ARTICLE

Activation of «COVID» Words in the Kazakh Language: Statistical Analysis

Zhazira Agabekova [1], Aigul Amirbekova [2], Kalamkas Kalybekova [3], Dana Ospanova [4], Bayan Abylbayeva [5], Bolat Khassenov [6]*

[1] *Department of Kazakh Language and Turkic Studies, Nazarbayev University, Astana, Kazakhstan*

[2] *Institute of Linguistics named after A. Baitursynov, 050000 Almaty, Kazakhstan*

[3] *Karaganda Technical University Named after Abylkas Saginov, Karaganda, Kazakhstan*

[4] *National Scientific and Practical Center «Til-Qazyna» Named after Shaysultan Shayakhmetov, Astana, Kazakhstan*

[5] *Bolashaq Academy, Karaganda, Kazakhstan*

[6] *Buketov Karaganda University, Karaganda, Kazakhstan*

ABSTRACT

Some words in the Kazakh language became more actively used during the pandemic. This article describes which words were most often used in the Kazakh language during the COVID-19 pandemic. The research materials were taken from Google Trends and the National Corpus of the Kazakh language. The analysis showed that in the Kazakh language words like *koronavırýs-coronavirus* (noun); *covid* (noun); *COVID*-19 (noun); *pandemıa-pandemic* (noun); *ózin-ózi oqshaýlaý-self-isolation* (pronoun+gesture name); *lokdaýn-lokdaun* (noun); *áleýmettik qashyqtyq-social distance* (adjective+noun); *qashyqtyqtan oqytý-distance learning* (noun+gesture name) were frequently used. The article has developed a visualization of the frequency of use of words with statistical diagrams. As a result of the study, we came to the conclusion that new words appeared in the Kazakh language that were actively used during the pandemic, it was also noticed in March 2020 that the words coronavirus, covid, COVID-19, pandemic were quite popular in the Kazakh language.

Keywords: Pandemic; COVID-19; Kazakh language

1. Introduction

The coronavirus pandemic began in 2019, when in mid-December it was discovered that residents of the Huainan Animal and Seafood Market in Wuhan, part of Hubei Province, located in central China, fell ill with unexplained pneumonia. The COVID-19 pandemic in Kazakhstan began on March 13, 2020, when infected people were identified in the cities of Nursultan and Almaty. By April 3, COVID-19 had already spread throughout the country.

On March 15, President of Kazakhstan Kassym-Jomart Tokayev signed a decree introducing a state of emergency in the country. Since March 22, exits from these cities were completely closed, and air and rail traffic was suspended. The pandemic situation greatly affected not only the economy but also the education system and language changes (Balabekova et al., 2021).

The sudden emergence of coronavirus infection quickly impacted all world languages, and the Kazakh language was no exception. Words and phrases like lockdown, distance learning, Corona, restrictive measures, masks, etc., quickly became frequently used. The academic lexicography of countries around the world also paid close attention to the intensive process of 'coronavirus neology.' In some works, it is shown that the COVID-19 pandemic contributed to the emergence of lexical innovations, the introduction of borrowed words, and new associations (Adilova et al., 2023; Daminov, 2021; Tan et al., 2020; Voskresenskaya et al., 2021; Zholshayeva & Zhanabekova, 2020). During the coronavirus pandemic, many neologisms associated with COVID-19 emerged, showing new word-building trends. Al-Salman and Haider (2021) examined the nature of new words and phrases that appeared in the English language due to the pandemic. Their study identified the types of word formation underlying these neologisms. The authors compiled a corpus of 208 neologisms related to COVID-19, collected from various internet sources: social networks, search engines, blogs, and news articles. The analysis showed that these neologisms were based on various word formation approaches, including affixation, compounding, blending, clipping, and complex word formation models, particularly the combination of compounding and blending.

