Effects of anti-epidemic (quarantine) measures on people during the COVID-19 pandemic: applying social network analysis to identify the key topics

Pleshkova E.K.[1, 2], Rezanova Z.I.[1]

[1] *National Research Tomsk State University" (NR TSU)*
36, Lenina Av., Tomsk, 634050, Russian Federation

[2] *Siberian State Medical University (SSMU)*
2, Moscow Trakt, Tomsk, 634050, Russian Federation

ABSTRACT

The aim of this study was to examine the public reaction to the implementation of quarantine measures through a personality-oriented discourse.

Materials and methods. Text data were collected from a microblogging platform, resulting in a dataset of 86,750 texts related to the topics of "pandemic" and "quarantine measures". The lexical conceptualization of the pandemic and quarantine measures represented in the texts was analyzed through the lens of a personality-oriented discourse. Text lemmatization was conducted using the "snowball" library. A data feature matrix was then created based on the lemmatized tokens, which included 53 tokens with a frequency of use exceeding 1,300 times. The Social Network Analysis (SNA) method was used to create a keyword co-occurrence network consisting of undirected graphs. This analysis was performed using the free software R version 4.4.1, with the assistance of the Quanteda library, built-in "base" packages, and the gsub function.

Results. The resulting network consisted of 53 key lexemes, which actors used to respond to quarantine measures in the personality-oriented discourse. The central node of the network was "coronavirus", which was used 79,838 times between March 1 and April 30, 2020. The nearest nodes were "test" (used 4,663 times) and "Russia" (used 5,848 times). This network had high centrality, indicating that despite strict restrictive measures, the focus of the general public was on the pandemic itself and its impact on society rather than on the restrictions imposed.

Conclusion. The implementation of these anti-epidemic measures has created a unique sociolinguistic world view, reflecting the interaction between society and the outside world in a time of uncertainty and health risks, affecting the analysis of information and the behavioral strategies chosen by society.

Keywords: coronavirus, personality-oriented discourse, natural language processing, social network analysis, sociolinguistic world view

---

✉ *Pleshkova Ekaterina K.,* pleshkova.ek@ssmu.ru

Влияние противоэпидемических (карантинных) мероприятий в условиях пандемии COVID-19 на население: выявление ключевых тематик с помощью социально-сетевого анализа

Плешкова Е.К.[1, 2], Резанова З.И.[1]

[1] *Национальный исследовательский Томский государственный университет (НИ ТГУ)*
634050, г. Томск, пр. Ленина, 36

[2] *Сибирский государственный медицинский университет (СибГМУ)*
634050, г. Томск, Московский тракт, 2

РЕЗЮМЕ

Цель исследования – изучение реакции общества на введение карантинных мер по данным личностно ориентированного дискурса.

Материалы и методы. Произведен сбор текстовых данных на платформе микроблогов. Датасет состоял из 86 750 текстов, объединенных тематикой «пандемия», «карантинные меры». Проведен анализ лексической концептуализации пандемии и карантинных мер в личностно ориентированном дискурсе, репрезентированной в собранных текстах. Выполнена лемматизация на основе библиотеки Snowball, построена матрица datafeature matrix на основе лемматизированных токенов, включавшая 53 токена, частотность употребления которых превышала 1 300 раз. Методом социального-сетевого анализа построена сеть соприсутствия ключевых лексем, состоящая из неориентированных графов. Анализ был выполнен в свободном программном обеспечении R версии 4.4.1 с использованием библиотеки Quanteda, встроенных пакетов base и функции gsub.

Результаты. Получена сеть из 53 ключевых лексем, с помощью которых акторы личностно ориентированного дискурса реагировали на карантинные мероприятия. Ядро сети – узел «коронавирус» употреблено 79 838 раз в период с 1 марта по 30 апреля 2020 г. Ближайшие узлы: «тест» (употреблено 4 663 раза) и «Россия» (употреблено 5 848 раз). Сеть имеет высокую центральность, центральный узел сети «коронавирус» свидетельствует о том, что, несмотря на введение жестких ограничительных мер, население фокусировалось не на введенных ограничениях, а непосредственно на пандемии и ее влиянии на жизнедеятельность общества.

