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 definitions of the key terms.

A. Background of the Research

There have been numerous researches that registered as studies based on the idea of using electronic corpora over the last five decades. Corpus or corpora in plural form, refers to collection of text, written or spoken, which are organized systematically according to the purpose of its construction. As what had been conveyed by Linquist (2009:01) that the idea of using electronic corpora was began around 1960s, since the development of computers started becoming the powerful reason. Afterwards, corpus investigation have been already used in many general linguistics and also especially in English linguistics even in many different aspects of language teaching researches. Linquist (2009:03) also stated that W. Nelson Francis and Henry Kucera at Brown University were the first pioneer who compiled the first electronic collection of English texts to be used for linguistic research in 1961. They compiled 500 samples of 2.000 words each from a selection of various text genres in present-day American English, and the collection of texts was commonly known as Brown Corpus. The total of words in Brown Corpus attained 1 million words length. Thus, it's known as first

representative general purpose corpus, carefully designed, sampled, and prepared. It gave huge contribution to the new era of corpus linguistics. It is available to any researcher who want to investigate and became the most cited resource in the field of corpus linguistics.

The great effort committed by the pioneer of electronic corpora organiser, Francis and Kucera, in building Brown Corpus opened a great way to the other leading linguists in following the similar tried. The existence of Brown Corpus inspired the others linguists to conduct the same study. The Brown Corpus later followed by arose a number of other corpora, such as the Lancaster-Olso/Bergen (LOB) Corpus, which used the same format as the Brown Corpus and made it possible to compare different varieties of English (Grigaliuniene, 2013:09). Afterwards, followed by the first corpus of spoken English named the London-Lund Corpus of Spoken English or LLC which was developed in 1975. According to Leech (1992) cited by Kholkovskaia (2017) these two corpora became the model for other comparable corpora of the so called "Brown Family" that contained of seven corpora covering the period between 1931-2006. Along with the advance of computer technology, other numerous corpora which contain larger amounts of text or commonly known as second generation corpora appeared. The corpora such as which represented by John Sinclair's Birmingham Collection of English Texts (The Cobuild project), the Longman/Lancaster English Language Corpus, the British National Corpus (BNC), the International Corpus of English (ICE), etc. (Grigaliuniene, 2013:09).

As corpus is organized which contains of collection of various texts, written or spoken, that produced in real use of the language, it can be used to study certain parts of any particular language. As stated by Linquist (2009:01) that if we are interested in any workings of particular language and we want to study it, for example English, it is a good idea if we study the language in its use. Thus, the fact that corpus is build based on the authentic data of language workings, it can be used to study about many facts about language, for example the language description, grammatical features, and many others. Corpus can give us many information how language is used naturally. By using corpus based study linguist can observe how language is used. In addition, the academic researchers also can use corpus study to examine how corpus could be used in teaching and learning particular language. By using corpus, it is possible to be shown to the students how language is used naturally. Hence, it would make the teaching of language become more effective to reach the students' needs in learning particular language. In conclusion, the existence of corpora has many contributions in language research by the linguists or used in language teaching and learning.

Nowadays, the idea of corpus study have been developed in many language research even in more various aspects of practical language. Corpus linguistics allows us to see how language is used today and how language is used in different context (Bennett: 2010). There were numerous corpus based study that conducted in various particular language workings, such as fiction, magazines, conversations, newspapers, academic papers, literatures and so on. For example, a corpus-based analysis conducted

by Kartal (2017) which aims to determine the most frequent adjectives used in academic texts and investigate whether the adjectives have difference of frequency and function in social sciences, technology, and medical sciences. In this study, the Contemporary Corpus of American English (COCA), a contemporary and genre based corpus, was used. This corpus included over 450 million words which covers the years between 1990 and 2012. COCA contains five sub-corpora namely spoken, fiction, magazines, newspaper, and academic.

Corpus based study also can be used to examine the language used in particular part of media mass publications. Several researchers had been conducted corpus study of media mass products. They might observe the language description and the grammatical features of the media mass texts. For 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 verbs of the horoscopes. For this study, the corpus consisted of horoscope from 108 issues of 3 magazines namely CLEO, Female, and the Malaysian Women's Weekly from the year 2011-2013. From the result of this study, Ling found that the most frequent word which appear in horoscope is pronoun "You" with 5.01 % of total occurrences. Then, the most frequent modal auxiliary verb used in horoscope was modal of probability "will" and followed by "may". Modal "will" was counted 76.46 % on its occurrences of total modal verbs occurrences in the corpus. It could be concluded that the writers of horoscope try to predicts and state the possible future events on their horoscopes.

Corpus study also developed in observing the comparative analysis of the news coverage between two media mass in different countries. One of the comparative studies was conducted by Chen (2013). He tried to investigate a comparative analysis of the news coverage of the 2011 Libyan civil war in two national media, those are China Daily and The New York Times. The selected articles were published between March 1, 2011, and October 23, 2011. The data analysis was used to generate and compare the word frequencies and analysed collocations of lexeme "Gadhafi/Qaddafi". Word "Qaddafi" was chosen because of the principal role of Muammar Qaddafi during the Libyan civil war. From the study could be revealed the results that the two corpora had a lot of similarities in word frequency, but there found differences in several high ranking lemmas too. In addition, the result of the collocation analysis of the two corpora revealed that there was a contrastive images of word "Qaddafi". In China Daily corpus, the collocations of "Qaddafi" included 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.

