CHAPTER III

RESEARCH METHOD

This chapter discusses about research design, data and data resource, technique of data collection method, technique of data verification and data analysis.

A. Research Design

This study aims to find out the most frequent verb and their role in fairy tales, so the research design that is used in this study by corpus-based study. Bianchi (2012:31) says that corpus (plural:corpora) is collection of texts in electronic form. Further, Kaewphanngam et.al. (2002:1) stated that corpus-based study is a way to analyze a naturally occurring collection of languages either in the form of recorded utterances or written text. Thus, the corpus is a collection of languages stored in electronic media, which can be used later to analyze large quantities of words, which are difficult to do manually.

This research is using corpus as subject as well. The corpus that used in this research is the collection of fairy tales in written form. Further, the data analysis is also done by using corpus tool, such as identifying the frequency of the words most frequently used in fairy tales, and marking parts of speech to see each part of the word that was frequently used. Grigaliuniene (2013: 43) says the frequency list is a list of all words that appear in the corpus which can be sorted alphabetically or by frequency to find the words that appear the most and least frequently in the text corpus. By using a frequency list, it can help researchers find the most frequently

used words, especially verbs. Meanwhile, part of speech tagging is used to make it easier for researchers to find verbs accurately and also categorize verbs into their respective roles.

B. Data And Data Resource

This study is a corpus-based study. Therefore, the corpus (particularly in the written form) is the only data used in this study. Bianchi (2012:33) says that the design of a corpus depends on the purpose for which the corpus is created. Consequently, it requires the application of selection criteria and sampling according to the purpose of the analysis. The data that is used in this research is collection of fairy tales on internet in written form, so it is the definite population because there are so many fairy tales on the internet. However, the researcher selects 60 fairy tales as sample. The sample are chosen by using simple random sampling. This sampling method allows every member of population has same chance to be chosen as sample. The researcher close the eyes, then pointed to screen that show the title of fairy tales. Whatever story that are chosen, they are the sample in this research. The sample of this research is presented as below:

Table 3.1 List of Fairy Tales

No	Author	Title	Source
1	Templeton Moss	A Quiet Boy	https://fairytalez.com/
2	Karen Melik (Shakhnazar)	A Kind Heart	
3	Raahul Pundit	Little Champ	
4	Shereen Sawalha	Catching Magic	
5	J.d. Howse	Far Far Away	
6	Yaser Faris	The Bonsai Tree	
7	April Eight	Grammy and the Star Soup	
8	Victoria Stasiv	Darynka	

9	Marjorie Ayen	Philip and His Good Deeds	
10	Johnny Wordsworth	The Boy Who Had Lost His Voice	
11	Jemma Hathaway	The Aria and the Ivy	
12	Lousille Ryman	Puss in Platforms	
13	Auntie Millie	The Troubles of Mrs. Tweedy	
14	Owen Meyerowitz	Milk Snake Mayhem on Monday	
15	Nina Adinou	The Snow Bunny	
16	Jordan Williams	Orange Flames	
17	Sergey Nikolov	Princess Rose and the Golden Bird	
18	Yaser Faris	The Six Gum Balls	
19	Cice Rivera	Morwenna and the Monkey	
20	Sharon Jonnalagadda	The Night of a Thousand Hours	
21	Brothers Grimm	The Elves and The Shoemaker	https://www.storynory.com/
22	Brothers Grimm	The Brave Little Tailor	
23	Bertie	Birdy's Halloween	
24	Hans Christian Andersen	The Tinderbox	
25	Hans Christian Andersen	The Swineherd	
26	Bertie	The Missing Bunny	
27	Joseph Jacobs	The Three Sillies	
28	Bertie	The Wild Man	
29	Bertie	The Sorcerer's Apprentice	
30	Letitia Maclintock	Halloween at the castle	
31	Hans Christian Andersen	The Flying Trunk	
32	Hans Christian Andersen	The Nightingale	
33	Bertie	The Witch Who Got into Trouble at School	
34	Bertie	Katie and The Forgetfulness Spell	
35	Hans Christian Andersen	The Princess and the Pea	
36	Bertie	The Moscow Cat	
37	Brothers Grimm	Rumpelstiltskin	
38	Andrew Lang	Goldilocks and the Three Bears	
39	Andrew Lang	Ali Baba and the Forty Thieves	

