COLLOCATION OF *ISOLATION* IN THE CORPUS OF ENGLISH WEB 2018 AND COVID 19: SEMANTIC APPROACH

Siti Lutfiah Rabiyatul Adawiyah
Universitas Negeri Yogyakarta
Sitilutfiah.2021@student.uny.ac.id

Annisa Nurul Firdausi
Universitas Negeri Yogyakarta
Annisa401fbs.2021@student.uny.ac.id

Abstract

When studying language, it is not only limited to studying the meaning of the word for word but also how one-word collocates with another word. This study was conducted to know the differences in collocations from the word *isolation* since the beginning of covid 19. This study used a corpus-based approach and qualitative research methods. The research data was collected from two different corpora, the English Web 2018 and Covid 19. Six collocations were taken from each corpus with the highest MI value. The results show differences in collocations in the word isolation since the emergence of Covid 19. In corpus English Web 2018, one collocation shows a negative meaning and four collocations indicate positive meanings. While in the covid 19 corpus, two collocations show a negative meaning, and three collocations show a positive meaning.
Keywords: semantic preference, semantic prosody, corpus

Introduction

Language will always be interesting field to study. Many scholars worldwide have made language the object of their research studies. Since then, many theories of language have developed rapidly and facilitated research in language study. When studying language, it is not only limited to studying the meaning of the word for word but also how one-word collocates with another word. A linguist from England named Firth

1950 first coined the term collocation (Hu 2015). Collocation revolves around that certain words combine with specific words in specific contexts. An analysis of the collocation must first be carried out to analyze the semantic prosody. Many scholars examine collocation as the focus of their research (Emery 1991), (Rodríguez and Goddard 2013), and (Gelbukh and Kolesnikova 2012).

Many scholars have also carried out studies on the study of semantic preference analysis and prosody (Zhang 2010), (Cheng 2013), and (Bednarek 2008). Sinclair, in 1991 was the first to come up with the terms preference semantics and prosody (Begagić 2013). Semantic preference is a relationship, not between individual words, but between an entry or word form and a series of semantically related words. *Semantic preference* is also defined as the relationship between words and words that have a meaning relationship in the field of meaning (Lindquist 2018). A word or phrase is said to have negative or positive semantic prosody if the word or phrase is accompanied by lexical units with a negative or positive meaning (McEnery and Hardie 2011).

Semantic preferences and semantic prosody will always be related to the corpus, which is indicated by several studies related to these three things (Liu 2020) and (Begagić 2013). The corpus is a collection of authentic texts, both written and conversational transcripts in large numbers that are stored electronically. It shows that language studies have progressed with the advent of computer technology. The existence of a corpus, especially in linguistics, has helped linguists carry out research in large amounts of words.

The emergence of the coronavirus in 2019 disrupted the human immune system worldwide and caused many changes to the study of linguistics.(Asif et al. 2020). Many new words appear as well as the use of different old words. It allows for changes in the collocation of each word which then gives a different meaning. For this reason, this study was conducted to determine whether there has been a collocation change in the word isolation since the emergence of covid 19. This study used the 2018 English Web corpus

and compared it with the Covid 19 corpus. Then, to find out the lexical meaning of node isolation by studying semantic preferences and prosody.

Method

This study uses a corpus-based approach and qualitative methods to analyze data from the corpus. The data is taken from the Sketch Engine corpus, which can be accessed via https://www.sketchengine.eu/ with the English Web 2018 and Covid 19 subcorpus. The collocation of the word isolation is used as data for analysis from the two subcorpus. There are several data collection procedures: first, determining the two subcorpus to be compared, namely the English Web 2018 corpus and Covid 19; second, finding the collocation with the highest MI value; third, analyzing each collocation semantically to determine the differences between collocations from the node isolation.