The comparative study of two news report corpora was also 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 Times and the 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. The data compiled in this study included a total of 53.638 news articles, 1.992 from the Daily Nation (750.720 words) and 51.646 articles from The Times (19.285.619 words). Moon compared the two word lists by its parts of speech, specifically for the pronouns, adjectives, nouns, and verbs in its top 500 words, then examined the characteristics of the two corpora which represent the news style of press systems in different political and social environment. From the analyses, Moon agreed 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 that 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. Corpora nowadays built in different language uses with different sizes based on different purposes of its studies. Such as what McCarthy (2004:02) argued that corpora now appear for many languages and compiled for different purposes. Corpus might be possible to use in organising dictionaries, grammar reference materials, and vocabulary learning materials, even to make a language course book. It can be seen that the study of corpus has given many contributions on the language fields, even in language

teaching. The advantage of using corpora is materials which are developed by using corpora can be more authentic and represent the real use of language.

But, according to several previous studies that investigated above, only few researchers who tried to study the most frequent verbs that used in news report articles. Thus there were not much investigations that can be used to find any source about verbs that mostly used in news reports. Therefore, at this present study the researcher wanted to investigate about the most frequently used verbs are in the news report articles. The knowledge of the most frequent verbs can be used to improve the students' vocabulary mastery. Knowing a word or vocabulary means knowing about meanings, word use, word formation, and word grammar (Harmer: 1991). The finding of this study was expected can be used to help the students in mastery various verb words and the function in English as vocabulary is an important component in foreign language learning. The mastery or language components has an important role to develop the language skills, such as speaking, listening, writing and reading. As stated by Richard and Renandya (2002: 265) that vocabulary is a core component of language proficiency and provides much of the basis for how well learners speak, listen, read, and write. Therefore, the mastery of various vocabulary has a great role in learning English.

While, particularly the knowledge of vocabulary of verbs also has a great important role in learning English too. In having a well communication, the students are mastering nouns, adverbs, and adjectives vocabulary only is not enough to make their communication in English well. They need to knowledge how verbs are used in

communications too. Verb is used to indicate, show, and state what they want to talk while communicating. Thus they need to use verb to communicate or transfer their idea to conduct a communication. Therefore, the researcher was interested to study about verbs. Talked about improving vocabulary, there are some effective sources that can be used to improve vocabulary knowledge of the language learners, for example from the practical language text such as news article. The reason why the researcher chose news article because news article is English written text that used by native speakers in their daily life. News article is one of reading materials that naturally used by the native of English. By studying the words, for example the verbs, which used in news article can help the students in improving their English proficiency effectively. Verbs have a significant role in building a news report. As we know that reporting news need to use appropriate verbs to deliver the information properly and make the readers enable to understand the information. Thus, it can be concluded that by knowing the vocabulary, especially verbs, in news article, can be used by the students to improve their understanding in reading authentic text that naturally used by the native and help their practice to used verbs in real communication.

This present study was conducted to find out the most frequent verbs that used in the news reports. According to Klavans and Kan (1998), verbs in news article are important in providing an event profile. By knowing the most frequent verbs used in news articles, it can be used to help the students in understanding an event description properly. This present study was conducted by analysing 317 news articles which

reporting some events or people from around the world to find out the verbs that are most frequently used there. The researcher is interested in conducting a research entitles **"The Corpus Analysis of The Most Frequently Used Verb in News Articles"**.

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 verbs in the news articles?"

C. Purpose of the Study

Based on the research question above, the purpose of the study is to know the most frequently used verbs in the news articles.

D. Significant of The Study

The researcher expects that this study will give contribution for:

1. The Students

    The researcher expects that the students will be able to acknowledge the most frequent verb used in news articles, so it would help the students in understanding explanation or description of events.

2. English Teachers

Teachers can provide an authentic material based on the most frequently used verbs in the new articles. So that, teachers are able to give an example about the proper verbs to describe events.

3. The other researchers

The result of the study can be used as the reference for another researcher to conduct a research with similar study or to find the deeper, more detailed result and also be able to investigate and correct the shortages from the previous study.

E. Scope and Limitation of The Research

In order to avoid the deviation from the purpose of the study, this study focuses only to analyse the most frequent verb that used in the news articles based on an American broadcasts websites http://www.foxnews.com/world.html. The weakness in the study is the small amount of the news articles around 317 titles and 103.203 words. So for the next researcher hopefully they can provide a large amount of the news articles and build up a bigger corpora.

F. Definitions of the Key Terms.

1. A corpus is a collection of texts, written or spoken, that represents language in the real use, and usually stored in a computer database (McCarthy, 2004: 1).

2. News article is article that used to extend some facts or information or news about an event or something that published by broadcasting company. In this study the news article was published online in a website.

3. Verb is a word or phrase that used to indicate an action, an event, or a state (Oxford Advanced Learner's Dictionary, 1995: 1323).

4. Frequency in this study refers to the rate of something, when it is happened or repeated.