ANALYSIS OF LEXICAL COHESION

Khairur Rijal, Sudarsono, Luwandi Suhartono
English Education Study Program, and Languages and Arts Education Department, Teacher Training and
Education Faculty of Tanjungpura University in Pontianak
Email: khairurrijalkr@gmail.com

Abstract
This research aims to find out types of Lexical Cohesion Devices (LCD encountered in research background of students' theses and how often these LCD were used in theses submitted to English Language Education Study Program, Languages and Arts Education Department, Teacher Training and Education Faculty, Tanjungpura University. This research was a descriptive study. The corpus of this research consisted of ten research backgrounds of undergraduate thesis submitted to English Language Education Study Program. This research focused on reiteration. The data were sorted from the corpus by taking the vignettes that contained LCD. Of 7410 words corpus, there were 101 LCD encountered. They are 69 repetitions, 25 synonyms, and 7 hyponyms. The most frequently used LCD was repetition while the least frequently used was hyponymy. On the other hand, antonym and metonymy were not found. Besides, the research also shows that some students applied various LCD in their thesis while some others failed to use them.
Keywords: Thesis Writing, Lexical Cohesion, Lexical Cohesion Devices

INTRODUCTION

A thesis is a piece of writing which is categorized as an academic form. Shannon (2011, p. 2) refers thesis to an academic field which focuses on a particular study. In English Language Education Study Program, Languages and Arts Education Department, Teacher Training and Education Faculty, Tanjungpura University, thesis writing is one of the important requirements for students to complete their study (Buku Pedoman Akademik, 2011, pp. 21-26). In order to be accepted as an academic writing, there are several important elements that students have to take into account. One of them is producing cohesive texts.

A text is an extended structure of words, phrases, clauses that has a unified meaning built up from a topic and purpose. A text can be spoken or written. A thesis is classified as a written text. Renkema (2004, p. 35) classifies the criteria to understanda text into six elements. One of them gives significant influence on the text quality in terms of its interpretation, which is cohesion. A good text

is cohesive. Cohesion does refer to the connection of one sentence to another in a text. Halliday & Hassan (1976, p. 8) refer cohesion to "a semantics relation between an element in the text and some other elements that are crucial to the interpretation of it". Cohesion can be expressed with lexical cohesion devices which depict a variety of vocabulary used in it; using synonyms to avoid monotonous sentences, producing academic range vocabulary and accurate diction based on a field of research. Halliday & Hassan (1976, p. 274) state that lexical cohesion is the cohesive effect achieved by selection of vocabulary. Therefore, the selection of vocabulary builds the connection within the text.

In reference to the pre-observation, students in English Language Education Study Program, Languages and Arts Education Department, Teacher Training and Education Faculty, Tanjungpura University stated that one of the obstacles that they encountered in writing a thesis was producing a minimum number of lexical cohesion devices. As an

example, students use 'find' along the sentence as opposed to using 'encounter' as a synonym of the word. Hence, it leads to monotonous sentence production in their thesis writing. Despite its less significant role in conducting a research, the variety of vocabulary affects the production of words used in aresearch. Production oflexical cohesion devices in students' writing can affect the quality of their research in terms of its words variety. A number of lexical cohesion devices that students produce indicate their proficiency in connecting one sentence with another one by using repetition, synonym, or antonym which constitutes cohesive writing that readers can interpret well. Therefore, it is interesting to analyze lexical cohesion to see the quality of vocabulary that students produce in their research background writing. Research about lexical cohesion has been conducted by some researchers. Frankhouser (2005) presented a system for linguistic exploration and analysis of lexical cohesion in English text. Sari (2012) conducted a research entitled Analysis of Lexical Cohesion in Applied Linguistics Journals in Andalas University, Padang. The aforementioned research focused on the journals and texts produced by English native writers. Therefore, this research focuses on analyzing the language produced by English students in English Language Education Study Program, Languages and Arts Education Department, Teacher Training and Education Faculty, Tanjungpura University, as non-native English speaker.

In this research, the writer investigates two main aspects. They are types of lexical cohesion which are used in thesis background and how often the lexical cohesion devices are used in the thesis. The corpus consisted of ten theses. Then, the data are sorted out from the corpora. Finally, the data are classified into the Lexical Cohesion Devices (LCD). This research is expected to be a beneficial source of information about lexical cohesion device use.

