CHAPTER I

INTRODUCTION

This chapter presents some topics that related to the study. This topic covers Background of the Research, Statement of the Research Problems, Research Objective, The Significance of The Study, Scope and Limitation of the Research and Definitions of the Key Term.

A. **Background of the Research**

The idea of using electronic corpora began around 1960s. According to Meyer (2004 : 1) the first machine-readable corpus, the Brown Corpus was created by W. Nelson Francis and Henry Kucera and dominated by the generative grammar. (McEnery and Gabrielatos 2006: 33) Since then, electronic corpora have come to prominence as a resource used by linguists. Lindquist (2009: 3) also states that when the first electronic collection of English text or Brown corpus compiled at Brown University US, the creator wanted to create a balanced and representative picture of printed and proceeded to collect 500 samples of 2.000 words each from selection of various texts genres.

Since then, the Brown corpus became a pioneer for many leading linguists. They were inspired in create balanced corpus, i.e. a corpus that give a fair picture of the whole of the language. (Grigaliuniene 2013: 4) The model is especially important since it was later followed, first in a British version called the Lancaster–Oslo/Bergen Corpus (LOB) for 1961, which used the same format as the Brown

Corpus and made it possible to compare different varieties of English. And much later followed by a team in Freiburg who created "clones" based on material published thirty years later: the Freiburg Brown and the Freiburg LOB (FLOB). In fifty years since 1961 number of corpora arose, Linguistics has gradually extended its scope and influence. Soon after computers technology became more powerful, it is give a big impact in the field of corpus linguistics to develop faster and faster.

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 Kenji Kitao (2014) conducted a study focused in analyzed the responses to apologize. The research used 320 apologize from a corpus made up of dialogue from 72 episodes from the US situation comedy *Modern Family.* The result of this study show that about half of the apologies studied, there was no response. In half of the apologies, there was a response. The majority of the responses – four out of five – fell into one of three categories: minimizing the offense, focusing on the offense, and responding to the justification/explanation/ question. Of these, the most common was minimizing the offense.

Then, Moinnvaziri (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 of this study show that the vocabulary presented in the reading passages of the textbook is mostly suitable 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 words 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 in explored the used of phrasal verbs in English language documents of the European Union (EU). The study was part of a larger-scale project examining the use of English in EU texts 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 Lexical Frequency and Range computer programs by heatley is a tool that used by the researcher to establish the frequency list and to compare it to the list of the most frequent adverbial particles and lexical verb + adverbial particle constructions in the written part of British National Corpus (BNC). In order to examine the most frequently occurring lexical verbs, the researcher used the concordance function of the Wordsmith Version 2 computer software. The result of the research for the most frequent phrasal verb show that the top 25 phrasal verb (PV) proved to be the most productive in forming verb + particle constructions in the Corpus of EU English (CEUE). A total of 130 lexical verb-types were identified in 187 verb + AP types (e.g., point + out, cut + down, etc.) and in a total of 1031 phrasal verb constructions. Interestingly, about half of the verbs on the list are among the 20 most frequent

lexical verbs forming phrasal verb combinations in the BNC Written (set, carry, put, make, find, move, take, work, break, bring, go, point).

A study conducted by Kartal (2017) aim to determine the most frequent adjectives used in academic texts and to investigate whether these adjectives differ in frequency and function in social sciences, technology, and medical sciences. This study focuses on the academic texts of The Corpus of Contemporary American English (COCA), a contemporary and genre-based corpus. COCA was used for the research because it is free to access, very comprehensive and highly representative data which includes over 450 million words. In selecting the adjective from COCA corpus, the researcher accessed the data from http://www.wordfrequency.info. A free list of the 5,000 most frequent words in COCA was used, and 839 of the words in the list were adjectives. In other words, 17% of the most frequent words in COCA are adjectives. Then, from the list of 839 adjectives, the ones most frequently used in the academic division were extracted. The new list, which is the focus of the study, included 334 adjectives. The results of the study revealed that 334 of the 839 adjectives in COCA were more frequently used in the academic sub-corpus when compared to adjectives used in spoken language, fiction, magazines, and newspapers. All the adjectives which are more frequently found in the academic register were used more in social sciences (history, education, geography/social science, law/political science, humanities, philosophy/religion) than in medicine and technology. The only exception was environment, which is used 11,872 times in social sciences and 21,992 times in science/technology and medicine.

Tran (2013) investigated the use of two synonymous evaluative adjective Interesting and Nice and in which structure they mostly appear in academic writing and speaking English. The researcher used four Corpora to analysis the real-life usage of the evaluative adjective. The first was the Corpus of Contemporary American English (COCA) created by Mark Davies. Data in COCA were collected from a wide range of authentic contexts and divided into spoken, fiction, magazines, newspaper, and academic texts. The second corpus was Michigan Corpus of Academic Spoken English (MICASE). MICASE is a specialized corpus and focuses on the language that occurs in academic speech. The third corpus was the Google Books collection, accessed by Google's corpus tool N-gram Viewer. And the fourth was small corpus of 56 daily conversations and talk shows from TalkBank and the David Letterman show. It was analyzed by using Antconc 3.2.4 software. The result of the research show that the evaluative adjective of Interesting and Nice near synonyms, they differ in grammatical behaviors, lexical collocations, and pragmatic functions. The researcher also designed the teaching material to increased students awareness in the used of these adjective. The result of the study shows that students' attention was drawn to the grammatical patterns with interesting and the functions of nice in speech.

