

CHAPTER II

LITERATURE REVIEW

2.1. The Concept of Beauty

The concept of beauty becomes an interesting topic to be discussed from time to time, because the concept itself is different from one society to another, and also changes from time to time. Based on the *Oxford Advanced Learner's Dictionary 8th edition* (2010), the word 'beauty' has four senses. They are (1) [uncountable] the quality of being pleasing to the senses or to the mind. (2) [countable] a person or thing that is beautiful. (3) [countable] an excellent example of its type. (4) [countable] a pleasing feature. The use of word 'beauty' in *Oxford Advanced Learner's Dictionary 8th Edition* (2010) ranges from reference to humans and objects to landscape action, and the description of sensual experiences.

As stated previously, the concept of beauty is different from one society to another. In China, the trend was Foot Binding, Chinese women focuses on the small feet. In Thailand, the Kayan tribe focussed on neck-elongation to be beautiful. In Kenya, the Masai women pierce their earlobes and then stretch them out by using materials like slices of elephant tusk. In South Korea, the trend is to have wide, round eyes, and many people are going under the knife to achieve them. In Europe, women focussed on making an hourglass body, so women in that era used a very tight corset to made their body like an hourglass. Thus, the particular set of physical characteristics perceived as beautiful and desirable can vary across cultures.

The concept of beauty is not just different across culture, but different across time. The concept of beauty is changing from time to time, it depends on the trend of the era. For the

example is the label for the most beautiful woman in the magazine, the woman chosen is different in year by year. In the United States, according to the *People's* magazine (www.people.com), in 1990, the one chosen as the most beautiful woman is Michelle Pfeiffer. In that year, she was described as a blonde American talented actress and singer. In 1995, the one chosen as the most beautiful woman is Courteney Bass Cox. She was described as a black hair American actress, producer, and director. In 2000, the one chosen as the most beautiful woman is Julia Robert. She was described as a skinny brunnete American actress who had become one of the highest-paid actresses in the world. In 2006, the one chosen as the most beautiful woman is Angelina Jolie. She was described as a brunnete American actress, film, director, screenwriter, and author. In 2014, the one chosen as the most beautiful woman is Lupita Nyong'o. She was described as a black skin woman, she is an actress, film, and music video director of dual Kenyan and Mexican citizenship. From the list of the most beautiful women according to *People's* Magazine, there is no specific standard in the concept of beauty, the concept is changing across time and it can be an unique things to be analyzed.

From the explanation above, the concept of beauty is relative. The definition itself is different from one culture to another, each culture has their own definition about the concept of beauty and the concept itself is changing across time and genres. Since the definition of beauty is different from one culture to another, the writer just focusses in one country, the United States of America.

The United States of America is a big country which has a big influence in the world. Since the concept of beauty in the United States of America may influenced the world, the writer wants to analyze the concept of beauty in American society. Based on Merriam-Webster (merriamwebster.com), the word 'beauty' has four senses, (1) the quality or aggregate of

qualities in a person or thing that gives pleasure to the senses or pleasurably exalts the mind or spirit, (2) a beautiful person or thing, (3) a particularly graceful, ornamental, or excellent quality, (4) a brilliant, extreme, or egregious example or instance.

As a big country, the United States of America also influences the education, music, films, fashion, etc. The writer wants to focus in analyzing the concept of beauty in American society, because we can find Hollywood films, American musics and also American books and magazines easily. There are a lot of Hollywood films which are shown all over the world, that is why there are a lot of actresses and actors that have influences not only in the United States of America but all over the world. As we can see, Brad Pitt. He is an American actor and film producer, and also get a lot of award nominations in acting and also in film production. Then, we can also see that American musics are really famous to us. As we can see there are a lot of American bands or singers have a concert in Indonesia. For the example is Bruno Mars had a concert at Jakarta on March 24th, 2014. Not only films and musics, the United States of America also popular in books and magazines. One of the famous book is *Forrest Gump*, *Forrest Gump* is a novel which was written by Winston Groom in 1986 and was filmed in 1994. Then, the popular American magazine which has localized is *Cosmopolitan*. *Cosmopolitan* is first published in the United States of America in 1886, then Indonesia is published *Cosmopolitan* in Indonesia language in 1997.

The writer wants to analyze the concept of beauty in American Society from 1911 to 2010, because the concept of beauty in the United States of America is different across time. It happens because of the changing of the era in the United States of America. As we can see in the book of *Outline of U.S History* by Bureau of International Information Programs U.S Department of State 2011, United States of America has some important events in years by

years, that is why there is a changing in the social condition of the society and it impacts to the trend of the era, so it makes the concept of beauty in the United States of America is different across time. In here, the writer tries to divide the analysis into 10 sequences of time, such as: 1911-1920, 1921-1930, 1931-1940, 1941-1950, 1951-1960, 1961-1970, 1971- 1980, 1981-1990, 1991- 2000, 2001-2010.

