UNIVERSITY OF OSLO

Master thesis

FUCK THIS SHIT: Swearing in Different Varieties

In Movies & on TV

Hiba Kassem

MA Degree in English Language and Linguistics

xxx 60 Credits

University of Oslo

Department of Literature, Area Studies and European Languages

Table of Contents

Abstract

This corpus-based study explores swear words, investigating their frequency, collocations, and implications in various linguistic contexts. Analysing 11 selected swear words uncovers four key patterns: intensifier reduplication, similar collocates for near-synonyms, function vs. lexical words, and diverse word classes. Comparisons with prior research, including Lutzky and Kehoe (2016), reveal both commonalities and distinctions. Semantic prosody illuminates nuanced meanings. Grice's Maxims of Conversation, addressing politeness and impoliteness, highlight potential face-threatening acts in swearing. This research enriches our understanding of the multifaceted world of swear words and their linguistic significance.

Keywords: Swearwords, swear, relation- situation vs literal meaning, function, Grice's Maxims, collocation, frequency, time, placement

Acknowledgements

I extend my heartfelt gratitude to my supervisor, Associate Professor Jacob Thaisen, for his unwavering support, invaluable guidance, and constructive feedback throughout the entirety of this thesis. His expertise and encouragement have been instrumental in making this research possible.

Chapter 1 Introduction

1.1 Good or Bad

Within a language, there are various combinations and nuances created based on the context, speakers, and listeners. The construction of sentences leads to dialogues, and by scrutinizing the word choices within these sentences, we can delve into the speaker's intentions and manner of speech, which may categorize language as either "good" or "bad." However, the determination of what qualifies as "good" or "bad" language is a topic of ongoing debate among scholars. One reason for this ongoing debate is the evolving moral standards over time, which subsequently influence the definitions of these terms. Take, for example, societal attitudes toward sexual activities. In the past, there were strict laws and moral guidelines against engaging in sexual activities outside of marriage, with greater consequences for women. Today, these norms have shifted. Similarly, language is subject to evolving societal values, and what constitutes "bad" language is less clear-cut in contemporary times. The line that once separated language into these two categories has gradually blurred in recent centuries.

1.2 History

Swearing and the usage of "bad" language is something that has been in practice for thousands of years. Of course, the variation of it was different than the ones used today. Society was not as accepting of the usage of this type of language, and it was considered taboo to make use of certain words. Having a duel was the way for the men to solve an argument that they may have. Society was divided into classes, be it between the upper class, which consisted mostly of nobles and royals, or the upper middle class, which consisted of rich commoners without any titles or merchantries, or the lower class, which consisted of the common people. In the case of England, this division into upper, middle, and lower/working classes came into being in the high Middle Ages. Considering the importance society had put on good relations between the upper and middle-upper classes, swearing or anything considered "bad language" was a taboo between them. Even when it was done, it was never in a public setting and more in private.

As for women, their image was the most important. Usage of foul language was considered a disgrace and therefore, swearing was not a desirable habit that any family wanted present.

Women were meant to be gentle, kind, graceful, and generous. This meant that they would have to be polite in their behaviour and always use polite speech, and that was what was always expected of them, no matter the gender they were speaking to. The higher their status in society was the more careful they had to be. Of course, that may not have been the case when they were in their own home, as it was not possible to document this in the privacy of their home, but in public, it was most certainly frowned upon should a woman behave and speak in that manner that they considered foul language.

1.3 Swearing is Spoken

Swearing and foul language can be performed in different ways. One of those ways can come in the form of writing. Today, swearing is not that unusual and there are not as many restrictions as there once were in the past. Using technology and posting on social media as a platform, we can tweet, text, and even send messages to express our thoughts. As most communication is done there most people tend to be as informal as possible. When sending a text to a sibling or a friend, for example, there are times when the texts contain a lot of swearing. This does not mean that it needs to be a negative thing as nowadays swearing is rather used to express the closeness between people and to show the comfortableness between them. The fact that modern novels and biographical books often contain swear words is the biggest prove that swearing has become more acceptable in society today.

There is, however, another way that swearing can be performed in, which is the focus of this thesis. This way comes in the way of speech. Seeing as swearing is seen mostly as a negative thing in a language, and this is the thought of most people, therefore, swearing is often done in situations that are considered bad or negative as well. For example, when we are angry or fighting verbally with someone, cursing and swear words tend to come out impulsively. This is an automatic response. When we are angry or fighting with someone, we tend to want to hurt their feelings, which is where swearing comes in. One of the ways it can come in, is in the form of name callings, such as *bitch* or *bastard*, where both swear words are meant to degrade the other person. There are other situations, however, where swearing can come out automatically without much thought. One of these situations is when one is in pain. For example, if I accidentally knock my forehead with someone else, I will most likely express my pain in a Yelp followed by a couple of swear words to show a mixture of both pain, anger,

and frustration. Some of the swears most people, who speak English as their first language, would most likely use in this kind of situation would be *fuck, hell,* or *shit*.

1.4 Negative or positive

Swearing and usage of "bad or foul" language does not necessarily mean that it is taken negatively. Depending on how the terms are defined, the outcome of them may be positive or negative. Reaction is another vital turning point to determine whether the two terms are taken in which direction. The bond and connection between the speaker and listener are something that plays a big role in how each of the swear words are interpreted. The way we talk to our friends and siblings, for example, is different than the way we tend to talk to our parents. The same way is for when swearing is used. Depending on the relationship between the two parties involved in the conversation where swearing is involved, the interpretation and intention behind the swear words used may be seen as a positive, or as the traditional negative for swearing.

1.5 Theory

When investigating a topic, in any study done, having a clear path throughout the paper is the most useful way to do it. By deciding on a specific method to carry out the investigation, the entire paper will be bound together and that makes it easier for the reader to read and understand exactly what the point of the paper is. The method helps the speaker investigate the thesis questions, or as we call them, research question or questions if there is more than one thing that the speaker wants the thesis to convey. There are many different methods that we, linguists, tend to use. One of the most popular methods is the corpus method. This method is one of the methods where the data is already gathered for us, which is why many linguists chose it. There are times one we cannot find a corpus that is suitable for some topics, which is why compiling and creating our very own corpus is an option as well. Although, it does take time, the results are that this makes it more efficient to use.

Another method is to investigate if the data that we are using applies to a theory coming from another author or linguist and test it using the theory or theories that are relevant to our topic. The theory that we chose for the test would have to be relevant to the topic chosen for our research so that we can avoid confusion as to why we are using a theory that cannot be used in an investigation. This thesis, for example, will investigate words that may be considered as

"bad", and that is why the theory that needs to be used would need to be able to create an opportunity for an investigation of the topic for this thesis.

1.6 Aim

This thesis delves into the intricate landscape of profanity and expletives as portrayed on television and in movies. Consequently, I will delve into the following research inquiries by using the lens of Grice's Maxims of Conversations (Kroeger, 2018, p. 142) as the groundwork for the theory. Additionally, I will employ frequency and collocation analyses to scrutinize potential shifts or variations that the English spoken language has experienced over a specific timeframe, from the 1930s to the 2010s in both the TV corpus and the Movies corpus.

The research questions that formed the groundwork for this thesis are as follows:
1. How do the findings of the present master's thesis compare with Robbie Love's findings?
2. Which swear words are the most frequent, in each corpus, over time, and in which variety of spoken English?
3. Which words collocate with the swear words and on that basis what semantic prosody do the swear words have?
4. Using Grice's maxims of Conversation, does using swear words violate or break any of the maxims? How?

Chapter 2: Literature Review

"Swearing draws upon such powerful and incongruous resonators as religion, sex, madness, excretion, and nationality, encompassing an extraordinary variety of attitudes including the violent, the amusing, the shocking, the absurd, the casual, and the impossible."

Geoffrey Hughes

This chapter aims to provide a comprehensive exploration of the concept of "bad" language, encompassing its definition, criteria for classification, and contextual interpretation. An examination of relevant prior research within this domain will be included, highlighting its relevance to the research questions addressed in this thesis. Furthermore, this chapter will elucidate the term "swearing" from multiple linguistic perspectives, drawing insights from at least two distinct linguists' viewpoints, and culminating in my understanding of the term. A compilation of the profane vocabulary employed in this thesis will be presented, complete with corresponding meanings and definitions. Additionally, alternative expressions that can substitute for these profanities will be outlined. From a linguistic standpoint, the thesis will expound upon Grice's Maxims, elucidating their significance and outlining instances where speakers deviate from these principles. This chapter will also delve into the interconnectedness between swearing, the usage of profanity, and Grice's Maxims (Grice, 1975), shedding light on the intricate relationship between language usage and communicative norms.

2.1 "Bad" language

What is it that counts as "bad" language? For many years linguists have investigated this topic and each time they come up with different terms and definitions to try to describe this topic. However, this topic has only received increased attention in the last two decades. Even though many of them could agree, many not agree but come close to definitions and terms close to it. One of the reasons behind this is the fact that each society is different, and the same case is for each generation. Byrne (2017) mentions the incident when she used a "bad" word against her little brother and was smacked as a result. At the time, she did not know the meaning, but she did learn that words are powerful. This situation is an example of how bad language is perceived and how there will be consequences that will follow should this type of language be used. What comes with the term bad is the usual negative perspective on the manner in which

the language is conducted. For example, the choice of the words and the attitude of the speaker.

Byrne's (2017) incident with bad language, shows that society still does not completely. Although the adults forbid the children around them from using profanity, that does not mean that it would stop them completely. In Byrne's (2017) case, it only intrigued her more. This proves that human nature cannot be controlled, even though this incident happened to her as a child, it still left an impression on Byrne. From a very young age, our parents tend to forbid us from using foul language, and more times than none would give out punishments as a consequence if we do use it after all. This is not only the case with parents, but also with all the adults around the children, such as teachers, relatives, family friends, etc.

One of the reasons why grown-ups forbid children from using such language may be for fear of them hurting others and themselves. Usage of foul language may offend others, be it their peers or the general other people around them. Parents do not wish for their children to appear rude or have the image of someone that hurts others, and that is one of the reasons that they stop the children and forbid them from using language that may hurt their image. But controlling children is easy only when they are small. Once the children get older and form their own friendships and social circles, they get more independent and start using foul language in those intimate social circles before they expand it as they get older. The more someone is forbidden from doing something, the more intrigued they get, and they end up wanting to do it even more. In adolescence, teenagers tend to form a language for themselves, in which case the language consists of the forbidden and what the adults considered the "foul" or "bad" language use. Love (2021, pp. 743–744) found that McEnery (2006, p. 38) discovered that in terms of age, "adolescence and young adults are more likely to use bad language words". This confirms the theory that concludes that people in adolescents tend to make use of bad or foul language more than that of grown adults.

2.2 Swearing

In his paper, Love (2021, p. 1) explains that swearing "has the potential to offend and abuse, and yet entertain and create humour in equal measure". This means that although *swearing* can offend and hurt someone if used in certain ways, it can bring some "seasoning" into our everyday lives. Love (2021) gathered two to three definitions from other linguists trying to

define swearing in their own words. Fägersten defined swearing as "having the potential to be offensive, inappropriate, objectionable, or unacceptable in any given social context"(Beers Fägersten, 2012, p. 3). The fact that he thinks that it is not acceptable in any social context does not seem to be entirely true. Most teenagers tend to use swearing as a way or form of bonding. Whether they are out together or are in social events together with friends or anyone their age, it is normal to be informal with each other, and swearing is one of the ways they choose to do it. This can be proven by Stapleton (2010) as he stated, "Swearing can also perform functions like expressing humour, creating social bonds and constructing identity"(Love, 2021, p. 741).

"Despite such attention in the literature, swearing is notoriously difficult to define" (Love, 2021, p. 741). The reason for this may be because swearing comes in different forms and different styles. Each generation modifies the already existing swear words and ends up creating a new version of them. This results in many "old" swear words not being used as much. An example of this is the swear word *fanny*. This swear word was most frequently used in British, Australian, and New Zealand spoken English from 1930 to 1960. After that, the swear word *arse* became more popular and was used more than *fanny* even though both swear words have the same meaning, as will become apparent from the data Chapter 5. Within the realm of lexical evolution, a phenomenon known as "semantic change" manifests itself through various principal types: widening, narrowing, amelioration, and pejoration. In this context, an examination of the swear words "fanny" and "arse" reveals intriguing instances of such semantic transformations. According to the Oxford Learner's Dictionary, the official denotation of the swear word "fanny" refers to "a person's bottom," predominantly within the informal context of North American English. However, the term also undergoes semantic change, acquiring an alternative meaning in British spoken English, where it signifies "the female sex organ." Notably, within the Oxford Learner's Dictionary, this swear word is categorized as both taboo and slang, indicating its contentious and informal nature (Oxford Learner's Dictionary, s.v. «fanny». Conversely, the term "arse" is semantically defined as "the part of the body that you sit on; your bottom" within the same dictionary. This definition underscores its anatomical connotation, highlighting its reference to the physical rear (Oxford Learner's Dictionary, s.v. «ass»,).

Even with all the definitions given, the question remains, what exactly is swearing? According to the definition of Cambridge Dictionary, swearing is defined as the usage of

"rude or offensive language that someone uses, especially when they are angry" (Cambridge Dictionary, s.v. "swearing"). This means that more often than none, the speaker's emotions, mostly anger, are the ones to control the impulses for swearing. On top of that, being "rude" and "offensive" are two actions mostly done to hurt the listener. However, depending on who the speaker is talking to and swearing at, the intention may be changed. For example, if the speaker is swearing at a third person with the intention of agreeing with their friend, but that said person does not know, then no one gets hurt. Another example would be when the speaker is amidst relatives or someone close to them and they use swearing to bond, then the swearing would neither hurt nor offend someone. This case would be the same as how Love (2021, p. 740) described and defined swearing, where it is offensive and rude most of the time, but humorous and entertaining some of the time.

Other definitions from different authors within pragmatics may define swearing. Jay and Janschewitz defines swearing as "Swearing is the use of taboo language with the purpose of expressing the speaker's emotional state and communicating that information to listeners" (Jay & Janschewitz, 2008, p. 268). This raises the following question: If swearing is taboo, then why use it to express the speaker's emotional state? The most obvious answer to that would be that using taboo language is the best way to express strong emotions, and by using taboo language the speaker gets the attention that they want considering that the usage of such language does give attention to the speaker. The attention may not necessarily be positive, which is something that can be seen in the situation with Byrne (2017) and her parents' reaction to her usage of taboo language as a child. In their book "Bad Language"(1990), Lars Anderson and Peter Trudgill delve into the social and cultural aspects of swearing and bad language. They investigate how societal factors such as class, age, and gender influence language use, particularly concerning taboo words. Their work emphasizes that bad language is not purely a linguistic phenomenon but is intricately tied to broader social structures and norms. Michael Adams, in his book "In Praise of Profanity" (2016), presents a contrasting perspective on swearing and bad language. Adams challenges conventional notions by arguing that profanity serves various linguistic functions beyond mere offensiveness. He highlights swearing as a dynamic and expressive tool for conveying strong emotions, creating humour, and building solidarity among speakers. Adams (2016) advocates for a more nuanced understanding of profanity's cultural and social significance in communication. Keith Allan and Kate Burridge, authors of "Forbidden Words" (2007), also contribute to the discourse on taboo language and swearing. They explore how swearing, while often viewed as a breach of

social norms, plays a vital role in communication, allowing individuals to express emotions, release frustration, and establish social bonds. Allan and Burridge (2007) shed light on how society's perception of forbidden words reflects broader cultural and moral values, making swearing a subject of both linguistic and sociocultural interest.

2.3 Listing the Swear Words

The first of the research questions, RQ1, asks whether Love's pattern is accurate or not. To answer this question, I listed some of the swear words that he has used for his paper and used them in this thesis to compare with his. Love uses the forms of swear words as the starting point for his paper, and that is also what I will do in this thesis as well. Many swear words are used today. This thesis will focus on just a few of the most popular swear words used from the 1930s until the 2010s. Table 1 is a list over the swear words that this thesis will focus on.

Table 1: List of swear words that Love (2021) and I use in our papers

Love's Swear words	My Swear words
Fuck	Fuck
Shit	Shit
Crap	Crap
Arse (ass)	Ass
Dick	Dick
Bitch	Bitch
Bastard	Bastard
Bloody	Bullshit
Bollock	Fanny* (older term for "ass" or female genitals)
Twat	twat
cunt	pussy

The table above displays a selection of swear words employed by Love (2021) in his paper. However, it's worth noting that several other swear words, such as *bugger, cock, cunt, piss, shag,* and *wank,* are predominantly of British origin. In contrast, This paper aims to encompass a broader spectrum of six varieties of spoken English, hence the decision to exclude these primarily British swear words. Love (2021) uses the BNC in his paper, which focuses on one variety of spoken English, and that is British English. The reason that Love

uses this corpus is because his paper focuses on the usage of swear words in spoken British English, and this variety of spoken English is the area that he explores in his paper. The swear words chosen for this thesis are a mixture of the most popular swear words that are used in all the varieties of spoken English that the two corpora, the TV, and Movies corpus, provide.

As evident from Table 1, it is important to note that the selection of swear words utilized in this thesis differs from those employed by Love in his paper. This discrepancy arises not only from the distinct corpora utilized by Love (2021) and me but also from variations in research objectives. Love's (2021) study primarily centers on the frequency of swear words from the British variety of spoken English. This thesis on the other hand, aside from focusing on showing the differences and similarities of the different varieties of spoken English provided by the two chosen corpora, time is another factor that is the focus. In Love's paper, the difference in the usage of swear words between the genders and the age are also another difference in comparison to this thesis. The reason for this is due to the limited information that the corpus provides.

2.4 Robbie Love

Love (2021) discovered a pattern of speech with the usage of swear words in spoken informal English. This pattern is distributed into three separate sections: gender, age, and socio-economic status. When it comes to gender, Love (2021) realized that there is a pattern where females swear less than males and that it was not expected as the theory is that there is an equal possibility of swearing between the two genders, especially since freedom of speech for all was present. When it comes to age, the pattern is that there are fewer people that swear once they reach adulthood, with the reason of becoming a parent. There is, however, a high frequency of swearing among adolescents and between teenagers. When it comes to age, Love (2021) found a pattern of less swearing happening when people reach adulthood as a result of getting to work and making work relationships, such as when speaking to clients and customers, and as a result of becoming a parent. When people become a parent, they tend to want to set a good example in front of their children because children tend to imitate the adults around them. If the parents do not swear and curse in front of the children, then the children will not swear either. With that mindset, Love (2021) believes is the reason for the decline in swearing in adulthood, between the ages of late 20's and above.

2.5 Type of swear words.

There are times when swear words are used to exaggerate to make a point. For example, *stop fucking around!* This is another way to say *stop messing around!* The two sentences have the same meaning, but when swear words are involved, the impact that the sentence or in this case, the command given, gives has a much stronger impression on the listener. The way the speaker uses swear words depends on the point that the speaker wishes to make. The swear words in Table 1 may look different than each other, but some of them have the same meaning behind them. For example, the swear word *fanny* is an old word for the swear word *ass*. Many swear words occur by combining swear words that are already known with human or animal's body parts and things. For example, by combining the swear word *shit* with the animal *bull*, the results will be a new swear word *bullshit*. This makes the swearing more specific and much more imaginative.

"This is just a load of bullshit." This sentence is another way of saying "*This is silly* or *this is just nonsense."* There are many times when simple words in sentences, mostly commands and exclamation sentences, are taken out of the sentences and replaced by swear words. This is done so that the sentence would have a stronger impression and impact on the listener. People usually pay more attention to swear words rather than simple "nouns" or "adjectives" and swearing causes a bigger scene than other words.

Swearing manifests through two distinct modes: the Literal Usage and the Figurative Usage, each delineated by specific linguistic characteristics. The Literal Usage aligns with the examples, where swear words are employed in their standard and direct meanings. Alternatively, this mode can be described as instances where profanities can be substituted with non-taboo near-synonyms. Certain scholars question the categorization of such cases as swearing, positing that their interchangeability with less offensive terms dilutes their classification. The Figurative Usage of swearing resists facile substitution with non-taboo near-synonyms, solidifying its status as a form of swearing. This mode encompasses instances where profanities are harnessed to convey emotions, intensify expressions, or punctuate rhetoric. Due to the inherent difficulty in replacing these terms with benign alternatives, the figurative usage retains a distinct potency that aligns more closely with the concept of swearing.

