

CHAPTER II

LITERATURE REVIEW

The literature review of the thesis consists of the definition of corpus linguistic and its relation in the sociolinguistic research in general, the lexical reference analysis steps that the writer takes to examine the adjectives occurred, the types of adjectives in English in general, and the last is the review of related studies.

2.1. Corpus Linguistic and Sociolinguistic

Corpus linguistic is a method of carrying linguistic analysis (Nesselhauf, 2005). The method of corpus linguistic can be used in the investigation of many sub-fields of linguistic using analysis based on the corpus. A corpus can be defined as a “systematic collection of naturally occurring texts” (Nesselhauf, 2005) in both written and spoken form of language. The data for linguistic analysis can be divided into two sub-parts: data gained by intuition and naturally occurring language. Corpus linguistic method can be used to carry out new findings about language and has become one of the most method commonly used by linguist in their investigation nowadays (Nesselhauf, 2005). Corpora consists of text that make linguist to elaborate their analysis toward a field of language that is being studied by the linguist (Meyer, 2002).

In the sub-field of sociolinguistic, especially the study of language and gender, the study was prominently initiated by Lakoff (1973). In his research

entitled language and woman's place, Lakoff (1973) is elaborating how language spoken by a woman may give a glimpse about how the woman's behavior. For example, there are words that are considered as *woman only* words. The words considered as woman only words are *adorable, charming, sweet, lovely, and divine*. These words are more commonly used by woman than by man. Lakoff (1973) also portrayed the words that may portray a woman's social statuses in the society, such as the usage of the word *lady* and *woman*. Although it may have a similar meaning, the usage of these two words is different. According to Lakoff (1973), for some speakers, a job that is considered as a low social class job in the society will more likely be described as a *lady*. For example, *cleaning lady* may sound as common as *cleaning woman*, but for *woman doctor*, if one is trying to say *lady doctor*, it will be considered as an insult. But there is no such thing in *men*, where *garbage man* or *salesman* is the only choice of word to refer to the job, and there is no *garbage gentleman*. And because male is not marked in the profession, there is no such thing as *man (male) doctor*. Lakoff (1973) here pointed out how the language usage, both on the side of the speakers and the listeners may give a different stereotype on the gender or subject referred. By this statement, there is clear that gender, in addition to the other social value, also plays an important role in determining the choice of words that is used by different gender in the way they speak.

Corpus linguistic researchers have already tried to elaborate the questions of how different types of people use language differently using spoken corpus data. According to Baker, the language use of a person or identity groups is generalized

by examining the presence or non-presence of specific linguistic variables: can be prosodic, phonetic, lexical, grammatical, discorsal, or pragmatic. Both written and spoken language can be examined by the researcher, even though most sociolinguists tend to focus on the spoken language one. The user of the language often divided into one or more categories based on their identities, such as the sex of the speaker. Most sociolinguist taken the speakers into multiple groups, such as a combination of sex, age, social class, occupation, geographic location, sexuality, etc (Baker, 2010)

An example of how the corpus linguistic method is used to carry out the language and gender sub-field is the research done by Schmid (2003) that analyze the sex differences in the 10 million words spoken corpora of the British National Corpus (BNC). Schmid (2003) searched for some words that reflected the state of “women’s words” and terms that is expected to be male preffered terms. Another example is the research done by Harrington (2008) that built a corpus of conversation occured between the male and female in British to distinguish the tendencies of how people will traditionally gender their language. Harrington found that several female speakers did use the “woman preferred” speech, and the rest of the female has the feature of speech which is similar to the males.

However, Baker (2010) stated in his book that the important point to remember in analyzing the sociolinguistic context throughout the corpus linguistic method is that the researcher must be careful in determining the findings, so that the findings can be categorized as an absolute findings rather than just tendencies. In looking at the frequency of the “woman only” words as stated by Lakoff

(1973), Baker (2010) also suggested that at some point the gender of the speaker is the only variable, but other variable such as age and social class can also has a role in analyzing the words usage.

2.2. Lexical Reference Analysis

The lexical items usage can be different according to who is the speaker and under what circumstances is that lexical items used. According to Savoy (2009), words that is used by one in one's talk can be analyzed and narrowed and later used to extract the terms that can lead into one's characteristics. In analyzing the lexical items, the researcher will need to understand how words can construct each other to make meanings and referential to the object referred. The meaning of a word will be vary, depends on the way and the context where the word is used. McEnery and Hardie (2012) stated that the lexical items interaction and placement in the structure of the sentence is the one that will develop the overall meaning of the sentences. Lexical analysis will focused on how the interaction and placement of the lexical items of the sentence may develop a different meaning.

