A SEMANTIC PROSODY ANALYSIS OF SWEAR WORDS IN A CORPUS OF ENGLISH SONGS

HAZRI SHAHREEN HASHIM

FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR

2020

A SEMANTIC PROSODY ANALYSIS OF SWEAR WORDS IN A CORPUS OF ENGLISH SONGS

HAZRI SHAHREEN HASHIM

DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER IN LINGUISTICS

FACULTY OF LANGUAGES AND LINGUISTICS UNIVERSITY OF MALAYA KUALA LUMPUR

2020

UNIVERSITY OF MALAYA
ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **Hazri Shahreen Hashim**

Matric No: **TGC150011**

Name of Degree: **Master of Linguistics**

Title of Dissertation: **A Semantic Prosody Analysis of Swear Words in a Corpus of English Songs**

Field of Study: **Corpus Linguistics**

 I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature   Date:

Subscribed and solemnly declared before,

Witness's Signature   Date:

Name:

Designation:

A SEMANTIC PROSODY ANALYSIS OF SWEAR WORDS IN A CORPUS OF

ENGLISH SONGS

ABSTRACT

This study looks at the semantic prosody of swear words found in a corpus of English songs by looking at their collocations using a corpus software; AntConc. A total of 545 songs were chosen based on the Billboard year-end chart from 2011 to 2016 with 243,689 number of tokens and 8,139 number of word types. Word lists and concordance lines were generated from the corpus for the data analysis. The analysis on the concordance lines showed that not all swear words with negative-based meaning possessed negative semantic prosody. Despite possessing negative semantic prosody, 10 out of 15 swear words possessed neutral semantic prosody and only one word had a positive semantic prosody. In summary, this study argues that swear words are not entirely negative as perceived, rather it depends on context to be regarded as negative.

Keywords: Swear words, Semantic Prosody, Collocation, Corpus, English songs

ANALISIS PROSODI SEMANTIK PERKATAAN KESAT DI DALAM KORPUS LAGU-LAGU INGGERIS

ABSTRAK

Kajian ini melihat prosodi semantik kata-kata kesat yang terdapat di dalam korpus lagu-lagu Inggeris melalui kolikasi dengan menggunakan sebuah perisian korpus; AntConc. Sebanyak 545 lagu yang dipilih berdasarkan carta lagu hujung tahun Billboard daripada tahun 2011 sehingga 2016 dengan jumlah token sebanyak 243,689 dan sebanyak 8,139 jumlah jenis perkataan. Senarai perkataan dan garis konkordans perkataan dari korpus tersebut telah dijana untuk dianalisis. Analisa garis konkordans mendapati tidak semua perkataan kesat yang maknanya berasaskan negatif mempunyai prosodi semantik negatif. Walaupun mempunyai prosodi semantik negatif, 10 daripada 15 perkataan kesat mempunyai prosodi semantik neutral dan hanya satu perkataan kesat mempunyai prosodi semantik positif. Secara ringkasnya, kajian ini menegaskan bahawa tidak semua perkataan kesat adalah negatif kerana penggunaan kata-kata kesat ini bergantung kepada konteks untuk dianggap sebagai negatif.

Kata Kunci: Perkataan Kesat, Semantik Prosodi, Kolokasi, Korpus, Lagu-lagu Bahasa Inggeris

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of Symbols and Abbreviations

BNC : British National Corpus

CoES : Corpus of English Songs

HFNS : Holy, Fucking, Shit, and Nigger Principle

UK : United Kingdom

US : United States

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1   Introduction

In this chapter, the background of the research will be provided (Section 1.2). This is followed by the research problem (Section 1.3), research aims (Section 1.4), research questions (Section 1.5) and significance of the research (Section 1.6). This chapter concludes with the limitation of the study (Section 1.7) and the organisation of the thesis (Section 1.8).

1.2   Background to the Research

The term 'swearing' is usually referred to the use of swear words. As stated by Oxford Dictionary 9th edition (2008), swearing refers to being rude or the use of offensive language. Horan (2015) adds that swearing refers to the use of offensive or taboo language for expressive purposes. Jay and Janschewitz (2008) define swearing as the implementation of taboo language aimed to convey a speaker's emotional judgement and to convey that message to the listeners. Similarly, Bergen (2016) states that swear words are used to express frustration, anger, or emphasis because these words elicit physiological reaction. Crystal (1997) as cited by Lutzky and Kehoe (2016) also delineates that swearing is a form of spoken language that portrays a release of frustration and suppression of emotion. Generally, swear words are used to express pent-up emotions or negative feelings such as anger. Depending on the context of use, it can also be used to convey positive feelings such as joy and excitement (Lutzky & Kehoe, 2016, and Byrne, 2018) for instance, swearing helps users to withstand pain, diffuse stress, and bond with colleagues (Baudin & Paramasivam, 2014, and Byrne, 2018). In the case of PewDiePie, a famous YouTuber, swear words are used to simulate casual conversations between friends, and swearing reduces social distance to create some sort of intimacy (Beers, 2017).

The use of swear words is not only taboo or considered rude in daily conversations, but it is also prohibited in literature. In the late 17th and 18th centuries, a moral reform movement was developed in England where the focus was on the use of bad language (see McEnery 2006a, and McEnery 2006b). Established by Edward Stephens on the Strand, London (McEnery, 2006b), the group aimed to alter the morals of Britain by actively seeking to pursue any person that the society thought was guilty of immorality. During its lifetime for 47 years from 1691 to 1738, the Society for the Reformation of Manners brought 101,683 prosecutions where it was estimated that 14,192 were prosecuted for swearing and cursing (McEnery 2006a, and McEnery 2006b). Other than that, the society expressed its disapproval of bad language by disseminating literature to encourage moral reform. In addition, through an associated organisation named Society for the Promotion of Christian Knowledge, the society provided free moral education for the children of the poor (McEnery, 2006b). The means that the society used to convey such views about bad language through prosecution, publication, and education.

Ljung (2011) divides swear words into two categories; (i) based on the functions and (ii) the themes. The first category is defined into two types namely (i) the stand-alone and (ii) the slot fillers. The first type is a condition when swear words can stand alone on their own, neither depending on any words following nor preceding it. He explains that this function includes expletive interjection for example *fuck!* and *shit!* to show disappointments, oaths to express swear words by referring to God such as *God!*, *Christ!* and *hell!*, curses to refer to swearing in general to express punishment on someone such as *fuck you!*, affirmation and contradiction that are the addition of swear words to express something such as *the hell it is!* (*it* refers to things like a door which cannot be opened), unfriendly suggestions to express aggression on someone such as *go to hell!*, ritual insults such as *your mother …* that is usually used for sexual exploitation towards someone's mother or sister, and name-calling such as *idiot!*, *moron!*, and *nigger!*.

The second type of swear words is the slot fillers that function when swear words have additional sentences or expressions. The expressions include adverbial/adjectival intensifiers such as *bloody marvellous* and *he is driving damn fast*, adjectives of dislike such as *he is a bloody fool*, emphasis such as *absobloodylutely!* and *infuckingcredible!*, modal adverbials such as *I bloody well drank my beer* and *They fucking bought me one drink between them*, anaphoric use of epithets such as *asshole/arsehole*, *bastard*, and *bugger*, and noun supports such as in *John is a boring son of a bitch* and *Philips is a hard-working son of a bitch*.

The second category of swear words by Ljung (2011) is based on its theme that is further divided into five types. The first type is the religious/supernatural theme in which the swear words are related to religion such as *God's blood*. The second type is the scatological theme that is related to unpleasantness and taboo words such as *shit*, *asshole*, and *poo*. The third type is the sex organ theme that is related to sex organs such as *balls* and *bollocks*. The fourth type is the sexual activities theme that uses swear words related to sexual intercourse for instance the swear word *fuck* in *I don't give a fuck*, *for fuck's sake*, and *she knows fuck all about it*. The last theme is the mother (family) theme in which the swear words refer to family members such as *son of a bitch* or *motherfucker*.

According to Bergen (2016), there are four categories of swear words namely *holy*, *fucking, shit,* and *nigger*. In his book, Bergen (2016) proposed the categories of swear words to be called as the Holy, Fucking, Shit, and Nigger Principle (henceforth, HFSN). The first category of English profanity comes from words or acts that desecrate the Holy. The word *profanes* in Latin literally means 'outside the temple' and refers to *holy*, *hell*, and *goddamn*. Blasphemy – a sin against religious doctrine – occurs for some people when religious words are used in secular ways for example the names of religious figures: *Jesus Christ*, *Jehovah*, or *Mohammad* (Bergen, 2016).

The second category of swearing relates to sex and sexual acts (Bergen, 2016). It includes the acts themselves; *fuck*, sex organs such as *pussy* and *cock*, the doer of those acts for instance *cocksucker* and *motherfucker*, and artefacts related to those acts such as *spooge*; semi-liquid gunk, and *dildo*.

The third category involves other bodily functions – products excreted from the body (Bergen, 2016). This includes robust cohorts of words delineating *faeces*, *urine*, and *vomit*. A word like *douchebag* is also included in this category since it is associated with body parts.

Lastly, there are the slurs (Bergen, 2016). The words under this category are a derogatory reference to people based on how a group of people perceive the world, for instance, how the superiors perceive the inferiors, and derogatory words that refer to them such as *nigger*, *faggot, homo*, and *retard*. Apart from that, the words are also perceived in terms of sex, sexual orientation, ethnicity, religion, and so on. Bergen (2016) reiterates that there is something about the swear words that is conceived as bad.

Up until this point, there are a number of terms to refer to the words under HFSN. Apart from the mentioned terms, Thelwall (2008) lists other similar terms including *obscenity*, *foul language*, *coarse language* or *vulgar*, *Billingsgate*, *naughty*, *expletive*, *bad language*, *rude words* and *oath*. To cater for the abundant terms that can be used interchangeably to refer to swear words, this study will use the term *swear words*.

Based on the researcher observation on English songs being broadcasted in local Malaysian radio upon conducting this study, swear words can be found in songs. These words however, are censored when one compares the songs broadcasted with the lyrics found online. Although the observation was purely based on personal experience, the usage of swear words in songs is however, against the guideline by Malaysian

Communications and Multimedia Commission (2018). According to the guideline, any words that have the sense of vulgarity, offensive or swearing should be censored for instance in the case of the word *hell* in a song by Taylor Swift entitled Wildest Dream. The swear word was censored throughout the song in a radio broadcast and was replaced with a bleep. Another example of a song containing swear words is Anaconda by Nicky Minaj where the word *fuck* was censored. On top of that, a verse of that song was removed from public because of the profanities used. This was observed when the line that contains the word *hell* and *fuck* was removed from the songs broadcasted on radio and it was noticeable because removing the words disrupts the melody of the song. However, the song flows without any disruption on the melody on YouTube.

In an unpublished article by Hall (2009) of Lehigh University, censorship has been enforced to enforce morality. The censorship essentially began when the Radio Act of 1927 was introduced in which it prohibited the use of obscene, indecent or profane language on the radio. A blog writer, Martin, in his entry on May 2012 (https://blogs.lt.vt.edu/claireyw/2012/05/02/why-is-music-censored/) reiterated that the implementation of censorship on songs was based on a claim that the music was clashing with moral values, racial motivations, generational value gaps, and fear. However that may be, the public could still access the missing verse and words with any search engines available such as Google, Firefox, or Yahoo. The initiative to censor the swear words is one of the ways to inculcate good moral, but removing swear words from lyrics does not remove them from our language. This is because the swear words are important in which people use the words to express the strongest human emotions such as anger, fear, and passion (Bergen 2016, and Lutzky and Kehoe 2016).

In language learning, songs can be utilised as one of the tools to teach English language. Some songs can be used in language teaching because of the clear utterance

and simple sentence structures where learners can learn how to pronounce words and use simple vocabulary. Like any other literature, songs can convey love and emotions for example "All of Me" by John Legend released in 2013 and "A Thousand Years" by Christina Perri released in 2011 (https://www.billboard.com/articles/list/6792625/top-50-love-songs-of-all-time). Some songs are also political in a sense that they criticise countries for example "This is America" by Childish Gambino released in 2018 and "Fight the Power" by Public Enemy released in 1989 that criticize government treatments towards black community (https://www.billboard.com/articles/business/chart-beat/9396644/protest-songs). Taking the song "This is America" as an example, a professor of music history at the University of Pennsylvania, Guthrie Ramsey, explains the metaphors behind the song in an online TIME newsletter reported by Mahita Gajanan (2018). In the music video, there are four key moments (i) the first gun shot, (ii) Gambino dancing with schoolchildren amid violence, (iii) the gunned down choir, and (iv) Gambino running away in the closing moments. Each of the moment signifies certain events that happened to the black community in America.

The video starts with a man strumming a guitar matching to a group of choir. After a brief moment, Gambino shoots the man who has been tied up with a head covered. The gun used to shoot the man is then handed over to another man who wraps the gun in red cloth as the dead body is dragged away. The moment goes right into the first line of rapped chorus that is "This is America/Don't catch you slippin' up". This portrays the contradictions of trying to get money, the idea of being black in America as explained by Ramsey. The next moment is when Gambino and a group of schoolchildren in school uniform dance throughout, smiling through with faultless moves as violence erupts behind them. In the same newsletter reported by Mahita Gajanan (2018), the mentioned act according to Ramsey could be interpreted in numerous ways, for example the dancers could be there just to distract the viewers similar to the way how black art is used to keep

the people occupied from the real problems plaguing America. However, he also adds that it is best to look at the video as a whole because the country itself is of "very strange juxtapositions". The next moment depicts a choir singing enthusiastically in a happy tone before Gambino shoots all of them. Ramsey explains that the incident is a reference to the 2015 Charleston shooting in which Dylann Roof, a white supremacist, killed nine black people in a church basement. The final moment of the video shows Gambino running, terrified, down a long dark hallway away from a group of people as a singer, Young Thug, sings "You just a Black man in this world/You just a barcode, ayy". The sprint goes back to a long tradition of black Americans having to run just to save their lives from slavery. According to Ramsey, this is a reference to one song dating back to slavery in the 19th century entitled 'Run N – Run" which refered to a black person running for his or her life.

1.3   Research Problem

Although there have been numerous studies conducted on swear words (see Bergen 2016, McEnery 2006a, McEnery and Xiao 2004, Van Lancker and Cummings 1999, Hughes 1991), there is still a general lack of understanding on the functions of swear words. Generally, studies conducted on swear words only focus on which gender swears more and which gender uses what swear words (see Van Lancker & Cummings 1999, McEnery and Xiao 2004, McEnery 2006a, Jay & Janschewitz 2008, Thelwall 2008, Ljung 2011, Butler & Fitzgerald 2011, Lutzky & Kehoe 2016, and Bergen 2016). Previous studies have shown that males swear more than females (McEnery, 2006a) and a more recent study has shown that female swears more than males in certain context (Thelwall, 2008).

There are also studies conducted on how swear words are perceived. This kind of study usually involves the rating of the swear words based on a Likert scale (see Bergen, 2016 and Goddard, 2015). Bergen (2016) conducted several studies on swear words on participants from New Zealand, Britain, and America on how these participants viewed and rated a list of given swear words. The problem about rating the swear words was that there was no consistency on how it was done and it was mostly based on assumptions. Looking at the data presented by Bergen (2016), there is an inconsistency on the results. Taking the word *fuck* from the data presented by Bergen (2016), 80% of the participants in Britain rated *fuck* as very severe but the data still showed that some participants rated *fuck* as fairly severe with as much as 20%. In America, the participants rated the swear word as very offensive with the mean offensive score of 6.5 out of 7 where max score 7 is most offensive. On the other hand, the swear word *fuck* scored 4 out of 7 by participants in America. In most studies, swear words are often ranked (see Bergen 2016, McEnery 2006a, Hughes 1991) based on perception in which it is unreliable because people can be biased when judging the swear words individually. Beers (2007) agrees that offensiveness ratings of isolated swear words are unreliable because how the participant interprets the words individually is still unknown.

Instead of rating the swear words like the previous researchers, this study looks at the collocations to determine the semantic prosody as done in research studies by Yusuf (2010) and Rahim (2012). Hence, highlighting the notion by Firth (1957), one must examine the collocations of a word instead of the word in isolation in order to understand it. This study focuses on the semantic prosody of swear words found in the chosen English songs using a corpus linguistics approach. Bergen (2016) supports that there is something about the swear words themselves that we would conceive as bad. For instance, *shit* falls under the category of bodily functions whereby words like *faeces*, *urine*, and *shit* are conceived as swear words. However, even if excretion is culturally taboo, it does not

mean that any available words describing it will be as well, for instance, *shit* is more profane compared to *poop* (Bergen, 2016).

This study is significant since it offers a new perspective on investigating swear words and not merely rating how people perceive it to be (Goddard, 2015 and Beers, 2012). Based on the mentioned studies, swear words are often being judged in isolation and because there is no actual method known on how the participants rate the swear words (Beers, 2007), the results on the swear words which are oftentimes regarded as showing negative prosody are unreliable. Therefore, this study attempts to provide a more empirical examination of swear words and their meanings according to their collocations with other words in context (see Bergen 2016, McEnery 2006a, Hughes 1991) by incorporating the notion of semantic prosody by Firth (1957).

1.4 Research Aims

The aim of this study is to identify the frequent swear words found in a corpus of English songs (henceforth, CoES) compiled from Billboards from the year 2011 to 2016 (https://www.billboard.com/charts/hot-100). In addition, the present study examines identified swear words in more detail by looking at their notion of semantic prosody.

1.5 Research Questions

Specifically, this study seeks to find answers to the following research questions:

1. What are the swear words found in a corpus of English songs?

2. What are the top 10 swear words found in the sub-corpus of English songs in from 2011 to 2016?

3. What is the semantic prosody of swear words found in a corpus of English songs?

1.6   Significance of the Research

Apart from shedding light on the notion of swear words, this study is to elucidate Stewart's (2010) criticism on the idea of semantic prosody as proposed by Firth (1975). One of the characteristics of semantic prosody is the *hidden* criteria as described in the body of literature. Stewart (2010) lists the words that are often used to describe semantic prosody and the words carry the sense of hidden, subliminal, and 'unconscious' nature. Partington (2004) describes semantic prosody as evaluative meaning and the meaning is concealed to the naked eyes for example in the case of *happen* or *sheer*. In addition, Louw (1993) claims that semantic prosody can disclose the speaker's attitudes although the speaker attempts to conceal them. This suggests that semantic prosody can function subliminally without the speaker's conscience. However, Stewart (2011) views that there is not enough evidence to say clearly that semantic prosody has a *hidden* feature. Therefore, this study is hoped to shed light on semantic prosody focusing on the *hidden* feature of semantic prosody.

1.7   Limitation

The scope of the study is limited to the number of swear words found within the collection of English songs in the billboard from 2011 to 2016. It is not possible to analyse all the swear words as contended by Biber, Conrad and Reppen (1998), as the representation of the language provided in the corpus itself is only limited to the design of the corpus (Biber, Conrad, & Reppen, 1998). Though this study can only examine swear words found within a corpus of English songs, different corpus would yield different results.

1.8 Organisation of the Thesis

This chapter is organised into 5 chapters. Chapter 1 provides the introduction to the thesis, while Chapter 2 presents a review of related literature. Chapter 3 highlights the research methodology of the thesis, followed by Chapter 4 which presents the analysis of the study. Finally, Chapter 5 presents the conclusion of the study.

CHAPTER 2: LITERATURE REVIEW

2.1   Introduction

In this chapter, relevant literature is presented in relation to the study. This chapter begins with studies conducted on English songs (Section 2.2). The next section focuses on swear words (Section 2.3). This section is followed by a review on corpus linguistics (Section 2.4) with two sub-topics – collocation (Section 2.4.1) and semantic prosody (Section 2.4.2). The chapter finally concludes with a summary of the chapter (Section 2.5)

2.2   English Songs

The search of linguistic studies/analyses on songs is rather difficult because most studies on songs focus on English language teaching; usually the implementation of English songs in English classroom. Notwithstanding, studies on second language acquisition have found that the implementation of songs can reduce the affective filter of a leaner (Schoepp 2001, Lin 2008, Boothe & West 2015). The affective filter is a wall that reduces learners' capability in absorbing knowledge. Hence, the songs used in the English language classroom are found to reduce the affective filter that help increase the capability of a learner to absorb more knowledge in a particular lesson. Likewise, Krashen (1981) states that affective learning will occur when the affective filter is low.

