

A Corpus-Based Analysis of Prepositional Phrase-Lexical Bundles in Linguistics Research Articles Written by Indonesian Writers

Nurinsani Yaman¹✉, Ni Gusti Ayu Roselani²

^{1,2}Master of Linguistics, Universitas Gadjah Mada, Indonesia

Article Info	Abstract
<i>Article History:</i> Received 04 March 2026 Approved 13 April 2026 Published 30 April 2026	This study examined the frequency and functional categories of 4-word PP-based bundles in linguistics research articles employed by Indonesian writers. The data were taken from a corpus comprising 116 articles with 602,421 tokens. The PP-based bundles were identified with the assistance of AntConc. Cut-off frequency and dispersion were applied to generate the list of bundles from the corpus. The identified bundles were then classified into functional categories proposed by Hyland. The findings reveal that only two out of three categories are identified: research-oriented and text-oriented. In terms of types, research-oriented and text-oriented bundles occur in nearly identical proportions. However, text-oriented bundles occur more frequently when token frequency is considered. Within the text-oriented category, framing signals emerge as the most dominant subcategory. Participant-oriented bundles are entirely absent, a result that may be influenced by the structural focus of the study.
Keywords: Academic Writing, Corpus Linguistics, English Preposition, Indonesian Writers, Lexical Bundles	

© Copyright 2026

✉ Corresponding author:
B3 Building, Kampus Sekaran, Gunungpati Semarang, Indonesia 50229
E-mail: nurinsaniyaman@mail.ugm.ac.id

INTRODUCTION

English prepositions are one of the most challenging aspects for many L2 English learners. Unlike many other grammatical categories, prepositions are characterized by their high frequency of use and have notoriously polysemous behaviors in sentences (Saravanan, 2015). In the Corpus of Contemporary American English (COCA) (Davies, 2010). Some prepositions occur more frequently than the most frequent noun, and the few extremely frequent ones are multifunctional (Schneider et al., 2016). Therefore, L2 English learners usually struggle to use English prepositions accurately in context.

In the academic writing context, one possible way to improve accuracy in the use of prepositions is through formulaic language that

features prepositions, particularly lexical bundles (Geluso, 2022). Lexical bundles are sequences of three or more words that co-occur frequently in a register (Cortes, 2004). Some examples of lexical bundles in academic prose are the context of the, on the other hand, and as a result of. Biber et al. (2021) stated that these word combinations are not idiomatic, which means that their meaning is transparent. In their study, it is also mentioned that most of lexical bundles do not represent a complete structural unit. Biber et al. (2021) found that only less than 5% of lexical bundles represent complete structural units.

Among different structural types of lexical bundles, prepositional phrase-based (hereafter PP-based) are especially prominent in academic writing. Previous studies on lexical bundles suggest that the most frequently used lexical bundles in

academic writing are parts of noun or prepositional phrases (Biber et al., 2021; Hyland, 2008). The use of PP-based bundles in academic writing is important to enhance the writing quality. They are effective to increase information density in writing (Geluso, 2022) and serve rhetorical functions (Lu et al., 2021). In addition, PP-based bundles can also help writers to develop their academic writing proficiency (Geluso, 2022). However, mastering such bundles can be particularly challenging for L2 English writers as it requires sensitivity to both grammatical and discourse functions.

Many studies have extensively highlighted challenges encountered by L2 English writers in employing lexical bundles (Bychkovska & Lee, 2017; Chen & Baker, 2016; Geluso, 2022; Pan & Liu, 2019). Bychkovska & Lee (2017), for instance, compared L1 English and L1 Chinese undergraduate students' use of lexical bundles in argumentative essays and found that the L1 Chinese students rely more on VP-based bundles than PP-based bundles, unlike their counterparts. Additionally, they found several misused lexical bundles produced by L1 Chinese students and most of the misused bundles in their writing are related to grammatical mistakes, particularly with prepositions (e.g., in the same time) and articles (e.g., the large amount of). Similarly, in their comparison of preposition-based phrase frames (PP-frames) in bachelor's theses, master's theses, and doctoral dissertations in the field of Applied Linguistics among Chinese EFL learners, Zhou & Li (2024) found that Chinese EFL learners use PP-frames less frequently and often make errors in their writing. However, the learners with a higher degree (PhD level) use a higher numbers and more variable PP-frames than their counterparts. This can happen because the PhD students have extensive writing experience and exposure to the target PP-frames.

