CHAPTER II

REVIEW OF RELATED LITERATURE

This chapter deals with the related literature about lexical, lexical density, lexical diversity, thesis conclusion, AntConc software, Type-Token Ratio (TTR), and some related previous studies.

A. The Nature of Lexical

According to Merriam-Webster Dictionary, the word lexical has meaning relating to words or the vocabulary of a language distinguished from its grammar and construction. The lexical concept is innate, and only complex lexical meanings might be learned. The point is that words are not explicitly learned by experience but lexical-semantic implicitly learned by experience. Lexical field theory had several forerunners in the 19th century and the first decades of the twentieth-century paper[15]. An analyzer of the lexical field is the first phase of the language compilation process. It deals with the processing of the input language. The lexical analyzer can be a convenient place to carry out some other chores like stripping out comments and white space between tokens and perhaps some macros and conditional compilation features. Discussion is also extended to the multi-core environment.

Language has traditionally been understood as a hierarchical system of phonology, morphology, syntax, etc. Lexical concepts may have a

---

[15] Bernd Kortmann. *Theories and Methods in Linguistics.* (New York: De Gruyter , 2013).

compositional structure, and meanings are composed of smaller parts across different sensory information and lexical domains. However, it should be emphasized that students do not make any claims about the cognitive mechanism underlying these compositions. The automatic and effortless lexical acquisition process shows that these processes are not conscious or intentional inferences.[16]

B. Lexical Density

Lexical density is a measure of the amount of content information in a text. The lexical density describes the proportion of lexical words to the total number of words in any given text in spoken or written.[17] A text with a lower density is easily to understand, while a higher density can be understood more difficult. In lexical density, the measurement of a text with a high proportion of content words contains is more information than a text with a high proportion of function words (prepositions, interjections, pronouns, conjunctions, and count words). If the text has more grammatical items than the lexical items, it is categorized as having lower lexical density. On the contrary, if the text has more lexical items than the grammatical items, the text is organized to the high lexical density. To find the lexical density is calculated by dividing the number of content words in a clause complex.[18]

Readers will receive information with a simple measure of how informative the text is by investigating lexical density. Lexical density has a more significant impact on the memory process. Texts with a higher density

---

[16] Dieter Hillert. "The Nature of Language: Lexical Concept," *Springer Science*, (2014), 89-97.
[17] Ure, *Lexical Density and Register*.
[18] Signes and Arroitia, *Analysing*.

are more difficult to be comprehended. On the contrary, the texts with low lexical density are easily to be comprehended even it contains high information. But, the highest level of lexical density not always give the readable text to readers. When readers cannot understand the content of a text, the readers might not get the point and information from there.

C. Lexical Diversity

Based on Signes and Arroitia, the lexical richness or also called lexical diversity of a text. [19] Lexical diversity measures how many different words are used in a text. On the other side, lexical density measures the proportion of lexical items (i.e., nouns, verbs, adjectives, and some adverbs) in the text.[20] For a text to be highly lexically diverse, the speaker or writer has to use many different words, with little repetition of the words already used. [21] Moreover, diversity has something to do with the range of vocabulary displayed and different word types. Counting other words is the most method of use in measuring diversity. When calculating a relative range is common, and in general, there have been two approaches to solving the research problem in language development research. The first is to standardize the size of the samples, and the second is to consider the ratio.[22]

---

[19]Carmen Gregori Signes, and Begoña Clavel Arroitia., "Analysing lexical density and lexical diversity in university", *Procedia - Social and Behavioral Sciences*, (2015), 546 – 556.

[20] Johansson, *Lexical Diversity And Lexical Density.*

[21] Ibid.

[22] Pilar Duran, David Malvern, Brian Richard, and Chipere Ngoni, "Developmental Trends in Lexical Diversity"*, Applied Linguistics*, (2004), 220-242.

Based on Gelderen, there are nouns, verbs, adjectives, and adverbs as lexical categories. Those categories carry meaning, and often words with a similar (synonym) or opposite meaning (antonym) can be found.[23]

1. Noun: a noun generally indicates a person, place, or thing. For instance, *chair, table*, and *book* are nouns since they refer to things.[24]

2. Verb: can be said as an event or activity where the only distinction, specific nouns such as *action* and *destruction* would be verbs since they imply action. Some of the morphological characteristics of verbs are that they can express tense, e.g., past tense ending in *-ed*, and the verb ends in *-s* when it has a third-person singular subject and is present tense; and that it may have an affix typical for verbs, namely *-ize*, e.g., in *modernized*.[25]

