CHAPTER 3
RESEARCH METHOD

The present chapter describes the methodology to analyze the focus of the study about lexical bundles. The descriptions represent in some sub-sections. They are; research design, data source, instrument, data collection, and data analysis.

3.1 Research Design

The present study contains quantitative analyses and qualitative analyses. First, a computer software or AntConc automatic analysis for the lexical bundles facilitated the quantitative part of this study by providing a frequency list. The study of the existence of lexical bundles in the corpus are considered to be corpus-driven analysis. Moreover, the results after running AntConc based on the settings also provided the information required for the qualitative interpretation for the description of structural and functional types of lexical bundles identified.

In general, this research orientates exploratory approaches. The exploration begins in the frequent lexical bundles. The data processed in an automatic software system. The tool read all specifics four-gram involved in the corpus. After the results out, it goes to structure analysis of the bundles. As the final research approach, the functions of the lexical bundles are added.

3.2 Data Source

Twelve preferred Indonesian journals are promoted to be a main data of this study. Data regarding the research article as analysis material for the lexical bundles are gathered from that selected Indonesian Journal publication. Actually, there are masses of Indonesian journal on English language studies instead of twelve chosen journal. But the journal's accreditation of SINTA specific rank very potential to change. So, every journal is checked individually in the SINTA database about when the journal is included as one member of specific SINTA rank.

In addition, the journals are not chosen by chance. There are some criteria to be a factor of selecting the right journal. Decision was made on the basis of Nwogu's journal selection criteria. Those criteria are; representativeness indicates

that the journals are carefully and systematically selected to ensure they represent the language of the members of the targeted discourse community. Reputation is the second criterion which refers to the state of being held in high esteem and honor, which members of readership hold for the publication of a particular peer-reviewed journal. The final criterion is the selected journals could be accessed online via electronic databases or libraries by all researchers.[50]

From above information about the detailed categories in choosing the journal is defined in the present study. Regarding representativeness, from the selected journal are clearly represent the language of the members of the targeted discourse community. The journals were considered as established journals in the studies of English language. As for reputation, the journal is reported as nationally accredited in indexing database SINTA. Moreover, researchers can browse the journals easily through available website.

The editor policy mentions a specific purpose and scope to related journal. The chosen journal with specific descriptions is presented below;

a. **TEFLIN** dedicated to enhancing and disseminating scholarly work in the field of English Language Teaching (ELT). The journal publishes both conceptual and research-based articles within the fields of teaching English as a second or foreign language, English language teaching and learning, and English language teachers' training and education.

b. **IJAL** aims to promote a principled approach to research on language and language-related concerns by encouraging enquiry into relationship between theoretical and practical studies. The journal welcomes contributions in such areas of current analysis in; first, second, and foreign language teaching and learning, language in education, language planning, language testing, curriculum design and development, multilingualism and multilingual education, discourse analysis, translation, clinical linguistics, literature and teaching, and forensic linguistics.

c. **EduLite**: Journal of English Education, Literature and Culture focuses on publishing manuscripts within the following areas; teaching English as a first, second or foreign language, English language teaching and

---

[50] Nwogu. "The medical research paper: Structure and functions." *English for Specific Purposes*, 16(2), 1997

learning, English language teachers' training and education, postcolonial literature, modern literature, popular teen literature, literature and teaching, modern culture, postmoderism, trans-national culture, popular culture, folk culture, and culture and language

d. **English Review** include the following topic areas; English Language Pedagogy, English language Literacy, English for Specific Purposes, ELT Materials Development and Evaluation, English Language Testing and Assessment, Teaching English to Young Learners, English Literature, Language Policy and Planning, Second Language Acquisition

e. **Indonesian EFL Journal** welcomes contributions in such areas of current analysis as First and Second Language Teaching and Learning, Language in Education, Language Planning, Language Testing, Curriculum Design and Development, Multilingualism and Multilingual Education, Discourse Analysis, Systemic and Functional Grammar, Translation, Prose Analysis, Drama Analysis, Poetry Analysis.