This nuanced categorization finds scholarly validation in linguistic literature. A seminal work by Timothy Jay and Kristin Janschewitz titled "The Pragmatics of Swearing" (2008) examines the multi-faceted nature of swearing, delving into its literal and figurative dimensions. The authors highlight how the figurative usage, often marked by emotional or idiomatic intent, distinguishes itself from the more straightforward literal usage. This insight substantiates the differentiation proposed here between the two modes of swearing, grounded in the intricate interplay of linguistic and pragmatic factors. In summary, the bifurcation of swearing into Literal and Figurative Usages finds credence in scholarly discourse, particularly in works such as Jay and Janschewitz's comprehensive exploration of swearing's pragmatic nuances.

Different times mean different styles. The society that we live in changes its ways thoughout history. From worshipping nature and different gods to believing in one religion, and then to move on evolving that religion. With society changing over time, so does the attitude of the people in that society. This leads to the perception that how we speak depends on the society that we live in, which leads to the next perception. The society we have is reflected in our swearing. This in turn reflects on the way we swear and on the timing of our swearing.

There are many different types of categories that we can split swearing and swear words into. One example of these categories is in religious context. There are times when the way we swear is done in relation to religious context. Swearing according to religion depends on the kind of society that we live in. Some of the swear words related to religion are, for example, *hell, Christ, Jesus, heaven, holy,* etc. This kind of swearing is mostly used when people are shocked, in pain, or when they want to express disbelief. There are times when the swear word is shortened for example, *Jesus* becomes *gee*. Or when the swear word is used in connection with another word, like for example the word *holy* can be used with the other swear word *shit*. Swearing encompasses various categories, extending beyond its connection with religion. Another significant category of swear words revolves around sexual and bodily functions, often regarded as among the most offensive forms of swearing. This classification stems from the potent and taboo nature of these words, which frequently evoke strong emotional reactions. Examples of such profanities include terms like *fuck, bitch, twat, ass, balls, prick,* and *cunt*. These words target intimate aspects of human anatomy and sexuality, directly challenging societal norms and sensitivities. Their explicit and graphic nature contributes to their classification as highly offensive swearing. The categorization of sexual and body parts swearing as one of the most offensive types finds validation in linguistic and

sociolinguistic studies. For instance, Steven Pinker explores the dynamics of taboo language and its cultural implications in his book "The Stuff of Thought" (2007). Pinker's (2007) analysis underscores how the use of explicit sexual and bodily terms often elicits strong reactions due to their transgressive nature. This aligns with the perception of these swear words as particularly offensive. The classification of sexual and body parts swearing as one of the most offensive types of swearing is grounded in linguistic analysis and scholarly discussions (Steven Pinker, 2007). This understanding acknowledges the emotional, cultural, and societal implications of these explicit and provocative words (Pinker, 2007). But at the same time, some of these swear words have been documented for their use for bonding between friends. The swear words *bitch* and *twat* are widely popular and are known to be used on many social occasions between friends. The example below gives two similar scenarios where the two swear words are used in such social interaction between friends.

Example:
A: That's just how I roll, girl.
B: *Bitch,* please. You wish!
A: Yo, *twat*! Where you at?

Even amid the categories previously outlined, one specific category stands out as the most universally offensive across all genders, social contexts, and age groups. This category pertains to racism and is uniquely positioned to inflict deeper levels of insult and emotional harm compared to other swearing categories. The subject of racism constitutes a substantial and ongoing discourse, spanning multiple dimensions and still under active consideration today. To facilitate a focused analysis within this thesis, I will partition this expansive category into three distinct facets: gender, race, and sexuality. Table 2, presented herein, provides illustrative examples of swear words pertinent to each subset within the broader realm of racism. This classification finds support within scholarly discussions. For instance, Deborah Cameron in her work "Verbal Hygiene" (1995) expounds on the dynamics of language and power, highlighting how racist language operates as a mechanism of discrimination and perpetuates unequal power dynamics. Cameron's exploration validates the assertion that racist swearing holds a unique potency, inciting harm on a different level compared to other swearing categories.

Figure 1: List of examples of swear words for each of the categories within racism.

Gender	Race	Sexuality
Bitch, whore, Dick, cunt, Bastard, pussy, sissy etc.	Nigger, gypsy, Ching Chong, terrorist (religion, Muslims) etc	Fuck, faggot, homo, fag, balls, etc.

Although the list of swear words listed in the gender category shown in Figure 1 is gender-specific, it is also sexual in nature. Robbie Love (2021) reports that the likelihood of swearing between genders is low when done across gender lines, but it occurs more naturally within the same gender. He also found that males tend to swear more than females. However, the discovery he made was that, overall, the usage of swear words, according to the BNC corpus, was lower in the 2010s compared to the 1990s.

2.6 Grice's Maxims of Conversation

There is a linguist named Paul Kroeger. What he wanted was to try to find the relation between saying and meaning. He analyzed the meaning of phrases and sentences that were said and divided them into their categories. An example of such categories is the speech acts. This is where the addressee, or the hearer, must answer three questions(Kroeger, 2018, pp. 9–10). The speech act is an utterance that depends on the speaker's intentions and the effect or meaning it has on the hearer or listener (Cole & Morgan, 1975). Because Kroeger investigates the relation between a saying and a meaning, the best way to show it is by using one of his examples (Kroeger, 2018, p. 10).

Example 1:
   A. *Please* pass me the salt.
   B. Can you *please* pass me the salt?

In the following example, Kroeger (2018) explains that although the word "please" does not contribute semantically, it does have an important role in the utterance meaning and for the hearer. It does, however, play a role in terms of politeness. The same theory may be applied to the usage of swear words as well.

What Kroeger tries to point out is that language has some specific rules and conditions that language follows and that should there be a time that language does not follow these rules and conditions, what can the results be in these cases? One thing that needs to be made clear is that these maxims are there to point out characteristics of language. This means that so long as the speaker can communicate with the hearer, by either obeying the maxims or breaking them, there are various ways to effectively convey meaning. "A speaker may communicate either by obeying the maxims or by breaking them" (Kroeger, 2018, p. 142). This flexibility in communication strategies underscores the dynamic nature of language and the adaptability of speakers in diverse contexts.

Another one of these categories is the theory of the relation between a saying and implication. This theory is, however, a citation from another person. This person is a scholar, and he is named Grice. According to Kroeger (2018, p. 141), "Grice was not only the first scholar to describe the characteristic features of implicatures, but also the first to propose a systematic explanation for how they work". His maxims of conversations are about the conditions that need to be met for us to investigate the relation between what is said and what is implied.

> Billy: Can you please pass me the salt?
> Sarah: The sauce is in the way.

The example above shows one of the things that Grice and Kroeger point out (Grice, 1975; Kroeger, 2018). In the example, Billy wants the salt and asks Sarah to pass it. Sarah's replay was an implicature, as what we might have wanted her to say was 'No, I cannot'. There are times when a direct answer to a question may be taken as rude and give off a negative feeling to the hearer. In these cases, giving out an answer with an implicature could allow the speaker to imply the answer in a way that would neither insult nor give off a negative reaction in the situation. In the case of the example above, by Sarah telling Billy that 'The sauce is in the way', Sarah is trying to refuse Billy's request in a way that does keep the politeness between them. Instead of the blunt 'No', the implicature gives the hearer, who in this case is Billy, a soft denial of the request.

Grice's Maxims of Conversation theory, reported by Kroeger (2018, p. 142), is divided into two main parts. The first part is The Cooperative Principle (Grice, 1975, p. 45) theory. This

theory is explained by Kroeger (2018) "Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged". One of the ways that this statement may be interpreted is that when you give a response or say something, the best way that you do it is to make it on point, answer with accordance to the topic, and state what you wish to say in accordance with the situation. According to Kroeger (Kroeger, 2018, p. 142), this theory is meant to be a kind of background assumption between the speaker and the hearer where they both can find a common ground.

Next, I will address the second main part of Grice's theory, which is called the Maxims of Conversation (Grice, 1975, pp. 45–46; Kroeger, 2018, p. 142). This theory is divided into four different parts: the maxim of quality, the maxim of quantity, the maxim of relation or relevance, and the maxim of manner. Each of these maxims has its own set of guidelines that should be followed for the maxims to be fulfilled. Grice's theory lays the groundwork and is of paramount importance for the topic of politeness.

2.6.1 Maxim of Quality

Before we start with the investigation, we need to explain what these maxims entail, and what their requirements to be fulfilled are. We start with the maxim of Quality. This maxim is the maxim that talks about the truth and contains two requirements for it to be fulfilled.

1. The first requirement for this maxim is to not say anything that the speaker believes to be false.

There is no point in giving out information to the hearer if the speaker already knows that that information is false and could lead the hearer. The following example will point out exactly how this maxim may be observed and how the speaker can break it.

Example 2:
A: What time is it?
B: It is 3pm.

If B already knew that it was 2pm but answered that it was 3pm instead then B is stating something that they know and believe to be false. The condition of this maxim is, therefore,

broken when B gives a false statement. If the aim is to fulfil this condition of the maxim, then all B had to do was say the truth which is that the time is 2pm.

2. The second requirement is for the speaker to not say anything that lacks adequate evidence for.

This requirement speaks for itself. The difference between the first requirement and this requirement is that in the first the speaker knows the answer but chooses to give an incorrect one or knows the right knowledge but chooses to give out incorrect knowledge. This requirement, on the other hand, is about when the speaker does not have the exact correct answer to the topic spoken but still gives a detailed answer not knowing if that is the correct answer or not, then the speaker breaches this condition of the maxim.

A: What time is it?
B: It is 4:33pm?

The example above demonstrates one of the possible scenarios where the speaker breaks this condition, and consequently, this entire maxim. If we say that in this scenario, B is sure that it is sometime in the afternoon but is not sure of the exact time and still gives out a specific answer, that means that B is breaking this requirement and therefore, breaching this maxim as a result. If B wanted to follow this maxim in compliance with the cooperative principle, then B could have answered in a differently. If B had answered that it is afternoon without giving a specific time, then B would not have breached this condition and this maxim would have been fulfilled.

2.6.2 Maxim of Quantity

1. Make your contribution as informative as is required (for the current purposes of the exchange). This means that you should share enough information to fulfill the needs and expectations of the conversation. You shouldn't leave out crucial details if they are necessary for understanding or context.

Example 3:

When someone asks, "What's the weather like?" on a rainy day, a good response would be, "It's raining heavily, and there's a thunderstorm warning." This response provides informative details relevant to the weather question.

2. Do not make your contribution more informative than is required. This implies that you shouldn't overload the conversation with excessive information that's unrelated or unnecessary. Stick to what's relevant to avoid overwhelming or confusing the listener.

Example 4:

If someone asks, "What time is it?" responding with, "Well, it's 2:15 PM, and the train to downtown leaves in 10 minutes, and I had a delicious sandwich for lunch" would be violating this maxim. A simple "It's 2:15 PM" suffices in this context.

2.6.3 Maxim of Relation

Another name that this maxim is known for is the maxim of relevance. This maxim is the one maxim out of the rest that is the most straightforward. It is simple and contains only one requirement from the respondent, BE RELEVANT. Any element in the equation that is not relevant to the topic must be cut out.

A: Would you like to join me to the concert?
B: Look, there is a hippo!!

In the example provided, it is evident that the maxim of relevance is violated in terms of the usage of swear words. When person A asks if person B would like to join them at a concert, person B's response, "Look, there is a hippo!!" is entirely unrelated to the topic of the concert invitation. This lack of relevance in B's response constitutes a violation of the maxim of relevance, as it does not contribute to the ongoing conversation or address A's question effectively.

2.6.4 Maxim of Manner

For this maxim, the most important thing to note if we wish to observe here is to be perspicuous. What is meant by that is for the speaker to be clearly expressed and easily

understood in the communication to the hearer as that will lessen or prevent any possibility for a misunderstanding or miscommunication to happen between them. This maxim has four requirements that needs to be followed for either the speaker or respondent to fulfill the maxim.

1. The first condition is to avoid obscurity of expression. Another way to interpret this condition is by explaining it as such, to avoid uncertainty of meaning or expression. In other words, the hearer has to be able to understand what is meant clearly. Another way to interpret this is thinking like Grice. The way he interpreted 'obscurity' is that speakers should in general use words and phrases that are easier to understand rather than harder ones. One of the examples that he gives is a situation where it is possible to break this condition of the maxim. Supposing two adults are speaking to each other, A and B, and there is a child present in the room or within hearing, C, the adults, A and B, would then try to avoid using the same words that are familiar with the child, C, or even start to use words and expressions that are unknown to the child. We can say that A and B are both deliberately obscure and hope that the child does not understand the contents of A and B's conversation. This way the speaker might use a word that the child probably hasn't yet learned, or a word in a foreign language, or a phrase that is hard to interpret.

Example:
A: This way we'll **shoot** two birds with one stone.
B: It's not going to be easy.

The example above shows how 'obscurity' may be used in a conversation. We can see that A and B are the ones engaging in a conversation. Let us say that there is a third party in the room. A and B can relate, the deliberate use of shoot instead of kill.

2. Avoid ambiguity. What is meant by this is to avoid anything that has the quality of being open to more than one interpretation. Through this condition of this maxim, the speaker must not leave any room for misunderstandings to the hearer.
3. Be brief!
4. Be orderly. (neat)

2.7 Breeching of the maxims

Why do we breach the maxims? Considering that the maxims are clear, why do people then keep breaching the maxims? There is a reasonable explanation for this. The phenomenon of breaching Grice's maxims is a complex one, even though the maxims themselves might appear clear on the surface. This deviation from literal adherence is influenced by various pragmatic, social, and psychological factors. Authors like Deborah Cameron and Timothy Jay, among others, provide valuable insights into the reasons behind these breaches and their relevance to swearing.

Deborah Cameron, in her work "Verbal Hygiene" (Cameron, 1995), highlights how language use reflects and reinforces social norms. When it comes to swearing, breaches of conversational maxims can be tied to larger power dynamics. Cameron argues that language serves as a tool for the negotiation of power relationships within society. Swearing, as a form of linguistic rebellion, often involves the breach of maxims to challenge established norms and assert individual or group identities. Timothy Jay and Kristin Janschewitz, in "The Pragmatics of Swearing" (2008), delve into the pragmatic, emotional, and social motivations behind the breach of conversational norms, particularly in the context of swearing. They emphasize that swearing is not just about linguistic transgression but also a communicative strategy with specific goals. Swear words, often breaching maxims for the sake of emotional impact or emphasis, can serve as potent tools for conveying strong emotions or challenging listeners' assumptions. Considering these perspectives, the breach of maxims becomes especially relevant within the realm of swearing. Swearing is characterized by heightened emotional content and distinct communicative goals. Norman Fairclough's insights from "Language and Power" (1989) also resonate here, emphasizing how language is intricately tied to power dynamics. Swear words can breach the maxims to exert dominance, provoke reactions, or subvert social norms. Such breaches not only serve the pragmatic function of achieving communication goals but also reflect the broader power struggles and social dynamics within which swearing occurs.

The breach of Grice's maxims, as illuminated by authors like Cameron (1995) and Jay (2008), finds relevance within the context of swearing due to its intricate connection with power dynamics, emotional expression, and communicative strategy. These scholars shed light on

the multifaceted motivations behind such breaches and how they contribute to the nuanced landscape of linguistic interaction.

2.8 Politeness and impoliteness

Politeness refers to the use of social conventions and manners to show respect and consideration for others. It involves using polite words and expressions, maintaining appropriate body language, and following cultural norms and expectations for social behaviour. Politeness is important in building positive relationships with others, as it shows that you value their feelings and opinions. It can also help to prevent misunderstandings and conflicts by communicating clearly and respectfully (Saeed, 2016).

Examples of polite behaviour include using "please" and "thank you" when making requests, holding the door open for others, using appropriate titles and honorifics when addressing someone, and showing interest and attentiveness when speaking with others. Overall, politeness is an important aspect of social interaction that can help to promote understanding, respect, and harmony between individuals and groups. Politeness is one of the main factors to mention when talking about swearing. As it was established previously, swearing and the usage of swear words are considered bad or foul language. Bad language is, therefore, considered to be the opposite of politeness, which makes swearing and the usage of swear words threaten the concept of politeness (Saeed, 2016).

Impoliteness, on the other hand, refers to the use of language or behaviours that violate social norms, conventions, or expectations of politeness. It involves intentionally causing offense, disrespecting others, or challenging social norms for various purposes, such as asserting power, expressing frustration, or establishing social distance.
Impoliteness can take many forms, including direct insults, offensive language, rudeness, sarcasm, mocking, and aggressive behaviour. It often involves face-threatening acts that challenge or undermine someone's positive or negative face. Unlike politeness, which aims to preserve social harmony and positive relationships, impoliteness can disrupt social interactions and create tension or conflict. Impoliteness strategies can be influenced by cultural, situational, and individual factors. Cultural norms and expectations shape what is considered impolite in different societies. Situational factors, such as power dynamics, context, and social relationships, can also influence the perception and interpretation of

impoliteness. Understanding impoliteness is important in studying communication and discourse analysis. It helps researchers analyse how language is used to express and negotiate power, conflict, and social dynamics. Additionally, understanding impoliteness can provide insights into how individuals navigate social interactions and manage face-threatening situations in various contexts.

Here's an example of impoliteness in linguistics:
Person A: "I just finished reading that book you recommended. Honestly, it was quite boring."

In this example, Person A is expressing their opinion about the book recommended by Person B. However, the use of the word "boring" directly and negatively can be considered impolite. It violates the social norm of politeness, as it can potentially hurt Person B's feelings or undermine their taste in books. Instead, Person A could have expressed their opinion more politely by using mitigating language or focusing on aspects they didn't enjoy without directly labelling the book as "boring". Impoliteness can vary in intensity and formality depending on the context and cultural norms. The example above demonstrates a relatively mild form of impoliteness, but more severe instances can involve insults, offensive language, or personal attacks (Saeed, 2016).

In linguistics, a face-threatening act (FTA) is a type of speech act that threatens someone's social identity, self-esteem, or reputation. FTAs can occur in various forms, such as criticisms, accusations, rejections, and requests that challenge someone's autonomy or social status. FTAs are considered a fundamental aspect of face theory, which explains how people interact and maintain social relationships through face-saving behaviours. When an FTA occurs, the speaker risks damaging the listener's face, which can lead to discomfort, tension, and conflict (Brown & Levinson, 2016).

Examples of FTAs include:
- Direct criticisms or insults that attack someone's competence, intelligence, or character.
- Requests that impose an obligation or expectation on the listener, such as asking for a favour or help.
- Refusals or rejections that challenge the listener's autonomy or sense of control.

- Disagreements or contradictions that challenge the listener's beliefs or opinions.

To minimize the risk of face-threatening acts, individuals often use language and behaviours that preserve positive social relationships and show respect for others' faces. This can include using politeness strategies, indirect speech acts, or mitigating language that softens the impact of the FTA. The concept of face-threatening acts is important in understanding social interactions and communication, as it highlights the importance of face-saving behaviours and strategies that people use to mitigate the impact of FTAs. For example, people may use humour, apology, or avoidance to manage face-threatening situations and protect their social identity and reputation.

Understanding face-threatening acts is important in effective communication, as it allows individuals to navigate social situations with sensitivity and respect for others' face needs. By being mindful of FTAs, individuals can use language and behaviours that preserve positive social relationships and minimize the risk of negative social interactions. Overall, understanding face-threatening acts can help individuals navigate social situations more effectively and build positive relationships with others by being mindful of how their actions and words may impact others' face needs.

The Face-threatening Act (FTA) mentions two different faces, the negative face, and the positive face. Both faces affect the speaker and the hearer in certain degrees. But before we get into the effect of the two faces, we need to first understand what the two faces are. First, the negative face. In linguistics, negative face refers to a person's desire to be free from imposition, interference, and obligation from others. Negative face is a fundamental aspect of face theory, which explains how people interact and maintain social relationships through face-saving behaviours.

A negative face can be expressed in a variety of ways, such as through indirect requests, hedging, and politeness strategies that allow the speaker to preserve the listener's autonomy and sense of control. For example, using phrases such as "If it's not too much trouble" or "Would you mind" can mitigate the imposition of a request and reduce the threat to the listener's negative face. The violation of a negative face can cause face-threatening acts and lead to social discomfort, tension, and conflict. For example, making a direct request or giving

unsolicited advice can be perceived as a threat to the listener's negative face and result in resistance or rejection (Saeed, 2016, p. 464).

Understanding negative face is important in effective communication, as it allows individuals to navigate social situations with sensitivity and respect for others' autonomy and needs. By being mindful of negative face, individuals can use language and behaviours that preserve positive social relationships and minimize the risk of face-threatening acts.

