 1. Average Positivity of the Texts of the Kammerer Trilogy

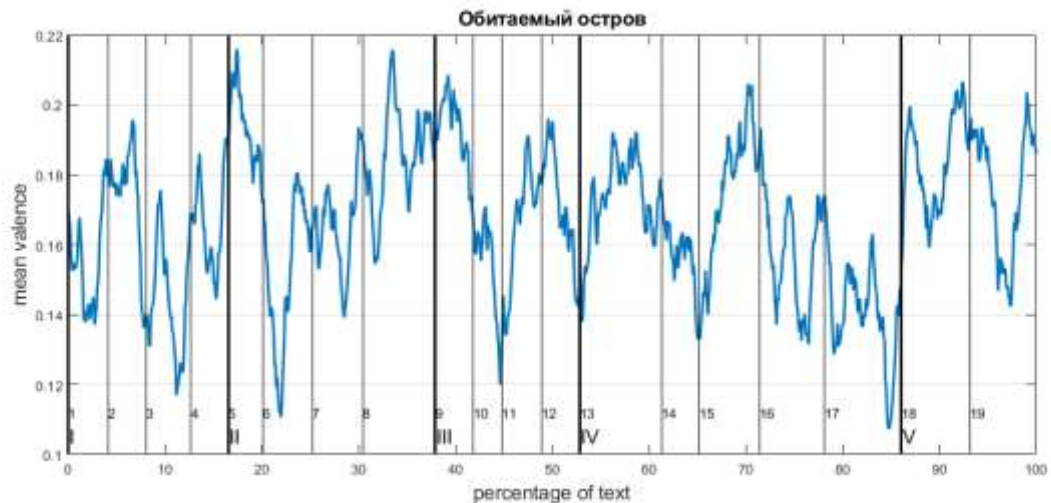


If we talk about each individual text and the positivity indicators for them, the curve breaks and changes in the trend direction in graphs 2, 3 and 4 very often occur at the boundaries between chapters, which seems natural and further confirms the applicability of the positivity analysis methodology to this type of texts.

In particular, the narrative and compositional elements of “The Inhabited Island” correlate with the positivity indicators in this way: the greatest growth of this category is observed at the beginning of the second part of the novel and the fifth chapter in it. This is the beginning of the narrative, the plot, in which the action is just beginning, and the protagonist is not yet aware of all the specifics of the world in which he has fallen. Accordingly, the conflict has not yet entered the open stage and the plot, on the whole, is relatively positive. As an example of such a positive text from Part II, Chapter Five (hereinafter translated by the authors of the article):

“Guy clicked his heels and walked out. Outside the door he allowed himself a smile. Still, the old warrior couldn't help himself and took responsibility. Good things are always good. Now one could in good conscience consider Maxim his friend. Maka Sima. I can't pronounce his real name. Either he invented it while he was in a delirium, or he really came from those mountaineers... what was the name of their ancient king... Zaremchakbeshmusarai... Gai went out to the training ground and looked for his section with his eyes. The indefatigable Pandey was chasing the boys through the upper window of a mock-up of a three-story building. The guys were sweating, and it was bad, because there was only an hour left before the operation” (Strugatsky 2024).

Diagram 2. Positivity in the novel “The Inhabited Island”



An example of the least positive segment in the text of this novel can be the end of Chapter 17 in Part IV, which describes one of the most poignant moments of the protagonist's struggle along with the corresponding setting characterized by images of death and war:

“The tank sped through a thick jet of black smoke: someone was burning on the left. We jumped through, and had to make a sharp turn at once, so as not to run over a dead man flattened by the tracks. A shrunken boundary sign came out of the smoke and disappeared, followed by tattered, crumpled wire fences. A man in a strange white helmet leaned out of an inconspicuous ditch for a moment, shook his fists violently and immediately disappeared, as if he had vanished into the ground. The smoky veil ahead gradually cleared, and Maxim saw the brown round hills, quite close, and the mud-stained stern of a tank, crawling at an oblique angle to the general movement, and another burning tank. Maxim turned to the left, aiming his car at a deep, bushy saddle between two higher hills. He was already close when fire spurted toward him, and the whole tank rumbled with a terrible thud. From surprise Maxim gave full throttle, the bushes and the cloud of white smoke above them leaped toward him, white helmets flashed, faces contorted with hatred, fists raised, then something cracked iron under the tracks, breaking, Maxim clenched his teeth, took a sharp right and drove the car away from this place, down the slope, lurching heavily, almost turning over, rounding the hill, and finally drove into a narrow hollow, overgrown with young trees” (Strugatsky 2024).