Заключение. Введение противоэпидемических мероприятий сформировало уникальную социолингвистическую картину мира, отражающую взаимодействие общества с внешним миром в условиях неопределенности и риска здоровью, влияющую на анализ информации и выбор поведенческой стратегии обществом.

Ключевые слова: коронавирус, личностно ориентированный дискурс, обработка естественного языка, социально-сетевой анализ, социолингвистическая картина мира

INTRODUCTION

New challenges have emerged with the advent of a new era. The once-familiar and orderly world, which many believed to be predictable, has been replaced by a volatile and unpredictable VUCA world, which has now evolved into a BANI world: brittle, anxious, nonlinear, and incomprehensible [1]. Each letter of the acronym BANI (brittle, anxious, nonlinear, incomprehensible) defines a new world. The term

was coined in 2020, when society faced a pandemic that brought about radical changes and triggered transformation that affected all aspects of society. The study of the COVID-19 pandemic is not only of interest for researchers in the fields of medicine, biomedicine, and economics, but also for linguists and philologists, as it is crucial to understand how the pandemic has transformed not only society, but also language as a means of expressing collective consciousness [2].

In this article, we will examine how society responded to the implementation of quarantine measures during the COVID-19 pandemic, based on personality-oriented discourse, as this discourse serves to illustrate and shape public opinion [3]. In recent years, the effectiveness of such studies has been enhanced by the use of natural language processing (NLP), computational linguistics, artificial intelligence, and social network analysis (SNA) on large corpora of texts, which allows for structuring information contained in texts. However, the focus of research has primarily been on social networks, with the social media platform X (formerly known as Twitter, which is blocked in the Russian Federation) being the leading platform until recently. By combining mathematical analysis and sociolinguistic analysis of texts, researchers are able to analyze various phenomena, such as the spread of diseases, the growth of protest movements in society, social inequality, the influence of social networks on youth communities, and much more [4]. For instance, Jinghao Wang et al. noted the use of social networks as a means of identifying how the effects of negative temperature dynamics, local air pollution, and natural disasters on individuals are interpreted [5].

Studying the pandemic through the lens of its manifestations in the language, it is important to consider specific stages of its development, particularly the implementation of quarantine measures and a public response to them, which was actively expressed. Researchers have interpreted the quarantine measures introduced in March 2020 in the Russian Federation as the largest psychological or psychosocial experiment, the study of which will continue for a long time in the social sciences using various methods [6].

MATERIALS AND METHODS

Social Network Analysis (SNA) is a method used to study various structures and relationships by applying graph theory to visualize and analyze connections between organizations or individuals in a network [7]. The studied network can be represented by any data, since the method allows for analyzing relationships between actors in a community, events related to a specific topic [8], relationships between scientific publications within a particular organization [9], and connections between lexical units in different types of discourse [10]. Essentially, SNA maps the relationships between different actors, providing a visual representation of how they interact with each other. This allows researchers to identify patterns and assess the strength of connections between network actors.

The analysis was conducted using the free software R version 4.4.1 and the Quanteda library [11], as well as the built-in base package and the gsub function for replacing string sections. In this study, we analyzed the lexical conceptualization of the pandemic and quarantine measures in personality-oriented discourse. At the first stage, the text data was collected from the social media platform X (formerly known as Twitter, which is blocked in the Russian Federation) to form the dataset.

