

DOI: <https://doi.org/10.18454/RULB.2020.24.4.13>**К ВОПРОСУ ОБ ИСПОЛЬЗОВАНИИ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ В ЛИНГВИСТИЧЕСКОМ АНАЛИЗЕ ХУДОЖЕСТВЕННОГО ТЕКСТА**

Научная статья

Южакова Ю.А.^{1*}, Иванова Д.С.², Лаврентьев В.А.³, Сомова М.В.⁴¹ ORCID: 0000-0002-5050-582X,^{1, 2, 3, 4} Рязанский государственный университет имени С.А. Есенина, Рязань, Россия

* Корреспондирующий автор (yu.yuzhakova[at]mail.ru)

Аннотация

Статья посвящена исследованию возможностей анализа идиостиля писателя с использованием информационных технологий. Описана попытка выявления, изучения и анализа лексических единиц со значением субъективной оценки, использованных И. С. Тургеневым при описании природы. В статье указаны информационные ресурсы (НКРЯ, Russian Morphological Dictionary, Mystem, ПСИ-Офис, контент-анализ текстов ВААЛ-2000, программа Textarc) и представлены их возможности при работе с художественным текстом писателя.

Ключевые слова: информационные технологии, лингвистический анализ, идиостиль, лексика, эмоционально-экспрессивная окраска.

ADDRESSING THE ISSUE OF EMPLOYING INFORMATION TECHNOLOGY IN CLOSE READING

Research article

Yuzhakova Yu.A.^{1*}, Ivanova D.S.², Lavrentev V.A.³, Somova M.V.⁴¹ ORCID: 0000-0002-5050-582X,^{1, 2, 3, 4} Ryazan state University named after S. A. Yesenin, Ryazan, Russia

* Corresponding author (yu.yuzhakova[at]mail.ru)

Abstract

The article investigates the issue of employing information technology in the analysis of the writer's individual style. It describes an attempt to investigate the Russian writer I. S. Turgenev's use of lexical units transmitting evaluative connotation in his descriptions of nature. The article enumerates information resources (RNC (Russian National Corpus), Russian Morphological Dictionary, Mystem, Psyoffice, VAAL-2000 Content Analysis, Textarc) and demonstrates their potential use in close reading.

Keywords: information technology, close reading, the writer's individual style, lexical units, affective and expressive connotations.

Introduction

Close reading has traditionally relied on a researcher's subjective perception of the writer's individual style. Modern linguistics, however, often resorts to information technology, which secures an objective analysis of the writer's choice of linguistic means, stylistic devices and figures of speech.

There is a number of close reading techniques that can be employed even by a technically incompetent philologist. Among such techniques are search techniques (search for data sources, search for texts and references to them), dictionaries, text generators. Modern philologists utilize information technology to analyze, transform, and linguistically process literary texts [10, P. 12]. They employ software for psycholinguistic and linguistic research to ensure a more comprehensive and efficient text analysis. A key electronic resource is the Russian National Corpus, which incorporates electronic versions of Russian texts [8].

The paper attempts to demonstrate the potential of modern information technology utilization in linguistics at the example of the Russian writer I. S. Turgenev's individual style. We intend to analyze I. S. Turgenev's novellas and short stories to find lexical units transmitting evaluative connotations in the writer's descriptions of nature. Processing I. S. Turgenev's works we have singled out the following groups of words actively used by author:

1. Words with emotional and intensifying connotations *наслаждаться, любоваться, славный, любимый*, etc., whose meanings imply subjective evaluations. The author uses emotionally-charged words to express positive or negative assessments: «Я, признаюсь, **не слишком люблю** это дерево – осину <...> **вообще я не люблю** этого дерева ...» [11, P. 42]. The words *восторг, радость, восхищение* express positive emotions. To perform a quantitative analysis of these lexical units, we resort to the Russian National Corpus (further referred to as the RNC) [3, P. 20]. The method of statistical sampling and data analysis shows that the word *наслаждаться* appears in I. S. Turgenev's prose 210 times, the word *славный* is used 114 times, the word *любимый* is registered 327 times. To facilitate statistical sampling, one can use the RNC together with such information resources as the Russian Morphological Dictionary, the Mystem, and the Psyoffice. The statistical analysis of emotional vocabulary which is used by the author to describe people's inner state, shows that I. S. Turgenev's descriptions of nature are deeply psychological.