Some research suggested that certain sounds used in world languages produce more droplets compared to others, which can significantly impact the spread of viruses depending on the language spoken. For example, Asadi et al. (2020) investigated the effect of voice and articulation mode on the emission of aerosol particles during human speech. They measured the particle separation rate of 56 healthy individuals who isolated sounds and spoke aloud. The results showed that some vowels (e.g., /i/) produced more particles than others (e.g., /ɑ/), and voiced plosive consonants (e.g., /b/) produced more particles than voiceless fricatives (e.g., /f/). Abkarian and Stone (2020) provided new data on the mechanisms of droplet formation in the oral cavity. They recorded high-speed video of a volunteer producing various sounds. The results showed that consonants like /b, d, p, t/ produced more saliva because they involved expelling air through a narrow space filled with saliva. In contrast, consonants such as /m/ produced only a few droplets as air passed through the nose. All consonants producing many droplets in speech had a similar articulation method; they were stop or plosive consonants, which occur when the articulators (e.g., lips, tongue) completely close, preventing air from escaping the mouth. When the articulators separate, air is released with a slight explosive sound. Inouye (2003) developed a controversial hypothesis suggesting that Japanese tourists in China were less likely to suffer from SARS than American tourists. The author proposed that the use of aspirated consonants increased the chances of SARS transmission from person to person, as such consonants produced more droplets compared to other sounds. He pointed out the possibility of Chinese vendors speaking Japanese with Japanese tourists, where aspiration is weaker, and English with American tourists, where aspiration is stronger; this might explain the zero infection rate among Japanese tourists. A subsequent study by Inoue and Sugihara (2015) provided additional support for this hypothesis, showing that the wind pressure and articulatory force were weaker for Japanese compared to English and Chinese. Aspiration refers to the burst of air that occurs after the articulation of a consonant and before the beginning of a vowel (Abkarian & Stone, 2020). Thus, aspirated consonants involve a release of air.

Research has been done in a specific language, as well as comparative studies that cover several language materials. For example, Radbil's research (2021) focused on the vocabulary of COVID-19 of the Russian language from a linguocognitive point of view. He investigated the semantic,

lexical and word-formation features of the emergence and spread of new phenomena in the speech practice of Russian-speaking people during the COVID 19 pandemic; Oluwate-niola et al., (2021) analyzed in a sociolinguistic aspect the lexical innovations and variations in Nigerian English lexemes that appeared during the COVID-19 pandemic; and in Danilenko's work (2022) new words in Ukrainian, Russian and Czech languages that appeared during the pandemic were considered (Danilenko, 2022). That is, this research would focus on the facts of multiple languages.

The COVID-19 pandemic led to the emergence of new incentives and behaviors. In this context, Kazakhstani scientists (Adilova et al., 2023) conducted an associative experiment, which revealed that some concepts had acquired new associative meanings. This indicated that the collective experience of the pandemic had changed how people perceived and connected with certain ideas and terms. There were also studies (Balabekova) on newspaper materials, which revealed many borrowings in the Kazakh language.

This study is aimed at studying the activation of words during coronavirus infection. Through Google Trends and the National Corpus of the Kazakh language (https://qaz-corpora.kz/) , we would like to demonstrate that during the Covid-19 pandemic, some words were activated.

2. Materials and methods

2.1 Google trends

Google Trends is a publicly available web application provided by Google, which allows analyzing the search frequency of specific terms across different regions worldwide based on the overall volume of search queries.

Selection of keywords

For this study, key terms and phrases related to the COVID-19 pandemic in Kazakh language were selected, such as "coronavirus", "COVID", "pandemic", "quarantine", "self-isolation", "lockdown", and others.

Analysis of time series

The Google Trends function was utilized to analyze time series data of the popularity of selected keywords. The aim was to identify temporal peaks in the activation of these terms during different phases of the pandemic.

2.2 National Corpus of the Kazakh language

The National Corpus of the Kazakh Language is an electronic database of texts in the Kazakh language, facilitating the search for linguistic phenomena in natural contexts.

Selection of textual corpora

Available textual corpora from the National Corpus of the Kazakh Language were used for analysis, comprising texts of various genres and styles, spanning from the beginning of the pandemic.

Content analysis

Content analysis methods were applied to study the context and frequency of new words and expressions related to the pandemic in natural textual data. This enabled an assessment of how new words integrated into various linguistic contexts, including news articles, scientific papers, social media discussions, and others.

2.3 Statistical data analysis

The following methods were employed for data processing and analysis:

Graphical representation of data

Temporal graphs and diagrams were constructed to visualize changes in the frequency of key terms over time.