f. **Vision**: Journal for Language and Foreign Language Learning focuses on first, second and foreign language learning. The scope of the journal covers first, second and foreign language acquisition, translation and language learning, language learning and identity, gender and equality in language learning, language skills, Islamic religious studies and language learning, linguistics and language learning, literature and language learning, and language teaching and learning. All disciplinary perspectives are welcomed, from linguistics and psychology to anthropology, cognitive science, neuroscience, sociology, and religious studies.

g. **JEES** cover the topics area of English Educational Research, English Linguistic, English Literature, Innovative Teaching and Learning Method, Curriculum and Material Development, English for Children, Translation, English Language Assessment, and English for Specific Purpose.

h. **Pedagogy: Journal of English Language Teaching** has the perspectives of English language teachings. This journal aims at presenting and discussing some outstanding issues dealing with English Language Teaching (ELT). This journal encompasses original research articles, and short communications, but not

limited to; Semantics & Pragmatics, Psycholinguistics, Sociolinguistics, Discourse Analysis, Translation Studies, English Language Teaching (ELT), English as Foreign Language (EFL), English for Specific Purpose (ESP)

i. **CELTIC** welcomes articles in the form of original research articles or reviewing research article on Culture, English Language Teaching, Literature and Linguistics.

j. **JALL** (Journal of Applied Linguistics and Literacy) presents research-based articles in the study of applied linguistics and literacy. The journal welcomes contributions in such areas of current analysis in; Language Pedagogy, Language Planning and Policy, Language Teacher Education, Language Testing and Assessment, Second Language Acquisition, Classroom Discourse Analysis, Pragmatics in Language Education, and Literacy in Language and Culture Education.

k. **ELT Forum** aim to promote a principled approach to research on language and language-related concerns by encouraging inquiries into relationship between theoretical and practical studies. The journal welcomes contributions in such areas of current analysis in the context of English language teaching encompassing; first, second, and foreign language teaching and learning, language in education, language planning, language testing, curriculum design and development, multilingualism and multilingual education, discourse analysis, translation, clinical linguistics, literature and teaching, and forensic linguistics.

l. **ELT in Focus** is a freely accessible, full text, peer reviewed journal allowing for the dissemination of ELT in these main areas; (1) Methodology in ELT, (2) Technology Enhanced Language Learning (TELL), (3) Language Policy and Curriculum, (4) Teacher Professional Development, (4) Literature in ELT, (5) Language Assessment, (6) Language Material Design and Evaluation, (7) Psychological and Sociological Influences on English language learning (8) Global/World Englishes and, (9) Other related disciplines or areas of research.

3.3 Instrument

3.3.1 Corpus

In the study of lexical bundles, the researcher must take corpus to fulfil this study. The aim of corpus is to obtain empirical data and to analyze the

collected data. The research article corpus is composed of collection of research article from twelve Indonesian journals in the present study. The research article is free to download from the official websites of target journals. The selected journals are all under English language studies and have different topic coverage. From the entire journals, the topic of English language teaching and learning can be found. The rest topics of interest of each target journal have multiple types.

The publication years of journal articles were limited in the same period between the years 2018-2020. The selected articles should be published in the last three years because to avoid possible new developments and changes in the structure of research articles over a lengthy time period. Moreover, the study carried out in that specific period indicate the current issue. As Ebrahimi and Heng stated that the chosen recent year of publication of research article is important to be respected in order to maintain currency of publication.[51]

Journal used in the corpus with the number of types word is shown in table below;

Table 3.1: The Total of Word Types from each Journal

No.	Name of journal	Types
1.	IJAL	152,319
2.	TEFLIN	144,744
3.	EduLite Journal	119,506
4.	English Review	120, 680
5.	Indonesian EFL Journal	127, 140
6.	VISION Journal	123, 471
7.	JEES	78, 344
8.	Pedagogy	112, 531
9.	CELTIC	99, 684
10.	JALL	90, 033
11.	ELT Forum	116, 164
12.	ELT in Focus	92, 231
	Total	**1,376,847**

The total number of 360 research articles were randomly chosen, and each of them represents different domains of research from all twelve target journals.