2. Evaluative and gradational adverbs:

– evaluative words – *хорошо, отрадно, славно, томно, вольно, любо*, etc.: «**Свежо, весело, любо!** Вы в тени, вы дышите пахучей сыростью; **вам хорошо**, а против вас кусты раскаляются и словно желтеют на солнце. Кто, кроме охотника, испытал, как **отрадно** бродить на заре по кустам?» [11, P. 44];

– gradational words – *изредка, слабо, тускло, глубоко, сильно* и проч.: «**Влажный ветерок изредка** набегает легкой волной, деревья **слабо** шумят ...» [11, P. 175]. An evaluative analysis discloses the author's preferences, reveals the degree of

his emotional expressiveness and highlights I. S. Turgenev's skills of describing nature as a manifestation of universal harmony where everyone can find inner peace.

3. Evaluative adjectives, which are used by I. S. Turgenev to describe nature and to express his storyteller's attitude to the places he visits and to the scenes he witnesses: «*Люди они были небогатые; дом их, весьма старинный, деревянный, но удобный, стоял на горе, между заглохшим садом и заросшим двором*» [11, P. 132].

It is relevant to compare I. S. Turgenev's works and works of his contemporaries. A comparative analysis can be performed on the basis of the Lingualyzer software, a tool of quantitative linguistics [6]. This electronic resource enables a researcher to identify similar texts. Moreover, the Link Grammar Parser software can be used to process translated texts and compare texts in different languages, Russian and English, for example.

4. Relative pronouns *как*, *какой* and etc. used in exclamatory sentences for the sake of emphasis: *Какой вид! Река вьется верст на десять, тускло синяя сквозь туман; за лугами пологие холмы, <...>, ясно выступает даль... Как вольно дышит грудь, как бодро движутся члены, как крепнет весь человек, охваченный свежим дыханьем весны!*» [11, P. 175].

The RNS analysis of I. S. Turgenev's texts shows that the pronoun *какой* forms collocations with such words as *вид*, *восторг*, *закат*, *восход*, *тишина*. The pronoun *как* is often accompanied by evaluative adverbs. On the one hand, this is a unique feature of I. S. Turgenev's individual style, but on the other hand, it is a speech feature of the epoch.

5. Emotive vocabulary is used in descriptions of nature together with such means of expression as:

– body parts which are endowed with emotions or emotional reactions (*сердце*, *душа*, *голова*): «*...Сердце то вдруг задрожит и забьется, страстно бросится вперед, то безвозвратно потонет в воспоминаниях. С вет так и хлынет потоком; сердце в вас встрепенется, как птица. Голова томно кружится от избытка благоуханий. Как вольно дышит грудь, как бодро движутся члены...*» [11, P. 178];

– time of the day: «*Какое наслаждение выехать весной до зари! Солнце быстро поднимается, небо чисто. А летнее, июльское утро! Еще свежо, но уже чувствуется близость жары*» [11, P. 176]. The descriptions of the time of the day are full of lyricism and are a part of the author's experience, which makes them appealing to common people;

– shades of color: «*На темно-сером небе кое-где мигают звезды; Зеленой чертой ложится след ваших ног по росистой, побелевшей траве. <...> Вдали стоит дубовый лес и блестит и алеет на солнце; Кое-где разве вдали желтеет поспевающая рожь, узкими полосками краснеет гречиха...; Река вьется верст на десять, тускло синяя сквозь туман*» [11, P. 176];

– verbs of motion used in descriptions of feelings and inner states: «*Сердце в вас встрепенется, как птица; Голова томно кружится от избытка благоуханий. На душу находит странная тревога. Сердце то вдруг задрожит и забьется, страстно бросится вперед, то безвозвратно потонет в воспоминаниях. Вся жизнь развертывается легко и быстро, как свиток*» [11, P. 178].

We can see that in the majority of cases neutral lexemes acquire emotional and expressive connotations in the context. Contextual and stylistic connotations of a word are formally manifested through changed grammatical properties and altered collocability, which can be analyzed by means of information technologies. A researcher can resort to the automatic processing of a text, which incorporates grammar, morphological, and semantic analysis for the English, German and Russian languages, the linguistic search system, various thesauruses and glossaries and the MediaLingua software, a system of automatic annotation, classification, and text processing [3], [4].