Statistical metrics

Statistical metrics were computed to assess changes in the popularity of words and phrases during different phases of the pandemic.

Content analysis

A qualitative assessment of changes in the meaning and use of key terms was carried out based on their contextual analysis in text corpora.

3. Results and discussion

During the pandemic, words originating mainly from the word classes of nouns, adjectives, and verbs have been prevalent. This shift in language reflects the rapid adaptation of society to unprecedented circumstances, bringing new terminology into everyday use.

The nouns introduced during the pandemic covered a

wide range of new and adapted terms: Sinovac Biotech, Sputnik light, Pfizer/BioNTech, spacesuit, mask, blood pressure measuring devices, alcohol, antiseptic, CVI, posts, protective glasses or screens, vaxirs, antivaxirs, infection, incubation.

The vast majority of "coronavirus" nouns in the Kazakh language are borrowed from other languages, primarily through Russian. Some words, like "mask" and "infection," existed previously but became more active, while others, like "Sputnik light" and "vaxers," were newly introduced. According to Klimchenko (2021), these borrowings have enriched the Kazakh lexicon, reflecting the global nature of the pandemic's impact.

Adjectives and Noun-Adjective Combinations. Examples of adjectives and noun-adjective combinations include: Symptoms and medical terms: Fever, dry cough, infectious disease, inpatient care; Healthcare facilities and measures: Inpatient quarantine, clinical trial, intensive care, special contingent, mobile vaccination points, preventive vaccinations; Pandemic-related zones: Dangerous process, red zone, yellow zone, green zone, coronavirus traffic light, epidemiological.

Most adjectives are translated from Russian as calque words, meaning they were directly borrowed and translated. Terms like "green zone" and "red zone" became essential in everyday language, designating areas with different levels of pandemic-related restrictions. The activity of suffixes forming adjectives from nouns has increased, indicating a dynamic language evolution in response to the crisis.

Verbs. Examples of verbs include: 1) Isolation and distancing measures: Self-isolation, social distancing; 2) Educational changes: Distance learning; 3) Emergency actions**: Evacuation. 4) Social media campaigns**: #StayAtHome.

Calque variants of verbs dominate, showing direct influence from other languages. Many of these verbs were rarely used before the pandemic but became common as new behaviors and policies emerged. This adaptation highlights the agility of language in meeting communicative needs during crises.

Many pandemic-related words saw a surge in usage during specific periods but have since declined as the pandemic's intensity lessened. Using the Kazakh corpus and Google Trends, we analyzed the frequency of selected words (**Table 1**) to illustrate this trend.

Table 1. *The words that were selected for analysis.*

№	Words
1	*coronavirus*
2	*covid*
3	*pandemic*
4	*quarantine*
5	*self-isolation*
6	*lockdown*
7	*distance learning*
8	*social distance*
9	*red zone, yellow zone, green zone*

3.1 Coronavirus, covid

The term *coronavirus* has been used in the media since 2000. However, it was widely used in 2020. As a result of the actualization of the concept and clarification of the meaning (Onikbaeva & Davis, 2021), there were also changes in the field of using the word: if earlier it was related to the field of medical terminology, then the new meaning was fixed in the colloquial language Dr, both in the language of official documents, and in the language of the media and the internet, which A synonym for the noun coronavirus is covid. This word is the absolute neologism of 2020. It is derived from the English acronym COVID-19 (from coronavirus Disease 2019). This means an infection caused by a certain new type of coronavirus-SARS-CoV-2 (Urazova & Serdaly, 2021). Initially, this word was mainly professional slang for medical workers. During the pandemic, it quickly entered the media and eventually became widespread in colloquial speech.

