

THE LINGUISTIC POTENTIAL OF THE COVID-19 LEXEME IN ENGLISH AND SLOVAK (A CORPUS-BASED STUDY)

MARTA LACKOVÁ

University of Žilina, Slovakia

marta.lackova@fhv.uniza.sk

Abstract

The aim of this article is to investigate the morphological and lexicosemantic traits of the COVID-19 lexeme as demonstrated in the framework of the English Web 2021 (enTenTen21) and the Slovak National Corpus (web-6.0). The exploration of the word-formation potential of the lexeme is carried out, too. A preliminary step of the research procedure is to recognize recurring patterns in which the studied lexical unit occurs so that we could inquiry into its functions in the given discourse. To achieve its goals, the combination of the lexical analysis and corpus analysis methods is employed. The findings of this study reveal that the lexeme COVID-19 entered all spheres of communication within both the studied languages. Furthermore, it has been adapted to their morphological and lexical systems and can access regular word-formation processes existing in them.

Key words: lexeme, COVID-19, morphological traits, lexicosemantic traits, word-formation, textual corpus

1. Introduction

The Covid-19 pandemic has caused unprecedented physical and mental health problems, life losses, and serious economic, political, and social consequences. It has also led to changes in the languages of the world and their structures. Both the negative and positive pandemic changes in people's lives resulted in the appearance of neologism associated with the Covid-19 pandemic and modifications of communication patterns (Ivanov et al., 2023).

Terminology from the field of medicine represents the main source of the Covid-19 neologisms. The terms have rapidly penetrated everyday communication owing to the mass media and social networks (Katermina, & Lipiridi, 2020). The processes of identifying and analysing these new lexemes are of high significance from several points of view – adding new vocabulary is important in relation to the synchronic

depiction of a language together with the description of sociolinguistic factors that shape the pandemic and post-pandemic eras.

The aim of this study is to describe the position of the lexeme COVID-19 in the English and Slovak languages in the span of 2019 – 2022 years as recorded in the textual corpora. Then, to understand different morphological and lexicosemantic possibilities of their functioning, which are direct manifestations of language vitality and become intriguing subjects of linguistic research.

2. Literature review

Research in the field of lexicology concentrating on the investigation of various aspects and features of the COVID-19 lexeme has been gaining an increasing popularity since it is associated with the intensive creation of new words during the COVID-19 pandemic together with their immense and rapid global spread thanks to the fact that people are nowadays digitally connected more than ever before (Fitria, 2021; Saleh, 2021).

Also, the presented study is related to previous explorations into various aspects of the COVID-19 lexeme. In her lexicological study, Samigoullina (2021) investigated models of word-building associated with the COVID-19 lexeme. The researcher considered substitution blending and overlap blending within the example of the COVID-19 vocabulary; she emphasizes that 63% of the studied lexemes were created by the substitution blending method. Moreover, Samigoullina (2021) recognized 11 groups of COVID-19 neologisms together with their semantic functions, features, and connotations.

The theoretical framework on which a study that focused on the phenomenon of new lexical units in order to examine the appearance of coinages during the COVID-19 outbreak was based on the three fundamental components of neologism – word formation, borrowing, and lexical deviation (Asif et al., 2021). Utilizing the procedures from the field of statistics, the study proves that there are also numerous shortened forms of the studied words utilized globally, e.g.: *corona*, *covid*, *rona* and *rone*. Their occurrence is connected specially with social media.

Regarding discourse aspects of the COVID-19 lexeme, Ivanov. A. et al. (2023) devoted their attention to linguistic potential of COVID-19 neologisms in the framework of socio-political register. They argue that the English language is the principal donor of neologisms semantically associated with COVID-19. In addition to this, the study categorizes these neologisms into four groups, while each of the categories is distinguished by a certain number of lexemes, the semantics, and functions of which are in concord with the general characteristics of the group.

Research in this specific branch of lexicology was carried out also by scholars who concentrated on the borrowing of neologisms of English origin with COVID-19 semantic relations by other European and world languages. As a way of example, Grzegza (2021a) observed that positive connotations which are created mirror that the coronavirus topic has a considerable impact on our way of living; it further shows people's lack of reliable information. According to Grzegza (2021b),

such lack arises if people access information containing lexically and conceptually confusing concepts. Haddad Haddad, A. & Montero-Martinez, S. (2020) add that the extension of the term worldwide may cause confusion, they advocate the co-existence of an Arabized neologism with the English loan to facilitate international communication.

While studying sociocultural background of derivate neologisms in the contemporary Russian and Czech languages, Samylicheva and Gazda (2020) prove that the COVID-19 vocabulary is based on occasional word-formation processes. Moreover, they argue that borrowings from the English language represent the primary source of new lexical units in Russian and Czech.