Next, the positive face. In linguistics, positive face refers to a person's desire for social recognition, approval, and inclusion. Positive face is a fundamental aspect of face theory, which explains how people interact and maintain social relationships through face-saving behaviours. Positive face can be expressed in a variety of ways, such as through compliments, agreement, and solidarity with others. Positive face is important because it allows individuals to feel valued and respected in social situations, and it reinforces positive social relationships. Maintaining positive face can involve using language and behaviours that show appreciation, recognition, and interest in others. For example, asking for someone's opinion or thanking them for their contribution can increase positive face and strengthen social bonds.
The violation of positive face can cause face-threatening acts and lead to social discomfort, tension, and conflict. For example, criticizing someone or dismissing their ideas can be perceived as a threat to their positive face and result in defensiveness or resentment. Understanding positive face is important in effective communication, as it allows individuals to build positive social relationships and avoid face-threatening acts. By being mindful of positive face, individuals can use language and behaviours that enhance social rapport and minimize the risk of negative social interactions (Saeed, 2016, p. 244).

Chapter 3: Corpus

3.1 What is a corpus?

In this chapter, I will discuss how I plan to investigate the research questions. Before I begin my investigation, I will need a method that will help me find the data that will help complete my investigation. Considering that the data required for the investigation consists of changes over time and geographical placement, the method chosen must be able to provide the thesis with data that includes dates to the period and place required for my research questions to be answered. The method that I chose is, therefore, something called for a corpus. This is the tool or program that I use in this thesis to gather the data and investigate the research questions. A corpus is a collection of written texts, and scripture of spoken language, that is collected into an electronic program that linguists or anybody can use to investigate a language or a pattern existing in a language.

The place that I use to find my corpora is called English-corpora.org. This website provides access to many different corpora, such as the BNC (British National Corpus), the COCA (Corpus of Contemporary American English), etc. My thesis research is based on data about swearing that can be found from the 1930s to 2010 in the spoken corpus. Therefore, the two corpora that I have decided to use for my investigations in this thesis are the TV corpus and the Movie corpus. Another reason why these two corpora were chosen is that my thesis investigates swear words in the informal spoken English language just like Love's (2021) paper did. Love used the BNC from the 1990s to 2014. The BNC is a corpus that includes transcript data from all kinds of spoken conversations, TV, movies, dialog, etc. But the TV corpus and Movies corpus focus on just them.

3.2. What is the size of the corpus?

The TV corpus consists of 325 million words. It consists of data gathered from TV shows from the period of 1950s to the year 2018. The data from this corpus is gathered from 6 English-speaking countries. That means that this corpus provides data from 6 different varieties of English spoken language. The 6 varieties of English are US (American English), CA (Canadian English), UK (British English), IE (Irish English), AU (Australian English) and NZ (New Zealand English). These 6 different varieties of English are divided in the corpus by geographical placements. For example, the results of the search for the US verity of

the English language, the search results for the CA verity are included. The reason is because of their geographical placement. The two verities, the US, and CA are considered neighbours in terms of their geographical placements. The same is true the UK verity of English and the IE verity of English. When the results for either one are requested, the results for the other variety are included in the search.

The Movies corpus, on the other hand, consists of only 200 million words. The data from this corpus consists of transcripts from the period back from the 1930s to the year 2018. This is one of the differences that the two chosen corpora have. The Movies corpus has data gathered from 20 years earlier than the TV corpus. The data from the Movies corpus is gathered from 6 different English-speaking countries, the same as in the TV corpus. The different English varieties in this corpus are the same as the ones presented for the TV corpus, which are US/CA, UK/IE, and AU/NZ. One of the reasons the Movies corpus has data from the 1930s is that the movies were available to the public easier than the Television.

3.2.1 What can be found in a corpus?

Two types of information or data can be found in these two corpora, **external linguistic** information, and **internal linguistic** information. I will start by explaining what external linguistics data is. What is meant by external linguistics is the external data that we can find using the corpus, such as topics on gender, age, etc. For example, the lemma or search word *fuck*. By searching up this lemma, we can find out the gender of the search results. We can also find out from the search results which period each gender has been recorded for using the swear word, as well as the placement of the swear word in the sentence that the results shown. But because of the limitations and the limited time frame the topic of gender will not be included in my investigation.

Internal linguistic information, in contrast, encompasses data that delineates, clarifies, and enables the scrutiny of language dynamics. This data operates through the grammar of the search word, the functional role of the lemma within sentences based on search results, the frequency, and occurrences of the search lemma, and more. These facets facilitate the exploration of various linguistic topics rooted solely in language-based aspects. For instance, by examining the lemma *twat*, insights can be gleaned about its historical usage patterns, potential shifts in spelling, its frequency in both spoken and written corpora and other

language-centric attributes. Internal linguistics allows for an inquiry into the inherent structure and nature of language itself, distinct from external linguistic considerations tied to societal influences. In external linguistics, the impact of society on language is explored. For instance, studies often reveal how words associated with femininity, like "girl," are linked to notions of appearance and sexuality, while words associated with masculinity, like "boy," are connected to concepts of activity and strength. In internal linguistics, the focus shifts to investigating linguistic configurations and changes within language itself. For instance, it enables us to ascertain whether a word has undergone alterations over time, if its frequency has diminished, or if a new word has emerged as an evolution of an older term. To supplement this understanding, the work of Penelope Eckert and Sally McConnell-Ginet in their book "Language and Gender" (2013) offers insights into the dynamic interplay between language and social constructions of gender. Their study underscores the importance of internal linguistic analysis in understanding how language evolves intrinsically, shedding light on linguistic shifts that transcend societal influences.

However, there is one thing that cannot be found or will be hard to investigate using the two chosen corpora, the TV, and the Movies corpus. One of the disadvantages that I get using these two corpora is that I cannot locate a specific search result for the gender of the search words. Being able to tell which gender the search results refer to is an advantage, however, one of the reasons that this is a disadvantage is that the results do not group the results and do not show for example the grouping of instances where people using the swear word *fuck* are males and how many of the instances are referring to females. Considering that the two corpora, the Movies corpus, and the TV corpus, are corpora that depend on scripts, it will be hard to determine if the search is pointing at a specific gender, be it male or female. For us to find this information in the two chosen corpora, we would have to go through every single search result for the lemma. The same is true the topic concerning age. The ages of the people in the search results of the lemma *fuck*, do not show the grouping of the ages of the people in the transcripts. If the topic concerning age needs to be investigated, each example and transcript provided by the corpora and its search results must be examined and that is time-consuming. However, since both corpora have the data gathered from TV and movies, and both are scripted it is difficult to accurately point out the data. The characters are simply following the script that another person has written for them. In these situations, it is possible that there is more than one writer that is involved and is responsible for each character, and the gender of these writers determines their writing, and this affects the ability to accurately

analyse the data for both gender and age. As this study focuses mostly on investigation relevance to time, frequency, and collocation, gender and age become somewhat irrelevant.

3.2.2 What is possible to do with the corpus?

Even though there are some disadvantages to using the TV corpus and the Movies corpus, there are, however, a lot of things that are possible to do and investigate using these two corpora. One of the things that we can do using the corpora is to find out which swear words are most frequently used. The frequency of the swear words found in the corpus can provide us with the data needed to figure out which of the swear words are most frequently used. For example, we can find out if the swear word *shit*, is more frequently used than the other swear word *crap*. This raises the question of why this is the case, considering the meaning of both swear words are the same.

Another thing that is possible to investigate using these corpora is to find out what the function of the swear words is in different sentences. In the corpora, it is possible to find out using the search results from a lemma to discover the grammatical function of the lemma. Does the lemma have a specific role? For example, *Hey!* **Bastard***, what are you doing?* This example shows that the swear word *bastard* is a vocative realized by a noun, which answers the question above, that yes, the swear word may have a specific role to act in a sentence. But does it always have a function? We need the corpus to investigate and answer questions such as those.

The corpora also provide the information on which and what year the word that is being investigated, occurs in and what the frequency of the intended word is. Using the frequency function in the corpus allows for the investigation of possible disappearance of a lemma, and a possible appearance of a new lemma based on the old lemma. A change of the lemma over time becomes a possibility for investigation as well, for example, that the lemma is spelled differently than it is originally spelled. It is also possible to investigate if there are any changes in the conjugation of the lemma. The corpus allows information on the lemma and the occurrences that the lemma occurs with. Using the corpus, we can investigate which words follow or come before the intended lemma. For example, the search word is *fuck*, after investigation, the following word is known to be *off*. But the corpus also allows for the investigation of words connected with the intended search word. For example, the swear word

shit connected with the word *bull* makes for the swear word *bullshit*. The corpus is not limited to animals. For example, the swear word *shit*, conjoined with the word *face* gives the swear word *shitface*. It is possible to investigate all these combinations using the two chosen corpora.

3.2.3 What does the data look like?

The corpus that I use has many functions that are useful for the investigation of both frequency and collocation. One of these functions is concordance. This can be found by simply typing in the search word or lemma in the search field and activating the search. If we wish to narrow down or expand the search, it is possible in the 'sections' field. Using this function, we may be able to control the search to make the investigation as relevant as possible. For example, we may be able to find all occurrences available in the corpus related to the topic of the thesis, which in this case is swear words. By typing in the swear word in the search field, the result for the different occurrences will show. We are also able to control the position of the searched swear word, for example, by instructing the settings to show the occurrences of five occurrences to the left or right of the sentence. This allows for the investigation of functions for said swear word in different sentences, which in return, allows for an investigation of a discovery of a pattern in the language, which in the case of this thesis is the informal spoken English language.

Another question to answer would also be, what can I search for in this corpus? The word that is searched in the corpus is often referred to as a Lemma. Through this corpus, it is possible to investigate several lemmas and the changes that have happened to those said lemmas over time. Another thing that we may find is the gender and age. What is meant by that is that it is possible to search for which age and which lemmas are most relevant through this investigation. This process is, however, very time-consuming as it requires us to go through every search result in detail to find the age and gender of the search results.

The data that can be found in these two corpora will be explained next. As mentioned, this thesis uses the two corpora from the TV show genre and the Movies genre. The data provided in these two corpora go back to the early 1900s, which makes the investigation over time a possibility. Another advantage of the two chosen corpora is that there are 6 different varieties of spoken English available. This advantage opens the opportunity for the investigation of

geographical placement. But just like there are advantages to the chosen corpora, there are also a few disadvantages.

3.3 Why was this corpus chosen?

This thesis will give a comparison on Robert Love's paper, *Swearing in Informal Spoken English: 1990s-2010s*. In his paper, Love (2021) uses the BNC (British National Corpus) as he focuses on the British English variety. He also focuses on the forms of the swear word in his list. The BNC consists of data from a mixture of genres, spoken, fiction, newspaper, academic, non-academic, etc. This means that Love's data was taken from different genres and that his findings were based on a mixture of them, which in one way makes the data from Love's paper, vast and a little generalized.

This thesis, on the other hand, is more specific. That is why the data is gathered from two genres only, the TV genre and the movies genre. This narrows the search and specifies the topic and makes the investigation of the different swear words more relevant., The TV and Movies corpus was one of the reasons I chose to make my investigation similar to Love's (2021) theory. Another reason why these two corpora were specifically chosen for the data investigation for this thesis is that the thesis focuses on the informal spoken English language. Although Movies and TV shows are mostly scripted, especially the movies, however, the shows are written so that there will be a resemblance to a natural flowing dialog. There are often many times when the actors or entertainers improvise, and it becomes a natural dialog embedded into the already scripted one as well. This situation is more common on TV variety shows rather than series or in movies.

3.4 What variables does are allow for investigation?

The usage of a corpus allows for an investigation of many different topics over many genres. For example, we can investigate the form of a word or the grammar changes that have happened to the specified word in fiction novels. This allows for the investigation of different topics in the fiction genre.

Another variable that the corpus allows for investigation is changes over time. Depending on which corpus is used, we can investigate if a word or a lemma has changed over certain periods. The changes may be grammatical, as the word itself has changed the way it is being

written or conjugated, or it can be in the function and the role it has in a sentence. Another type of change over time that the corpus would allow for investigation is the pragmatic change. For example, the swear word *bastard*. This swear word was used to call an illegitimate child, born out of wedlock, in the past. Over time, the way this word is used is for insulting or degrading someone, or in a negative way. Not to mention that the investigation of such words in the corpus can be done over many genres. For example, we can investigate the usage of the word *bastard* in TV and Movies. How the swear word is used, when it is used, and what pragmatic function and meaning it brings, are some of the questions that can be answered using the corpus.

Different corpora may allow for different varieties of the spoken languages that are being investigated. As mentioned before, this thesis makes use of the TV corpus and the Movies corpus, which allows for the investigation of informal language in 6 different varieties of spoken English language. By having more than one variety of spoken English available for the investigation of a topic, another variable that the corpus allows for investigation comes up. That variable is the ability to make a comparison between the 6 different varieties of spoken English language. By making a comparison between the 6 different varieties, another question opens for investigation; why does one language have so many different varieties?

Considering that the investigation of the comparison of the 6 varieties of spoken English is allowed using the two corpora, another variable is allowed for investigation based on a continuation of the last variable. A comparison of geographical placement is the variable that opens for investigation based on the variable mentioned above. The 6 different varieties of spoken English in the TV corpus and the Movies corpus come from 6 different countries and three different continents. Because the English varieties come pairs in the results, the placements that are open for investigation are within the three continents, America, Europe, and Australia/Oceania.

Chapter 4: Methodology

This study utilizes a quantitative research method to analyse a specific corpus of data. The choice of this corpus and the quantitative approach is based on several factors, including the research objectives, the availability of data, and the feasibility of conducting the study. This section will provide an overview of the rationale behind the selection of these corpora, the quantitative method employed, and a general observation justifying their usage.

4.1 Rationale for Corpus Selection

The selection of a corpus is crucial as it forms the foundation of the study and provides the data necessary for analysis. In this research, the chosen corpora were selected based on their relevance to the research objectives and the availability of suitable data. As mentioned before, the chosen corpora are the TV corpus and the Movies corpus. These two corpora may be found on the website english-corpora.org. The TV and movies corpora provide the data extracted from TV shows and movies dating back from the early 1900's to the 2010's period. (See Chapter 3 for more information on the corpora.)

The selected corpora offer a diverse range of information related to the research topic, encompassing various geographic locations and periods. The inclusion of different geographical regions allows for a comprehensive analysis of the phenomenon under investigation, considering potential variations and regional influences. Moreover, considering multiple periods enables the examination of temporal trends and changes, providing valuable insights into the research question.

4.2 Justification for Quantitative Method

The decision to employ a quantitative research method stems from its suitability for addressing the research objectives and analysing the selected corpora. Quantitative methods involve the use of numerical data and statistical techniques to identify patterns, relationships, and trends. This approach facilitates the objective analysis of large datasets and enables researchers to derive generalizable findings. By utilizing a quantitative method, this study aims to uncover empirical evidence and establish statistical relationships within the selected corpora. This approach allows for the identification of significant patterns or associations between variables, providing valuable insights into the research topic. Additionally, the quantitative method provides a structured framework for data analysis, enhancing the rigor and replicability of the study.

4.3 General Observation and Justification

Based on a preliminary examination of the selected corpora, it is evident that they contain a substantial amount of relevant data for the research question at hand. The available information within the corpora exhibits a wide range of geographic locations and temporal coverage, thereby offering a comprehensive representation of the phenomenon under study. Furthermore, the quantitative nature of the data within the corpora aligns well with the objectives of this research. The availability of numerical and measurable variables allows for quantitative analysis, enabling the exploration of patterns, trends, and statistical relationships. By employing statistical techniques, the study aims to provide empirical evidence and contribute to a deeper understanding of the research topic. In conclusion, the selection of the chosen corpora and the adoption of a quantitative research method are justified by their alignment with the research objectives, the availability of relevant data, and the potential for generating robust and generalizable findings. These decisions will enable a comprehensive analysis of the selected corpora and facilitate the achievement of the research objectives.

4.4 Why these corpora

The choice of using a TV and movie corpus for studying swear words is justified for several reasons:

1. Records of Speech: TV and movies provide a rich source of recorded speech, capturing authentic, natural language use in various contexts. Swear words are primarily spoken rather than written, making speech-based corpora more suitable for studying their usage, context, and sociolinguistic aspects.

2. Diverse Texts and Genres: TV and movies encompass a wide range of genres, settings, and characters, allowing for a diverse analysis of swear words across different contexts. By including multiple genres such as dramas, comedies, documentaries, or reality shows, researchers can examine how swearing varies in different storytelling formats and character portrayals. But because the corpora that I am using is not annotated for these variables, I decided I could not investigate this variable within the time frame I had.

3. Regional and Production Information: TV and movie corpora often provide metadata that includes information about the origin and region of the production. This enables researchers to investigate regional variations in swearing practices, dialectal differences, or cultural influences on the usage of swear words.

4. Spoken Language Features: Swearing in speech exhibits distinct features related to pronunciation, prosody, intonation, and non-verbal cues like gestures and facial expressions. TV and movie corpora capture these multimodal aspects, allowing for a comprehensive analysis of swear words beyond textual representation. But because the corpora that I am using is not annotated for these variables, I decided I could not investigate this variable within the time frame I had.

While written corpora have their own merits, they may not provide the same level of authenticity and contextual information for studying swear words. Written texts, such as books, newspapers, or online forums, may contain less explicit swearing or lack the contextual cues necessary for understanding the pragmatic functions and social dynamics associated with spoken swearing.

However, it's important to note that combining both spoken and written data can offer a more comprehensive understanding of swearing phenomena. Researchers may choose to supplement their analysis with written sources, such as social media data or literary works, to explore additional dimensions of swearing in different domains. Ultimately, the choice of corpus depends on the specific research objectives and the linguistic phenomena under investigation. TV and movie corpora are particularly valuable for exploring the spoken aspects of swearing due to their authenticity, context, and the availability of multimodal data.

4.5 Quantitative method:

Next, I will explain the **Quantitative** method. This method is when we gather the data and investigate patterns within the language based on the data provided to us. For this method, the best way to explore and make use of this method is by looking at the frequency of the search words. This study will be looking at the frequencies of each swear word listed in Table 1 and they will be compared to each other. If gender information is not included as a variable in a quantitative study due to restrictions, such as limited access or time constraints, it can be

attributed to practical considerations and the availability of data. Here are some reasons why gender may be excluded in a quantitative method:

1. Data limitations: Sometimes, the dataset being used for the research may not contain readily available or comprehensive gender information. Data collection processes might not prioritize collecting gender-specific data or may not have captured it consistently across the sample. In such cases, researchers may have to work with the available data and focus on variables that are more accessible.

2. Resource constraints: Collecting gender data, particularly in large-scale quantitative studies, can require additional resources, such as survey design, participant recruitment, and data coding. Depending on the scope of the study, the time and effort required to gather gender data may exceed the available resources or the timeline of the research project.

3. Research scope and objectives: The research objectives and focus of the study may not directly involve gender-related factors. If gender is not a significant aspect of the research question or hypothesis being investigated, researchers may choose to exclude it as a variable to maintain focus on other key factors. Because of the limitations of this paper, I will not include the topic of gender in my research.

While the exclusion of gender as a variable might be necessary due to practical limitations, it is important to acknowledge that gender can be an important dimension in understanding various phenomena, including social, economic, and health-related outcomes. If gender is a critical aspect of the research topic, alternative methods like qualitative research or additional efforts to gather gender-specific data could be considered to provide a more comprehensive analysis. In the context of the restriction mentioned in the paper, where the necessary information is available but not readily accessible and obtaining it would require additional time, it's important to acknowledge that research studies are often subject to practical constraints and limitations. These limitations may arise due to factors such as resource availability, time constraints, or data accessibility, which is the case in this paper and the reason for excluding the topic of gender.

When faced with restrictions, researchers often must make strategic decisions regarding the variables they include in their analysis. In some cases, certain variables, such as gender, may

not be readily available or easily accessible within the given timeframe. This is the case for my thesis. Collecting gender-specific data may require additional efforts, such as conducting surveys or obtaining data from external sources, which may not be feasible within the scope of the study.

4.6 Method

The following are the results of the method chosen for this paper:

1. We collect the frequency of occurrences for each swear word by decades. Analysing the data by decades allows for a broader perspective and captures long-term trends or patterns, minimizing the impact of year-to-year fluctuations. This approach provides a more consolidated view and facilitates the identification of significant shifts over time.

2. Picking words based on lexical function: To obtain more reliable results, certain words may need to be excluded, such as "bullshit" for lack of sufficient numbers of occurrences. Swear words and offensive language can introduce bias or emotional connotations that may affect the analysis. By focusing on words with a more neutral or informative lexical function, you can obtain results that are less influenced by emotional language.

