Frequency Analysis: Lexical Bundles Comparation in Medical Journal Abstract and Law Journal Abstract

Brama Kumbara Omar and Riskayani Djamaleng[(✉)]

Master of Linguistics Program, Warmadewa University, Denpasar, Indonesia
riskayanidj14@gmail.com

Abstract. Academic writings are widely used as research corpus of lexical bundles because they are productive and have strict writing and systematics, so it is likely that high frequency use of lexical bundles will be found. Based on this, the researchers tried to identify and compare lexical bundles in legal articles and health articles by using a predetermined corpus of articles to see the frequency of their use in both articles. Many researches on lexical bundles have been done before, for example, the research conducted by Haq, et al. (2021) which specifically tried to examine abstracts from PhDs, to determine initial bundles and the function of lexical bundles in PhD abstracts. Thus, this study focuses on identifying and comparing the frequency of lexical bundles in legal articles and health articles, which are more specific in their abstracts. The data used as the corpus in this study came from abstracts from legal journals and health journals. In this study, researchers used 50 abstracts from legal articles, while for comparison data came from 50 abstracts from health. In this study, the number of texts in the abstract data of legal articles amounted to 8,674 words, while the number of texts in the abstract data of health articles amounted to 9,611 words. Through the abstract data presented, the researcher wants to show that there is use of lexical bundles in both types of journals which are then compared with the frequency of their use.

Keywords: Lexical bundles · Antcont · comparation · Frequency analysis

1   Introduction

Nowadays language patterns have become an interesting thing to be studied by linguists in the field of linguistics. In English, one of the language patterns that is quite common to observe is multiword sequences. Basically, these word sequences are words that coincidentally follow each other more often by meeting certain criteria. There are a number of terms used to describe this series of words. [1] use the term "lexicalized sentence stem", [2] uses the word "clusters", and [3] calls it "N-grams" [4]. Meanwhile, Finegan et al. (1999) call this series of words as lexical bundles [5].

The term lexical bundles was first used in the English grammar book "Longman Grammar of Spoken and Writing English" (Chapter 13). This book is written using a

corpus-based approach to studying vocabulary collections in conversational and academic writing. A vocabulary set is defined as an extended collection of collocations that show statistical tendencies that occur together [5]. Collocation is a permanent relationship between a word and another word throughout the sentence [6].

To make it easier to understand, lexical bundles are a series of three or more words that often occur in a register. Register is a language variation based on the domain or field [7]. One example of registers I don't think so, as a result of, and I thought that was. In general, lexical bundles are not idiomatic in terms of meaning and incomplete in terms of structure. In terms of meaning, lexical bundles are not the same as idioms because their meaning can be derived from each word that forms them. Meanwhile, in terms of structure, most of the lexical bundles do not represent a complete structure, either in the form of phrases or clauses.

Academic writings are widely used as research corpus of lexical bundles because they are productive and have strict writing and systematics, so it is likely that high frequency use of lexical bundles will be found. However, apart from academic writing, there are also other types of texts that are productive and have strict writing and systematics because of their enormous influence in their respective fields. Based on this, the researchers tried to identify and compare lexical bundles in legal articles and health articles by using a predetermined corpus of articles to see the frequency of their use in both articles. Many researches on lexical bundles have been done before, for example, the research conducted by Yuliawati et al. (2021) which specifically tried to examine abstracts from PhDs, to determine initial bundles and the function of lexical bundles in PhD abstracts [8]. Thus, this study focuses on identifying and comparing the frequency of lexical bundles in legal articles and health articles, which are more specific in their abstracts. In general, abstracts are simple writings that are not too long, the presentation of which is fast, short, dense, and easily understood by the reader. It is an important part of compiling scientific papers. Abstract is a brief summary of the contents of scientific papers, such as thesis, dissertations, or research reports/journals. This paper is one of the things that stands out in a scientific paper. The presentation is placed in the first part of a written work.

2 Method

This research is a corpus-based linguistic research. In the book Exploring Corpus Linguistics Language in Action (Cheng, 2011), corpus linguistics is a collection of texts that have been compiled for later analysis with a specific purpose [9]. The data used as the corpus in this study came from abstracts from legal journals and health journals. In this study, researchers used 50 abstracts from legal articles sourced from an international journal, Scinedirect, while for comparison data came from 50 abstracts from health articles sourced from an international journal, Mendeley.

