

CORPORUM JOURNAL OF CORPUS LINGUISTICS

HEC “Y” CATEGORY
ISSN: 2617-2917

Lexicogrammatical Features of Covid-19: A Syntagmatic and Paradigmatic Corpus Based Analysis

Sadia Irshad
Sadia Arshad
Kaukab Saba

To cite this article: Irshad, S., Arshad, S., & Saba, K. (2021). Lexicogrammatical Features of Covid-19: A Syntagmatic and Paradigmatic Corpus Based Analysis. *Corporum: Journal of Corpus Linguistics*, 4(2), 76-94

Corpus Research Centre
Faculty of Social Sciences
Air University
ISLAMABAD



Lexicogrammatical Features of “Covid-19”: A Syntagmatic and Paradigmatic Corpus Based Analysis

Sadia Irshad

Assistant Professor
Dept. of English
Air University, Islamabad
sadia.irshad@mail.au.edu.pk

Sadia Arshad

Lecturer (English),
NUST,
School of Civil & Environmental
Engineering, Islamabad
sadia.arshad@nice.nust.edu.pk

Kaukab Saba

Lecturer (Linguistics),
International Islamic University,
Islamabad
kaukab.saba@iiu.edu.pk

Correspondence:

Sadia Arshad
sadia.arshad@nice.nust.edu.pk

Received: 15th Dec, 2020

Accepted: 8th Jun, 2021

Published: 23rd Dec, 2021

Keywords: Covid-19,
Lexicogrammar, Paradigmatic
relations, Syntagmatic relations
and Systemic Functional
Grammar

Abstract

The outbreak of Covid-19 has influenced people worldwide and has led to generate a bulk of literature in the form of awareness campaigns, research reports, blogs, etc. The text and corpus produced, thus is of interest to the language researchers and linguists. According to Halliday and Matthiessen (2014), the corpus of a language helps to theorize the language as it provides authentic, representative, and quantitative data. The word *Covid-19* since its coinage appears in the corpora as a lexeme with a particular grammatical structure. This study focuses on the lexicogrammar of the node word *Covid-19* in the Covid-19 corpus of 224,061,570 words available on the Sketch Engine. The design of this research focuses on the perspective of systemic functional grammar to analyze the paradigmatic and syntagmatic relations of the word *Covid-19*. We found that it occurred 423× times as a noun in the corpus of 1.51 per million tokens. The word-sketch of *Covid-19* includes *infect, spread, fight* as syntagmatically related and the words *pandemic, epidemic, disease, outbreak, case* as paradigmatically related. However, the same lexeme may enter into both relationships like *Covid-19 and disease, Covid-19 and epidemic*. They enter into a systemic contrast, and they also collocate. The lexicogrammatical environment of *Covid-19* in the corpus shows the dichotomy of emergency and safety; chaos and health care responsibility; and pandemic and health care systems. This analysis helps in understanding how lexicogrammatical systems and structures relate to meaning-making process.

1. Introduction

Pandemics are not new and have occurred at different stages in human history (Ferguson et al., 2020). Multiple instances of pandemic outbreaks in past are proof for this. The most famous of them are Bubonic, Cholera, Influenza, smallpox, yellow fever, Spanish flu, wine flu,

and the list goes on (Brodeur et al., 2020). The year 2020 introduced a new pandemic, named Coronavirus Disease-19 or COVID-19 (Qiu et al., 2020), to the world and changed the whole dynamics, socially, culturally, politically, and economically. The virus affected many lives, and the data kept on changing and multiplying each passing day because the rate of infection was skyrocketing (Brodeur et al., 2020).

Due to the impact of Covid-19 on every walk of life of human beings, the term *Covid-19* remains the most researchable topic in various disciplines, especially linguistics. The research on Covid-19 in the field of linguistics includes the study of phonetics (Cavallaro et al., 2020), morphology (Anyanwu & Udoh, 2020), semantics (Shymko & Babadzhanova, 2020), syntax (Astia & Yunianti, 2020), sociolinguistics (Ibrahim et al., 2020) and critical discourse analysis (Syartanti, 2020). The present study also focuses on the study of language use from the perspective of functional grammar. Lexicogrammar, according to Systemic Functional Grammar emphasizes the nexus between lexical items and their grammatical usage (Halliday & Matthiessen, 2014). Although the term lexicogrammar integrates grammatical and lexical patterns, the grammatical perspective is central. Sinclair (2004) defines that lexicogrammar “does not integrate the two types of patterns as its name might suggest—it is fundamentally grammar with a certain amount of attention to lexical patterns within the grammatical frameworks; it is not in any sense an attempt to build together grammar and lexis on an equal basis... Lexico-grammar is still firmly a kind of grammar, laced, or perhaps spiked with some lexis” (p. 1). The word classes in lexicogrammar occur at two axes, that is vertical and horizontal. The former refers to paradigmatic, and the later axis is syntagmatic. A paradigmatic relationship defines the systems, choices and substitution of words, while a syntagmatic relationship elaborates structure, chains and positioning (Chiu & Lu, 2015). Halliday’s theory of Systemic Functional Grammar to study this paradigmatic organization and the relationship between the paradigmatic and the syntagmatic axes “is one of the breakthroughs in twentieth-century theoretical linguistics” (Halliday & Matthiessen, 2014, p.xvii).

Keeping this trend in mind the current research is an endeavour to explore the linguistics perspective of Covid -19 through corpus analysis. Language analysis involves a study of data that requires energy and time. The corpus-based study helps to analyze the language of the selected corpus from various linguistic perspectives. This study has opted for lexicogrammatical analysis of the Covid-19 corpus available on Sketch Engine. It has various tools and functions which can be used in linguistic analysis, for example, identifying frequency, concordance lines, word sketch, thesaurus, keywords, N-grams, and collocations. We argue that it is significant to study corpora for identifying lexical sets around Covid-19 from “roundabout” (Halliday & Matthiessen, 2014 p.59). In other words, paradigmatic and syntagmatic relations define the lexicogrammatical structure of a lexeme. Therefore, this study postulates a relationship between the syntagmatic lexical collocations and syntagmatic grammatical structures. The lexemes operate by establishing the relationship between paradigmatic lexical sets and paradigmatic grammatical systems.