In more recent research, Yunjung (2025) specifically examined the use of PP-based lexical bundles among three different groups: English learners from Indo-European (IE) languages, non-Indo-European (NIE) languages, and native speakers (NS) of English. The result of this study revealed a tendency among non-native speaker

groups (IE and NIE) to use a high frequency with a more limited variety of PP-based bundles compared to the NS group. This suggest that the non-native speakers often rely on the lexical bundles they are already familiar with. Additionally, the non-native speaker groups favored colloquial expression, contrasting with NS group's preference for academic and formal language. This finding lends support to Lee et al.'s (2020) research. They analyzed the English preposition in lexical bundles in essays written by Korean university students and found that the students rely heavily on a small number of PP-based bundles and underuse those that are characteristic of academic prose.

In the Indonesian context, many studies have examined the use of lexical bundles in academic writing, such as theses (Fitrianasari et al., 2018; Samodra & Pratiwi, 2018; Wachidah et al., 2020), essays (Oktavianti & Prayogi, 2022; Oktavianti & Sarage, 2021; Ulfa & Muthalib, 2020), and research articles (Fajri et al., 2020; Insani & Hardjanto, 2022; Insani & Roselani, 2023). These previous studies on lexical bundles focus on the structures and functions of lexical bundles. The findings of many of these studies consistently indicate that Indonesian writers still struggle to use the PP-based bundles in their writing. Despite the recurring identification of this problem, research that specifically analyzes PP-based structure remains limited. To address this gap, this study focuses on examining the four-word PP-based bundles and its functions among the Indonesian writers. The findings are expected to contribute to a better understanding of Indonesian writers' fluency and competence in utilizing PP-based bundles. Therefore, this research attempts to answer these questions:

1. What are the most frequent four-word PP-based bundles in the linguistics research articles written by Indonesian writers?
2. What are the functions of four-word PP-based bundles in the linguistics research articles written by Indonesian writers?

METHODS

Description of the corpus

The corpus used in this study, referred to as INCO (Indonesian Corpus), consists of 602,421 tokens compiled from 116 research articles published from 2024 to 2025. All articles were derived from five Indonesian linguistics journals (see Table 1). These journals were selected because they hold accreditation from the Indonesian Ministry of Research, Technology, and High Education (Sinta 1-3) and only publish articles in English. Regarding the writers, they were classified as Indonesian writers based on two criteria: (1) their first and last name are considered as typical Indonesian names, and (2) they are affiliated with Indonesian institutions. Thus, any articles that did not meet both criteria were excluded from the corpus.

Table 1. Data Sources

No	Journals	Rank
1	Script Journal: Journal of Linguistics and English Teaching	SINTA 1
2	Jurnal Arbitrer	SINTA 1
3	NOTION: Journal of Linguistics, Literature, and Culture	SINTA 2
4	JoPR: Journal of Pragmatics Research	SINTA 3
5	Lexicon	SINTA 3

Identification of PP-based bundles

This study investigated four-word PP-based bundles. These four-word bundles are more prevalent and offer clearer structures and functions compared to three-word bundles (Chen & Baker, 2010; Hyland, 2008). To generate a list of lexical bundles, this study employed two criteria. The first criteria is the cut-off frequency, which determines the number of lexical bundles to be included in the analysis. The second criteria is dispersion, which refers to the distribution of bundles in multiple texts in the corpus. Both the frequency and dispersion threshold vary from study to study because it depends on the corpus size. Biber et al. (2021) set a minimal cut-off of at least 10 times per million words across 5 texts for a four-word sequence to be considered a lexical bundle. However, the corpus in this study is relatively small and only examined one structural classification (i.e., PP-based bundles). Therefore,

this study set a lower threshold of 7 times across 3 different texts.

The PP-based bundles were extracted using the n-gram tool in AntConc 4.3.1 (Anthony, 2024). The retrieved bundles were then manually checked. Any context-dependent bundles such as *from the Jakarta Post, in the United States, as a foreign language, on social media platforms*, were excluded from the analysis.

Data Analysis

The data that have been collected were classified based on the functional taxonomy proposed by Hyland (2008). The framework is divided into three broad categories (i.e., research-oriented, text-oriented, and participant-oriented), with subcategories grouping more specific functions. This study employed Hyland's taxonomy because it aligns closely with the academic writing context.