Principally, thesis is a piece of academic writing. Hogue, Hoshima, & Ann (2007, p. 3) refer academic writing to the kind of writing used in college classes. Academic writing is different from other types of writing, such as creative and personal writing whose purposes are to write stories, letters, and e-mails to particular people. Besides, Shannon (2011, p. 2) adds that academic writing is found in essays, course papers, research papers, book reports, translations, theses, dissertations, books, articles, and abstracts.

Wiyanto (2014, p.8) states that thesis is an academic writing to obtain a master degree in Indonesia. The undergraduate thesis is called *skripsi* while the doctoral dissertation is called *disertasi*. In general, those three terms are usually called as *tugas akhir* (final assignment) which is mandatory for the completion of a degree. These three forms of final assignment are aimed to require students to be able to present the result of their research in a structured and standardized form based on the requirements of college or university. Shannon (2011, p. 2) states the purpose of academic writing is to present information about a specific subject precisely and objectively. In addition, it is usually geared towards students demonstrating their mastery of content and the analytical and writing processes requisite for entry into their professions. Generally, academic writing applies either expository or argumentative prose to complete the requirements. Undergraduate students usually begin to write their final assignment in their third and fourth enrollment years, depending on the requirements of their respective disciplines and universities. In some universities, students are required to write a *skripsi proposal* and *thesis proposal* before they are permitted to write their final assignment (Wiyanto, p.9, 2014). The students are permitted to proceed to write their final assignment if the thesis proposal is considered to have fulfilled the qualification by the academic examiners.

In writing a thesis, background of research becomes an important element that students have to take into account. Background is a broad overview of students' research. Phillips & Pugh (2005, p. 15) state that background of a study is the state of art review of the field of study which includes current developments,

controversies and breakthroughs, previous research and relevant background theory.

In order to be accepted as an academic writing, there are some elements that have to be fulfilled in writing thesis backgrounds. They are a composition of idea and academic range of vocabulary. Wortham (2008, pp. 14-15) states that writers have to follow the elements until the writing achieves certain level of acceptance and authority. In addition, Bailey (2006, pp. 8-9) states there are some elements that need to be taken into account in academic writing; the paragraph construction, the use of compound and complex sentence, and the reference. Therefore, academic writing has its characteristics in terms of writing structure, form of language, as well as the sentences. In addition to the aformentioned elements, in order to be accepted as an academic writing, there are also several expectations that need to be fulfilled.

A text is an extended version of unit: words, sentences, clauses which has meaning and rules. Cornbleet & Carter (2001, p. 3) refer a text to a stretch, an extract or complete piece of writing or speech. A text generally adheres to broad conventions and rules which determine the language and structure used in particular text types. Halliday & Hassan (1976, p. 1) claim that text which is used in linguistics refers to any passage, spoken or written of certain length, that does form a unified whole. A text is a unit of language in use. It is not grammatical unit, like a clause or sentence and it is not defined by its size. Verdonk (2002, p. 2) adds that one recognizes a piece of language as a text, not because of its length, but because of its location in a particular context. If one is familiar with the text in particular context, one knows what the message is intended to be. Furthermore, Halliday & Hassan (1976, p. 17) state that a text always exists in the context.

Texts and contexts are inseparable. In addition, context of situation is feature which is relevant to speech that is taking place. A text can be divided into two types: spoken and written (Halliday & Hassan, 1976,

p. 3) classifies text into its general types which are spoken and written.

A text can be divided into two types: spoken and written (Halliday & Hassan, 1976, p. 3) classifies text into its general types which are spoken and written. Spoken texts include oral stories, interviews, dialogues, monologues, phone conversations, discussions, role plays, or any other piece of spoken language. When people are speaking to each other, their interaction is made up of series of utterances, for example questions and replies, comments and suggestions, requests and responses. On the other hand, written texts include stories, comic strips, instructions, recipes, PowerPoint presentations, emails, text messages on mobile phones, newsletters, posters, scripts for plays and performances, factual texts and explanations, or any other piece of written language. When people are writing to/for each other, their interaction may be made up of series of sentences, paragraphs and connected ideas.