The period of popularity, relevance of the word coronavirus is the first of March 22 and April 2020 (**Figure 1**). And the dynamics of the relevance of its synom is as follows (**Figure 2**):
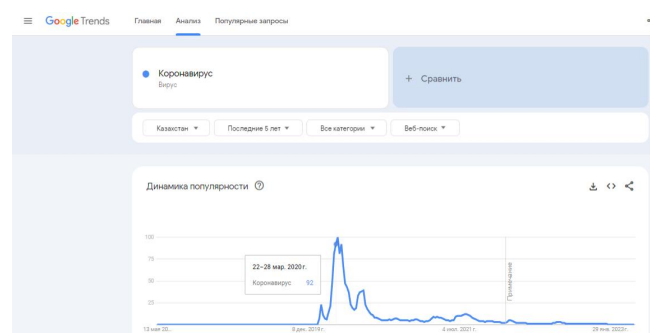

Figure 1. *Coronavirus.*

In the first picture, it seems that the word coronavirus first became an asset and then began to fall out of popularity, which can be associated with the emergence of the jargon of the synonym covid.
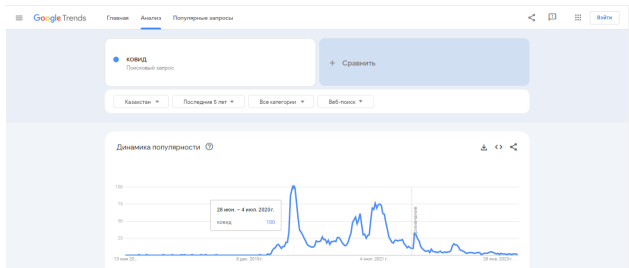

Figure 2. *Covid*.

In the National Corps (**Figure 3**), it is clearly seen that the word *coronavirus* was very often supported in 2020, as opposed to the word *covid*. *Covid* is very little exposed because of its use as slang.
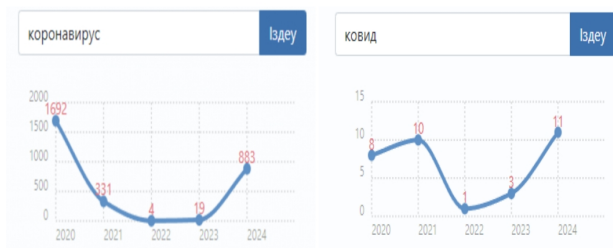

Figure 3. *National Corps (coronavirus and covid).*

3.2 Pandemic

The word *pandemic* is also one of the lexemes used by asset in the Kazakh language. The time of popularity is March 8–14, 2020 (**Figure 4**). Pandemic (Greek. πανδημία "The whole population") is the highest degree of development of the epidemic process, a very strong epidemic in which an infectious disease affects a significant part of the population of many countries in a relatively short time, sometimes on different continents. A pandemic is the most dangerous type, that is, an epidemic that covers most of the world.

3.3 Quarantine, self-isolation, lockdown

One of the words activated during the coronavirus pandemic is *quarantine*. Activation time-March 29–April 4 (**Figure 5**). Quarantine (Fr. quarantaine, ital. quarantena -

"forty days")-a set of restrictive and regime anti-epidemic measures aimed at limiting the contacts of an infected person, animal, cargo, goods, vehicle, settlement, territory, districts, regions, etc. In some cases, quarantine implies complete isolation of the epidemic focus with armed security around the perimeter. Quarantine is aimed at disrupting the mechanism of transmission of infection.
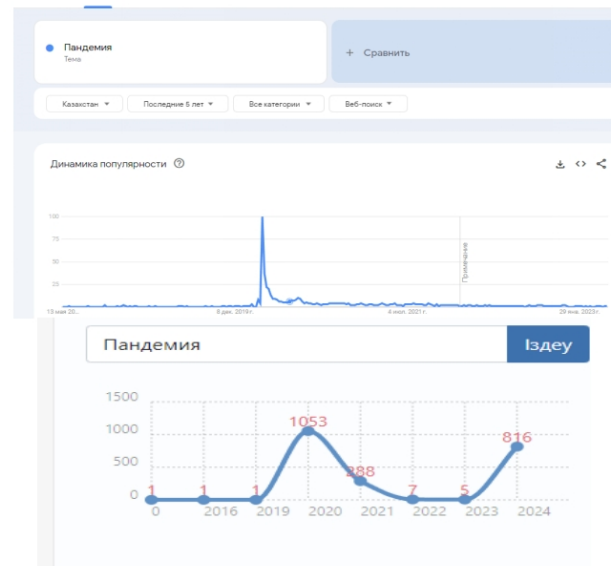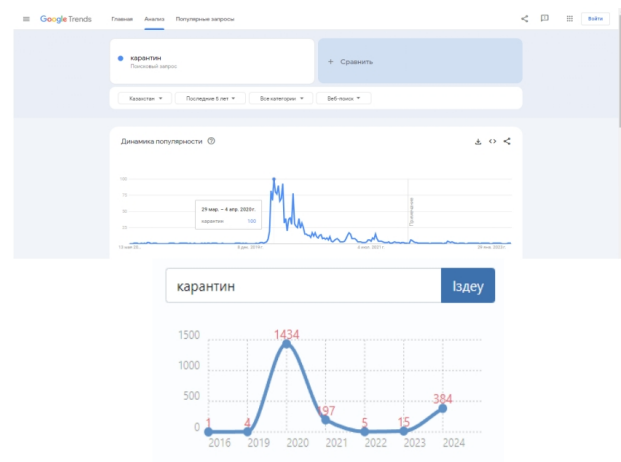