According to the book of Outline of U.S History, in 1911-1920, the United States of America was in war (World War I) and was prospering in industry. In 1921-1930, the United States of America was found an increasing of the literateurs. In 1931-1940, the new era of the United States of America, the development of agricultural, industry and labor. In 1941-1950, the United States of America entered the World War II. In 1951-1960, the United States was in the civil rights movement. In 1961-1970, the United States launched two-person Gemini spacecraft and Neil Armstrong became the first human to walk on the moon. In 1971-1980, the United States of America was in the women's movement. In 1980-1990, the United States of America entered the increasing of computer year. In 1991-2000, the United States of America was in the Democratic era. In 2001-2010, the United States entered the millenium era, the important event was Barack Obama was eleced as the 44th American president ans he is the first black president of the United States.

The changes of the concept of beauty in American society was influenced by many factors, such as political and social factors in a certain period of time. These changes can be analyzed by using various approaches. Since the writer is interested in linguistics, she uses the Corpus Linguistics approach. In general, Corpus Linguistics analysis can be divided into synchronic and diachronic. A synchronic corpus focusses on one particular period of time, while a dichronic corpus covers a long period of time. Since this study covers the years from 1911 to

2010, the specific approach selected is the diachronic corpus linguistics approach. The description and explanation of diachronic corpus is presented in the following sub chapter.

2.2. Diachronic Corpus

Before explaining what is a diachronic corpus is, it is necessary to explain what a corpus is. According to Bussmann (1996:260), a corpus is “a finite set of concrete linguistics utterances that serves as an empirical basis for linguistic research.” Based on the definition of corpus, it is appropriate to use a corpus in a linguistic research. In the case of analysis using a corpus, the method is called corpus linguistics. McEnery and Hardy (2012:1) defines a corpus linguistics as a study of language data on a large scale, aided analysis of very extensive collections of transcribed utterances or written texts.

Hunston (2002:16) differentiates between a synchronic corpus and a diachronic corpus. A synchronic corpus is a corpus of texts from a particular period of time, while a diachronic corpus is a corpus of texts from different period of time. Furthermore Hunston (2002) mentions that a diachronic corpus is used to trace the development of aspects of a language over time. This is suitable with this study as the time in the corpus will be divided into several period.

One example of a diachronic corpus is The Corpus of Historical American English (COHA). The major sources of COHA are fiction, magazines, newspapers, and non-fiction. COHA allows us to see how words, phrases and grammatical constructions have increased or decreased in frequency, how words have changed meaning over time, and how stylistic changes have taken place in the language. COHA helps us to search more than 400 million words of text of American English from 1810 to 2009.

The major sources of the words included in COHA are as follows:

- Fiction: Project Gutenberg (1810-1930), Making of America (1810-1900), scanned books (1930-1990), movie and play scripts, COCA (1990-2010).
- Magazines: Making of America (1810-1900), scanned and PDF (1900-1990), COCA (1990-2010). In each decade, the magazines are balanced across at least ten magazines (with equivalent sub-genres for the 1900s).
- Newspapers: PDF > TXT of at least five newspapers (1850-1980), COCA etc (1990-2010).
- Non-fiction: Project Gutenberg (1810-1900), www.archive.org (1810-1900), scanned books (1900-1990), COCA (1990-2010). In each decade, the non-fiction is balanced across the Library of Congress classification system.

Since the corpus in COHA ends in 2009 it is necessary to combine it with another corpus. The other corpus that becomes the source of data is The Corpus of Contemporary American English (COCA). COCA is the largest freely-available corpus of English, and the only large and balanced corpus of American English. The corpus contains more than 410 million words of texts and is equally divided among spoken, fiction, popular magazines, newspapers, talk shows, and academic texts. There are about 20 million words of texts are being added each year to enrich the corpus data. It also provides a balance of availability, size, genres, and currency (how recent it is) that is not found in other corpora, such as the American National Corpus (ANC), the British National Corpus (BNC), and the Bank of English (BOE).

According to Davies (2008), the texts in COCA come from a variety of sources, they are:

- Spoken: (85 million words [84,818,571]) Transcripts of unscripted conversation from more than 150 different TV and radio programs (examples: *All Things Considered* (NPR), *Newshour* (PBS), *Good Morning America* (ABC), *Today Show* (NBC), *60 Minutes* (CBS), *Hannity and Colmes* (Fox), *Jerry Springer*, etc).