In this study, the number of texts in the abstract data of legal articles amounted to 8,674 words, while the number of texts in the abstract data of health articles amounted to 9,611 words. In addition, the search for the lexical cluster words also adjusts to the research, namely three word clusters. This was done because the research was the first research on lexical clusters in English found by the researcher so that the researcher made it a reference for conducting this research. AntConc will return results based

on those constraints. Furthermore, an analysis of these results is carried out using an inductive approach. The steps taken are grouping lexical groups based on their frequency of occurrence, structure, and function. The following is a recapitulation of the number of texts and words used as the corpus of data in this study.

3   Result and Discussion

In this section, the frequency of lexical bundles has been calculated automatically using the antcont application. From the data provided by the researcher, there are 50 abstracts of legal articles and 50 abstracts of health articles. From these data, it was found that the comparison of the number of different words by the two abstract journals was found. Table 1 provides information about this comparison.

Through the abstract data presented, the researcher wants to show that there is use of lexical bundles in both types of journals which are then compared with the frequency of their use. After determining the articles and keywords that are used as comparisons, the researcher makes a table to describe them. The range shows the amount of texts that use the bundle. To find the comparison between the abstract of the law journal and medical journal, the researcher sorts the two journals by determining the articles and keywords that will be used to determine the bundles contained in legal journal abstracts and medical journal abstracts.

Table 2 shows that the lexical bundles contained in legal journal abstracts. The data above shows that from the articles that have been determined by the researcher, only a few articles are contained in the abstract of the law journal, namely, the, with, that, this, in, at, as, an/a, can, for, is, are, it., and law. From the data above, it shows that phrases with articles the, in, with which have the highest frequency and phrases with articles this, an, and law are the lowest use of articles.

Table 3 provides information about the use of lexical bundles in abstracts of health journals using the corpus. The data above shows that the articles as, to, in, health are the articles with the highest number of frequencies. Then, the articles with the lowest frequency were generated by the, to, in, health, is, are, and this with the same number of frequencies, namely 3.

Table 1. The size of the scopus

No	Disciplines	Number of articles	Number of words
1.	Law journal Abstract	50	8.674
2.	Medical Journal Abstract	50	9.611

Table 2. The frequency of lexical bundles in the law journal abstract

Lexical Bundles of Law Abstract Journal	Occurrences
The rule of	10
In conflict with	8
With the law	8
In order to	6
The development of	6
The number of	6
A power law	5
As well as	5
In public restrooms	3
In this paper	3
As a whole	3
An annual mean	3
A deviation from	3
A dual state	3
Can be used	3
For gay men	3
Has become a	3
Is associated with	3
Are in conflict	3
It has been	4
The house of	4
The omnibus act	4
The role of	4
The united status	4
Law in the	4
The creation of	3
The criminal law	3
The distribution of	3
The effects of	3
The employment of	3
The European union	3
The importance of	3

(*continued*)

Table 2. (*continued*)

Lexical Bundles of Law Abstract Journal	Occurrences
The law is	3
The Malaysian common	3
The need for	3
The pre-law	3
The rise of	3
The study of	3
The supreme court	3
The use of	3
That it is	3
This article discusses	3
This article examines	3
This article explores	3
This study is	3
Law enforcement offices	3
Law is a	3
Law on the	3
Law period to	3
An annual mean	3

Table 4 shows combined data between lexical bundles in legal journal abstracts and health journal abstracts. From the data then look for the comparison between the two. Comparison of lexical bundles of legal journal abstracts and medical journal abstracts. From the results of the comparison above, it can be concluded that there are several articles that appear in legal journal abstracts and health journal abstracts, namely as well as, in order to, is associated with, the need for. This comparison reflects that the lexical bundles contained in legal and health abstracts have a unique identity in the choice of words, especially in the selection of the use of articles. Although there was a common bundle used in both corpora, a typical bundle with high frequency and range scores was found to characterize the different groups of authors. The typical lexical bundles found are not coincidental but they show that there is a pattern of bundle use in a particular group of writers and disciplines namely linguistics.

Scalar *variables* and *physical constants* should be italicized, and a bold (non-italics) font should be used for **vectors** and **matrices**. Do not italicize subscripts unless they are variables. Equations should be either display (with a number in parentheses) or inline. Use the built-in Equation Editor or MathType to insert complex equations.