Principally, a text functions as a means of communication. Through a text, one is able to transfer messages to others. As a means of communication, texts play a very important role in getting the meaning across others (Renkema, 2004, p. 61). Having very different types which are literary/expressive, scientific/informative and so on, they can meet the expectation of different-purposed readers. In other words, no matter what type it may be, every text ought to address certain receivers who read it for a specific purpose. Therefore, text functions as a mean of information transmitter or conveying ideas.

A text can be classified based on its genre. Texts are classified into genres on the basis of the intention of the communicator. Macken (1990, p. 12) classifies genre into story genre and factual genre. Story genre consists of five genres. They are narrative, news story, exemplum, anecdote, and recount. Factual genre is divided into procedure, explanation, report, exposition, and discussion. Mark and Kathy (1997, p.3) categorize genre into two types, that is, literary genre and factual genre. Literary genre consists of narrative, poetry,

and drama. Factual genre has seven types of genre. They are recount, explanation, discussion, information report, exposition, procedure, and response.

Discourse Analysis is a term to refer to study of the ways in which language is used in texts. Stork (1976, p. 69) states that discourse analysis is the linguistics analysis of language stretches longer than a sentence that aims to find sequences of utterances with similar environments (equivalence classes) and to establish regularities in their distribution. Yule (1996, p. 139) claims,

"In the study of language, some of the most interesting questions arise in connection with the way language is 'used', rather than what its components are. We are asking how it is that language-users interpret what other language-users intend to convey. When we carry this investigation further and ask how it is that we, as language-users, make sense of what we read in texts, understand what speakers mean despite what they say, recognize connected as opposed to jumbled or incoherent discourse, and successfully take part in that complex activity called conversation, we are undertaking what is known as discourse analysis."

In short, discourse analysis concerns with the study of how a language is used in context, despite of whether the language is correct. Discourse analysis involves many things such as delivering style, appropriateness in delivering, cohesiveness and rhetorical force of the discourse, topic/subtopic structure, as well as grammatical properties. Besides, discourse analysis can be in a form of either spoken or written. Discourse analysis is the study of language in the context larger than sentences. It analyzes not only the relationship among paragraphs, sentences and phrases in term of the flow of ideas, but also the meaning of the discourse that is intended by the speaker or writer.

Coherence is a term of text linguistics. It is used to refer to relations (sentences or propositions) of a text. Coherence deals with the unity of ideas in a text. Besides, cohesion is the network of semantic relations which organizes and creates a text by establishing continuity of sense. Yule (1996, p. 140) points out that something which exists in people is like the key concept of the coherence, not something which exists in the language. People are the decision-makers of what is logically acceptable and what is not based on their understanding or interpretation. In addition, Yule (1996, p.141) also says that this ability to rationalize of what one reads or hears is probably only a small part of the general ability one has to rationalize of what one perceives or experiences in the world. Jane E et al (2007) state that coherence is achieved through some elements (as cited in Jayetta Slawson et al (2011, pp. 42-45). The elements are repetition, transitonal expression, pronoun, synonyms, and paralel structure to link ideas, sentences, and paragraphs. To achieve coherence, a text should contain the aformentioned elements.

The example is shown in (1).

(1) **The research** aims to provide general picture of education. Besides, **it** also provides guidelines to write essay, analyze articles, and conduct a research.

In (1), the sentences are considered coherent. The coherence is achieved because the first and second of occurrence are linked by elements which make the text meaningful and understandable. In second occurrence, the word **it** is a pronoun of the word **the research** in the previous occurrence.

Cohesion is relation that indicates connection exists among texts and its meaning. Halliday & Hassan (1976, p. 8) refer "cohesion" to "a semantics relation between an element in the text and some other elements that are crucial to the interpretation of it". Therefore, cohesion plays a vital role in discourse because it enables the readers and listeners comprehend the speaking or listening, get involved in the flow of ideas, and interpret the semantic or pragmatic meaning of the discourse. Hence, in order to produce understandably good discourse, both spoken and written discourse should be cohesive. Cohesion emphasizes on the semantic relation rather than the pragmatic relation. In order to be cohesive, a text must

be connected by devices. Bussman (1998, p. 199) states that cohesion is often manifested through the use of such "devices" as pronouns (e.g. I, he, she, it)

An example of a cohesive sentence is written bold in (2)

(2) I have **a very pretty house.** *It* is red-painted.