The lexical analysis itself has a connection with the word frequency. As stated by Bybee (1985), the high-frequency words may form more remote lexical connections compared to the low-frequency words, and in the case of morphologically complex words, high-frequency words will experience less analysis and are less dependent on their base words compared to low-frequency words. And in order to analyze the lexical items that may shaped the gender-

stereotyping point in this research, the researcher will need to focused on the high-frequency words and also trying not to neglect the low-frequency words.

Freitas, Moscarola, and Jenkins (1998) have conducted research on context analysis and lexical analysis, emphasizing their content and use in analyzing a language's feature. They said that the context analysis and lexical analysis may presents a major common rational characteristic, being more or less intuitive, personal, and subjective. In this research, Freitas, Moscarola, and Jenkins areelaborating the qualities, types, application, stage, and problem, and has concluded that content and lexical analysis allows us to pass simple description and to reach the objective of every scientific research: the discovery of explanations and causal relationships

According to Freitas, Moscarola, and Jenkins (1998), there are seven levels of lexical analysis:

1. The summary lexicon approach: Reduce the body of the original text to the "high" or top of the lexicon.
2. The controlled lexicon approach: Reduce the text to its lexicon, and to control, through lexicon surfing or navigation, the validity and the foundations of the interpretations elaborated, starting from the lexicon
3. The selective lexicon approach: Work on a reduced lexicon after having eliminated or deleted the tool words.
4. Lexicon statistics and text quantification: The Research of the lexicon characteristics, to establish the statistics of the words of the text according to an external non-textual variable.

5. Codifying the lexicon and generating measures: Describe the text by a nominal variable, created by a codification based on the presence in the answers of a group of selected among the words from the lexicon.
6. Quantification of text: Calculate the number of times certain words occur, starting from a “lemmatized” text and already with the tool words taken out of the lexicon.
7. The multivariate analysis of the textual data: Data analysis made on the variables extracted from the lexicon.

Hence, not all 7 steps in this lexical analysis steps made by Freitas, Moscarola, and Jenkins (1998) are needed by the writer. The writer later decided to simplify the 7 levels into just 4, by selecting the most suitable levels needed for analyzing the gender stereotyping features in “*The Simpsons*”.

The 4 levels of lexical analysis that is going to be used by the writer are:

1. The summary lexicon approach
2. The selective lexicon approach
3. Quantification of text
4. The multivariate analysis of the textual data

In the summary lexicon approach step, the frequency of word occurrence is examined, which of those lexical items that can be used in analyzing the gender stereotype features. In the selective lexicon approach step, the examined lexical items will be reduced, and the research will more focused on the adjectives, with the dictionaries as the reference. In the quantification of text step, the frequency of words will be calculated, and then grouped up into the groups that may resemble

the research. In the multivariate analysis of the textual data step, the data is analyzed and perceived by using the statistics of the words.

2.3. Adjectives in English

In order to create understandable sentences, the sentences need to be constructed by the right words. In English, there are several parts of speech, or also often called “syntactic categories,” or word classes (Harley, 2006). There are several word classes: noun, verb, adjective, adverb, with suffix and prefix as the modifier of the words. Noun meanings can refer to: people, places, concrete things, abstract things, properties, and activities. On the other hand, adjectives and verbs is not as free as nouns. Adjectives always refer to properties, even when the adjectives are derived from nouns. Verbs always refer to events or states, even when the verbs are derived from nouns (Harley, 2006) but in English, one word can have more than one word class, depending on its position on the sentence.

In English, besides the word classes, the words are divided again into two categories: content word and function word. According to Harley (2006), content words are the words that considered as the core of the messages delivered, where function words are the words that is used to bond and connect the content words, to create a structured and understandable sentences. Or in other word, a sentence’s main message will always brought by the content words, but to create a more specific and complex sentence, the function words is needed to accompany the content words.

Content words consist of nouns, verbs, and adjectives. Harley (2006) stated that noun meanings can refer to: people, places, concrete things, abstract things, properties, and activities. On the other hand, adjectives and verbs is not as free as nouns. Adjectives always refer to properties, even when the adjectives are derived from nouns. Verbs always refer to events or states, even when the verbs are derived from nouns.

Harley (2006) also divided function words, into several subclasses:

1. Conjunctions: words that connect two same typed words. (e.g.: and, or, but, and although)
2. Determiners: words that sometimes placed in front of the nouns, or the nouns that has been affected by adjectives or any other words.
3. Pronouns: words that can be used to substitute a noun or a noun phrase.
4. Complementizers: words that refers to a complex clause or another sentence.

In short, content words are bond-free words, or can be analyzed within its own single word, where function words need a connection with another words in order to be analyzed, and the meaning is not exact, depends on in what content the function words are being used.