Findings and Discussions

Collocation of *Isolation* in The Corpus of English Web 2018

The node in this study is the word "isolation", followed by several different collocations in the corpus of English Web 2018. Many collocations follow node "isolation" in this corpora, but chosen collocations below are taken from the top five words often used as collocations from the word "isolation". Five collocations were selected based on the highest MI value, as shown in the table below:

Collocate	Freq	Coll. freq.	T-score	MI
Loneliness	4525	76074	67,2527	12,0925
Characterization	2558	162513	50,5329	10,1745
RNA	2490	280493	49,8233	9,34826
Reproductive	2036	217931	45,0563	9,42194
Transformer	1314	88820	36,2158	10,0851

Table 1. The collocation of "isolation" in the corpus of English web 2018

The data in table 1 shows that the word isolation in the corpus of English web 2018 collocates with several words such as loneliness, characterization, RNA, reproductive, and transformer. The word "lonelines" has the highest frequency, with a value of 4,525. The second position is the word "characterization," which has a frequency with a value of 2558. The third, fourth, and fifth positions are RNA, reproductive, and transformer, with frequencies of 2490, 2036, and 1314, respectively.

The following figure will show the application of each collocation of "isolation" in the corpus of English Web 2018.

Left	Kwic	Right
the savings or sufficient government pension to finance a comfortable retirement. Studies have shown that social	isolation	and **loneliness** among the elderly are killers; contact with younger people is good for their health. As a result,
Enhanced Biological Trace Organic Contaminant Removal (EBTCR) during wastewater treatment. This project involved	Isolation	and **characterization** of bacteria capable of degrading various trace organic contaminants, bioaugmentation of
calculated as described earlier [24]. Quantitative (real-time) RT PCR studies For quantitative PCR analysis, **RNA**	isolation	was carried out using SpectrumTM plant total RNA isolation kit (Sigma) from neem callus, suspension, leaf, flower,
it tends to split into branching systems of subgroups or varieties. Ecological differentiation and **reproductive**	isolation	result in some of these varieties becoming new species. This process is normally a gradual one that occurs under the
building infrastructure and industrial applications. Eaton 9355 40kVA Uninterruptible Power Supplies Options:	Isolation	**transformer**, long-life batteries, external battery cabinets, integrated manual bypass and external maintenance

Figure 1. Concordance of "isolation" in the Corpus of English Web 2018

Based on figure 1, most collocations of 'isolation' are placed on the right, and two collocations are on the left side. According to the online oxford dictionary, *loneliness*,

characterization, RNA, and *transformer* are nouns, while only *reproductive* is an adjective. The most collocation used by the node isolation in the corpus English Web 2018 is part of speech noun. There is only one collocation part of speech adjective. When the node isolation is on the right side, the collocation act as a modifier to modify noun after it. For example, RNA is a noun that modifies noun isolation. Another example is reproductive as an adjective to modify noun isolation. However, when the node is on the left side, the node act as a modifier to modify noun after it. For example, noun transformer is modified by isolation. The other two nouns are the collocations of the node isolation that always come together using conjunction and.

Left	Kwic	Right
the savings or sufficient government pension to finance a comfortable retirement. Studies have shown that social	isolation	and **loneliness** among the elderly are killers; contact with younger people is good for their health. As a result,

Data 1. The combination of isolation and loneliness

According to Online Oxford Learner's Dictionary, loneliness means feeling sad because you do not have people to talk to or have no friends. Loneliness is a noun, and isolation is a noun. The combination of those two nouns creates a negative meaning. Data 1 shows that feeling of loneliness and isolation among the elderly negatively affects them, so they need to converse with the youth. Those two words combine and make a negative meaning.

Left	Kwic	Right
Enhanced Biological Trace Organic Contaminant Removal (EBTCR) during wastewater treatment. This project involved	Isolation	and **characterization** of bacteria capable of degrading various trace organic contaminants, bioaugmentation of

Data 2 The combination of isolation and characterization

According to the online oxford dictionary, characterization is how something is described or defined, while isolation is the act of separating something. Characterization and isolation are nouns. Both words are usually combined and used in biological study. Microbial isolation separates one type of microbe from another from various microbial mixtures intending to obtain pure cultures. The biologist will know the bacteria's morphological, biochemical, and molecular characteristics. The combination of characterization and isolation based on data 2 has a positive meaning.