A single lexical bundle may have multiple functions. In cases where bundles displayed multifunctionality, classifying them according to their most frequent use might be considered good practice (Biber et al., 2004). Therefore, the PP-based bundles in the present study were also categorized under the functional category in which they were most frequently used. To determine the primary function, the concordance lines of all potentially multifunctional bundles were examined. After assigning all the PP-based bundles to their functional category and subcategory, the analysis proceeded to examine the frequency of types and tokens of each functional category. The findings were then compared to the previous studies on lexical bundles.

RESULTS AND DISCUSSION

Results

Frequency of four-word PP-based bundles used by Indonesian writers

Based on the identification, there are 99 four-word PP-based bundles (1909 tokens) found in the Linguistics research articles written by Indonesian writers. Table 2 presents the 20 most

frequent bundles found in INCO along with their frequency and range.

Table 2. Top 20 PP-based bundles

No	PP-based bundles	Freq	Range
1	in the form of	163	55
2	in the context of	131	60
3	on the other hand	90	48
4	in the final position	64	3
5	as a form of	53	31
6	at the end of	45	22
7	in the field of	42	21
8	at the same time	38	25
9	in line with the	33	26
10	in the use of	33	25
11	in this case the	33	15
12	as a means of	33	23
13	in this study is	32	21
14	of this study is	28	25
15	of this research is	26	19
16	as a result of	25	16
17	as part of the	24	18
18	to the use of	24	13
19	in the face of	23	12
20	of the use of	23	13

Among all the PP-based bundles, *in the form of* and *in the context of* hold the highest number in terms of both frequency and range. The two bundles occur over a hundred times in more than 50 different texts in the INCO corpus. These figures indicate that both bundles are not only frequent but also widely distributed across different writers.

In PP-based bundles, there are two main categories that can be identified: prepositional phrase with embedded of-phrase and other prepositional phrase (fragment). Table 2 shows that these two categories appear in a relatively balanced proportion. This result is also reflected in the overall dataset. Among the 99 PP-based bundles, 52 bundles contain an embedded of-phrase, whereas the remaining 47 belong to the second category (see Table 3).

Table 3. Frequency of PP-based bundles

PP-based	Types	Tokens
Prepositional phrase with embedded of phrase	52	1067
Other prepositional phrase (fragments)	47	770
Total	99	1837

In addition to the frequency and structural analysis, the data also reveal the presence of overlapping words across multiple bundles. These overlaps occur when the same initial word sequence forms the base of several related bundles that differ only in their final word (except for *the use of* which starts with 8 different prepositions). Table 3 lists the PP-based bundles with overlapping words in INCO.

Table 4. PP-based bundles with overlapping words

Overlapping words	PP-based bundles	Tokens
in this research (3)	in this research is	11
	in this research were	9
	in this research are	8
in this study (4)	in this study is	31
	in this study are	10
	in this study was	9
	in this study were	8
as a result (2)	as a result of	25
	as a result the	16
as a tool (2)	as a tool for	20
	as a tool to	20
of this study (3)	of this study is	28
	of this study are	13
	of this study lies	8
of this research (2)	of this research is	26
	of this research are	9
the use of (8)	in the use of	33
	to the use of	24
	of the use of	24
	on the use of	19
	with the use of	19
	by the use of	11
	as the use of	9
	for the use of	7
Total 24 types (397 tokens)		

As shown in the table, among the 99 PP-based bundles used by Indonesian writers in INCO, approximately 24% (24 bundles) shared some overlapping words with one to seven other bundles. The total token of these overlapping words was 397, which represents about 22% of the total token counts in INCO. This finding suggests that Indonesian writers employed a more diverse PP-based bundles in their writings because nearly 80% of bundle use consists of different expressions rather than repeated instances of the same phrases.

Functions of PP-based bundles

The functional classification used in this study was the one developed by Hyland (2008).

The lists of PP-based bundles in the table below are organized alphabetically by preposition.