3. Adjective: adjective modifies a noun, the quality it describes will be one appropriate to a noun, e.g., nationality/ethnicity (*American, Dutch, Iranian*), size (*big, large, thin*), age (*young, old*), color (*red, yellow, blue*), material/personal description (*wooden, human*), or character trait (*happy, fortunate, pleasant*).[26]

4. Adverb: adverbs often modify actions and will then provide information typical of those, e.g., manner (*wisely, fast, quickly, slowly*), or duration (*frequently, often*), or speaker attitude (*fortunately, actually*), or place

---

[23]Elly van Gelderen. *An Introduction to the Grammar of English.* (Philadelphia: Arizona State University, 2010),12.
[24] Ibid,13.
[25] Ibid.
[26] Ibid, 14.

(*there, abroad*), or time (*then, now, yesterday*). Also, negatives such as *not* and *never* are adverbs in that they usually modify the verb.[27]

D. Thesis Conclusion

A group of paragraphs written about a single topic and a control main idea is called an essay.[28] An essay must include at least three paragraphs but commonly consist of five paragraphs for academic writing. Academic essay text written by the students aimed to know students' writing skills and quality. According to Whitaker, there are ten points of good academic essay characteristics, and there are clear purpose, audience engagement, clear point of view, single focus, logical organization, strong support, clear and complete explanations, effective use of research, correct citation style, writing style.[29]

There are three main points of the essay they are an introduction, main body, and conclusion.[30] The first paragraph is explaining the topic in general ideas. Here is an introduction paragraph that has the thesis statement as the main idea. The thesis statement is the sentence that tells the main idea of the whole essay given by the author to state the important opinion about the topic. Some paragraphs explain and support the thesis statement. These paragraphs are called the main body of an essay since they come between the introduction and conclusion. The paragraphs in the main body of an essay should always elaborate on the thesis statement. In each paragraph, there is only a thesis statement that discusses a topic. In the last paragraph, there is a

---

[27] Ibid.
[28] Zemach, and Rumisek. *Academic Writing*.
[29] Ibid.
[30] Whitaker, *Academic Writing Guide*.

conclusion. This paragraph summarizes or restates the thesis statement and the supporting ideas of the essay.

According to Oshima and Hogue, the conclusion is the final paragraph, it has three puposes. There are:[31]

1. It signals the end of the essay. Author should begin the conclusion with a transition signal (*in conclusion, in summary, in brief, in short, etc.)*

2. It reminds the reader of the main points from the author. It can do in one of two ways: summarize the subtopics or paraphrase the thesis.

3. It leaves the reader with with final thoughts on the topic. This is opportunity to convey a strong and effective message that will remembered by the reader.

The purpose conclusion part is to summarize the main points of the research. It is the last opportunity to bring together what the results has been found and make the opinion and understanding of the topic clear to the reader or even the examiner. For this reason, it is a crucial component of the research.

In writing conclusion section, Oshima and Hogue give some techniques that can used by the author to make the conclusion memorable:[32]

1. Make a prediction:

2. Suggest results or consequences

3. Suggest a solution, make a recommendation, or call for action.

4. Quote an authority on the topic.

---

[31] Alice Oshima and Ann Hogue, *Writing Academic English: Fourth Edition*, (New York: Pearson Education, 2006), p.72.
[32] Ibid.

Concluding paragraph can be arranged by summarizing the main poin of an essay or result of a research. It can be a restatement of the thesis statement in different word that consits of message from the outhor to the reader.

E. AntConc Software

The interaction between computer processing data and linguists interpreting them provides enormous benefits in language research. Using concordance software, the corpus or corpora offers many benefits to the world of language, both in teaching and research.[33] AntConc offers a word clusters/bundles tool for studying multiword units such as collocations, phrasal verbs, and idioms. Word clusters are found by using word clusters tool and arranged alphabetically or as frequency groups. The search terms can be substrings, words, phrases, or regular expressions.[34]

AntConc software was released in 2002 as a simple Key Word in Context (KWIC) concordancer program designed by a professor at Waseda University Japan named Anthony Laurence. The purpose of this software is to be used in the classroom as an analysis instrument, utilizing a powerful concordancer, word and keyword frequency generators, tools for cluster and lexical bundle analysis, and a word distribution plot.[35]

AntConc is a freeware based on multiplatform application and can be run either on Windows or Linux/Unix systems. Another unreserved

---

[33] Arum, and Winarti. "The Use of Antconc in Providing ", 106-112.

[34] Anthony Laurence,. "AntConc: Design and development of a freeware corpus analysis toolkit for the technical writing classroom." *International Professional Communication Conference Proceedings.* Piscataway: IEEE, (2005), 729-737.