The corpus research technological tools and “corpus linguistics has made the identification of lexicogrammatical patterns much easier than it once was,” (Pearce, 2007, p. 214). This study on both axes helps to provide information on collocation and grammar of *Covid-19* as well as illustrates its lexical sets and its grammatical systems. This study aims to

analyze the lexicogrammatical features of lexeme *Covid-19* in a corpus on Sketch Engine. This corpus includes texts released under COVID-19 Open Research Dataset (CORD-19).

2. Literature Review

While there have been many outbreaks and human catastrophes, there has been a notable rise in the frequency of pandemics from the year 2000 and thereafter (Brodeur et al., 2020). History is witness to the fact that all these pandemics were handled and stopped and the efforts of WHO, advanced medical field, and precautions played an operative role in putting a full stop to these pandemics (McCauley, Minsky & Viswaanth, 2013 cited by Singer, (2020). The Covid-19 pandemic broke in Wuhan, China, and spread like jungle fire adversely affecting comforts, businesses, jobs, studies, personal and public life. The pandemic has become a hot topic in every walk of life including drawing rooms, media, politics, sociology, psychology, research, and economics (Singer, 2020).

Covid-19 is not the first-ever pandemic on planet Earth but the first pandemic of the digitized world (Singer, 2020) as the previous pandemics happened before the digital era. That is why the approaches toward this pandemic are different as compared to the earlier pandemics when one corner of the world was not as easily connected to the other corner of the world. Furthermore, Singer (2020) has argued that Coronavirus is not a new pandemic as multiple infections outbreak in the past too but still none of them created a panic as worst as created by Covid-19. One of the reasons for making this panic worst could be the fast communication of the fear associated to this pandemic even before it reached physically. Social distancing and lockdown are considered to be the only temporary solution for flattening the curve of Corona positive cases and since the vaccine is still not prepared for the virus the leaders across the globe had to rely on this temporary solution of social distancing. But this solution has adversely affected the businesses and economy. The exceptional nature of coronavirus and strategy against it through social distancing led to an immense decline in financial activity (Alstadsæter et. al., 2020).

The research on Covid-19 is not only limited to scientific papers but the issue was explored from various angles such as economy (Herat, 2020) politics (Imhoff & Lamberty, 2020), sociology (Singer, 2020), psychology (Kazak, 2020), and linguistics perspectives. Linguistics analysis has remained part of research since the emergence of the field of linguistics but to handle the huge data requires extra energy and time by the researcher. Corpus linguistics has made the task easier for the researchers to conduct computer-based data analysis. The corpus-based analysis is not only limited to linguistics rather expanded to multiple disciplines as it is “the approach” (Louw & Milojkovic, 2016) and “methodology” (Brezina, 2018) to analysis which is part of any research irrespective of disciplines boundaries.

Multiple corpora such as British National Corpus (BNC), American National Corpus (ANC), Corpus of Contemporary American English (COCA), International Corpus of English (ICE) are available for corpus analysis. Moreover, researchers use certain corpus tools such as Antconc, Wordsmith, CLAWS, corpora CoCo, Sketch Engine for analyzing the language of the selected corpora from different linguistic perspectives depending upon the research objectives. Since Covid-19 has cast a great impact on every walk of life and has changed the world

dynamics, the researchers and academics are not freed from this topic. Numerous researchers use the corpora of Covid-19 on Sketch Engine from various angles and multiple perspectives to do the detailed analysis in their research because the Sketch Engine is a “user friendly” and “wider used” corpus tool in the field of lexicography domain (Purwitarini, p, 214, 2020). Muslimah (2020) has compared the HIV and Covid-19 with modifier *critical* on Sketch Engine. Sulalah (2020) has conducted a study to explore the semantic prosody of the verb ‘increase’ in the corpus of Covid-19 available on Sketch Engine. Astia & Yuniarti (2020) have analysed the most frequent adjectives of Covid-19 used in the corpora of academic writing through Sketch Engine. Another study was conducted by Purwitarini (2020) to identify the frequency of the positivity or negativity of the words “suspect” and “patient” in the corpus of Covid-19 on Sketch Engine through collocation. Covid-19 corpus of 280 million words in Sketch Engine consists of text from various genres and it is publicly accessible (Tan et.al, 2020). The research thus shows that the Sketch Engine has various tools and functions which can be used in linguistic analysis, for example, identifying frequency, concordance lines, word sketch, thesaurus, keywords, N-grams, and collocations.

The term collocation was first defined by Palmer (1933), as “a succession of two or more words that must be learned as an integral whole and not pieced together from its parts” (p.1). Later on, Firth (1957, p.11) elaborated *collocates* with his famous definition “you shall know the word by the company it keeps”. This definition is further extended by Sinclair (2000) defining collocates as neighboring words existing together and influencing each other. Words have denotative as well as connotative meanings for that reason semantic definition of each word is influenced by the neighboring words it is surrounded by. As highlighted by Xiao & McEnery (2006) in their study that collocates contribute to the meaning making process within a text. It leads to the discussion of partnership of collocates within a text hence it is not wrong to say that collocates are the lexical partners of the node words (Joharry & Turiman, 2020). Collocation is the combination of words grouped within a text hence, this combination can be syntagmatic or paradigmatic and the corpus is a handy tool to study Systemic Functional Linguistics (SFL) as presented by Halliday (1993). SFL is an approach to study the function, meaning, and structure of language altogether instead of focusing on only structure. SFL is the study of the relationship between language and its function within a social setting (Halliday, 1993). SFL has introduced Grammar with a different perspective of linking it with semantics and functional grammar and as a meaning-making resource (Liu & Jang, 2009). Halliday (2004) introduced another term in the field of SFL named lexicogrammar strengthening the relationship between lexis and syntax and it believes that each lexeme has its own grammar i.e., noun is a lexeme which has syntactical use in term of singularity and plurality, masculinity and femininity. Similarly, the syntactic value of the singular or plural noun decides its placement of the verb i.e. we add *s* or *es* with the verb if the noun is singular.