Table 5. Functional categories and subcategories

Category	Subcategory	PP-based bundles
Research-oriented	Location	at the end of, at the same time, at the beginning of, for a long time, in the final position, in this study is, in the initial position, in this research is, in this study are, in the study is, in this research were, in this study was, in this research are, in this study were.
	Procedure	as a means of, as a tool for, as a tool to, as a means to, as a way to, as the use of, as a medium for, by the use of, by looking at the, for the use of, in the use of, in a way that, in the process of, in analyzing the data, in the implementation of, of the use of, on the use of, on the analysis of, on the role of, to the use of, to the development of, with the use of, with the help of.
	Quantification	as part of the, as one of the, as part of a, in a variety of, with a total of.
	Description	as a form of, as the object of, as a source of, as a reflection of, as a symbol of, as a type of, of the relationship between, of the role of.
Text-oriented	Transition signals	in line with the, in other words the, in accordance with the, in addition there is, in addition to that, in line with this, in addition this study, on the other hand, on the one hand.
	Resultative signals	as a result of, as a result the, from the results of.
	Structuring signals	in the following table, in the following examples, in the example above, in the table below.
	Framing signals	as a sign of, as the foundation for, for this research is, for the benefit of, from the perspective of, in the form of, in the context of, in the field of, in the face of, in this case the,

in terms of the, in the aspect of, in the case of, in relation to the, in the realm of, in the fields of, in this case is, of this study is, of this research is, of this study are, of this research are, of the research is, of this study lies, on the results of, on the quality of, on the representation of, to the fact that, to the lack of, to the field of, with a focus on, with the aim of, within the context of, within the framework of.

Participant-oriented	Stance features	-
	Engagement features	-

To provide a more in-depth discussion, each category (research-oriented, text-oriented, and participant-oriented) is presented below with examples drawn from INCO.

1. Research-oriented

Research-oriented bundles help writers to present and discuss the content in texts. There are four subcategories in this function and each of them is presented below.

1.1 Location

This function refers to bundles that indicate time or place in the texts. For examples, bundles *at the end of* and *at the same time* are used for locating events temporally as in (1) and (2). There were also bundles referring to external textual location such as in (3) and (4).

- (1) The observation will be carried out *at the end of* July 2024.
- (2) ...that you can do two or more jobs *at the same time*, resulting in multiple benefits.
- (3) ...the vowels that experience changes are located *in the final position* of the word...
- (4) ...the message will be considered polite or polite enough if there is a greeting *at the beginning of* the message...

1.2 Procedure

Bundles in this subcategory indicate how something is done or what something is for. Some variants of PP-based bundles found in INCO were *by the use of* and *with the help of*.

- (5) This particular role is often signaled explicitly *by the use of* verbs like see, note, and observe.
- (6) The chi-square tests in the present research were calculated *with the help of* statistical software...

1.3 Quantification

Quantification bundles indicate the quantity of something. Some of the examples in INCO were *as part of the* and *as one of the*.

- (7) ...the identity of the Sasak people *as part of the* Austronesian ethnolinguistic family.
- (8) *Merantau* is described *as one of the* stages in the life of Minangkabau men to...

1.4 Description

Bundles in this subcategory are to describe an abstract or physical description of something. For instance, *as a form of* and *of the relationship of*.

- (9) According to Hamka, as a Muslim scholar, he defines love *as a form of* mercy, based on sincerity and purity, as it is a blessing from God.
- (10) ...*gunungan hasil bumi* is a symbol *of the relationship between* humans and nature and also humans and God.

2. Text-oriented

Text-oriented bundles help writers to organize the text and its meaning as a message or arguments. All four subcategory of text-oriented are presented below.

2.1 Transition signals

Bundles in this subcategory are used to establish additive (11) or contrastive (12) between elements in texts.

- (11) *In addition to that*, external modifications were also identified.
- (12) Boys are usually portrayed as active, adventurous, and forceful, sometimes taking on leadership roles or embarking on

risky missions. Girls, *on the other hand*, are frequently represented as passive, focused on domestic activities, and merely supporting male characters.

2.2 Resultative signals

Resultative signals mark inferential or causative relations between element. One of the examples from INCO was *as a result of*.

- (13) Many viewers reevaluated the standards that society upholds *as a result of* this exposure.

2.3 Structuring signals

Text-reflective markers which organize stretches or discourse or direct reader elsewhere in the text. Some bundles found in INCO were *in the following table* and *in the example above*.

- (14) The metabases or prosodic terms of gender issues found in the national media are presented *in the following table*.
- (15) The usage of the faithful translation method is evident *in the example above*.