The dataset consisted of 86,750 tweets (short texts) related to the topics of "pandemic" and "quarantine measures", posted between March 1 and April 30, 2020. During the preparation stage, hyperlinks, hashtags, and user tags were removed, as well as punctuation marks and numbers. A corpus of texts was put together based on the dataset, and then it was tokenized. Stop words were removed, and lemmatization was performed based on the Snowball library. A data feature matrix was then created based on the lemmatized tokens. The matrix only included tokens that were mentioned more than 1,300 times. So, the matrix comprised 53 tokens. A feature co-occurrence matrix was then built, including the most frequently occurring words in texts related to the pandemic and quarantine measures. Using this matrix, a feature co-occurrence matrix was created using the textplot_network function. Finally, social network analysis methods were used to study how information about the introduction of quarantine measures was distributed and functioned in personality-oriented discourse.

RESULTS

The analysis resulted in a network of 53 key lexemes that were used at least 1,300 times (Fig. 1). These lexemes were frequently used by actors in personality-oriented discourse to express their attitude towards the introduction of quarantine measures. The resulting network illustrates the lexical landscape of discussions concerning quarantine measures from

March 1 to April 30, 2020. This is a single-core network with high centrality. The network nodes are the most commonly used lexemes related to the pandemic and quarantine measures, while the edges of the graph represent the connections between these lexemes. The weight of the graph edge indicates the strength of the connection between the nodes: the greater the weight, the more frequent the use. This pointed out that these lexemes were frequently used together, thus forming a semantic unity.
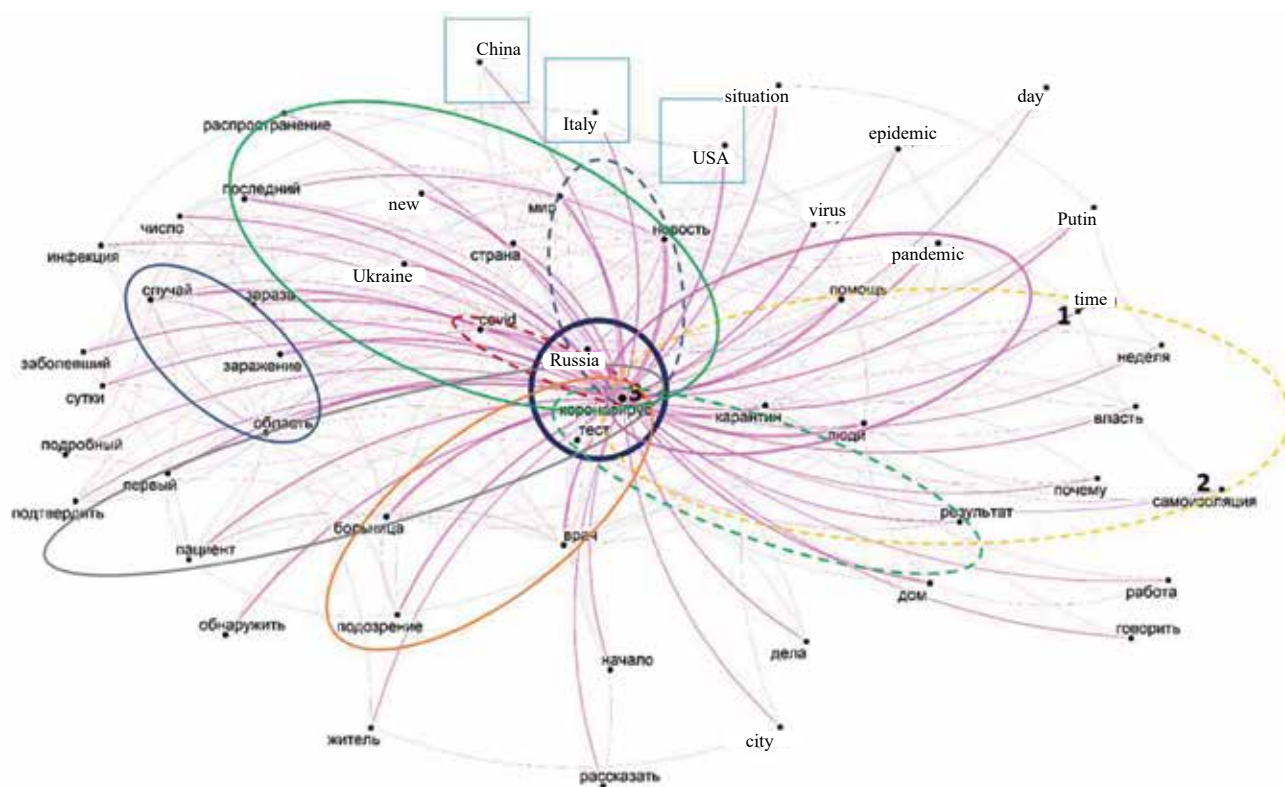