There are two types of relationships between the lexical items that can be viewed “at their level” i.e., paradigmatic (vertical axis) and syntagmatic (horizontal) (Halliday, 2004, p. 59). These two relationships were initially introduced by Ferdinand De Saussure, but Halliday expanded this concept further in terms of lexicogrammaticality. Paradigmatic is concerned with the semantic relationship of the lexical items whereas syntagmatic relation reveals the structure (Wanying & Kun Lu, 2016). The paradigmatic analysis of lexical items reveals their semantic features whereas common patterns of collocations would be revealed through syntagmatic relation (Halliday & Matthiesen, 2014). Collocates are the neighboring words of the lexical items

to be explored within a text and 'true collocation can be shown to be not only statistically but also semantically relevant' (Halliday, 2004). Paradigmatic and syntagmatic relationships can be studied from two angles, lexical and grammatical. This study presumes that to study these relationships, a study of the corpus may reveal the lexico-grammatical structure of any selected word. As explained earlier, the discourse on Covid -19 has been studied from various angles in numerous disciplines but the current study focuses on the question that what are the paradigmatic and syntagmatic relations of the node word Covid-19 in the corpus created on Sketch Engine? For this purpose, the study aims to investigate the relationship between the lexical collocations and grammatical structures; and it also studies the relationship between lexical sets and grammatical systems.

3. Research Methodology

3.1 Research Design

Use of Corpus technological tools is an innovative approach for linguistic analysis. As a methodology corpus-based research is an empirical method that utilize an authentic and meaningful data to achieve the objective outcomes. However, results need to be analyzed and interpreted in the light of theory and linguistic hypothesis (Hunston, 2002, as cited in Purwitarini, 2020). As the purpose of the study was to analyze the lexicogrammatical features of the node word *Covid-19* using COVID-19 corpus through Sketch Engine, the present study applied a qualitative method. The corpus taken was the *Open corpus* of Covid-19 available at Sketch Engine with 224,061,570 words. It is the combination of different genres of both pure and social sciences. Sketch Engine provides several other Open corpora with different functions for analysis such as word sketch, concordance, wordlist, keywords, N-grams and many more. For the present research Word Sketch function was used to study lexicogrammatical features of the node word *Covid-19*. COVID-19 corpus on Sketch Engine can easily be accessed through <https://app.sketchengine.eu/#open> or http://ske.li/covid_19.

3.2 Data Analysis Procedure

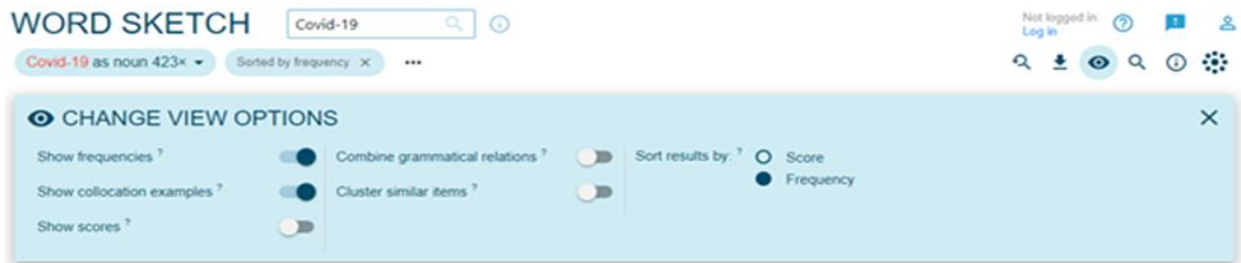
This study design includes few steps to collect and analyze the data using the Sketch Engine tool. There are several options available on Sketch Engine (Figure 1). However, researchers have only used those options that aided to answer our research questions. As a first step, the researchers typed the word *Covid-19* in the Word Sketch option which display collocation, word combinations and concordances. They then analyzed the words that collocate most frequently with the keyword *Covid-19*. Next, they studied the collocates based on POS tagging (identifying the part of speech of a word and labelling it). Sketch Engine automatically distributes the lexical items into their categories or tag them based on part of speech. This feature helped researchers to study the lexical items based on the parts of speech.

Secondly, in a *change view options*, the option of *show frequencies, frequency* and the cluster similar items were selected for analyzing the occurrences in detail (Figure 1). Thirdly, visualization of the tabulated data was accessed through *show visualization option* to get the most comprehensive picture of the results. Finally, we studied the most frequently occurring collocates around *Covid-19* through their concordance lines to establish the syntagmatic and paradigmatic relationship between the lexical items. Paradigmatic relationship defines the

systems, choices and substitution of words whereas syntagmatic elaborates structure, chains and positioning (Chiu & Lu, 2016). Halliday and Matthiesen (2014) illustrate the relations of paradigmatic and syntagmatic where former displays delicacy and latter rank.

Figure 1

Change view options available on Sketch Engine



4. Results

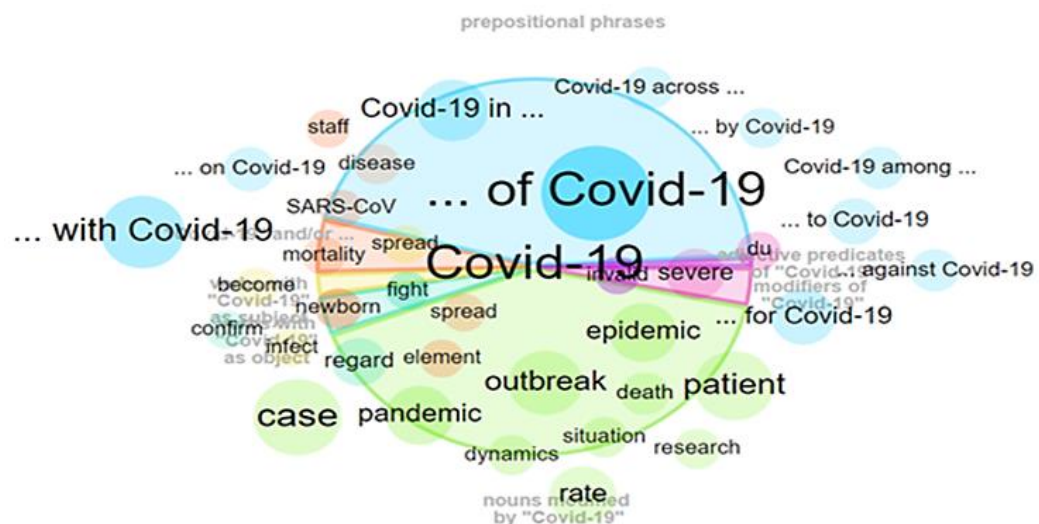
In this section we present the results of our analysis of the word *Covid-19* on both axes in the selected corpus.