The phrase **'a very pretty house'** is a noun phrase. In the second sentence, this phrase is replaced by the pronoun**'it'** which functions to repeat the noun phrase '**a very pretty house**'. In the above example, the interpretation of **it** depends on another element in the text (i.e **a very pretty house**).

Lexical cohesion deals with the meaning in the text. To complete the picture of cohesive relations it is necessary to take into account also lexical cohesion, lexical cohesion is the cohesive effect achieved by the selection of vocabulary(Halliday & Hassan, 1976, p. 274). It means that writing composed by selection of vocabulary and lexical cohesion concerns with connection among words used. Halliday & Hassan (1976, pp. 276-278) divide the lexical cohesion into reiteration and collocation. In this research, the researcher only focuses on reiteration.

Reiteration is the "repetition of a lexical item, or the occurrence of a synonym of some kind, in the context of reference; that is where the two occurrences have the same referent" (Halliday & Hassan, 1976, p. 278). Reiteration can be repetition, synonymy, antonym, hyponymy, and metonymy.

Repetition refers back to preceding word. Repetition is a part of lexical cohesion that involves the repetitions of lexical items.

(3) **Science students** need to master the use of present continuous tense because *these students* are required to make sentences by using present continuous tense.

In (3), 'science students', the subject of the main clause of (3), is repeated in the sub clause with 'these students'. It is categorized as repetition because the meaning of 'these students' is still related with 'science students' in the main clause.

Synonymy is the experiential meaning of two lexical items which are identical. This is

not a form of overlapping meanings; but this expresses the similar meaning of different lexical items.

(4) Some students **found** difficulties in learning English. Besides, they also *encountered* difficulties in grasping the materials.

In (4), the words 'found' and 'encountered' are synonymous. The author used different words simply to make variation in the sentence.

Halliday & Hassan (1976, p. 312) define antonym as "the oppositeness of meaning; the members of our co-extensional tie". An example of antonym is shown in (5).

(5) The materials which are given today are so **difficult**. Students told the teacher that previous materials were *easy.*

In (5), the word 'difficult' and 'easy' are antonymous. This pair is lexical item that is denial of the one implies the denial of other.

Hyponymy is the relation that holds between a general case and its sub-classes (Halliday & Hassan, 1976, p. 80). An item that relates to general class is called subordinate and the item that relates to the sub-classes is called hyponymy.

(6) Of all the **students** in the school, the researcher took *science and language students* as the samples.

In the above sentence, it can be seen that students is a subordinate and its hyponymy are science and language students.

Metonymy refers to a part-whole relation or a concept used in the linguistic community which specifically deals with part and whole relation.

(7) This research will be conducted in one of the public **schools** in the city. This research cannot be separated from the existence of students*, school staffs, curriculum,* and *facilities.*

In the above sentence, the words 'students', 'school stuffs', 'curriculum', and 'facilities' are metonymy of school,

which students, school staffs, and facilities are lexical set of school.

Halliday & Hassan (1976, pp. 39-40) define collocation as "a word that is in some way associated with another word in the preceding text because it is a direct repetition of it, or is in some sense synonymous with it, or tends to occur in some lexical environment, coherence with that word and so contributes to the texture".

(8) 'school' and 'students'
(9) 'country' and 'president'
(10) 'college' and 'lecturers'

METHOD

This research is a descriptive study. Cohen, Manionand, & Morrison (2007, p. 205) deal descriptive research with "conditions or relationships that exist; practices that prevail; beliefs, points of views, or attitudes that are held; processes that are going on; effects that are being felt; or trends that are developing. At times, descriptive research is concerned with how *what is* or *what exists* is related to some preceding event that has influenced or affected a present condition or event". Descriptive research is designed to obtain information about the current status of phenomena. Based on the purposes and problems of this research, the linguistic analysis within the descriptive method is used. The linguistic analysis focuses on the linguistics use or the language structure. There was not any administration or control of a treatment as is found in an experimental research.