Figure 4. *Pandemic*.

Figure 5. *Quarantine*.

In addition to the noun *quarantine* to denote a set of measures and restrictions imposed to counter coronavirus infection, the lexeme of self-isolation (оқшаулау) is also widely used in modern language. The noun self-isolation has been used in Russian since the 30s of the twentieth cen-

tury. For example, the policy of self-isolation (a direction in foreign policy), self-isolation as a psychological term (a symptom in psychiatry), magnetic self-isolation as physical terms. Temporary self-isolation of people from society. The general concept of "self-isolation" (self-isolation) was introduced by the Japanese even before the onset of this epidemic (Gromenko et al., 2021). In Japanese, this is "Hikikomori", which means a rapid change of skills, traditions of life, personal behavior, voluntary spontaneous spontaneous alienation from society. Self-isolation is an operational action of the government of the country, announced to citizens in connection with the outbreak of the epidemic, and temporary home quarantine, regardless of age. The word self-isolation has been updated, which in the meaning of "restrictions on preventing the spread of infectious diseases" has become used as a synonym for the noun quarantine. However, if the concept of quarantine implies isolation of sick or related people, the noun self-isolation also includes new semantic components. This word can be defined as 'isolating oneself from other citizens in order to prevent illness'. The phrase self-isolation was not found in Google Trends. When we were only looking for the name of the isolating gesture, the following dynamic appeared.

The time of popularity of this word is February 21–27, 2021. A synonym for the words quarantine, self-isolation is lokdaun (the asset was used on June 28 and July 4) (**Figure 6**). This is a word from English. The meaning is strict isolation. Common in the American version of the English language.
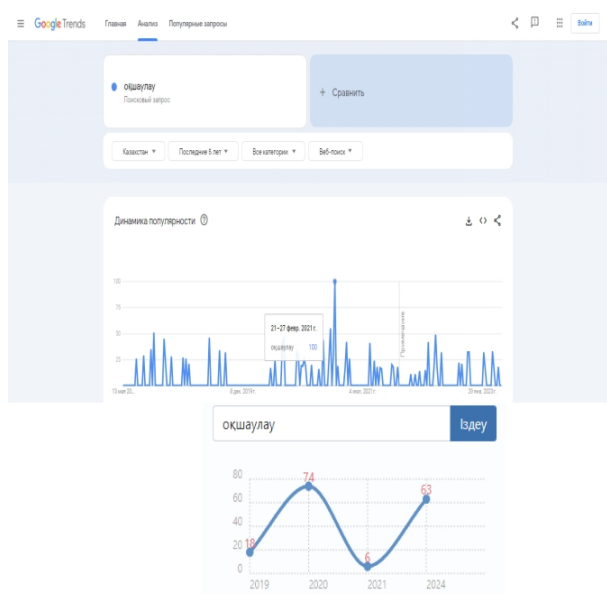

Figure 6. *Self-isolation.*

The noun *lockdown* means not only isolation of citizens (**Figure 7**), but also the whole range of measures to combat coronavirus (closure of public institutions and shops, social distancing, etc.).
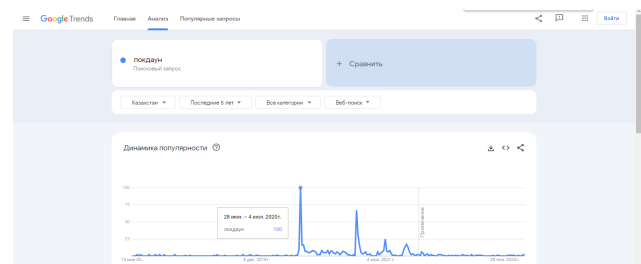

Figure 7. *Lockdown.*

Another difference between the noun *lockdown* is that it is used more often than the words quarantine, self-isolation, to describe restrictive measures introduced in some countries states.