2.4 Framing signals

Framing signals are bundles which situate statements by specifying a context or limiting conditions. This subcategory comprised a high portion of text-oriented bundles in INCO. Some examples found in INCO were *in the form of* and *in the context of*. These two examples were in the top two most frequent PP-based bundles found in the corpus (see Table 2).

- (16) It often appears at the end of the website pages *in the form of* clear call-to-action statements,
- (17) ...examination of the relationship between language and power, particularly *in the context of* media narratives surrounding Indonesian democracy.

3. Participant-oriented

Participant-oriented bundles focus on the writer or reader of the text. There are two subcategories in this function: Stance features and Engagement features. Stance features are bundles that express the writer's attitude and evaluation,

while engagement features are bundles that address readers directly. Both stance and engagement bundles were absent in INCO.

To provide a comprehensive discussion, the frequency of type and token of each functional categories and subcategories were provided in the table below, except for the stance and engagement features in participant-oriented category that are not found in INCO.

Table 6. Frequency of functional category

Category	Subcategory	Types	%	Tokens	%
Research-oriented	Location	14	14.14	285	15.51
	Procedure	23	23.23	331	18.02
	Quantification	5	5.05	75	4.08
	Description	8	8.08	124	6.75
	Total	50	50.51	815	44.37
Text-oriented	Transition signals	6	6.06	143	7.78
	Resultative signals	3	3.03	49	2.67
	Structuring signals	4	4.04	41	2.23
	Framing signals	36	36.36	789	42.95
	Total	49	49.49	1022	55.63
Participant-oriented	Stance features	0	0.00	0	0.00
	Engagement features	0	0.00	0	0.00
	Total	0	0.00	0	0.00

As seen in Table 6, the data indicate a balanced distribution of types and tokens between the first two categories (i.e., research-oriented and text-oriented). Within the research-oriented, procedure bundles are the most frequent in term of types frequency, then followed by location bundles. Description and quantification appear less frequently. In text-oriented category, framing signals is the subcategory that makes up the largest proportion both in types and tokens frequency. The second most frequent is transition signals. Notably, this subcategory shows a relatively small number of types but a higher proportion of tokens. Meanwhile, the other two subcategories, resultative and structuring signals, account for only a share of both types and tokens.

Discussion

This present study attempts to investigate four-word PP-based bundles in linguistics research articles produced by Indonesian writers. The first research question of this study identified the frequency of PP-based bundles. The results showed that *in the form of* and *in the context of* are the most frequent PP-based bundles used by the writers. Structurally, *in the form of* and *in the context of* follow a PP with embedded of-phrase pattern, and both begin with the preposition *in*. PP-based bundles beginning with preposition *in* are particularly frequent in INCO (39 out of 99). Interestingly, the majority of these bundles (25 out of 39) are classified under “other prepositional phrase (fragment)” pattern rather than the PP with embedded of-phrase. This result is in accordance with Biber et al. (2021), who stated that lexical bundles beginning with the preposition *in* are especially common within that pattern. Several of them are used to identify a particular location or time and others are used to specify a particular context.

The second research question of the study examined the functions of PP-based bundles written by Indonesian writers. The result reveals that only research-oriented and text-oriented bundles are present in INCO, while participant-oriented bundles are entirely absent (see Table 5). In terms of types, research-oriented bundles are slightly more frequent than text-oriented bundles. This finding is in accordance with the previous study conducted by Oktavianti & Prayogi (2022), who found that research-oriented bundles made up the largest proportion in the writing of Indonesian EFL learners. In contrast, Insani & Hardjanto (2022) reported that Indonesian writers tend to rely on text-oriented bundles in their writing. The contrasting finding can also be found in the study conducted by Fajri et al. (2020), who claimed that text-oriented bundles are significantly used in applied linguistics journal articles. Despite this discrepancy, it is important to note that the types difference observed between research-oriented and text-oriented in this present study is only one percent. This small percentage suggest that both categories are nearly equally represented in terms of types. Interestingly, when token frequency is

examined, text-oriented appears more frequently than research-oriented.

In addition, the majority of overlapping bundles found in INCO fall within this category. Similarly, Yunjung (2025) also found that most of overlapping lexical bundles consisted of research-oriented bundles (or referential bundles). This indicates that the shared recurrent expressions across texts are mostly associated with describing the time, place, procedures, and quantity.

