

THE LANGUAGE OF COVID-19

PhD Can. Zana Prekazi

University of Prishtina “Hasan Prishtina”, Faculty of Philology, Kosovo

Abstract

This paper examines the linguistic changes that have occurred since the emergence of the coronavirus at the end of 2019. Specifically, it focuses on the comparison of the vocabulary prior and during the pandemic to track the impact that this crisis is having on the English language. In order to carry out this research, a list of words taken from the April update of the Oxford English Dictionary has been extracted and analyzed. Two corpora, Corpus of Contemporary American English and The Coronavirus Corpus, have been used to put the words in context and analyze the frequency of the words used and their new meanings as a result of the pandemic. Through this analysis, we will be able to see trends of language change that reflect our knowledge for the virus during different periods of the year.

Keywords: Sociolinguistics, linguistic development, Covid-19, pandemic vocabulary

1. Introduction

According to the well-known model introduced by the Russian-American linguist, Roman Jakobson (1960), there are six main functions of language: referential (or informational), emotive, conative, phatic, and metalingual. Of all these functions, the informational one is most important (Leech, 1974) because it depends on truth and value: it relates to the thing “spoken of” (Jakobson, 1960, p. 355). Covid-19 has become the new reality for the entire world and as a result a new language is needed to describe this new phenomenon. Due to the emergence of the coronavirus at the end of 2019, significant linguistic developments have been taking place. Such developments are occurring so rapidly that linguists are taking all measures possible to record and track the impact that this crisis is having on the English language as it happens. According to Robert Lawson, a sociolinguist at Birmingham City University, there are different factors that may have contributed to the unprecedented speed of these linguistic changes. The most common ones are the spread of the virus, its dominance in the media, and interconnectivity. Although it first emerged in China, the virus soon spread all over the world, with now almost one hundred and twenty million people infected, one year after its emergence. Another factor is the life-threatening effect it has on people’s lives. With more than 2.6 million deaths as a result of the virus, this pandemic has become the most talked about topic in 2020. One other aspect that urged changes in the language was that Covid-19 was a novel virus. There was no prior knowledge about it, no descriptions and discussions about it as it had not existed before 2019. This paper aims to research the change of meaning in the English language due to the new reality of the coronavirus and how the meaning has changed to adapt to this new reality.

2. Methodology

The list of words selected to be analyzed in this study has been extracted from the April update of the Oxford English Dictionary (OED). To put those words in context, we used the Corpus of Contemporary American English (COCA) and the Coronavirus Corpus. COCA provides record of the English language in different genres and during different periods of time (before the pandemic). The Coronavirus Corpus, on the other hand, records the social, cultural, and economic impact of the coronavirus in 2020. It shows what people are saying in online newspapers and magazines in 20 English-speaking countries. The corpus released in May 2020 is currently about 692 million words in size. These two corpora will help us compare the vocabulary prior to the pandemic and now, one year on, tracking the linguistic trends and change as it happens. In order to look at particular nuances or senses, we look at the contexts in which a word is used. Given that these two corpora differ in size, we base our statistical framework in the contingency table discussed in Kilgarriff (2001), reprinted below:

| | <i>X</i> | <i>Y</i> | |
|--------------|--------------|--------------|--------------------------|
| <i>w</i> | <i>a</i> | <i>b</i> | <i>a + b</i> |
| not <i>w</i> | <i>c</i> | <i>d</i> | <i>c + d</i> |
| | <i>a + c</i> | <i>b + d</i> | <i>a + b + c + d = N</i> |

Table 1. Basic contingency table. For word *w* in texts *X* and *Y*, there are *a* occurrences of *w* in text *X* (which contains *a + c* words) and *b* in *Y* (which has *b + d* words).¹ *X* and *Y* in our case represent COCA and Coronavirus Copus; *w* represents the word of interest.