6. Dialectisms, professionalisms and other words stylistically limited in their use have an expressive and affective potential and can characterize a writer's individual style. These words usually require an explanation of their origin, use, grammatical nuances, etc. For example, depicting a wide pond in the Orlov Region, I. S. Turgenev uses the professionalisms *чирок*, which is often used by hunters in relation to a game bird, and the professionalism *майер*, which denotes reed thickets: «*Если даже какой—нибудь отсталый и неопытный чирок и подвергался нашим выстрелам и лишился жизни, то достать его из сплосного майера наши собаки не были в состоянии*» [11, P. 32]. The analysis of words is performed by means of electronic dictionaries and computerized systems for searching, classifying and comparing information [1, P. 35].

7. Synonyms and antonyms also possess an expressive potential. Using synonyms, an author can avoid tautology and fully express denotative and connotative shades of meaning. Here is an example of I. S. Turgenev's masterful use of contextual synonyms to describe an old forest that has lost its former grandeur: «*Дубы поднимали кверху свои безжизненные, обломанные ветви; у других из листвы, еще довольно густой, хотя не обильной, не избыточной по—прежнему, торчали толстые, сухие, мертвые сучья...*» [11, P. 96]. Here is another example depicting a stiflingly hot summer day. Note the author's use of synonyms: «*Солнце разгоралось на небе, как бы свирепя; парило и пекло неотступно; воздух был весь пропитан душной пылью*» [11, P. 103].

A software-based analysis shows that chains of synonyms are characteristic of I. S. Turgenev's individual manner of writing. The writer uses them to depict a quality more intensively: «*Солнце так и било с синего, потемневшего неба...*», «*крупные капли дождя резко застучали, зашлепали по листьям, сверкнула молния, и гроза разразилась*», «*их [дубов] статные, могучие стволы великолепно чернели на золотисто-прозрачной зелени орешников и рябин*», «*прохваченная морозом земля потела и оттаивала на солнце*» [11, P. 20]. A software-based analysis can be used to investigate individual synonyms and groups of synonyms both in the works of a particular author and in the literary heritage of a certain epoch [5, P.102].

8. Antonyms can also be used to increase the expressive potential of descriptions. The juxtaposition of antonyms makes descriptions more expressive and affective: «*Душный жар внезапно сменился влажным холодом*» [11, P. 113].

Using antithesis, i.e. juxtaposing antonyms, I. S. Turgenev manages to compare the scenery of two different villages (a village which is situated in the Orlov Province, and a village located in the Kaluga Province), highlighting the prosperity of one village and the degradation of the other: «*Орловская деревня <...> кроме немногих ракут, всегда готовых к услугам, да двух—трех тощих берез, деревца на версту кругом не увидишь; изба лепится к избе, крыши закиданы гнилой*

соломой... Калужская деревня, напротив, большею частью окружена лесом; избы стоят вольней и прямой, крыты тесом...» [11, P. 50].

Employing antithesis in his descriptions of nature, the writer shows movement. The juxtaposition of the verbs *просыпался* – *утихал* and the verbs *разыгрывается* – *замер* enables the reader to feel the wind on a summer day, warm and soft, but occasionally impetuous: «*Легкий ветерок то просыпался, то утихал: подует вдруг прямо в лицо и как будто разыгрывается, – всё весело зашумит, закивает и задвигается кругом, грациозно закачаются гибкие концы папоротников, – обрадуешься ему... но вот уж он опять замер, и всё опять стихло»* [11, P. 53].

I. S. Turgenev employs antithesis to depict changing weather. The reader can clearly see thunderous clouds dispel to be followed by light clouds: «*в отдалении еще толпились тяжелые громады туч, <...> звездочки мерцали сквозь жидкие, быстро летеющие облака»* [11, P. 77].

The RNC and other software can be used to statistically assess the results of close reading of I.S. Turgenev's works [2, P.210]. For instance, I.S. Turgenev's individual choice of synonyms and antonyms can be compared against individual preferences of another writer belonging to the same literary epoch. The analysis can result in a hypothesis that expressive linguistic means are more extensively used by this or that writer. The collected data are processed by means of Microsoft Excell, SPSS, and Stadia and are presented in spreadsheets [6], [9].