Other than lowering learners' affective filter, songs are also used to enhance learners' vocabulary. The natural characteristics of songs enable learners to learn how native English speakers use certain words. Learning vocabulary in the first place is an essential part of the language learning process. Rodhe and Tiefenthal (2000) outline that in acquiring new vocabulary, a three-step procedure is involved namely (i) isolating the word from the input, (ii) creating potential meaning, and (iii) mapping meaning on to

form. In the case of young language learners, Coyle and Gomez Grazia (2014) state that young learners can acquire a word even by hearing it once. Since song lyrics are repetitive in nature, these young learners can acquire vocabulary more easily.

In spite of the benefits of using songs in enhancing learners' vocabulary, a group of researchers have tried to prove the effectiveness of using songs in the English language class. Setia, Rahim, Nair, Husin, Sabapathy, Mohamad, and Kassim (2012) conducted a study on the effectiveness of songs in the English language classroom. They found that songs not only enhanced the learners' vocabulary, but they helped to improve the learners' language proficiency. However, the researchers employed a survey method and this could be merely descriptive and the study would be better if pre-test and post-test were conducted on the learners. The pre-test should be used to assess the participants before they received a treatment – to know their proficiency level. The post-test assessed the participants after the treatment to identify any significant changes. Then it can be clearly seen that songs did help in enhancing learners' vocabulary. Nonetheless, this study put forward that using songs in English language classroom could help learners in acquiring new vocabulary.

2.2.1   English Songs in Linguistic Studies

Research on the use of songs to study language usually employs corpus techniques that involve a digital collection of songs that are later analysed through the use of a computer software. Songs are packed with ideologies that the language possesses. In a study by Katznelson, Gelman, Lindblom, and Caput, (2010), other than looking at the diachronic aspect of semantic change in words, they use a corpus of songs that covers a twenty-year span i.e. from 1989 to 2009 comprising top 10 hits from every two years of different genres of music; 433 songs and 178,982 words to inspect the ideologies the songs carry.

Katznelson *et al* (2010) generated a word list and a frequency list to analyse the data. The annotated corpus then was divided into semantic/grammatical categories to generate data. Before the analysis began, Katznelson *et al* (2010) hypothesized that these genres would exhibit certain qualities that reflected perceptions of the music. Based on the analysis, it was found that each genre had a specific concept (Katznelson *et al*, 2010). Rock songs reflected physicality imbued with the concept of pain. In pop songs, it was deduced that the genre portrayed an event with a loud sound or rhythm of the dance floor because the top noun generated was only the word "boom". Country genre on the other hand portrayed top temporal nouns that it was deduced that this genre was mainly descriptive in nature. Lastly, Hip Hop genre reflected possessions and meta-narrative since the top nouns were *ass*, *music*, *bitch*, and *money*. Hence, the list of words generated showed that top 10 nouns found in a corpus of songs could reveal certain concepts or contexts (Katznelson *et al*, 2010, Falk, 2012) that could assist in the identification of semantic prosody of the swear words. This is similar to a study conducted by Rahim (2012) in her study of semantic prosody where she defined a context in her corpus made up of articles in newspapers by identifying the top 10 nouns.

Falk (2012) on the other hand, conducted a study on the linguistic features of rock song lyrics from 1950 to 1999 aiming to investigate the trends of the years. The corpus for the study comprised of the top 60 songs for each decade where roughly 53,000 words were investigated synchronically and diachronically. Findings showed that rock lyrics depicted a high level of involvement – interaction between the singer and listener. This is because the linguistic features of rock lyrics tended to be more spoken-like, than written-like (Falk, 2012).

Kuhn (1999) conducted a study that applied speech act theory to the analysis of seductive strategies in blues lyrics. It was found that songwriters added the same type of

language requests to songs that would be found in real-life situation. Focusing on the language request, the lyrical request for making love for instance contained internal and external modifiers such as flattery, promises, and reasons for submitting to requests (Kuhn, 1999). A study by Kuhn (1999) is one of the studies that looks at song lyrics and found that it possesses similar manner of communication such as political speeches and ordinary conversations (Pettijohn & Sacco, 2009).

Another linguistic analysis on songs was carried by Cole (1971) that surveyed the top 10 Billboard songs for each year in the 1960s. He conducted a content analysis of the song lyrics for themes such as mood, love-sex, religion, violence, and social protest. Since the study was conducted solely on lyrics, the popularity of the songs was not highlighted. However, the study revealed that love-sex was identified as a predominant theme during the 1960s.

Numerous studies have been conducted by using songs as the source of data for example as discussed in the field of education where songs can be used to aid language learners and enhance their vocabularies. Songs are also useful in lowering the active filter to enable language learners to absorb more than their normal learning capacity. Although the mentioned studies used songs as the source of data in the field of linguistics, studies on swear words in songs could not be found. Nevertheless, the next section will be covering the studies on swear words in various fields.

2.3   Swear Words

The studies on swear words according to Bryne (2017) were put on hiatus for almost a century long since early 1880s to 1980s. Although there were studies conducted on brain-damaged patients in the United States (henceforth, US), the United Kingdom (henceforth, UK), France, Germany, and Italy, most swearing among the aphasics patients

were not included in the data (Bryne, 2017). This could be because swearing was not counted as "real" speech in that period, instead swear words were considered to be incomprehensible sounds, more like an animal's howl than human expression (Bryne, 2017).

Nowadays, some researchers would still avoid looking at swear words because the negative impression against these words still exist. For instance, Hitchcock and Cairns (1973) in their study on amygdalae where two Scottish doctors had a conversation with a thirty-four-year-old man that was undergoing a surgery. The conversation that contained swear words was snot included in the transcription although the study's highlight was not on swear words. Even though the swear words were not included, readers would know the missing swear words by reading the transcription.

A number of studies for instance by Van Lancker and Cummings (1999), McEnery and Xiao (2004), McEnery (2006a), Jay and Janschewitz (2008), Thelwall (2008), Ljung (2011), Butler and Fitzgerald (2011), Lutzky and Kehoe (2016), and Bergen (2016) focused on swear words. In most of the studies of swear words, it is often speculated that males swear more than females (McEnery, 2006a). A study on swear words by McEnery (2006a) looks into how the swear words are used and the associations between the swear words with gender, social class, and age. By using the spoken section of the British National Corpus (henceforth, BNC), he takes a socio-historical approach to discourses about swear words in English by looking at a wide range of data. The study explores how the general public's perception of swear words has changed over the past 400 years in three periods; 1586 to 1690, 1690 to 1745, and 1960 to 1980. Based on the study, the attitudes towards swear words were established by the moral reform movements of the late 17th and early 18th centuries to form a discourse of power for the growing middle classes in Britain. Focusing on swearing and gender, McEnery (2006a) stated that the act

of swearing was proven in the 1970s research where males used more strong swear words compared to females. Gauthier, Guille, Rico, and Deseille (2015) and Coates (2004) also expressed that males swore more than females. However, a more recent study proved that the strong swear words users were no longer dominated by males (Thelwall, 2008).

The study conducted by Thelwall (2008) focused on swearing in MySpace page by users in UK and US. The home pages were checked for the occurrence of swear words that were known, commonly strong, and moderate swear words for the comparison with the occurrence with gender, age, and nationality. For the analysis, Thelwall (2008) constructed a list of swear words and counted the number of MySpaces that contained the swear words. In the study, Thelwall (2008) distinguished the swear words in a sense that weak swear words were only thought as swearing only when the words were used in the context of abuse. In the case of reclaimed words, for instance *queer* was used by a gay man in describing himself which established a positive-non-offensive context and it was abusive if the word was from a stranger (Thelwall, 2008). Lastly, there are cases where offensive swear words are used in a context as endearment for instance in the case of males calling one another as *motherfucker* (Jay, 1992). This usage is not offensive because they are not intended to offend. The study by Thelwall (2008) found that most MySpaces of 16-year olds and 15% of middle-aged users contained strong swear words for both males and females. In UK, there was no substantial gender difference for strong swear words particularly for younger users (16 – 19). On the other hand, US male MySpaces contained more strong swear words than females.

The studies by McEnery (2006a) and Thelwall (2008) both focus on swear words and gender. The studies highlighted that regardless of age, social class, and gender, everyone swears. A more recent study by Gauthier *et al* (2015) looks at swear words by using a corpus compiled from Twitter; a social media. Both studies have set the participants' age

ranging from 12 years old to 60 years old (Gauthier *et al*, 2015) and 16 years old to 40 years old (Thelwall, 2008). The study focuses on the sociolinguistic features, for instance, the gendered uses of swear words on Twitter by using corpus methodology.

In the study, Gauthier *et al* (2015) use three software to assist with the quantitative data analysis namely Named-entity Recognition, Twitter Application Programming Interface, and Mention-Anomaly-Based Event Detection. Each of the software was used to yield data to answer the research questions in relation to sociolinguistics, age, and gender. Based on the analysis, it was found that there was an immense imbalance in the representation of the different age group, with a majority of the users reported being between the age group of 12 to 30 years old. Within the age group; 12 to 18 and 19 to 30, men have the highest number of tweets compared to women. However, this could be because the number of male users was higher than the number of female users. Although there was an imbalanced number of tweets between male users and female users, the study found that at least one swear word was found in male tweets by 5.8% compared to female which was only 4.8%. According to Gauthier *et al* (2015), the following are the common swear words published by male users in descending order; *fuck*, *shit*, *hell*, *cunt*, *piss*, *tit*, *bloody*, *dick*, *bitch*, *damn*, and *bastard*. As for female tweets, the common swear words in descending order are *fuck*, *shit*, *bitch*, *piss*, *bloody*, *damn*, *dick*, *tit*, *crap*, and *cunt*. From here, Gauthier *et al* (2015) calculated the log-likelihood score which was grounded by the null hypothesis that there was no dissimilarity amongst the relative frequencies of a (swear) word in the two corpora that referred to the tweets by males and females. They rejected the null hypothesis at the level of $p < 0.01$ when the log-likelihood value was greater than 6.63 as proposed by Rayson, Berridge, and Francis (2004) as cited by Gauthier *et al*, (2015). The calculation of the log-likelihood found the significant swear words for both male and female. For male users, the swear words that were significant

for them were *fuck*, *tit*, and *cunt*, whereas *bitch*, *hell*, and *bloody* were significant swear words for female users.

The findings were similar to McEnery (2006a) as his study on swear words used by men and women on MySpace revealed that *fucking*, *fuck*, *jesus*, *cunt*, and *fucker* were mostly used by males, whereby *god*, *bloody*, *pig*, *hell*, *bugger*, *bitch*, *pissed*, *arsed*, *shit*, and *piss* were mostly used by females. However, Gauthier *et al* (2015) further reiterated that though the swear words found were similar, the frequency of usage of the swear words was different. This study found that *cunt* was the most significant words used by men and *bitch* was the most significant swear word for women in which it could suggest a development in the gendered preferences of swear words (Gauthier *et al*, 2015).

Most of the studies conducted on swear words usually focused on gender and age by implementing corpus linguistics methodology. It can be concluded that the usage of swear words is contextual. Although there were no instances of swear words given because the focus was only on the occurrence of swear words in certain context, Gauthier *et al* (2015) reiterate that male and female users of Twitter swear more when talking about politics. However when male users talk about sports, less swear words are found as well as when female users talk about Princess Kate's royal baby. This suggests that the uses of swear words are contextual.

2.3.1   Semantic Studies on Swear Words

As reiterated by Goddard (2015), only a small number of studies contribute to the field of semantics focusing on swear words. This is because the use of swear words to a great extent is context-sensitive. Furthermore, linguists see swearing as part of pragmatics rather than semantics. Nevertheless, the separation line between these two fields lies in one's assumption and method (Goddard, 2015).

Provided by the limited number of literature available, the review shall encompass three frameworks in terms of semantics and swear words. The three frameworks are (i) Natural Semantic Metalanguage (Goddard, 2015), (ii) Stereotypes semantics of slurs (Jeshion, 2013), and (iii) Kaplan's framework (Hedger, 2012).

In a study by Goddard (2015), he integrated a combination of semantics and pragmatics approach on swear words. The study proposed semantic explications for expression and swear words used in American English and Australian English. The swear words that Goddard (2015) explicated are in Table 2.1 below;

Table 2.1 The swear words explicated by Goddard (2015)

Category	Swear words
Exclamation	*Shit!, Fuck!, Damn!, Christ!*, and *Jesus!*
Abuse formula	*Fuck you!* and *Damn you!*
Interrogative	*Who the fuck do you think you are?*
Imperative	*Get the hell out of here!*
Expressive adjectives	*fucking* and *goddamn*

The approach is used to show how swear words usage can be influenced by perceptions of familiarity, mutuality, and solidarity. In other words, before the utterance of the swear word *fuck* for instance, the speaker is influenced regardless of being conscious or unconscious by the three perceptions mentioned – familiarity, mutuality, and solidarity.

These terms: Natural Semantic Metalanguage, cultural script, and semantic explication, were integrated in the study that can be explicated as follows. Natural Semantic Metalanguage looks at swear words as expressions that can be attributed to particular expressive meanings (Goddard, 2015). Goddard (2015) defines Natural Semantic Metalanguage as an approach to cultural analysis and language by looking at metalanguage that is made up of a small collection of simple cross-translatable words or known as semantic primes, alongside their associated grammar of combination. Goddard

(2013) also delineates that the metalanguage of semantic prices is a model that is adapted for "expressive" meanings, social cognition, and subjective construal. Semantic explication on the other hand, is a paraphrase that is put together in a simple cross-translatable expressions extracted from Natural Semantic Metalanguage that is aimed to represent the mental representation of the meaning that a speaker utters by using specific phrase, words, or other lexicogrammatical unit (Goddard, 2015). Cultural scripts are presentations of cultural assumptions, norms, and values that are greatly formulated in the metalanguage. However, on certain occasion, it includes culture-specific words assigning culturally important concepts or category.

From an ethnopragmatic perspective, Goddard (2015) uses semantic analysis on the terms *curse word* and *swear word*. As stated previously (Section 1.1), this study assumes that these terms fall under the same category – swear word. Notwithstanding, there is no harm in explicating the terms. According to Goddard (2015), based on a figure extracted from the COCA corpus of American English, the terms are accessible and often used in both American English and Australian English. The terms are lexical keys to understand the formulations of ethnopragmatic of average Anglo English speakers (Goddard, 2015). By deconstructing the terms, it provides a platform that helps in developing scholarly understanding of swearing phenomena per se. The following tables show the explications of the terms *swear word* and *curse word*.

Table 2.2 The explication of *swear words*

words of one kind	Category
one of these words is 'shit', another one is 'fuck', there are many other words of this kind	Exemplars
many people think about these words like this: "it is bad if someone says these words" many people feel something bad when they hear words of this kind	Metapragmatic status
at some times someone can want to say words of this kind when it is like this: this someone feels something bad in one moment	Prototypical context-of-use
words of one kind	Category

Table 2.3 The explication of *curse words*

words of one kind	Category
one of these words is 'damn', another one is 'fuck', there are many other words of this kind	Exemplars
many people think about these words like this: "it can be very bad if someone says these words if if someone says these words, something bad can happen to someone because of it" many people feel something bad when they hear words of this kind	Metapragmatic status
at some times someone can want to say words of this kind when it is like this: this someone feels something bad towards someone else	Prototypical context-of-use

Based on Table 2.2 and Table 2.3, the category section is the claim that people see both terms in a taxonomic manner as "word of a kind" (Goddard, 2015). The second section recognises identified exemplars – *shit* and *fuck* for *swear words* and *damn* and *fuck* for *curse words*, whilst indicating that there are "many other words of this kind". The latter magnifies the difference between the two terms that is the metapragmatic status. Both explications begins with "many people think about these words like this:…" in which according to Goddard (2015), it spells out a widely circulated social attitude that is not similar. For *swear words*, it is claimed that the attitude is simply "it is bad if someone says these words" (Goddard, 2015). On the other hand, *curse words* are more disapproving because of the presence of the intensifier – "it can be *very* bad if someone says these words". Additionally, *curse words* have the potential of inflicting bad consequences – "something bad can happen to someone because of it", though "someone" and "bad consequences" are left unspecified. Regardless, both share the same component that indicates offensiveness.

With regards to the *offensiveness*, though both have connections to the negative feelings, there is another difference between the two terms. For *swear words*, the general state is that "someone feels something bad in one moment". On the other hand, for *curse words*, the general state is "someone feels bad towards someone else". To put it

differently, *swear words* are thought of mainly conveying a speaker's quick negative feelings. Differently, *curse words* are regarded as being used "against" someone else in which it is congruent with the grammar of the speech-act verb *curse* – a person as a direct object (Goddard, 2015).

Goddard (2015) then explicates the swear words that are used as exclamations (refer to Table 2.4 and Table 2.5), abuse formulas, question formulas, and imperative. The following is the explication of *shit!* that is used as exclamations.

Table 2.4 The explication of *Shit!* as exclamation

I know: something happened a moment before	Cognitive trigger
I feel something bad because of it	Reaction
I want to say something bad now because of this	Expressive
I want to say it in one moment	Impulse
Because of this, I say this word: {shit}	Word utterance
I think about this word like this: "some people can feel something bad when they hear this word some people think like this: "it's bad if someone says this word"	Metalexical awareness

Table 2.5 The explication of *Fuck!* as exclamation

I know: something happened a moment before	Cognitive trigger
I feel something bad because of it	Reaction
I want to say something bad now because of this	Expressive
I want to say it in one moment	Impulse
Because of this, I say this word: {fuck}	Word utterance
I think about this word like this: "some people can feel something very bad when they hear this word some people think like this: "it's very bad if someone says this word"	Metalexical awareness

Tables 2.4 and Table 2.5 are made of four sections namely (i) Cognitive trigger – an element based on semantic prime of *KNOW* and/or *THINK* that depicts situation, (ii) Reaction – that depicts *FEEL* that a model of the speaker's response towards bad feeling, (iii) Expressive impulsive – the strong restless desire to say something bad particularly,

and (iv) Word utterance – the speaker's performative utterance of a specific word. Throughout the study, Table 2.4 is applied for abuse formulas, question formulas, and imperative (refer to Table 2.1). The highlight of the study by Goddard (2015) is that the *metalexical awareness* (Table 2.4 and Table 2.5) section where it delineates the nature of swear words and how people view swear words. Based on the study, it reveals that the user is "attached" – aware, of the use of swear words regardless of the form (Goddard, 2015).

A study by Jeshion (2013) looks into the conventional or formulaic conception or image of slurs. In stereotypes semantics of slurs, the views differ in various crucial respects. According to Jeshion (2013), slurs are motivated by some of the following considerations. Firstly, the uses of slurs would bring stereotypes of the referred groups into the light effortlessly. It means that people would know whom and no prior knowledge is needed to actually know that the word *nigga* refers to a black person – *the* black person. Secondly, slurs are considered as extremely harmful that words like *jerk* or *idiot* affect the target's self-conception and self-esteem in ways that are common to the social group as a whole (Jeshion, 2013). Thirdly, as widely known that slurring is strongly taboo in society as well as in several contexts, societal taboos using slurs seem greater and it is differently originated, hence stereotypes semantics of slurs can furnish for why it is more strongly forbade. Lastly, by appealing to stereotypes, the notion possesses an obliging explanation of slurs derogatory variation in which it can explain why *nigger* is more offensive than any racial slurs like *honkey* and *limey*.