40	Lafcadio Hearn	The Old Woman Who Lost Her Dumpling	
41	James Baldwin	Androclus and the Lion	https://americanliterature.com/
42	Hans Christian Andersen	The Little Match Girl	
43	Oscar Wilde	The Selfish Giant	
44	Charles Perrault	Puss in Boots	
45	Frank Stockton	Huckleberry	
46	Frank Stockton	The Sprig of Holly	
47	Hans Christian Andersen	The Ugly Duckling	
48	Brothers Grimm	Hansel and Gretel	
49	Joseph Martin Kronheim	The Little Old Woman Who Lived in a Shoe	
50	Peter Christen Asbjornsen	East of the Sun and West of the Moon	
51	Parker Fillmore	The Forest Bride	
52	Brothers Grimm	Hans the Hedgehog	
53	Hans Christian Andersen	The Little Mermaid	
54	Joseph Jacobs	The Sprightly Tailor	
55	Hans Christian Andersen	The Travelling Companion	
56	Brothers Grimm	The Girl Without Hands	
57	Brothers Grimm	The King of the Golden Mountain	
58	Charles Perrault	Little Thumb	
59	Joseph Jacobs	The Sea-Maiden	
60	Grace James	Reflections	

Data sources are sources that the researcher has obtained in collecting the data, and corpus is the only source of data in this study. There are sixty fairy tale titles used in the corpus. The researcher chose online source for the data, because online source can be accessed anytime and anywhere. In this study used some website for sources of the data are https://fairytalez.com/,

https://www.storynory.com/, and https://americanliterature.com/. They are trusted website on internet which presenting a lot of fairy tales from around the world.

C. Technique of Data Collection

The data collection method of this research is documentation. Flick (2018:378) says that documentation in form of paper or digital is the method to collect data when the data is spread out. Further, he also says that documents include both the written elements of texts and the extra elements that are attached in documents. In this study, the researcher used fairy tales as the data used to compile the corpus. The fairy tales are collected from several websites. Some of the websites chosen in this study provide a complete collection of fairy tales. The first, researcher collects the data by copying the fairy tales from source to PC software, namely Microsoft Word. Then, moved to notepad. It is done to create UTF-8 encoded file or file in form of text (*.txt). The story are the sample that has been chosen randomly. After that, the researcher save it in a folder so that makes the verifying data done easily.

D. Technique of Data Verification

In the study, the data must be valid or have been through verification. In verifying data truths in the verbs, the researcher uses TagAnt 1.2.0 which is a simple freeware tagging tool from Laurence Anthony. TagAnt is used to tagging the part of speech in every word of a collection of text in corpus fairy tales before evaluated using AntConc. Each tagged word is included in the word class tagging, for example, nouns, verbs, adjectives, and others. After tagging part of speech in corpus fairy tales, the researcher sorted and analyzed the words tagged as verbs.

Thus, the data used in this study is truly a verb and does not include the researcher's presumption.

E. Data Analysis

Data analysis is a crucial process because the researchers faced with large amounts of data collection and examination which will later be used to analyze and interpret data. According to Schwandt (2007:6), analyzing data is a series of processes in organizing, reducing, describing data, and proceeding with drawing conclusions or interpreting data, then guaranteeing the interpretation of the data. In analyzing the data on this study, there are several processes including:

Firstly, the researcher downloaded and collected the fairy tales from the internet which were to be used to make the corpus. Selecting the websites which provide a complete collection of fairy tales. The collection of fairy tales were stored and collected in Microsoft Word format files, this can make it easier for researchers to check for errors in writing words in the text easily before being analyzed. Previously, the corpus software here can only be read with a TXT format file. So the fairy tales that have been collected are copied to a TXT format file so the data can be read. In converting corpus files to TXT, the researcher uses Notepad++ 7.9.3 free software program which can be accessed through https://notepad-plus-plus.org/.