Left	Kwic	Right
calculated as described earlier [24]. Quantitative (real-time) RT PCR studies For quantitative PCR analysis, **RNA**	**isolation**	was carried out using SpectrumTM plant total RNA isolation kit (Sigma) from neem callus, suspension, leaf, flower,

Data 3 The combination RNA and isolation

Data 3 shows that isolation collocation is placed on the right side. According to Online Oxford Dictionary, RNA is a chemical in all living cells, like DNA. RNA is a noun, and its combination with isolation is widely used in biological study by looking at data 3. RNA isolation is the first step biologists must take in the genetic engineering process before moving on to the next procedure. Just like data 2, the combination of RNA and isolation shows positive meaning in biological study.

Left	Kwic	Right
it tends to split into branching systems of subgroups or varieties. Ecological differentiation and **reproductive**	**isolation**	result in some of these varieties becoming new species. This process is normally a gradual one that occurs under the

Data 4 The combination of reproductive and isolation

Data 4 shows isolation is collocated with the adjective 'reproductive' in the right place. Reproductive is connected to reproducing based on the online oxford dictionary. Like the last two data, reproductive collocated with isolation is used in biological study. Reproductive isolation is the ability to live things to interbreed with each other, but not with members of other species, which causes speciation. the sentence in data 4 explains that speciation occurs because of reproductive isolation. Just like data 2 and 3, the combination of reproductive and isolation shows positive meaning in biological study.

Left	Kwic	Right
building infrastructure and industrial applications. Eaton 9355 40kVA Uninterruptible Power Supplies Options:	**Isolation**	**transformer,** long-life batteries, external battery cabinets, integrated manual bypass and external maintenance

Data 5 The combination of isolation and transformer

Data 5 shows that isolation is collocated with 'transformer' as a noun. Based on the online oxford dictionary, a transformer is a tool to reduce or increase the voltage of an electric power supply. The sentence in table 5 shows that isolation transformer is an option needed when building infrastructure. The combination of isolation and transformer has a positive meaning.

Collocation of *isolation* in the corpus of covid 19

The node in this study is the word "isolation" followed by several different collocations in the Corpus of Covid 19. Many collocations follow node "isolation" in this corpora, but chosen collocations below are taken from the top five words often used as collocations from the word "isolation". Five collocations were selected based on the highest MI value, as shown in the table below:

Collocate	Freq	Coll. freq.	T-score	MI
Quarantine	1317	11619	36,243	9,5768
Rooms	995	6762	31,5118	9,9533
Precautions	800	7849	28,2431	9,4235
Virus	765	18928	27,5571	8,089
Kit	596	13903	24,3286	8,174

Table 2. The collocation of "isolation" in the corpus of Covid 19

The data in table 2 shows that the word isolation in the corpus of Covid 19 collocates with several words such as quarantine, rooms, precautions, virus, and kit. The word "quarantine" has the highest frequency, with a value of 1317. The second position is the word "rooms," which has a frequency with a value of 995. The third, fourth, and fifth positions are precautions, virus, and kit with frequencies of 800, 765, and 596, respectively.

The following figure will show the application of each collocation of "isolation" in the corpus of covid 19.