Moving to the stylistic features of these lexemes, Khalfan et al. (2020) divided COVID-19 neologisms into two main categories: scientific terminology and pop culture. Ivanov et al. (2023) partially agree with this opinion as medical terminology became an important part of the vocabulary of general public during the pandemic. However, they claim that the terms *Covid-19*, *coronavirus*, *covid*, and *corona* directly refer to the principal source of neologisms; in other words, these lexemes are categorized in line with semantic varieties that match with the statuses of people in the pandemic context.

Similarly, Atabekova et al (2022) strived to investigate verbal means that are utilized by English-speaking Twitter accounts to describe the pandemic with the focus on extralinguistic factors as primary catalysts for linguistic changes. Their results show that the lexeme COVID-19 is a productive ground for derivation into neologisms, abbreviations, substantive word combinations, and anthropomorphic metaphors.

Taking into consideration the corpus investigation of the COVID-19 lexeme, Gustilo et al. (2021) carried out a corpus study focusing on word-formation processes and meanings of lexemes related to the pandemic across varieties of English. For the needs of their research, they created a new corpus – the News on the Web English corpus where they identified and classified COVID-19 terms into new formations, blends, and revitalized compounds.

Furthermore, Cartier, E. et al. (2022) executed a long-term cross-linguistic newspaper corpus analysis of COVID-19 discourse and their revealed several key insights. First, there were found differences in the distribution of various competing denominations for the virus. Denominations relating to the disease caused by the virus (*covid* and *covid19*) are more popular within the English and French data; on the other hand, denominations relating to the virus itself (*corona*, *coronavirus*) prevail in German and Dutch. Secondly, they observed cross-linguistic stability in the way it is discussed despite this variation in the denomination of the disease. In all four languages, more factual categories like features, measures, and effects are complemented by semantic categories providing an evaluation or intensity of the disease.

3. Methodology

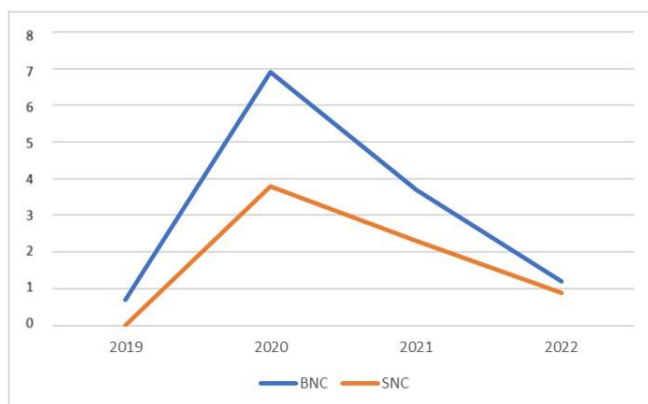
For the needs of the study, two methods will be employed: a method of corpus analysis involving a qualitative and quantitative investigation of the presence and role of the COVID-19 lexeme in the framework of textual corpora and a lexical semantics method.

The criterion for the selection and determination of the COVID-19 neologism in our study was primarily the functional point of view. In this respect, Orgoňová, Bohunická (2011) claim that a neologism should be considered as any enrichment of the lexical fund (qualitative or quantitative) that satisfies new communication needs, for example it reflects changes in the cultural context, in the development of the way of life, scientific and technical development with the emergence of new realities, the impact of globalization and the strengthening of multiculturalism, etc.

From the point of view of corpus linguistics methodology, our investigation concentrates on concordances, collocations, and word combinations in which the lexeme COVID-19 typically appears. The principal positions of the research interest are represented by the first lemma in the direction to the right and the first lemma in the direction to the left from the investigated lexeme. In other words, the aim is to recognize repeated patterns of its occurrence within the textual corpora and extralinguistic factors in the background of its functioning in the language.

We investigate the selected lexical unit within the monolingual synchronic annotated textual corpora English Web 2021 (enTenTen21) with the help of the text analysis software Sketch Engine and the Slovak National Corpus (web-6.0) applying the bonito.korpus.sk tool. In the English Web 2021 (enTenTen21) corpus, there are 61 585 997 113 tokens (52 268 286 493 words) are placed while the size of the (web-6.0) is 4 373 231 228 tokens (3 127 101 759 words). It is remarkable that the occurrence of the lexeme COVID-19 in the corpora is changing over the years in question (Figure 1).