Figure. Network displays the most commonly used lexemes in microblogs collected from the X platform (formerly known as Twitter, blocked in the Russian Federation). These texts were published in response to lockdown measures in March–April 2020. The network consists of 53 key lexemes that were used at least 1,300 times

The core of the network, the most frequently used word, is "coronavirus" which was used 79,838 times in texts and had the highest frequency of use during the period of quarantine measures (March 1 – April 30, 2020). The cluster of key lexemes closest to the network core includes "coronavirus" («коронавирус») (used 79,838 times), "test" («тест») (used 4,663 times), and "Russia" («Россия») (used 5,848 times) and is highlighted by a dark blue circle with a solid line. The nodes "Russia" – "coronavirus", and "test" – "coronavirus" have the smallest geodesic distance in the network, as well as the highest weight of the edges connecting them, indicating frequent co-occurrence of these lexical units. It is likely that in their texts, actors of personality-oriented discourse discussed the incidence (increase in incidence) of the novel coronavirus infection and the availability of diagnostic tests for COVID-19 (e.g. "Today, anyone can take a

coronavirus test, even without symptoms" or "What rapid tests for COVID-19 exist?" («Сегодня тест на коронавирус может сдать любой желающий, даже при отсутствии симптомов», «Какие существуют экспресс-тесты на COVID-19?»), etc.).

It is worth noting that in the Russian segment of the social media platform X (formerly known as Twitter blocked in the Russian Federation), people predominantly use the Russian name for the virus, compared to the English name "COVID" (3,970 vs 79,838 uses of the lexeme "coronavirus").

The lexeme "COVID" was often used together with "coronavirus" as seen in the graph (Figure) in the area highlighted by a red dashed line ellipse. A graph with nodes "COVID" and "coronavirus" connected by a high-weight edge confirms the frequent co-occurrence of these lexemes in texts related to the pandemic and quarantine measures. The high edge weight and small

geodesic distance between the nodes also indicate a fairly frequent co-occurrence of these lexical units. Most often, the co-occurrence of "coronavirus" and "COVID" was observed in tweets and hashtags (e.g. "Yes, people wear masks, but the Chinese, even without coronavirus, always wore masks because of the air #covid #coronavirus" (Да, народ ходит в масках, но китайцы даже без короновируса на моей памяти всегда ходили в масках из-за воздуха #covid #coronavirus)). It is worth noting that despite the negative impact of restrictive measures, society still focused on discussing the original source of the problem – the new virus – giving it a predominant role in their discussions within personality-oriented discourse.

The graph shows a cluster of lexemes, including "world" («мир») (3,154 uses), "news" («новость») (3,179 uses), and "coronavirus", highlighted by a blue dashed line ellipse. This suggests that the Russian users on the X platform (formerly known as Twitter, which is blocked in Russia) discussed information from both Russian and international media (represented by the "world" node). This aligns with the findings of international and Russian researchers, who have noted that media discourse is the main source of information about the pandemic, the novel coronavirus infection, and quarantine measures[5, 12, 13]frequent global measurements of affective states to gauge the emotional impacts of pandemic and related policy interventions remain scarce. Using 654 million geotagged social media posts in over 100 countries, covering 74% of world population, coupled with state-of-the-art natural language processing techniques, we develop a global dataset of expressed sentiment indices to track national- and subnational-level affective states on a daily basis. We present two motivating applications using data from the first wave of COVID-19 (from 1 January to 31 May 2020. This connection indicates that society primarily receives information from media discourse rather than other sources, such as business discourse, which may provide unbiased information (orders, regulations, laws, etc.). The graph confirms this statement since its edge connecting the "coronavirus" and "news" nodes has high weight, which indicates a strong connection and a high degree of co-occurrence of these lexemes. This may contribute to the growing infodemic in personality-oriented discourse, as media discourse is not always unbiased.