3. Literature Review

The impact of Covid-19 is an ongoing process. As we learn more and more about the pandemic, our vocabulary changes to reflect that knowledge. By analyzing the corpora, we can see those trends of language change that reflect our knowledge for the virus during different periods of the year. Considering this new reality, linguists have set up teams to record the covid language. In this study we will make use of the so-called *The Coronavirus Corpus*, which comprises of more than 40 million tokens of English language from articles related to Covid-19.

Oxford English Dictionary, on the other hand, seeing the great impact that Covid-19 is having on the English language, published two special updates dedicated solely to covid language, which fall outside their usual quarterly publication cycle. In their April update, the new word entries are mostly related to the definition of the disease like *Covid-19*, *infodemic*, and words describing how people can protect themselves from the virus, such as: *self-isolate* and *self-quarantine* (and their inflected forms), *social distancing*, *elbow bump*, etc. In their July update, the language is shifted more towards social and economic impacts, with most frequent words being *PPE*, *lockdown*, *pandemic* (April); *reopen*, *lockdown*, *Covid-19* (May); and *defund*, *Juneteenth*, *brutality* (June)².

Even within the two OED updates there are differences with regard to the same expressions. For example, in their April update, Covid-19 is defined as “an acute respiratory illness”, but the editors chose to change that definition in their July update by adding other effects of the virus, such as “fever and cough”, “blood coagulation abnormalities”, and “death”. According to editors, this change in definitions reflects the new information about the disease (Trish 2020).

4. Findings and Discussion

The following list of words has been extracted from the April update of OED. Below are analyzed their frequency and meaning change during 2020.

| |
|--------------------------|
| Self-isolation |
| Infodemic |
| Shelter-in-place |
| Social-distancing |
| Elbow bump |
| WFH |

Table 2. Coronavirus words added to OED in 2020

4.1. Self-isolation

¹ Table and notes taken from Kilgarriff (2001: 234)

² The top words of July prove that Black Lives Matter protests were another major topic during that month.

Firstly, in the 1800s *self-isolation* used to mean when countries chose to isolate themselves politically from the rest, now self-isolation means “self-imposed isolation to prevent catching or transmitting an infectious disease”³ This is supported by the difference in use of *self-isolation* in 2016 (cf. [1]) and in 2020 (cf. [2]):

- [1] Yayla said the irresponsible charge is part of Turkey's *self-isolation* from the West in the wake of the coup attempt.⁴
- [2] The patient lives in Toronto with his wife. So far, she is not displaying any symptoms, but has been put in *self-isolation* and is being monitored.⁵

Secondly, in addition to the meaning change, our data proves that there is also a difference in frequency. Table 3 below illustrates the significant statistical change in terms of instances per million: *self-isolation* is used 251 times per million tokens in the Coronavirus Corpus and only 0.04 times per million in COCA.

| | The Coronavirus Corpus | COCA |
|--------------------|------------------------|------|
| Frequency | 59727 | 42 |
| Words (M) | 237 | 993 |
| Per million | 251.32 | 0.04 |

Table 3. Distribution of inflected forms of *self-isolation* in both corpora

This high use of *self-isolation* during the pandemic raised the need to use it in all its forms. Consequently, there are 9 inflected forms of *self-isolation* in the Coronavirus Corpus (*self-isolation*, *self-isolate*, *self-isolating*, *self-isolated*, *self-isolates*, *self-isolations*, *self-isolator*, *self-isolators*, and *self-isolaters*) compared to only 4 forms in COCA (*self-isolation*, *self-isolating*, *self-isolated*, *self-isolate*). In addition, the last three forms were not part of English dictionaries due to their low frequency of use before 2020. Now, as the chart below illustrates, these forms have surged in frequency and as a result lexicographers added them to the English dictionary in April 2020.