Figure 1: The number of occurrences of the COVID-19 lexeme as recorded in English Web 2021 (enTenTen21) and in the Slovak National Corpus (web-6.0) in the year span 2019 – 2022 (calculated per thousands of words)



The relevance of the investigation has been strengthened through the introduction of interlingual connections between the semantic features of the COVID-19 lexeme in two typologically diverse languages. English is the dominant language while the Slovak language has been chosen for the needs of comparison. The description of the data will be supported by numerous examples and represented in the form of graphs and tables.

As the preliminary step of our research, we elaborated and compared the frequencies of the lexeme COVID-19 and other frequent lexemes from the field of medicine to highlight the significant position of the lexeme for the needs of the recent communication. In the English Web 2021 (enTenTen21) corpus, there are 1 521 405 hints recorded for the COVID-19 lexeme, with the representation of 32.47 hints per million tokens and 0.0037% percent per the corpus (Table 1). In the Slovak National Corpus (web-6.0), there are 95 365 hints recorded for the COVID-19 lexeme, with the representation of 7.28 hints per million tokens and 0.0013% percent per the corpus (Table 1).

Table 1: Number of concordances of the chosen lexemes and their frequency per million in English Web 2021 (enTenTen21)

lexeme	the number of hints	the number of hints per million tokens	the percentage % per the whole corpus
cancer	4 622 429	102.79	0.01
COVID-19	1 521 405	32.47	0.0037
flu	454 581	10.11	0.001
asthma	394 343	8.77	0.00088
heart attack	213 306	4.74	0.00047

Table 2: Number of concordances of the chosen lexemes and their frequency per million in English Web 2021 (enTenTen21)

lexeme	the number of hints	the number of hints per million tokens	the percentage % per the whole corpus
cancer	832 418	65.81	0.02
COVID-19	95 365	7.28	0.0013
flu	124 851	32.12	0.001
asthma	394 343	8.77	0.00072
heart attack	423 006	39.73	0.0057

The relatively high differentiated incidence of these lexical units prove that they contribute to the fulfilling of functional needs of communication; moreover, their position in the core of the lexical stratification grants the communication flow. The frequency belongs to the factors effecting the lexeme functioning from the perspective of lexicology. However, scholars understand frequency as a relative property; various corpora might provide a researcher with diverse frequency lists. According to Mahlberg (2005), isolated lexical units in frequency lists provide a researcher with information about the content of a text; yet their place in the list does not determine its semantics.

Subsequently, the methods of morphological and lexical semantic analysis will be applied to identify the most relevant grammatical and semantic categories associated with COVID-19 lexeme. This approach valid in lexical semantics makes it possible to consider the lexeme denoting COVID-19 as an example of a global neologism that was introduced into numerous world languages (Kupolati et al., 2021; Saleh, 2021).

4. Research results

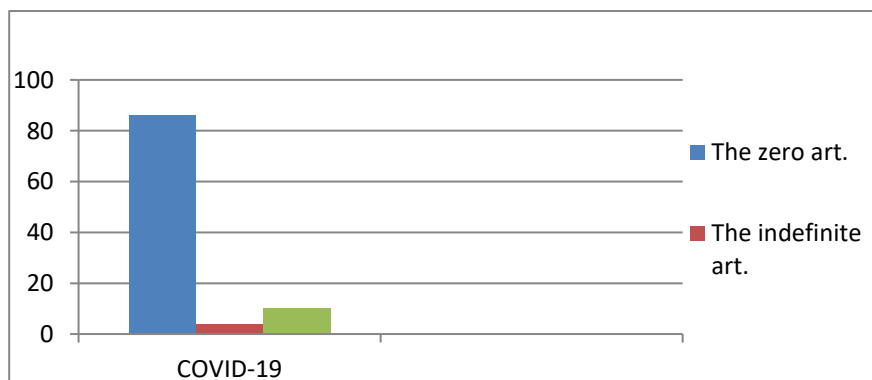
When studying the features of the COVID-19 lexeme within both the textual corpus, we must point out that the corpus substantiation displays its occurrence in a wide scope of structures. The following patterns connected with word classes are apparent: the investigated lexeme prevails in structures that contain determiners – further senses are added to it by various types of modifiers varying from adjectives to nominal and verbal prepositional phrases.

- 1) But, due to Covid 19 The Southern Alberta Art Gallery is thinking out of the box so the twenty-eighth annual ArtFrenzy will be an online auction of 44 artworks and 16 gift packages. (thelabeat.ca)

The utilization of the lexeme in question for the communication needs of English speakers within various registers is demonstrated by formal expressive means of the grammatical category of determination that are applicable with it, too. When discussing the frequencies of these determiners, it is apparent that the zero article is the most frequent in the position of the first lemma to the left. The definite or indefinite articles in these instances demonstrate their determinative effectiveness, and in this way they modify the COVID-19 lexeme according to the context needs. Figure 2 displays the percentage representation of the above-mentioned.