Jeshion (2013) aims to investigate the manner of stereotype expressed and the manner of the encoding of it would affect the truth conditions by using the theories developed by Hom (2008) and Camp (2011). Hom (2008) reiterates that a slur's offensiveness originates from semantic content. The semantic content is characterized into

nonindividualistically and externally that it is similar to that natural kind terms possess externally characterized contents. The meaning of slurring terms is not characterized solely by the speakers' belief about the socially applicable group referred by slurs' indifferent counterparts but it is also by the relation that the speaker has to societal institutions of homophobia, sexism, racism, and others. The societal institution contains ideology that is a set of beliefs that is normally negative about certain group and a set of practices concerning how the groups are dealt with the ideology typically proposing to make such practices legal. Semantic content that slurs possess express the properties that is in the ideology alongside with the menace of discriminatory exercises towards the group. Refer to the following example used by Hom (2008):

(1) Yao is a Chink.

Sentence (1) contains a slur term that is *chink*. Hom (2008) claims that the word expresses a complex socially constructed property for example that Yao being a *chink* is subjected to higher college admission standards and is a subject to the exception from advancement to managerial positions. He adds more that this is due to the fact that Yao possesses slant-eye, devious, good at laundering, and many more just by being a Chinese. With regards to the racist that utters sentence (1), says something false, let it be for neither Yao nor any Chinese person of being a *chink*, no one ought to be a subject to such practice just by being a Chinese. For Hom (2008), he agrees to this because intuitively what is uttered by a racist as example (1) must be false. Highlighting the offensiveness of a slur, it is originated and derives from the social institution in which it expresses the semantic content that does not require ones to be "conscious", "in the mind", and "internal to" for the content is externally characterized. This is also chastity according to Hom (2008) because it grants the slur to take over "derogatory autonomy in which it is the capacity to offend the irrespective attitudes of the user. This semantic externalism permits the theorist

to explain the manner of a speaker who is not aware of the particular practices and ideology to be at any rate a minimal competent user of such term.

The notion of stereotypes semantics of slurs was criticized in a way that it was not important and offensiveness was widely known to reside in slurs even in negations, modals, conditionals, and other various forms. Refer to the examples by Jeshion (2013);

    (2) Yao is not a Chink. He's a Jap.

    (3) If there are too many Chinks in the kitchen, my father won't eat there.

    (4) Is he a Chink?

    (5) Get that Chink in here.

    (6) Chink!

According to Jeshion (2013), if the slurs are used by a racist, the source of offensiveness should be the same in sentence (1) to sentence (6). In sentence (2), it was denied that Yao is a Chink. However, according to Hom (2008), the offensiveness in sentence (1) is because of the speaker predicating something offensive about Yao. Regardless, he adds that there is no explanation to account for the first sentence in sentence (2) because it expresses that Yao is not a Chink semantically. However, Hom (2008) reiterates that sentence (2) is still offensive pragmatically because though it is not specifically Yao, there is someone other than Yao that ought to be subjected to Chink because of being Chinese.

Another study on slurs that looks specifically on racial slurs was conducted by implementing Kaplan's framework to provide theory of meaning of the racial slurs (Hedger, 2013). Kaplan's framework was adopted for differentiating expressive and descriptive content of the racial slurs. According to Hedger (2013), this is rather difficult for truth-conditional semantics due to the racial slurs' projection behaviours. A little bit of insight on truth conditional, it can be summarised that the meaning of an argument is

characterized by the meaning of its' parts (Hedger, 2013). Refer to the following examples used by Hedger (2013):

(7) David is intelligent

Based on sentence (7), according to Hedger (2013), meaning is either is, or what determines, a truth value. Hence, sentence (7) is true if David is intelligent and is false otherwise. Hedger (2013), the extension of sentence (7) will be truth and the way to understand the intension is based on the possible states of affairs to truth values. Consider sentence (8), (9), and (10);

(8) David is intelligent and Canada is cold.

(9) If David is intelligent, then so is Judith.

(10) It is not the case that David is intelligent.

If someone asserts sentence (8), he/she also asserts sentence (7). However, if someone asserts sentence (10), means he/she does not assert sentence (7). This is because the truth value of sentence (8), but not sentence (9), depends on the truth value of (7). On the other hand, sentence (10) is out of the discussion since sentence (10) is of course false whenever sentence (7) is true. Consider the following sentence containing swear word but in the article by Hedger (2013), he avoided from using any racial related swear words;

(11) If Obama is an S, then so his wife.

In conditional semantics, the projection behaviours of swear words are a challenge to truth conditional semantics (Hedger, 2013). He adds more to the account that the use of swear words regardless of any circumstances stays offensive although it is laid under negation. Thus, sentence (11) without further explanation is already offensive. In this sense of sentence (11) contrasts with sentence (9), because asserting sentence (9) does not mean the speaker to predicating the intelligence of David or Judith. Sentence (9) in

this case could still be true even if neither both of them are intelligent. Nevertheless, if someone utters sentence (11), it is still something offensive (Hedger, 2013).

In explaining Kaplan's framework, the understanding on truth conditional is crucial because Kaplan adopts a technique from truth conditional semantics for him to bring in the difference between two different semantic contents. In the case of two statements that are similar in semantic content, valid inference can be made from one to the other (Hedger, 2013). Refer to sentences below used as examples by Hedger (2013);

(12) David is polite.

(13) David is courteous.

(14) Kaplan was promoted.

(15) That damn Kaplan was promoted.

In sentence (12) and (13), Hedger (2013) assumes that the word "courteous" and "polite" are synonyms and approximately match in semantic content to infer sentence (13) from (12) or (12) from (13) validly. However, according to Kaplan (2005), he notes that intuitively sentence (14) and (15) cannot be inferred validly but we can validly inferring sentence (14) from (15). This is because although sentence (15) contains the semantic content of sentence (14), sentence (15) possesses something extra that displays an attitude of contempt on the part of the speaker. Sentence (14) is composed of purely descriptive content in which it constitutes the world as being in a certain way that it can be false or true. However, with the extra expression such as *damn* (used as an adjective), it contains only expressive content (Hedger, 2013). Expressions as such are simply exhibit an attitude of the speaker in which these expressions are not truth-apt. Kaplan adds more that the use of *damn* is the same as when someone utters "óuch" that they express (roughly) an abrupt pain and it makes no sense to say that what was said is true or false.

In a way, sentences that contains swear words are not true because the sentences are purely expressive.

2.4   Corpus Linguistics

Corpus linguistics is the study of language through computational analyses of large collections of texts known as a corpus. The corpus consists of naturally occurring data usually analysed with specialised software programmes. To assist the study in corpus linguistics, common tools used include AntConc, Wmatrix, Wordsmith, and others. The application of these tools helps the researchers to generate word lists, concordance lines, and word frequencies among others.

In the study of corpus, it involves empirical and analysing actual patterns of language use in texts. It also utilises a large and objectively defined collection of natural text; corpus, as the foundation for the analysis. Since corpus linguistics analysis involves the use of tools as mentioned, computers are used to conduct the analysis by using both interactive techniques and automatic that includes both qualitative and quantitative analytic techniques. Having said that, computers help in identifying and analysing complex patterns of the language used in which it allows the analysis of a larger corpus that would be rather difficult to handle by hand.

2.4.1   Collocation

Numerous studies in corpus linguistics have been using collocation to conduct their investigation by looking at collocation. However, this study will not use collocations to identify the semantic prosody. Instead, it is important to note that collocation is part of the study in corpus linguistics that several researchers used collocation to identify semantic prosody like Rahim (2012). To investigate collocation, the nature of such study

necessitates the researcher to decide few matters. Firstly, a researcher has to decide the number of collocates to be studied on each side of the node word. As cited by Martinez (2008), the number of words within the span varies as in 4 span for Clear (1993) and 10 span for (Martin *et al*, 1983). Secondly, the researcher also has to decide whether to include or to exclude the grammatical words. If the investigation is about phrasal verbs of prepositions adjectives, it tends to and would evidently demand the presence of the grammatical words. Thirdly, a researcher has to decide for the corpus is to be pre-processed or not. In the case of Stubbs (1996), he observed that dissimilar word forms can have dissimilar collocates and he provides example as in the case of *educate* that co-occurs with *enlighten*, *help*, *inform*, and *train*, *education* that collocates with words referring to institutions, and also *educated* that repeatedly followed by *at*.

To define collocation, let's start with Martinez (2008). According to Martinez (2008), collocations are the co-occurrence of words within a short span in a text that is statistically significant. Additionally, the collocations occur possibly because of grammatical rules for instance in the case of *the cat* in which it involves *determiner + noun*. Another case is, it could be because of the words are a component of an idiom, saying or proverb. Sinclair (2003) defines collocation as two or more words occurring near each other in a text. According to Sinclair (2003), the usage varies between researchers on lexis; some would reserve it for statistically significant and use the term co-occurrence to describe simple physical event, and others would require it to be mapped precisely onto a complete grammatical structure for instance a noun phrase. Hunston (2002) on the other hand reiterates that collocation is the statistical tendency of words to co-occur.

In the definition provided, Martinez (2008) and Hunston (2002) describes collocation in a sense that it involves statistics. This is because the collocation is identified by using statistical measurement for example by implementing T-score, Mutual Information score

(henceforth, MI), or Z-score among many others. Based on the definition provided, collocation examines a word; node word, by looking at words that are close to it. In this case, it is the span of the words and it depends on the researcher on determining the span (Martinez, 2008).

Directing the discussion to the statistical measurement, the following discussion focuses on T-score, MI score, and Z-score based on Martinez (2008). She further reiterates that T-score considers frequency in which the score provides more accuracy when dealing with low frequency words. MI, as cited by Martinez (2008), is determined by comparing the probability of word *x* to co-occur with word *y*; the joint probability, by observing the words independently; chance, (Church *et at*, 1991). Z-score on the other hand, though very similar to T-score, it compares the observed frequency between a node and its' collocates to the expected on and then evaluates the difference between the values by using standard deviation. The three significance measures are most commonly used in collocational studies as in Stubbs (1996) and Martinez (2008).

To compare the values by using the measurements, Martinez (2008) uses two corpora published by Oxford University Press; Corpus Collections A and Corpus Collection B. The corpora both consist approximately one million words; 2, 047, 903 words, made up of written and spoken language samples of various domains. Based on a significant number of frequency in the corpus, the chosen node word is *time*.

However, it is not mentioned in the study the tool(s) used to generate the occurrences. By reasonable assumption, most of the tools available nowadays have been incorporated with such measurements. Taking into consideration a tool named GraphColl, it has the feature of 14 statistical measurements that includes frequency, MU, MI, MI2, MI3, Log-likelihood, Z-score, Dice, LogDice, T-score, LogRatio, MinSens, DeltaP, and Cohen. Other tools such as AntConc also has the feature but less comprehensive in which it has

MI and T-score. Wmatrix on the other hand has 11 statistical measurements that include MI, MI2, MI3, Phi-square, Log-likelihood, Ochiai (OCH), McConnoughy Coefficient (MCC), Yule Coefficient (YUL), Fager and McGowan Coefficient (FAG), Kulczinsky Coefficient (KUC), and Simple Matching Coefficient (SMC). Regardless of the list of the statistical measurements, aforementioned, the most commonly used statistical measurements are T-score, MI, and Z-score in which all of the mentioned statistical measurements aim to identify significant collocates of a node word.

Shifting the attention to the three most commonly used statistical measurements; T-score, MI, and Z-score, Martinez (2008) identifies the collocation of *time* within a span of five words to the left and right of the node word. In the corpus, there are 3, 372 occurrences of *time* with a total of 712 different collocates.

Table 2.6 The significance measures compared

T-score	MI	Z-score
For	Zurvan	Same
First	Ripe	Has
Same	Cues	Space
Has	Cyclical	Cues
This	Before	About
Space	Spends	First
About	Spend	Before
The	Much	Much
When	Second	Spent
Long	Space	Second
Was	Finite	Could
Much	Lapse	For
Some	Wasted	Spend
Before	About	Into
Could	Kerr	Zurvan
Cues	Spent	Ripe
Into	Measuring	Long
Spent	Has	This
Full	Occurring	Full
Half	Waste	When
Second	Eternal	Use
Since	Use	Such
Come	Could	Party
Such	Cue	Waste

Spend	Into	Cyclical

Table 2.6 above shows the result of the collocations ordered based on their values for each measurement. The words in red are words occurring in the three columns, the words in blue are words occurring in Z-score and MI columns and the words in green are words occurring in Z-score and T-score columns (Martinez, 2008). Based on the table, there are two outstanding differences identified; (a) words which are found in a list but absent in another list and (b) words whose ranks are different between the lists. The table shows that 44% or 11 words occur in all lists; in red. Having a closer look at the table, 20% or 5 collocates in both Z-score and MI columns cannot be found in T-score list namely *cyclical*, *ripe*, *use*, *waste*, and *zurvan*. These words, however, have low frequencies. On the other end, there are a number of words that possess high frequency which present in the T-score list but absent in the other two lists for example *come*, *half*, *since*, *some*, *the*, and *was*. To compare the lists similarity, MI and Z-score results are closer in which 64% of the words are similar whist between MI and T-score is only 44%.

In terms of frequency, both MI and Z-score have words with low frequency listed as significant collocates for instance the word *ripe* and *zurvan* are uncommon occurrences in the corpus. Appearing in the top position for MI list is puzzling because the words only have 5 occurrences in the case of *zurvan* and 8 occurrences for *ripe*. However, words with high frequency are absent in both of the list in the case of *the*.

Martinez (2008) delineates that MI and Z-score amplify the significance of low frequency co-occurring words that left a more reliable statistical measure that is T-score, when dealing with collocates of low frequency in a corpus. Furthermore, the MI and Z-score also diminish collocates that possess high frequency in the whole text, whilst T-score on the other hand highlights words with high frequency in the case of *the* that shows

high frequency in the corpus studied. Generally, T-score is likely to highlight frequent recurring words in which the words include grammatical words such as personal pronouns, prepositions, particles, and determiners alongside lexical words that strongly associated with the node word. MI, however, highlights the technical phrases, idioms, proverbs, and fixed compounds. Martinez (2008) concludes that the differences amongst the three statistics as contrasted and the information they have that both Z-score and MI values are almost similar than T-score.

2.4.2 Semantic Prosody

Generally, semantic prosody evaluates lexical items in a way that the lexical items can be graded in terms of favourableness – favourable or unfavourable. According to Stubbs (2002), semantic prosody is relevant to the collocational meaning that rises from the interaction of a node and its' typical collocates. Louw (2002) reiterates more that it is a form of meaning that is established through the closeness of a consistent series of collocates. Numerous studies have been conducted to look at the semantic prosody of words, for instance *budge* or *brook* by Sinclair (2004), *happen* or *sheer* by Partington (2004), and *perempuan* (female) and *wanita* (woman) by Rahim (2012). According to the studies cited by Hunston (2007), semantic prosody is neither described as 'good' nor 'bad' but rather as gradable; a word may have a 'more or less' either favourable or unfavourable prosody (positive or negative) (Partington, 2004). The studies by Sinclair (2004), Partington (2004), and Rahim (2012) exemplify the idea that synonymous words are seemingly positive but may actually be used in a negative light (Stewart, 2010). This is especially true in the case of *cause* discussed in Stubbs (1995).

Sinclair (2003) reiterates that semantic prosody arises from corpus linguistics, and in particular the "phraseological" tradition that centralises the typical behaviour of single

lexical item as observed using "key word in context" concordance lines. Various researchers have adapted the concept in their studies, for example Stubbs (1996, 2001), Togini-Bonelli (2001), Partington (1998, 2004), McEnery (2006), Xiao and McEnery (2006), McEnery (2006), Rahim (2012), and Hashim and Rahim (2016). Though there are various studies that apply the notion of semantic prosody, it was criticised by Whitsitt (2005) and Stewart (2010). This section focuses on two parts of the criticism briefly by Stewart (2010) that is the *hidden* criteria of semantic prosody and the aspect whether the study on semantic prosody should either be diachronic or synchronic.

One of the characteristics of semantic prosody is the *hidden* criteria as often described in the body of literature. The words to describe semantic prosody as listed by Stewart (2010) are in a sense of *hidden*, subliminal, and 'unconscious' nature. Partington (2004) describes semantic prosody as evaluative meaning that is concealed to the naked eyes. In the same sense, Louw (1993) argues that semantic prosody can disclose the speaker's attitudes although the speaker's attempt to conceal them. This suggests that semantic prosody functions subliminally without the speaker even being conscious about it. There are a number of researchers that describe semantic prosody as mentioned such as Hunston (2001) and McEnery, Xiao, and Tono (2006). However, Stewart (2011) has his own views on this. To him, there is not enough evidence to say clearly that semantic prosody needs to be regarded as *hidden*. This, however, will be discussed in depth in the discussion (Chapter 5).

Stewart (2010) argues in the case of a word to acquire new meaning, this process itself takes a period of time rather than a fixed period of time; synchronic rather than diachronic. This is based on the metaphors used by scholars in defining semantic prosody. The metaphors are 'imbued' (Louw, 1993), 'shift' (Bublitz, 1996), and 'takes on' (Hunston & Thompson, 2000). These metaphors used to describe semantic prosody do have the sense

of over a period of time rather than a fixed time. Due to this, he argues that the concept is not well defined that it falls into the diachronic framework. Hence, current study shall provide the information needed whether it should fall under the diachronic framework or synchronic framework.

Based on the definition provided earlier, Stubbs (2002) and Louw (2002) define semantic prosody by incorporating the word collocation; collocation (Stubbs, 2002) and constant series of collocates (Louw, 2002). To look at semantic prosody, the study needs to look at the collocation that appears constantly. In other words, to determine the semantic prosody of a word, *perempuan* (Rahim, 2012) for instance, looked at the frequencies that appear with the node word in determining the semantic prosody. In the study, the word *perempuan* is appeared to collocate with nouns that have the sense of family. The left collocates of *perempuan* include *adik* (sister), *anak* (daughter), *cucu* (granddaughter), and *menantu* (daughter-in-law). On the right side, the word *perempuan* collocates with words that have the sense of age and adjectives that are considered negative for example *sulung* (first child), *malang* (unfortunate), *mangsa* (victim), *pelacur* (prostitute), and *simpanan* (mistress). Hence, the identification of semantic prosody depends on both collocation and frequency.

Certain words possess positive or negative semantic prosody depending on the collocational environment or the sharing of words (Rahim, 2012). In a study by Stubbs (1995), the word *cause* possesses negative semantic prosody and the word *provide* possesses positive semantic prosody. This conclusion was drawn based on the collocational analysis by using Cobuild corpus that owns 120 million words. By observing the collocations, Stubbs (1995) found that *cause* possesses negative collocation because most of the words that collocate with *cause* are less favourable for instance *accident*, *concern*, *damage*, *death*, and *trouble*. On the other hand, the word *provide*

possesses positive semantic prosody because the collocates are nouns that possess positive meaning for instance *aid*, *assistance*, *care*, *employment facilities*, *food*, *funds*, *jobs*, *money*, and many more.

This brings us to the next issue of connotation that is inherited because of its' collocation as addressed by Rahim (2012). To define connotation, according to Jackson (2014) and Jackson and Ze Amvela (2000), it is rather problematic to be defined because of the disagreement due to the distinction between connotation and denotation. Lyons (1977) and Leech (1981) on the other hand do not accept the differentiation and they prefer using connotation and denotation in a rather particular sense. Regardless of the disagreement, connotation is closely associated with synonymy (Jackson and Ze Amvela, 2000). In other words, synonyms have the possibility to share the same denotation for instance cognitive or conceptual meaning, but this is rather different in the case of connotation. However, it is rather difficult to draw the distinct line between connotation, which is essentially stylistic, and denotation. In Scottish English, for instance the word *loch* (vs. lake) is either a dialectal marked variant of the Standard English word or as a different lexeme with a particular descriptive meaning or denotation, and additionally a connotation of Scottishness (Lyons, 1977). Jackson and Ze Amvela (2000) defines denotative meaning in which it refers to the relationship between a linguistic sign and its' denotatum or referent. Connotation on the other hand, constitutes additional properties of lexemes for instance baby language, slang, and poetic, rhetorical, casual, humorous, formal colloquial, legal, literary, and biblical (Jackson and Ze Amvela, 2000). Using Rahim's (2012) example in the case of *matahari* (sun) and *suria* (solar radiation), both denote the same meaning that relates to the sun. According to Oxford Dictionary 9th edition (2008), the sun is the star that shines in the sky during the day and it provides the earth both light and heat and solar is connected with the sun; solar radiation. In this

context, the word *suria* is mostly found in literary texts such as poems, songs, and novels that the collocations bring forth a romantic connotation in which the word *matahari* lacks.