4.1 Syntagmatic/Lexical (collocation)

For this study, *Covid-19* was chosen to be the node word. It occurred 423× times as a noun in the corpus of 1.51 per million tokens. The number of collocates increases or decreases by increasing or decreasing the span of collocates.

Figure 2

Visualization of Covid-19 collocates (Source: Sketch Engine)



The greater the frequency, the bigger and bold the word will be in the visualization display. We found a total of 7 categories of collocates in the vicinity of *Covid-19* (Figure 2) that are as follows:

- a. Modifiers of *Covid-19*
- b. Nouns modified by *Covid-19*
- c. Verbs with *Covid-19* as an object
- d. Verbs with *Covid-19* as subject
- e. Conjunctions with *Covid-19*
- f. Prepositional Phrases containing *Covid-19*
- g. Adjective predicates of *Covid-19*

Table 1

Collocates of Covid-19 and their occurrence probability on Sketch Engine

modifiers of "Covid-19"	nouns modified by "Covid-19"	prepositional phrases	
severe 9 patients with severe Covid-19	case 38 confirmed Covid-19 cases	... of "Covid-19" 72	
du 2 du Covid-19	patient 28 severe Covid-19 patients	... with "Covid-19" 33	
"Covid-19" and/or ...			
newborn 3 Covid-19 and 28 newborns	outbreak 22 the Covid-19 outbreak	"Covid-19" in ... 19	
disease 3 disease , Covid-19	pandemic 16 the Covid-19 pandemic	... for "Covid-19" 13	
element 2 current outbreak of Covid-19 , such elements may not be	epidemic 16 of the Covid-19 epidemic	... to "Covid-19" 6	
staff 2 unsettling information on Covid-19 , staff providing psychological or	rate 15 Covid-19 growth rates	... on "Covid-19" 5	
spread 2 Province . Regarding Covid-19 , the spread seems to have	death 4 numbers of confirmed Covid-19 deaths	"Covid-19" across ... 4	
SARS-CoV 2 Covid-19 , SARS-CoV	situation 3 Covid-19 pandemic situations	... against "Covid-19" 4	
mortality 2 clinical severity of Covid-19 , neonatal mortality , admission to	dynamics 3 global patterns of Covid-19 early outbreak dynamics during January-March	"Covid-19" among ... 4	
verbs with "Covid-19" as object			
regard 7 studying related information regarding Covid-19 in written and	research 3 on Covid-19 research	... by "Covid-19" 3	
fight 2 fight Covid-19	hospitalisé 2 Covid-19 hospitalisés	... in "Covid-19" 3	
confirm 2 pregnant women with confirmed Covid-19	staff 2 unsettling information on Covid-19 , staff providing psychological or	... from "Covid-19" 3	
	mortality 2 clinical severity of Covid-19 , neonatal mortality , admission to	"Covid-19" on ... 3	
	element 2 current outbreak of Covid-19 , such elements may not be	"Covid-19" from ... 2	
		adjective predicates of "Covid-19"	
		invalid 2 the information regarding Covid-19 is invalid and is produced	
		verbs with "Covid-19" as subject	
		spread 2 Covid-19 has spread	
		infect 2 infected by Covid-19	
		become 2 Covid-19 has become	

Seventy-four (74) collocates of prepositional phrases occur around *Covid-19* which is the highest frequency among all the categories of collocates. Seventy-two (72) instances of *of* “Covid-19”, 33 *with* “Covid-19”, 19 hits of “Covid-19” *in*, 6 *to* “Covid-19”, 5 *on* “Covid-19”, 4 hits of “Covid-19” *across, against* “Covid-19”, “Covid-19” *among*, 3 instances of *by* “Covid-19”, *in* “Covid-19”, *from* “Covid-19”, “Covid-19” *on*, and 2 hits of “Covid-19” *from* were identified with the node word Covid-19 (Table 1). It means that when you hear or read the word *Covid-19*, there is a heightened expectancy that it will occur in the vicinity of prepositions. Now, when we further study the collocates of Covid-19 in the prepositional phrases, we find that the frequency of the collocates increases by increasing the range (span). Second highest collocates were nouns modified by the keyword Covid-19 with the total number of 154 collocates.

When we further studied the collocates of nouns modified by Covid-19, we found that collocates in the same vicinity decrease or increase by decreasing/increasing the span, from +1-1 to +5 -5 (Table 2& 3).

Table 2

Change in number of collocations by increasing span (range)

Range [?]												
		-5	-4	-3	-2	-1	KWIC	1	2	3	4	5
	Word	Cooccurrences [?]					Candidates [?]	T-score	MI	↓ LogDice		
1	<input type="checkbox"/> ranged	3					10,592	1.73	11.03	3.21 ...		
2	<input type="checkbox"/> periods	4					15,596	2.00	10.89	3.07 ...		
3	<input type="checkbox"/> confirmed	13					65,495	3.60	10.52	2.70 ...		
4	<input type="checkbox"/> cases	36					211,856	6.00	10.29	2.48 ...		
5	<input type="checkbox"/> incubation	5					36,588	2.23	9.98	2.16 ...		
6	<input type="checkbox"/> growth	7					65,689	2.64	9.62	1.80 ...		
7	<input type="checkbox"/> rate	5					118,594	2.23	8.28	0.47 ...		
8	<input type="checkbox"/> reported	7					196,419	2.64	8.04	0.22 ...		
9	<input type="checkbox"/> data	5					310,652	2.22	6.89	-0.92 ...		
10	<input type="checkbox"/> between	3					357,467	1.70	5.95	-1.86 ...		
11	<input type="checkbox"/> from	7					1,138,769	2.59	5.51	-2.31 ...		
12	<input type="checkbox"/> 1	4					818,225	1.94	5.17	-2.64 ...		
13	<input type="checkbox"/> of	27					8,766,274	4.97	4.51	-3.31 ...		
14	<input type="checkbox"/> The	4					1,585,718	1.89	4.22	-3.60 ...		
15	<input type="checkbox"/> were	4					1,722,066	1.88	4.10	-3.72 ...		
16	<input type="checkbox"/> to	9					4,465,166	2.80	3.90	-3.92 ...		
17	<input type="checkbox"/> the	20					10,795,472	4.15	3.77	-4.04 ...		
18	<input type="checkbox"/> for	3					2,222,910	1.56	3.32	-4.50 ...		
19	<input type="checkbox"/> in	7					5,264,881	2.38	3.30	-4.52 ...		
20	<input type="checkbox"/> a	4					3,157,149	1.79	3.23	-4.59 ...		