- 2) Sales have increased since the Covid 19 outbreak! (dealstream.com)

Figure 2. The application of the definite, indefinite, and zero articles with COVID-19 lexeme within English Web 2021 (enTenTen21) (expressed as a percentage)



Frequent collocations of the COVID-19 lexeme have been observed when it turns, because of the conversion process, into an adjective modifying another noun as a part of a noun phrase – we investigate the first lemma to the right from it (e.g.: *pandemic, crisis, outbreak, vaccine, cases, lockdown, restrictions, virus, infection, testing, patient*).

- 3) Due to COVID 19 restrictions, there are a limited number of people who can attend the service. (hogle.ca)

The collocations might be of twofold nature: either they copy notions already established in dictionaries or they have the capacity to display authentic word combinations. From the point of view of sociolinguistics, the latter are noteworthy when trying to understand humans, their behaviour, and institutions. Moreover, solely a limited number of these nouns possesses the potential to function in front of the studied lexeme. This possible double position illustrates their meaning realized in the communication processes between patients and health specialists.

- 4) The virus causes the respiratory disease COVID 19, manifesting in worst cases as pneumonia. (l-a-k-e.org)

As a rule, the negative influence of COVID-19 is implied with the help of linguistic means on various spheres of life. It can be expressed explicitly with noun combinations where the first element possesses a common sense of result or reason (e.g.: *the costs of covid, the outcome of covid-19 for the climate, the consequences of covid-19*). Indeed, the application of passive constructions where COVID-19 is introduced as the doer of the action conveys the impression of influence. Syntactically, the lexeme Covid-19 often takes the role of the subject of a clause which might result in the personification of the phenomena.

Adjectives preceding the lexeme COVID-19 in the English Web 2021 (enTenTen21) Corpus denote both subjective and objective qualities (e.g.: *severe, unprecedented, confirmed, asymptomatic*). It is evident that these adjectives cover the semantics of a wide scope of negative phenomena that arose during the pandemic; the exaggeration of the adverse qualities directly reflects the mental state of the language users. The global character of the pandemic is often specified thanks to the application of the given lexeme and adjectives, such as *world-wide, world, global* in the same context.

- 5) During the unprecedented COVID 19 outbreak, we are only able to answer telephone calls when staffing the warehouse for delivery loading. (sofabedgallery.co.uk)

When considering verbs that alter the semantics of the COVID-19 lexeme, the specimen gathered from the corpus entail that there have arisen more semantically relevant collocations than the ones that regular dictionaries display (e.g.: *to prevent, to confirm, to combat, to cause, to regard, to spread, to catch*). The verbs cover a wide semantic range of variations from positive to strongly negative connotations and they communicate procedural meanings within which the inspected lexeme is realized.

Moreover, they demonstrate the ability to introduce the idiomatic and metaphorical contexts for the existing problems – it is considered metaphorically as a threat or killer (e.g.: *the scars of covid-19*, *the ill wind of covid-19*, *Covid dries up a cash cow*, *the sting of covid-19*, *covid-19 spreads like wildfire*). In the given context, the disease is presented as a test of existing beliefs or rules: (e.g.: *to face a stern test from the covid-19*, *covid-19 is a challenge to the generation of politicians*) or a filter of the existing global problems. In this case, it is combined with the verb *to expose*.

- 6) The nationwide lockdown to combat COVID 19 has forced all of us to stay indoors. (jagranjosh.com)

Prepositions creating prepositional phrases with the COVID-19 lexeme are found frequently in the position of the first left lemma from it; semantically, they are defined as explanations of various external conditions associated with the disease, which are mostly of temporal or spatial nature. At the same time, the presentation of the COVID-19 image of might be followed by the expressions with the semantics of danger, risk, or fear (e.g.: *the threat posed by covid-19*).

- 7) What appears below are arrangements in force before Covid 19. (hastingsonlinetimes.co.uk)

As a subsequent step of our research, we identified the systematically predictable word-formations in relation to the COVID-19 lexeme, nominal compounds with this non-adapted term represent the most frequent neologisms in the research sample. Semantically they communicate the meanings connected to or caused by COVID-19. Two ways of writing of these compounds have been recorded: either together or with a hyphen (e.g.: *covid-positive*, *covid-infected*, *covid-stricken*, *covid-related*, *covid-afflicted*, *covid-battered*).

- 8) For months, Cuomo had faced questions about a decision made early in the pandemic to force long-term care homes to accept COVID-positive patients from hospitals. (cbc.ca)

Another group of these neologisms is presented by blends. We understand them as lexical units that are distinctly sociolinguistically motivated bearing the features of stylistic effectiveness, wit, and jocularly (e.g.: *covexit*, *covideo party*, *covidiot*, *covidivorce*, *covidpreneur*, *covidient*).