The word *suria* acquires its' connotation through the frequency found in the literary texts such as poems, songs, and novels that it acquires the connotation of romantic. This directs to the next important element that influences language use. Using Rahim's (2012) example in the case of *terror*, in English "He is a terror" simply means that the *he* is a person that inspires fear or dread or a very troublesome person. On the other hand, in Malaysia and still referring to the same word, the word *terror* means the opposite thing that is *remarkable* or *noteworthy*. This happens because there is a chain of meaning that exists because of the frequent usage of the word in such way (Rahim, 2012). Nevertheless, the highlighted issue arises because of behaviour of collocation. The meaning; connotation, is fluid because it depends on the context as in the case of *suria* and *matahari* discussed (Rahim, 2012).

In the case of *perempuan* and *wanita* (Rahim, 2012), *wanita* possesses positive collocation, hence positive semantic prosody and negative semantic prosody for *perempuan*. This conclusion was drawn based on the frequency of the collocations – let it be positive or negative. The number of concordance lines found in the corpus for *wanita* is six times more than *perempuan*. It shows that the word *wanita* is mostly used in newspapers reports (Rahim, 2012). In terms of usage, early speculation was made by Rahim (2012) that the word *wanita* is mostly used. To acquire the context of the analysis, Rahim (2012) looks at the collocations in the span of five words to the left and five words to the right; 5L and 5R. According to Sinclair (1991), the implemented analysis provides the semantic analysis of a node word. Sinclair (1991) further explains that if *a* is the node word and *b* is the collocate, the is called the downward collocation. However, in the case of where *b* is the node word and *a* is the collocate, this is called the upward collocation.

In the study, Rahim (2012) only looks at the nearest collocates in the case of *a* and *b*; direct collocation. The tool used to assist the analysis is WordSmith Tools that enables the researcher to look at concordance lines hence analysing it in which only words that collocate with the node word twice and above was considered in the research. Based on the analysis, there are 863 direct collocations for *wanita* and 114 collocations for *perempuan*. Be that as it may, the study only takes content or lexical words into account that are nouns, verbs, and adjectives. This is because this study focuses on semantic rather than syntax; *semantic* prosody. Hence, function words are excluded. This study, however, looks at content words as well in determining the semantic prosody of swear words.

To determine the semantic prosody of *perempuan* and *wanita*, Rahim (2012) groups the content words into three groups; nouns, verbs, and adjectives, for both *perempuan* and *wanita*. From there, the collocates of the words then analysed for example nouns class for *perempuan*, the words that collocate to the left mostly made up of words that shows the theme of family for example *adik* (little sister), *anak* (daughter of), *cucu* (grandchild of), *menantu* (in-law), and *saudara* (cousin). The word *perempuan* also frequently collocates with words that describes gender for example *budak* (child) *jantina* (gender), *kanak-kanak* (children), *pelajar* (student), and *murid* (pupil). Based on the data, the word *perempuan* has the tendency to collocate with words that is neutral in terms of gender (Rahim, 2012). It was also found that the word *wanita* has the same features with *perempuan* when it comes to gender for example *golongan* (group), *kaum* (race), and *remaja* (teenager) in which this is not as frequent as *perempuan*.

In her study, Rahim (2012) identifies the collocations for both *perempuan* and *wanita* for both verbs and adjectives.. For adjectives, the word *perempuan* is considered negative because of the word collocates with words that are considered negative; *malang* (unfortunate), *mangsa* (victim), and *simpanan* (mistress). Based on the identified

collocates, Rahim (2012) identifies the semantic prosody by merely looking at the meaning of the collocations whether it is positive or negative in the given context. The reason why collocation is not used in this study is because the study conducted by Rahim (2012) was on Malay language in which the adjectives always follow the nouns for example *perempuan malang* (unlucky female) in most cases. This however could not be conducted with current study because the collocations identified at first did not show any *hidden* meaning the swear words possess. Thus the semantic prosody could not be identified.

In Rahim's (2012) explanation, she "considered" or perceived that *wanita* possesses positive semantic prosody based on the perception of the collocations that is perceived positive for instance *baik* (good woman), *berani* (daring), *beriman* (pious), and *bijak* (smart). In the case of *Islam* (Hashim and Rahim, 2016), the semantic prosody of the word *Islam* is identified by looking at the collocations. The theme arises from the collocations determine the semantic prosody for instance these words; *said*, *told*, *implement*, *represent*, and *entitled*, "suggest" neutral prosody of *Islam* in which it is supported by the analysis on concordance lines of the verbs which portrays that they have neutral prosody (Hashim and Rahim, 2016). The study concludes that the list of collocations "suggest" that the word *Islam* has a positive prosody whereas the verb collocates; *menghina* (insult/degrade), *mencerminkan* (reflect), *mengaitkan* (associate), *hina* (insult) and *meletakkan* (place), "suggest" that Islam is being taken advantage of and interpreted falsely (Hashim and Rahim, 2016).

Similarly, Hunston (2007) reiterates similar notion concerning semantic prosody. Hunston (2007) states that the discussions on semantic prosody tend to encompass on one of two types of consistency: (a) consistent co-occurrence of (types of) linguistic items as

in Partington (2004) or (b) consistency in the discourse function of a sequence of such linguistic items as in Sinclair (2004).

Numerous researchers have adopted the notion that focuses on consistency in the co-occurrence of items especially when it concerns the behaviour of a word or of a short multi-word unit for instance phrasal verb. Similarly, Rahim (2012) and Hashim and Rahim (2016) conducted the study by implementing the same notion in which the phrase or word is observed to co-occur frequently with a wide range of items that do not belong to a semantic set, rather they have in common a particular attitudinal meaning (Partington, 2004). Based on a study by Stubbs (1995), the words; collocates, indicating what entity is *caused*, "considered" to be undesirable for instance the words that collocates with the word *caused* are *confusion*, *a rift*, *anger*, *audience fragmentation*, *a kidney stone*, and so on in which it depends on context. Nevertheless, the example on *cause* exemplifies that the attitudinal context of a lexical item may depend on other aspects of items in the immediate co-text.

Focusing on the notion of context as discussed only on the surface level, the following discussion concerns about the notion in depth. For that purpose, I shall take a step backward to add more on semantic prosody. Aforementioned, semantic prosody is an aspect of evaluative meaning in which it is the attitude of the speaker, writer, or stance towards viewpoint or feeling towards entities and propositions that is being mentioned about (Hunston and Thompson, 2000). In the case of *happen* and *set in*, Sinclair (1987) first noticed the items usually associated with unfavourable events in the data in hand. This in a way states that a particular node possesses particular semantic prosody depending on the context that it is being used. Hoey (2004) reiterates that a word is deemed for collocational use in which he adds a learner comes across a word regardless of through reading or interaction; speaking, it is packed with knowledge of the word that

it co-occurs with another words; collocation. In the case of swearing, the words are normally found in unfavourable environment in which resulting the taboos across regions. *Set in* in this case is not normally found in a favourable environment.

Stubbs (1995a) looks at the lemma CAUSE and found that the lemma is used vastly in a context where the cause and effect are unpleasant. In the other words, there are also CAUSE that is used in a pleasant context. Stubbs (1995a) delineates more by arguing that the size of a corpus influences the positivity/negativity/neutrality of a word. The lemma CAUSE appears in all genres represented in Lancaster-Oslo-Bergen (henceforth LOB) corpus where it is negative presumably because newspapers' report that focuses on crises and disaster. Hence, semantic prosody as reiterated depends on frequency of words that is in a context in which it reflects the components of the corpus.

2.5 Summary

In this section, relevant studies have been reviewed pertaining to the research. There are also studies found that analyse songs linguistically by focusing on the language features in songs. Apart from that, some of the studies also incorporate corpus linguistics methodology to identify context of certain genres in songs and this is useful in assisting my study especially in the identification of semantic prosody of swear words in English songs. Studies on swear words usually focus on gender, age, and the use of swear words. In addition, there are also studies conducted on swear words to identify the offensiveness of swear words usually by using Likert scale. However, merely rating the swear words ignores the complexity of the swear words. This is the part where the notion of semantic prosody being incorporated that is to look at a word in collocations to understand how certain words behave within certain context.

CHAPTER 3: METHODOLOGY

3.1   Introduction

This chapter mainly focuses on the methodology that is used to conduct the study. This chapter is divided into several sections that restate the research questions (Section 3.2), the CoES (Section 3.3), the corpus design (Section 3.4), data collection (Section 3.5), ethical consideration (Section 3.6), data analysis procedure (Section 3.7), and lastly summary of the chapter (Section 3.8).

3.2   Research Questions

To recap, this study seeks to find answers to the following research questions;

1. What are the swear words found in a corpus of English songs that comprises of songs collected from Billboard starting at year 2011 to 2016?

2. What are the top 10 swear words found in the sub-corpus of English songs in from 2011 to 2016?

3. What is the semantic prosody of swear words found in a corpus of English songs?

To assist with the data analysis, AntConc 3.4.4w was used to identify the most frequent swear words for research question one. To answer the second research question, the word list tool in the software counted all the words in the corpus and the software presents the data in an ordered list to be analysed. Concordance lines were generated by using AntConc 3.4.4w to generate data to answer the third research question.

3.3   Corpus of English Songs

The corpus used for this study primarily consists of English songs. The songs were compiled from 2011 to 2016 in which each year, the top 100 songs from the Billboard are chosen. The Billboard is a platform established by an American entertainment media

brand that tracks the most popular songs and albums in different genres. Apart from posting news about singers, Billboard has charts that rank famous songs and the famous charts are *Top Hot 100* and *Billboard 200*. The chart selected for the songs selection was the Hot 100 songs year-end chart (https://www.billboard.com/charts/hot-100). There are some repetitive songs found on the chart, for example songs on 2011 chart can be found on 2012 chart. To mitigate this, the repeated songs were not included in the list of the subsequent year for instance if the songs on 2011 chart were found on 2012 chart, the songs on 2012 chart were not included. In other words, they were counted only once. The total number of songs that were omitted because of repetition were 51 songs. Apart from the repetitive songs, few songs were omitted because the songs were not in English for instance Gangnam Style by Psy (Korean song) found in the list of songs in 2012 and 2013. A total of 4 songs were omitted because the songs were in Korean and Spanish.

Table 3.1 The distribution of CoES

Year	Number of tokens	Number of word types	Total songs
2011	48,082	3,299	100
2012	37,869	2,759	91
2013	38,782	3,003	89
2014	39,159	3,055	88
2015	39,159	3,055	91
2016	40,260	2,832	86
Total	**243,689**	**8,139**	**545**

Based on the criteria, the total number of the chosen songs from the Top 100 year-end chart are 545 songs from 2011 to 2016. A total of 55 songs were omitted from 600 songs because the songs did not fulfilled the required criteria mentioned that is either repeated or the songs were not in English. Table 3.1 shows the total number of tokens, word types according to the year, and total songs. Each sub-corpus contains more than 35,000 words as shown in Table 3.1. This is because it is influenced by the total number of songs each year. However, sub-corpus 2013 has more number of tokens compared to sub-corpus

2012 although sub-corpus 2012 has more number of songs. This happens due to the number of words a lyric might have and each lyric would have different number of tokens.

3.4   The Corpus Design

There are four aspects of corpus design that will be discussed here; size, content, balance and representativeness, and permanence (Hunston, 2002). The corpus size for this study is relatively small compared to the BNC that has over 1 million words. The question of size can be contentious. A small corpus can be useful if the software used cannot store large amount of data. However, a small corpus yields a small amount of data. To mitigate that, researchers need a large corpus. Carter and McCarthy (1995) as cited by Hunston (2002), argue that for the purpose of studying grammar in spoken language, a small corpus is sufficient. Hence, the size of the corpus depends on the study itself. In this study, the corpus comprises of songs in which the minimum number of words in a song is estimated around 300 to 500 words. This leads to the next point that is the content of the corpus.

Hunston (2002) raised a few questions concerning the content of the corpus such as what to be included and how will the materials be selected. As mentioned in Section 3.3, the songs chosen for this study are based on Billboard Hot 100 songs year-end chart that contains songs across different genres. As mentioned by Katznelson *et al* (2010), each genre has specific context and it could yield different results compared to the study conducted by Falk (2012) that only focuses on a specific genre. The various songs of different genres are ranked by a company named Nielson Company that monitors the online streaming, radio airplays, and music consumer behaviour.

The focus of this study is swear words, hence it would be difficult to look for materials that contain many swear words. In the same vein, McEnery and Xiao (2004) reiterated

that it is indeed difficult to gather large-scale natural language swearing data due to the fact that written texts are usually written in language registers that would exclude swearing. In other words, the written texts are context governed. In this case, the amount of swear words found in songs is very limited because there are no songs that are written solely by using swear words.

3.5   Data Collection

The corpus of this study consists of lyrics from 545 songs based on the Billboard Hot 100 songs (Table 3.2). Since the study looks at songs from 2011 to 2016, the collection of songs began from 2011 chart and so on to identify repeated songs as well as songs that were not English (see Section 3.2).

Table 3.2 The list of songs in 2012

No.	Title	Artist(s)
1	Somebody That I Used to Know	Gotye featuring Kimbra
2	Call Me Maybe	Carly Rae Jepsen
3	We Are Young	Fun featuring Janelle Monáe
4	Payphone	Maroon 5 featuring Wiz Khalifa
5	Lights	Ellie Goulding
6	Glad You Came	The Wanted
7	Stronger (What Doesn't Kill You)	Kelly Clarkson
9	Starships	Nicki Minaj
10	What Makes You Beautiful	One Direction

The lyrics were copied from AZLyrics (https://www.azlyrics.com/) and transferred into Microsoft Words (doc) for it to be saved into txt format. The txt files were then labelled according to its' number for example 5.Ellie – Lights based on Billboard chart 2012 as shown in Table 3.2. To keep the files neat, the files were separated according to year. This study used materials that are available online and can be assessed anywhere at any time by using any available search engine. In this case, the lyrics were browsed by

using Google Chrome version 63.0.3239.132 (Official Build) (64-bit). Using the search engine, song lyrics are searched by typing in the keywords from the song title and the name of the singer followed by the word "lyrics" for example, "Ellie Goulding Lights lyrics". This process was repeated for all of the songs chosen for this study. To retrieve the lyrics, AZLyrics was chosen because it ranked the first in the search engine and it is also listed as one of the sites mentioned by Falk (2012) in compiling songs for her study.
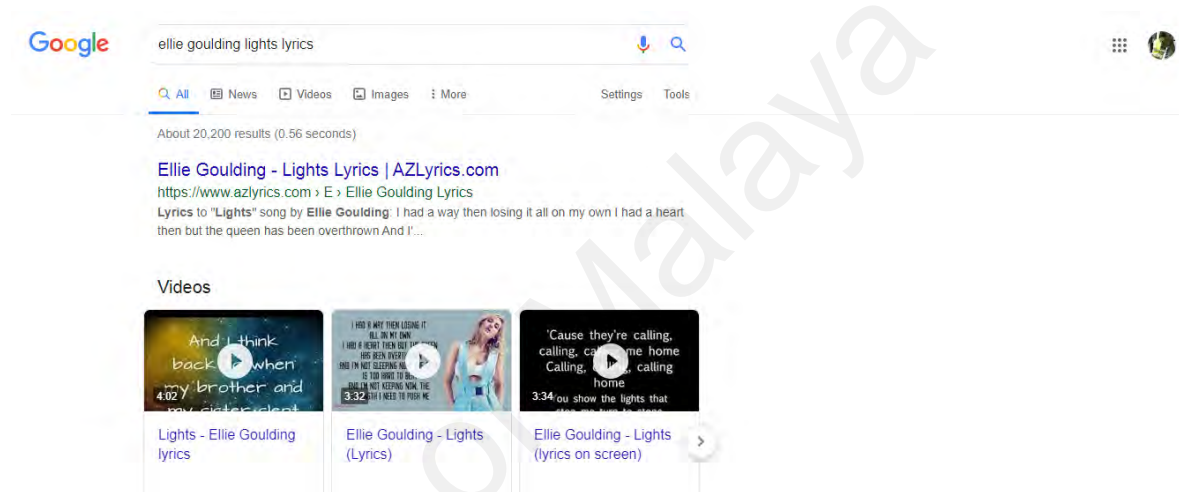
Figure 3.1 Screenshot of the search result for the lyrics

3.6   Ethical Consideration

Since the current study acquired the lyrics online, there is a case of considering ethical issues in the use of publicly available materials. In terms of consent, a number of researchers consider that the materials made public; online, needs no consent (Garton, 1997) as cited by Sixsmith and Murray (2001). However, this moves to the next issue that is the copyright regulations.

Figure 3.2 Screenshot of the copyright of AZLyrics

The policy in Figure 3.2 (https://www.azlyrics.com/) stated that all lyrics are the property and copyright of their owner and the lyrics provided online are only for educational purposes and personal use only. Hence, there is no major ethical issue faced by this study.

3.7   Data Analysis Procedures

Figure 4.1 below summarises the flow of the data analysis. A list of words was firstly generated by AntConc to assist with the study. The swear words were then identified by using the HFSN notion of swear words highlighted by Bergen (2016). The swear words were chosen by considering both the frequency of the swear words and the number of songs that the swear words occurred in. It is important to note that some swear words may be highly frequent, but they are not widely distributed in a number of songs. The songs, however, are still chosen to be included in the analysis because of the scope of the study that is top 10 lemmas – in the case of *hoe* that has similar frequency as *stupid* that is 43 but it was found in 11 songs for *hoe* and 25 songs for *stupid*. This closely corresponds to what Sardinha (1999) proposes in reducing the set of words to be analysed that is by selecting a simple majority. The example provided by Sardinha (1999) is plus one to the half of the total number of words. To provide simple mathematics calculation, if the total number of swear words found in the corpus of English songs is 10, the total number of swear words to be analysed will be 6 since half of 10 is 5 and plus one. Ultimately, the

cut-off point is important to take out underused (Rayson, 2008) swear words to focus more on the significant swear words.
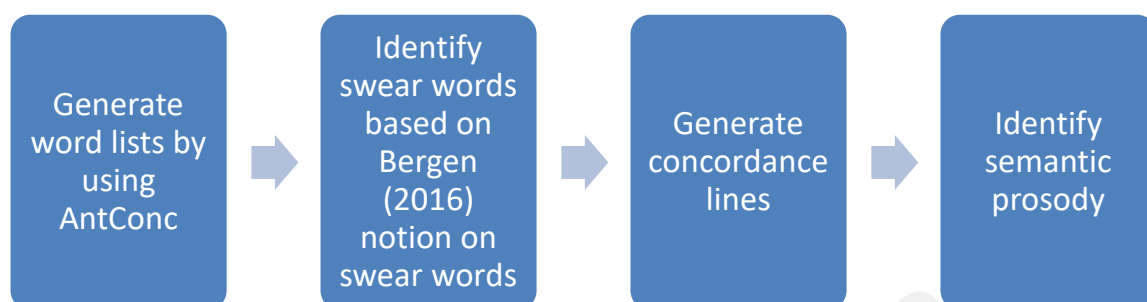
Figure 3.3 Flowchart of data analysis procedures

Stubbs (1996) contends that some words have a greater influence of negative prosody, a couple of words possess a positive prosody, and many words are neutral. In this case, if the collocates possess strong negative semantic features, the node word is regarded to possess strong negative prosody in the case of *cause* that often collocates with words such as *accident*, *concern*, *damage*, *death*, and *trouble*. If the collocates are made up of words with positive features, the node word is then regarded to possess a positive prosody in the case of *provide* that often collocates with words such as *aid*, *assistance*, *care*, *employment facilities*, *food*, *funds*, *jobs*, *money* and other words that have no relation to negative things. Lastly, if the collocates are equally made up of words with positive and negative features, the node word is then regarded to possess a neutral prosody. However, this notion proposed by Stubbs (1996) does not define what is negative, neutral, and positive specifically, rather the classification provided only focuses on the node words' classification based on the negativity (neutrality and/or positivity) of the surrounding (collocates) rather than definition of the semantic aspect of negative, neutral, and positive. To start, the definition of negative, neutral, and positive will be based on the Oxford Dictionary 9th edition (2008). Negative means bad or harmful, neutral means not

supporting or helping either side in something or belonging to either, and positive means good in situation or feeling or good or useful. In this case, negative prosody is labelled to any instances that bring harm to somebody or something, neutral is labelled to any instances that support neither negative nor positive, and positive is labelled to any instances that bring good or being useful to something or somebody. The definition based on the dictionary corresponds to Partington's (2004) favourable, neutral, and unfavourable prosodies. Any favourable (good in situation) affective meaning was labelled as positive while unfavourable (bad or harmful) affective meaning was identified as negative, and neutral (not supporting either side) was labelled to any instances that showed either it contained both prosodies or it did not contain any prosodies.