In the second column of Table 1, word *cases* with the similar items *patient* and *rate* clustered together around the node word *Covid-19* occur 38 times, word *outbreak* with the similar items *pandemic*, *epidemic*, *situation* and *research*, 22 times, *death* along with *mortality*, *dynamics* with *element* 4 and 3 times respectively, *hospitalize* and *staff* 2 times, with the keyword *Covid-19*. Thus, these occurrences indicate that all these words will be found with the highest expectancy in the surroundings of node word *Covid-19*.

Table 3

Change in number of collocations by decreasing span (range)

COLLOCATIONS												
BASIC		ADVANCED		ABOUT								
Attribute *		Range *										
word		-5	-4	-3	-2	-1	KWIC	1	2	3	4	5
	Word	Cooccurrences ?	Candidates ?	T-score	MI	↓ LogDice						
1	<input type="checkbox"/> confirmed	11	65,495	3.31	10.28	2.46	...					
2	<input type="checkbox"/> cases	35	211,856	5.91	10.25	2.44	...					
3	<input type="checkbox"/> of	8	8,766,274	2.41	2.75	-5.06	...					
4	<input type="checkbox"/> the	5	10,795,472	1.58	1.77	-6.04	...					

The third highest frequent words were conjunctions “and” and “or”. There is a greater probability that the node word *Covid-19* will occur within the boundary of and/or. Word *newborn* (Figure 3), *disease*, and *mortality* (similar items cluster) are the most frequently occurring words that occur 3 times, *element*, *staff*, *spread*, *SARS-CoV* occur 2 times in the neighborhood of *Covid-19* (Table 1).

Figure 3

Covid-19 and newborn collocation pattern

> Thirty-three pregnant women with Covid-19 and 28 newborns were identified. </s:
 case series of pregnant women with Covid-19 and newborns born to these women.
 anagement of pregnant women with Covid-19 and their newborns . </s><s> In conc

Verbs used as an object of Covid-19 are *regard*, *flight*, and *confirm* with the frequency of 7 and 2. On the other hand, verbs used as subjects are *spread*, *infect*, and *become* with the frequency of 2 each (Table 1). Modifiers that are used most frequently with Covid-19 are *severe* and *du* (a French word) with the frequency of 9 and 2 respectively. The word *invalid* is used as an adjective predicate with a frequency of 2 (Table 1).

4.2 Syntagmatic/Grammatical (structure)

Another analysis on the syntagmatic level is to study the realization of the structure, which is the organic configuration of the elements in functional terms. According to Halliday (1996), such a sequence of classes is known as syntagm. How different classes have been chained within the corpus with the node word Covid-19 is investigated. The first 4 concordance lines of Prepositional Phrases containing Covid-19 (Figure 4) show that the most common structural pattern in which Covid-19 exists is noun + preposition + noun (line 1), adjective + noun+ preposition+ noun (line 2), determiner + adjective + preposition+ noun (line 3), determiner + adverb+ adjective + noun+ preposition+ noun (line 4). But this structural classification tells us very little about what it means unless analyzed in functional terms.

Figure 4

Commonly used structural patterns

expected to accelerate the spread of Covid-19 .	</s><s> No strong intervention w
be critical for pandemic prevention of Covid-19 ,	including prioritization the surveill
onsistent with the rapid spreading of Covid-19 .	</s><s> Nevertheless, by taking
hina </s><s> The first known case of Covid-19	has been traced back to Decembe

From commonly used structural patterns in Figure 4, *spread* (concordance line 1) denotes the entity being referred to, its function here is as a thing, an abstract noun, the occurrence of which is visible in the deteriorating change in the behavior of human beings (Table 4).

Table 4

Syntagmatic structural & functional analysis of entities

The			spread	of Covid-19
Deictic			Thing	Qualifier + Classifier
		Pandemic	Prevention	of Covid-19
		Qualifier	Thing	Qualifier + Classifier
The		Rapid	Spreading	of Covid-19
Deictic		Post-deictic	Thing	Qualifier + Classifier
The	first	known	Case	Of Covid-19
Deictic	Deictic + Quantifier	Post Deictic	Thing	Qualifier + Classifier

Covid-19 here functions as a qualifier, as we have to find out what kind of spread is meant here. At the same time, *Covid-19* is also functioning as a classifier, as it also denotes the class within the general category of disease. Further, *the* has a pointing out function, known as Deictic, as it signals some particular attribute that is linked with the classifier or a disease. In concordance line 2 *pandemic prevention of Covid-19* and 3 *the rapid spreading of Covid-19*, similar structures can be seen, except in concordance line 2, deictic *the* is not used. In concordance line 4, *first* is functioning as a quantifier, pointing towards the first case of Covid-19. Thus, the grammatical sequences of entities and classes indicate the underlying functional relationship that makes the sentence meaningful or vice versa, for example, if we try to jumble up the sentence structure of line no.4 as **to first known the case Covid-19 of* does not make any sense. It clearly indicates the words have a structural relationship.

4.3 Paradigmatic/lexical (the lexical set)

Paradigmatic analysis of lexical items shows that they share semantic features and common patterns of collocation (Halliday & Matthiesen, 2014). If we look at the cluster of words in Picture no 1, we find that *Covid-19*, *disease*, *SARS-CoV* share the same features, and the corpus shows that they have in common the tendency to collocate with words such as a *pandemic*, *epidemic*, *outbreak*. When it is asked what words come into mind on hearing the word Covid-19, people might come up with the lexical sets that can be found on both the axes. For example, if we look at word-sketch of Covid-19, we see that the words *infect*, *spread*, *fight* are related to Covid-19 syntagmatically and word *pandemic*, *epidemic*, *disease*, *outbreak*, *case* are linked paradigmatically. However, the same lexeme can be related in both ways like *Covid-19* and *disease*; *Covid-19* and *epidemic* they enter a systemic contrast (Covid-19 is a type of disease) but they also collocate as in Figures 5 & 6.