3.4 Words related to education and learning

We also see the actualization of words related to education and learning. The pandemic period has become a difficult test for people and has changed almost all areas of human life and activity. Economics, medicine, science, public relations, changes have not bypassed the sphere of Education. The whole society has switched to self-isolation mode, and education has switched to an online format. For several decades, distance learning has been systematically introduced into the education system, many universities have taken over online platforms for lectures, seminars and discussions, gradually introducing tests, programs for completing control tasks and many attributes necessary for the educational process. However, with the onset of the pandemic, the transition from traditional education to distance education has become an urgent need, which has created a number of problems.

In this regard, it is important to consider the problems of distance learning identified in the course of the implementation of educational programs and find ways to solve them. In addition, a number of tasks that need to be taken into account in the research process come to the fore:

- analysis of the problem of rapid and large-scale transfer of the educational process to an online format by the education system;
- identify the main difficulties of students associated with

the transition to distance learning;

- consider the main difficulties of the transition to online for teachers;
- identify possible ways to solve the identified problems.

The problem of rapid and large-scale transition to the online format on the part of the general education system is as follows.

Distance learning is the interaction of a teacher with students not in face-to-face and live communication, but remotely through IT technologies. Also, for the implementation of online training, the following components are required: goals, methodology and technology, methodological materials, forms of organization and means of transmitting knowledge. If within the framework of Higher Education, the distance learning format was not new, then school education was not ready for the transition to such a new regime in a short time, when teachers had to "adapt" to the role of TV presenters, and schoolchildren, together with their parents, had to master the convenience of new educational internet platforms and resources. In this context, we can note the main difficulties that all participants in the educational process Face: teachers, schoolchildren, their parents, as well as power structures.

Distance learning plays a huge role in the educational process. Modern education is impossible without the internet, which has firmly entered our lives. Specialists of all professions must constantly improve their knowledge, skills and abilities in order to meet the high professional level required by the needs of our modern society. In the context of the pandemic and strict quarantine, vital issues related to education and education in general have arisen. This need led to the search for new methods of education and teaching technologies.

This phrase was used very often between March 29 and April 4, 2020 (**Figure 8**). At the same time, there was a time of mass online training in Kazakhstan.
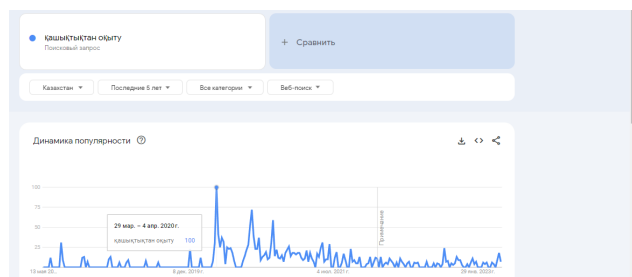

Figure 8. *Distance learning.*

In the spring months of 2020 (see: **Figure 9**), since the 70s of the XX century, the meaning of the sociological term *social distance*, denoting "hierarchical differences between social strata or classes in society," has radically changed. During the pandemic, this official term meant "a distance of at least 1.5 M, which citizens must observe to reduce coronavirus infection" (Gromenko et al., 2021, 61 p.). This word was used to insure people on many TV channels, in supermarkets, in the market, etc. In some large retail outlets, this word is still preserved.
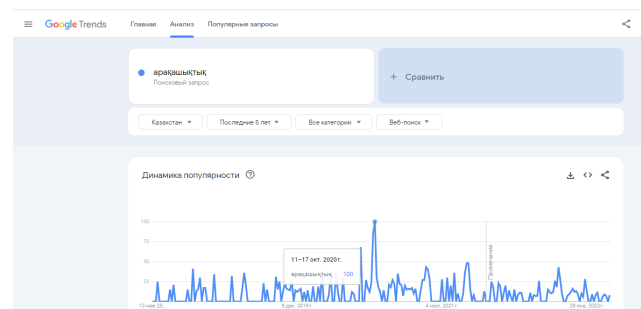