To identify the semantic prosody, for each of the swear words identified, *fuck* for instance, concordance lines for *fuck* were generated and the lines were analysed manually to look for significant collocates that would bring any meanings mentioned earlier.

Hit	KWIC	File
1	I'll run in ya house, then I'll fuck your hoe 'Cause Remy Boyz or nothing, Re-Re-	4.Fetty Wap - Trap Queen.txt
2	't they tell you that I was a savage? Fuck your white horse and a carriage Bet you never	13.Rihanna - Needed Me.txt
3	't they tell you that I was a savage? Fuck your white horse and a carriage Bet you never	13.Rihanna - Needed Me.txt
4	eyes hurt Niggas dying back where I was birthed Fuck your iris and the IRS Get the hell up	22.Jay Z feat Justin Timberlake - Holy Grail.txt
5	go I mean for real, fuck how you feel Fuck your two cents if it ain't goin' towards	47.Big Sean feat. E-40 - I Don't Fuck With You.txt

Figure 3.4 Screenshot of concordance lines of *fuck*

Based on Figure 3.4, all the instances of *fuck* + NOUNS were identified as possessing negative semantic prosody because the instances were identified as negative (bad or harmful) because *fuck* + NOUNS is used to as an abrupt emphatic exclamation expressing emotion.

33	And we laying the club like we don't give a fuck You can't have my heart And	84.Lady Gaga - Do What U Want.txt
34	ompliments Overdosed on confidence Started not to give a fuck and stopped fearing the consequence Dr	85.Drake - Headlines.txt
35	the hood, tatted like a Mexican Car too fast, give a fuck about pedestrians (In my section less	86.Chris Brown - Ayo.txt

Figure 3.5 Screenshot of concordance lines of *fuck*

Figure 3.5 shows the instances of *fuck* that possess positive semantic prosody. The swear word *fuck* in line 35 is used together as a phrase *give a fuck* in which it is used to show care about someone or something (good in situation). However, when the phrase *give a fuck* is preceded with a negation *don't* for instance like 33 *don't give a fuck*, that particular instance has a negative semantic prosody.

3.8   Summary

This chapter has presented the methodology used to conduct the study. The chapter also describes how the corpus of English songs was built. Issues involved in constructing the corpus were also clarified in this chapter. This chapter also describes few matters that should be considered in constructing a corpus. I have also delineated how the data were collected by using AntConc 3.4.4w. The corpus software helps in generating word lists and concordance lines to make it possible for qualitative analysis and making sense of the data. Any ethical issues were also clarified in this section. The final section of this chapter describes in detail on how the data were analysed in order to answer the research questions.

CHAPTER 4: DATA ANALYSIS

4.1 Introduction

This chapter presents the data from the corpus of English songs. Section 4.1 presents the frequent swear words found in the corpus, followed by Section 4.2 which looks at the sub-corpora to identify the frequent swear words based on the respective years. Section 4.3 identifies the semantic prosody of the swear words found in the corpus and sub-corpora together with a bird eye view of the semantic prosodies across the years.

4.2 The Frequent Swear Words found in CoES

This section presents the frequent swear words found in a corpus of English songs. Table 4.1 shows the frequency list of the most frequent swear words with its' frequency and rank order.

Table 4.1 Swear words found in CoES

Rank	Frequency	Word
94	387	nigga
123	301	fuck
140	260	bitch
146	252	shit
159	244	niggas
213	148	ass
244	124	bitches
288	102	god
295	100	damn
332	85	fuckin
349	80	fucking
418	65	hell
517	47	pussy
555	43	hoe
560	43	stupid
578	41	hoes
582	41	sex
611	38	dick
636	36	fucked

869	23	motherfucker
965	20	hella
1,070	17	motherfucking
1,434	11	motherfuckin
1,847	7	cock

It is interesting to note that the word *nigga* tops the list of swear words identified in the corpus, ranked 94 in the corpus. This is inconsistent with the study by McEnery (2006) on swear words, where *nigga* is not a swear word that is preferred by males and females in the BNC's top 15 swear words. In addition, Bergen (2016) states that the swear word *nigger* (spelling used by Bergen) is in the top 3 of totally unacceptable words in New Zealand. However, it could be a sign that society has changed in a sense that they are more tolerable towards swear words especially the famous racial slur, *nigga*. Even in the US, *nigga* is now thought as a positive term for self-identification and in many cases, endearment between the blacks (Hoffman, 2009). In the case of songs based on CoES, *nigga* can be found mostly in hip hop genre for instance a song entitle My Hitta by YG (Kenny, 2011). In another study mentioned by Bergen (2016), the swear word *nigger* was rated 7 out of 7 in the scale of offensiveness of American English words where 1 is considered as the least offensive while 7 is considered as the most offensive. However, in CoES, the most frequent swear word found is *nigga* despite being rated as most offensive.

Moving to the least frequent swear word found in CoES for example *cock* ranked 1847, according to Bergen (2016), the swear word is considered acceptable/fairly acceptable, similar to *fuck*. In a small scale study by Bergen (2016) on the offensiveness of American English words, though *nigger* is absent in the small scale study, *fuck* is rated 6 out of 7 and *cock* on the other hand is rated 5 out of 7. In a bigger scale study, *cock* is rated 3 out of 7. At the early stage of analysing the swear words, it was hypothesised that the more

offensive the swear word is, the least it will be used. However, this present study proved differently that *nigga* though it was rated as totally unacceptable and as most offensive, the usage in the corpus of English songs top the list of the swear words.

4.3 Frequent Swear Words Found in the Sub-corpora of CoES

This section looks at the frequent swear words found in the sub-corpora of CoES. Table 4.2 below shows the top 10 swear words found in the sub-corpora according to year (2011 to 2016) based on the criteria mentioned in Section 3.7. The swear words were identified by using the HFSN notion by Bergen (2016).

Table 4.2 The swear words in sub-corpora of CoES

2011			2012		
Rank	Frequency	Word	Frequency	Rank	Word
141	51	fuck	118	46	shit
164	43	niggas	121	45	bitch
257	24	bitch	125	43	ass
260	24	fuckin	166	35	niggas
263	24	nigga	182	32	nigga
289	21	god	247	22	hoe
303	20	hell	253	22	stupid
314	19	damn	259	21	fuck
478	10	bitches	354	14	damn
484	10	dick	358	14	god
2013			**2014**		
Rank	Frequency	Word	Frequency	Rank	Word
115	53	nigga	46	168	nigga
140	41	fuck	122	51	niggas
142	41	shit	128	46	fuck
165	34	bitches	129	46	shit
224	32	bitch	131	45	bitch
273	19	fucking	178	30	bitches
275	19	god	222	24	fucking
288	18	ass	256	20	ass
347	14	hell	293	17	damn
392	12	demons	313	16	god
2015			**2016**		
Rank	Frequency	Word	Frequency	Rank	Word
88	77	fuck	104	65	fuck
89	75	bitch	112	56	nigga
116	54	nigga	136	47	bitch

147	42	niggas	141	46	shit
157	38	ass	205	29	niggas
159	38	shit	280	19	fucking
215	26	bitches	281	19	god
232	24	fuckin	292	18	fuckin
253	21	damn	353	14	damn
271	19	fucked	423	11	pussy

Based on Table 4.2, the frequency distribution of the swear words is uneven in the corpus. However, it is clear that there are 3 common swear words that occur in different ranks across the years namely *fuck*, *nigga*, and *bitch*. Looking at the frequencies of the swear words, only a handful of swear words were analysed based on the criteria mentioned. In 2014, the swear word *nigga* had the highest frequency among the other swear words. One of the contributing factors is because the nature of lyrics of songs that tend to have a lot of repetition. This swear word was found in 16 songs and it occurs 97 times in a song entitled My Hitta by YG. It also indicates that the word is more popularly used in songs based on its' frequent usage throughout 2011 to 2016.

The swear word *fuck* (77) has the highest frequency in the 2015 chart. This is similar to *nigga* in way that it occurs 25 times in a song entitled *I Don't Give A Fuck With You* by Big Sean. This is also a sign that the swear word *fuck* is more popular to be used in songs because apart from the song by Big Sean, *fuck* was also found in another fifteen songs (excluding Big Sean) in 2015.

The swear word *bitch* (75) also has the highest frequency in 2015 that is a rank below *fuck*. The swear word *bitch* was found mostly in a song similar to *fuck* that is a song entitled *I Don't Give A Fuck With You* by Big Sean. The swear word *bitch* occurs 21 times in the song. Apart from that, *bitch* also occurs 33 times in a song entitled *Bitch Better Have My Money* by Rihanna. This is also similar to *nigga* due to the nature of songs which has a lot of repetition.

63

Apart from looking at the frequency, the rank clearly shows how frequent the swear words are being used in the sub-corpora. Most of the swear words are found at the 100th rank and above. However, *nigga* in sub-corpus 2014 is found at the 64th rank that is significantly higher than other swear words. If we are to look at the range of the ranks, swear words in sub-corpus 2015 rank higher compared to other sub-corpora. It can be concluded that the songs in sub-corpus 2015 use more swear words compared to the other sub-corpora.

4.4   Semantic Prosody of the Swear Words

This section analyses the semantic prosody of the swear words identified by using the HFSN notion highlighted by Bergen (2016). Prior to the analysis, the context of CoES was identified following Rahim (2012) in her study of semantic prosody on "female" and "woman". By generating the word lists, the top 10 words were used to establish the context (see Katznelson *et al*, 2010 and Falk, 2012). Table 4.3 below shows the top 10 nouns found in the corpus of English songs.

Table 4.3 The top 10 nouns in CoES

Top 10 nouns	Frequency
love	1,555
baby	1,212
girl	670
night	621
time	600
life	463
nigga	387
money	354
heart	351
man	314

Comparing the top 10 nouns found in CoES with what Katznelson *et al* (2010) found, 7 out of 10 nouns are similar namely *love*, *baby*, *girl*, *heart*, *time*, *life*, and *man* (in

descending order). Based on the words, the songs reflect *carefree* love or *carefree* existence (Katznelson *et al*, 2010). Oxford Dictionary 9th edition (2008) defines *carefree* as having no worries or responsibility in which it is a neutral word. Based on that, the context of CoES is a neutral setting where it supports the idea that this is the way life is or the way life should be (Katznelson *et al*, 2010). Once the context was established, by looking at the concordance lines, the semantic prosody of the swear words was identified (see Partington, 2004, Hunston, 2007, Yusuf, 2010, Begagic, 2013, Rahim 2012, and Hashim and Rahim 2016). Any semantic prosody that conveys the notion against the idea of how life should not be is regarded as negative; if it supports the idea of how life should be, then it is regarded as positive. Lastly, any semantic prosodies that supported neither, is regarded as neutral in the context of the corpus of English songs.

4.4.1 NIGGA

The swear word NIGGA has the highest frequency compared to the other swear words found in the corpus. By definition, NIGGA is a taboo slang that is a very offensive word for a black person (Oxford Dictionary 9th edition, 2008). Focusing on the semantic prosody of NIGGA, the swear word generally possesses negative semantic prosody. This swear word is negative because it is mostly used in a way that it *brings harm to others*, *owns weapons*, and *womanizer* for *nigga*, and a *faker*, and a *thief* for *niggas* (See Table 4.4). However, *nigga* also possesses neutral semantic prosody when it is used as a self-identification that is a *friend*.

Table 4.4 The semantic prosody of NIGGA

Swear Words	Concordance Lines	Collocates	Semantic Prosody	File
Nigga	igga bigger than gorilla 'Cause I'm killing every **nigga** that try to be on my shit Better cuff	-	A person who does harm to others	File21_2011

	Living the life Vanilla wafers In a villa *Illest* **nigga** alive Michael Jackson's, Thriller			File22_2013
	Yeah they stereotyping 'Cause they know a **nigga** keep ten rifles And they know a nigga keep	-	Owns weapons	File30_2016
	a nigga keep twenty snipers And they know a **nigga** keep ten wifeys Sniper, sniper, sniper,	-	Womanizer	File30_2016
	nigga Fucked my first bitch passed her to my **nigga** Hit my first lick pass with my nigga Fuck	-	A slang for friend	File58_2014
	ain't going in, unless I'm with my nigga My **nigga**, my nigga (dope boys, my niggas) My			File58_2014
	fronted me, he the reason I'm straight My **nigga** got a hoe, my nigga got a lady My			File58_2014
Niggas	two days (Guess what!) I don't fuck with **niggas** 'cause they two fake (I swear I swear) I	-	Faker	File49_2015
	rolling up so I get love for that Them **niggas** stole my swag but I don't want it	-	Thief	File73_2012
	was on 14 I stole his keys Me and my **niggas** was gone Stealin' portions of his liquor,			File92_2014

In the sub-corpora, not much can be extracted about the swear words due to the less number of occurrence. However, faint patterns can be seen from the swear words. To begin with, NIGGA is one of the three swear words that does not have fixed collocations and the semantic prosody is identified by looking at how the swear words behave that is by looking at the attitudinal meaning (Partington, 2004). Moreover, this makes the semantic prosody of NIGGA not be identified for the following years; 2011, 2014 (only *niggas*), and 2015. Although it has the highest frequency throughout CoES, NIGGA does

not fulfill the requirement in terms of frequency of occurrences in the sub-corpora because the distribution of it is uneven.

NIGGA, however, still possesses the same semantic prosodies identified earlier. The semantic prosodies for each year are mostly different from one another except for the 2014 and 2016 sub-corpora. In the 2012 sub-corpora, *nigga* has the sense of inflicting pain to others in which this was classified as negative. In 2014 sub-corpora, there are two semantic prosodies of NIGGA identified that are negative and neutral; own weapons (File30_2016) and a slang for a friend (File58_2014). Still, similar negative semantic prosody is identified in 2016 for NIGGA that was owning weapons. Compared to the semantic prosodies identified by using the English corpus of English songs, some of the semantic prosodies identified are not found due to the cut-off point of the swear words to be analysed.

Although there are over 200 concordance lines of NIGGA in CoES, only selected few were chosen because most of the concordance lines did not portray any expression (Smith and Nordquist, 2012). These concordance lines, rather than expressing the instances in Table 4.4, NIGGA is just a slang that is a very offensive word for a black person. Based on the analysis, NIGGA possesses negative semantic prosody as a black person that does harm to others (File21_2011), owns guns (File30_2016)), a womanizer (File30_2016), a fake (File58_2014)), and a thief (File73_2012). Generally, the term itself is universally viewed as pejorative and historically violent that NIGGA possesses a sense that it is used often times coincided with violence towards blacks males (Fogle, 2013). However, NIGGA shows a neutral semantic prosody when it is used to mean *friend* (File58_2014) in which according to McEnery (2006), swear words used by males mostly are not linked to abuse rather, they are associated with intensification and in this case, it is used as some

sort of manly endearment. This is also supported by Fogle (2013) that it is indeed a positive term for self-identification – endearment between the blacks.

4.4.2 FUCK

According to Oxford Dictionary 9th edition (2008), *fuck* can be either a verb, a noun, and can be used to form idioms. As a verb, *fuck* is defined as to have sex with someone, or a swear word that most people find offensive that is used to express anger, disgust, or surprise. As a noun, *fuck* is an act of intercourse. If it collocates with *the* as in *the fuck*, it is used to emphasize or to show someone is annoyed or surprised. On the other hand, *fucking* is an adjective or an adverb that is considered a taboo slang that most people find offensive and it is used to emphasize a comment or an angry statement. In CoES, FUCK possesses all three semantic prosodies; negative, neutral, and positive.

Table 4.5 The semantic prosody of FUCK

Swear Words	Concordance Lines	Collocates	Semantic Prosody	File
fuck	in my pocket Wasn't enough, I'm like **Fuck** you And forget her too Said, if I was	*fuck* + NOUNS	An abrupt emphatic exclamation expressing emotion	File7_2011
	ration black and white diamonds, **fuck** segregation fuck that shit, my money up, you			File41_2011
	't they tell you that I was a savage? **Fuck** your white horse and a carriage Bet you never			File13_2016
	got sober, felt 10 years older But **fuck** it, it was something to do I'm living			File15_2016
	cool but I may just go ape shit Say " **fuck** y'all" to all of y'all faces It			File19_2016
	Chandelier swinging, we don't give a **fuck** Film star, yeah I'm deluxe Classic, expensive	don't give a *fuck*	To not care about	File4_2014

	done And I don\x92t really give a **fuck** , and my excuse is that I\x92m young		someone or something	File47_2011
	with us ('scuse me, 'scuse me) She gonna **fuck** the squad, what else? I'm a fuck her	gonna/wanna + *fuck*	To treat someone or something badly	File30_2016
	wanna see a nigga trap She wanna **fuck** all the rappers Black girl with a big booty	*fuck* + with		File30_2014
	Stand ofr it And Nicki if you ever try to **fuck** Just give me the heads up so I	to + *fuck*		File51_2015
	they don't wanna party, Tell them shut the **fuck** up So DJ turn it loud and watch me	shut the *fuck* up	To shut someone up (harsh)	File49_2011
	's due I'mma worry about me, give a **fuck** about you Nigga, just as a reminder to myself		To care about someone or something	File32_2013
	me (Yeah, yeah) (Yeah, yeah) I can give a **fuck** 'bout no hater Long as my bitches love me	give a *fuck*		File39_2013
	tatted like a Mexican Car too fast, give a **fuck** about pedestrians (In my section less niggas,			File86_2015
	100,000 just in two days (Guess what!) I don't **fuck** with niggas 'cause they two fake (I		To not treat someone or something badly	File49_2015
	who I fuck with', mad bitches we don't **fuck** with' I don't fuck with' them chickens unless	negation + *fuck* with		File51_2015
	just said She's singing Ah lahmlahlah Don't **fuck** with my love That heart is so cold All			File52_2014
fuckin	And you know I'm rollin' weed that's **fuckin'** up the ozone I got a bitch that text	*fuckin* + up	To cause problem to someone or something	File47_2015
	ever, ever feel like you're nothing You're **fuckin'** perfect to me You're so mean when you			File19_2011
	much 'Cause I never knew love would hurt this **fuckin'** bad The worst pain that I ever had	*fuckin* + ADJECTIVES	Intensifier	File77_2012
	shit 0 to 100, nigga, real quick Real quick, real **fuckin'** quick, nigga 0 to 100, nigga, real			File97_2014
	pain, it ain't nothing but pain You just **fuckin'** complain, you ain't tough as you claim Jus	*fuckin* + NOUNS	Emphasize	File68_2016

	I\x92m your leader, you're supposed to **fuckin'** be my mentor I can endure no more I	*fuckin* + VERBS		File51_2011
	gon' either wanna fight when I get off this **fuckin'** mic Or you gon' hug me, but I'm			File51_2011
fucking	trend it Jeff The Don doing business Zana Ray **fucking** up shit and she doin' her business I	*fucking* + up	To cause problem to someone or something	File6_2016
	But I don't think none of these bitches **fucking** with me Want a billboard bitch stop runnin	*fucking* + with	To treat someone or something badly	File87_2014
	this interview It ain't nothing new I been **fucking** with you None of them bitches ain't taking			File31_2015
	city Travel to another town, you can bet they **fucking** with me I be killing this shit Pray to			File50_2015
	again You\x92re beautiful, and your mind is **fucking** beautiful And I can\x92t pretend that	*fucking* + ADJECTIVES	Intensifier	File81_2013
	'm ducking bucking them out here I'm looking **fucking** fantastic, I am up in a classic Now I			File92_2014
	to get to you, baby But this song so **fucking** dope, girl, it's hard not for me to			File96_2014
	'm hunting, looking for a come-up This is **fucking** awesome Nah, walk up to the club like	*fucking* + NOUNS	Emphasize	File1_2013
	played him I ain't here to save the **fucking** children But if one kid out of a hundred	*fucking* + VERBS		File16_2014
	Started from the bottom now my whole team **fucking** here Started from the bottom now			File32_2013

Based on Table 4.5, the swear word FUCK basically possesses negative semantic prosody when it is used as an abrupt emphatic exclamation to express emotion, to show lack of care to someone or something, to treat someone or something badly, and to cause problem. Although *fuck* does not possess neutral semantic prosody, both *fuckin* and

fucking possess neutral semantic prosody used as an intensifier and to emphasize a subject. Table 4.5 also reveals that both *fucking* and *fuckin* share the same semantic prosody; and collocates, that these two swear words are used in the same manner despite having different spelling. On the other hand, *fuck* (File32_2013, File29_2013, and File86_2015) possesses positive semantic prosody that shows care to someone or something and to treat someone or something good.