According to Sassoon and Gurion (2010), there are two kinds of adjectives: positive adjective and negative adjective. Positive adjectives has a correlation with the adjective's basic meaning, while negative adjective are the opposition of the positive adjective's meaning, and the positive and negative value of the adjectives are equal (Sassoon and Gurion, 2010). In the case of this research, the

writer rely more on the position and the context of the adjective usage in the sentences and its meaning on the sentences to determine whether the adjective is used as a positive adjective or a negative adjective.

Beside the positive and negative adjective, the writer also wanted to take a closer look to the distribution of the adjectives from each character. The distribution of the words is divided into two, distribution based on the age of the speaker and distribution based on the gender of the speaker. In order to take a look on whether a character is using the words suitable for the age or gender of the character, the writer used the corpus of British National Corpus Web Edition (later called BNC).

The BNC is a collection of approximately 100 million words which data is collected from both sample of written and spoken language from many sources, with the distribution percentage of ninety percent (90%) written source and ten percent (10%) spoken source. The writer use BNC to see the characterization of the words found and comparing the words to the characterization of the characters in *The Simpsons*, since the BNC provided several data and distribution of the words searched in the database.

However, the BNC provided the data of the words usage mostly on the source that is written on British English, while *The Simpsons* mostly used American English. The practical reason for choosing the BNC is that there is no American corpus which shows the distribution of age and gender. On the other side, BNC here concerned on the British English language, the page of type of

corpus provided by BNC mentioned that non-British English and foreign language is also included in the corpus.

2.4. Review of Related Studies

There are several research concerned on the area of language and gender. In this thesis, the writer chose only two researches which are closest to this thesis topic: “*Gender in The Simpsons: A Bachelor Thesis*” by Anna-Karin Mattsson (2009) and “*The Simpsons, Gender Roles, and Witchcraft: The Witch in Modern Popular Culture*” by Sarah Antinora (2010).

Anna-Karin Mattson (2009) in “*Gender in The Simpsons: A Bachelor Thesis*” analyzed the pictorial of the gender role in terms of language, female and male activities, and plot, using four episodes with equal lengths, and using six features for the analysis: names, adjectives, occupations, verbs, representations and gender roles. Using four episodes with equal length, this research is aimed to portray the way gender may be represented in the Television, especially in the series of *The Simpson*. The study carried out with the method of quantitative and uses the model of corpus statistics to carry out the result and statistics of the data with six features to be analyzed: names, adjectives, occupations, verbs, representation, and gender roles.

Mattson (2009) found that the female characters are more physically described by the series, and the female characters tend to be called by shortened names, compared to the male characters. Also, there are more verbs that describing the male works compared to the female works, and from the four

episodes analyzed, the female characters is described as underprivileged compared to the male characters.

The result of the research that was conducted by Mattson (2009) is there is an immense difference between female and male visibility and representation. The women is depicted as the one with better imagery in the researcher's eye according to six features of analysis that she had already conducted, compared to the men in the serial, even though the men speaks more than the women.

Compared with this research, there are several similarities and differences found. Both of the researches are concerned with how *The Simpsons* may depict a gender through the feature used in the series. The differences between the research done by Mattson (2009) and this study is that in Mattson's research, there were several features used by the writer to analyze the gender portrayal: names, adjectives, occupations, verbs, representation, and gender roles, while in this research, the writer is more concerned on how each gender may use the adjective that may shows the characteristics of each character itself. Also, Mattson only use four episodes for the research, while this research take each episode from the first to the twenty-fourth season, allowing the writer to analyze the development of adjective usage from time to time.

The other one is "*The Simpsons, Gender Roles, and Witchcraft: The Witch in Modern Popular Culture*" by Sarah Antinora (2010) that examined the animated series by using a character considered as "the witch" which is a term for a women that is so intelligent and brilliant, and how it reflects the role of women in modern American society, both from its sentences and gestures from the others. Carried

out with the qualitative method and using the sociocultural as its main theory, research conducted by Antinora (2010) shows the sarcasm toward the social gender stereotyping culture, where women seem to be more underestimated than men in the American society.

Antinora (2010) found that the characterization of "the witch" portrayed in the episodes analyzed is somehow a depiction of how females are portrayed to be less able than males, and if the females is showing a determination toward the males, the females will be described as "the witch." Antinora (2010) later concluded that *The Simpsons* shows some issues in the society is still constructing the figure of the females in our society, and how they should behave and act, which is portrayed as a figure of "the witch."

Compared to this research, it may be found that both of the researches done are concerned on how *The Simpsons* depict each gender role. But there also several differences found between the research done by Antinora (2010) and this research. Antinora focused only on one character on one episode, which is the *witch*. Although Antinora deeply analyzed the characterization of the *witch*, it is insufficient to generalize the characterization of gender in *The Simpsons* only based on one character that appear only on one episode, while this research use four main characters on twenty four episodes on *The Simpsons*. Also, this research is also more concerned on how each character uses adjectives, in result of conclusion of the characterization of each gender based on its adjective usage.