Table 3. The frequency of lexical bundle in medical journal abstract

Lexical Bundles of Medical Abstract Journal	Occurrences
As well as	6
The island ecosystem	6
The mind me	6
To mental health	5
In mental health	5
In the field	5
Health care professionals	5
The field of	4
The public health	4
Health care needs	4
Health literacy is	4
Health system responses	4
In order to	4
In this population	4
In health care	4
The extent of	3
The health impacts	3
The importance of	3
The need for	3
The Social determinants	3
To account for	3
To achieve the	3
To access the	3
To health care	3
In need of	3
In public health	3
In relation to	3
Health education programs	3
Health promotions behaviors	3
Is associated with	3
Are less likely	3
This paper examines	3

Table 4. The comparation of the lexical bundles of law journal abstract and medical law journal abstract

Lexical Bundles	Medical	Law
A deviation from		✓
A dual state		✓
A power law		✓
An annual mean		✓
An annual mean (Law)	✓	
Are in conflict		✓
Are less likely	✓	
As a whole		✓
As well as	✓	✓
Can be used		✓
For gay men		✓
Has become a		✓
The criminal law	✓	
Health care professionals	✓	
Health education programs	✓	
Health literacy is	✓	
Health promotions behaviors	✓	
Health system responses	✓	
In conflict with		✓
In health care	✓	
In mental health	✓	
In need of	✓	
In order to	✓	✓
In public health	✓	
In public restrooms		✓
In relation to	✓	
In the field	✓	
In this paper		✓
In this population	✓	
Is associated with	✓	✓
It has been		✓
Law enforcement offices		✓
Law in the		✓
Law is a		✓
Law on the		✓
Law period to		✓
That it is		✓
The creation of		✓

4   Conclusion

Academic writings are widely used as research corpus of lexical bundles because they are productive and have strict writing and systematics, so it is likely that high frequency use of lexical bundles will be found. Based on this, the researchers tried to identify and compare lexical bundles in legal articles and health articles by using a predetermined corpus of articles to see the frequency of their use in both articles. Many researches on lexical bundles have been done before, for example, the research conducted by Haq, et al. (2021) which specifically tried to examine abstracts from PhDs, to determine initial bundles and the function of lexical bundles in PhD abstracts. Thus, this study focuses on identifying and comparing the frequency of lexical bundles in legal articles and health articles, which are more specific in their abstracts. The data used as the corpus in this study came from abstracts from legal journals and health journals. In this study, researchers used 50 abstracts from legal articles sourced from an international journal, Scinedirect, while for comparison data came from 50 abstracts from health articles sourced from an international journal, Mendeley. In this study, the number of texts in the abstract data of legal articles amounted to 8,674 words, while the number of texts in the abstract data of health articles amounted to 9,611 words. This was done because the research was the first research on lexical clusters in English found by the researcher so that the researcher made it a reference for conducting this research. From the data provided by the researcher, there are 50 abstracts of legal articles and 50 abstracts of health articles. Through the abstract data presented, the researcher wants to show that there is use of lexical bundles in both types of journals which are then compared with the frequency of their use. To find the comparison between the abstract of the law journal and medical journal, the researcher sorts the two journals by determining the articles and keywords that will be used to determine the bundles contained in legal journal abstracts and medical journal abstracts. The data above shows that the articles as, to, in, health are the articles with the highest number of frequencies.

References

1. Pawley, Andrew and Frances Hodgetts Syder. 1983."Two puzzles for linguistic theory: Nativelike selection and nativelike fluency." In: Jack J. C. Richards.
2. Scott, M. (2001). Comparing corpora and identifying key words, collocations, and frequency distributions through the Word Smith Tools suite of computer programs. Small Corpus Studies and ELT, 47–67.
3. Stubbs, S. (2002). Inclusive Education: Where there are few resources.
4. Novita, H., & Kwary, D. A. (2018). Comparing the use of lexical bundles in Indonesian- English translation by student translators and professional translators. The International Journal of Translation and Interpreting Research, 10(1), 53–74.
5. Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). Longman Grammar of Spoken and Written English. England: Pearson Education Limited.
6. Kridalaksana, H. (2013). Kamus linguistik (4th ed.). Gramedia Pustaka Utama.
7. Chaer, A. (2011). Tata Bahasa Indonesia Praktis. Renika Cipta.

8. Yuliawati, azka saeful haq; R. M. A. S. (2021). Lexical Bundles of Indonesian and English Research Articles: Frequency Analysis. JALL (Journal of Applied Linguistics and Literacy), 5(1).
9. Cheng, W. (2011). Exploring corpus linguistics. Routledge.