Then, the results of TXT are analyzed to find out the frequent verb by using PC software, namely AntConc 3.5.9. AntConc is corpus analysis tool kit by Laurence Anthony, Ph.D. This software is used to know the frequent POS, after each POS is tagged by TagAnt 1.2.0. TagAnt is corpus analysis tool kit that created

by Laurence Anthony, Ph.D. It is an easy freeware tagging instrument based around the TreeTagger tool (developed by Helmut Schmid) that takes either a input content or an input list of content files (UTF-8 encoded) or in form of .txt and tagging the Parts-Of-Speech (POS). Those software are easy to use and also freeware. The POS of the context is marked as follow:

Table 3.2 List of TagAnt's Part of Speech Code

POS Tag	Description	Example	POS Tag	Description	Example
CC	Coordinating conjunction	and, but, or, &	VB	Verb be, base form	be
CD	Cardinal number	1, three	VBD	Verb be, past	was/were
DT	Determiner	The	VBG	Verb be, gerund/participle	being
EX	Existential there	there is	VBN	Verb be, past participle	been
FW	Foreign word	d'œuvre	VBZ	Verb be, pres, 3rd p. Sing	is
IN	Preposition/subord. Conj.	in,of,like,after, whether	VBP	Verb be, pres non-3rd p.	am/are
IN/that	Complementizer	That	VD	Verb do, base form	do
JJ	Adjective	Green	VDD	Verb do, past	did
JJR	Adjective, comparative	Greener	VDG	Verb do gerund/participle	doing
JJS	Adjective, superlative	Greenest	VDN	Verb do, past participle	done
LS	List marker	(1),	VDZ	Verb do, pres, 3rd per.sing	does
MD	Modal	could, will	VDP	Verb do, pres, non-3rd per.	do
NN	Noun, singular or mass	Table	VH	Verb have, base form	have
NNS	Noun plural	Tables	VHD	Verb have, past	had
NP	Proper noun, singular	John	VHG	Verb have, gerund/participle	having
NPS	Proper noun, plural	Vikings	VHN	Verb have, past participle	had
PDT	Predeterminer	both the boys	VHZ	Verb have, pres 3rd per.sing	has
POS	Possessive ending	friend's	VHP	Verb have, pres non-3rd per.	have
PP	Personal pronoun	I, he, it	VV	Verb, base form	take

PP$	Possessive pronoun	my, his	VVD	Verb, past tense	took
RB	Adverb	however, usually, here, not	VVG	Verb, gerund/participle	taking
RBR	Adverb, comparative	Better	VVN	Verb, past participle	taken
RBS	Adverb, superlative	Best	VVP	Verb, present, non-3rd p.	take
RP	Particle	give up	VVZ	Verb, present 3d p. Sing.	takes
SENT	End punctuation	?, !, .	WDT	Wh-determiner	which
SYM	Symbol	@, +, *, ^, |, =	WP	Wh-pronoun	who, what
TO	To	to go, to him	WP$	Possessive wh-pronoun	whose
UH	Interjection	Uhhuhhuhh			

The step to use TagAnt and AntConc is quite easy to follow. The first thing to do is operating TagAnt. The data that have been saved imported on TagAnt. After that, click start button on software. Then, the software start to duplicate the file, and automatically tag every part of speech on duplicated file. The last move is operating AntConc. The file that have been tagged imported on AntConc. The researcher opens the menu Cluster/N-grams tool, which serves to find out the frequencies from the most appear to the least appear in fairy tales. In addition, there is a range and word type of verbs.

From the results of the frequency list of verbs that are often used in fairy tales in the AntConc software, the researcher chose the top 130 verb lists for analysis. The list of verbs which is in AntConc copied to Microsoft Office Excel. Then, sorted according to their frequency, and calculate the percentage of the frequency list of verbs. Furthermore, researchers of the classification of the role of verb based on the English course book are entitled English Verb Classes and Alternations by Beth Levin (1993). After knowing the classification of the role of

verb, the researcher analyzes the data by comparing the explanation in the book with the source data. The steps are quite complex, so the researcher does it carefully.