Figure 5

Paradigmatic/lexical relationship of lexical set (Covid-19 & disease)

, is the aethiological agent of a pandemic **disease** , **Covid-19** , causing severe pneumonia outbreaks at the global scale (7) .
t, or Mann-Whitney U test, as appropriate. </s><s> **Covid-19** , coronavirus **disease** 2019. </s><s> All rights reserved. </s><s>
udies will establish a possible relationship between **Covid-19** and coronary artery **disease** . </s><s> Only the awareness of s

Figure 6

Paradigmatic/lexical relationship of lexical set (Covid-19 & epidemic)

Wuhan Province where the **Covid-19 epidemic** started were between 0.5
tions in 2015. </s><s> The **Covid-19 epidemic** is still in its early stages in
sive patients. </s><s> The **Covid-19 epidemic** is disrupting not only our
obilised in response to the **Covid-19 respiratory illness epidemic** . </s><s>

Members of lexical sets are also linked with each other through semantic features such as synonymy, antonymy, hyponymy and meronymy (Halliday & Mathiessen, 2014). Some may be alike or opposite in meaning; some are subtypes of the same type, and some may be the parts of the whole. For example, words like *Covid-19*, *coronavirus*, *coronary artery disease*, *SARS-CoV* can be co-hyponymy (a word whose semantic field is included within another word) of the hypernym (a word constituting a category) *disease*. While *epidemic*, *pandemic*, *outbreak* are synonymous creating an environment of global disaster. Moreover, *mortality*, *newborn cases*, *death*, *patients* are lexical sets of co-hyponymy related to life. Thus, here these complex relationships that are set up by the users of language within the corpus of Covid-19 contribute significantly to unfolding meanings of the text that helps in creating a huge impact on the social, economical, political, and cultural future of the world. These lexicogrammatical relationships, created by the writers and speakers within the text sometimes change the complete facade of human existence.

4.4 Paradigmatic/grammatical (grammatical system)

Grammatical categories are also organized in systems (Halliday & Mathiessen, 2014). Covid-19 is treated as a singular neutral noun; thus, no system of gender is attached with this entity in the corpus and the pronoun “*it*” is used for Covid-19 (Fig. 7).

Figure 7

Use of “it” for Covid-19 in the corpus in Sketch Engine

normal individuals, suggesting that **it** might not be a latent virus in humans(11). </s><s>
 coronavirus the crown-jewel of pandemics? </s><s> No. </s><s> **It** is a serious infectious disease,

Selection of words in the clause describes the environment of process type. The most frequently occurring process types in the corpus of Covid-19 are *spread*, *infect*, *become* in the vicinity of *Covid-19* and *fight*, *regard*, *confirm*, around *Covid-19*. In the current study, Halliday and Mathiessen's, (2014) categorization of material and mental processes is used to identify and analyze the environment of process types.

From the analysis of concordance lines (Figure 8) in which these processes occur around Covid-19, it is found that in all the clauses most frequently occurring selections in the system of process type are mental processes. However, Covid-19 corpus can be analyzed in detail to study the function and occurrence of material processes. From the results it can be inferred that the spread and pandemic of Covid-19 has created a world-wide suffering and fear, which is highlighted by the writers and speakers using mental processes.

In the material clauses the central participants around the environment of *Covid-19* are pandemic, safety measures, WHO, government agencies, health care systems, and public authorities, while in the mental clauses the participants are sensor and phenomenon i.e., fear, suffering and responsibility.

Figure 8*Exploring Process types in the corpus of Covid-19*

Left context	KWIC	Right context
Since January 2020, the	Covid-19	has spread rapidly from Wuhan
</s><s> Like SARS-CoV,	Covid-19	spreads by respiratory droplets
who are or may be	infected by	Covid-19 and who
:2 that he had been	infected by	Covid-19 following
• Three months later,	Covid-19	has become a worldwide pandemic
l that the epidemic of	Covid-19	had become a pandemic. </s><s> E
workers can help	fight	Covid-19 from a safer distance than t
/ countries are	fighting	Covid-19 with all strategies. </s><s>
information	regarding	Covid-19 is invalid and is produced
ce. </s>	<s>	Regarding Covid-19 , the spread seems to hav
t women with	confirmed	Covid-19 , 39.4% of women
t women with	confirmed	Covid-19 , so we believe tha

Figure 9*Examples of semantic relationship in Covid-19 corpus construed by the grammatical systems*

domestic spread of a pandemic, and mitigate **disease** , **suffering** and death; a

, and COVID-19 is still a big **fear** of the people over the world.

numbers of confirmed **Covid-19 deaths** , the Hubei province

or hypertension among **Covid-19 deaths** in Hubei </s><s> The

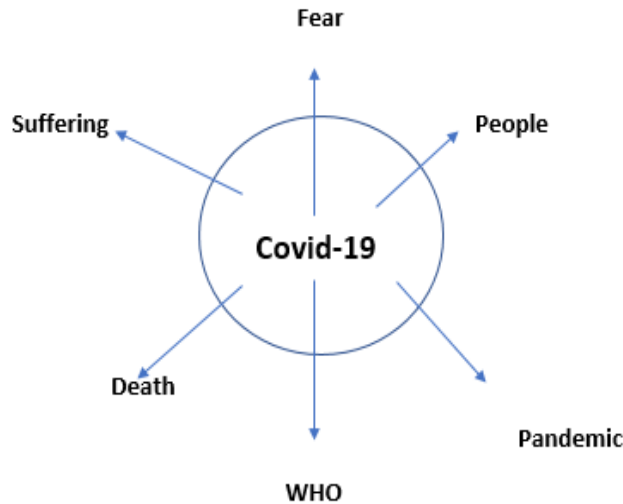
Of all those **people dying** within 1 year, it is likely that COVID-19

than 12,000 **people** have **died** according to data from WHO (https

Thus, a highly patterned selection of these choices creates an overall environment of emergency that calls for safety measures throughout the world due to pandemic, a sense of responsibility among the individuals related to different fields specially health departments. A network of semantic relationships is emerged from the data is given in Fig. 9 above. This relationship can also be represented diagrammatically as in Fig. 10 below.