In the sub-corpora, only the swear word *fuck* met all the criteria mentioned. Hence, both *fucking* and *fuckin* were not included due to the lack of number of occurrences. This is because the occurrence of both *fucking* and *fuckin* was less than 30 occurrences. However, it was lucid that throughout the years *fuck* mostly possesses negative semantic prosody from 2011 to 2016. The swear word was usually used as to treat someone badly (File87_2014), to not care about someone (File32_2013), and an abrupt emphatic exclamation expressing emotion (File7_2011). Notwithstanding, *fuck* showed a positive semantic prosody in File86_ 2015 that it was used to care about someone.

FUCK in the first place can function in a lot of grammatical categories (Pujol, 2006). In the current study, FUCK is used as an intensifier (File77_2012), to emphasize (File68_2016), and phrasal verbs to state a few. It possesses negative semantic prosody and used to express emotions; anger (Pujol, 2006), to treat someone badly (Stenstrom, 2006), and it has the tendency to inflict emotional pain as well as to provoke violent disagreement (Bergen, 2016) – hence negative. On the other hand, it possesses positive/neutral semantic prosody because FUCK is used without the intention of insulting (Stenstrom, 2006).

4.4.3   SHIT

Oxford Dictionary 9th edition (2008) defined *shit* as a swear word that most people find it offensive that is used as exclamation to show that one is angry or annoyed. As a noun, *shit* has multiple meanings that it refers to solid waste matter from bowels or excrement, an act of emptying waste matter, stupid remarks, an unpleasant person that is treating others badly, and an unfair treatment. As a verb, *shit* is more to the action of emptying bowels and is more neutral compared to *shit* as a noun. Oxford Dictionary 9th edition (2008) also defined *shit* that means very bad as an adjectives. An example of usage of adjective *shit* is *They're a shit team*. Generally, *shit* possesses negative base meaning.

Table 4.6 The semantic prosody of SHIT

Swear Words	Concordance Lines	Collocates	Semantic Prosody	File
shit	a big cock!" I'm so pumped about some **shit** from the thrift shop Ice on the fringe	-	A slang for drugs	File1_2013
	all over you Never mind, we only poppin' **shit** Man I been getting high with these fools And			File90_2016
	all good I'm still sippin' this bubbly This **shit** is lovely, this shit ain't random, I didn'			File19_2016
	like this All those fairy tales are full of **shit** One more stupid love song, I'll be sick	full of *shit*	Someone or something is useless	File4_2012
	you can go and take that little piece of **shit** with you I'm at a payphone	piece of *shit*		File4_2012
	world on the floor You know we're running **shit** tonight on the floor Brazil, Morocco,	running *shit*	Doing something bad	File11_2011
	The Don doing business Zana Ray fucking up **shit** and she doin' her business I be	*fucking* up *shit*		File6_2016
	I'm a kid my ego is big I don't give a **shit** and it goes like this Take me by the	don't give a *shit*	A statement of not acknowledging something	File9_2011
	in the kitchen saying, "How the hell did this **shit** happen?"	-	A trouble or a difficulty	File35_2014

	Oh baby, drunk in love we be all			
	it can't get worse, I've had a **shit** day (NO!) Have you had a shit day? (NO!)			File37_2012
	had a shit day (NO!) Have you had a **shit** day? (NO!), we've had a shit day (NO!)	-	Bad day	File37_2012
	had a shit day? (NO!), we've had a **shit** day (NO!) I think that life's too short			File37_2012

The swear word SHIT is interesting to look at because in the corpus of English songs, none of the instances found literally mean *shit*; faeces. However, based on Table 4.6, *shit* was mostly found to possess negative semantic prosody. The collocations that SHIT has revealed that the swear word is usually used as a slang to refer to drugs (File1_2013), someone or something is useless (File4_2012), doing something bad (File11_2011), and a statement to not acknowledging something (File9_2011). On the other hand, SHIT also possesses neutral semantic prosody where the usage shows a trouble or a difficulty (File35_2014), and another way of saying a bad day (File37_2012).

Looking at SHIT throughout the years beginning from 2012, the swear word is used to show someone or something is useless, as an exclamation to express emotion, and frequently used as a reference to drugs. Even so, neutral semantic prosody was also found in the sub-corpus 2012 that SHIT is used to mean bad day and in sub-corpora 2014 and 2015 that SHIT is used to represent a trouble or a difficulty.

The words for faeces are typically taboo and the dirtiest among the dirty words in English is *shit* (Andersson and Trudgill, 1990). Simpson (2004) adds that it is also not appropriate to be used. *Shit* is something that has little or no value as in rubbish or trash (Ayto and Simpson, 1996). Hence, comparing a man to a *shit* is degrading him. Also, *shit* refers to drugs such as heroin, cannabis or marijuana (Ayto and Simpson, 1996). Apart

from making reference to the very worst, *shit* is also used to make reference to something good (Fernandez Dobao, 2006). In this case, it is something neutral that *shit* only refers to a troubled situation and a bad day – no harm was done.

4.4.4   ASS

As defined by Oxford Dictionary 9th edition (2008), *ass* (British English *arse*) is a taboo or slang that refers to part of the body that you sit on or your bottom. In British English, *ass* means a stupid person. Additionally, *ass* can be used as idiom for instance *get your ass in gear* or *move your ass* that is a rude way of telling someone to hurry and *get your ass over/in here* that is a rude way of telling someone to come here.

Table 4.7 The semantic prosody of ASS

Swear Words	Concordance Lines	Collocates	Semantic Prosody	File
ass	niggas Pose for your class picture Now kiss my **ass** if you hate I'm getting ass so I'	Kiss my *ass*	Verbal abuse	File39_2013
	well, now wish me well tell 'em kiss my **ass** , call it kiss and tell Word to my mama			File41_2011
	the fuck we been? You can kiss my indecisive **ass** crack, maggots And the crackers' ass, little			File51_2011
	I don't fuck with you You little stupid **ass** bitch, I ain't fuckin' with you You little	Stupid/dumb *ass*		File47_2015
	't fuckin' with you You little, you little dumb **ass** bitch, I ain't fuckin' with you I got			File47_2015
	All Saints for my angel Alexander Wang too **Ass** -tight Denim and some Dunks I'll show you	*Ass*-tight	Skin-tight	File20_2013
	's a stick up, no more makeup Get that **ass** on the floor Ladies put your lipstick up Doubl	PRONOUNS + *ass*	Colloquial universal pronouns	File32_2015
	know, where you re going Just get your **ass** back home No one hold me down			File51_2012

let me play with his rifle Pussy put his **ass** to sleep, now he calling me NyQuil Now that			File36_2014
be late about a week ago I worked my **ass** off, but I still can't pay it though			File39_2011
you The NBA players fuck with you The bad **ass** bitches doing makeup and hair fuck with	bad *ass*	Intensifier	File51_2015

Similar to SHIT, none of the instances found in the corpus of English songs mean literally buttock. The swear word *ass*, however, possesses both negative and neutral semantic prosody. Based on Table 4.7, it is negative when *ass* is used to downgrade someone in verbal abuse (File41_2011). It also collocates with other swear words for example *dumb* and *stupid* that actually the source of the verbal abuse. This is because, in this semantic prosody, *ass* acts as a replacement for nouns the means *someone* to be specific. This is due to the fact that *ass* can be used as pronouns as seen in the colloquial usage of *ass*. On the other hand, *ass* is neutral when it is used as another way to say a skin-tight clothing (File20_2013), and it is used a colloquial universal pronouns (File32_2015) (JSTOR Daily, 2019).

In the sub-corpora, *ass* can only be analysed in 2012 and 2015. It is apparent that *ass* is normally used as a neutral colloquial universal pronouns in both sub-corpora and in 2015 sub-corpus it shows negative semantic prosody that shows undesirable quality.

In conclusion, apart from showing size; skin-tight clothing, the swear word *ass* also has other meaning. JSTOR Daily (2019) stated that the construction of *ass* with another pronouns is reckoned as colloquial American English. Other than that, *ass* can also be used to replace intensifiers such as *very* and *really* in which the function is similar (JSTOR Daily, 2019).

4.4.5 BITCH

BITCH as a noun refers to a female dog, an offensive way of referring to a woman, a slang for a thing that causes problems or difficulties, or a complaint about someone or something (Oxford Dictionary 9th edition, 2008). As a verb, it means to make unkind and critical remarks about someone or something especially when someone or something is absent.

Table 4.8 The semantic prosody of BITCH

Swear Words	Concordance Lines	Collocates	Semantic Prosody	File
bitch	in physics (right, right) You should want a bad **bitch** like this (huh?) Drop it low and pick it	ADJECTIVES + *bitch*	Having undesirable or negative quality	File4_2014
	all in the car, talking 'bout you the baddest **bitch** thus far Talking 'bout you be repping that			File35_2014
	for a T-shirt - that's just some ignorant **bitch** (shit) I call that getting swindled and pimp			File1_2013
	ride that dick Man I wouldn't trust that **bitch** No! Come on, come on, girl Why you	-		File30_2014
	my religion Got my baby mamma and my side **bitch** kissing I turn the Ritz into a lean house	side *bitch*	Mistress	File30_2016
	bitch out of your league too, ah Side **bitch** out of your league too, ah House so empty			File58_2016
	comes around like a hula hoop Karma is a **bitch** ? Well just make sure that bitch is beautiful	-	Unkind remarks	File95_2011
	your unconscious Feel me when you get a fine **bitch** Just don't forget to read the fine print	fine *bitch*	Good bitch (ironic)	File47_2015
bitches	lost a few good bitches Met some more bad		Having undesirable	File39_2013

	bitches And I be schooling them niggas Pose	ADJECTIVES + *bitches*	or negative quality	
	The NBA players fuck with you The bad ass **bitches** doing makeup and hair fuck with you			File51_2015
	entertainers Now I got tattoos on my body Psycho **bitches** in my lobby I got haters in the	psycho *bitches*		File22_2013
	that good kush and alcohol I got some down **bitches** I can call I don't know what I	down *bitches*	Up for anything or something bad	File39_2013

The corpus of English songs has revealed that BITCH mostly possesses negative semantic prosody. Similar to NIGGA, BITCH does not have fixed collocation. Based on Table 4.8, BITCH is negative when it collocates with negative adjectives such as *bad* and *baddest* (File35_2014) to show the negative quality of BITCH. The swear word is also used negatively to mean a mistress (File30_2016), up for anything bad (File39_2013), and as unkind remarks (File95_2011). However, an ironic instance of *bitch* was found that collocates with positive adjective; *fine* (File47_2015), that makes *bitch* to have neutral quality.

Starting from sub-corpus 2012 to sub-corpus 2016, BITCH shows that it possesses negative semantic prosody usually having negative quality, up for anything bad in sub-corpus 2013, and as mistress in sub-corpus 2016. The neutral semantic prosody was only found in sub-corpus 2015 that is ironic.

In a study by Kleinman, Ezzell, and Frost (2009), they have described the harms of BITCH. They conclude that the usage of BITCH is negative especially when it is used to refer to humans; woman, that it is an insult. Additionally, its' usage carries a connotation of sexual libel that is a *bitch* is accused of being worse than a prostitute (Kleinman *et al*,

2009). Bergen (2016) also reiterates that BITCH usually denotes an aggressive or an unpleasant person. In this study, however, BITCH portrays a neutral semantic prosody but it is ironic when it is paired with adjectives like *good* and *fine*.

4.4.6   GOD

Oxford Dictionary 9th edition (2008) defined *god* as a being that is worshipped and is believed to have created the universe, a spirit who is believe to possess power over a particular part of the nature, a person who is admired or loved by other people, something to which too much importance or attention is given, and informally in British English, it refers to the seats that are high up at the back of a theatre. This swear word basically possesses neutral meaning because of the religious sense imbued with it.

Table 4.9 The semantic prosody of GOD

Swear Words	Concordance Lines	Collocates	Semantic Prosody	File
god	rumbling Castles crumbling I am trying to hold on **God** knows that I tried Seeing the bright	swear to *god*	Religious remarks	File15_2012
	feels me, till' it kills me I swear to **God** I'll be the fucking illest in this music	*god* almighty		File34_2011
	ld Throwin' little cherry bombs into my fire Good **god** almighty Girl, you make my speakers	dear *god*		File53_2012
	she gon' let it burn, baby, burn, baby Dear **God** , if you're here God Make the fire disappear	pray to *god*		File49_2013
	with me I be killing this shit Pray to **God** they forgive me They said, "When you blow up			File50_2015
	other tongue Skeet skeet skeet: water gun Oh my **God** , Becky, look at her butt! Tunechi!	oh my *god*	An abrupt emphatic exclamation expressing emotion	File20_2012
	Man now we stepping out like, "Whoa" (Oh **God**) Cameras point and shoot (shoot) Ask me			File34_2012

	But look what that shit did to Hammer (uh) **God** damn it I like it Bright lights is enticing	*god* damn (it)		File22_2013
	packed It's too turned going up like gas **God** damn pull out my racks Mike Mike Jackson			File45_2012

Based on Table 4.9, GOD only possesses two neutral semantic prosody. The swear word *god* usually used as religious remarks (File15_2012) and exclamation to express emotion (File20_2012) based on the collocation. Although *god* could not be analysed throughout the sub-corpora because of the low number of occurrences; less than 30 occurrences, it is clear that the instances can be found in each sub-corpora.

This swear word possesses neutral semantic prosody because of the religious sense it bears since it is related to religion (Amrullah, 2016) and it was categorised as religious oath by McEnery (2006a). This word is also used to express emotion that is evoked with surprise or astonishment (Amrullah, 2016). Above all, *god* possesses neutral semantic prosody because it usage shows that it neither brings harm nor good to the user that it has lost the sense of taboo (Stenstrom, 2006).

4.4.7 DAMN

According to Oxford Dictionary 9th edition (2008), *damn* is used as an exclamation informally to show that one is annoyed or disappointed. As an adjective and adverb, *damn* is a swear word that people use to show the annoyance towards someone or something. This swear word is also used by people to emphasize what they are saying. As a verb, *damn* is informally used when swearing at someone or something to show anger. It is also used in phrase for instance *damn someone* (of god) to decide that someone must suffer in hell and *damn someone/something* to show criticism very strongly.

Table 4.10 The semantic prosody of DAMN

Swear Words	Concordance Lines	Collocates	Semantic Prosody	File
damn	I'm so pretty I'm too hot hot **damn** Call the po-lice and the fireman I'm	hot *damn*	An abrupt emphatic exclamation expressing emotion	File1_2015
	lice and the fireman I'm too hot hot **damn** Make a dragon wanna retire, man I'm too			File1_2015
	dragon wanna retire, man I'm too hot hot **damn** Say my name you know who I am I'			File1_2015
	But look what that shit did to Hammer (uh) God **damn** it I like it Bright lights is enticing	god *damn*		File22_2013
	packed It's too turned going up like gas God **damn** pull out my racks Mike Mike Jackson			File45_2012
	'm loving you, oh you know You're so **damn** pretty If I had a type then baby it	*damn* + ADJECTIVES	Intensifier	File16_2011
	and slamming the door You and I get so **damn** dysfunctional, we start keeping score			File18_2012
	but I am up in space You're so **damn** hard to please, we gotta kill this switch You			File28_2013
	t gonna solve it all But it's a **damn** good place to start No law is gonna change			File43_2013
	Of everything I got You\x92re looking so **damn** hot And I don\x92t know what road			File93_2012
	line The irony I fuck 'em at the same **damn** time She eyeing me like a nigga don't	*damn* + NOUNS	Emphasize	File41_2013
	's only one flo, and rida I'm a **damn** shame Order more champagne, pull it down hell			File17_2012
	care you never did You don't give a **damn** about me Yeah all alone I watch you watch	don't give a *damn*	To not care	File38_2016
	of my lil friends They don't give a **damn** how a nigga been All they wanna know is			File57_2016

Based on Table 4.10, it is clear that DAMN only possesses neutral semantic prosody.

The swear word DAMN is used as an abrupt emphatic exclamation expression

(File1_2015), intensifier (File18_2012), to emphasize (File41_2013), and another way of saying to not care about someone or something (File38_2016). The abrupt emphatic exclamation expression, *damn* collocates with another swear word; *god* (refer to Table 4.10 File22_2013 and File45_2012). In this case, it shares the same religious remarks that *god* has due to its' usage which is usually concerned with sacred matters. Throughout sub-corpora 2011 to 2016, the swear word *damn* could not be identified because *damn* did not fulfil the requirement stated. However, from Table 4.10 it is visible that the neutral semantic prosody can be found in almost all the sub-corpora except sub-corpus 2014.

The analysis of the swear word *damn* revealed that the swear word mainly possessed neutral semantic prosody in the corpus of English songs. Neither of the instances in the corpus were found negative nor positive. The swear word was used to intensify adjectives and it was used to emphasize nouns (topic) (Karjalainen, 2002). The next expression was thought to be negative, however the word *care* (File57_2016).mentioned in the instances were found to be related to a person accepting another person as who they were. Hence, in this case it was identified as neutral. In addition, the instances found did not show that the swear word was used to directly insults other. According to Jay (2009), the act of directing insults towards others is negative.

4.4.8   HOE

Apart from being a tool with a flat blade attached at right angles to a long (usually) wooden handle, *hoe* is a short form of *whore* (Oxford Dictionary 9th edition, 2008). In the dictionary, *hoe* is spelled as *ho* or the plural is *hos* or *hoes*. Regardless, *hoe* (plural *hoes*) is a slang that is referred to a female as a prostitute and it is an offensive word used about a woman, especially one who you think has sex with a lot of men.