One of the most discussed topics during the lockdown is visualized in an interesting way. It includes the following nodes: "coronavirus" + "test" + "result" («результат»), which naturally should have formed a semantic unity with high co-occurrence. The resulting network contains this graph, the nodes of which are lexemes "coronavirus" (79,383 uses), "test" (4,663 uses), and "result" (approximately 1,300 uses). However, the nodes "test" and "result" are connected by an edge with very low weight, which suggests that people hardly ever mentioned these lexemes together. Instead, the lexeme "test" was more frequently used with the lexemes "coronavirus" and "result", since the edge connecting it with the former lexeme is of high weight and a small geodesic distance. The edge connecting "test" and "result" has lower weight than that connecting "test" and "coronavirus", but its weight is higher than that connecting the lexemes "result" and "test", with a large geodesic distance between the nodes.

When planning to study the way lockdown is reflected in personality-oriented discourse, we assumed that such units as "restrictions" («ограничения») and "self-isolation" («самоизоляция») would be among the most frequently used and form a large number of combinations (semantic unities). However, the node "restrictions" was not included in our network of 53 most commonly used lexemes. The network contains the node "self-isolation", which is located on the periphery of the network, far from the core node ("coronavirus"). This suggests that "self-isolation" was not a commonly used term in lockdown-related texts. Additionally, the node for "self-isolation" ($\geq$ 1,300 uses) is part of a graph that includes the following nodes: "coronavirus", "time" («время») ($\geq$ 1,300 uses), and "self-isolation". This graph is highlighted by a yellow dotted line ellipse, and the nodes are labeled in the following way: "time" – 1, "self-isolation" – 2, and "coronavirus" – 3. The nodes "time" and "self-isolation" are connected by a weak edge, and the geodesic distance between them is quite large, indicating that these terms were not often used together. It is worth noting that the nodes "time" and "coronavirus" and the nodes "self-isolation" and "coronavirus" also have a low degree of connectivity (with weak edges and a large distance between them).

On the other hand, the graph consisting of the nodes "coronavirus" (used 79,383 times), "help" («помощь») (used 3,923 times), "pandemic" («пандемия») (2,048 times used), and "people" («люди») (3,556 times used) most likely represents a discussion on the topic of medical care provided during the pandemic. The topic related to diagnosing the novel coronavirus

infection among the population by presentation to medical facilities is represented by a cluster of words associated with the "coronavirus" node, which reflects the development of the topic devoted to the pandemic and its impact on social and demographic phenomena. This cluster is highlighted by an orange solid line ellipse and includes the words "coronavirus" (79,383 times used), "suspicion" («подозрение») (2,130 times used), "hospital" («больница») (2,762 times used), and "doctor" («врач») (2,365 times used). It is interesting to note that the edge connecting the nodes "coronavirus" and "hospital" is stronger than the one connecting "coronavirus" and "suspicion", indicating a higher level of connectivity and confirming the joint use of these key lexemes in personality-oriented discourse. The same can be observed with the nodes "coronavirus" and "doctor". This suggests that the actors in personality-oriented discourse discussed the medical community (represented by the node "doctor") and the incidence statistics (represented by the joint use of "coronavirus" and "hospital"), which allows us to identify these clusters in the resulting network.