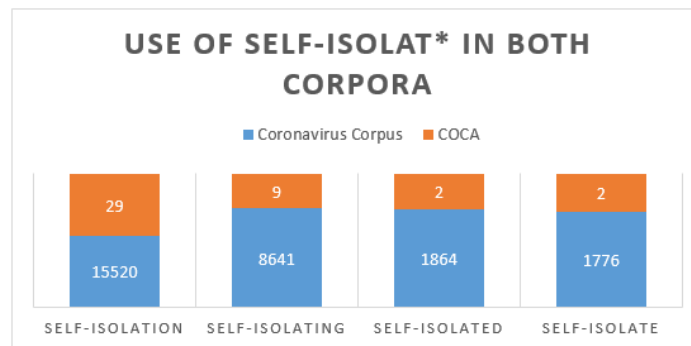


Chart 1. Four most frequent inflected forms of *self-isolation* in the the Coronavirus Corpus and COCA

4.2. Infodemic

Infodemic, a portmanteau word from *information* and *epidemic*, was so rarely used that we find no instances of it prior to the pandemic (although it was coined in 2003 for the SARS epidemic). Now it is being used to describe the current proliferation of news around coronavirus. Its peak of use was mid-February. Out of 100 articles related to Covid-19, published during February and March 2020, in 77 of them *infodemic* is enclosed in quotation marks (cf. [3]), and in 23 of them in no quotation marks (cf.[4]). Regardless of the presence of quotation marks or lack thereof, in half of the occurrences the word *infodemic* was accompanied by a description of its definition, as the below examples prove:

³ Definition taken from OED.

⁴ MAG: Fox News, 2016 (16-07-25) Turkish newspaper accuses US general of planning failed coup, COCA

⁵ CA: CTV News, 2020 (20-01-25), Canada's first 'presumptive positive' case of coronavirus ..., The Coronavirus Corpus

- [3] However, she raised alarm about what she described as an “*infodemic*” -- an epidemic of misinformation on the current outbreak.⁶
- [4] ... the World Health Organization (WHO) had to approach the social media and internet companies in Silicon Valley in the USA on February 13 for help in controlling the -- *infodemic*, the term given by WHO for the flood of fake news on COVID 19 epidemic on internet and social media.⁷

4.3. Shelter-in-place

Shelter-in-place is a protocol instructing people to find a place of safety in the location they are occupying until the all clear is sounded in the event of a nuclear or terrorist attack⁸. Now this expression has the meaning of staying indoors to protect themselves and others from coronavirus. The pair of examples below illustrate these different meanings of *shelter-in-place*:

- [5] Areas extending 1 mile from the fire have been declared a *shelter-in-place* zone, meaning windows should be closed, outdoor activities stopped, and air circulation systems shut down.⁹
- [6] Miguel Salas was arrested on suspicion of selling and transporting methamphetamine. Salas's narcotics delivery business does not fall under the essential business category. Salas will now *shelter-in-place* in jail.¹⁰

Before the coronavirus, *shelter-in-place* was used only 0.01 times per million, as Table [4] proves.

| | The Coronavirus Corpus | COCA |
|-------------|------------------------|------|
| Frequency | 7415 | 14 |
| Words (M) | 237 | 993 |
| Per million | 31.2 | 0.01 |

Table 4. Distribution of inflected forms of *self-isolation* in both corpora

4.4. Social distancing

Social distancing, first used in 1957, was originally thought as the action or practice of maintaining a degree of remoteness or emotional separation from another person or social group¹¹. Now the term has physical connotation, namely the distance we need to keep with others to avoid infection. This is clear from the following examples:

- [7] Arrange for social interaction during virtual and in-person meetings so that new members of the community become familiar with others and *social distance* overall is kept to a minimum. REDUCING VIRTUAL DISTANCE THROUGH CO-ACTIVATING LEADERS PHYSICAL DISTANCE¹²

As seen from example [7], before the covid emerged, social distance implied the following definition of Merriam-Webster dictionary:

the degree of acceptance or rejection of social interaction between individuals and especially those belonging to different social groups (such as those based on race, ethnicity, class, or gender)

⁶ IN: Down to Earth (20-02-05), Novel coronavirus outbreak not a pandemic: WHO, The Coronavirus Corpus

⁷ IN: Deviscourse (20-02-14), Handling fake news Infodemic in time of Coronavirus epidemic, The Coronavirus Corpus.