Mazdayasna and Firouzi (2013) conducted a study to investigate the frequency distribution and syntactic functions of adjectives across literary and technical texts. The literary and technical texts were analyzed from the register perspective as well as genre perspective. The data were collected from two corpora, namely novels and engineering academic books in English. Each corpus consisted of

five literary novels and five engineering academic books. Then, fourteen paragraphs were randomly selected from each text. A total number of seventy paragraphs were randomly extracted from five engineering academic books as well as seventy paragraphs were randomly drawn from five novels and the frequency distribution of adjectives with respect to their position (attributive, predicative) and syntactic functions (descriptive, verbal, numeral, etc.) were tallied and summed. The result of the study shows that there is a significant difference in the frequency use of adjectives across the two corpora. From a register perspective, the high frequently use of adjectives in technical texts in comparison to literary texts (67.3% and 32.7%, respectively).

Saad (2016) tried to investigate three the Protagonists characters in Harry Potter and Deathly Hollows series. According to the researcher, the Protagonist characters play important roles in any novel. So in J.K. Rowling's Harry Potter novels, the characters were presented in the most vivid of ways so as to make the characters alive and memorable. Based on the reason above the researcher conducted a study o find out the frequency of adjective that was used to describe three Protagonists characters in the novel Harry Potter and the Deathly Hallows. Then, the data from the novel Harry Potter was analyze with CLAWS and Antconc 3.2.4 software. The result of the study shows a very distinctive recurring adjective for each character which suggests that each one of them is a different unique entity. Further examination suggests that by look at the character's adjectives alone would not be intuitive in finding the deeper characterization of these individuals.

Sun (2016) focused on the analysis of vocabulary coverage in children 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 settings in Taiwan. Teachers, students, and parents perceive extensive reading as impractical due to the lack of immediate measurable gains in test scores. So, to overcome these difficulties, reading a large quantity of target language materials can aid learners in acquiring new words, and incidentally strengthen previously learned ones. The researcher used 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 word list published by The Ministry of Education in Taiwan. This three word list is a guideline for book developers to follow. The findings indicate that extensive reading of the Magic Tree House series can support word list learning due to its wide coverage of vocabulary from the three lists, especially the *Middle School Basic 1,200* list. The 48 books of the series cover 91.4% of *Basic 1,200* words, 36.4% of *Advanced 800* words, 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 word lists. 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 will not only expose learners to a large portion of words from the lists, but many of them repeatedly. According to all

previous studies, there are still a few studies focus on the analyzing the most frequently used adjective in the children short story for middle school students.

From the statement above, we know that corpora have been developed rapidly over time. Based on the some 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 languages and are collected for different purposes. Learner's dictionaries, grammar reference materials, vocabulary learning materials, and, more recently, course books have all benefited from the information in corpora. But based on the some previous studies, only few researchers try to analyze the most frequently adjective in the short story for middle school. A study conducted by Kartal analyzed the most frequent adjective but the researcher focus in academic texts of COCA Corpus. Then, study conducted by Saad also analyzed the most frequent adjective that described the protagonist character but the researcher focus in the novel of Harry Potter. So that only a few journal or research which can be used to support the theory of corpus of children short stories. Therefore this present study is intended to investigate what are the most frequently used adjective in the children short story for middle school students?

Thus, this present study was conducted to find out the most frequent adjective that used in the short story. As Jara (2012) said students needed to understand the important of the adjectives when they wrote a description about people and objects. Bartan (2017) in his study said that in order to improve students writing skill through short stories there are some list of criteria to select short story. One of the criteria is the effective and powerful adjective. So that, it was considered that short story needed

to contain the example of effective use of adjectives. Therefore, this present study is done by analyzing sixty eight short stories to find the most frequently used adjective. 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 adjective in the short stories for middle school students. The researcher is interested in conducting a research entitled **"The Corpus Analysis of the Most Frequently Used Adjective in the Children Short Stories for Middle School Students".**

B. **Statement of the Research Problems**

Based on the background of the study above, the questions that formulated by the researcher is "What are the most frequently used adjective in the children short stories for middle school students?"

C. **Research Objective**

Based on the research question above, the purpose of the study is to know the most frequently used adjective in the children short stories for middle school students

D. **Significance of The Study**

The researcher expects that this study will give contribution for:

1. The Students

Students are able to understand about the most frequent adjective that used in the short stories to help them in giving the description about people and objects.

2. English Teachers

Teachers can provide an authentic material based on the most frequently used adjective in the short stories. So that, teachers are able to give an example about the proper adjective to describe people and objects.

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 focus only to analyze the most frequently used adjective in the short story. The short stories were collected and downloaded from the particular websites which provide a complete collection of short story. The researcher only selected the short story which is aimed and appropriate for middle school level or child for ages 13 years old and over. (https://www.rethinkela.com/2014/05/40-excellent-short-stories-for-middle-school/, https://americanliterature.com/middle-school-short-stories, https://www.weareteachers.com/best-short-stories-for-middle-schoolers/). The

weakness in the study is the small amount of the short story around sixty eight short stories which the total of the word token are 201442 words. So for the next researcher hopefully they can provide a large amount of the short story and build a larger size corpus.

F. **Definitions 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 adjective that happens or repeats many times in the text.

3. The short story is a concise form of prose fiction (M. Klarer, 2004: 13). According to Cavanagh et al. (2010: 103) the characteristic of the short story is that it is a story which is short. Thus it can be distinguished from a novel, which is a story that is long. So from the statement above, short

story is prose fiction that is shorter in length than a novel. Therefore short story has a unique characteristic such as focus on one incident, has a single plot, a single setting, and small number of characters.

4. 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