Table 4.11 The semantic prosody of HOE

Swear Words	Concordance Lines	Collocates	Semantic Prosody	File
hoe	I got it, ooh, ooh, ooh, ooh Give that **hoe** some X, she gon' wanna sex every nigga in	-	Activities associate with sexual intercourse	File49_2015
	run in ya house, then I'll fuck your **hoe** 'Cause Remy Boyz or nothing, Re-Re-Remy Boyz			File4_2015
	that ass, walk her out the club (Yeah, **hoe**) Lap dance for the first date Bet I threw			File55_2015
	m red leather, this cocaine, I'm Rick James **hoe** I'm bill droppin', Ms. Pacman is pill poppin'	-	Takes drugs	File27_2012
	bill droppin', Ms. Pacman is pill poppin' ass **hoe** I'm poppin' too, these blue dolphins need			File27_2012
hoes	And I done did everything but trust these **hoes** (CB fuck with me!) When a rich nigga want	-	Dishonest	File30_2014
	your nigga can't do nothing for ya These **hoes** ain't loyal These hoes ain't loyal Yeah	-	Unfaithful	File30_2014
	nothing for ya These hoes ain't loyal These **hoes** ain't loyal Yeah, yeah, let me see Just			File30_2014

HOE is another swear word that shares similar properties as NIGGA and BITCH in which these swear words do not possess fixed collocates. Nevertheless, the semantic prosody of HOE is identified by looking at how the swear word behaves in the corpus of English songs (Partington, 2004). Based on Table 4.11, HOE only possesses negative semantic prosody. The swear word is usually associated with sexual activities (File49_2015), takes drugs (File27_2012), being dishonest (File30_2914), and unfaithful (File30_2014). Since the HOE occurs less than 30 occurrences in the sub-corpora, the semantic prosody of HOE could not be identified in the sub-corpora throughout 2011 to 2016. However, the concordance lines in Table 4.11 shows that HOE was found in sub-corpora 2012, 2014, and 2015.

In a study by Kleinman *et al* (2009), they stated that the usage of HOE is similar to BITCH that is as an insult and a woman that is considered to be worse than a prostitute. It is indeed reflected in this study that the senses mentioned by Kleinman *et al* (2009) are similar especially HOE in the singular form – *hoe*, that both BITCH and *hoe* are always engage in sexual intercourse as prostitute.

4.4.9   HELL

Based on the definition provided by Oxford Dictionary 9th edition (2008), *hell*, in some religions is believed to be the home of devils and where bad people go after death. This swear word also refers to an unpleasant experience or situation. Also, it is a swear word that is used when one is annoyed or surprised or to emphasize something.

Table 4.12 The semantic prosody of HELL

Swear Words	Concordance Lines	Collocates	Semantic Prosody	File
hell	and man tell your friends We'll raise some **hell** where the black top ends Yeah I'm chillin'	-	Referring to violence and needless disturbance	File43_2011
	x91em. Hell, we can all go raise some **hell** on any other night Girl, I don\x92t			File43_2011
	It's like I'm always causing problems, causing **hell** I didn't mean to put you through this			File80_2016
	good but now Ooohhh I'm thinking "What the **hell** ?" All I want is to mess around, and I	-	Enhancement of WH-words	File62_2011
	back And make me forget my name What the **hell** you do I won't remember I'll be			File82_2013
	doubt it! And I\x92m like, what the **hell** they talkin\x92 \x91bout? And if I got			File92_2012
	I've been good but now whoa "What the **hell** ?" La la la la la la la la whoa			File62_2011
	We woke up in the kitchen saying, "How the **hell** did this shit happen?" Oh baby, drunk in			File35_2014

	Wayne ain't got nothing on my fringe game, **hell** no I could take some Pro Wings, make			File1_2013
	not ready to go home Can I get a " **Hell** , no! "? (Hell no) 'Cause we're gonna go all	-	Emphasize	File17_2013
	ain't sorry I ain't sorry No no **hell** nah Now you want to say you're sorry			File71_2016
	ain't sorry I ain't sorry No no **hell** nah Looking at my watch he should've been			File71_2016
	Know I'm great but I'm broke as **hell** Having dreams that I'm folding cake All my			File19_2016
	'll be waking up feeling satisfied but guilty as **hell** But baby there you go again, there you go	*hell* + ADJECTIVES	Intensifier	File18_2012
	't even make it out And he dumb as **hell** and I swear his ass don't think, ayy			File44_2015
	my partner Roscoe like bruh I'm drunk as **hell** Can't you tell, goose we been spent fifty			File45_2011
	Though I've never been through **hell** like that I've closed enough windows to were innocent			File76_2013
	Took this heart and put it through **hell** But still you're magnificent I, I'm a	-	Difficulty and suffering	File60_2012
	makes my spirit shake I had to go through **hell** to prove I'm not insane Had to meet			File93_2015
	can run the mile You can walk straight through **hell** with a smile You can be the hero You			File85_2013
	more bottles cause you know it don't stop (808) **Hell** Yeah Drink it up, drink-drink it up,	-	An abrupt emphatic exclamation expressing emotion	File72_2011

Table 4.12 shows the semantic prosody of HELL. The swear word *hell* possesses negative and neutral semantic prosody. It is negative because the instances found show that *hell* refers to violence and needless disturbance (File43_2011) and it is used to enhance the WH-words (File82_2013). The swear word *hell* can also be used to

emphasize (File17_2013) and to intensify (File18_2012). In addition, it is used to show difficulty or suffering (File60_2012) and to express abrupt emphatic exclamation to express emotion (File72_2011). Since the swear word does not meet the criteria to be included in the analysis, the semantic prosody of *hell* could not be identified in the sub-corpora. However, Table 4.12 shows that the instances were found in all sub-corpora.

Based on the analysis, this swear word mainly possesses neutral semantic prosody. However, it is negative when it is used as a reference to create needless violence that could affect others. The swear word is also used as an enhancement of Wh-words though the context is more of a hostile expression (Gati, 2015) as well as to express demand, urgency, or anger (Goddard, 2015). On the neutral side, it is used to emphasize, intensify, as a reference of hardship or difficulty to reach a destination and to express emotion. As an expression to express emotion, the swear word is considered as an expletive swearing in which it is not directed to other people (Prawinanto, 2016).

4.4.10 PUSSY

Instead of being a domesticated cat that is commonly kept as a pet, *pussy* as defined by Oxford Dictionary 9th edition (2008) as a slang that refers to female sexual organs especially the vulva.

Table 4.13 The semantic prosody of PUSSY

Swear Words	Concordance Lines	Collocates	Semantic Prosody	File
pussy	diamonds really froze Put that dick up in her **pussy** bet she feel it in her toes I'm	-	External parts of female genitalia	File34_2016
	threat Bought a jet, what do you expect? Her **pussy** 's so good I bought her a pet			File6_2014
	leave with me Said that I can get that **pussy** easily Said that I can hit that shit so			File34_2016

	hand on that steering wheel Right hand on that **pussy** I say fuck you, 'less I'm with'			File54_2012

This swear word is the third swear word after NIGGA and HOE that does not have fixed collocates. Hence, the semantic prosody of *pussy* is identified by looking at how it behaves in the corpus (Partington, 2004). Notwithstanding, there is only one semantic prosody found for *pussy* that it refers to the external parts of female genitalia (File34_2016). Bergen (2016) shares the same view that one does not simply (ever) use the word *pussy* to refer to a cat or a rooster as a *cock*. Throughout the years, the expression is found in sub-corpora 2012, 2014, and 2016 that only refers to the female genitalia.

4.5 Summary

Findings indicate that swear words with negative base meaning can possess neutral (or even positive) semantic prosody in which it is context dependent for instance in the case of FUCK that is positive when it is used to show care and negative when it is used to express emotion. However, some of the swear words remain negative. The following is a summary of the analysis of the semantic prosody of the swear words.

Table 4.14 The summary of the semantic prosody of the swear words

Swear Words	Negative	Neutral	Positive
nigga	1. Inflict pain to others 2. Own weapons 3. Womanizer	1. A slang for friend	-
niggas	1. Faker 2. Thief	-	-
fuck	1. An abrupt emphatic exclamation expressing emotion 2. To not care about someone or something	-	1. To care about someone or something 2. To treat someone or something good

	3. To treat someone or something badly 4. Shut up (harsh)		
fuckin	1. To cause problem to someone or something	1. Intensifier 2. Emphasize	-
fucking	1. To cause problem to someone or something 2. To treat someone or something badly	1. Intensifier 2. Emphasize	-
shit	1. A slang for drugs 2. Someone or something is useless 3. Doing something bad 4. A statement of not acknowledging something	1. A trouble or a difficulty 2. Bad day	-
ass	1. Verbal abuse	1. Skin-tight 2. Colloquial universal pronouns 3. Intensifier	-
bitch	1. Having undesirable or negative quality 2. Mistress 3. Unkind remarks	1. Good bitch (ironic)	-
bitches	1. Having undesirable or negative quality 2. Up for anything or something bad	1. Having a desirable or positive quality (ironic)	-
god	-	1. Religious remarks 2. An abrupt emphatic exclamation expressing emotion	-
damn	-	1. An abrupt emphatic exclamation expressing emotion 2. Intensifier 3. Emphasize 4. To not care about negativity	-
hoe	1. Activities associated with sexual intercourse 2. Takes drugs	-	-
hoes	1. Dishonest 2. Unfaithful	-	-
hell	1. Referring to violence and needless disturbance	1. Emphasize 2. Intensifier 3. Difficulty and suffering	-

	2. An enhancement of Wh-words	4. An abrupt emphatic exclamation expressing emotion	
pussy	1. External parts of female genitalia	-	-

Based on Table 4.15, 2 out of 15 swear words possesses only neutral semantic prosody namely *god* and *damn*. Additionally, there is 1 out of 15 swear words that possess positive semantic prosody namely *fuck*. The other swear words either possess only neutral semantic prosody or negative semantic prosody.

The swear words possess negative semantic prosody due to the senses of inflicting pain upon others, verbal abuse, consuming harmful substances, putting someone into trouble, treating someone badly, and many more. On the other hand, the swear words that possess neutral semantic prosody is either possessing both negative and positive semantic prosody together or the swear words do not possess any negative senses. Stubbs (1996) reiterates that if the collocates that a node word attracts are mostly of strong negative semantic characteristics, the node word possesses a strong negative prosody. However, if both positive and negative collocates exist in the context, the node word can be said to possess a neutral or mixed prosody. This is similar to Partington (2004) that a pleasant or favourable affective meaning is labelled as positive whilst an unpleasant or unfavourable affective meaning is identified as negative. The instances are labelled as neutral when the context provided has no evidence of neither positive nor negative semantic prosody.

The swear words that possess neutral semantic prosody is due to the senses of ironic, negation, a slang, religious remarks, used as intensifier, exclamation, emphasize, and many more. Lastly, the swear words also possess a number of positive semantic prosody. The senses are to show that one is care about someone or something, to treat others good, and exposing the truth.

CHAPTER 5: CONCLUSION

5.1   Introduction

In this chapter, the summary of the findings will be presented in section 5.2. The discussion of the findings is presented in section 5.3 followed by recommendations for future studies (Section 5.4).

5.2   Summary of Findings

Based on the analysis, swear words with negative meaning can possess either neutral or positive semantic prosody. The swear words that possess neutral semantic prosody in the study are *nigga*, *fuckin*, *fucking*, *shit*, *ass*, *bitch*, *bitches*, and *hell*. There are 2 swear words that only possess neutral semantic prosody namely *god* and *damn*. There is only 1 swear word that possesses positive semantic prosody that is *fuck*. Lastly, there are 3 swear words that only possess negative semantic prosody namely *niggas*, *hoe*, and *hoes*. The semantic prosody is however, context dependent.

Table 5.1 Summary of the semantic prosody

Negative	Neutral	Positive
1. Inflicting pain to others	1. A slang for friend	1. To care about someone or something
2. Own weapons	2. Intensifier	2. To treat someone or something good
3. Womanizer	3. Emphasize	
4. Faker	4. A trouble or a difficulty	
5. Thief	5. Bad day	
6. An abrupt emphatic exclamation expressing emotion	6. Skin-tight	
7. To not care about someone or something	7. Colloquial universal pronouns	
8. To treat someone or something badly	8. Good bitch (ironic)	
9. Shut up (harsh)	9. Having a desirable or positive quality (ironic)	
10. To cause problem to someone or something	10. Religious remarks	
11. A slang for drugs	11. An abrupt emphatic exclamation expressing emotion	
12. Someone or something is useless	12. To not care about negativity	
13. Doing something bad		

14. A statement of not acknowledging something 15. Verbal abuse 16. Having undesirable or negative quality 17. Mistress 18. Unkind remarks 19. Up for anything or something bad 20. Activities associated with sexual intercourse 21. Takes drugs 22. Dishonest 23. Unfaithful 24. Referring to violence and needless disturbance 25. An enhancement of Wh-words 26. External parts of female genitalia		

Table 5.1 above shows the summary of the semantic prosody. Generally, swear words that possess negative semantic prosody are mainly because of the unfavourable meanings that are either bad or harmful. On the other hand, swear words that possess positive semantic prosody are because of the meaning which shows something that is good such as care and good treatment (good in situation). Swear words that possess neutral semantic prosody are generally because the meaning supports neither negative nor positive meanings.

5.3   Discussion of Findings

Primarily, this study seeks to answer the 3 research questions as follows:

1. What are the swear words found in a corpus of English songs that comprises of songs collected from Billboard starting at year 2011 to 2016?

2. What are the top 10 swear words found in the sub-corpus of English songs from 2011 to 2016?

3. What is the semantic prosody of swear words found in a corpus of English songs?

To answer research questions one and two, AntConc was used to generate a word list to identify the swear words based on the notion by Bergen (2016) on HFSN. The third research question was also answered by using AntConc by generating concordance lines. The semantic prosody of swear words was then identified manually by adapting the definition of semantic prosody as proposed by Partington (2004).

Usually, semantic prosody is found to have associated negative meanings with few bearing positive meaning (Xiao and McEnery, 2006). In prominent studies on semantic prosody, it was found that a lot of phrases possess negative semantic prosody for instance BREAK *out*, HAPPEN, and SET *in* by Sinclair (1991), *bent on*, *build up of*, END *up* V-ing, and *symptomatic of* by Louw (1993), ACCOST, CAUSE, and *signs of* by Stubbs (1995, 1996), and COMMIT, PEDDLE/*peddler*, and *rife* by Partington (1998, 2004). There are also a few phrases that possess positive semantic prosody namely BUILD *up a* by Louw (1993), PROVIDE and *career* by Stubbs (1995, 1996), and *impressive* by Partington (1998, 2004). Similarly, the data from the analysis shows that 13 swear words possess negative semantic prosody. However, there are also instances where the collocation forms ironic meaning as portrayed in BITCH for example *good bitch*. According to Louw (1993), collocates that have the sense of irony, insincerity or humour violate semantic prosody and thus, would abrogate the negative meaning it would normally cause. In this case, the ironic BITCH possesses neutral semantic prosody. The data in this study shows that the only swear word that possesses positive semantic prosody is *fuck* in which it is used to show care and to treat someone good although the swear word *fuck* possesses negative base meaning just like the other swear words. Louw (1993) and Xiao and McEnery (2006) reiterate that semantic prosody is inaccessible to the speaker's conscience. In other words, the analysis in this study shows that semantic prosody does portray a hidden quality in which swear words with negative base meaning

91

could also possess positive and neutral semantic prosody. This is however subject for a more thorough analysis by using a larger corpus.

In most studies on swear words, researchers usually use survey to rate the offensiveness of swear words (see Baudin & Paramasivam, 2014, Dewaele, 2015, Kapoor, 2016, Burgen, 2016, and Bryne, 2017). The act of rating is classified into 2: (i) rating the act of swearing as a whole, and (ii) rating the swear words individually. In the case of rating the act of swearing, the act was rated without any context provided (see Bryne, 2017). Based on a data in a study by Kapoor (2016), the perception on swearing varies among the participants regardless of context. In a study by Baudin and Paramasivam (2014), the participants were asked to rate swearing as a whole also without any context nor swear words provided. Although both swear words and context were provided for the British English participants and American English, the perception on the same swear words (*bugger*, *thick*, *bollocks*, and *wanker*) is different (see Dawaele, 2015). The British English participants rate the mentioned swear words as more offensive than the American English participants. This portrays that the use of survey to rate swearing/swear words is unreliable (Beers, 2007). Hence, context is needed for the rating of swear words. Although semantic prosody can identify whether a swear word possesses negative, positive, or neutral semantic prosody, it cannot identify the intensity of a swear word. In this case, the degree of offensiveness a swear word is used towards others as an insult cannot be measured. Then again, Likert scale (or survey) is useful in measuring the degree of offensiveness as illustrated in a few studies (see Bergen, 2016 and Goddard, 2015).

5.4   Recommendation for Future Studies

Studies involving swear words are recommended for data derived from a corpus of natural language for example a corpus that is made of daily conversation. However, it is meticulous to do so. This is because a lot of time is needed to build such corpus, for example Murphy (2009) took over eight months to compile a spoken corpus of Irish English. The time needed in the compilation of the spoken corpus is to ensure the representativeness of the focused language (Biber, 1990). Apart from time, the availability of participants and duration of conversations should be considered (Murphy, 2009) as it play a major role in obtaining data. Nevertheless, it is still doable. Hence, to have a more authentic semantic prosody of swear words with authentic examples (Louw and Chateau, 2010), a spoken corpus comprises of natural conversation is recommended.

5.5   Summary

Based on the study of semantic prosody of swear words, it can be concluded that a negative based swear word can possess positive semantic prosody. This is however, context dependent. Hence, context plays an important role in determining how a word behaves. As coined by Firth (1957), one must examine the collocations of a word instead of the word in isolation in order to understand it. However, as this study has shown/revealed, a larger corpus would be ideal to investigate how swear words truly behave in specific contexts/register, particularly those found/occurring in popular songs (lyrics).

REFERENCES

Allan, K., Burridge, K. (2006). *Forbidden words: taboo and the censoring of language*. Cambridge: Cambridge University Press.

Amrullah, L. (2016). English swear words by Indonesian learners. *Journal of English Language Teaching and Linguistics*, *1*(1), 1-12.

Andersson, L. and P. Trudgill. (1990). *Bad Language*. Oxford: Basil Blackwell.

Anderson, L., & Lepore, E. (2013). What did you call me? Slurs as prohibited words. *Analytic Philosophy*, *54*(3), 350-363.

Archer, D., Wilson, A., & Rayson, P. (2002). Introduction to the USAS category system. *Benedict project report.*

Ayto, J., & Simpson, J. (1996). *The Oxford Dictionary of Modern Slang*. Oxford: Oxford University Press. 308 pp.

AZLyrics. (n. d.). *Welcome to AZLyrics*. Retrieved from https://www.azlyrics.com/.

Baker, P., Hardie, A. & McEnery, A. (2006). *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Baudin, N., & Paramasivam, S. (2014). Swearing In English Among A Group Of Female Malaysian Teenagers. *International Journal of Contemporary Applied Sciences*, *1*(3), 14-25.

Beers Fägersten, K. (2007). *A sociolinguistic analysis of swear word offensiveness*. Universität des Saarlands. Saarland Working Papers in Linguistics.

Beers Fägersten, K. (2012). *Who's swearing now? the social aspects of conversational swearing*. Cambridge Scholars Publishing.

Bednarek, M. (2008). Semantic preference and semantic prosody re-examined. *Corpus linguistics and linguistic theory*, *4*(2), 119-139.

Begagić, M. (2013). Semantic preference and semantic prosody of the collocation make sense. *Jezikoslovlje*, *14*(2-3), 403-416.

Bergen, B. K. (2016). *What the F: what swearing reveals about our language, our brains, and ourselves*. New York: Basic Books.

Biber, D. (1990). Methodological issues regarding corpus-based analyses of linguistic variation. *Literary and linguistic computing*, *5*(4), 257-269.

Biber, D. (1991). *Variation across speech and writing*. Cambridge University Press.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus Linguistics: investigating language structure and use*. Cambridge: Cambridge University Press.

Billboard, (2017). Hot 100 song year-end chart. Retrieved from
        http://www.billboard.com/charts/year-end/2016/hot-100-songs.

Boothe, D., & West, J. (2015). English language learning through music and song
        lyrics–The performance of a lifetime. In *Conference Proceedings. The Future of
        Education* (p. 248). libreriauniversitaria. it Edizioni.Bouma, G. (2009).
        Normalized (pointwise) mutual information in collocation extraction.
        Proceedings of GSCL, 32-40.