Figure 9. *Social distance.*

3.5 The colour symbolism of the pandemic

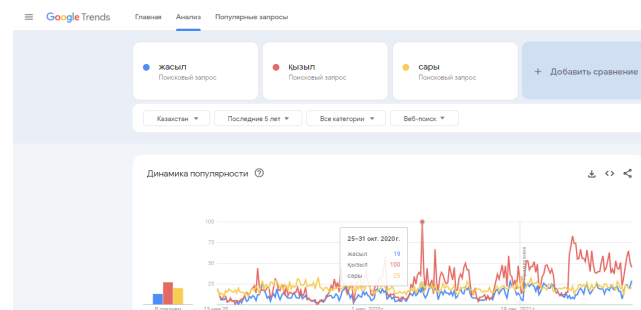During the COVID-19 pandemic, new phrases such as red zone, yellow zone, green zone appeared (**Figure 10**).

Figure 10. *Actualization of the words red, yellow, green.*

In order to indicate the level of danger zones during the pandemic, the colors red, yellow, green were used. Red represents a very dangerous area, yellow represents a moderately dangerous area, and green represents a safe area.

These concepts were novel for Kazakhs, who historically used color designations for spatial orientation. For example, in ancient Kazakh culture, colors symbolized different directions: white for the west, red for the south, black for the north, and blue for the east. The introduction of color-coded zones during the pandemic provided a new layer

of meaning to these traditional associations, helping people navigate the complex landscape of pandemic restrictions.

4. Conclusion

The COVID-19 pandemic has introduced a host of new words into the Kazakh language, while also updating existing terms to reflect the new reality. The key additions and updates include:

1. Коронавирус (koronavırýs) - Coronavirus (noun)
2. Ковид (kovid) - Covid (noun)
3. COVID-19 - (noun)
4. Пандемия (pandemiya) - Pandemic (noun)
5. Карантин (karantin) - Quarantine (noun)
6. Өзін-өзі оқшаулау (özın-özı oqshaulau) - Self-isolation (pronoun + gesture name)
7. Локдаун (lokdaun) - Lockdown (noun)
8. Әлеуметтік қашықтық (áleýmettik qashyqtyq) - Social distance (adjective + noun)
9. Қашықтықтан оқыту (qashyqtyqtan oqytý) - Distance learning (noun + gesture name)

Some of these words existed prior to the pandemic but have gained new significance and frequency of use. Others have only recently entered the language. For instance, terms like "coronavirus," "covid," "COVID-19," and "pandemic" were virtually unknown in everyday Kazakh vocabulary before March 2020. The sudden rise in their usage underscores the profound impact of the pandemic on language and communication.

This linguistic evolution reflects how societies adapt to new challenges by expanding and modifying their lexicons to include terms that help articulate the shared experiences and measures necessitated by the pandemic. The incorporation of these terms into the Kazakh language signifies a broader trend of global linguistic adaptation in response to the COVID-19 crisis.

In further research, it would be beneficial to include several Turkic languages for a deeper and more comprehensive study. The diversity of Turkic languages can provide a unique perspective on how the COVID-19 pandemic has impacted associative links and the perception of concepts across different cultural and linguistic contexts. This will enable a fuller understanding of how global events can alter language and cultural practices among Turkic-speaking communities.

Author Contributions

Zhazira Agabekova and Aigul Amirbekova reviewed the literature, Kalamkas Kalybekova and Dana Ospanova wrote the research methodology, also dealt with the draft version of the manuscript, so Zhazira Agabekova and Aigul Amirbekova, Bayan Abylayeva was engaged in data collection and visualization, Bolat Khassenov edited the article, corresponded with the editorial board of the journal.

Conflict of Interest

The authors stated that there are no conflicts of interest.

Data Availability Statement

Access to the research materials can be accessed through these links: https://trends.google.ru/trends/ and https://qazcorpora.kz/graphic.