It is worth noting that the beginning of the pandemic was a frequently discussed topic in personality-oriented discourse, which is evidenced by the presence of a coherent block of co-occurrent terms: "first" («первый») + "patient" («пациент») + "coronavirus". It is obvious that the topic discussed was the first patients with coronavirus, which can be visualized as a graph with the nodes "coronavirus" (79,383 times used), "first" ($\geq$ 1,300 times used), and "patient" (2,029 times used). The cluster is highlighted by a gray solid line ellipse. Additionally, it is worth mentioning that the cluster also includes the node "region" («область») (3,127 times used). This node is part of another cluster, forming a semantic unity with the nodes "case" («случай») (2,602 times used) and "infection" («заражение») (2,418 times used). However, its proximity to the cluster reflecting public attitude towards the beginning of the pandemic suggests that people learning about the first people infected with COVID-19 looked up the statistics of their region. The co-occurrence and high frequency of the key lexemes "region", "case", and "infection" indicate that people actively sought information about the statistics of incidence in their region and expressed their own attitude towards the situation through short text messages. This hypothesis is further supported by the fact that the key lexemes "infection" and "case", which are mentioned with high frequency, form a semantic

unity "case of infection" («случай заражения»), which is also frequently mentioned (596 times used) (Table 2). For example, "New cases of COVID-19 infection have been registered in the Kostanay region" («Новые случаи заражения Covid-19 зарегистрированы в Костанайской области»).

The cluster, including the key lexemes "news", "latest" («последний»), and "coronavirus", located at a large geodesic distance from each other, confirms that participants of personality-oriented discourse relied on information obtained from the news when expressing their attitude towards the pandemic and the introduction of restrictive measures. The presence of this cluster in the network, as well as the high frequency of mentions of the key lexeme "news" (3,879 times used) and the phrase "latest news" («последние новости») (822 times used) highlights the close connection of personality-oriented discourse and media discourse. This connection became even more significant during the COVID-19 pandemic, as society was seeking new sources of information to make informed decisions in the face of increasing risks to life and health.

The lexemes "virus" («вирус») (2,245 uses) and "epidemic" («эпидемия») (2,248 uses) formed graphs in the network with the core node "coronavirus", which may indicate that the actors of personality-oriented discourse, as expected, focused on everyday problems and restrictions brought by the pandemic, without discussing the phenomenon of the pandemic as a whole. The absence of connections with the node "virus" may indicate a close connection of personality-oriented discourse and media discourse. People received information from the media and expressed their attitude to the information received, using the terminology specified in media discourse ("coronavirus", and not just "virus" or "disease").

The nodes "China" («Китай») ($\geq$ 1,300 times used), "Italy" («Италия») ($\geq$ 1,300 times used), and "USA" («США») (2,182 times used) indicate that users of the Russian segment of the X platform (formerly known as Twitter, blocked in the Russian Federation) took into account world news when forming their own attitudes towards the new coronavirus pandemic and the introduction of lockdown but were still more focused on the situation in the Russian Federation. This is confirmed by the remoteness of the nodes from the network core, as well as the fact that the nodes are linked to other key lexemes by edges of lower weight, demonstrating a low degree of connectivity of the key lexemes. In addition, the lower interest in foreign

scenarios for the development of the pandemic is confirmed by the frequency of mentions of countries: "China" (≥ 1,300 times used), "Italy" (≥ 1,300 times used), and "USA" (2,182 times used) in comparison with the key lexeme "Russia" (5,848 times used), which ranks second in the frequency of use after the lexeme "coronavirus".

The incoming centrality of the network is estimated by the number of connections included in the node; it can be said that the centrality corresponds to the concept of "popularity". Evaluating the constructed network, we can say that the most popular topic discussed during the period of introduction of quarantine measures was not restrictions but coronavirus, since it is this lexeme that represents the core of the network with the largest number of connections included in it.