3. Extracting data in a meaningful way: To extract data meaningfully, consider the following steps:

a. Define objectives: Clearly define the research question and objectives of the analysis to guide the data extraction process.

b. Select relevant sources: Identify appropriate sources that align with the research question, such as books, articles, social media data, or other text corpora.

c. Set inclusion/exclusion criteria: Determine the criteria for including or excluding specific data points based on relevance, quality, or other factors.

d. Choose appropriate data analysis techniques: Select methods that align with the research question and the type of data available. This could include statistical analysis, natural language processing, sentiment analysis, or topic modelling, among others.

4.7 Why those swear words?

The choice to focus on specific swear words can stem from various reasons:

1. Sociolinguistic significance: Swear words often carry cultural and sociolinguistic importance. Studying their usage can provide insights into the evolution of language, cultural shifts, and societal attitudes.

2. Colloquial language: Swear words like "shit" and "crap" are commonly used in informal or colloquial speech. Analysing their frequency or context of use can reveal trends in language usage, slang, or changes in popular culture.

3. Emotional impact: Swear words are known for their strong emotional connotations. Analysing their usage can provide insights into emotional expression, attitudes, or changes in the perception of taboo language.

Chapter 5 Frequency

This Chapter will contain the analysis and the reflection on the analysis of the data that was gathered for this thesis. One of the main parts of this chapter will go through the frequency and concordance lines of the swear words listed in Chapter 2, in Table 1. This analysis will try to answer the research questions based on the data found in the investigation as well. In this Chapter, we will look at the frequencies of the chosen swear words and see which of these swear words are most frequent. This will be illustrated through a chart and afterwards explained in texts. The progress of the swear words throughout the timeline, between 1930's to the 2010's will be shown here.

5.1 Frequency

The frequency of a lemma or word is found by finding out how many times it is repeatedly used in any context or function. By finding out the frequency of a word or lemma, we can investigate the various ways that it was used in said language and how the usage of it has been developed over the years. We may also find out which functions that said word or lemma plays in a sentence, and what impact that has on the language. Love (2021) uses frequency to investigate how swearing and swear words impact and affect a language. He is, however, not the only one to use this method to investigate language. Jay (2008) had similar findings to Love (2021), but as mentioned before, Jay's (2008) investigation was not as restricted as Love's (2021), which makes the comparison between the two studies somewhat have similarities but also many differences.

The first thing that I will start with when presenting the frequency is to show the frequency of the chosen swear words for this thesis using the Chart function. One of the difficulties of presenting this Chart is the number of swear words that are allowed on one chart before it becomes hard to read, and the data of the swear words collocated from both corpora, which are too large for them to fit in one chart. This means that the data would need to be broken down and divided into more than one chart so that the chart can show accurate and clear lines, and the reading of those charts would not be hard. The first thing to do when we find the frequency of a word or lemma is to find the function named 'chart'. There are two steps to find the frequency using this function:

1. Find the 'chart' tool function above.

List **Chart** Collocates Compare KWIC

swear word   [POS] ?

See frequency by section   Reset

☐ Sections   Texts/Virtual   Sort/Limit   Options

2. Place the intended researched word or lemma in the search field and press the 'find frequency by section' button.

This tool is good at searching for a specific period and/or a specific variety of spoken English. To search for this, the part that says section would have to be used. Here we can specify and modify the search to suit the topic that the thesis is about or create and/or edit our corpus to suit the topic. This function is available in both chosen corpora for this thesis, the TV corpus, and the Movies corpus (See the Charts in Appendix I and Appendix II) .

5.2: Analysing the Charts

Something important to note is that this analysis is based solely on the provided data and doesn't consider the broader context or factors that may have influenced the usage of swear words in movies and TV during these decades. The data presented for the analysis of each of the swear words is available on the last pages of the paper (See Appendix I & II). The data presented is the frequency per million words for the swear words presented in Table 1 in chapter 2. To avoid confusion, chapter 5 will use "" and CAPITAL letter when referring to the swear words.

5.2.1: Fuck

In the 1930s and 1940s, movies were the dominant form of entertainment, with a relatively high percentage of occurrences of the swear word in movies (0.06 and 0.05, respectively). TV had not yet gained significant popularity, as indicated by the negligible values (0.00) during these decades. In the 1950s, movies continued to be a popular choice, with a slight increase to 0.07. However, TV started to emerge as a contender, capturing 0.04 of people's leisure time. In the 1960s, both movies and TV saw a rise in popularity. Movies reached their peak during this decade with 0.08, while TV also had a notable increase to 0.05. In the 1970s and 1980s, movies remained relatively popular, but their dominance began to decline. The percentage of leisure time spent on movies increased to 0.12 in the 1970s and further rose to 0.19 in the 1980s. Meanwhile, TV gained ground and accounted for 0.03 of leisure time in the 1980s. In the 1990s, the trend continued with movies losing some popularity (0.28) while TV grew in importance, capturing 0.10 of people's leisure time. By the 2000s and 2010s, the decline of movies and the rise of TV became more evident. Movies accounted for 0.33 in the 2000s and 0.39 in the 2010s, while TV's share increased to 0.21 and 0.24, respectively.

In summary, this chart shows the shifting trends for the swear word "FUCK" in entertainment over time. Movies were initially dominant, but TV gradually gained popularity and started to compete with movies as a primary form of entertainment. By the 2000s and 2010s, TV had become more popular than movies in terms of people's leisure time.

5.2.2: Shit

In the 1930s and 1940s, the usage of the word "SHIT" in movies was relatively low, with percentages of 0.04 and 0.07, respectively. TV had negligible values (0.00) during these decades. In the 1950s, both movies and TV saw an increase in the usage of the word "SHIT." Movies remained at 0.07, while TV showed a value of 0.06. In the 1960s, both movies and TV maintained the same level of usage as the previous decade, with 0.07 and 0.01, respectively. In the 1970s, there was a significant increase in the usage of the word "SHIT" in movies, reaching 0.27. TV also showed an increase to 0.06. The 1980s continued the trend of increased usage, with movies reaching 0.36 and TV remaining at 0.06. In the 1990s, there was a slight decrease in the usage of the word "SHIT" in movies, dropping to 0.33. TV had a similar decrease to 0.11. In the 2000s and 2010s, both movies and TV maintained relatively stable usage levels. Movies had a slight decline to 0.32 and 0.33, respectively, while TV

increased to 0.23 and 0.29. The usage of the word "SHIT" in both movies and TV gradually increased from the 1950s through the 1980s, with a peak in movies during the 1980s. In the subsequent decades, there were fluctuations, but the term remained commonly used in both mediums.

5.2.3: Crap

In the 1930s and 1940s, the usage of the word "CRAP" in movies was relatively low, with percentages of 0.05 and 0.05, respectively. TV had negligible values (0.00) during these decades. In the 1950s, there was a noticeable increase in the usage of the word "CRAP" in both movies and TV. Movies reached 0.13, while TV had a higher percentage at 0.15. In the 1960s, both movies and TV saw a decrease in the usage of the word "CRAP." Movies dropped to 0.06, and TV decreased to 0.04. The 1970s and 1980s continued the downward trend in the usage of the word "CRAP." Movies had relatively low values of 0.04 and 0.03, respectively, while TV slightly increased to 0.07 and 0.12. In the 1990s, there was a further decrease in the usage of the word "CRAP" in both movies and TV. Movies dropped to 0.02, and TV reached 0.13. In the 2000s and 2010s, the usage of the word "CRAP" remained at relatively low levels. Movies and TV both maintained percentages of 0.03 and 0.10 in the 2000s and 0.02 and 0.07 in the 2010s, respectively.

5.2.4: Ass

In the 1930s and 1940s, the usage of the word "ASS" in movies was relatively low, with percentages of 0.10 and 0.07, respectively. TV had negligible values (0.00) during these decades. Later, in the 1950s, there was a noticeable increase in the usage of the word "ASS" in both movies and TV. Movies reached 0.12, while TV had a percentage of 0.04. In the 1960s, both movies and TV saw a slight decrease in the usage of the word "ASS." Movies dropped to 0.09, and TV decreased to 0.02. In the 1970s, there was a significant increase in the usage of the word "ASS" in movies and TV. Movies reached 0.18, while TV had a higher percentage of 0.14. The 1980s and 1990s continued the trend of increased usage of the word "ASS." Movies and TV both had percentages ranging from 0.07 to 0.20 during these decades. In the 2000s and 2010s, there was a slight decrease in the usage of the word "ASS" in both movies and TV. Movies dropped to percentages of 0.11 and 0.08, respectively, while TV had percentages of 0.18 and 0.13. The usage of the word "ASS" in movies and TV during the mid-20th century showed a notable increase from the 1950s, with both mediums experiencing

fluctuations in the subsequent decades. While the 1970s witnessed a significant surge in usage, the 1980s and 1990s maintained higher percentages. The 2000s and 2010s marked a slight decrease, but "ASS" remained a frequently used term, especially in TV shows during the 21st century.

5.2.6: Dick

Starting in the 1930s, Movies prominently featured the word "DICK," with a high prevalence of 0.49, suggesting its frequent usage in that era. On the other hand, TV had negligible values, indicating minimal or no usage of the word during this period. The same happened during the1940s, when the usage of the word "DICK" decreased in movies to 0.32, but it still maintained a relatively significant presence. TV remained at negligible levels.

In the 1950s both movies and TV experienced an increase in the usage of the word "DICK." Movies reached 0.42, indicating continued prominence, while TV notably surged to 0.62, suggesting a substantial increase in the portrayal or use of the term. In the period 1960s, there was a significant drop in the usage of the word "DICK" in movies, declining to 0.11. This decrease might reflect evolving societal norms or changes in censorship. Conversely, TV witnessed a sharp rise to 0.74, indicating a notable increase in the presence or acceptance of the word during this decade. In the period 1970s and 1980s, the prevalence of the word "DICK" remained relatively low in both movies and TV. Movies and TV maintained percentages of around 0.03 during these two decades, suggesting a decline or stricter regulations on the usage of such explicit language. In the 1990s, the usage of the word "DICK" continued to be relatively low in movies (0.04) and TV (0.13), indicating a continuation of the trend observed in the previous two decades. In the 2000s and 2010s, the usage of the word "DICK" remained consistently low across movies and TV, with percentages ranging from 0.03 to 0.04. This suggests a continued adherence to censorship guidelines or a decrease in the prevalence of explicit language.

5.2.7: Bitch

The first period available starts in the 1930s. The usage of the word "BITCH" in movies was relatively low, with a prevalence of 0.04. TV had negligible values, indicating minimal or no usage of the word during this period. We move to the 1940s, where Movies saw a slight increase in the usage of the word "BITCH" to 0.05, suggesting a slightly more frequent

presence. TV remained at negligible levels. In the 1950s, both movies and TV experienced a further increase in the usage of the word "BITCH." Movies reached 0.08, indicating a continued rise in its prevalence, while TV had a low percentage of 0.01. A shift was made in the 1960s. The usage of the word "BITCH" notably increased in movies, reaching 0.16. This suggests a significant rise in the portrayal or use of the term during this decade. TV also saw a slight increase to 0.02. In the 1970s, both movies and TV maintained relatively high percentages, with movies at 0.17 and TV at 0.10. This indicates a continued prevalence of the word "BITCH" in both mediums during this decade. Then another shift occurred in the 1980s. The usage of the word "BITCH" decreased in both movies and TV, with movies reaching 0.13 and TV dropping to 0.06. This decline might reflect evolving societal norms or changes in censorship.

This continued into the 1990s. The prevalence of the word "BITCH" decreased further in movies (0.10) and TV (0.16), indicating a continued decline or stricter regulations on the usage of such explicit language. Then during the period of 2000s and 2010s, the usage of the word "BITCH" remained relatively consistent and low across movies and TV. Movies and TV had percentages ranging from 0.09 to 0.13, suggesting a continued adherence to censorship guidelines or a decrease in the prevalence of explicit language.

5.2.8: Bastard

Starting in the periods of 1930s and 1940s, the usage of the word "BASTARD" in movies during these decades was relatively low, with percentages of 0.04 and 0.04, respectively. TV had negligible values, indicating minimal or no usage of the word during this period. Then 1950s, both movies and TV experienced an increase in the usage of the word "BASTARD." Movies reached 0.07, suggesting a slightly more frequent presence, while TV had a percentage of 0.04. In the 1960s, there was a significant increase in the usage of the word "BASTARD" in movies, reaching 0.24. This indicates a notable rise in the portrayal or use of the term during this decade. TV also had a percentage of 0.04. But then in the 1970s, the prevalence of the word "BASTARD" decreased in movies to 0.11. TV, however, experienced a notable increase to 0.15, suggesting a rise in the usage or acceptance of the word during this period. In the 1980s, both movies and TV saw a decrease in the usage of the word "BASTARD." Movies dropped to 0.06, while TV remained relatively high at 0.14, indicating a continued prevalence in TV shows.

The situation continued in the 1990s. The usage of the word "BASTARD" continued to decrease in both movies (0.05) and TV (0.12), reflecting a decline or stricter regulations on the usage of such explicit language. In the periods of 2000s and 2010s, the prevalence of the word "BASTARD" remained consistently low across movies and TV. Movies and TV had percentages ranging from 0.02 to 0.06, suggesting a continued adherence to censorship guidelines or a decrease in the prevalence of explicit language.

5.2.9: Bullshit

In the 1930s, both movies and TV had negligible values for the usage of the word "BULLSHIT," indicating minimal or no usage during this period. Then in 1940, Movies saw a slight increase in the usage of the word "BULLSHIT" to 0.01, suggesting a slightly more frequent presence. TV remained at negligible levels. This carried on in the 1950s. Both movies and TV had a low prevalence of the word "BULLSHIT," with percentages of 0.01 and 0.00, respectively, which indicates that the usage of the word "BULLSHIT" is almost non-existent in the Movies and non-existent in the TV. A shift started in the 1960s. There was a small increase in the usage of the word "BULLSHIT" in both movies and TV, with movies reaching 0.02 and TV having a value of 0.00. Although the data shows an increase, however, the data indicates that this swear word is still not so vary actively used in this period, neither in the Movies nor the TV. The data starts to show some progress around the 1970s, when Movies experienced a notable increase in the usage of the word "BULLSHIT" to 0.05, suggesting a rise in the portrayal or use of the term during this decade. TV also saw a small increase to 0.01. In the 1980s, both movies and TV maintained percentages of 0.05 and 0.01, respectively, indicating a continued prevalence of the word "BULLSHIT" in both mediums during this decade. In the 1990s, the usage of the word "BULLSHIT" remained relatively stable in Movies (0.05) and slightly increased in TV (0.02). Then in the 2000s and 2010s periods, the prevalence of the word "BULLSHIT" decreased in both movies and TV. Movies dropped to 0.04 and further decreased to 0.03 in the 2010s. TV, on the other hand, saw a slight increase to 0.03 in the 2000s and reached 0.04 in the 2010s.

5.2.10: Fanny

The data starts from the period of 1930s, when the swear word "FANNY" had a relatively high prevalence in movies during the 1930s, with it being used in approximately 17% of the

analysed Movies from that decade. On the other hand, TV shows had negligible instances of the word, indicating minimal or no usage of it during this period. In the 1940s, the usage of the word "FANNY" experienced a significant increase in Movies during the 1940s, with it being present in about 31% of the Movies analysed. However, TV shows still had negligible usage of the swear word. A change happened in the 1950s. There was a notable decline in the usage of the swear word "FANNY" in Movies during the 1950s, with it being used in only 2% of the analysed Movies from that decade. TV shows also had a low prevalence of the word, with it appearing in approximately 1% of the shows. In the 1960s, Movies saw a slight increase in the usage of the swear word "FANNY" during the 1960s, with it appearing in about 12% of the analysed Movies. TV shows had a relatively higher prevalence, with the word being used in approximately 8% of the shows. In the 1970s, however, the prevalence of the swear word "FANNY" decreased significantly in Movies during the 1970s, with it being present in only 1% of the analysed Movies. However, TV shows had a relatively higher usage, with the word appearing in approximately 12% of the shows. In the 1980s, both movies and TV shows had very low percentages of the usage of the word "FANNY" during the 1980s. It appeared in approximately 1% of the analysed Movies and 18% of the analysed TV shows, and this continued throughout the 1990s to 2010s: The usage of the word "FANNY" further decreased in both Movies and TV shows during the 1990s to 2010s. It was not present in any of the analysed Movies or TV shows across these decades.

5.2.11: Twat

The data for this swear word was incredibly little to the point that the highest frequency is around 3%. The swear word "TWAT" had negligible or non-existent usage in both Movies and TV shows across the analysed decades. In Movies, there were no instances of the swear word "TWAT" being used from the 1930s to the 2010s. Similarly, in TV shows, the prevalence of the word "TWAT" was extremely low, with only a minimal appearance in the 1950s (0.01%) and 1970s (0.03%), but otherwise not present in any other decades.

5.2.12: Pussy

This is the last swear word to be analyzed. Similarly, to the swear word "TWAT", the swear word "PUSSY" had limited usage in both Movies and TV shows across the analysed decades. We started in the 1930s. In Movies, the swear word "PUSSY" was used in approximately 1% of the analysed films from the 1930s. A slight shift occurs in the 1940s, where the prevalence

of the word "PUSSY" in movies increased slightly to around 3%. In the 1950s, the usage of the word "PUSSY" in Movies remained relatively low, with it appearing in approximately 2% of the analysed films. A small shift happens in the 1960s, when Movies saw a slight increase in the usage of the word "PUSSY" to around 3%. Both movies and TV shows had a low prevalence of the word "PUSSY" during the 1970s, with movies having it in approximately 1% of films, while TV shows increased to around 6%. In the 1980s, the usage of the word "PUSSY" remained low in Movies, appearing in approximately 1% of films. TV shows had a slightly higher prevalence of around 4%. Then in the 1990s, both Movies and TV shows had a consistently low prevalence of the word "PUSSY," with it appearing in approximately 1% of films and shows. Moving to the period 2000s, the prevalence of the word "PUSSY" remained relatively consistent in Movies and TV shows, with both having it in approximately 2% of their respective content. The usage of the word "PUSSY" continued to be consistent in Movies and TV shows, with it appearing in around 2% of films and shows in the period of 2010s.

5.3 How this affects the language:

The usage of explicit or offensive language, including swear words, in movies and TV shows can have various effects on language and communication:

- Desensitization: The frequent or casual use of these explicit words in media can lead to desensitization, where people become less shocked or offended by their usage. This can result in a more relaxed attitude towards profanity in everyday language.

- Normalization: When explicit language is frequently used in popular media, it can contribute to the normalization of such language in everyday conversations. People may become desensitized to these words and may use them more casually or without considering their impact.

- Social Acceptability: The portrayal of explicit language in movies and TV shows can influence social acceptability and attitudes toward the use of such language. If these words are frequently used and accepted in media, it may contribute to a more permissive attitude towards their use in society.

It is worth noting that the impact of explicit language in movies and TV shows is not universally negative or positive. It depends on individual perspectives, cultural norms, and societal values. Additionally, there are considerations of context, intent, and audience when it comes to the usage of such language in media. It is also important to note that the effects of explicit language on language and society are complex and multifaceted. They can vary depending on cultural, social, and individual factors. Furthermore, the impact of explicit language should be considered within the broader context of language evolution, societal norms, and the dynamic nature of communication.