- 9) The two doctors had stepped out of their houses to buy fruits, and our covidiot accused them of spreading the virus. (dailyo.in)

Eventually, the lexeme COVID-19 is combined with prefixes of both domestic and foreign origin to indicate temporal relations: *pre-covid*, *post-covid*.

- 10) Finally, he will provide insights into the post-covid consumer: their values, priorities and needs and what it means for marketing in general and brand communications in particular. (mii.ie)

So far, the investigation of the lexeme COVID-19 in the Slovak textual corpus has not been a topic of interest; for this reason, we try to contribute to unveiling of this linguistically provoking issue. Our motivation to work within a corpus framework has been inspired by the fact that the linguistic material appears in a natural discourse. To analyse the linguistic behaviour of the lexeme in the Slovak National Corpus, the same procedure as for the English Web 2021 has been carried out.

The considerable influence of English on the contemporary Slovak language is reflected in the manner how easily the lexeme and its derivatives have been adopted by Slovak speakers although the lexeme with its features does not fit into the Slovak grammatical paradigms. When integrating into the system of Slovak nouns, it keeps the original categorical meaning; it means that the semantics of materiality copies syntactical and morphological traits of nouns.

The final procedure of the lexeme morphological adaptation is its incorporation into the valid morphological paradigm; it can acquire the double spelling *covid/kovid*, and grammatical inflections of Slovak singular nouns: *covid-e*, *covid-u*; *covid-om*:

- 11) Po prekonaní Covidu tak máte podľa vedcov minimálne na pol roka pokoj. (vysetrenie.zoznam.sk)

One way of inspecting the keyword COVID-19 is to study the nouns that modify it within textual corpus. Except for their persuasive functions, the nouns have the capacity to perform nominative and informative functions:

- 12) Cieľom realizovaného prieskumu bolo zmapovať vplyv šírenia covidu 19, ako aj kroky, ktoré prijímajú manažmenty v dôsledku tejto novej ekonomickej reality. (targetnews.sk)

The eight most recurrent nouns that are recorded in the closest surrounding of the lexeme COVID-19 are pandémie (pandemic), vírus (virus), ochorenie (disease), nákaza (infection), obdobie (period), infekcia (infection), príznaky (symptoms), priebeh (course). The utilization frequency of these nouns is changing over time, while in 2020 the noun liečba (treatment) prevailed in 35, 42 % of instances, in 2022 its occurrence decreased to 21, 55%. On the other hand, the noun prekonanie (overcoming) increased from 1, 78 % to 49, 12 % over the same period.

- 13) Neexistuje dôvod na to, aby sme obišli súkromné zdravotné poisťovne, ktoré sú rovnako postihnuté pandemiou covidu. (tyzden.sk)

The psycholinguistic power of the COVID-19 lexeme might become even stronger when it is applied as a modifying element of another noun. In this case, it does not undergo derivational processes typical of Slovak adjectives, the lexeme has not been fully assimilated into the system of the Slovak language; in fact, we deal with the process of conversion which is a peripheral phenomenon also in case of lexemes of English origin taking into consideration the superiority of Slovak inflectional elements:

- 14) Okres Prievidza je podľa COVID automatu aktuálne v druhom stupni varovania. (myhornanitra.sme.sk)

Hyponyms coexisting with the lexeme COVID-19 in the discourse (e.g.: *ochorenie, choroba* (*illness, disease*)) are semantically related to other health circumstances which emphasize its place in the system of illnesses:

- 15) Až do chvíle, kým túto nezrovnalosť úrad vlády napraví a v stredu dokument aktualizuje, majú kontrolné orgány rešpektovať potvrdenia o prekonaní ochorenia covidu, ktoré nie sú staršie ako 180 dní. (refresher.sk)

The expressive function of the adjectives that in the Slovak National Corpus modify the lexeme COVID-19 allow the reader to identify the traits that make the described situation unique. The category of the adjective *nový* (*new, novel*) has been identified as very relevant since it denotes the novelty of the phenomena. It is noteworthy that speakers of the Slovak language do not utilize possessive pronouns with the lexeme, this fact illustrates that they want to keep a certain distance from the disease.

- 16) Počet hospitalizovaných s potvrdeným COVIDom sa zdvojnásobil za menej než 2 týždne. (debata.pravda.sk)

Further instrumental, spatial, and temporal relations of the lexeme COVID-19 are communicated with the help of prepositional structures; the most persistent structures where a preposition appears together with a verb is *zomrieť na covid* (*to die of covid*), and with a noun is *test na covid* (*a covid test*).

