

CHAPTER I

INTRODUCTION

In this chapter, the researcher presents background of the research, statement of the research problem, research objective, the significance of the study, scope and limitation of the research and definition of the key term.

A. Background of the Research

Nowadays corpus studies are successfully integrated into language research in practically every discipline. Many of studies and analysis have been conducted in corpus. Here are some previous studies related to corpus. Kathleen and Kenju Kitao (2014) conducted a study focused in analyzed the responses to apologize. The research used 320 apologize from corpus made up of dialogue from 72 episodes from US situations comedy *Modern Family*. The result of this study show that about half of the apologies studied, there was no respond. In half of the apologies, there was a respond. The majority of the responses-four out five-fell into one of three categories. Then corpus based study can be used to examine the language used in particular part of media mass publications. Several research has been conducted corpus study of media mass product. They might observe the language description and the grammatical features of media mass text. For the example, the study was conducted by Ling (2016). She conducted a corpus analysis in magazine horoscopes. She tried to investigate the use of modal auxiliary verb of the horoscopes. For this study, the corpus consisted

of horoscopes from 108 issue of 3 magazines namely CLEO, Female, and Malaysian Women's Weekly from the year 2011-2013. The result of this study, she was found that the most frequent word which appear in horoscopes is pronoun, then the most frequent modal auxiliary verb used in horoscopes was modal of probability "will" and followed by "may".

Corpus study also developed in observing the cooperative analysis of the news coverage between two media mass in different countries. One of the cooperative studies was conducted by Chen (2013). He tried to investigate a cooperative analysis of the news coverage of the 2011 Libyan civil war in two national media, those are China Daily and The New York Times. This article was published between March 1,2011 and October 23,2011. The data analysis was used to generate and compare the word frequencies and analyzed collocations of lexeme "Gadhafi/Qaddafi". Word "Qaddafi" was chosen of the principal role of Muammar Qaddafi during the Libyan civil war. From this study could be revealed the results that the two corpora had a lot of similarities in word in word frequency, but there found differences in several high ranking lemmas too. In addition, the result of the collocations analysis of the two corpora revealed that there was a contrastive images of word "Qaddafi". In china daily corpus, the collocations of "Qaddafi" includes many neutral and even positive description, in other hand, negative description and highlighted the pressure on Qaddafi were mainly appeared in The New York Times corpus.

For some studies, example Moinvaziri (2012) tried to investigate vocabulary used in general English textbook and used a computer program called Vocabprofile (VP). The reason for the researcher to select the textbook because it was the only textbook which is taught for General English course nationally which has been taught for more than ten years. The result shows that the vocabulary presented in the reading passages of the textbook is mostly for beginner and low intermediate levels but the level of textbook cannot be an efficient one for the university students. And for the frequency of word presented in the textbook does not seem to provide enough instances of new items and neither does it provide enough opportunity for practicing these items. The only vocabulary items which are acceptably rehearsed are the words of first 1000 frequent words which are not suitable and challenging for the level of the students. Trebits (2009) conducted a study to explore the use of phrasal verbs in English language documents of the European Union (EU). The study was part of a large-scale project examining the use of English in EU text from various aspects. The specific purpose of the research was to generate findings that may serve as a basis for the compilation of teaching materials and syllabi designed to develop the necessary language skills of those who work with English language EU documents.

The second previous study, Sun (2016) focused on the analysis of vocabulary coverage in children's story series "The Magic Three House". The study was conducted based on the problem in second language learning of

extensive reading. According to the researcher, extensive reading is still not commonly practiced in EFL setting in Taiwan. Teachers, students, and parents perceive extensive reading as impractical due to the lack of immediate measurable gain in test scores. So, to overcome the difficulties, reading a large quantity of target language materials can aid learners in acquiring new word, and incidentally strengthen previously learned ones. The researcher use a corpus analysis approach to analyze the vocabulary coverage within a single-author story series, its repetition of vocabulary, and the incidental and intentional vocabulary learning opportunities therein. Then, the result of word list from The Magic Three House compared with three words list published by The Ministry of Education in Taiwan. This three word list a guideline for book developers to follow. The findings indicate that extensive reading of The Magic Three House series can support word list learning due its wide coverage of vocabulary from the three list, especially the Middle School Basic 1,200 list. The 48 books of the series cover 91.4% of Basic 1,200 word, 36.4% of Advanced 800 word, and 36.4% of High School 4,700 words. In addition, the rich Magic tree House corpus provides ample repetition of vocabulary from all three words list. In particular, the findings indicate that the more basic the vocabulary list is, the higher the percentage of repetition. Therefore, extensive reading of only one story series we not only expose learners to large portion of word from the list, but many of them repeatedly.