Corpus and data are inseparable elements in corpus linguistics. Kennedy (1998, p. 19) defines a corpus as a systematic collection of naturally occurring texts (of both written and spoken language). Corpus is a systematic collection. It means that the structure and contents of the corpus follow certain extra-linguistic principles. For example, a corpus is often restricted to certain text types, to one or several varieties of English, and to a certain time span. If several sub-categories (e.g. several text types, varieties etc.) are represented in a corpus, these are often represented by the same amount of text.

Corpus is available to the researcher (including the number of words in each category).

In this research, the corpus consists of ten research backgrounds of theses submitted to English Language Education Study Program, Languages and Arts Education Department, Teacher Training and Education Faculty, Tanjungpura University. These theses consist of 7410 words. The corpora are selected purposively. The theses were written by students whose GPA is equivalent to and more than (≥) 3.51. They were selected from different years of graduation. The research does not focus solely on certain types of research. The researched

After the corpus is selected, the words of corpus are counted. They consist of 7410 words corpus. Afterwards, the data are sorted out. To sort out the data, vignettes of the words corpus are collected. Vignettes are short depictions of typical scenarios intended to elicit responses that will reveal values, perceptions, impressions, and accepted social norms. In this research, vignettes are quoted form of sentence in undergraduate thesis. In collecting the vignettes, the table of specification is used to list all vignettes that contain lexical cohesion devices.

The table of specification consists of number of vignettes that are obtained through theses and Lexical Cohesion Devices (LCD) that are encountered the obtained vignettes. In addition, the table of specification also consists of the frequency of the encountered Lexical Cohesion Devices.

Procedure of Data Analysis

There are four steps of data analysis. They are collecting the backgrounds of the selected theses, identifying the lexical cohesion, and classifying the lexical cohesion.

Collecting the Corpus

After the researcher determines the selected theses, the backgrounds are collected. In collecting the backgrounds, the researcher conducts two steps which are explained in (1) and (2).

(1) **Requesting Permission**

In this step, the researcher asks permission of library staffs of Tanjungpura University. In order to be permitted, the researcher submits the letter of conducting the research.

(2) **Collecting the Backgrounds**

After the researcher is permitted, the researcher collects the backgrounds by photocopying them.

Sorting out the data (LCD) out of the Corpora

After collecting the backgrounds, the data are sorted out from the corpus. Then, the data are identified. In identifying the data, the researcher lists the Lexical Cohesion Devices which are encountered in vignettes.

Classifying the Lexical Cohesion Devices (LCD)

Lastly, after the data are identified, the data are classified. The data are classified into types of LCD used in undergraduate theses, the frequency of each LCD their correct and incorrect use.

Computing the Frequency of Each Type

After the data are classified, the data then are computed in a form of table to count the number of encountered LCD.

RESEARCH FINDINGS AND DISCUSSION
Research Finding

Referring to the research questions which are: 1). What types of the lexical cohesion are used in research background of students' undergraduate thesis?, and 2). How often are lexical cohesion devices used in thesis, The findings of this research show that 101 LCDs were used of 7410 words corpus. The types of lexical cohesion devices used in undergraduate thesis consisted of 69 repetitions, 25 synonyms, and 7 hyponyms. On the other hand, antonym and metonymy were not found. In addition, repetition is the most frequently-used type of Lexical Cohesion Devices while hyponymy is the least frequently-used device. Furthermore,

some devices of lexical cohesion were still incorrectly used in backgrounds of undergraduate theses.

Of all the LCDs, repetition was the most-frequently used device in backgrounds of undergraduate theses. Of 7410 words corpus 68 devices were used. These 68 devices were used to cohere sentences.

Synonymy is the second most frequently-used Lexical Cohesion Devices. Out of 7410 words corpus 25 synonyms were used.

Table 2. Lexical Cohesion Devices (LCD) found in 7410 words Corpus

LCD	Frequency
Repetition	69 devices
Synonymy	25 devices
Hyponymy	7 devices

Discussion

The research findings show that there are 3 types Lexical Cohesion Devices that were used in theses. They are repetition, synonym, and hyponymy. Of 7410 words corpus, 100 devices were found. These devices consist of 68 repetitions, 25 synonyms, and 7 hyponyms.