5.4: Frequency in Different Varieties

Figure 2: *Frequency per mill in different varieties of spoken English*

	fuck	shit	crap	ass	fanny	dick	Pussy	twat	Bastard	Bitch	bullshit
TV:UK/IE	228.57	232.69	27.49	16.74	14.98	33.30	8.18	15.18	88.35	53.69	18.21
TV:US/CA	178.61	215.04	80.90	151.34	4.09	44.43	11.83	0.84	30.36	118.13	33.79
TV:AU/NZ	552.86	554.42	72.56	34.04	2.14	54.66	15.95	1.36	74.70	101.16	107.97
MV:UK/IE	657.78	357.95	30.41	63.79	23.31	47.49	15.57	14.53	118.92	84.37	36.35
MV:US/CA	691.47	741.43	51.08	238.73	7.65	89.10	38.26	1.36	53.98	202.95	87.43
MV:AU/NZ	1,177.22	830.07	59.59	91.31	2.37	85.08	25.20	2.08	187.06	143.19	145.26

Figure 2 is a table describing the frequency of all the swear words presented in Table 1 from Chapter 2. Figure 2 shows the frequency across the different varieties of spoken English from the TV corpus and the Movies corpus. The numbers marked in yellow show the highest frequency of the swear word in both corpora. The results shown in Figure 2 can be said to be expected but some presented unexpected results. First, we will start with the swear word "FUCK" and analyze the findings for each swear word in the order of Figure 2. The swear word "FUCK" is one of the most popular swear words used today, therefore, the high frequency presented in the figure was expected. In the TV corpus, the AU/NZ variety of spoken English shows the highest frequency for this swear word, and the same is the case in the Movies corpus. Having a high frequency was the expected result, but the variety that the swear word had the highest frequency was not. Based on the popularity of the US/CA variety of spoken English, we would expect this variety to dominate the other four varieties. Instead,

the AU/NZ variety dominated the US/CA variety by around 400 words in the TV corpus and around 500 words in the Movies corpus. The second swear word is "SHIT". This swear word is just as popular as the swear word "FUCK" and the results were the same in the TV corpus where the frequency of both swear words is in the 550's per million words. One thing to comment on is the difference in the frequency among the different varieties. The difference between the highest frequency results and the others is not that large. There is around 300 word difference between the highest frequency and the lowest in the TV corpus, while there is a 500 word difference in the Movies corpus. Figure 2 shows that between all the swear words, the swear words "FUCK" and shit have the highest frequency.

The next swear word to be discussed is "CRAP". This swear word is related to "SHIT" as they both have similar meanings. Having similar or the same meaning, we would expect the results for the two swear words to be similar or even identical. This was however not the case. While the highest frequency for the swear word, "SHIT", is as high as 554.42 words per million in the TV corpus, the highest frequency for the swear word, "CRAP", is only 80.90 per million words in the TV corpus. The same is the case for the results in the Movies corpus. The highest frequency for the swear word, "CRAP", is 59.59 per million words. There is also a difference between which variety that the highest frequency is presented in. While the highest frequency is presented in the AU/NZ variety of spoken English for the swear word, *shit*, there are four varieties that the swear word, "CRAP", is presented in. In the TV corpus, the swear word, "CRAP", is most frequent in the US/CA variety of spoken English, and in the Movies corpus the AU/NZ variety presents the highest frequency among the other varieties.

Next, is the swear word "ASS". The heist frequency for this swear word is presented in the US/CA variety of spoken English. This is because of the usage of compound words e.g. "Big ass car, smartass, dumbass", which increase the frequency of the swear word. This swear word is not always used with the intention of swearing. Another usage for this swear word can be in, for example, intensifying measurements, such as in, e.g. "big ass car". There is an older and different word for this swear word, which brings us to the next swear word, "FANNY". This swear word has the same meaning as the swear word, "ASS". The swear word, "FANNY", however, is less frequent than the swear word, "ASS", as the frequency for this swear word is in double digits as shown in the figure, while the frequency for the swear word, "ASS", is in triple digits. The reason behind the low frequency may have to do with the fact that this swear word is old and mostly used in names. The swear word, "FANNY", is

most frequent in the UK/IE variety of spoken English in both corpora, which is an expected result considering the origin of this swear word is British. The next three swear words are "DICK", "PUSSY", and "TWAT". These three swear words are words that refer to the female and male sex organs. The highest frequency for all three swear words is in the double digits and is considered, along with the swear word "FANNY", to be the lowest among all swear words on the list. The difference between the three swear words is that for the swear words "DICK" and "PUSSY", the highest frequency in the TV corpus presents itself in the AU/NZ variety of spoken English, while the highest frequency is presented in the US/CA variety of spoken English in the Movies corpus. As for the swear word, "TWAT", the highest frequency in both corpora is in the UK/IE variety of spoken English, which is an expected result considering the swear word, "TWAT", is a British word. One of the expected results discovered in Figure 2 is the fact that the frequency for the swear word, "TWAT", showed such a low frequency for the US/CA variety, which leads to the fact that this swear word is not used in US/CA variety and is almost non-existent.

The difference between the highest and the second highest frequency is the next thing to discuss. For example, for the swear word, "BASTARD", the highest frequency is 88.35 per million words in the TV corpus, and this is presented in the UK/IE variety. But looking at Figure 2 we can see that in the TV corpus, the AU/NZ variety had a frequency close to the highest frequency with 74.70 per million words, which means that the two/four varieties are similar in the usage of this swear word. The same is happening for the frequency in the Movies corpus, only the highest frequency is presented in the AU/NZ variety, while the second highest is the UK/IE variety. The same difference can be seen in the swear word, "BITCH". The highest frequency is presented in the US/CA variety and the next highest frequency is presented in the AU/NZ variety in both corpora. The frequency results for the last swear word, "BULLSHIT", show that this swear word is most frequently used in the AU/NZ variety of spoken English, but the frequency presented is not among the lowest nor the highest.

5.3.1 Similarities and Differences

It is common for people to use the same words. But there are times when variation is needed in language so that the language does not become too bland to follow. The same thing may be said to happen with swearing and the usage of swear words. Many swear words exists, and

even though the words are different and consists of different letters, their meaning is one of the things that they have in common. A few examples of that are the swear words "ASS" and "FANNY", where the two both mean a person's butt or behind, but both are written differently and come from two different varieties of English and from two different periods but both have or are used in the same contexts. The question that I will try to answer is, even though both swear words have the same meaning, do they have the same frequency or not, and if the answer is no, then why is it that one swear word has a higher frequency than the other?

First, I will compare the frequency of the sets of swear words that have the same meaning as one another. The first set with two swear words with the same meaning to compare are "SHIT" and "CRAP". In the Movies corpus, the swear word "SHIT" has a frequency of 670.81 per million words throughout time, dating from 1930's to 2010's in all six different varieties of spoken English. The frequency for the TV corpus for the swear word "SHIT" for the same period, on the other hand, is much lower with 224.33 per million words. As for the swear word "CRAP", in the Movies corpus, this swear word has a frequency of 47.42 per million words. In the TV corpus, the frequency of the swear word "CRAP" is 72.04 per million words. Just by looking at the frequencies of both swear words, we get the result that the swear word "SHIT" has a higher frequency in both corpuses overall in the period of 1930's to 2010's. Both of the swear words have similar meanings. Cambridge Dictionary has concluded that gives both the swear word "SHIT" and the swear word "CRAP" the definition being the slang word used to describe "solid waste of an animal or person" or as we would say in normal not linguistical speech, poop (Cambridge Dictionary, s.v. "crap"; "shit").

The next set that I will look at is the set with the two swear words "FANNY", "PUSSY", and "TWAT". The same case can be seen happening here. The swear word "FANNY" is a swear word most frequently used in the UK/IE variety of spoken English. The meaning of this word in this variety of English is, "a woman's sexual organ or genitals" (Cambridge Dictionary, s.v. "fanny"). The overall frequency of this swear word in the Movies corpus is 10.08 per million words over the period of the 1930's to the 2010's. The frequency of this swear word in the TV corpus for the same period is 5.84 per million words. The difference in frequency between the two corpora is halved in the TV corpus. As for the meaning of the swear word "PUSSY", Cambridge Dictionary concluded that in the case of offensiveness, the meaning behind this word is "a woman's vagina", or just a term used for the female genital organ. The frequency

of this swear word in the Movies corpus is overall 34.01 for the period from the 1930s to the 2010s. As for the frequency of this swear word in the TV corpus, the data show that it is 11.34 per million words. This swear word is mostly used in the US/CA variety of spoken English in the Movies corpus, while it is most frequently used in the AU/NZ variety of spoken English in the TV corpus. Based on the frequency presented in both corpora, the swear word "PUSSY" is much more frequently used than the swear word "FANNY", even though both swear words have the same meaning and are both considered offensive. The last swear word on my list that have the same meaning as the previous two is the swear word, "TWAT". According to Cambridge Dictionary, this word is offensive, and the meaning is "sex organ" and it is "an offensive (term) for the vagina" (Cambridge Dictionary, s.v. "twat"). Cambridge Dictionary further explains that this swear word is used as a noun to call someone a "stupid or unpleasant person". According to both the TV corpus and the Movies Corpus, the swear word, "TWAT", is dominantly used in the UK/IE variety of spoken English. The frequency of this swear word in the Movie corpus, from the period of 1930's to the 2010's, is 3.29 per million words, while in the TV corpus, from the same period, the frequency is 3.53 per million words. Just by comparing the frequency of all three swear words in both corpora, the results are that the swear word "PUSSY" is the swear word that is used the most, while the swear word "TWAT" is the least popular of all of them.

Lastly, I will compare the swear words, "BULLSHIT" and "SHIT" to one another. Although, technically these two swear words are connected in a different way than that of the last two sets of swear words. The reason behind that is that one of the two swear words, "BULLSHIT", is a compound taken from the two nouns, *bull*, and a swear word, *shit*. This gives the swear word a more specific picture to the hearer and adds variation to the language by including a known and general thing in it. Despite the swear words likeness to each other, the question of which one of them is most frequent remains. The frequency for the swear word "SHIT" in both corpora has already been presented and therefore, the only thing remaining is to present the frequency of the swear word "BULLSHIT" and compare the findings to one another. According to the data collected from both corpora over the period of 1930's to 2010's, the frequency of the swear word "BULLSHIT" in the TV corpus is a total of 32.55 per million words, while the frequency is 3,29 per million words in the Movies corpus. The data presenting the frequency from the TV corpus for "BULLSHIT" shows that the AU/NZ variety of spoken English is the variety that has the highest frequency for this word, while the data presenting the frequency from the Movies corpus shows that the UK/IE

variety of spoken English has the highest frequency for this word. When comparing the sets of swear words presented above, another noteworthy observation emerges. The data consistently indicates that the frequency of these swear words is significantly higher in the Movies corpus compared to the TV corpus. However, it's worth noting that the last set of swear words, particularly the swear word "BULLSHIT", exhibited a higher frequency in the TV corpus in contrast to the Movies corpus. The observed difference in frequency between the Movies corpus and the TV corpus for most swear words can be attributed to several factors:

1. Contextual Differences: Movies often have more freedom in terms of language use and content compared to TV shows, especially those aired on network television. This greater creative liberty allows for more explicit language, including the use of swear words, in movies. In contrast, TV shows, particularly those on broadcast television, are subject to stricter content regulations and censorship, which can limit the frequency and intensity of swear words.

2. Audience Demographics: The target audience for movies and TV shows can vary significantly. Movies may cater to a more diverse and potentially mature audience, while TV shows may be tailored to specific demographics or age groups. This can influence the frequency of swear words to align with audience expectations and ratings.

3. Genre Differences: Different movie genres may naturally incorporate more profanity than TV shows of various genres. For example, action movies or adult comedies might feature more swearing as part of their typical dialogue, whereas family-oriented TV shows would restrict such language.

4. Creative Choices: The writers, directors, and producers of movies and TV shows make creative decisions about language use that align with their artistic vision and storytelling goals. These creative choices can influence the inclusion of swear words and their frequency based on the desired tone and impact of the content.

5.4 Summery

The swear words "FUCK", "SHIT", "ASS", and "BASTARD" all exhibit triple-digit frequencies in both the Movies and TV corpora. However, "BITCH" and "BULLSHIT" only register double-digit frequencies, with "BULLSHIT" falling into this category exclusively within the Movies corpus. The remaining swear words, including "CRAP", "FANNY",

"DICK", "PUSSY", "TWAT" and another instance of "BASTARD" yield double-digit results in their respective data. The data implies that words like "FUCK", "BITCH", "BULLSHIT" and "SHIT" maintain higher frequencies and are more popular than swear words referencing anatomical or sexual terminology, as indicated in Figure 2. Among the 11 swear words listed in Table 1 (chapter 2), the AU/NZ variety of spoken English boasts the highest frequency for 8 of them. The US/CA variety follows closely, with 5 out of the 11 swear words achieving the highest frequency, while the UK/IE variety ranks last, with 4 out of the 11 swear words having the highest frequency within the corpora.

Chapter 6 Collocation

Next, I will discuss the collocation that my set of swear words presented. Collocation is the discovery of the different words that co-occur with the researched word. The researched words or lemmas that I will try to find the collocation to in this case are the swear words presented in Table 1 in the list named "My swear words" in Chapter 2, which is a list of eleven different swear words that are considered some of the most popular in general. Most of these swear words were taken from Robbie Love's article "Swearing in Informal Spoken English" (2021). The collocates of these swear words do not specify from which period they are taken from but rather give a summary of the period of the whole corpus. This will be presented as tables for each of the eleven swear words. Through these tables, the results from the data provided from both the TV corpus and the Movies corpus will lead to the discovery of which collocations collocate the most with the swear words. On top of that, I will try to investigate if there are any words found in each list that are present and is common for two swear words or more if not all of them.

The first thing that needs to be done in to find the collocation of the swear words. To find the collocates in the corpus, we need to follow the following steps:

1. Find the collocates function in the front row at the top.
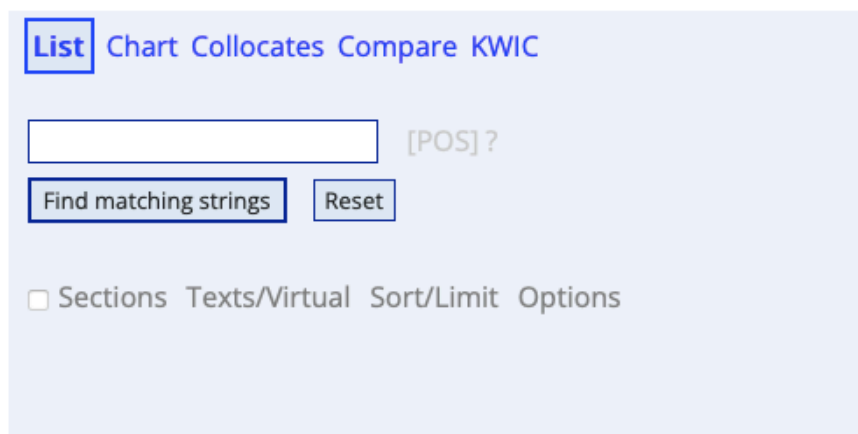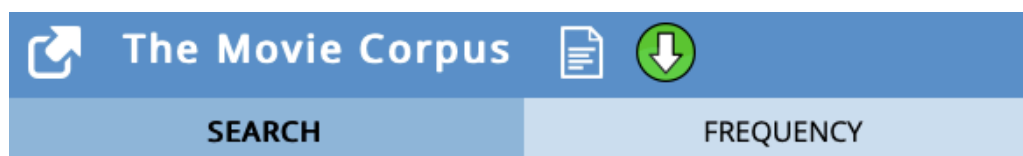
2. Next, we will find two boxes. In the first search box, which is named 'word/phrase, we type in the swear word, while in the second search box, which is named 'collocates', we type this star sign, (*). This sign makes it possible to get all the collocates that come with the search word that we are looking for.

3. Before we press the 'find collocates' button, we can control which position the collocates are in, if they are on the left or the right, and by how many, for example, two to the left and two to the right, or any order that is preferred. I find the collocates of the swear words by setting the settings to one to the left, 1L, and three to the right, 3R. This means that as long as the swear word or lemma is in the position between 1L and 3R, the collocation list will include all the collocates for them.

4. The last of the steps is to press 'find collocates' and we get a whole list of words and signs that collocates with the search word that we are looking up.

In the tables, following the collocates on the list, small signs such as (L) and (R) will be found. The reason for this is to indicate the position of the collocate if it occurs to the of the swear words or to the right of them. For example, *You **fucking** idiot!* The collocate, *you*, is on the left in terms of position to the swear word, which will be indicated with the sign (L) in the table. To avoid confusion, the swear words will be written in *italics* while the collocations will be written in CAPITAL LETTERS.

6.1: Collocation Tables

Table 2: *List of collocates that collocate with the swear word "fuck"*

	Fuck		
Ranks	Movies corpus	TV corpus	Examples:
1	Fuck (L&R)	Fuck (L&R)	Fuck fuck fuck fuck fuck. Fuck!
2	Man (R)	Sake (R)	Fuck **man**! Talk to me for fuck **sake**!
3	Shit (R)	Shit (R)	Listen to me, you fuck **shit**. I could fuck **shit** up for you…
4	Sake (R)	Hey (R)	For fuck **sake**, Marcus! Fuck**! Hey**!
5	Hey (L)	Holy (L)	**Hey** fuck you! **Holy** fuck.
6	Just (R)	Wrong (2R)	**Just** fuck' em, I to go to my dad's. What the fuck is **wrong** with you?
7	Wrong (R)	Right (R)	Fuck **wrong** with you? Yeah fuck **right** off!
8	Fucking (R)	Away (R)	**Fucking** fuck it! Get the fuck **away** from me!
9	Right (R)	Fucking (R)	Fuck **right** off! How the **fucking** fuck did this happen?
10	Away (3R)	Well (R)	Fuck the pain **away**. Fuck. **Well**, are they safe?

Table 2 provides a comprehensive overview of the top ten words that commonly collocate with the swear word *fuck* within both the Movies corpus and the TV corpus. Upon careful examination, it becomes apparent that the collocates associated with *fuck* exhibit a striking similarity across both corpora, with only one set of words differing between them. An aspect to note is the variation in the order of these collocates between the two corpora. For instance, the collocate MAN secures the second position in the Movies corpus list but does not feature in the TV corpus list. Delving deeper into the original data I collected, it becomes evident that the majority of the collocates that emerge in the collocation list belong to the category of function words. These encompass articles such as "A" or "THE," pronouns like "YOU," "HE," and "YOUR," adverbs including OUT or IN, and verbs such as "WAS," among others. This observation sheds light on the multifaceted usage of the swear word *fuck*. It primarily serves as a reference to someone, the sexual act itself, or a reference to said action, in addition

to functioning as a verb within a sentence. However, an intriguing revelation stemming from the original collocation lists from each corpus is that *fuck* predominantly assumes the role of a noun in most sentences and contexts. This multifaceted usage underscores the versatility and adaptability of this swear word across various linguistic and contextual dimensions.

There is, however, another discovery to make note of, which is the fact that while most of the collocates mentioned in Table 2 are lexical words, these collocates are filtered from the original collocation list for this paper. The reason why I decided to filter the collocation list that is presented in this dissertation so that it focuses on lexical words instead of function words collocates is because lexical words contain more meaning in them. The lexical words give a better distinction for the investigation of patterns in the language using the different functions that the collocates for the researched swear words provide. The only knowledge that the different function words in the original collocation list provide is that the swear words have the function of nouns. This is different from the lexical words where different conclusions are provided because of the different collocations. For example, for the function words, if we say the collocates are THE, A, IN, OF, and so forth, all that tells us is that the word is a noun. But we already know that all the words are nouns, except for FUCK whose variants include verbal, FUCK IT, FUCK YOU, and adjectival, FUCKING IDIOT, uses in addition to the nominal ones, DON'T GIVE A FUCK. Since we already know the word class, having a collocation list full of function words only reinforces what is already known, and therefore, what is of interest is what lexical words occur in conjunction with the given swear word.

The collocation list that contains lexical collocates mentioned starts with the collocate FUCK, which is number 1 in the rank. In this case, the collocate, FUCK, is replicated many times, as seen in the example in the Table. Because of the replication in the way that this collocate is used in the example, the collocate is acting as an intensifier. This means that the context of the example is putting the weight on the collocate, FUCK. When the collocate is used as an intensifier, FUCK is emphasized even more than when used as a noun. In the case of this collocate, both corpora use the same example from the Table. This is mostly used in the US/CA variety of spoken English. None of the other collocates on the collocation list is used the same way as this collocate is used, as shown in the example. This means that in the case of the lemma *fuck*, the only collocate that uses replication as an emphasis is the collocate

FUCK. The usage of this collocation is the same way as when the speaker keeps repeating the word "no, no, no, no!" to the hearer for a firm emphasis.

The rankings of the collocates in the tables are indeed significant, as they offer insights into the popularity and frequency of these collocates in the two distinct corpora. A prime example is the collocate SAKE, which holds the second position in the TV corpus but is ranked fourth in the Movies corpus within Table 2. This disparity suggests that SAKE sees more extensive use in the TV corpus compared to the Movies corpus. Despite these differences in rankings, the usage of the collocates remains consistent, as evidenced by the examples provided. This pattern is mirrored across various other collocates that appear in both corpora; they maintain the same or similar examples but exhibit varying rankings. The positioning of these collocations within the examples is also noteworthy, as they consistently appear adjacent to the swear word *fuck.* This placement underscores the flexibility of the collocate FUCK, demonstrating that it can occupy any of the positions indicated in the settings, ranging from 1L to 3R. The rationale behind selecting these settings lies in their ability to capture the diverse contextual nuances and grammatical variations in which this collocate is employed across the two corpora. Because I know that most of the swear words are nouns, I choose 1L setting to capture nouns and adjectives remodifying a noun head. As for the reasoning behind choosing 3R settings is that for those of the swear words that are verbs, the interesting collocates often come within 3 positions to the right, as in for example "fuck off".