Text-oriented bundles are particularly dominant in terms of tokens, especially framing signals. Unlike previous studies that claim that L2 writers employed framing signals less frequently than L1 writers (Bychkovska & Lee, 2017; Pan & Liu 2019), this study shows the opposite. This result resonates with findings from Fajri et al. (2020), which revealed that framing signals are used by Indonesian professional scholars more frequently than any other subcategories. The high number of framing signals might be related to the field under study, linguistics. In this kind of discipline, writers have to pay greater attention to contextualizing their research and specifying limitations. This tendency echoed in Hyland's (2008) findings, which identified about 50% of text-oriented bundles in social science texts worked to frame arguments by highlighting connections, specifying conditions, or pointing to limitations.

Regarding participant-oriented bundles, Hyland (2008) argued that some functional categories are strongly associated with structural patterns. For examples, NP + of-phrase structures are prominent in research-oriented functions, PP-based structures in text-oriented functions, and anticipatory it, which is under VP-based structures, frequently occur in participant-oriented functions. Therefore, the absence of participant-oriented bundles in this study is not unexpected since this study focuses only on PP-based, a structure that Hyland associates primarily with text-oriented functions.

CONCLUSION

This study attempted to analyze the frequency and functions of four-word PP-based

bundles in linguistics research articles produced by Indonesian writers. In terms of frequency, of the 104 bundles retrieved, *in the form of* and *in the context of* stand out as the most frequently used PP-based bundles in INCO. Functionally, the findings show that research-oriented and text-oriented bundles are present in the corpus, whereas participant-oriented bundles are entirely absent. In terms of types, research-oriented bundles (50.51%) and text-oriented bundles (49.49%) is almost balanced, differing by one percentage point. This indicates that the variety of PP-based bundles used by Indonesian writer for presenting research content and for organizing the text is nearly equal. However, when frequency of occurrence is considered, text-oriented bundles are more dominant (55.63%) compared to research-oriented bundles (44.37%). This contrast suggests that bundles used to organize the discourse appear more frequently throughout the texts. More specifically, framing signals emerge as the most prominent subcategory. The high frequency (types and tokens) of framing signals in INCO is closely connected to the disciplinary context of the study. Finally, the absence of participant-oriented bundles can be understood in relation to the structural factor. As the study focuses on PP-based bundles, it naturally captures functions commonly realized through this structure.

REFERENCES

- Anthony, L. (2024). *AntConc (Version 4.3.1 [Computer Software]* [Computer software]. <https://www.laurenceanthony.net/software/AntConc>
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at ...: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25(3), 371–405. <https://doi.org/10.1093/applin/25.3.371>
- Biber, D., Johansson, S., Leech, G. N., Conrad, S., & Finegan, E. (2021). *Grammar of spoken and written English*. John Benjamins Publishing Company. <https://doi.org/10.1075/z.232>
- Bychkovska, T., & Lee, J. J. (2017). At the same time: Lexical bundles in L1 and L2 university student argumentative writing. *Journal of English for Academic Purposes*, 30, 38–52. <https://doi.org/10.1016/j.jeap.2017.10.008>