⁸ Definition taken from OED.

⁹ NEWS: Oregonlive, 2018 (18-03-12), Black column of smoke rising from NE Portland fire, COCA

¹⁰ US: Sfchronicle (20-03-03), Coronavirus live updates: six dead in Washington, the Coronavirus Corpus

¹¹ Definition taken from OED.

¹² ACAD: Mechanical Engineering Nov2009, Vol. 131 Issue 11, p30-34, 5p leading the DISPERSED WORKFORCE. Reilly, Richard R. Lojeski, Karen Sobel, COCA

As example [8] proves, in its new meaning *social distance* is equivalent with physical distance. This is evident from the modifier *two-meter*:

- [8] While the World Health Organization (WHO) has advised that it isn't necessary for people to wear masks unless they have developed flu symptoms, population density in Hong Kong makes it very difficult to maintain the recommended two-meter *social distance* between people, and in spaces such as apartment building lifts, social distance is reduced to nearly zero.¹³

As for the frequency, before the coronavirus, *social-distancing* was used only 0.08 times per million, now it is used 900.81 times per million, as Table [5] illustrates.

| | The Coronavirus Corpus | COCA |
|-------------|------------------------|------|
| Frequency | 214079 | 76 |
| Words (M) | 237 | 993 |
| Per million | 900.81 | 0.08 |

Table 5. Distribution of inflected forms of *social-distancing* in both corpora

4.5. Elbow bump

Elbow bump was in its earliest manifestation (1981) a way of conveying celebratory pleasure to a teammate, rather than a means of avoiding hand-touching when greeting a friend, colleague, or stranger (Paton, 2020). Since we could not find any instance of *elbow-bump* before the coronavirus, example [9] is taken from Miami News, as cited in the OED. In 2020, on the other hand, *elbow-bump* is not used as a celebratory pleasure, but to reduce the risk of spreading or catching an infectious disease, as example [10] proves.

- [9] Gene Banks of the Duke basketball team speaks of the hand slap, high five, elbow bump and other varieties of shakin'.¹⁴
- [10] Inslee said he's doing the *elbow bump* with people instead of shaking hands to prevent the spread of germs, and that his visit to the store was to encourage people to keep patronizing businesses during the COVID-19 Coronavirus outbreak.¹⁵

WFH

At the beginning of the pandemic, OED notes, terms regarding the disease and definitions are used. *COVID-19*, *Covid-19*, etc. abbreviations emerge later. There is no significant use of the abbreviation C-19 prior to July 2020. This term is now used mostly as a modifier.

In *Elements of Style*, Strunk and White (2009) suggest to "start your article by writing out names in full, and then, later, when your readers have got their bearings, to shorten them" (p. 76). *Covid-19* is itself a shortening of *coronavirus disease 2019*. According to OED it is the only new word added to the dictionary due to the coronavirus disease. But for the rest of abbreviations used during the pandemic, based on corpus data, abbreviations emerge in a later period, confirming that people got familiar with the words, and abbreviations were used to save space and time. The abbreviation *WFH* (Work From Home) is used 1703 times in coronavirus articles and only 19 times prior to the pandemic.

¹³ HK: Hongkongfp (20-02-14) Semi-quarantined: Living with the coronavirus outbreak in ..., The Coronavirus Corpus