There are, however, two exceptions. The collocates FUCK and SHIT have the same rankings in both the Movies and the TV corpus. This means that no matter which corpus the collocates are in, the amount of their usage with the search word *fuck*, is the same. The example for the collocation SHIT is not the same entirely, whereas the example for the TV corpus is figurative. This can be proven by seeing if the sentence may still have the same meaning even if we change the swear words into non-swear words, which in this case is possible. The example from Table 2 is as follows, "I could fuck shit up for you". Changing the swear words here into non-swear words and still maintaining the same meaning is possible. The example would then be as follows, "I could mess things up for you". The meaning stands the same as the speaker is stating their intentions for the speaker. The swear word *fuck* here is the replacement for *mess* and the collocation SHIT is acting as the replacement for the word THINGS.

It is also noticeable that the swear word *fuck* can be used in various forms, including the present participle. An example of that is the transaction from fuck to fucking. The following phrases can illustrate this clearly, "I don't **fuck** around" as opposed to "Don't you keep on **fucking** around". Both phrases may be taken figuratively, but only one of them may have a literal meaning to it (Coulson & Oakley, 2005).

There are some other differences present in the collocation list in Table 2. For example, collocate number 2 MAN and number 6 JUST in the Movies corpus are not present in the TV corpus, while collocate number 5 HOLY and number 10 WELL in the TV corpus are nowhere to be found in the Movies corpus. This means that each collocate is available only in their respective corpus.

Table 3: List of collocates that collocate with the swear word "shit"

Shit			
Ranks	Movies corpus	TV corpus	Example:
1	Holy (L)	Holy (L)	**Holy** shit!
2	No (L) !!!	Shit (R&L)	**No** shit **Shit shit shit shit shit!**
3	Shit (L/R)	Yeah (R)	Oh Jesus, **shit shit**. Oh, shit **yeah**, lady…
4	Man (L&R)	Right (L&R)	Holy fucking shit **man**. Some grown-ass **man** shit. It ain't **righ**t shit. That was really solid shit right there.
5	Hey (L)	But (L&R) !!!	Tara, **hey** shit head, watch it. …he knew shit **but** didn't know shit…
6	Right (L&2L)	Man (L)	You did all the **right** shit, all the **right** police shit… Once I had a **man** shit in the back seat.
7	Yeah (3L)	Just (L)	**Yeah**, I talk shit. I **just** shit my pants.
8	Fucking (L)	Hey (R)	No **fucking** shit! **Shit** hey!
9	fuck (L) !!!	Little (L)	You know how you fuck shit up… …you stupid **little** shit.
10	Just (L)	Fucking (L)	I think I **just** shit my pants. Holy **fucking** shit!

Table 3 has the same scenario as Table 2, where the top ten words that collocate with the searched swear word, *shit*, are presented from both the Movies corpus and the TV corpus. In Table 3, like Table 2 almost all the collocates from each corpus have the exact set of words, just that all the collocates are in different order, except for the first set. The collocation set is in the first rank for both corpora and is the only set that fulfils both requirements of being in the same rank and being the same. The collocation HOLY is the collocation in question. The usage of this collocation is usually done with context to religion. A typical example of the usage of this collocation is "holy bible or holy Quran". In contrast to this, the example makes use of this collocation in pair with the swear word that has the meaning of 'bodily waste'.

This matches with the case that was for Table 2, where the collocations FUCK and SHIT also fulfilled the same two requirements. This means that in the case of the lemma or swear word, *shit*, the collocation HOLY is considered the most used word in combination with the swear word in both the TV and Movies corpus. Just like the case in Table 2, the collocations gathered in this list are the lexical words that are taken from the original list that is available in the data. The reason for the lexical words to be filtered out from the lexical words is for a better understanding and a deeper investigation. As mentioned before, the function words do not provide us with any other information and knowledge of the swear words other than their word class, and they are used as nouns, which is old knowledge, whereas the lexical words may provide new knowledge.

Another similarity that Table 3 has with Table 2 is the usage of the collocation SHIT ranking number 2 in the TV corpus on the list from Table 3 to the collocation FUCK ranking number 1 on the list for both corpora from Table 2. It is clear from the example on the table that the collocation SHIT is being replicated five times. This replication is identical to the replication of the collocation FUCK. The fact that both collocations are so high on the list leads to the connection that both collocations, in with their usage of their respective swear word, have the same usage and are more popular than the other collocations on the lists. What is interesting to point out is that in Table 2 the collocation FUCK has the same ranking in both corpora, while in Table 3 the collocation SHIT in the Movies corpus is one rank under. This means that even though both corpora use the same collocation, and both are high up on the list, the usage of the collocate SHIT as a replication in connection with the swear word *shit*, is still slightly more used in the TV corpus than the Movies corpus. However, the main point is that

both collocations are used as an intensifier by using replication as an emphasis to make the point even more clear.

Table 3 has two sets of collocations that can be found in the list for one corpus but not the other. The first set is collocation NO, which is number 2 on the list, and the collocation FUCK, which is number 9 on the list for the Movies corpus. The other set is found on the list for the TV corpus but not on the list for the Movies corpus. This set contains of the collocation BUT, which is number 5 on the list, and the other collocation is LITTLE, which is number 9 on the list. There is another thing that needs to be pointed out. Collocation number 9 is a collocation that describes quantity or in the case of the example in Table 3, size. By using the collocation LITTLE, the speaker belittles the hearer and makes them believe that they are smaller than them. At the same time, the collocation contrasts with the word 'big', which is also a word that describes size. Keeping in mind that the collocation is used in contexts with the swear word *shit*, the speaker used the collocation LITTLE with the intent to insult a person rather than to describe something or just a situation that the speaker is in like many of the other collocations do. When the speaker intends to insult and hurt, and it is directed at the hearer, adding adjectives describing the size or quantity of the said insult, can increase emphasize the insult.

The collocation MAN, ranking number 4 in the list for the Movies corpus on the collocation list in Table 3, shows a different usage when collaborating with the swear word *shit* than with the other collocations on the table. The examples on the table show two different ways that the collocation MAN may be used with the swear word. The first example shows the collocation as an added afterthought. If we were to rewrite the example without the collocation, the example turns to "holy fucking shit". Without the collocation, the example becomes only an exclamation, not knowing if it is directed at or to a person or at a situation, while the example including the collocation, allows the example to be directed to the hearer. In addition, by including the collocation, it allows for the observation that the person that the sentence is pointed to is someone that the speaker may be close to, such as a friend or relative from the same or close age group, and by using the collocation MAN, it can be lead to the assumption that the person that the speaker is speaking to is of the male gender, because of the meaning and word class of the collocation. The other example for the collocation, on the other hand, has a similar example to the collocation MAN ranking number 6 in the TV corpus on the collocation list. The example makes use of the collocation to describe a person, who is

referred to as "a man shit". This is a typical way to describe someone with bad behaviour. The thing that must be pointed out is the difference in the ranking between the collocation from the Movies corpus to the TV corpus. The collocation has a higher rank in the Movies corpus than the TV corpus by two ranks. This leads to the question as to why one word that is used in two corpora and that has a similar use, has two different ranks. The answer to that is simple. The collocation MAN in the Movies corpus used with the lemma *shit* is used as a direct object and as an adjective in a sentence, which both usages are popular, and that is why the collocation MAN in the Movies corpus is that high on the collocation list. The collocation used in the TV corpus, on the other hand, is used as part of an adjective mostly in a sentence, which is more of an insult, and even though it is popular and frequently used, it does not have the same frequency as the collocation in the Movies corpus.

The theory that I come to expect for the pattern is that the swear words with the same or similar meaning have a high chance of having similar sets of collocates with similar examples. To make a better comparison, swear words with similar meanings will, therefore, be placed after each other to test this theory. Making sure that the words with the same or similar meanings are presented after each other, makes it easier to see any differences and similarities between them. The collocate HOLY gives a different direction for interpretation than the rest of the collocates on the table. The reason behind that is that, like in Table 2 where the collocate FUCK has a different function, it is the same case for the these collocates as well.

Table 4: *Collocates of the swear word "crap" in the Movies and TV corpora, by rank.*

		Crap	
Ranks	Movies corpus	TV corpus	Example:
1	Holy (L)	Holy (L)	**Holy** crap!
2	Game (L/R)	Crap (L&R)	Enough of this dating **game** crap here. Pete, it's a crap **game**. Crap crap crap crap!
3	Crap (L/R)	Just (L&R)	Oh, crap crap crap! This song is crap **just** like everything… If I don't, principle gibbson will **just** crap all over my future
4	Know (3R)	Know (L)	…watch YouTube videos of people falling on their asses than go to some crap movie they **know** is fake. (M) They don't **know** crap.
5	Got (2R)	Got (2R)	All of the crazy crap you **got** stuffed into that belt and you didn't think to bring bug spray All this crap has **got** to go, all of it.
6	Come (R)	Right (R)	Crap! **Come** on! Because that is some Jesus crap **right** there
7	Right ()	Some (L)	None of that pussy football princess crap, **right**? CONFIRMATION Somebody just wanted to stir up **some** crap.
8	Some (L)	Up (R)	Yeah, nothing major, just... **some** crap about perseverance, personal struggles… The crap **up** her nose means she was most likely alive when they dumped her
9	Off (R)	Look (R)	Come on, let's get this crap **off** you. Holy crap, **look** at all this closet space.
10	Man (R)	Man (R)	You're a piece of crap, **man**! How could you talk all this crap, **man**?

The results of the collocation for the swear word *crap*, show that the collocate HOLY is at the top of the list for each of the two corpora. One thing to note in the findings that this result shows is that when it comes to the collocate HOLY, whether the data is found in the TV corpus or the Movies corpus, the instance written in the examples column in Table 4 is the only instance where the collocate *holy* is presented. There are no other examples of the usage of this collocate with the lemma or search word *crap*. There is a similarity between Table 3 and Table 4 in the case of this collocation. The collocation, HOLY, is ranked number 1 in

both the Movies and the TV corpora in the collocation list in Table 4, and that is the same as the results for Table 3. Another likeness that Table 2, Table 3, and Table 4 have in common is the finding of the search words as a collocation, and the usage of it in the examples. For example, in Table 2, the search word is *fuck* and in the collocation list, FUCK is one of the collocations. The same is the case for Table 3 where the search word is *shit* and SHIT is also found on the collocation list. The pattern continues in Table 4, where *crap* is the search word and the collocation in rank numbers 2 and 3 is also CRAP. The pattern continuous on into the example of the usages of these collocations and the fact that they are identical. This leads to the discovery that if the search word is also present in the collocation list, then it is most likely going to be used as an intensifier to emphasize.

Another collocation to pay attention to is the collocation MAN. This collocation is ranked number 10 on the TV corpus list in Table 4, but it is ranked number 10 on the collocation list for the TV corpus on the same Table. This collocation is ranked number 4 in the Movies corpus in Table 3 and number 6 in the TV corpus. The ranking is not so very similar for each of the tables, even though both tables contain a search word with the same or similar meaning to each other. However, the theory that expects search words with the same or similar meaning to have similar sets of collocations is partially proven to be true. The collocation MAN is in rank 10 as a position in the Movies corpus and the TV corpus in Table 4 with identical and similar examples to Table 3 with the search word *shit*, which confirms that the usage of this collocation for both swear words, *shit,* and *crap*, to have the same effect usage from the collocation MAN.

The results of the data show that this swear word may be used in the case of expression. For example, for the collocate HOLY, the expression that may be expressed here is the expression of shock or perhaps even anger. Table 3 shows an example where the collocate is used in a similar manner, which makes the theory of similar swear words stick. The theory goes as follows: swear words with the same or similar meanings tend to be used in similar situations. If we can assume that the speaker just heard something unbelievable, the speaker may express this by uttering "Holy crap!". This means that the collocate *holy* would have to be one word to the left of the swear word. Table 4 shows that in the case of the swear word *crap*, the sets of collocates for both corpora are similar but not so much alike, and the collocates that can be found on the list for both corpora have the priority of the collocates in a different but similar position. Some collocates have the same rank position in both corpus lists on the table, which

can be seen in the cases of for example the collocates, KNOW, and HOLY. These collocates are in the same rank position on the table for both corpora. We can also see that, like the case for the previous tables, there are some sets of collocates that differ from each other. The collocates can be divided into their word classes, nouns, adjectives, adverbs, etc.

The function word collocates show the word class options, while the lexical words give the option of providing more meaningful data. All word classes belong in this category, nouns, adjectives, verbs, and adverbs. This corresponds to table 4, where the word classes are as such.

Table 5: *List of collocates that collocate with the swear word "bullshit"*

Bullshit			
Ranks	Movies corpus	TV corpus	Example:
1	Fucking (L)	Fucking (L)	Just didn't know it was all **fucking** bullshit. Thom, it's **fucking** bullshit.
2	Bullshit (L/R)	Just (L)	Bullshit. Bullshit bullshit! That's **just** bullshit.
3	Man (R)	Know (3R)	… that's some bullshit **man**. That is bullshit and you **know** it.
4	Know (2R)	About (R)	Just tired of all the bullshit you **know**. (relation) It's all that bullshit **about** looking up to me.
5	Just (L)	Right (2R)	**Just** bullshit on a stick if you ask me. Yeah, bullshit is **right**.
6	About (R)	Man (R)	... and you're gon na fucking read a lot of bullshit **about** me. This is bullshit, **man.**
7	Right (R)	Bullshit (R/L)	You do know they're all bullshit, **right**? (confirmation) Bullshit. Bullshit! I know what goes on over there.
8	Like (L)	Such (L)	Well, that smells **like** bullshit.(comparison) Yeah. **Such** bullshit.
9	Such (L)	More (L)	This decision is **such** bullshit. No **more** bullshit.
10	Total (L)	Total (L)	**Total** bullshit. Ugh, it was **total** bullshit.

The next word investigated on the list from Table 1 is the swear word *bullshit*. Table 5 compiles the top 10 words that collocate with this swear word. It's worth noting that *bullshit* shares a similar meaning with two other swear words, *shit* and *crap*. These two swear words

essentially convey the same meaning but employ different words to describe it. In the case of the swear word *bullshit*, it's composed of two distinct parts. Both "bull" and "shit" are referred to as stems in linguistic terms. The first part, "bull," serves as the stem, and its meaning refers to the animal—a bull, which is a type of cow. The second part is derived from one of the two swear words mentioned previously, "-shit," which denotes "the solid waste released from the bowels," applicable to both humans and animals. Interestingly, in the case of "bullshit," the specific animal, the bull, is explicitly specified, resulting in a clearer mental image compared to the other swear words. However, despite sharing one of its components with the other swear words, we might expect the collocations of the swear word, *bullshit*, to be somewhat like those of the other swear words, but this is not the case (Cambridge Dictionary, s.v. "bullshit").

One of the similarities between results for this list of collocates that this table shows from the ones above is that this list has an almost identical set of collocates, with the order of the collocates in different priorities on the list. There is only one set of collocations that Table 5 has that do not match, the collocations LIKE from the Movies corpus and MORE from the TV corpus. The results that Table 5 shows is that between the Movies corpus and the TV corpus, there is almost no difference in the usage of the searched word, *bullshit*. Even though some of the collocates are in different priority order on the list, there is not a big difference between them as their position on the list has around one or two position differences. An example of that is the collocates KNOW and SUCK. In the TV corpus, the collocate KNOW is ranked number 3 and is followed by the collocate ABOUT, while in the Movies corpus, this collocate is ranked number 4 and is followed by the collocate JUST.

There are some collocations, on the other hand, that have the same position in both corpora. The first collocation is FUCKING. This collocation is at the top of the list for both corpuses and is ranked number 1. The fact that the collocation is in this position indicates that the collocation FUCKING is used most frequently with the swear word *bullshit*. This collocation can be found in Table 3, but the number of the ranking is different than that for Table 5. In Table 3, this collocation is ranked number 8 in the list for the Movies corpus, while it is ranked number 10 in the TV corpus. The swear word that is investigated in Table 3 is the swear word *shit*. Even though the swear word for Table 5 is similar in both the writing and the meaning to Table 3, the data show that the results are different for the two swear words. The collocation HOLY is ranked number 1 in Tables 3 and 4, which indicates that the two swear

words, *shit* and *crap*, are similar in that aspect. But while Table 3 and Table 4 are alike in the sense that both tables contain the collocation HOLY as their number 1 on the ranks, Table 3 and Table 5 have the collocation FUCKING in common. Although this collocation is ranked lower in Table 3, it is still a common ground between the tables, which in turn gives another common ground and similarity between the two swear words *shit* and *bullshit*, aside from their similar meaning and spellings. The examples from Table 5 show that the usage of this collocation is the same for both corpora. But this collocation is used in a way to put weight and pressure on the swear word *bullshit* that follows afterward. The usage of this collocation is similar to the usage of the word "absolutely". This can be seen by replacing the collocation and the swear word with this word and testing out if the same point comes across. We take the example from the table and convert it as follows:

A: Thom, it's **fucking** bullshit.
B: Thom, it's **absolutely** ridiculous.

In the example provided by A, both the collocation and the adjective "FUCKING" are employed together, as demonstrated in Table 5. Conversely, the example illustrated by B reveals that, within this context, the adjective "FUCKING" can be substituted with the word "absolutely". However, it's important to note that the adjective "FUCKING" is the present participle of the swear word, *fuck*. As discussed earlier, the utilization of swear words and swearing, in general, serves to intensify and draw attention to the statement being made. In this context, if the adjective "FUCKING" were to be replaced with what is considered a non-swear word, the impact and the edgy quality of the phrase might diminish. The fact that both the collocation and the adjective can be interchanged with non-swear words suggests that this example operates within a figurative framework, emphasizing the figurative, expressive dimension rather than adhering to a strict literal interpretation.

The same holds for the collocation "TOTAL". In Table 5, this collocation maintains the same position on the list for each corpus, consistently ranking at number 10. However, what sets this collocation apart from "FUCKING" is its absence from the lists in Table 3 and Table 4. However, the example featuring the collocate "TOTAL" demonstrates its utilization of the figurative meaning rather than adhering to the literal interpretation (See Table 5). This becomes evident when we modify the example by substituting the swear word *bullshit* with a non-swear word, thus gauging if the same point remains effectively communicated. Consider

the example from the TV corpus in Table 5: "Ugh, it was **total bullshit**". Now, when we replace the swear word with the term "nonsense", the sentences are transformed as follows: "Ugh, it was total nonsense". Both sentences maintain a similar tone and convey the same point. However, the weight and emphasis of the point are noticeably more pronounced compared to the usage of the non-offensive word "nonsense". This accentuation underscores the impact and intensity of the statement.

Another collocation to make note of is the collocation MAN. In the list for the Movies corpus, this collocation has the position ranked number 3, while in the TV corpus, the collocation MAN is ranked number 6. The reason that this collocation is important, is because this word is presented in the top 10 list of the most frequently used collocation among all three swear words presented. The collocation MAN is found in Tables 3,4 and 5. This collocation is another connection that binds all three swear words together and enforces the theory that similar words have similar results. Just like the case for Table 3 and Table 4, this collocation is used more as an independent word, a constituent, rather than as a pair with the swear words.

The verbs presented in the collocation list show the two different functions that they are presented in. The thing that catches the eye about these verb collocations is the fact that they may stand by themselves as they are independent and are not standing in context with the swear word, *bullshit*, which is the same case as the collocate MAN. Most of the rest of the collocations, on the other hand, have a role to emphasize and put more weight on the swear word. For example, the collocation TOTAL, emphasizes making sure to give a clear picture of the swear word, *bullshit*, and the swear word itself. In the case of the verb collocations, however, such as the collocation KNOW and LIKE, the collocations can stand alone without it affecting the swear word in any way. The two functions and roles that these two verb collocations focus on are relation and comparison. As it can be seen in the examples in Table 5, for the collocation KNOW in the rank number 4, the speaker is looking for a confirmation by trying to get the hearer to relate to them. As shown in the example, the speaker is speaking and expressing a feeling that the hearer is likely able to relate to, but the collocation KNOW is used as an afterthought and as a weak confirmation for the hearer to respond to. As for the verb collocation LIKE, the speaker is using this verb as a comparison for the hearer to get a better picture of the situation.