Figure 10

Semantic relationship in Covid-19 corpus construed by the grammatical systems



5. Discussion

In this section, we first show the relationship between the syntagmatic lexical collocations and syntagmatic grammatical structures; and then between paradigmatic lexical sets and paradigmatic grammatical systems of the word “Covid-19”. We analyzed the lexicogrammatical features of the selected words in the corpus using Systemic Functional Grammar model (Halliday & Matthiesen, 2014).

As pointed out in Halliday and Matthiesen (2014), syntagmatic order in systemic theory is the compositional aspect of language or rank. On analyzing the grammar of “Covid-19”, this study finds that the structure of this unit in the corpus is an organic configuration whereby each part on either side has a distinctive function concerning the whole corpus. According to Halliday and Matthiesen (2014), the word-classes can be viewed semantically and at their level where they enter paradigmatic relations about the choices and the syntagmatic relations entailing the company they keep. The analysis shows *Covid-19* has 7 categories of collocates including modifiers of *Covid-19*, nouns modified by *Covid-19*, verbs with *Covid-19* as an object, verbs with *Covid-19* as a subject, conjunctions with *Covid-19* like “or”, prepositional phrases, and adjective predicates of *Covid-19*. The prepositions collocate most frequently with this word followed by nouns. The study shows that the number of collocates increase or decrease by increasing or decreasing the span of collocates. As discussed earlier the organic configuration of a word describes the function of units on either side of the selected word (node). The first 4 concordance lines containing the term *Covid-19* show that the most common structural pattern in which *Covid-19* exists is noun + preposition + noun (line 1), adjective + noun+ preposition+ noun (line 2), determiner + adjective + preposition+ noun (line 3), determiner + adverb+ adjective + noun+ preposition+ noun (line 4). This sequence of classes is referred to as syntagm, and it requires a further description of functions of such organic configuration. For example, in

“the spread of Covid-19” *Covid-19* functions as a qualifier, as we must find out what kind of spread is meant here. At the same time, *Covid-19* is also functioning as a classifier, as it also denotes the class within the general category of disease.

Paradigmatically, lexemes function in sets with shared semantic features (Halliday & Matthiesen, 2014). In the selected corpus the set of lexemes *Covid-19*, *disease*, *SARS-CoV* tends to collocate with words such as a *pandemic*, *epidemic*, *outbreak*. This shows the semantic relationship of synonymy between pandemic, epidemic, and outbreak; Covid-19 as a hypernym of disease; and Covid-19 and SARS-CoV are co hyponyms. The following shows the lexical patterns in the COVID-19 corpus on sketch engine.

1. disease [hyponymy] = Covid-19 =SARS-CoV [co hyponyms]
2. Pandemic=epidemic =outbreak [synonymy]

These complex lexical relationships within the corpus of Covid-19 contribute significantly in unfolding meanings of the text and help understanding an impact on the social (Singer, 2020), economic (Alstadsæter, 2020), political (Imhoff & Lamberty, 2020) and cultural future of the world. We illustrate the selection of words in the system of process types in the clause to view paradigmatic grammatical systems. The corpus shows Covid-19 in subject position with the verbs like *spread*, *infect*, *become* and *fight*, *regard*, *confirm*, when it occurs as an object. This paradigmatic environment of Covid-19 resonates strongly with the severity of the Covid-19 around the globe and described in earlier research too (Purwitarini, 2020). The results show that in the material clauses the central participants around the environment of *Covid-19* are pandemic, safety measures, WHO, government agencies, health care systems, and public authorities whereas in the mental clauses the participants are sensor and phenomenon i.e., fear, suffering and responsibility. These lexemes create the environment of Covid-19 in the corpus that shows the dichotomy of emergency and safety; chaos and health care responsibility; pandemic and health care systems. Thus, this study enabled us to gain insight into the Covid-19 corpus. Concluding the corpus analysis, we establish that it helps unfolding meanings by studying patterns in the syntagmatic and paradigmatic context in the system of language.

6. Conclusion

This study investigates the lexicogrammar of the lexeme *Covid-19* by analyzing its syntagmatic and paradigmatic relations-following systemic functional grammar approach in a corpus-based study. As corpus-based research is inductive, in this study the lexicogrammatical constructs of *Covid-19* emerged from the analysis of the corpus of *Covid-19* available on Sketch Engine.

The analysis reveals that the node *Covid-19* occurs predominantly as a noun i.e. 423× times in the corpus of 1.51 per million tokens. Syntagmatically *Covid-19* has 7 categories of collocates including modifiers of *Covid-19*, nouns modified by *Covid-19*, verbs with *Covid-19* as an object, verbs with *Covid-19* as the subject, conjunctions with *Covid-19* like *or*, prepositional phrases, and adjective predicates of *Covid-19*. The grammatical structure of *Covid-19* in the corpus describes its usage in noun specific slots and sequences of entities and classes indicate the underlying functional relationship that makes its usage meaningful in a sentence where it occurs as a noun. Word-sketch of *Covid-19* shows that the words *infect*, *spread*, *fight* are related to *Covid-19* syntagmatically and word *pandemic*, *epidemic*, *disease*, *outbreak*, *case* are linked

paradigmatically. Nevertheless, the lexemes can be related in systemic contrast and collocation like *Covid-19 and disease*; *Covid-19 and epidemic*. Finally, this study found that the paradigmatic analysis of lexical items around *Covid-19* showed that they share semantic features and common patterns of collocation. Therefore, collocates as neighboring words existing together and influencing each other (Sinclair, 2000) contribute to the meaning making process within a text (Xiao & McEney, 2006). Hence it is not wrong to say that collocates are the lexical partners of the node words (Joharry & Turiman, 2020).

Further work may explore similar paradigmatic and syntagmatic relations of other content words to identify lexeme specific lexicogrammatical systems and structures. As such analysis helps in understanding how lexico-grammatical systems and structures relate to meaning-making process. The future research may also explore *Covid-19* following similar or different research designs by comparing two or more corpora.