¹⁴ Miami News 13 Feb. (Home ed.) c1/2

¹⁵ US: Columbian (20-03-05), School district north of Seattle closes over virus ... The Coronavirus Corpus

| SECTION | ALL | 20-01-01 | 20-02-01 | 20-02-11 | 20-02-21 | 20-03-01 | 20-03-11 | 20-03-21 | 20-04-01 | 20-04-11 | 20-04-21 | 20-05-01 | 20-05-11 | 20-05-21 | 20-06-01 | 20-06-11 | 20-06-21 |
|------------------------------|------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| FREQ | 1992 | 0 | 0 | 2 | 0 | 16 | 93 | 182 | 201 | 227 | 161 | 77 | 173 | 145 | 61 | 60 | 63 |
| WORDS (M) | 237 | 7.3 | 4.8 | 4.0 | 5.7 | 17.6 | 26.8 | 55.5 | 38.4 | 35.8 | 33.8 | 31.3 | 30.5 | 36.1 | 29.3 | 27.1 | 26.8 |
| PER MIL | 8.38 | 0.00 | 0.00 | 0.50 | 0.00 | 0.91 | 3.47 | 3.28 | 5.23 | 6.35 | 4.76 | 2.46 | 5.67 | 4.02 | 2.08 | 2.21 | 2.35 |
| SEE ALL SUB-SECTIONS AT ONCE | | | | | | | | | | | | | | | | | |

Table 6. WFH in the Coronavirus Corpus

An interesting fact is that *WFH* is now used also as a modifier, the most frequent examples including: *WFH model*, *WFH environment*, *WFH policy*, *WFH diaries*, and *WFH employees*.

The reason for the drop of using Work from Home or WFH can be attributed to the fact that after May it became the norm.

Conclusion

Covid-19, just as any other social change, is impacting language in general and English language in particular. Linguists have seen this pandemic as a great opportunity to record relevant terms used for its description and discussion as they emerge and change. In this study we have made use of the list of new words added to the Oxford English Dictionary in April 2020, and the hundreds of articles related to the coronavirus, collected in the Coronavirus Corpus in order to study the impact that Covid-19 is having on people's language. As our study proves, the Covid-19 impact was reflected in the English language in two main ways: meaning change and frequency surge. Such addition of new meanings to the existing words was done to specifically describe the coronavirus situation, whereas due to the spike and surge of words, the greater use of certain words led to additions in different parts of speech thereof.

References

- Davies, Mark. (2008-) *The Coronavirus Corpus*. Available online at <https://www.english-corpora.org/corona/>
- Davies, Mark. (2008-) *The Corpus of Contemporary American English (COCA)*. Available online at <https://www.english-corpora.org/coca/>.
- Elbow-bump. (2020). *Cambridge English Dictionary*. Retrieved from <https://dictionary.cambridge.org/dictionary/english/elbow-bump>
- Jakobson, R. (1960). Linguistics and Poetics, in T. Sebeok, ed., *Style in Language*, 350-377. Cambridge, MA: M.I.T. Press.
- Kilgarriff, A. (2001). Comparing corpora. *International Journal of Corpus Linguistics*, 6(1), 97–133.
- Leech, G. (1974). *Semantics*. Harmondsworth: Penguin books.
- Merriam-Webster. (n.d.). Social distance. In Merriam-Webster.com dictionary. Retrieved November 24, 2020, from <https://www.merriam-webster.com/dictionary/social%20distance>
- Oxford English Dictionary (2020, April). *Updates to OED*. Oxford University Press. <https://public.oed.com/updates/new-words-list-april-2020/>
- Paton, B. (2020, April 9). Social change and linguistic change: the language of Covid-19. [Blog post]. Retrieved from <https://public.oed.com/blog/the-language-of-covid-19/>
- Ro, Ch. (2020, May 25). From 'coviidiots' to 'quarantine and chill', the pandemic has led to many terms that help people laugh and commiserate. *BibGuru*. Retrieved from <https://www.bbc.com/worklife/article/20200522-why-weve-created-new-language-for-coronavirus>
- Social-distancing. (2020). *Cambridge English Dictionary*. Retrieved from <https://oed.com/view/Entry/88377097>
- Steward, T. (2020, July 15). July 2020 update: scientific terminology of Covid-19. [Blog post]. Retrieved from <https://public.oed.com/blog/july-2020-update-scientific-terminology-of-covid-19/>
- Strunk and White. (2009) *Elements of Style*. Boston: Allyn and Bacon.