Left	Kwic	Right
practice may then be expected to include contact tracing, and monitoring and support of people in **quarantine** or home	**Isolation**	. Other responsibilities may include prescribing and dispensing antiviral drugs and participating in mass
of SARS-CoV-2 was even detected in environmental samples taken from the surface of toilet bowl and sink in infection	**Isolation**	**rooms.** 26 Moreover, SARS-CoV-2 remained viable in the stool of infected patients as reported by some case reports. 18
it will be mandatory to keep a high index of suspicion for cases that fit the general criteria of SARS and to follow	**Isolation**	**precautions** for known cases. SARS has been described by the World Health Organization (WHO) as a rapidly progressive,
each province. Cases of COVID-19 included diagnosed cases confirmed by at least one of the following three methods:	**Isolation**	of COVID-19 **virus** , at least two positive results for COVID-19 virus by real-time reverse-transcription polymerase
guidelines from the World Health Organization based on the use of the MagNA Pure total nucleic acid large volume	**Isolation**	**kit** for the extraction of SARS-coronavirus RNA. The modified protocol was compared with a column-based extraction

Figure 2. Concordance of "isolation" in the Corpus of Covid-19

Based on figure 2, most collocations of 'isolation' are placed on the right, and two collocation is on the left side. According to the online oxford dictionary, *quarantine*, *rooms, RNA*, and *transformer* are nouns. The collocation used by the node isolation in the corpus Covid 19 is part of speech noun. When the node is on the left side, the node act as a modifier to modify noun after it. For example, noun kit is modified by isolation. Another noun is the collocation of the node isolation that always come together using conjunction or.

Left	Kwic	Right
practice may then be expected to include contact tracing, and monitoring and support of people in **quarantine** or home	**Isolation**	. Other responsibilities may include prescribing and dispensing antiviral drugs and participating in mass

Data 6 The combination of quarantine and isolation

Based on the Online Oxford Dictionary, quarantine is when an animal or person has a disease isolated from others to stop spreading the disease. Quarantine and isolation are both nouns. The words isolation and quarantine are collocated to make negative meaning. The sentences in data 6 indicate support needs for people in quarantine or isolation. People have to be quarantined because they are infected with the virus or in contact with people infected with it.

Left	Kwic	Right
of SARS-CoV-2 was even detected in environmental samples taken from the surface of toilet bowl and sink in infection	**Isolation**	**rooms**. 26 Moreover, SARS-CoV-2 remained viable in the stool of infected patients as reported by some case reports. 18

Data 7 The combination of isolation and rooms

Data 7 shows that the word isolation is collocated with 'room'. Based on the online oxford dictionary, the room is a noun and has meaning as a part of a building, own walls, floor, and ceiling, and used for specific occasions. The sentence from data 7 shows that people infected by a virus use an isolation room. The combination of isolation and room creates a negative meaning.

Left	Kwic	Right
it will be mandatory to keep a high index of suspicion for cases that fit the general criteria of SARS and to follow	**Isolation**	**precautions** for known cases. SARS has been described by the World Health Organization (WHO) as a rapidly progressive,

Data 8 The combination of isolation and precautions

Data 8 shows that isolation is collocated with precaution. Both precaution and isolation are nouns. Precaution means to do something in advance to avoid problems or prevent problems. The sentence from data 8 shows that it is necessary to follow isolation precautions to prevent SARS from spreading. The combination of isolation and precaution creates a positive meaning.

Left	Kwic	Right
each province. Cases of COVID-19 included diagnosed cases confirmed by at least one of the following three methods:	**Isolation**	of COVID-19 **virus** , at least two positive results for COVID-19 virus by real-time reverse-transcription polymerase

Data 9 The combination of isolation and virus

The online oxford dictionary says that a virus means a tiny living thing that can not be seen without a microscope and causes disease. Based on data 9, the words isolation and virus are collocated as nouns. The sentence from data 9 shows that an isolation virus is needed to diagnose or identify whether the coronavirus has infected someone. The word virus has a negative meaning, but it creates a positive meaning when collocated with the word isolation.

Left	Kwic	Right
guidelines from the World Health Organization based on the use of the MagNA Pure total nucleic acid large volume	**Isolation**	**kit** for the extraction of SARS-coronavirus RNA. The modified protocol was compared with a column-based extraction

Data 10 The combination of isolation and kit

Data 10 shows that isolation is collocated with kit. Both kit and isolation are nouns. The kit is the equipment used for a specific occasion. The sentence from data 10 shows that an isolation kit is used to extract RNA viruses according to WHO guidelines. The combination of isolation and kit creates a positive meaning.