- Fiction: (80 million words [80,251,038]) Short stories and plays from literary magazines, children's magazines, popular magazines, first chapters of first edition books 1990-present, and movie scripts.
- Popular Magazines: (86 million words [86,000,281]) Nearly 100 different magazines, with a good mix (overall, and by year) between specific domains (news, health, home and gardening, women, financial, religion, sports, etc). A few examples are *Time*, *Men's Health*, *Good Housekeeping*, *Cosmopolitan*, *Fortune*, *Christian Century*, *Sports Illustrated*, etc.
- Newspapers: (82 million words [81,861,570]) Ten newspapers from across the US, including: *USA Today*, *New York Times*, *Atlanta Journal Constitution*, *San Francisco Chronicle*, etc. In most cases, there is a good mix between different sections of the newspaper, such as local news, opinion, sports, financial, etc.
- Academic Journals: (82 million words [81,840,348]) Nearly 100 different peer-reviewed journals. These were selected to cover the entire range of the Library of Congress classification system (e.g. a certain percentage from B (philosophy, psychology, religion), D (world history), K (education), T (technology), etc.), both overall and by number of words per year

The number of words in COHA and COCA is too big to be analyzed in each entirety. Therefore, the writer needs to take one particular word class. Since the words to be focus on is the adjectives 'beautiful', the collocate in the English language is a noun. Consequently, the word class to be analyzed is noun. The discussion on collocations and nouns are presented in the following sub chapter.

2.3. Collocations and Nouns in The English Language

In corpus linguistics the most common analysis is on the collocation of a word. According to Mcenery and Hardie (2012: 123), the term of collocation has a meaning co-occurrence patterns observed in corpus data. There are two types of collocation in a Lindquist's recent book, collocations in a window and adjacent collocations. Collocation in a window is the words collocates which occur in the left or in the right of the keyword, and the term 'window' stands for the space to the left and the right of the keyword (Linqvist, 2009:73). Adjacent collocations are not only looked at in the collocations occur on the right or on the left of the keyword (or sometimes with an empty slot in between), adjacent collocations are closer to real linguistic structure. That is why the writer uses adjacent collocation to complete the research.

In order to understand a collocation of a word, we need to determine the word classes. In the English language, word classes can be classified into content words and function words. From the book of *An Introduction To Language* seventh edition, content words are called as the open class words and these words denote concepts such as objects, actions, attributes, and ideas that we can think about, so that we can and regularly do add new words to these classes. Function words are called closed class words and these words do not have clear lexical meaning or obvious concepts associated with them, so that they have a grammatical function. To make a meaningful analysis on words, researchers in corpus linguistics usually use the content words.

Content words consist of nouns, verbs, adjectives, and adverbs. Nouns are the syntactic category, also lexical category, of words that can function as the head of a noun phrase, such as book, Jean, sincerity. In many languages nouns have grammatical alternation for number, case, and gender and occur with determiners. Verbs are the syntactic category, also lexical category, of words that can be the head of a verb phrase. Verbs denote actions, sensations, and states (e.g., climb, hear, understand). Adjectives are the syntactic category, also lexical category, of words

that function as the head of an adjectives phrase, and that have the semantic effect of qualifying or describing the referents of nouns (e.g., tall, bright, intelligent). Adverbs are the syntactic category, also lexical category, of words that qualify the verb such as manner adverbs like quickly and time adverbs like soon. The position of the adverb in the sentence depends on its semantic type (e.g., John will soon eat lunch, John eats lunch quickly).

Since this study analysis an adjective that is the word 'beautiful', the collocates is a noun. Therefore, it is necessary to explain further about nouns. There are several ways to classify nouns in the English language. One of the ways is by using the WordNet. WordNet is a large lexical database of English which is developed by Princeton University, it can be accessed freely at <http://wordnet.princeton.edu>. According to Miller (1995:158), WordNet is like a standard dictionary, it is containing the definitions of words. Wordnet is not organized alphabetically, but conceptually because of the basic unit in WordNet is a synonym set (synset) which represents a lexicalized concept. In WordNet, the organization of nouns, verbs, adjectives, and adverbs are each organized differently, all are organized in synsets, the semantic relations among the synsets differ depending on the grammatical category. The structure of WordNet is a useful tool for computational linguistics and natural language processing.

The main relation among words in WordNet is synonymy. According to the wordnet.princeton.edu, each of WordNet's 117,000 synsets is linked to other synsets by means of a small number of "conceptual relations". The most frequently encoded relation among synsets is the super-subordinate relation (also called hyperonymy or hyponymy). According to Brown (2006:666), concepts which are expressed by nouns are interconnected by the hyponymy relation (or hyperonymy) which connected to the more general concepts. For the example, the synset (mailbox, letterbox) is a hyponym, or subordinate, of (box), which in turn is a hyponym of

(container). (Mailbox, letter box) is a hypernym, or subordinate, of (pillar box), which denoted a specific type of mailbox. Hyponymy builds hierarchical ‘trees’ with increasingly specific ‘leaf’ concepts growing from an abstract ‘root’.