---

[51] Seyed Foad Ebrahimi and Chan Swee Heng. Grammatical subject in results and discussion section of research articles: Disciplinary variations. *Journal of Teaching Language Skills*, 37(1), 2018.

Above informational table are gained from different volume and number and were calculated from the last three years. Mostly all journals are published its biannual research article except IJAL journal that published three times in a year. There was also considerable specificity in the amount of research article gathered from the twelve journals. 10 texts are taken from each year edition of the journals. The selection of the journals was developed from different issues and volume. In that way, it can balance the number of 360 research articles texts. Finally, this strategy providing the basis for accurate principled categories in which the minimum number of corpus research has to achieved 1,000,000 words. Based on this, the involvement of research article structure consisting; introduction, methodology, results, finding, discussion, conclusion, and abstract section have been in the focus of research.

In addition, the criteria word counts of each journal are varied from one another. In TEFLIN Journal and IJAL Journal as SINTA 1 the word counts is equally start from 5000 with 7000 maximum word count in IJAL Journal and 8000-word count for TEFLIN Journal. Indonesian EFL Journal take a Minimum length text for 6000 words. Whereas, in English Review journal 6000 words is the limitation of words. The rest journal required the author to write the article more or less 5,000 words for the minimum. For the maximum length of text, the average words required in other existed journal in this study is not more than 7000 words. According to Biber "a corpus must be large enough to adequately represent the occurrence of the features being studied".[52]

In addition, the articles were chosen based on certain criteria. First, they were empirical or experimental studies. Therefore, other document types that usually can be found in many journals such as review articles and book reviews were not included. Second, the research articles must follow the structure of IMRD/C (Introduction, Methods, Results, Discussion, Conclusion) format with abstracts at the beginning of the article. Third, for the linguistic analysis, irrelevant information other than the text including names of authors, headings, tables, graphs, references, page numbers, formulas, titles, figures, appendices, biographical information, and acknowledgment was eliminated from the corpus.

---

[52] Doughlas Biber. *University language: A corpus-based study of spoken and written registers.* Amsterdam: Benjamin. 2006

There are some problems in the second criteria about following the conventional structure of research article. In the author guideline, each selected journal differs in giving the rules of naming the label of heading as in the conventional structure. The researcher met that kind of variation in the research article. First, there is no label of introduction in the article but the introduction section exists in the article. The journal which requires the structure of the article must include introduction without heading are TEFLIN Journal. Even the point of literature review can be explained in the introduction section, but some journal separates this label and become a main heading. Second, there is results section but was labeled differently. For example, in IJAL Journal and TEFLIN the label of results changes with the label finding. A different label of the heading discussion also found in the research article. Some journal has guided the writer of the manuscript to combine finding and discussion into one label heading. The journals which follow those criteria are; TEFLIN and IJAL, EduLite and Vision journal. For the last section of research article or conclusion section, some SINTA indexed Journal enquired the author to add suggestion for the label heading. For the abstract section, some journals asked the writers to provide both English and Indonesian languages.

The different label heading does not influence seriously in the process of analyzing the lexical bundles in each structure of research article. New version of heading in the structure of research article is still represent the basic and intended point from the conventional structure. Thus, the researcher has to deal with the variations of headings sections in research articles.

3.3.2 Computer programs

1) AntConc. 3.4.4 programmed

In order to find the most frequent lexical bundles in the research article. The analysis will be processed by a computer software. Since four-word lexical bundles become the main concern of identification. The researcher decides to use a fast, simple and more importantly gaining valid result with Ant Conc. 3.4.4w (Windows) 2014. This computer software is developed by Laurence Anthony. The advantages of using AntConc software than other is a useful tool preference. Concordance tool, for example, this tool shows search results in a 'KWIC'

(keywords in context) format. The specific function of 'KWIC' is the researcher can see the target analysis in the form of clause or phrases of texts. It also has word and keyword frequency generators. Particularly, 'clusters/ N-grams' menu is appropriate to analyze the lexical bundles.

2) AntFile Converter 1.2.0

AntFile converter is a freeware tool to convert PDF and Word (DOCX) files into plain text for use in corpus tools like AntConc.