The example:

Well, that smells **like** bullshit.

The swear word *bullshit*, may be taken in the literal sense. When using comparisons in any text, or in this case, speech, the speaker tends to use things that are already known to the hearer as something to compare to. In this case, *bullshit* is something that most if not all people are can picture. The verb "smell" is also a very important factor to make sense of the example and the role of comparison that this collocation has. Since the smell of body waste of animals or people is something familiar, the hearer can understand the picture that the speaker is trying to paint, without the speaker having to say it directly. Even though the swear word itself is used in the literal sense, that does not mean that the sentence in its entirety is used in the literal meaning. The image that the body waste gives, is a negative one, as the stench from it makes everyone escape from it. The case from the example above shows that the speaker is in disbelief and does not believe what had been previously said to them, and therefore, made the comparison of what was said to that of *bullshit*. This means that in the case of this example, the usage of the swear word is used in the figurative meaning rather than the literal meaning (Coulson & Oakley, 2005). But the literal meaning does have the chance of being used for this example, depending on the intention and the situation that the speaker is in. If the speaker is somewhere and they can smell actual *bullshit*, then the speaker is using the literal meaning of the word rather than using it as a swear word. At the same time, the word itself is a compound of two words, one of them being a swear word. The expression 'smells like bullshit' is used when the speaker wants the hearer to know that they do not believe what was previously said strongly and seriously.

Another thing that this collocation list shows is that in the case of the swear word *bullshit*, there is a consistent pattern of usage for this search word. The collocations show that aside from two to three collocations, most of the words in Table 5 contribute to the swear word to put more weight on the point that the speaker is trying to make. In the table, there are two verbs, and around two nouns present, but the rest of the collocations are adverbs and adjectives.

Table 6: *List of collocates that collocate with the swear word "ass"*

Ranks	Movies corpus	TV corpus	Example:
		Ass	
1	Here (L)	Here (L)	LAURIE: Well, get your ass **here**. (literal) After busting my ass **here** for two years, I qualified for a scholarship
2	Back (L)	Kicked (R)	Get your ass **back** to L.A. We **kicked** ass together with the Russians.
3	Down (L)	Back (L)	Get our asses **down** there, please. We need to haul our ass **back** to that restaurant, or get your CB and then call the police.
4	Fucking (L)	All (L)	I'll shove them up your **fucking** ass! No titties, all ass. (literal) vs He's on my ass **all** the time.
5	Right (R)	Down (L)	You were made a temporary supervisor because you kissed the **right** asses. We're shuttin' your ass **down**.
6	All (R)	Right (R)	You like when they butt-naked, showing **all** ass? I mean you've come to the right... I've come to the **right** ass?'
7	Now (L)	Like (R)	The people after you, they're on my ass **now**! Yeah, bright, colorful, cute as a button, tastes **like** ass.
8	Get (L)	Now (L)	You let your ass **get** kicked by some 22-year-old no-name, underdog bitch from the streets. Oh, you defending' his lying ass **now**?
9	Fat (R)	Kick (R)	I don't know, I kind of like the **fat** ass. (literal) Oh man, I got to tell you, you **kick** ass, man. (figurative)
10	Like (R)	Some (R)	Kris, time to wake up. I feel **like** ass, dude. And she's here to kick **some** ass and take names if necessary.

Next, I will talk about the collocations lists for swear words representing body parts. The first swear word is *ass*. Table 6 shows the results of the investigation of this swear word and this table shows the top 10 words that collocate with the swear word. There are many ways to name a body part. This swear word is one way to refer to a person's behind or butt. What this word used to refer to is a donkey.

The only differences in the collocates on the list are the three sets of collocates, FUCKING, GET, and FAT from the Movies corpus, and the collocates KICKED, KICK, and SOME from

the TV corpus. If these three sets of collocates are overlooked, we will see that Table 6 shows that both the TV corpus and the Movies corpus, have the same collocations on their lists, with the difference in their ranking. The difference in the collocation's ranks is not all that different or large. The biggest difference in the ranks is two ranks up or down from each other. For the collocation BACK, for example, it is ranked number 2 on the list for the Movies corpus but is ranked number 3 on the list for the TV corpus.

Table 6 shows that the others above did not show the type of usage of the swear word in collaboration with the collocations on the list. Between the literal meaning and the figurative meaning, the swear word, *ass*, uses is different from the other Tables. In the case of the swear word, *ass*, the literal meaning is more of a possibility than the figurative meaning. The reason behind that is the fact that most people use this swear word to refer to someone's bottom rather than using the non-swear word.

As evident from the examples presented in Table 6, a dominant feature is the use of imperative sentences. For instance, collocates like HERE, GET, BACK, and DOWN are frequently employed in command-like contexts, as exemplified by the sentence, "Well, get your ass here!" Within many of these examples featuring the collocates, it becomes clear that the literal interpretation prevails in approximately half of these instances. This literal usage extends to collocates such as GET, NOW, FAT, and so on. Within this collection of examples, one instance stands out due to its introduction of a distinct sexual context. This example revolves around the collocate ALL, which is featured in both corpora. It showcases a linguistic phenomenon known as "metonymy." Metonymy involves using a term for a part to represent or signify the whole. An analogous instance of metonymy in everyday language includes referring to a "CAR" as "wheels" or a "GUITAR" as "strings." In the context of the collocates discussed, this metonymic usage can be seen in the way certain collocates are employed to symbolize the broader, often literal meaning associated with the swear word *pussy*.

Table 7: *List of collocates that collocate with the swear word "fanny"*

		Fanny	
Ranks	Movies corpus	TV corpus	Example:
1	Miss (L)	Pack (L)	And soon **Miss** Fanny Bright was seated by my side… So what if a tour came by and some **fanny** pack fatty took a picture…
2	Pack (R)	Madam (L)	The fanny **pack** is fine. No notes from **madame** Fanny.
3	Fanny (L/ R)	Miss (L)	Fanny. Fanny I do not quite know what to make of **Miss** fanny.
4	Aunt (L)	Fanny (L&R)	Here is your **aunt's** fanny. Fanny, fanny, fanny fanny fanny!
5	Price (R)	Packs (R)	Fanny **Price**! (name of a person) Fanny **packs** are out, bitch.
6	Dear (R)	Dear (L)	Come along, Fanny **dear**. **Dear** Fanny, …
7	Fun (R)	Brice (R)	A bit of fanny **fun**. (sexual) Yes, Fanny **Brice**?
8	Darling (L)	Aunt (L)	**Darling** Fanny, I know whatever you did…. My **aunt** Fanny.
9	Brice (R)	Price (R)	Miss Fanny **Brice**. Marry Fanny **Price**?
10	Hill (R)	Lady (L)	Martha dear, come and meet Fanny **Price**. Smart play, **lady** Fanny, you play the butler?

The findings of the data that Table 7 presents are almost entirely different from Tables that have been presented so far. The data show the collocations that collocate with the swear word *fanny*. As mentioned before, this swear word is an old word. As we saw from the data in chapter 5, the swear word *ass* is more frequently used than the swear word *fanny*. It appears that age or time may have been a factor influencing the data presented in Table 7. Even though the searched word is an old word that is mostly used in the British varieties of spoken English, the results show that the swear word *fanny* is mostly taken in context with people, and is used as the name of females, which can be clearly shown through the collocations such as MISS, MADAM, and AUNT.

One interesting fact that must be pointed out is the fact that by filtering the collocations and trying to include only the lexical words in the collocation list in Table 7, the less the

frequency hits became for this search word. This means that for example, even though the collocation MISS and PACK are on the top of the list in Table 7, these collocations are very low on the original, unfiltered collocations list from the corpora. The reason for this may have to do with the fact that the swear word *fanny* is used as a name, both a first name and a last name. Contrary to the meaning of the swear word, when used as a name, Senior Content Writer Avril Whelehan wrote in *The Bump* that the word is a diminutive of the French name "Estafany", which means "crown", and that "Fanny" may also be interpreted as "someone from France" (2023). The name has a nice ring to it and is ideal for naming a baby, who is bound to be "the crowning joy in your life". In English, the word is a name given to females and it has the meaning "free", which is why it is so popular, and this can be seen from the collocations list where most of the collocations refer to people, mostly females.

In the earlier Tables, we can see the repetition of the search word in the list. This is the case for Table 7 as well. The search word *fanny* is seen in Table 7, both in rank number 3 in the Movies corpus and in rank number 4 in the TV corpus. Considering the pattern, it is expected for this collocation to follow it as well. But these results show something else. As we can see in the examples, the collocation is repeated along with the search word. But the effect given from this repetition is different. According to the example, when the collocation FANNY is used with the search word *fanny* in repetition, it is to call the person with the name. In the example for the collocation in rank number 3, the context that the word was used in shows that the speaker was calling the name, "Fanny", in a scolding tone.

Two collocations do not refer to people in the collocations list in Table 7, one of them uses the search word as the first meaning, which refers to the bottom of a person. The first collocation is PACK. This collocation refers to a waist bag, and this can be seen in the example for both collocations, PACK, and PACKS, as they both are the same word with the difference of one being a singular and the other being plural. The examples that the collocations are written in lead to an object. The only other collocation that stands with the search word *fanny* with the meaning of a person's behind or bottom, is the collocation FUN. This collocation is number 7 in the ranks, and it is only mentioned in the Movies corpus. The fact that this collocation is this low on the list suggests that this meaning of the search word is not prioritized in the English spoken language anymore and that the word *fanny* may not be considered as a swear word in the possible near future for the lack of the usage of the word. The thing that is worth noting in Table 7 is that this search word is mostly used in UK/IE and

AU/NZ varieties of spoken English. There were hardly any hits for this search word for the US/CA variety of spoken English.

Table 8: *List of collocates that collocate with the swear word "dick"*

		Dick	
Ranks	Movies corpus	TV corpus	Example:
1	Dick (R&L)	Dick (R&L)	I called you a dick dick. Dick dick dick …. Dick dick dick dick dick…
2	Big (R)	Big (R)	Anyway, yeah I'm so Dominican, look it, I got a **big** dick and a bad temper. …or else I probably would've worn my **big** dick jeans!
3	Fucking (R)	Moby (R)	You're the one driving like a **fucking** dick. I love **Moby** dick.
4	Off (L)	Off (L)	Don't let him pull his own dick **off**. The monkey would bite the Rock's dick **off**.
5	Like (R)	Like (R)	God, it smells **like** dick. Now, I don't **like** dick any more than you do, but he does, but that's his business
6	Moby (R)	Know (R)	There is a great **Moby** dick illustrated… You **know** Dick's guilty, right?
7	Out (L)	Out (L)	Milton Berle unzips his pants and whips his dick **out.** You pulled your dick **out** and I touched your dick.
8	Know (R)	Fucking (L)	… and wonder boy Walter over here don't **know** dick about stage fighting. Suck on my **fucking** dick.
9	Get (R)	Total (R)	Your dick **get** bigger, too, or something? I was a **total** dick to you.
10	Hard (L)	Mouth (3R)	That'll make your dick **hard**. …she is a sensualist who sees no real difference between a pastrami sandwich and a dick in the **mouth.**

Table 8 shows a very different scenario than what Table 7 shows. In Table 8, the collocations presented in the table are the top 10 collocations that collocate with the swear word *dick*. The results that the data presents show that the top collocation for both corpora is the collocation DICK. This collocation is used in the same manner as the collocations FUCK, SHIT, CRAP, where they are using replication as an intensifier to emphasize, but instead of it being used as

an exclamation, it is used as a reminder for the hearer. The way this collocate is used may be in a manner of anger or teasing by repeating the same word over again to make a jab at the hearer. As it can be seen in the examples for this collocation, the speaker is using the collocation DICK and the search word *dick* to bring forth an insult upon the hearer. Being called a 'dick' may not necessarily mean that the speaker means the male genitals. The speaker may intend for the hearer to interpret the word in another way depending on the situation. If we replace the collocation DICK and the swear word in the example for the collocation in rank number 1, we get "I called you a jerk…". This reflects more on the behaviour of the person the speaker is talking to rather than indicating the sex organ itself. This is another example of figurative meaning rather than literal meaning. Since the figurative meaning is being used through this collocation, one of Grice's maxims may be broken as well. One of the requirements for the Maxim of Quality says, "Do not say what you believe to be false"(Kroeger, 2018, p. 142). The example for the collocation DICK, where the speaker is calling the hearer "a dick", is therefore, breaking this maxim as the speaker is talking to a person and not the sex organ. But this is only the case if the example has a literal meaning to it. Judging by the context of the examples, the speaker most likely used the collocation DICK in the figurative meaning. If assumed that the collocation DICK means the noun "jerk" instead then the maxim is not broken as the speaker is only voicing their opinion of the person, they are speaking to, and we can assume that they believe the statement to be true. The collocation TOTAL from the TV corpus is presented similarly to the collocate DICK from the Movies corpus. If the meaning used in the context is the literal meaning, then Grice's maxim is broken. However, as the example is presenting a common expression, it is best to assume that the meaning used in this example is the figurative meaning rather than the literal. The example shows the speaker admitting to being a dick which can be translated to being a jerk, and therefore has a similar effect as the collocate DICK, without the replication of this collocate.

The fact that most of the collocations in Table 8 are taken literal meaning is one of the main points to make. But there are a few collocations on the list in this Table that do not have anything to do with the actual sex organ, a dick, but are referring to people or to describe a situation. The collocation MOBY is mentioning a character in a famous novel, Moby Dick. The word, *dick*, does not always refer to the actual sex organ, as can be seen in Table 8, but may also refer to a certain attitude or behaviour. The first examples from the collocation DICK are proof of this. Another collocate that does not contribute to referring to the sex

organ itself but rather a person is the collocation KNOW. The collocate refers to the knowledge of a person's guilt over something who goes by the name Dick. These two collocates are the only two collocates that do not use the search word, *dick*, as a swear word in any way, either the literal meaning or the figurative meaning. Grice's Maxims are, therefore, not broken in the examples for the two collocations MOBY and KNOW. The rest of the collocates in Table 8 show a sexual context in the examples and the collocates are helping the swear word, *dick*, to be describing the state of the sex organ, like with the collocate HARD or the collocate BIG. The examples for these collocations show the usage of the literal meaning, therefore, making sure that Grice's Maxims of Communications (Kroeger, 2018, p. 142) are not broken. For the collocation HARD, the example goes as follows "That'll make your dick hard". The collocation is an adjective that describes the state of the noun it belongs to, which in this case is *dick*, and this example cannot be taken for anything other than it's literal meaning.

Table 9: Collocates list of the swear word "pussy"

Pussy			
Ranks	Movies corpus	TV corpus	Example:
1	Fucking (L)	Fucking (L)	You're a **fucking** pussy bitch. I can't even feel you, you **fucking** pussy.
2	Like ()	Like (R)	Smells **like** pussy. Great,' cause this place smells **like** pussy cream and I wish I was dead.
3	Some (R)	Little (R)	You had time to come get **some** pussy. Someone who's not a fucking **little** pussy! -
4	Little (R)	Cat (R)	You really are a **little** pussy, aren't you, huh? We're done, pussy **cat**.
5	Man (R)	Out (2L)	You're a pussy, **man.** It's a few good years before your pussy wears **out**.
6	Here (R)	About (2L)	**Here** pussy, pussy, pussy, pussy. The pussy you **about** to get, my [man].
7	Fat (L)	Some (R)	…and he's calling you a bitch, and a big **fat** pussy… Do you think I'd trust **some** pussy fanner to touch my head on a day like today?
8	Got (R)	There (2L)	Every time he **got** pussy, he thought he was in love. Yo, that's why I never licked no pussy right **there.**
9	Out (2L)	Broken (L)	I'll throw that loose ass pussy back **out** on the street! Where's the **broken** pussy?
10	About (R)	Big (r)	I like when Lil Wayne talks **about** pussy, because he's … Quit feeling sorry for yourself and step up to the plate, you **big** pussy.

The top 10 collocates that occur most frequently with the swear word, *pussy*, are presented in order from rank 1 to rank 10 in Table 9. Table 9 shows some similarities between the two lists for the TV corpus and the Movies corpus. First, the collocation FUCKING is in rank 1 for both corpora lists. Just like the other Tables, we can see the similarities and differences in Table 9 for the TV corpus list and the Movies corpus list. The four collocates, MAN, HERE, FAT, and GOT in the TV corpus list do not exist in the list for the Movies corpus. At the same time, the four collocates, CAT, THERE, BROKEN, and BIG from the Movies corpus list are not presented in the list for the TV corpus. Despite that, there are still some similarities between these different sets. The collocation HERE from the Movies corpus and the collocation THERE from the TV corpus are both adverbs. But in the example for the collocate HERE, it is shown that the speaker is calling the hearer over to them, or just trying to get their attention. The speaker in this example is using the swear word *pussy* as either a nickname for a cat or to call a cat over to them. As for the collocate THERE, on the other hand, it is clear from the example that the collocate is used to indicate a specific location or direction, rather than as an attention-getting device. The same is the case for the collocation FAT from the Movies corpus list and the collocation BROKEN from the TV corpus list. The link between them is their word class, which both are adjectives. From the example of the collocate FAT, one of the two interpretations of the swear word, *pussy*, may be considered. The first is the physical description of the female sex organ. The second is, rather than physical description of the female sex organ, the collocate FAT contributes to the other definition of the swear word, *pussy*, which is 'coward'. In the second definition of the swear word, pussy, the collocate FAT intensifies the second meaning of the swear word by giving a clear mental image from the speaker to the hearer.

If the word *pussy* is used to refer to a "cat", it indeed assumes a non-swear word sense, like how *fanny* and *dick* are not considered swear words when used as personal names. The distinction lies in the context and intended meaning of these words. The definition of semantic prosody, encompasses the inherent connotations or associations that words or phrases may carry beyond their literal definitions. Semantic prosody can range from sexual or negative connotations to more positive ones, depending on how a word is used within specific contexts.

The subsequent explanation regarding the collocate CAT being employed as a nickname or pet name within a romantic context and its prevalence in the US/CA variety of spoken

English offers valuable insight. This usage of CAT sheds light on the semantic prosody of the swear word *pussy* in this specific context, signifying a more affectionate or endearing application compared to its customary offensive connotations. Remarkably, this usage confers upon the swear word *pussy* a positive semantic prosody, despite its original negative connotations. Furthermore, reinforcing the linguistic analysis, the mention that the collocate MAN functions as a constituent underlines the structural significance of these collocations within the language and discourse.

The collocate MAN listed in Table 9 for the swear word *pussy* serves a function like that of other collocates found in previous tables. Much like in Table 2 for the swear word *fuck*, Table 3 for *shit*, Table 4 for *crap*, and Table 5 for *bullshit*, the collocate MAN functions as a constituent. In this context, the term "constituent" denotes that the collocate MAN is an integral part of the collocative structure surrounding the swear word *pussy*. It plays a role in shaping the overall meaning and usage of the expression, contributing to the rich tapestry of language dynamics explored in these Tables. This consistent pattern across various swear words and their respective collocates highlights the recurring linguistic mechanisms and strategies employed within the corpus.

The examples listed in Table 9 exhibit an interplay between the literal and figurative meanings of their respective collocates. In particular, the figurative meaning of the swear word *pussy* is equated with the non-swear word "coward". This figurative usage introduces a dimension to the communication, potentially involving the breaking of Grice's Maxims of Communication (Kroeger, 2018, p. 142) in these instances found in Table 9. For instance, in the case of the collocate LIKE paired with the verb "smell", as seen in the phrase "smells **like pussy**", a negative connotation is evoked. This construction implies an unpleasant or foul odour, aligning with the figurative interpretation of the word *pussy* as a derogatory term synonymous with "coward". However, it's essential to recognize that the context can vary, and depending on the situation, this expression might be intended literally in a sexual context, as demonstrated by the example in the Movies corpus. Conversely, the example from the TV corpus illustrates that the collocate LIKE is used to describe either a negative situation experienced by the speaker or the unpleasant quality of the air in their current location. The addition of the statement "wish I was dead" following this description underscores the intensity of the aversion, indicating that the speaker is so distressed by the odour that they wish for a drastic change in their circumstances, even to the point of wishing for death. These

varied uses of collocates and the figurative interpretations they invoke underscore the intricacies of language and the role of context in shaping meaning and impact within different communicative contexts.

Indeed, there are additional examples of the collocate LIKE in both the Movies corpus and the TV corpus that provide further insight into its usage. One such example is the sentence, "How could a pussy like you kill someone?" In this context, LIKE can be interpreted to liken the person in question to a "coward", effectively conveying, "How could someone who is perceived as a coward like you commit murder?" Here, the collocate LITTLE serves to draw a direct connection between the individual being addressed and the concept of cowardice.