9. Lexical repetitions also possess an affective potential. I.S. Turgenev resorts to lexical repetitions to draw parallels between his characters' feelings and nature, the inner and the outer worlds.

For instance: «*нежный отблеск белого шелка»*, «*березы стояли все белые, белые, как только что выпавший снег»* – при сопоставлении березы и осины; или при характеристике погоды: «*перепадал мелкий дождик, ... роца, влажная от дождя»*, «*мельчайший дождь»* [11, P. 43]. Linguistic software is very helpful when it is necessary to analyze lexical repetitions. The analysis of contexts shows that I. S. Turgenev characterizes different phenomena comparing and juxtaposing them.

The Textarc software is also very helpful [6]. It enables a researcher to visualize texts in a rather unusual manner. The text appears in the screen in the shape of a galaxy, where individual stars represent separate words. Words which are often used by the author are very bright, words which are more seldom employed are much paler, while rare ones are almost invisible. Text lines resemble spiral threads of a galaxy. The Textarc software enables a researcher to see the structure of a text, to visualize interconnections between words in the context, to see quantitative characteristics, concordance, and co-occurrence. The software enables a researcher to assess whether words are often used by a writer (I.S. Turgenev, for instance) to depict nature or can only be occasionally found in texts.

Masterfully depicting nature and being an expert of psychological prose, I.S. Turgenev has always been admired not only by Russian readers but also by readers in many other countries of the world. The so called Turgenev phenomenon can be investigated with the help of the VAAL-2000 software [6]. The software analyzes the potential influence of a text on the readers and investigates the author's personal and psychological characteristics.

Conclusion

It should be noted that linguistic software, which enables one to identify lexical means of expressiveness, can be used both by researchers and educators. Researchers can employ linguistic software to investigate how lexical means function in a literary text (how they highlight various shades of meaning, enable the readers to visualize the author's descriptions, intensify denotative meanings, stylistic and emotional connotations). Educators can resort to linguistic software to collect and systematize data which can be used to demonstrate how linguistic means function in speech.

Конфликт интересов

Не указан.

Conflict of Interest

None declared.

Список литературы / References

1. Зубов А. В. Информационные технологии в лингвистике: Учеб. пособие для студ. лингв. фак-тов высш. учеб. заведений / А. В. Зубов, И. И. Зубова. – М. : Академия, 2004. – 208 с.
2. Зубов А. В. Основы искусственного интеллекта для лингвистов: учеб. пособие / А. В. Зубов, И. И. Зубова. – М. : Университетская книга; Логос, 2007. – 320 с.
3. Информационные системы и применение компьютерной техники в профессиональной деятельности. Основные понятия // Электронный учебник по предмету ИТ [Электронный ресурс] – URL.: <http://it-uchebnik.ru/it/glava1/1.html> (дата обращения: 02.05.2020).
4. Компьютерная лингвистика // Онлайн Энциклопедия "Кругосвет" [Электронный ресурс] – URL.: http://www.krugosvet.ru/enc/gumanitarnye_nauki/lingvistika/KOMPYUTERNAYA_LINGVISTIKA.html (дата обращения: 02.05.2020).
5. Леонтьева Н. Н. Автоматическое понимание текстов: системы, модели, ресурсы: учеб. пособие для студ. лингв. фак. вузов / Н. Н. Леонтьева. – М. : Академия, 2006. – 304 с.
6. Логичев С. В. Каталог лингвистических программ и ресурсов в Сети / С. В. Логичев // Русская виртуальная библиотека [Электронный ресурс] – URL.: <http://www.rvb.ru/soft/catalogue/index.html> (дата обращения: 02.05.2020).
7. Мечковская Н. Б. История языка и история коммуникации: от клинописи до Интернета: курс лекций по общему языкознанию / Н. Б. Мечковская. – М. : Флинта: Наука, 2009. – 584 с.
8. Национальный корпус русского языка, разработанный при поддержке РГНФ, проект № 15-04-12018 «Развитие специализированных модулей НКРЯ» [Электронный ресурс] – URL.: www.ruscorpora.ru (дата обращения: 02.05.2020).
9. Персональная система автоматического анализа текстов TextAnalyst 2.0 // Microsystems, Ltd [Электронный ресурс]. — URL.: <http://www.analyst.ru/index.php?lang=eng&dir=content/downloads/> (дата обращения: 02.05.2020).