3.3.3 Biber et al. (1999) Structural Taxonomy

As a broadening section after having a list of four-word lexical bundles in the text, the structural patterns are necessary to investigate. This is a possible way to study grammatical patterns because Biber and Conrad state that only 15% of the lexical bundles in conversation are complete clauses or phrases, whereas less than 5% for academic prose.[53]

Next, the researcher follows the 12 structural classifications based on Biber et al. (1999) to the lexical bundle. The following table introduced structural types of lexical bundles (Biber et al., 1999).

Table 3.2: Structural Types of Lexical Bundles (Biber et al., 1999).

Structural Category	Example
The component of Noun Phrase	
Noun phrase with *of*-phrase fragment	*a combination of*
Noun phrase with other post-modifier fragments	*an important role in*
The component of Prepositional phrase	
Prepositional phrase with embedded of-phrase fragment	*as a result of*
Other prepositional phrase fragment	*at the same time, on the other hand*
The component of Verb phrase	
Be + noun/adjective phrase	*is one of the*
Passive verb+ prepositional phrase fragment	*is based on the*
Anticipatory it + Verb / adjective phrase	*it is important to*
Pronoun/noun phrase+ be (+…)	*there are a number of*
Clausal	
Verb / adjective+ to-clause fragment	*to be able to*
Verb phrase+ that- clause fragment	*should be noted that*
Adverbial clause fragment	*as shown in figure*
Other expressions	*as well as the*

---

[53] Susan M Conrad and Douglas Biber. "The Frequency and Use of Lexical Bundles in Conversation and Academic Prose." *Applied Linguistics Faculty Publications and Presentations* (2005)

3.3.4 Hyland's (2008) Functional Framework

One manual classification process of Hyland (2008) has an additional advantage to show the functional characteristic of the bundles. As mentioned earlier, the identified lexical bundles were analyzed structurally according to Biber. For analyzing functional feature of lexical bundles, Biber also proposed a functional framework that shaped the types of the lexical bundles. Biber et al. classify the function of lexical bundles into three major functions of stance expressions, discourse organizers, and referential expressions. However, for the functional analysis of the Corpus, Hyland's (2008) framework is preferred. Hyland states that their framework is more appropriate for huge spoken and written corpus rather than small genre-based analysis of written texts.[54] The general categories of Hyland's (2008) functional framework is specifically designed to reflect the concerns of research writing.

There are three broad types of function: Research-oriented, Text-oriented, Participant-oriented. each type entails several sub-types. The first functional scheme is Research-oriented. This function help writers to structure their activities and experiences of the real world. There are five sub-categories included in this group. Location is a sub category to indicate time and place, for example: in the present study. Procedure is indicating methodology or purpose of research, e.g., the purpose of this. Quantification describing the amount or number, e.g., is one of the. Description detailing qualities or properties of material, e.g., in the control group. The last sub category is Topic related to the field of research, e.g., in the United States.

The second broad types are text-oriented bundles concerned with the organization of the text and the meaning of its elements as a message or argument includes. Type two has four subtypes; Transition signals establishing additive or contrastive links between elements, e.g., on the other hand, as well as the. Resultative signals mark inferential or causative relations between elements, e.g., the results of the. Structuring signals text-reflexive markers which organize stretches of discourse or direct readers elsewhere in the text, e.g., as shown in fig.

---

[54] Hyland, "Academic Clusters: Text Patterning in Published and Postgraduate Writing".

Framing signals situate arguments by specifying limiting conditions, e.g., in the presence of.

Participant-oriented focuses on the writer or reader of the text. Sub-categories of participant-oriented bundles are; Stance features convey the writers' attitudes and evaluations.[55] According to Cortes, this category includes attitude markers, epistemic-certain, epistemic-uncertain and intention bundles, e.g., were more likely to.[56] Engagement features- address readers directly, e.g., it should be noted.