- Chen, Y.-H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology*, 14(2), 30–49.
- Chen, Y.-H., & Baker, P. (2016). Investigating criterial discourse features across second language development: Lexical bundles in rated learner essays, CEFR B1, B2 and C1. *Applied Linguistics*, 37(6), 849–880. <https://doi.org/10.1093/applin/amu065>
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23(4), 397–423. <https://doi.org/10.1016/j.esp.2003.12.001>
- Davies, M. (2010). *The Corpus of Contemporary American English (COCA)* [Dataset]. <https://www.english-corpora.org/coca/>.
- Fajri, M. S. A., Kirana, A. W., & Putri, C. I. K. (2020). Lexical bundles of L1 and L2 english professional scholars: A contrastive corpus-driven study on applied linguistics research articles. *Journal of Language and Education*, 6(4), 76–89. <https://doi.org/10.17323/jle.2020.10719>
- Fitrianasari, N. I., Sulistyorini, T., & Sugiarti, T. R. (2018). Lexical bundles in academic writing by undergraduate and graduate students of English language education program. *Loquen: English Studies Journal*, 11(2), 1–14. <https://doi.org/10.32678/loquen.v11i02.1102>
- Geluso, J. (2022). Grammatical and functional characteristics of preposition-based phrase frames in English argumentative essays by L1 English and Spanish speakers. *Journal of English for Academic Purposes*, 55, 101072. <https://doi.org/10.1016/j.jeap.2021.101072>
- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27(1), 4–21. <https://doi.org/10.1016/j.esp.2007.06.001>
- Insani, I., & Hardjanto, T. D. (2022). An examination of lexical bundles in linguistics' published research articles of by Indonesian writer. *International Journal of English Linguistics, Literature, and Education (IJELLE)*, 4(2). <https://doi.org/10.32585/ijelle.v4i2.2967>
- Insani, I., & Roselani, N. G. A. (2023). Bundle-move analysis in research article of science and technology: Connecting lexical bundles and rhetorical move. *SULUK: Jurnal Bahasa, Sastra, Dan Budaya*, 5(1), 109–130. <https://doi.org/10.15642/suluk.2023.5.1.109-130>
- Lee, Y. E., Yoo, I. W., & Shin, Y. K. (2020). The use of English prepositions in lexical bundles in essays written by Korean university students. *Journal of English for Academic Purposes*, 45, 100848. <https://doi.org/10.1016/j.jeap.2020.100848>
- Lu, X., Yoon, J., & Kisselev, O. (2021). Matching phrase-frames to rhetorical moves in social science research article introductions. *English for Specific Purposes*, 63–83. <https://doi.org/https://doi.org/10.1016/j.esp.2020.10.001>
- Oktavianti, I. N., & Prayogi, I. (2022). Discourse functions of lexical bundles in Indonesian EFL learners' argumentative essays: A corpus study. *Studies in English Language and Education*, 9(2), 761–783. <https://doi.org/10.24815/siele.v9i2.23995>
- Oktavianti, I. N., & Sarage, J. (2021). Lexical bundles in students' argumentative essays: A study of learner corpus. *Indonesian Journal of EFL and Linguistics*, 6(2), 509–534. <https://doi.org/10.21462/ijefl.v6i2.421>
- Pan, F., & Liu, C. (2019). Comparing L1-L2 differences in lexical bundles in student and expert writing. *Southern African Linguistics and Applied Language Studies*, 37(2), 142–157. <https://doi.org/10.2989/16073614.2019.1625276>
- Samodra, M. C., & Pratiwi, V. D. R. (2018). Lexical bundles in Indonesian and English undergraduate thesis abstracts. *Proceedings of the Fourth Prasasti International Seminar on Linguistics (Prasasti 2018)*. Fourth Prasasti International Seminar on Linguistics (Prasasti 2018). <https://doi.org/10.2991/prasasti-18.2018.35>
- Saravanan, J. (2015). The use of English prepositions: An empirical study. *Journal of NELTA*, 19(1–2), 158–168. <https://doi.org/10.3126/nelta.v19i1-2.12089>
- Schneider, N., Hwang, J. D., Srikumar, V., Green, M., Suresh, A., Conger, K., O'Gorman, T., & Palmer, M. (2016). A corpus of preposition supersenses. *Proceedings of the 10th Linguistic Annotation Workshop Held in Conjunction with ACL 2016 (LAW-X 2016)*, 99–109. <https://doi.org/10.18653/v1/W16-1712>
- Ulfa, N., & Muthalib, K. A. (2020). Lexical bundles in students' essay writing. *Englih Education Journal*, 11(3), 367–379.
- Wachidah, W. D. N. A., Fitriati, S. W., & Widhiyanto, W. (2020). Structures and functions of lexical bundles in findings and discussion sections of graduate students' thesis. *English Education* p-ISSN: 2252-6323 e-ISSN: 2721-4540

- Journal*, 10(2), Article 2.
<https://doi.org/10.15294/eej.v10i1.33994>
- Yunjung, K. (2025). A corpus analysis of prepositional phrase-lexical bundles in academic writing: L2 writers from Indo-European and Non-Indo-European languages. *Applied Corpus Linguistics*, 5(2), 100128.
<https://doi.org/10.1016/j.acorp.2025.100128>
- Zhou, H., & Li, T. (2024). "In the present study": An exploration of prepositional phrase-frames in Chinese EFL learners' theses and dissertations of applied linguistics. *Journal of English for Academic Purposes*, 70, 101400.
<https://doi.org/10.1016/j.jeap.2024.101400>