References

- Alstadsæter, A., Bratsberg, B., Eielsen, G., Kopczuk, W., Markussen, S., Raaum, O., & Røed, K. (2020). *The first weeks of the coronavirus crisis: Who got hit, when and why? Evidence from Norway* (No. w27131). National Bureau of Economic Research. <https://www.nber.org/papers/w27131>
- Anyanwu, E. C., & Udoh, V. C. (2020). Corpus Linguistic Analysis of Fear-Factor Lexemes of Selected Online Newspaper Headlines on Coronavirus Pandemic. *IDOSR Journal of Communication and English*, 5(1) 76-83
- Astia, I., & Yuniarti, S. (2020). Corpus-Based Analysis of the Most Frequent Adjective on Covid-19. *Indonesian Journal of EFL and Linguistics*, 5(2), 505-515.
- Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide*. Cambridge University Press.
- Brodeur, A. (2021). *A literature review of the economics of. November 2020*, 1007–1044. <https://doi.org/10.1111/joes.12423>
- Cavallaro, G., Di Nicola, V., Quaranta, N., & Fiorella, M. L. (2020). Acoustic voice analysis in the COVID-19 era. *ACTA Otorhinolaryngologica Italica*. <https://doi.org/10.14639/0392-100x-n1002>
- Chiu, W., & Lu, K. (2015). Paradigmatic relations and syntagmatic relations: How are they related? *Proceedings of the Association for Information Science and Technology*, 52(1), 1-4.
- COVID-19 Open Research Dataset (CORD-19). (2020). Version 2020-05-02. <https://10.5281/zenodo.3715505>
- Firth, J. R. (1957). A synopsis of linguistic theory, 1930-1955. *Studies in linguistic analysis*, 1-32. Basil Blackwell.
- Ferguson, N., Laydon, D., Nedjati Gilani, G., Imai, N., Ainslie, K., Baguelin, M., & Dighe, A. (2020). Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand.
- Halliday, M. A. (1993). Towards a language-based theory of learning. *Linguistics and education*, 5(2), 93-116.
- Halliday, M. A. K., & Matthiessen, C. M. (2013). *Halliday's introduction to functional grammar*. Routledge.
- Halliday, M., & Matthiessen, C. M. (2014). *Halliday's Introduction to Functional Grammar* (4th ed.). Routledge.

- Herat, M. (2020). "I feel like death on legs": COVID-19 isolation and mental health. *Social Sciences & Humanities Open*, 2(1), 100042. <https://dx.doi.org/10.1016%2Fj.ssaho.2020.100042>
- Ibrahim, E. R., Kadhim, S. A. H., Mayuuf, H. H., & Haleem, H. A. (2020). A Sociolinguistic Approach to Linguistic Changes since the COVID-19 Pandemic Outbreak. *Multicultural Education*, 6(4), 122-128.
- Imhoff, R., & Lamberty, P. (2020). A bioweapon or a hoax? The link between distinct conspiracy beliefs about the Coronavirus disease (COVID-19) outbreak and pandemic behavior. *Soc. Psychol. Personal. Sci.* (2020), 10.1177/1948550620934692
- Joharry, S. A., & Turiman, S. (2020). Collocation Networks and Covid-19 in Letters to the Editor: A Malaysian Case Study. *Asia Pacific Journal of Corpus Research*, 1(1), 1-30.
- Kazak, A. E. (2020). Psychology is an essential science: American Psychologist highlights the role of psychology in understanding and addressing COVID-19. *American Psychologist*, 75(5), 605.
- Louw, B., & Milojkovic, M. (2016). *Corpus stylistics as Contextual Prosodic Theory and subtext* (Vol. 23). John Benjamins Publishing Company.
- Matthiesen, C.M. (2014). Introduction. In Halliday, M., & Matthiessen, C. M. (2014). *Halliday's Introduction to Functional Grammar* (4th ed.). Routledge.
- McCauley, M., Minsky, S., & Viswanath, K. (2013). The H1N1 pandemic: media frames, stigmatization, and coping. *BMC Public Health*, 13(1), 1116. <https://doi.org/10.1186/1471-2458-13-1116>
- Muslimah, R. W. (2020). A Corpus-Based Analysis of Critical Strategies in Covid-19 Corpora. *Lire Journal (Journal of Linguistics and Literature)*, 4(2), 258-268.
- Palmer, H. E. (1966). Second interim report on English collocations. In *the Tenth Annual Conference of English Teachers under the Auspices of the Institute for Research in English Teaching. Tokyo, Institute for Research in English Teaching, 1933*. Harold E Palmer.
- Pearce, M., (2007) *The Routledge Dictionary of English Language Studies*. Routledge.
- Purwitarini, D. (2020). The Words "Suspect" and "Patient": A Corpus-Based Analysis of Semantic Prosody in COVID-19 Corpora. *OKARA: Jurnal Bahasa dan Sastra*, 14(2), 209-223.
- Qiu, Y., Chen, X., & Shi, W. (2020). Impacts of Social and Economic Factors on the Transmission of Coronavirus Disease 2019 (COVID-19) in China. *J Popul Econ*, 33, 1127-1172. <https://doi.org/10.1007/s00148-020-00778-2>

- Shymko, V., and Babadzhanova, A. (2020). Short communication: Linguistic Semantics of the Covid-19 Quarantine Concept Perceived by Ukrainians. *Advance*. Preprint. <https://doi.org/10.31124/advance.13102880.v1>
- Sinclair, J. (2000). Lexical grammar. *Naujoji Metodologija*, 24, 191-203.
- Singer, N. (2020). Coronavirus Media Discourse and Current Situation (QCA for Different Responses to Combat Covid-19). *Humanities & Social Sciences Reviews*, 8(3), 665-675.
- Sulalah, A. (2020). The Semantic Prosody analysis of 'increase' in Covid-19: a Corpus-Based Study. *Lire Journal (Journal of Linguistics and Literature)*, 4(2), 237-246.
- Syantanti, N. I. (2020, November). Humour in the midst of pandemic: critical discourse analysis on Covid-19 related memes. In *The 2nd International Interdisciplinary Conference on Language Affairs (IICoLA) 2020*.
- Tan, H. (2020). China's Xi rallies people back to work as country continues to battle virus outbreak. *CNBC*. Accessed April 11th 2020 at <https://www.cnbc.com/2020/02/24/chinas-xi-jinping-speaks-about-coronavirus-threat-on-economy.html>
- Wanying C. and Kun Lu. (2016). Paradigmatic relations and syntagmatic relations: How are they related? In *Proceedings of the ASIST*. 122:1-122:4.
- Xiao, R. & McEnery, T. (2006). Collocation, Semantic Prosody, and Near Synonymy: A Cross-linguistic Perspective. *Applied Linguistics*. 27. 10.1093/applin/ami045.