- 17) V januári totiž na Slovensku na covid zomrelo viac ľudí, než počas celého roka 2020. (extraplus.sk)

Regarding the word-formation capacity of the lexeme COVID-19, the compounding (e.g.: *covid-saniarka* (*covid ambulance*), *covid-pacient* (*covid patient*)) is followed by derivation – the lexeme COVID-19 is productive of the adjective *kovidový*, *kovidovým*, *kovidového*, *kovidovými*.

- 18) Aktuálne šírenie kovidovej infekcie vyžaduje obozretnosť a zvýšené nároky na dezinfekciu. (hoteldubnaskala.sk)

The etymological and stylistic traits of the COVID-19 lexeme predestine its ability to collocate with units of diverse origin. We marginally also detect adjectival ordinary derivation by prefixation with the borrowed prefix *anti* (e.g.: *antikovidový*).

- 19) Antikovidový balíček obsahuje dva certifikované respirátory FFP2, štyri trojvrstvové ochranné rúška a balenie vitamínu D3. (myorava.sme.sk)

In addition to the aforementioned, the high productivity among new COVID-19 formations is shown by blends (e.g.: *covidiot/kovidiot*). These are a more recent type of formation within the Slavic languages. As Sicherl and Žele (2018) point out there should be realized two fundamental conditions to justify this type of

word-formation: the overlap must be meaningful, semantically recognizable, and stylistically effective. Moreover, the level of back formation of individual words is to be adapted to the pronunciation probabilities of the given language. Blends in Slovak are formed with the purpose to influence the receiver of the information and to achieve a certain stylistic effect (Sicherl & Žele, 2018).

5. Discussion

The examples above taken from both English Web 2021 (enTenTen21) and the Slovak National Corpus (web-6.0) demonstrate the usage of the COVID-19 lexeme and types of its reference for individual word classes and results of word-formation as observed in the concordances; here we would like to emphasize that the specimen of language which were made available by the concordance material do not directly demonstrate linguistic phenomena in their most general way. This conclusion is in concord with research results provided scholars working in the field of corpus linguistics (Halliday, et al., 2004, Mahlberg, 2005; Leláková & Toman, 2023).

This study argues that the behaviour of the COVID-19 lexeme within the British and Slovak corpora follow the principles behind the general structures valid within English and Slovak morphological and lexical systems. This finding also supports the investigation results presented by Akut (2020), Gustilo et al. (2021), and Pollak et al. (2023) and the theoretical knowledge on the issue provided by the prescriptive grammars.

Further, the corpus evidence in both the investigated languages proves that the lexeme COVID-19 has penetrated common usage and colloquial structures, and thus the lexeme is a subject of determinologization. A similar process has been observed in the framework of the Slovenian language (Pollak et al., 2023) where the emergence of neologisms and occasionalisms connected with COVID-19 vocabulary has been subjected to an accelerated Slovenization process.

All the same, Biró et al. (2021) conclude that COVID-19 coinages and lexicalization processes connected with them are not solely recorded in the sphere of specialized language; ordinary people had to interpret the new vocabulary denoting the new reality, too. This research result is in line with ideas of our investigation. In both the studied corpora it is evident that speakers of the English and Slovak languages incorporated the lexeme COVID-19 with great ease into actual communication acts.

Consequently, our analysis also confirms the findings of Katermina & Lipiridi (2020) while the lexeme COVID-19 is widely represented in various genres in various ways, including novel word forms and word combinations. In this frame of reference, compounds, derivatives, and blends have been registered as the most relevant categories capable of acquiring further connotations in distinct contexts.

6. Conclusion

We can conclude that covid-19 pandemic has affected the language to a great extent, these changes are reflected in the textual corpora, too. Both the morphological and lexicosemantic features of the lexeme COVID-19 as manifested within English Web 2021 (enTenTen21) and Slovak National Corpus (web-6.0) are in line with valid principles functioning in the contemporary English and Slovak languages; as such, they reflect the demands of their users which naturally comprise various constituents outside the languages themselves. When discussing the semantics, it is not possible to describe the investigated lexeme in full detail, the representation as presented in this study tends to show that the lexeme shares elements of meaning with diverse terms applicable in the field of medicine and in general usage. The findings support the claim that these semantic elements defining nominal functions in the discourse serve as part of the semantic prosody of functional group samples (Mahlberg, 2005).

Based on the observations of the research sample, we can conclude that lexemes semantically related to Covid-19 are created mostly through the blending and compounding word-formation processes in both the studied languages; this result is also shared by other scientists who have scrutinized the neologisms (Katermina & Lipiridi, 2020). The lack of studies devoted to the exploration of these neologisms in the framework of Slovak textual corpora indicates further investigations is needed in the sphere.