2.4. Related Studies

There are found studies which related to the writer’s study. First related study is about how politic build beauty construction in a social life which is conducted by Dian Paramita Sastrowardoyo (2007), entitled “Komplek Industri Kecantikan: Sebuah Kritik Sosio Fiosofis”. The second related study is about diachronic corpus which is conducted by Kirana Wuri Anjar Arum (2012), entitled ‘A Diachronic Corpus Based Analysis of The Adjectival Collocates of [man] and [woman] in The American English’. Then, the third related study is about the synchronic corpus by comparing three languages, Russian, English, and Spain which is conducted by Anna Gladkova and Jesus Romero-Trillo (2013), entitled “Ain’t it beautiful? The conceptualization of beauty from an ethnopragmatic perspective”.

First, the studies of Sastrowardoyo (2007), entitled ‘Komplek Industri Kecantikan: Sebuah Kritik Sosio Fiosofis. This study shows that beauty is a part of social construction which is built by some political needs and played by industrial system, mass media, and patriarcy mass. According to Sastrowardoyo (2007), this study will become an academic comments to beauty complex itself which used sociology and philosophy. The result of this study shows that the concept of the beauty which is played in Beauty Industrial Complex is a beauty concept which is constructed by contemporer ideology and it works in a part of contemporer mass, that is why she try to make a deep research about the new concept of beauty according to the postmodern theory.

There are similarities and differences of the study by Sastrowardoyo (2007) with the writer's study. The similarity is both analyze the concept of beauty, and with a different point of view, Sastrowardoyo (2007) analyzes the concept of beauty in sociology and philosophy perspectives, while the writer's study analyzes the concept of beauty in linguistics perspectives. The other difference is in the time coverage, Sastrowardoyo's (2007) study is synchronic, while the writer's study is diachronic.

Second, the studies of Kirana Wuri Anjar Arum (2012), entitled 'A Diachronic Corpus Based Analysis of The Adjectival Collocates of [man] and [woman] in The American English'. This study uses COHA (The Corpus of Historical American English) and COCA (The Corpus of Contemporary American English) as the source to analyzed the adjectival collocates of [man] and [woman] in the American English from the Civil War period (1861) to Obama's presidency (2010). The objective of the study is to compare the changes in the adjectival collocates of [man] and [woman] from 1861 to 2010. In the analysis, the corpus evidence is also related to the facts written in American history references, to show how far they conform to one another.

The results of Arum's (2012) study shows that there are similarities and differences in the adjectival collocates of [man] in American English over time. For the example, the phrase 'white man' has been quite frequent since 1861, while the phrase 'black man' has just started to gain high frequency in 1941 and has even higher in frequency than 'white man' in 2001 and 2010. The similar case goes to the adjectival collocates of [woman]. The phrase 'American woman' can be found in all periods of time, while 'French woman' is only high frequency during the World War I and 'Muslim woman' has just started to gain high frequency during Obama's presidency. There also found the differences within the same period of time and across time, for

the example, the phrase 'fat man' occurred in three period of time, while the phrase 'fat woman' only occurred once.

The similarities between Arum's (2012) study and the writer's study are both of them analyze diachronic corpus and use COHA (The Corpus of Historical of American English) and COCA (The Corpus of Contemporary of American English) as the sources. The difference is on the object of the study. Arum's (2012) study focusses on two lexical items, they are [man] and [woman]. Whereas, the writer's study is on a particular concept that can be related to various lexical items.

The third related study is an article from Journal of Pragmatics written by Anna Gladkova and Jesus Romero-Trillo (2013), entitled 'Ain't It Beautiful? The Conceptualization of Beauty from an Ethnoprismatic Perspective'. This study aims to answer the question of the ethnoprismatic conceptualization of beautiful in three languages, English, Russian, and Spanish. The sources of the data are from three online corpora: Russian National Corpus (Russian), Cobuild Wordbanks Online (English) and Corpus de Referencia del Espanol Actual (Spanish). This article elaborates, describes and reveals similarities in their perceptions of 'beautiful' across the three languages in relation to the nouns modified by this concept. Gladkova and Romero-Trillo (2013) investigate the presence of the perception universals identified by NSM: SEE, HEAR, and FEEL, which in our data are central to the analysis of the aesthetics vocabulary along with the primitives GOOD, SOMEONE, SOMETHING and THINK. Both Gladkova and Trillo's (2013) study and the writer's study use corpus linguistics approach. Gladkova and Trillo focusses on a particular time, but covers three different languages. On the other hand, the writer's study is on one particular language, but across different time frames.