Bowers, J. S., & Pleydell-Pearce, C. W. (2011). Swearing, euphemisms, and linguistic
        relativity. *PloS one*, *6*(7), e22341.

Brezina, V., McEnery, T., & Wattam, S. (2015). Collocations in context: A new
        perspective on collocation networks. *International Journal of Corpus
        Linguistics*, *20*(2), 139-173.

Bublitz, W. (1995). Semantic prosody and cohesive company: somewhat predictable´.
        General and Theoretical Papers, 1-23.

Butler, C. W., & Fitzgerald, R. (2011). "My f*** ing personality": swearing as slips and
        gaffes in live television broadcasts. *Text & Talk-an Interdisciplinary Journal of
        Language, Discourse & Communication Studies*, *31*(5), 525-551.

Byrne, E. (2017). *Swearing is Good for You: The Amazing Science of Bad Language*.
        Profile Books.

Carston, R. (2011). Truth-conditonal semantics. In M. Sbisà, J. O. Östman, & J.
        Verschueren, (Eds.), *Philosophical perspectives for pragmatics* (Vol. 10). John
        Benjamins Publishing.

Coates, J. (2004). *Women men, and language: a sociolinguistic account of gender
        differences in language*. Pearson Longman.

Coyle, Y., & Gómez Gracia, R. (2014). Using songs to enhance L2 vocabulary
        acquisition in preschool children. *Elt Journal*, *68*(3), 276-285.

Croom, A. M. (2013). How to do things with slurs: Studies in the way of derogatory
        words. *Language & Communication*, *33*(3), 177-204.

Croom, A. M. (2014). The semantics of slurs: a refutation of pure
        expressivism. *Language Sciences*, *41*, 227-242.

Dewaele, J. M. (2015). British 'Bollocks' versus American 'Jerk': Do native British
        English speakers swear more–or differently–compared to American English
        speakers?. *Applied Linguistics Review*, *6*(3), 309-339.

Ellis, R. (1989). *Understanding second language acquisition.* Oxford: Oxford
        university press.

Fägersten, K. B. (2017). The role of swearing in creating an online persona: The case of
        YouTuber PewDiePie. *Discourse, context & media*, *18*, 1-10.

Falk, J. (2012). *We will rock you: a diachronic corpus-based analysis of linguistic features in rock lyrics* (Unpublished BA Thesis). Linnaeus University.

Fernández Dobao, A. M. (2006). Linguistic and cultural aspects of the translation of swearing: The Spanish version of Pulp Fiction. *Babel*, *52*(3).

Firth J.R., (1957a). *Papers in Linguistics*. London, Oxford University Press.

Flowerdew, L. (2012). Corpora in the classroom: an applied linguistics perspective. In K. Hyland, Chau Meng Huat, M. Handford, (Eds.), *Corpus application in applied linguistics* (pp. 208-224). Britain: Continuum.

Fogle, C. D. (2013). The Etymology, Evolution and Social Acceptability of 'Nigger', 'Negro', and 'Nigga', 1-28.

Gajanan, M. (2018). *An Expert's Take on the Symbolism in Childish Gambino's Viral 'This is America' Video*. Retrieved from https://time.com/5267890/childish-gambino-this-is-america-meaning/.

Garside, R. (1987). The CLAWS word-tagging system. In: R. Garside, G. Leech and G. Sampson (eds), *The computational analysis of English: a corpus-based approach*. London: Longman.

Gati, P. (2015). The use of swear words by women: a *study of single sex and mix sex Conversations* (Unpublished BA Thesis). Halmstad University.

Gauthier, M., Guille, A., Rico, F., & Deseille, A. (2015). Text mining and Twitter to analyse British swearing habits. *Handbook of Twitter for Research*, 27-43.

Goddard, C. (2013). The semantic roots and cultural grounding of 'social cognition'. *Australian Journal of Linguistics*, *33*(3), 245-256.

Goddard, C. (2015). "Swear words" and "curse words" in Australian (and American) English. At the crossroads of pragmatics, semantics and sociolinguistics. *Intercultural Pragmatics*, *12*(2), 189-218.

Hall, R (2009.). *Music Censorship*. Retrieved from https://web.archive.org/web/20160819103434/http://www.lehigh.edu/~infirst/musiccensorship.html.

Hashim, S. Z. M., & Rahim, H. A. (2016). Defying the Global: The Cultural Connotations of" Islam" in Malaysia. *KEMANUSIAAN: The Asian Journal of Humanities*, *23*, 81-98.

Hauser, D., & Schwarz, N. (2016). Semantic prosody and judgement. *Journal of Experimental Psychology General*, 1-39.

Hedger, J. A. (2012). The semantics of racial slurs: using Kaplan's framework to provide a theory of the meaning of derogatory epithets. *Linguistic and Philosophical Investigations*, *11*, 74.

Hitchcock, E., & Cairns, V. (1973). Amygdalotomy. *Postgraduate Medical Journal*, *49*(578), 894-904.

Hoey, M. (2004). *Lexical priming and the properties of text*. United Kingdom: Taylor & Francis

Hoffman, A. J., Wallach, J., Graham, M., & Sanchez, E. (2009). Changing perceptions in the justification of the use of the "N" word through community service work experiences: Offensive comments or terms of endearment?. *Journal of Praxis in Multicultural Education*, *4*(1), 3.

Hom, C. (2012). A puzzle about pejoratives. *Philosophical Studies*, *159*(3), 383-405.

Hom, C. (2008). The semantics of racial epithets. *The Journal of Philosophy*, *105*(8), 416-440.

Horan, G. (2013). 'You taught me language; and my profit on't/Is, I know how to curse': cursing and swearing in foreign language learning. *Language and Intercultural Communication*, *13*(3), 283-297.

Hornby, A. S. (2015). *Oxford advanced learner's dictionary*. Oxford: Oxford University Press.

Hughes, G. (1991). *Swearing: a social history of foul language, oaths and profanity in English*. Oxford: Blackwell.

Hunston, S., & Francis, G. (2000). *Pattern Grammar; A corpus-driven approach to the lexical grammar of English*. University of Birmingham: John Benjamin Publishing Company.

Hunston, S., & Thompson, G. (Eds.). (2000). *Evaluation in text: Authorial stance and the construction of discourse: Authorial stance and the construction of discourse*. United Kingdom: Oxford University Press.

Hunston, S. (2001). Colligation, lexis, pattern, and text. In M. Scott, G. Thompson, *Patterns of text: In honour of Michael Hoey* (pp. 13-33). Amsterdam: John Benjamin Publishing Company.

Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press. University Press, UK.

Hunston, S. (2007). Semantic prosody revisited. *International Journal of Corpus Linguistics*, 249-268.

Jackson, H. (2014). *Words and their Meaning*. London and New York: Routledge.

Jackson, H., & E. Ze Amvela. (2000). *Words, Meaning and Vocabulary: An Introduction to Modern English Lexicology*. London: Cassell.

Jay, T. (1992). *Cursing in America; A psycholinguistics study of dirty language in the courts, in the movies, in the schoolyards and on the streets*. Amsterdam: John Benjamins Publishing Company.

Jay, T., & Janschewitz, K. (2008). The pragmatics of swearing. *Journal of Politeness Research*, *4*, 267-88.

Jay, T. (2009). The utility and ubiquity of taboo words. *Perspectives on Psychological Science*, *4*(2), 153-161.

Jeshion, R. (2013). Slurs and stereotypes. *Analytic Philosophy*, *54*(3), 314-329.

JSTOR Daily. (2019). *In Which We Get to the Bottom of Some Crazy-Ass Language*. Retrieved from https://daily.jstor.org/in-which-we-get-to-the-bottom-of-some-crazy-ass-language/.

Kapoor, H. (2016). Swears in context: The difference between casual and abusive swearing. *Journal of psycholinguistic research*, *45*(2), 259-274.

Karjalainen, M. (2002). Where have all the swearwords gone: An analysis of the loss of swearwords in two Swedish translations of JD Salingers Catcher in the Rye. (Unpublished Thesis). University of Helsinki.

Katznelson, N., Gelman, J., Lindblom, K., & Caput, M. (2010). American song lyrics: a corpus−based research project featuring twenty years in rock, pop, country and hip−hop. San Francisco: San Francisco University

Kaye, B. K., & Sapolsky, B. S. (2009). Taboo or not taboo? That is the question: offensive language on prime-time broadcast and cable programming. *Journal of broadcasting & Electronic media*, *53*(1), ISS 1.

Kenny, A. (2011). A Qualitative Analysis of the History of the N-word and How It Is Used as a Term of Endearment. *2011 NCUR Password Protected Papers*.

King, S. A. (1996). Researching Internet communities: Proposed ethical guidelines for the reporting of results. *The Information Society*, *12*(2), 119-128.

Kleinman, S., Ezzell, M., & Frost, A. C. (2009). Reclaiming Critical Analysis: The Social Harms of 'Bitch.'. *Sociological Analysis*, *3*(1), 46-68.

Klimova, B. F. (2013). Developing thinking skills in the course of academic writing. *Procedia-Social and Behavorial Sciences*, *93*, 508-511.

Krashen, S. D. (1981). *Second Language Acquisition and Second Language Learning*. Retrieved from http://196.189.45.87/bitstream/123456789/58723/1/26pdf.pdf.

Kuhn, E. D. (1999). 'I just want to make love to you'-Seductive strategies in blues lyrics. *Journal of pragmatics*, *31*(4), 525-534.

Leech, G. (1981). Semantics (2nd edn). Harmondsworth: Penguin.

Lepore, E., & Ludwig, K. (2003). Outline for a Truth Conditional Semantics for Tense. *Tense,time and reference*, 49-105.

Lin, G. H. C. (2008). Pedagogies Proving Krashen's Theory of Affective Filter. *Online Submission*.

Ljung, M. (2011). *Swearing: A cross-cultural linguistic study*. Basingstoke: Palgrave MacMillan.

Louw, B. (2002). Contextual prosodic theory: Bringing semantic prosody to life in C. Heffer, H. Sauntson, and G. Fox (eds): *Words in context: A tribute to John Sinclair on his retirement*. Birmingham: University of Birmingham.

Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. *Text and technology: In honour of John Sinclair*, *240*, 251.

Louw, B., & Chateau, C. (2010). Semantic prosody for the 21st century: Are prosodies smoothed in academic contexts? A contextual prosodic theoretical perspective. In *Proceedings of 10th International Conference on Statistical Analysis of Corpus Study* (pp. 755-764).

Lutzky, U., & Kehoe, A. (2016). Your blog is (the) shit. *International Journal of Corpus Linguistics*, 165-191.

Lyons, J. (1977). *Semantics* (Vol. 2). London: Cambridge university press.

Malaysian Communications and Multimedia Commission. (2018). *Bahasa Terlarang*. Retrieved from https://www.mcmc.gov.my/skmmgovmy/files/22/2271ad9b-6921-4407-b5e3-0d244e4fa837/index.html#1/z.

Martin, (2012). *English 1106*. Retrived from https://blogs.lt.vt.edu/claireyw/2012/05/02/why-is-music-censored/.

Martinez, A. S. (2008). Collocation analysis of a sample corpus using some statistical measures: An empirical approach. In *Proceedings of the 25th International AESLA [The Spannish Society for Applied Linguistics] Conference* (Vol. 25, pp. 763-768).

McEnery, A., & Xiao, Z. (2004). Swearing in modern British English: the case of fuck in the BNC. *Language and Literature*, *13(3),* 235-268.

McEnery, T. (2006a). *Swearing in English: Bad Language, Purity and Power from 1586 to the Present*. New York: Routledge.

McEnery, T. (2006b). The moral panic about bad language in England, 1691-1745. *Journal of Historical Pragmatics*. 7:1, 89-113.

McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. London and New York: Routledge.

McGee, I. (2012). Should we teach semantic prosody awareness? *RELC Journal*, *43*(2), 169-186.

Murphy, B. (2009). "She's a F**king ticket': the pragmatics of f**k in Irish English – an age and gender perspective". *Corpora*, *4(1)*, pp. 85-106.

Nesselhauf, N. (2004). What are collocations? In D. Allerton, N. Nesselhauf, & P. Skandera (Eds.), *Phraseological units: Basic concepts and their application* (pp. 1-21). Basel: Schwabe.

Noor Baizura Adnan. (2013). A corpus-based study of time metaphors in the text simplification of 'around the world in eighty days' (Unpublished Thesis). University of Malaya.

O'Higgins, P. (1972). *Censorship in Britain.* Suffolk: Thomas Nelson Publisher.

Partington, A. (1998). *Patterns and meanings: Using corpora for English language research and teaching* (Vol. 2). John Benjamins Publishing.

Partington, A. (2004). "Utterly content in each other's company": Semantic prosody and semantic preference. *International Journal of Corpus Linguistics*. *9*(1), 131-156.

Pettijohn, T. F., & Sacco Jr, D. F. (2009). The language of lyrics: An analysis of popular Billboard songs across conditions of social and economic threat. *Journal of Language and Social Psychology*, *28*(3), 297-311.

Piao, S., Rayson, P., Archer, D., & McEnery, A. M. (2004). Evaluating lexical resources for a semantic tagger.In *proceedings of 4th International Conference on Language Resources and Evaluation*, *2*, 499-502.

Pragglejaz Group (2007). MIP: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol*, *22*(1), 1-39.

Prawinanto, A. (2016). A Socialinguistic Analysis of the Use of Swear Words in the Antologi Rasa Novel Written By Ika Natassa. *Yogyakarta: Sanata Dharma University. https://repository. usd. ac. id/3303/2/1112141 58_full. pdf*.

Pujol, D. (2006). The Translation and Dubbing of'Fuck'into Catalan: The Case of From Dusk till Dawn. *The Journal of Specialised Translation*, *6*, 121-133.

Rahim, H. A. (2012). 'Impak konotasi budaya terhadap leksis: satu kajian semantik berasaskan korpus, ke atas perkataan perempuan dan wanita'. *Jurnal Bahasa*, *5*(1): 83-111.

Rayson, P. (2008). From key words to key semantic domains. *International Journal of Corpus Linguistics*. *13*(4), 519-549.

Rohde, A., & Tiefenthal, C. (2000). Fast mapping in early L2 lexical acquisition. *Studia Linguistica*, *54*(2), 167-174.

Sardinha, T. B. (1999). *Using keywords in text analysis: practical aspects*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.582.9840.

Sbisà, M., Östman, J. O., & Verschueren, J. (Eds.). (2011). *Philosophical perspectives for pragmatics* (Vol. 10). John Benjamins Publishing.

Schoepp, K. (2001). Reasons for using songs in the ESL/EFL classroom. *The internet TESL journal*, *7*(2), 1-4.

SEO Mark UK. (2018). *How does google rank websites?*. Retrieved from https://www.seomark.co.uk/how-does-google-rank-websites/.

Setia, R., Rahim, R. A., Nair, G. K. S., Husin, N., Sabapathy, E., Mohamad, R., ... & Kassim, R. M. (2012). English songs as means of aiding students' proficiency development. *Asian Social Science*, *8*(7), 270.

Simpson, J., Weiner, E., & Durkin, P. (2004). The Oxford English Dictionary Today. *Transactions of the Philological Society*, *102*(3), 335-381.

Sinclair, J. M. (Ed.). (1987). *Looking up: An account of the COBUILD project in lexical computing and the development of the Collins COBUILD English language dictionary*. London: Collins Elt.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Sinclair, J. M. (2003). *Reading concordances*: *An introduction*. United Kingdom: Pearson Longman.

Sinclair, J. M. (2004). *Trust the text: Language, corpus and discourse*. London: Routledge.

Sixsmith, J., & Murray, C. D. (2001). Ethical issues in the documentary data analysis of Internet posts and archives. *Qualitative health research*, *11*(3), 423-432.

Smith, K. A., & Nordquist, D. (2012). A critical and historical investigation into Semantic prosody. *Journal of Historical Pragmatics*, *13*(2), 291-312.

Steen, G. J., Dorst, A. G., Herrmann, J. B., Kall, A. A., Krennmayr, T., & Pasma, T. (2010). *A method for linguistic metaphor identification. From MIP to MIPVU*. Amsterdam: John Benjamins.

Stenström, A. B. (2006). Taboo words in teenage talk: London and Madrid girls' conversations compared. *Spanish in Context*, *3*(1), 115-138.

Stewart, D. (2010). *Semantic prosody: a critical evaluation*. London: Routledge.

Stubbs, M. (1995). Collocations and semantic profiles: on the cause of the trouble with quantitative studies. *Functions of Language*, *2*(1): 23-55.

Stubbs, M. (1996). *Text and corpus analysis: Computer-assisted studies of language and culture* (p. 158). Oxford: Blackwell.

Stubbs, M. (2002). Two quantitative methods of studying phraseology in English. *International Journal of Corpus Linguistics*, *7/2*: 215-44.

TEDx Talks, (2015). *The Joy of Swearing*. Retrieved from https://youtu.be/_bt8SNdke4Y.

Teubert, W. (2005). My version of corpus linguistics. *International journal of corpus linguistics*, *10*(1), 1-13.

The Nielson Company, (2017). Music sales management. Retrieved from http://www.nielsen.com/us/en/solutions/measurement/music-salesmeasurement.html.

Thelwall, M. (2008). Fk yea I swear: cursing and gender in a corpus of MySpace pages. *Corpora*, *3*(1), 83-107.

Thommen, T., & RECANATI, S. B. F. (2014). *Response-Dependence and the Semantics of Slurs* (Doctoral dissertation, Master Dissertation, Paris, École Normale Superieure).

Tognini-Bonelli, E. (2001). *Corpus linguistics at work* (Vol. 6). John Benjamins Publishing.

UCREL, (2000). *POS-Tagging error rates*. Retrieved 22nd of December from http://ucrel.lancs.ac.uk/bnc2/bnc2error.htm#table1.

Van Lancker, D., & Cummings, J. L. (1999). Expletives: neurolinguistics and neurobehavioral perspectives on swearing. *Brain Research Reviews*, *31*, 83-104.

Waskul, D. (1996). Considering the electronic participant: Some polemical observations on the ethics of on-line research. *The Information Society*, *12*(2), 129-140.

Whissell, C. (2008). Emotional fluctuations in Bob Dylan's lyrics measured by the Dictionary of Affect accompany events and phases in his life. *Psychological reports*, *102*(2), 469-483.

Wiles, K. (2014). *On the origin of 'fuck"*. Retrieved from http://www.huffingtonpost.com/kate-wiles/on-the-origin-of-fuck_b_4784565.html.

Xiao, R., & McEnery, T. (2006). Collocation, semantic prosody, and near synonymy: a cross-linguistic perspective. *Applied Linguistics*, 27/1: 103-129.

Yusuf, Y. Q. (2010). A Corpus-based Linguistic Analysis on Spoken Corpus: Semantic Prosodies on "Robots". *Journal of Language and Linguistic Studies*, *6*(1), 49-64.

Zhang, C. (2010). An overview of corpus-based studies of semantic prosody. *Asian Social Science*, *6*(6), 190-194.

LIST OF PUBLICATIONS AND PAPERS PRESENTED

Papers presented in conference

Title	Author(s)	Affiliation	Conference
Statistical Aspect of Collocations Using Different Tools in a Corpus of English Songs	Hazri Shahreen Hashim Sheena Kaur Jaswant Singh	University of Malaya	Symposium on Corpus Research Repost: Malaysia in conjunction with Language for Specific Purposes International Conference and 10th Global Advances in Business Communication Conference, Johor, Malaysia
Statistical Aspect of Collocations Using Different Tools in a Corpus of English Songs	Hazri Shahreen Hashim Sheena Kaur Jaswant Singh	University of Malaya	Asian Universities Alliance Postgraduate Academic Forum at Tsinghua University, Beijing, China
An analysis of collocations using different tools in a corpus of songs	Hazri Shahreen Hashim Sheena Kaur Jaswant Singh	University of Malaya	International Conference on Language, Culture and Society 2018 at International Islamic University Malaysia, Malaysia