Out of 7410 words corpus, the most frequently-used LCDs are repetition. Of these 69 devices, 3 of them were used incorrectly. It is due to the incorrect lexical items between first and second occurrence. Repetition is used to refer back to the preceding word. Repetition is used by altering the previous lexical items with pronoun.

In relevance with Halliday & Hassan (1976, p. 278), interpretation of element in a text depends on another element in the text. In repeating the element in the first occurrence in the text, a correct pronoun must be used. Yet, 3 occurrences were not cohered by the correct pronoun.

Furthermore, 25 synonyms were used out of 7041 words corpus. Synynomys were used correctly. The students were able to give variation upon their theses by providing another diction. The researcher identified this

by checking the availabilty of the dictions on Oxford dictionary. Students varied their diction through two ways: writing another diction of a particular word or providing another choice of diction from first occurrence.

Lastly, 7 hyponyms were used out of 7041 words corpus. Hyponymy was produced by giving general class of words in the first occurrence and its sub-classes in second occurrence. Hyponymy is the least frequently-used Lexical Cohesion Device (LCD). Out of 7410 words corpus 7 Hyponyms were encountered.

CONCLUSION AND SUGGESTION
Conclusion

Referring to the discussion of this research, the researcher would like to point out a few conclusions which are explained as follows:1) Lexical cohesion consists of 2 types. They are reiteration and collocation. The researcher focused on reiteration which consists of 5 Lexical Cohesion Devices; repetition, synonymy, antonym, hyponymy, and metonymy. Among these 5 Lexical Cohesion Devices, repetition, synonymy, and hyponymy were only LCDs that were found out in the research while antonym and metonymy were not found out.2) Repetition is the most frequently used Lexical Cohesion Devices while Hyponymy is the least frequently used. 3) Out of 7041 words corpus, 101 LCD were used. They are 69 repetition, 25 synonyms, and 7 hyponyms. 4 Hyponymy and Synonym are used properly by the students. The Repetition, however, is not necessarily used properly. Some students are still unable to cohere the first occurrence with the second occurrence by a correct pronoun.

Suggestion

Suggestions will be given to the improvement of upcoming research, students' capacity regarding Lexical Cohesion, and English teaching in general. The suggestions are explained as follows: 1) It is suggested that upcoming research of Lexical Cohesion or research pertaining to it will improve the quality of this research in terms of its analysis. The upcoming research can focus on the contributing factors of how all Lexical Cohesion Devices do not occur. 2) Students of English Language Education Study Program, Languages and Arts Education Department, Teacher Training and Education Faculty, Tanjungpura University are suggested to be able to expand their knowledge about Lexical Cohesion and how to use Lexical Cohesion Devices properly in undergraduate thesis. 3) Despite not being formally taught in classes, lecturers are also suggested to be able to give a lecture about Lexical Cohesion and how to write coherent sentence by using Lexical Cohesion Devices.

REFERENCES

Bailey, S. (2006). *Academic Writing: A Handbook for Intrernational Students.* London: Routledge.

Cohen., L., Manionand., L., & Morrison, K. (2007). *Research Methods in Education.* New York: Routledge.

Flowerded., John., & Mahlberg, M. (2009). *Lexical Cohesion & Corpus Linguistics.* London: Routledge.

Halliday, M., & Hassan, R. (1976). *Cohesion in English.* London: Longman.

Hogue, Hoshima, A., & Ann. (2007). *Writing Academic English.* New York: Pearson Longman.

Phillips, & Pugh. (2005). *Research Studies.* London: Longman.

Renkema. (2004). *Discourse Analysis.* London: Longman.

Shannon, P. (2011). *A Guide to Academic and Scholarly Writing.* Ovieda: Baldwin Book Publishing.

Stork, F. C. (1976). *Dictionary of Language and Linguistics.* London: Applied Science Publisher.

Thompson, G. (2009). *Introducing to Functional Grammar.* London: Routledge.

Troike, M. (2006). *Introducing Second Language Acquisition.* Cambridge: Cambridge University Press.