In order to estimate the frequency of key lexemes that formed the network and demonstrated the main topics that concerned the population during the lockdown, we used the data feature matrix function in the Quanteda library. We calculated the sum of key lexemes in the columns of the data feature matrix, each row of which is a separate tweet (microblog). The results obtained are shown in Table 1.

Table 1

Numerical expression of nodes located in the network demonstrating key lexemes and their interactions in personality-oriented discourse during the lockdown period

coronavirus	79,838
Russia	5,848
test	4,663
quarantine	4,488
COVID	3,970
help	3,923
news	3,879
people	3,556
world	3,154
region	3,127
home	2,864
hospital	2,762
case	2,602
new	2,484
infection	2,418
doctor	2,365
virus	2,245
Putin	2,207
USA	2,182
suspicion	2,130
pandemic	2,048
patient	2,029

The results presented in Table 1 support the visual representation of personality-oriented discourse (Figure). "Coronavirus" is the most frequently used

lexeme, which formed the basis of all combinations expressing the public attitude towards the introduction of quarantine measures and the pandemic. "Russia" is the second most frequently used lexeme, as the analysis was conducted using information from the Russian-speaking segment of the social network. An interesting observation is that the lexeme "pandemic" has one of the lowest frequency rates. It can be assumed that the term "pandemic" is more common for media or business discourse, where the professional medical community expressed their views on the pandemic. In contrast, personality-oriented discourse is characterized by lexemes that reflect the topics concerning the general public during the specified period, such as "coronavirus", "Russia", "quarantine" («карантин»), and "test".

Table 2

The most frequently used phrases in personality-oriented discourse, including key lexemes

Collocation	Number of uses
coronavirus test	3,217
suspected coronavirus	1,883
coronavirus detected	1,136
coronavirus in Russia	1,004
latest news	822
infected with coronavirus	711
coronavirus infection	707
to overcome the coronavirus	599
cases of infection	596
coronavirus analysis	591
coronavirus has been detected	578
coronavirus has been diagnosed	566
the pandemic of coronavirus infection	553
coronavirus testing	550
spread of coronavirus	518
coronavirus news	458

In order to determine the most frequently used semantic units related to quarantine measures between March 1 and April 30, 2020, the textstat_collocations function from the Quanteda library was utilized. The top 20 phrases were extracted. The results are shown in Table 2. This analysis reveals the main topics that were discussed in personality-oriented discourse during the lockdown period from March 1 to April 30, 2020, represented as clusters in the resulting network.

DISCUSSION

The analysis demonstrates that the topic of the novel coronavirus infection, as well as the introduction of quarantine measures, consists of several interconnected key topics. These topics form visually identifiable clusters in the constructed

network, with a high centrality and a core "coronavirus". This confirms the presence of coherent links between the texts studied during the period of March 1 – April 30, 2020, in non-institutional type of discourse. Additionally, the resulting network indicates the integration of discourses in which there was an active discussion of the introduction of quarantine measures in the Russian Federation. The population received information about the pandemic and quarantine measures through media discourse, which formed their perception of the risk to their own health and life, which often led to growing panic due to both information overload and insufficient information. The key term "coronavirus", which forms the core of the network, encompasses discussions on public health, incidence statistics, incidence rates in different regions, the spread of the virus, and measures to contain it. It also includes discussions on measures to help and support the population. The analysis revealed that the key lexemes reflecting the main concepts and concerns of the general public during the pandemic and the introduction of quarantine measures have a sociocultural dependence. They allow us to form a linguistic worldview during this pandemic.