The growth of general positivity in the next novel of the trilogy, “The Beetle in the Anthill”, is connected, as we think, with the movement of the action to the World of Noon. And if we look at the peak of positivity in the novel, it falls in its middle. Which is also symptomatic. If “The Inhabited Island” is, on the whole, more tense in its plot, “The Beetle in the Anthill” is somewhat slower and more philosophical in its narrative. Therefore, the same clear movement of increasing negativity from the beginning to the end cannot be distinguished in it - Diagram 3.

Diagram 3. Positivity in the Novel “The Beetle in the Anthill”

An illustrative example from the chapter “June 2, '78. Some guesses about Lev Abalkin's intentions.”

“How did he know my number? That's easy. I used to leave the number with my teacher. Besides, Maya Glumova could have told him about me. So either he had another conversation with Maja Toivovna, or he decided to see the Teacher after all. In spite of everything. For twenty years he hadn't let me know about him, and now he suddenly decided to see him. Why?”

Why did he call me? Sentimental reasons, for example. Memories of his first real job. Youth, the happiest time of my life. Hmm. Doubtful... Altruistic desire to help the journalist (the discoverer of the beloved Golovanovs) in his work, laced with, say, a healthy ambition. Nonsense. Why would he give me a false address in this case? Or maybe not a fake one? But if not fake, it means that he is not hiding, it means that Exelenz is confusing something... Indeed, how does it follow that Lev Abalkin is hiding?” (Strugatsky 2024).

We see in this passage a clear dominance of lexemes with increased positivity: “youth”, “happiness”, “good” and others. They overlap the relatively neutral lexemes and lexemes with reduced positivity found in the same part of the text.

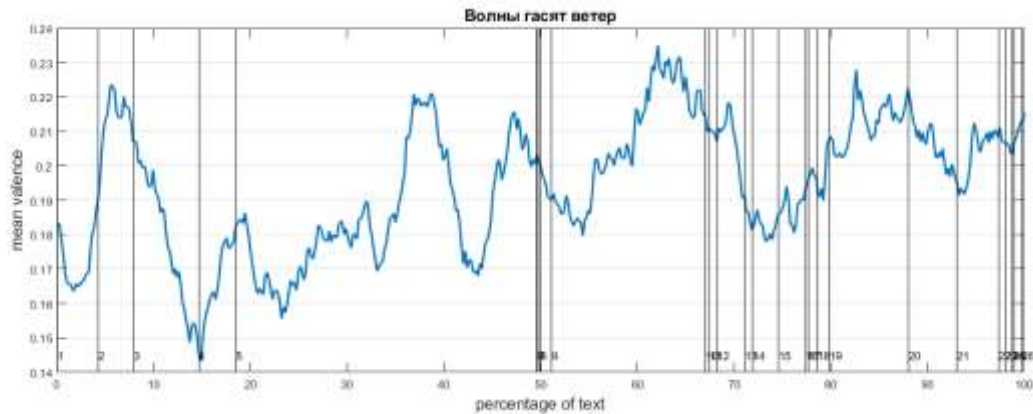
The greatest negativity of the text is presented in the short chapter “From Lev Abalkin's Report”. Firstly, here the World of Noon is shown through the eyes of someone who is probably able to evaluate it from the outside and see its imperfection. Second, Abalkin's image is deliberately tragicized in the text to emphasize the general pessimism and complexity of this character:

“Ahead on the right, a layer of plaster separates from the dirty-white balcony on the third floor and collapses noisily onto the sidewalk. Rats squeak indignantly. A mosquito pole bursts from a pile of trash and spins in the air. A huge snake streams across the street in a patterned metal ribbon, coils in front of Szczekn, and raises its rhombic head threateningly. Shchekn does not even stop - carelessly and briefly waves his front paw, the rhombic head flies off on the sidewalk, and he is already trudging on, leaving behind a wriggling ball of decapitated body” (Strugatsky 2024).