Funding

Acknowledgments

References

Abkarian, M., Stone, H.A., 2020. Stretching and break-up of saliva filaments during speech: a route for pathogen aerosolization and its potential mitigation. Physical Review Fluids. 5.

Adilova, A., Balabekova, M., Tuite, Y., et al., 2023. Lexico-Semantic Change in the Kazakh Language of the COVID Era. Theory and Practice in Language Studies. 13(5), 1172–1181. https://doi.org/10.17507/-tpls.1305.10.

Al-Salman, S., Haider, A.S., 2021. COVID-19 trending

neologisms and word formation processes in English. Russian Journal of Linguistics. 25(1), 24–42.

Asadi, S., Wexler, A.S., Cappa, C.D., et al., 2020. Effect of voicing and articulation manner on aerosol particle emission during human speech. PLoS One. 15(1), e0227699.

Balabekova, M., Khan, N., Tuite, Y., et al., 2022. The Spread of English Language Elements During the COVID-19 Period. Theory and Practice in Language Studies. 12(12), 2583–2590.

Balabekova, M.A. 2022. Language convergence in mass media texts. Bulletin of the Kara- ganda university. Philology Series. 106(2), 13–19. https://doi.org/10.31489/2022ph2/13-19.

Daminov, I., 2021. Content analysis of the covid-19 pandemic in Kazakhstan English-speaking SMI: creating idyllic facades or reading facts? In Kazakhstan and COVID-19: media, culture, politics. Almaty: Izdatelstvo "DELUXE printing". pp. 146–167.

Danilenko, L., 2022. The Coronavirus epoch and its new vocabulary and phraseology in Slavic languages. Slavia Centralis. 1, 327–346.

Gromenko, E.S., Pavlova, A.S., Priemysheva, M.N., 2021. On covid-coronavirus processes in the Russian language in 2020. Studia Slavica. 1(65), 51–70.

Inouye, S., 2003. SARS transmission: language and droplet production. The Lancet. 362(9378), 170.

Inouye, S., Sugihara, Y., 2015. Measurement of puff strength during speaking: comparison of Japanese with English and Chinese languages. Journal of the Phonetic Society of Japan. 19(3), 43–50.

Klimchenko, A., 2021. "Pragmatic speciality and prosperity of the current state of discourse 42,500 tenge in the context of government briefings in the period of Pandemic Covid-19" In Kazakhstan and COVID-19: media, culture, politics. Almaty: Izdatelstvo "DELUXE printing". pp. 118–145.

Oluwateniola, K., Adebola, A., Akinniyi, A., et al., 2021. "Someone has been coronated" Nigerian English lexical innovations in the COVID-19 Pandemic. Cogent Arts & Humanities. 8(1), 8-19.

Onikbaeva, A.B., Davis, B.M., 2021. Neologisms that appeared in the lexicon of Kazakhstanis during the coronavirus pandemic. Tiltanym. 4, 60–71. https://doi.org/10.55491/2411-6076-2021-4-59-69.

Radbil, T.B., 2021. Active processes in the vocabulary and word formation of the Russian language of the coronavirus era: linguocognitive aspect. Scientific dialogue. 1, 63–79.

Tan, K.H., Woods, P., Azman, H., et al., 2020. COVID-19 insights and linguistic methods. 3 L The Southeast Asian Journal of English Language Studies. 26, 1–23.

Urazova, D., Serdaly, D., 2019. Kazakh humor during the pandemic (covid-19): memes of internet users in Kazakhstan. In Kazakhstan and COVID-19: media, culture, politics. Almaty: Izdatelstvo "DELUXE printing". pp. 336–357.

Voskresenskaya, N.A., Gulik, O.O., Malysheva, T.S., 2021. Semantic Typology Pandemic Vocabulary: Covid Neologisms in English, French and German. Nauchnyi Dialog. 11(3), 48–61.

Zholshayeva, M., Zhanabekova, A., 2021. Linguopragmatics of reflection of the COVID-19 pandemic in the Kazakh media text. In Kazakhstan and COVID-19: media, culture, politics. Almaty: Publishing house "DELUXE Printery". pp. 94–117.