3.4 Data Collection

This study is based on corpus-driven research. Therefore, an analyst needs to include a particular targeted size of corpus to determines the number of bundles. In entering items of finishing corpus, the researcher does some ways; 1. Only research articles that published in SINTA 1-6 indexed journal were included in the corpus. Loi identified three main variables to control for when building corpus: genre, authorship, and journals under investigation.[57] Based on such guidelines, the researcher finally decides on twelve Indonesian journals to be investigated. 2.the research articles from twelve target journals were downloaded freely.

The original text is in the form of pdf. A change needed to prepare with the file original version before being uploaded to AntConc. The researcher then modified the pdf version into a Word document. The next step included conversion of the Word documents into a plain text format. The step is acted with the aim of making the file detectable and readable in the computer software. Before the file is in a plain text version, Swales suggests omitting all reference lists, tables, appendices, footnotes, and acknowledgments; and also, titles, authors' names, headers, footers, and mathematical equations.

The counted for the category of a frequency 40 times per a million words was based on the cleaned articles. After preparation is complete, the investigation

---

[55] Ken Hyland, "Academic Clusters: Text Patterning in Published and Postgraduate Writing. *International Journal of Applied Linguistics* 18(1), 2008

[56] Douglas Biber, Susan Conrad, and Viviana Cortes "If You Look at … Lexical Bundles in University Lectures and Textbooks," *Applied Linguistics* 25, 2004.

[57] Chek Loi, "Research article introductions in Chinese and English: A comparative genre-based study." *Journal of English for Academic Purposes*, 9(4), 2010

of the lexical bundle can be started by first examining the frequency of lexical bundle, structures, and then functions of lexical bundles in all parts of the research article.

The interpretation of manual classification might be biased, therefore for this reason to verifying the validity of the data, the researcher applies methods of triangulation. From the five types of triangulations; data triangulation, investigator triangulation, theory triangulation, methodological triangulation, and environmental triangulation. Investigator triangulation is implemented in this study.

Investigator triangulation involves using several different investigators or evaluators in an evaluation project. Typically, this would manifest as an evaluation team. In order to triangulate, each different evaluator would study the program using the same qualitative method (interview, observation, case study, or focus groups). The findings from each evaluator would be compared. If the findings from the different evaluators arrive at the same conclusion, then validity has been established. If the conclusions differ substantially, then further study is warranted to uncover the "true" and "certain" finding.[58]

3.4 Data Analysis

In terms of analysis, there are three series of work that were considered to be a priority in this study. Investigating the frequency of lexical bundle is the first thing to do in the present study. To manage that, the researcher must equip the corpus to explore the variations of lexical bundles. The corpus used for the analysis was 360 published research articles from twelve Indonesian journal.

The use of AntConc. 3.4.4w (Windows) 2014 computer programmed is required to identify four-word bundles finding in the research article. The frequency cutoff of 40 times per million words corpus with a requirement that this rate of occurrence is realized in at least 36 different texts or 10 % of the total published research articles was considered as criteria. Expressions that appeared in less than 36 texts are not considered to be lexical bundles.

---

[58] Lisa Guion, Triangulation: Establishing the Validity of Qualitative Studies*, University of Florida*, 2002

As lexical bundles can be briefly described as expressions of two, three, four, five, and even more six words. The four-word bundle is significantly studied. This is because practically four-word bundles occur more than others.

The next identification process is structural categorization. For the classification, Biber et al.'s taxonomy (1999) was used. the structural category consists of three main categories. First category is Phrasal bundles. Phrasal involve some sub-categories; Noun-phrase bundles, Verb-phrase bundles, and Prepositional-phrase bundles. Second category is clausal bundles and other expression of lexical bundles as the last structural classification.

The last series of work is the functional categorization of the lexical bundles identified in the corpus. Hyland's (2008) functional taxonomy was used. Hyland's taxonomy is based on academic registers. So, that typical framework is sufficient to use in distinguishing the functional category of the lexical bundle in the journal research article.

There are three primary categories; 1. Research-oriented, 2. text-oriented, 3. Participant-oriented. Those three categories have each specific function to structure writers' experiences or activities in the research-oriented, organize text and its meaning in the function of text-oriented, and participant-oriented to focus on writer or reader. And the three main general functional categories break down into different subcategories.