The findings of this study can be utilized for further elaboration of ideas connected with COVID-19 vocabulary in diverse languages since vocabulary of contemporary languages is constantly enriched and it requires a special consideration in the framework of its linguistic potential. Moreover, the results of the study might be useful for supplementing the practical base in the spheres of semantics and social linguistics since the factors of meaning and social conditions must be taken into account when studying the behaviour of selected lexemes in the corpus environment.

The results of the presented study reveal that the morphological and lexicosemantic features of the term COVID-19 reflect to a high degree extralinguistic factors of individual speech acts. For that reason, the outcomes from our corpus analysis might have further practical relevance for health professionals and educators working in the field of health care.

Possible limitations of the paper are perceivable when reflecting on the amount of linguistic material to be elaborated here. It needs to be noted that this study discusses only evidence available in textual corpora which are, on one hand, constantly enriched by new entries, on the other hand, they do not cover the language in its full range.

References

- Abdul-Mageed, M., Elmadany, A. R., Nagoudi, E. M. B., Pabbi, D., Verma, K., & Lin, R. 2021. Mega-COV: A Billion-Scale Dataset of 100+ Languages for COVID-19. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, 3402–20, Online. Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.eacl-main.298>
- Akut, K. B. 2020. Morphological Analysis of the Neologisms during the COVID-19 Pandemic. *International Journal of English Language Studies*, 2(3), pp. 1–7. <https://doi.org/10.32996/ijels.2020.2.3.11>
- Asif M., Zhiyong D., Iram A., & Nisar M. 2021. Linguistic Analysis of Neologism Related to Coronavirus (COVID-19). *Social Sciences & Humanities Open*, 4(1), 100201. <https://doi.org/10.1016/j.ssaho.2021.100201>
- Aslam, F., Awan, T. M., Syed, J. H., Kashif, A., & Parveen, M. 2020. Sentiments and emotions evoked by news headlines of coronavirus disease (COVID-19) outbreak. *Humanities & Social Sciences Communication* 7, 23. <https://doi.org/10.1057/s41599-020-0523-3>
- Atabekova, A., Lutskovskaia, L., & Kalashnikova, E. 2022. Axiology of Covid-19 as a Linguistic Phenomenon. *Journal of Information Science*. <https://doi.org/10.1177/01655515221091542>
- Biró, E., Kovács, G. & Nagy, K. 2021. Change(s) in Vocabulary (ies) – Hungarian and Romanian Lexical Phenomena During COVID. *Acta Universitatis Sapientiae Philologica*. 13(3), 62–78. <https://doi.org/10.2478/ausp-2021-0028>
- Cartier, E., Onysko, A., Winter-Froemel, E., Zenner, E., Andersen, G., Hilberink-Schulpen, B., Nederstigt, U., Peterson, E., & van Meurs, F. 2022. Linguistic Repercussions of COVID-19: A Corpus Study on Four Languages. *Open Linguistics*, 8(1). 751–766. <https://doi.org/10.1515/opli-2022-0222>
- Fitria, T.N. 2021. Word Formation Process of Terms in Covid-19 Pandemic. *Leksika*, 15 (1). <https://doi.org/10.30595/lks.v15i1.9248>
- Grzega, J. 2021a. From Coronaparty to Coronapaniek: Cross-linguistic critique of COVID-related contributions to Wiktionary. *Language@ Internet*, 19(1), 1.
- Grzega, J. 2021b. The linguistic impact of coronavirus on online service offers in Spain as well as France, Germany, Italy, the Netherlands, and the UK. *Linguistics and Culture Review*, 5(1), 31–42. <https://doi.org/10.37028/lingcure.v5n1.82>
- Gustilo, L., Pura, C. M., & Biermeier T. 2021. Coronalexicon: Meanings and Word-Formation Processes of Pandemic-Related Lexemes across English Varieties. *International Journal of English Language Studies*, 27(4), pp. 1–15. <https://doi.org/10.17576/3L-2021-2704-01>
- Haddad Haddad, A. & Montero-Martinez, S. 2020. COVID-19: a Metaphor-based Neologism and its Translation into Arabic *JCOM* 19(05), A01. <https://doi.org/10.22323/2.19050201>
- Halliday, M. A. K. et al. 2004. *Lexicology and Corpus Linguistics: An Introduction*. London: Continuum.
- Hu, Z., Yang, Z., Li, Q., & Zhang, A. 2020. The COVID-19 Infodemic: Infodemiology study Analyzing Stigmatizing Search terms. *Journal of Medical Internet Research* 22(11). <https://doi.org/10.2196/22639>
- Ivanov, A., Nikonova, Z., Frolova, N., & Muratbayeva, I. 2023. Linguistic Potential of COVID-19 Neologisms in the Metaphoric Language of Socio-political Discourse. <https://doi.org/10.1177/20570473231186475>
- Katermina, V. V., & Lipiridi, S. C. 2020. Reflection on the Social and Psychological Consequences of the Coronavirus Pandemic in the New Vocabulary of the Nonprofessional English Language Medical Discourse. *Proceedings of the research technologies of pandemic coronavirus impact* (pp. 44–49). <https://www.atlantis-press.com/proceedings/rtcov-20/125945666> <https://doi.org/10.2991/assehr.k.201105.009>