[35] Anthony, "AntConc: Design and development of a freeware corpus analysis toolkit ", 729.

advantage of AntConc software is its distribution as a single executable file easily copied onto the user's computer. Also, AntConc does not require much memory to be saved and run, approximately 2 MB of disk space is needed to launch it. Multiple-level sorting, Unicode support, along an easy-to-use manual makes AntConc a unique and extremely applicable analytical tool.[36]

F. Type-Token Ratio (TTR)

Many lexical analyzer tools were developed in the past and are available, which is best suited for sequential processing. It is better to standardize the number of tokens compared since two methods can be used in language developmental research. There is a way of determining both the standard number of tokens used and the method of selecting them. The second method is the use ratio, the ratio between the number of types and the total number of the word (token).

According to Schiffrin, Tanen, and Hamilton, a text is any stretch of naturally occurring language in use, spoken or written, produced independently of the analyst for some real communicative purpose.[37] A corpus is an extensive collection of computer-readable texts of different text types, representing spoken and/or written usage. No corpus can be a fully representative sample of the whole language. However, such collections can be designed to represent significant dimensions of language variation, such as spoken and written, casual and formal, fiction and nonfiction, British and American, intended for different age groups, experts and laypersons, and so

---

[36] Rozanov, and Tsybulsky. "Linguistic Analysis of Science Teachers' ", 211-230.
[37] Deborah Schiffrin, Deborah Tannen, and Heidi E. Hamilton. *The Handbook of Discourse Analysis.* (Massachusetts: Blackwell Publishers, 2001).

on. Effective means at least millions, and possibly hundreds of millions, of running words (tokens). Here is the measurement of lexical density using Type-Token Ratio:

$$\text{Lexical Density} = \frac{the\ number\ of\ lexical\ item}{the\ total\ words} \times 100$$

As a rule, texts with a higher density are more difficult understood, and written texts have higher lexical density levels than spoken texts.

G. Previous Studies

Some previous studies analyze lexical diversity using AntConc software and using Type-Token Ratio (TTR) to measure the lexical density. As mention before, the previous researchers examined various topics, such as English textbooks, narrative texts, and news texts. Fadhillah in 2018, analyzed English textbooks focusing on lexical density analysis for analyzing 15 reading texts provided in *Pathway to English* Textbook for Grade X of Senior High School and analyzed based on lexical density. As a result, most of the reading passages in this textbook are at a high level. It shows that from 15 reading texts analyzed, 9 texts with high lexical densities (complicated texts) and medium lexical densities (not complicated/the difficulty is just in the middle position).[38]

Besides that, in 2019 Arum and Winarti also conduct research by the title "*The Use of Antconc in Providing Lexical and Syntactical Information of the Textbook of Radiographic Positioning and Related Anatomy: A Corpus*

---

[38] Fadhillah, *Analyzing Lexical Density of English Reading*.

Linguistic Study." This study aimed to determine how AntConc can provide linguistic features at the syntactic level concerning vocabularies (word classes) and grammar (syntactical) that appear and are used in the textbook. This research showed that AntConc has successfully generated a list of vocabularies sorted by their occurrences. In this case, AntConc was able to assist researcher in providing necessary lexical and syntactical information of the textbook. [39]

Rozanov and Tsybulsky conducted the other research using AntConc software in 2019 with the title *"Linguistic Analysis of Science Teachers' Narratives Using AntConc Software."* This research reveals that AntConc software is a valuable tool of linguistic analysis, which can contribute to understanding various issues in educational research. Linguistic analysis enhanced by computational technologies (e.g., the AntConc software) can process large text corpora, highlighting patterns of implicitly conveyed thoughts and feelings, thereby providing a means for in-depth observations, leading to new insights into various educational phenomena. [40]

While research that analyzes news text entitled *"Lexical Density Analysis and its Function in BBC News"* was conducted by Aulia in the same year, 2019. This study's data are the lexical density of the reading online newspapers found on application of *BBC News*. Furthermore, the results showed that the distribution of lexical density in reading text in *BBC News* newspapers differed. The text's lexical density level had different levels: the reading texts had five texts containing lexical items dominant and three non-

---

[39] Arum and Winarti, "The Use of Antconc in Providing" 106-112..
[40] Rozanov, and Tsybulsky. "Linguistic Analysis of Science Teachers' ", 211-230.

content carrying lexical items. So it readers quite difficult to understand the vocabulary.[41]

---

[41] Aulia,. *Lexical Density Analysis*.