This passage is clearly dominated by lemmas with expressed lowered positivity: “noise”, “rat”, “mosquito”, “heap”, “garbage”, “snake”, “threat”, “decapitated”, and others. In general, the whole of this chapter is characterized by a kind of doubting view of everything that is happening, passing into a general pessimism.

The last novel of the trilogy, “Waves Extinguish the Wind”, is the most positive of all. Although its increased dystopia should have dictated the dominance of conditionally negative lexical units in it, the analysis showed that it is the most positive work of the whole trilogy (Diagram 4).

Diagram 4. Positivity in the Novel “Waves Extinguish the Wind”



The least relative and absolute positivity, according to the graph and numerical values, is characteristic of the beginning of the chapter “Document 3”:

“Date: March 20, '99. Author: T. Glumov, inspector. Subject 009: “Visit of an old lady”. Content: cosmophobia, “penguin syndrome”.

Analyzing cases of space phobia over the last hundred years, I came to the conclusion that within the framework of topic 009, materials on the so-called “penguin syndrome” may be of interest to us.

Sources:

A. Mebius, report at the XIV conference of cosmopsychologists, Riga, 84. A._Mebius. “Penguin syndrome,” PKP (“Problems of Space Psychology”), 42, 84.

A. Mebius. “Again on the nature of the 'penguin syndrome',” PKP, 44, 85.

Reference: Mebius Asmodeus-Matvey, Doctor of Medicine, Corresponding Member of the Academy of Medical Sciences of Europe, Director of the branch of the World Institute of Space Psychopathology (Vienna). Born. 26.04.36, Innsbruck. Education: Faculty of Psychopathology, Sorbonne; Second Institute of Space Medicine, Moscow; Higher Courses of Non-Aborne Aquanautics, Honolulu. Main areas of scientific interests: extra-industrial space and aquaphobia. From 81-91 - Deputy Chairman of the Main Medical Commission of the Space Fleet Administration. Nowadays he is a recognized founder and head of the school of so-called “polymorphic cosmopsychopathology” (Strugatsky 2024).

As we can see, the least positive in the novel and within the chapter itself was the part connected with the greatest formalization of language. On the other hand, thematically it is devoted to the description of negative medical indications - first of all, phobias. In turn, the greatest increase in positivity in the novel is recorded in the chapter “Document 8” and coincides with one of the gastronomic descriptions and the conventionally calm “evening” chronotope:

“TOIVO GLUMOV AT HOME. MAY 8, '99. LATE EVENING.

They dined in a room scarlet with sunset.

Asya was in upset feelings. Paszkowski's sourdough, delivered to the delicatessen factory directly from Pandora (in living bags of biocontainers covered with terracotta frost, stained with horny hooks of vaporizers, six kilograms of precious sourdough in each bag), this sourdough had rebelled again. Its flavor spontaneously went into the “sigma” class, and its bitterness reached the last allowable degree. The board of experts split. The master demanded that from now on, until clarification, to stop the production of famous for the entire planet “alapaichiki”, and Bruno - a bold talker, a boy, a cheeky - said: why should it? Never in his life had he dared to speak against the Master, but today he suddenly began to orate. Ordinary amateurs would not notice such a change in taste, but as for connoisseurs, he gave his head to the wind - at least one in five would be delighted by such a flavor variation... Who needs his severed head? But he was supported! And now it's unclear what will happen...” (Strugatsky 2024).

Thus, in all three novels of the Strugatsky brothers' trilogy about Maxim Kammerer, we record the growth of positivity from the first to the last novel. If we focus on the peculiarities of the narrative, correlated with the moments of the greatest rise and fall of positivity in these texts, we can record, on average, a decrease of this indicator in plot situations that are associated with the greatest dynamics - what can be conditionally called the genre signs of a fantastic action movie. In turn, “calm and quiet” philosophical or descriptive inclusions in novels are more characterized by increased positivity. It is perhaps this feature that is one of the keys to explaining the overall dynamic. The fast-paced action in “The Inhabited Island” “slows down” in the semi-detective “The Beetle in the Anthill” and conditionally fades in “Waves Extinguish the Wind”. Accordingly, the average positivity of these texts and the entire trilogy grows at the same time.