- Khalfan M., Batool H., & Shehzad W. 2020. COVID-19 Neologisms and their Social Use: Analysis from the perspective of Linguistic Relativism. *Linguistics and Literature Review*, 6(2), 117–129. <https://doi.org/10.32350/llr.v6i2.959>
- Kupolati O., Adebileje A., & Adeleke A. 2021. ‘Someone has been Coronated’ Nigerian English lexical Innovations in the COVID-19 Pandemic. *Cogent Arts & Humanities*, 8(1), 1947559. <https://doi.org/10.1080/23311983.2021.1947559>
- Leláková, E., & Toman, M. 2023. A Corpus Based Study of Commas Use in EFL Written Performance. *International Journal of Instruction*, 16 (1), 1-20. <https://doi.org/10.29333/iji.2023.1611a>
- Mahlberg, M. 2005. *English General Nouns. A Corpus Theoretical Approach*. Amsterdam/Liverpool: John Benjamins. <https://doi.org/10.1075/scl.20>
- Orgoňová, O., & Bohunická, A. 2011. *Lexikológia slovenčiny. Učebné texty a cvičenia*. Bratislava: Univerzita Komenského.
- Pollak, S., Voršič, I., Kern, B., & Ulčar, M. 2023. Novel Slovenian Vocabulary from the Perspective of Naming Possibilities and Word-Formation. *Electronic lexicography in the 21st century (eLex 2023): Invisible Lexicography. Proceedings of the eLex 2023 conference*, pp. 429-448, Brno, 27–29 June 2023. Brno: Lexical Computing CZ s.r.o.
- Roig-Marín, A. 2021. English-based Coroneologisms. a Short Survey of our Covid-19-related Vocabulary. *English Today* 37(4), 193–5. <https://doi.org/10.1017/S0266078420000255>
- Saleh A. S. 2021. COVID-19 Trending Neologisms and Word Formation Processes in English. *Russian Journal of Linguistics*, 25(1), 24–42. <https://doi.org/10.22363/2687-0088-2021-25-1-24-42>
- Samigoullina A. D. 2021. English Coroneologisms: Function and Semantics. *International Scientific Research Journal*, 3-3(105), 166–170. <https://doi.org/10.23670/IRJ.2021.105.3.092>
- Samylicheva N., & Gazda J. 2020. Derivative Neologisms as Sociocultural Dominants in the Russian and Czech Languages of the Modern Period. *SHS Web of Conferences*, 88, 01022. <https://doi.org/10.1051/shsconf/20208801022>
- Sicherl, E., & Žele, A. 2018. Prekrivanje v slovenščini z vidika vpliva angleškega jezika. *Jezik in slovstvo*, 63(1), pp. 75–88. <https://doi.org/10.4312/jis.63.1.75-88>
- Tan, K. H., Woods, P., Azman, H., Abdulah, I. H., Hashim, R. Z., H. Rahim, A., et al. 2020. Covid-19 Insights and Linguistic Methods. 3L: language, Linguistics, literature. *The Southeast Asian Journal of English Language Studies* 26(2), 1–23. <https://doi.org/10.17576/3L-2020-2602-01>
- Vargová, D. 2021. Aktuálne zmeny v slovnej zásobe anglického a slovenského jazyka – jazyk pandémie. *Lingua et vita*, 10(19): 93-105.
- Zhang, J., & Li, J. Eds. 2020. Linguistic Diversity in a Time of Crisis: Language Challenges of the COVID-19 Pandemic. *Special issue of Multilingua* 39(5). <https://doi.org/10.1515/multi-2020-0136>

Sources

- English Web 2021 (enTenTen21). [Online] Available from: <https://app.sketchengine.eu/>. [Accessed: 11th February 2023]
- Slovak National Corpus (web-6.0). [Online] Available from: <https://bonito.korpus.sk/>. [Accessed: 11th February 2023]