The last study is from the two news report corpora was conducted by Moon(2016) from University of Washington. She studied the similarities and differences in word use between The Time of London and Kenya's Daily Nation. The researcher used the news sections of The Time and Daily Nation during 2013 which all the content indexed by Factiva. The study focused on the equivalent comparison across outlets and excluded the most irrelevant items such as recipes, music reviews, opinion columns and wire stories. This study include a total of 53.638 new articles, 1.992 from Daily Nation (759.720 words) and 51.646 articles from The Times (19.285.619 words). Moon compare the two list by its parts of speech, specifically for the pronouns, adjective, nouns and verbs in its top 500 words, then examines the characteristics of the two corpora which represent the news style of press system in different political and social environment. From the analysis, Moon agree that the two news publications, the Daily Nation and The Times, show some key differences that indicate stylistic and topical variation. Greater word frequency of both politics and geography were generally devoted in the Daily Nation, and it tends to investigate the topics in terms of internal news and institutional power, whereas more space of international and global news and citizens were devoted in The Times. From the investigations above, we can see the study of corpora have been developed faster along with the improvement of computer technology. It makes corpora used in widely range of linguistics research and in various language fields.

From the statement above, we know that corpora have been developed rapidly over time. Based on the previous studies, corpora have a wide range of research. In the corpus study, many research conducted by the researcher based on their purposes. (McCarthy, 2004: 2) Corpora exist for many language and are collected for different purpose. Learner's dictionary, grammar reference materials, vocabulary learning materials, and more recently, course books have all benefited from the information in corpora. But based on the previous studies, only few researchers try analyze the most frequently verb in the short story in the middle school. In analyzing the short stories, the researcher used Antconc 3.2.4 and TagAnt software. The significance in this study will find the most frequently used verbs in the short stories for junior high school students. The researcher is interested in conducting a research entitled **“The Corpus Analysis of the Most Frequently Used Verb in the Children Short Stories for Junior High School Students”**.

B. Statement of the Research Problem

Based on the background of the study above, the questions that formulated by the researcher is “What are the most frequently used verb in the children short stories for junior high school students?”

C. Purpose of the Research

Based on the research question above, the purpose of the study is to know the most frequently used verb in the children short stories for junior high school students.

D. Significant of The Study

Following on resolving the research problems regarding the most frequently used verb in the children short stories for junior high school students, the finding discovered by this study can provide sort of contribution towards English second language education in Indonesia be it theoretically or practically. The result of this study that concerned on the most frequent verbs found in the corpus of selected children short stories can benefit not only for students in enriching their vocabularies knowledge and its proper application into sentences both when communicating orally or written, but also can benefit English teacher or lecturer in determining and selecting the appropriate teaching material containing numerous verbs that familiarly used in the children short stories for junior high school students. By the finding of this research, the information collected from numerous verbs can be used in sorting a database of verbs or dictionary. This study can also provide additional resource towards future researcher in deciding the topic based on the gap they can find. So, they will be able to do further investigation that can produce a new discovery relating to the corpus study and verbs in the future.

E. Scope and Limitation of the Research

In order to avoid the deviation from the purpose of the study, this study focus only to analyze the most frequently used verb in the short story. The short stories were collected and downloaded from the particular website which provide a complete collection of short story. The researcher only

selected the short story which is aimed and appropriate for junior high school level or child for ages 13 years old and over. The weakness in this study is the Small amount of the short stories. This small corpus might disadvantage to fully cover or represent accurately the short stories in general. Also through the small amount of corpus some of data needed might still lacking than other source for instance instructional materials such as English textbook. So for the next researcher hopefully they can provide a large amount of the short story and build a large size corpus.

F. Definition of the Key Term

1. A corpus is a collection of texts, written or spoken, usually stored in a computer database. (M. McCarthy 2004:1). According to Grigaliuniene (2013: 9) corpus refer to a collection of machine-readable, authentic texts, chosen to characterize or represent a state or variety of a language. In this study, the researcher concluded that corpus is collection of authentic texts stored in computer database and used by linguists or researchers with the purpose of research.
2. Frequency McCarthy (2004: 10) states that frequency is the most basic tool in analyzing the text in corpus. A frequency lists tells what words and phrases are used most often. In this study researcher concluded that frequency refers to total of verb that happens or repeats many times in the text.
3. Child short story Oberstein (1999: 15) mentioned that the definition of children literature is a category of books that existence of which

absolutely depends on supposed relationships with particular reading audience: children. While short story define as prose fiction that is shorter in length than novel, so it can be concluded that child short story is prose fiction with concise form which is absolutely aimed to particular reading audience i.e. children.