Discussion

In their trilogy, the Strugatsky create different models of the world, embodying the most typical features of the social atmosphere of the Soviet Union of the periods of the 1960s, 1970s and 1980s. The model of the world of each part of the trilogy is built according to the characteristic principle of binary oppositions and, according to (Nadezhkina, 2008), captures the most typical features of the epoch, and in the characteristic form of differentiation of one's own from another's, good from evil, ultimately harmonising the world or preventing it.

As researchers note, in general, the evolution of the Strugatsky brothers' artistic world follows the model of movement from lighter and more optimistic colours to pessimistic conclusions about the future and present society. However, the data of our quantitative analysis of the categories of positivity (valence), strength (dominance) and activity (arousal) in the texts showed that the average positivity of the texts is on the contrary increasing. A natural question arises - what can this be related to? As it seems to us, the answer lies in the peculiarities of the chronotope of this or that text. Chronotope, or the place and time of action, is an important part of the artistic world of a work and words related to the description of the chronotope constitute a significant lexical-semantic group. This correlates well with such elements of characterisation of Strugatsky's texts as utopianism and anti-utopianism (Kaspe, 2018, 241). If the world described in “The Inhabited Island”, in general, is more of a non-accepted world both by the narrator and the main characters, in the other two texts we see a greater immersion in the realities of the ‘ideal world’, which require an appropriate descriptive vocabulary.

The most telling is ‘The Inhabited Island’, in which the main action takes place on another planet, albeit in the realities of the same world. The descriptive narrative, the main task in this particular case is to depict the imperfections of the world of the planet Saraksh, is logically connected with the dominance of appropriate lexical units. At the same time, they are arranged in the text according to the development of plot and compositional elements. From the bright depiction of the World of Noon and Kammerer's equally positive expectations of the new world to the most condensed description in the middle of the novel of the main negative features of the planet Saraksh and the protagonist's further relatively optimistic struggle with circumstances and the system. Genre-wise, the novel is close to a sci-fi action film with a lot of action by the characters, which is evident in the increased number of verb forms in the text.

"The Beetle in the Anthill", detective in its basis, is much less connected with the theme of struggle and action and is more directed towards the resolution of socio-political problems. There is less 'struggle' in the text, it is slower, calmer and more philosophical, which is reflected in the overall increase in the number of positive lexemes.

"Waves Extinguish the Wind", on the other hand, is the most "contemplative" novel of the trilogy and the concluding text on the World of Noon. The external conflict in it is based on the opposition between 'good' and 'best', but not at all in the key of the Soviet production novel. The fantastic plot here is presented, on the one hand, in semi-memoir form, on the other hand, against the background of the authors' and characters' reflections on superhumans as a new stage in the evolution of the human species and some disillusionment with this idea.

Conclusion

There is a general movement towards greater positivity in those Strugatsky texts whose chronotope is connected with the World of Noon. In this case, we can distinguish a clearly expressed utopian beginning, which also wins at the lexical level when depicting the relevant realities of the world in which the characters live. However, it is worth separating the general positivity in the depiction of the surrounding world from the general mood of the texts. The large average positivity may be due to the fact that such lexemes dominate in describing the surrounding world, which is the contrast to which the dystopian content of the selected novels is based. If we look at the characteristics of the chronotope of these works, we will see that in the first of them chronologically everything takes place on another planet, the world of which is criticised to a greater extent, hence the overall positivity is obviously lower. In the second and third novels the action is transferred to the Earth with conditionally ideal living conditions, but in the second novel there is a specific criticism of the paradigm 'friend-alien', therefore the general positivity is obviously lower.

Acknowledgments

The research was supported by the Russian Science Foundation grant № 24-18-00570, <https://rscf.ru/en/project/24-18-00570/>.