10. Потапова Р. К. Новые информационные технологии и лингвистика: Учебное пособие / Р. К. Потапова. – Изд. 4-е, стереотипное. – М. : КомКнига, 2005. – 368 с.

11. Тургенев И. С. Полное собрание сочинений и писем в тридцати томах / И. С. Тургенев. – Т. 3. М. : Наука, 1979. – 283 с.

Список литературы на английском / References in English

1. Zubov A. V. Informacionnye tekhnologii v lingvistike: Ucheb. posobie dlya stud. lingv. fak-tov vyssh. ucheb. zavedenij [Information technology in linguistics: Textbook. manual for stud. lingual. fact-tov higher. study. Institutions] / A. V. Zubov, I. I. Zubova. – Moscow: Academy, 2004. – 208 p. [in Russian]

2. Zubov A.V. Osnovy iskusstvennogo intellekta dlya lingvistov: ucheb.posobie [Fundamentals of Artificial Intelligence for Linguists: Textbook] / A. V. Zubov, I. I. Zubova. – Moscow: Universitetskaya Kniga; Logos, 2007. – 320 p. [in Russian]

3. Informacionnye sistemy i primeneniye komp'yuternoj tekhniki v professional'noj deyatel'nosti. Osnovnye ponyatiya [Information systems and the use of computer technology in professional activities. Basic concepts] // Electronic textbook on the subject of IT [Electronic resource] – 2009. – URL: <http://it-uchebnik.ru/it/glava1/1.html>. (accessed: 02.05.2020). [in Russian]

4. Computational linguistics // online encyclopedia Krugosvet [Electronic resource]. – 2001-2009. – URL: http://www.krugosvet.ru/enc/gumanitarnye_nauki/lingvistika/KOMPYUTERNAYA_LINGVISTIKA.html (accessed: 02.05.2020) [in Russian]

5. Leontyeva N.N. Avtomaticheskoe ponimanie tekstov: sistemy, modeli, resursy: ucheb posobie dlya stud. lingv. fak. vuzov [Automatic understanding of texts: systems, models, resources: textbook for students. lingual. fac. universities] / N. N. Leontieva – Moscow: Academy, 2006. – 304 p. [in Russian]

6. Logichev S.V. Katalog lingvisticheskikh programm i resursov v Seti [Catalog of linguistic programs and resources on the Web] / S. V. Logichev // Russian virtual library [Electronic resource]. – 1999-2009. – URL: <http://www.rvb.ru/soft/catalogue/index.html> (accessed: 02.05.2020). [in Russian]

7. Mechkovskaya N.B. Istoriya yazyka i istoriya kommunikacii: ot klinopisi do Interneta: kurs lekcij po obshchemu yazykoznaniyu [History of language and history of communication: from cuneiform to the Internet: a course of lectures on general linguistics] / N.B. Mechkovskaya. – Moscow: Flinta: Nauka, 2009. – 584 p. [in Russian]

8. Nacional'nyj korpus russkogo yazyka, razrabotannyj pri podderzhke RGNF, proekt № 15-04-12018 «Razvitie specializirovannykh modulej NKRYA» [National corpus of the Russian language, developed with the support of the Russian Humanitarian Science Foundation, project No. 15-04-12018 "Development of specialized modules of the RNC"] [Electronic resource]. –URL: www.ruscorpora.ru (accessed: 02.05.2020) [in Russian]

9. Personal'naya sistema avtomaticheskogo analiza tekstov TextAnalyst 2.0 [Personal system of automatic text analysis TextAnalyst 2.0] // Microsystems, Ltd [Electronic resource]. – 2009. Accessed at: <http://www.analyst.ru/index.php?lang=eng&dir=content/downloads/> (accessed: 02.05.2020) [in Russian]

10. Potapova P.K. Novye informacionnye tekhnologii i lingvistika: Uchebnoe posobie [New information technologies and linguistics: textbook] / RK Potapova. - Ed. 4th, stereotyped. – Moscow: Komkniga, 2005. – 368 p. [in Russian]

11. Turgenev I.S. Polnoe sobranie sochinenij i pisem v tridcati tomah [Complete works and letters in thirty volumes] / I. S. Turgenev.- V. 3. Moscow: Nauka, 1979. – 283 p. [in Russian]