CONCLUSION

This study confirms the presence of a unique set of topics that reflect the sociolinguistic worldview that was shaped by the pandemic and the introduction of quarantine measures. These topics identified through key lexemes not only reveal public attitude towards the pandemic and quarantine measures but also shed light on the evolution of public opinion, sources of information, and factors influencing the interpretation of information. Considering the findings of other researchers studying the pandemic through the lens of sociology, philosophy, and linguistics, we can note a growing trend towards a comprehensive approach to studying the effects of the pandemic on society. This approach involves examining the transformation of society, the reactions and behaviors of the population, and the role of institutional and non-institutional discourses. What makes this study outstanding is its integration of mathematical and sociolinguistic research methods to analyze the complex impact of the pandemic and related measures on the general public. However, it should be noted that this analysis

alone cannot provide a complete sociolinguistic worldview of the public response to the pandemic and quarantine measures. To achieve this, it would be beneficial to supplement the existing study with frame analysis and sentiment analysis of texts related to the novel coronavirus infection. Additionally, comparing the representation of the pandemic and quarantine measures across different types of discourse would be an interesting avenue for future research.

REFERENCES

1. Wasserman S., Faust K. Social network analysis: methods and applications (Structural analysis in the social sciences, series number 8). Cambridge University Press; 1994. DOI: 10.1017/CBO9780511815478.

2. Maslova V. A. *Homo lingualis* in culture. Moscow: Gnosis; 2007.320 p. (in Russ.).

3. Zubkova Y. Discourse studies: new results (review of the book: Karasik V.I. Language plasticity of communication). Moscow: Gnosis; 2021.536 p. DOI: 10.18254/S294939000028978-4.

4. Carrington P.J., Scott J., Wasserman S. Models and methods in social network analysis. Cambridge: Cambridge University Press; 2005. DOI: 10.1017/CBO9780511811395.

5. Wang J., Fan Y., Palacios J., Chai Yu., Guetta-Jeanrenaud N., Obradovich N. et al. Global evidence of expressed sentiment alterations during the COVID-19 pandemic. *Nature Human Behaviour.* 2022;3(6):349–358. DOI: 10.1038/s41562-022-01312-y.

6. Pleshkova E. K. Strategies of risk communication in socially significant discourses during the pandemic of a new coronavirus infection: a social network analysis. *Russian Linguistic Bulletin.* 2023;42(6) (in Russ.). DOI: 10.18454/RULB.2023.42.8.

7. Barabási A.-L., Pósfai M. Network science. Camridge: Cambridge University Press; 2016.

8. Borgatti S.P., Everett M.G., Johnson J.C. Analyzing social networks. SAGE Publications Ltd.; 2018.

9. Basarab M.A., Glinskaya E.V., Ivanov I.P., Kolesnikov A.V., Kuzovlev V.I. Study into the structure of the scientific coathorship graph using social network analysis. *Cybersecurity Issues.* 2017;1(19):31–36. (In Russ.). DOI: 10.21681/2311-3456-2017-1-31-36.

10. Rezanova Z.I., Stepanenko A.A. Discursive variants of thematic modeling of COVID-19 (news media discourse VS social networks). *Tomsk State University Journal of Philology.* 2023;(86):84–101 (in Russ.). DOI: 10.17223/19986645/86/6.

11. Benoit K., Watanabe K., Wang H., Nulty P., Obeng A., Müller S. et al. quanteda: An R package for the quantitative analysis of textual data. *Journal of Open Source Software.* 2018;30(3):774. DOI: 10.21105/joss.00774.

12. Chew C., Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PloS One.* 2010;11(5). DOI: 10.1371/journal.pone.0014118. DOI: 10.51217/npsyresearch_2023_03_04_08.

Authors' information

Pleshkova Ekaterina K. – Junior Researcher, Laboratory of Linguistic Anthropology, NR TSU, Head of the Department of International Development, Siberian State Medical University, Tomsk, pleshkova.ek@ssmu.ru, https://orcid.org/0000-0002-9075-429X,

Rezanova Zoya I. – Dr. Sci. (Philol.), Professor of the Department of General, Computer and Cognitive Linguistics, NR TSU, Deputy Head of the Laboratory of Anthropological Linguistics, NR TSU, Tomsk, rezanovazi@mail.ru, https://orcid.org/0000-0002-0550-991X

(✉) **Pleshkova Ekaterina K.,** pleshkova.ek@ssmu.ru