References

- Amusin, M. (1996). *Brat'ia Strugatsky: Ocherk tvorchestva*. Ierusalim. Beseder.
- Bardasova, E.V. (1995). Kontseptsiiia «vozmozhnykh mirov» v svete esteticheskogo ideala pisatelei-fantastov A. i B. Strugatskikh. Kazan. Kazan state university.
- Bochkarev, V.V., Achkeev, A.A., Savinkov, A.V., Shevlyakova, A.V., Solovyev, V.D. (2024). Large Sentiment Dictionary of Russian Words. In: Calvo, H., Martínez-Villaseñor, L., Ponce, H. (eds) *Advances in Soft Computing. MICAI 2023. Lecture Notes in Computer Science*, vol 14392. 68–82. Springer, Cham. https://doi.org/10.1007/978-3-031-47640-2_6
- Bochkarev, V.V., Solovyev, V.D., Nestik, T. A., Shevlyakova, A.V. (2023). Variations in average word valence of Russian books over a century and social change. *Issledovaniia po prikladnoi matematike i informatike. II–1, Zap. nauchn. sem. POMI*, 529. 24–42.
- Breeva, T.N. (2015). Dekonstruksiia utopicheskogo diskursa v tsikle proizvedenii brat'ev Strugatskikh «Mir Poludnia». *Izvestiia Ural'skogo federal'nogo universiteta. Ser. 2, Gumanitarnye nauki. № 1 (136)*. 226–235.
- Dodds, P., et al. (2015). Human language reveals a universal positivity bias. *Proc. Natl. Acad. Sci.* 112(8), 2389–2394.
- Grave, E., Bojanowski, P., Gupta, P., Joulin, A., Mikolov, T. (2018). Learning word vectors for 157 languages. In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. European Language Resources Association (ELRA). Miyazaki.
- Howell, Y.H. (1990). *Apocalyptic Realism: The Science Fiction of Arkady and Boris Strugatsky*. The University of Michigan.
- Izotov, V.P. (2002). *Nekotorye filologicheskie podrobnosti fantasticheskogo tvorchestva brat'ev Strugatskikh*. Orel. Orion, 2002.
- Kajtoch, W. (1993). *Bracia Strugaccy W. Kajtoch*. Krakow. Universitas.
- Kaspe, I. (2018). *V soiuzе s utopiei. Smyslovyе rubezhi pozdnesovetskoj kul'tury*. Moskow. Novoe literaturnoe obozrenie.
- Kotel'nikova, E., Razova, E., Kotel'nikova, A., Vychezhzhanin, S. (2020). Sovremennye slovari ocenocnoj leksiki dlya analiza mnenij na russkom i anglijskom yazykah (analitich-eskij obzor). *Nauchno-tehnicheskaya informaciya. Seriya 2: Informacionnye processy i sistemy* 12, 16–33. <https://doi.org/10.36535/0548-0027-2020-12-3>.

- Kulagin, D. (2021). Publicly available sentiment dictionary for the Russian language KartaSlovSent. Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue". vol. 20. 1106–1119.
- Maslinsky K. (2022). Sto let schast'ia v detskoj literature (1920–2020): stalinskii kanon i ego dolgosrochnye posledstviia. Steps. T. 8. № 4. 2022. 226–247. DOI: 10.22394/2412-9410-2022-8-4-226-247.
- Mohammad, S. (2018). Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 English words. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Vol. 1: Long Papers). 174–184. <https://doi.org/10.18653/v1/P18-1017>
- Nadezhkina, T.O. (2008). Mifopoeticheskaia organizatsiia trilogii A. i B. Strugatskikh "Obitaemyi ostrov", "Zhuk v muraveinike", "Volny gasiat veter": diss. kand. filol. nauk : 10.01.01. Vladivostok, 2008.
- Potts, S.W. (1991). The Second Marxian Invasion: The Fiction of the Strugatsky Brothers. San Bernardino. California.
- Salvestroni, S. (1984). Ambiguous Miracle in Three Novels by the Strugatsky Brothers. Science Fiction Studies (Monreal). № 11. 291-303.
- Strugatsky, B. (2003). Kommentarii k proidennomu. Saint Petersburg. Amfora.
- Strugatsky, A. & B. (2024). Novels, short stories. <https://strugacki.ru>.
- Suvin, D.R. (1980). Introduction. Strugatsky A., Strugatsky B. The snail on the slope. New York. 1–20.
- Suvin, D.R. (1974). The Literary Opus of Strugatsky Brothers. Canadian-American Slavic Studies. Vol. 8, № 3. 454–463.