The Semantic Preference and Semantic Prosody of Node Word 'isolation'
in the corpus of English Web 2018

Semantic preferences are categorized based on collocations in each corpus of data. The semantic node isolation preferences in the corpus English Web 2018 are categorized into three types: emotion, science, and electricity. Loneliness is included in the emotion category. RNA, reproductive, and characterization are included in science category, and transformer is categorized as housing and architecture. Semantic prosody shows that the meaning of a word depends on its relationship to the words that accompany it. Node isolation on the corpus of English web 2018 will show a negative meaning when it collocates with the word loneliness, while node isolation will show a positive meaning if it collocates with the words characterization, RNA, reproductive, and transformer.

The Semantic Preference and Semantic Prosody of Node Word 'isolation'
in the corpus of Covid 19

Semantic preferences are categorized based on collocations in each corpus of data. The semantic node isolation preferences in the corpus of Covid 19 is categorized

into only one type because all collocations are related to Covid 19. Semantic reference from the corpus of Covid 19 data is word and our environment. Semantic prosody shows that the meaning of a word depends on its relationship to the words that accompany it. Node isolation on the corpus of Covid 19 will show a negative meaning when it collocates with the words quarantine, and room, while node isolation will show a positive meaning if it collocates with the words such as precautions, kit, and virus.

Conclusion

This study indicates differences in collocations in the word isolation since the emergence of Covid 19. There are 6 collocations with the highest MI values taken from two different corpora, Corpus English Web 2018 and Covid 19, to determine each collocation's use of the word isolation. The findings of this study in the corpus of English web 2018 are; that one collocation shows the negative meaning, and four collocations indicate positive meaning. While in the corpus covid 19 data are, two collocations show a negative meaning, and three collocations show a positive meaning.

References

Asif, Muhammad, Deng Zhiyong, Anila Iram, and Maria Nisar. 2020. "Linguistic Analysis of Neologism Related to Coronavirus (COVID-19)." *SSRN Electronic Journal* 4(December 2020):2–7. doi: 10.2139/ssrn.3608585.

Bednarek, Monika. 2008. "Semantic Preference and Semantic Prosody Re-Examined."

Begagić, Mirna. 2013. "Semantic Preference and Semantic Prosody of the Collocation Make Sense." *Jezikoslovlje* 14(2–3):403–16.

Cheng, Winnie. 2013. "Semantic Prosody." *The Encyclopedia of Applied Linguistics* 1–7.

Emery, Peter G. 1991. "Collocation in Modern Standard Arabic." *Zeitschrift Für Arabische Linguistik* (23):56–65.

Gelbukh, Alexander, and Olga Kolesnikova. 2012. *Semantic Analysis of Verbal*

Collocations with Lexical Functions. Vol. 414. Springer.

Hu, Hsueh-chao Marcella. 2015. "A Semantic Prosody Analysis of Three Synonymous Adjective Pairs in COCA." *Journal of Language and Linguistic Studies* 11(2):117–31.

Lindquist, Hans. 2018. *Corpus Linguistics and the Description of English*. Edinburgh University Press.

Liu, Weiqian. 2020. "A Corpus-Based Study of Semantic Prosody and Semantic Preference of 'Seem.'" *Open Journal of Social Sciences* 8(5):282–91.

McEnery, Tony, and Andrew Hardie. 2011. *Corpus Linguistics: Method, Theory and Practice*. Cambridge University Press.

Rodríguez, María A. Barrios, and Cliff Goddard. 2013. "'Degrad Verbs' in Spanish and English: Collocations, Lexical Functions and Contrastive NSM Semantic Analysis." *Functions of Language* 20(2):219–49.

Zhang, Changhu. 2010. "An Overview of Corpus-Based Studies of Semantic Prosody." *Asian Social Science* 6(6):190.