In Table 9, the usage of the collocate OUT is shown by two different examples, one from the Movies corpus list, and the other from the TV corpus list. The example for the collocate for the Movies corpus list is part of the verb, "to wear out". This means to use something until it is useless, or to do something and then becomes too tired to do it again. The former meaning is applied in the example for the collocate OUT from the Movies corpus list. The example for the collocate OUT from the TV corpus list, on the other hand, is used as an adverbial particle. The example goes as follows, "I'll throw that loose ass pussy back **out** on the street!". Both corpora do not limit the usage of the collocate OUT alongside the swear word, *pussy*, only as the two ways mentioned, but they provide other examples of the usage of the collocate OUT and how the collocate affects or gives effect to the swear word. One of those examples is, "Don't pussy **out** on me now".

The example for the collocate BIG shows one of the options of usage for the swear word *pussy*. The TV corpus illustrates another usage of the collocate in the same context as the swear word. The example for this collocation presented in Table 9 is intensifying the swear word. But the example uses the swear word, *pussy*, as a replacement for the non-swear word "coward". The speaker in this example is calling the recipient a coward but using the swear word for better effect. This is a case of the usage of the figurative meaning in action. The example presented for the collocation BIG focuses on emphasizing the speaker's opinion of the hearer. In Table 9, the collocate, BIG, is only presented on the list for the TV corpus. The list in Table 9 focuses on only the top 10 lexical collocates, which means that if this collocate is available in the Movies corpus, it is not so frequent that it makes the list for Table 9.

But that is not the case for every instance and concordance hit for the TV corpus. The corpus provides other examples where the collocation BIG. One of the instances of this concordance is by describing the female sex organ and using it in the literal sense. An example of this is presented as follows, "She has a **big** pussy. She fucks all the time". Aside from being taken for the literal meaning, this example also focuses on the sexual aspect of the swear word. Depending on the context, Grice's Maxims of Conversations (Kroeger, 2018, p. 142) may be broken here, even though this example uses the literal meaning. If the speaker intends to defame the person they are mentioning and is aware that what they are saying is false, the maxim of quality will be broken in this case (More details in Chapter 7).

Table 10: *List of collocates that collocate with the swear word "twat"*

			Twat
Ranks	Movies corpus	TV corpus	Example:
1	Fucking (L)	Twat (R&L)	You **fucking** twat! Twat. Twat. Twat, twat, twat, twat, twat, twat,
2	Little (R)	Fucking (2L)	So why the fuck did you behave like such a **little** twat yesterday? **Fucking** arsehole twat!
3	Get (L)	But (L)	Hey! See that skinny twat, **get** him out of there! Yes, I am twat, **but** I was like that before, I can't help it.
4	Mr (R)	Stupid (L)	Look, **Mr** Twat… I'm her dad, you **stupid** twat!
5	All (2L)	All (L)	What kind of a twat goes **all** in on two pair? Don't be a twat **all** your life.
6	Here (2L)	Some (R)	Get your freakin' minimum-wage, $10-an-hour twat over **here**. (ADVERB OF PLACE) You jeopardise everything because **some** twat takes a shot at us?
7	Old (R)	Fat (L)	Shut the hell up, you dried up **old** twat! I'll have your job away for this, you **fat** twat.
8	Complete (R)	Posh (R)	What am I, a **complete** twat? To Albert, you're just a **posh** twat...
9	Fat (L)	Daft (L)	Oi, knobhead! **Fat** twat! No, you **daft** twat!
10	Shut (3R)	Old (R)	… she is a stupid twat that should **shut** up.. Instead of heckling my friends, I am gon na call that **old** twat

Table 10 presents the top 10 collocates that collocate with the swear word *twat*. According to Cambridge Dictionary, the swear word *twat*, is a noun and has the meaning of the female sex organ, or as it can be called a non-swear word vagina (Cambridge Dictionary, s.v. "twat"). The Dictionary also mentions that the swear word, *twat*, is the offensive way word for the sex

organ, mostly used in the British varieties of spoken English (Cambridge Dictionary, s.v. "twat"). The lists presented in Table 10 show that there are a lot of differences between the Movies corpus and the TV corpus as there are only 4 collocates that the two lists have in common. The rest of the collocates on the list for Table 10 are not presented in both lists, which suggests two different reasons for it. Reasoning number one is that the collocates may not be presented in the other corpus. Reasoning number two, the collocates are present in the other corpus, but have a different ranking in the corpus, which in turn means that they did not make it on the lists in Table 10.

Table 10 shows that the collocates FUCKING, ALL, OLD, and FAT are the only collocates that are presented in both the TV corpus and the Movies corpus lists. The rankings of these four collocates are different. The only collocate that has the same ranking for both corpus lists is the collocate ALL, which is ranked number 5 in Table 10. The examples for this collocate show that the collocate ALL is mostly in connection with either the expression presented in the example in the Movies corpus or part of the phrase later phrase in the TV corpus. Either way, the collocate is not dependent on the swear word *twat* or affect it in a big way. The example from the Movies corpus shows the collocate as part of an expression, which is "to go all in". The example is using the swear word *twat* in figuratively rather than literally. What the example means after replacing the swear word with a non-swear word is, "What kind of idiot goes all in on a pair?". The example from the TV corpus, on the other hand, is part of an imperative clause, where the swear word is the object and the head of a noun phrase, which makes the collocate ALL independent from the swear word *twat*.

The collocation TWAT in Table 10 with the swear word *twat*, has the same function as the collocation FUCK in Table 2 with the swear word *fuck*. There is a case of reduplication, where the swear word gets repeated. By having both the swear word and the collocate being the same, the reduplication becomes clearer. Reduplication happens to emphasize the point. The repetition of the swear word and collocate TWAT is not just aimless curses at the air because of emotions. It is more used to call someone, even though it is offensive, which is most likely the point. The collocation FUCKING on Table 10 is ranked number 1 on the Movies corpus list but is ranked number 2 on the TV corpus list. The examples from both lists for this collocate show that the collocate FUCKING is acting as an intensifier. The difference between the effect of these collocates used as an intensifier and the effect of an intensifier using replication is the effect of repetition. The collocate FUCKING in Table 10 is used as a

determiner in the example sentences in Table 10. Table 10 presents more adjectives than the other tables. Another point to make is the fact that most of the hits for all the collocates in Table 10 show the concordances are mostly in the UK/IE variety of spoken English. The swear word, *twat*, is originally the offensive British version of the non-swear word "vagina", which makes the results of the data something expected from the beginning. There were, however, very few hits that showed concordance in the US/CA variety of spoken English. These hits, however, were more of a narration of what someone else said, for example, "What a twat, he said…" etc. The more we go down the list in Table 10, the less the frequency there is for the hits on the collocates. Most of the collocates that are shown in each corpus are function words. The reason for that is because the swear word, *twat*, is a noun. This makes the hits focus on the collocates, such as articles, to support the swear word as a word class.

Table 11: *List of collocates that collocate with the swear word "bitch"*

Bitch			
Ranks	Movies corpus	TV corpus	Example:
1	Fucking (L)	Little (L)	You stupid, **fucking** bitch!
			You are a great ungrateful **little** bitch!
2	Little (L)	Hey (L)	Milo is with this **little** bitch
			Hey, bitch, wake up!
3	Get (R)	Crazy (L)	Bitch **get** off my pole?
			You **crazy** bitch. (all hits, this example)
4	On (R)	On (R/L)	Get this bitch **on** a leash!
			Bitch **on** bitch Makes for very bankable entertainment.
5	Come (R)	Get (R)	You bitch **come** here!
			Do not let that bitch **get** the best of you, Mary Jane.
6	Hey (L)	Are (L)	**Hey** bitch, wake up.
			Attention, y'all **are** bitch ass hoes.
7	Are (R)	All (L)	Black Jack bitch **are** you fuckin' for real?
			Manny's crew **all** bitch about him
8	Bitch (R&L)	Fucking ()	I'm a boss-ass bitch Bitch, bitch, bitch, bitch
			She was a **fucking** bitch
9	Crazy (L)	Stupid (L)	You **crazy** bitch.
			You **stupid** bitch!
10	Stupid (L)	Bitch (R&L)	We took care of his **stupid** bitch wife.
			Bitch bitch bitch?

Table 11 presents the top 10 lexical words that collocate with the swear word *bitch*. This Table presents five verbs that collocate with the swear word, *bitch*. These sets are as follows, from the Movies corpus list we have the three collocates GET, COME, and ARE. From the TV corpus, we have the following verb collocates, GET and ARE. The verb collocation COME from the Movies corpus is not presented in the TV corpus in Table 11. This collocation is ranked number 5 on the list in the Movies corpus list in Table 11, which means that it is right in the middle in terms of frequent usage of the swear word, *bitch*. The collocate COME gives the sound of a command with it. As it can be seen in the example for collocate number 5 in Table 11, the speaker is giving a command to the hearer and names specifically where the speaker wants the hearer to come, 'You *bitch* **come** here!'. The verb collocation GET is ranked number 3 in the Movies corpus list in Table 11, while in the TV corpus list, it is ranked number 5. The collocate GET from the TV corpus list is the finite verb of the sentence, and does not effect the swear word, *bitch*. Looking at the examples of the usage of the collocate in the two corpora, the collocate GET from the Movies corpus list, on the other hand, gives the swear word, *bitch*, the same function as the collocate MAN from the previous Tables. The speaker is using the swear word to replace the hearer's name, and by refusing to call their name, the speaker intends to hurt the hearer. The last verb collocate to mention is the collocate ARE. This collocate is the present participle of the infinitive verb, to be. The usage of the collocate ARE in the example for the Movies corpus differs for the example from the TV corpus. The example from the Movies corpus shows that the collocate has the same effect on the swear word, *bitch*, as the verb collocate GET from the Movies corpus has, where the swear word is being used independently to call someone instead of their name. The collocate ARE from the TV corpus has also a similar function as the collocate GET from the TV corpus.

The next collocates are the collocates FUCKING, CRAZY, STUPID and LITTLE. These four collocates are presented in both lists in Table 11. The only difference between them is their different ranks on both the Movies corpus list and the TV corpus list. For example, the collocate FUCKING is ranked number 1 on the Movies corpus list, while it is ranked number 8 on the TV corpus list in Table 11. This collocate switches from being ranked at the top of the list for one list to being ranked third to last on the other list. The collocate CRAZY presents the same case scenario, where it is ranked number 9 on the Movies corpus list, while in the TV corpus list, it is ranked number 3. Although the difference in the rankings is not as wide as for the collocate FUCKING, it is still considered a big gap in the rankings for the

same collocate using two different but similar corpora. The two other collocates STUPID and LITTLE, on the other hand, do not have that big of a difference in their ranks between the two lists. The collocate STUPID, for example, is ranked number 10 in the Movies corpus list, and in the TV corpus list, it is ranked number 9. There is only one rank difference between the lists. The same case is for the collocate LITTLE. In the Movies corpus list, it is ranked number 2, while in the TV corpus list it is ranked number 1. The one rank difference shows that the difference between the collocates is not all that different between the two corpora for the collocates STUPID and LITTLE, and that is evident from the examples presented in Table 11. Aside from the collocate FUCKING, the other three collocates CRAZY, STUPID, and LITTLE are adjectives. Adjectives can be used as intensifiers. The collocates mentioned here are acting as intensifiers, and that is one of the things that the four collocates FUCKING, CRAZY, STUPID, and LITTLE have in common. The collocate BITCH, in the Movies corpus is ranked number 8, while in the TV corpus is ranked number 10. The collocate is used with the swear word, *bitch*, which makes it a repetition. Repetition is used as an intensifier, as can be seen in the examples for both the Movies corpus list and the TV corpus list in Table 11.

In the case of the collocate ALL, Table 11 shows the usage of this collocate with the swear word, *bitch*, in one way. This collocate is a determiner. The example presented for the collocate ALL from the TV corpus list shows that it has the function of a quantifier of the sentence. The example shows that the speaker is referring to 'everyone' mentioned in the context. The collocate ALL is, therefore, affecting the function of the swear word. The swear word, *bitch*, is being used as the finite verb in the sentence in the example from Table 11. The last collocate from Table 11 that needs to be mentioned is collocate ARE. This collocate is a verb. Despite being presented in both the Movies corpus list and in the TV corpus list in Table 11, the usage of the collocate ARE in the two examples differ from one another, and the effect that this collocate has on the swear word, *bitch*, differs as well. The collocate ARE presented in the Movies corpus list, shows the usage of the collocate as the finite verb in the sentence presented one position to the left of the swear word (L). But it is not affecting the swear word, *bitch*, as both the collocate and the swear word are independent of one another. This means that the swear word is a constituent. The swear word, *bitch*, from the example presented for the Movies corpus, shows the speaker is using it instead of calling the hearer's name. The example for the verb collocation ARE, presented in the TV corpus list, on the other hand, shows that the swear word, *bitch*, is used after the collocate (1R). Even though the collocate

ARE in both examples from both corpora have the function of the finite verb of the sentences, the position of the swear word, *bitch*, differs. In the example for the TV corpus list, the swear word is part of the direct object clause of the sentence and is used as an adjective of the clause. There is a common ground between all the collocations presented in Table 11. The examples for the collocations show that the speaker is referring to the female hearer. An example of that would be for the collocate STUPID, where the collocate and the swear word, *bitch*, both refer to someone that the speaker is referring to as 'his wife'. The swear word, *bitch*, according to Cambridge Dictionary, has the meaning of 'female dog'(Cambridge Dictionary, s.v. "bitch"), which means that when using this swear word, it is most likely pointed towards women rather than men. But there are cases in which men try to put each other down by calling a man a 'female dog'. Another important point to make is that most if not all of the hits are in the US/CA variety of spoken English. This means that the swear word, *bitch*, is most used in general American and Canadian rather than for example British.

Table 12: *List of collocates that collocate with the swear word "bastard"*

			Bastard
Ranks	Movies corpus	TV corpus	Example:
1	Fucking (L)	Poor (L)	Bleed it all out, you **fucking** bastard. I'm actually starting to feel sorry for the **poor** bastard.
2	Poor (L)	Sick (L)	She's gon na take the **poor** bastard for everything he's got. And you thought it was okay to let her handle that **sick** bastard by herself?
3	Sick (L)	Fat (L)	They were siblings, you **sick** bastard! Piss off, you **fat** bastard!
4	Crazy (L)	Lucky (L)	I'm not gon na let that **crazy** bastard eat him. You **lucky** bastard, you.
5	Bastard (2L/2R)	Lying (L)	In this world will always be bastard and bastard alone. You **lying** bastard!
6	Dirty (L)	Son (R)	You **dirty** bastard. You bastard **son** of a bitch!
7	Lucky (L)	Bastard (R/L)	I hate you, you **lucky** bastard. Bastard! Bastard! Bastard! No! Please, please. You bastard!
8	Fat (L)	Rat (L)	You put on some pounds, you big **fat** bastard. I need you to find and kill this ballsack, **Rat** Bastard.
9	Son (R)	Crazy (L)	I'm nothing but the bastard **son** of a prostitute. Why do you hang out with that fucking **crazy** bastard?
10	Lying (L)	Fucking (L)	I've got one **lying** bastard and one... it's him that's to blame... the filthy **fucking** bastard.

The last Table that is presented for the topic collocations of the swear words is Table 12. This table presents the top 10 collocations that collocate with the swear word, *bastard*. According to Cambridge Dictionary, the literal meaning of the word, *bastard*, is 'an illegitimate child born out of wedlock'(Cambridge Dictionary, s.v. "bastard"). This word is used as an insult today, directed towards anyone without prior knowledge of the person's birth. In Table 12, the examples for both the Movies and the TV corpora, show the usage of the swear word used as this type of insult. There are a few collocations and their examples in Table 12 that need to be explained further. First, the collocation FUCKING and its function. The collocate FUCKING is presented in Table 12 on both the Movies corpus list and the TV corpus list. The usage from both examples is identical. The collocate FUCKING is an adjective and is being used as an intensifier for the swear word, *bastard*. The speaker in both examples shows that the speaker is just throwing an insult at the intended person they are referring to. In the example for the Movies corpus, the speaker is calling the hearer a 'fucking bastard' directly to their face. In the case of the example presented in the TV corpus list, on the other hand, the speaker is referring to and describing a third party as a 'filthy fucking bastard' by adding an adjective in front of the collocate and swear word.

Another thing to mention is the ranking for the collocate FUCKING. There is a huge ranking difference for the collocate FUCKING in Table 12. In the Movies corpus list, the ranking for the collocate FUCKING is ranking number 1, which is the top of the list. The ranking for this collocation in Table 12 in the TV corpus list, on the other hand, has the ranking number 10, which is at the bottom of the list. This big of a gap shows that on tv people do not use the collocate FUCKING with the swear word, *bastard*, as much as in the movies. The further down the list in the corpora we go, the fewer frequency hits we get. The collocate FUCKING has less frequency hits in the TV corpus than in the Movies corpus.

There is a usage of repetition for the collocate BASTARD in Table 12. In the other Tables, there are instances where the swear word in question is collocating with the collocate on the list. For example, in Table 2, the collocate is FUCK and the swear word is *fuck*. The example turned out to be as follows '*fuck fuck fuck* etc'. In the case of Table 12, there are two instances where the collocate BASTARD is collocating with the swear word, *bastard*. In the Movies corpus, the collocate BASTARD is ranked number 5, while in the TV corpus, it is ranked number 7. There are only two rank differences between them, which means that the collocate BASTARD is almost equal in terms of frequency and usage with the swear word *bastard*. The

usage of this collocation with the swear word is demonstrated in two different ways. The example from the Movies corpus list shows the usages of the coordinating conjunction 'and' between the collocate and the swear word, '…bastard and bastard…'.

Most of the collocates on both the Movies corpus list and the TV corpus list in Table 12 are adjectives such as the collocates FAT, SICK, DIRTY etc. These are adjectives that describe the state or physical character of a person. Their semantic prosody is mostly negative. This is reflected in the example for each of their examples from both corpora. Semantic prosody analysis reveals that the majority of collocates associated with the swear word, *bastard* have a predominantly negative connotation. However, an exception is found with the collocate LUCKY, which carries a positive meaning. In contrast, the swear word *bastard* tends to be associated with a negative semantic prosody. Table 12 shows that the corpora have an identical set of collocates for both lists.

6.2 Summery

I will conclude this chapter by pointing out the main findings that the data has provided for this paper. Chapter 5 focused on presenting the data that was found in both the Movies corpus and the TV corpus. There are two main topics that the findings focused on. The first topic is the findings of the collocations of each of the swear words presented in Table 1 in Chapter 2, and the second is the frequency of the swear words from the same Table. The data collected is available in these two corpora.

6.2.1: Patterns

First, I will explore some patterns that emerged from the data analysis:

- Intensifier Use through Reduplication: A notable pattern is the use of intensifiers through reduplication, as seen in examples like "SHIT, SHIT, SHIT." This repetition of the swear word itself serves to amplify its impact in a sentence.

- Similar Collocates for Near-Synonyms: Another pattern observed is that near-synonyms often share similar collocates. For instance, both "SHIT" and "CRAP" tend to appear with comparable sets of words, indicating their interchangeable use to some extent.

- Function Words in Collocates: Collocates consisting of function words primarily reveal the word class and are generally less interesting. They provide limited contextual information compared to collocates with content words.

- Consistency Across TV and Movies Corpora: For most of the swear words, both the TV and Movies corpora exhibit similar collocates. This consistency suggests that these words maintain their usage patterns across different forms of media, reflecting their enduring prominence in informal spoken English.

The corpora are extensive to ensure statistical significance, resulting in a substantial overlap between the TV and Movie corpora in terms of collocates. However, discrepancies can arise due to factors such as the low frequency of individual words or inherent distinctions between the two corpora, especially considering that movies are typically scripted to a greater extent than television.

As we delve further down the collocation lists, the results tend to become more unpredictable. In the data collection process, it became evident that as we moved down the lists within the corpora, the frequency of the presented results decreased. Consequently, this variation in frequency is a key factor behind the disparities observed in the rankings and the presence or absence of different collocations between the Movie corpus and the TV corpus.

6.2.2: Ranking and Adjectival Usage

Another aspect of the analysis involves disparities in rank. For instance, a stark difference in rank is observed in the case of the collocate "FUCKING" when comparing its position in the TV corpus to that in the Movies corpus, as illustrated in Table 11. This discrepancy in ranking points to variations in the usage and frequency of specific collocates between the two corpora, shedding light on the nuanced differences in language patterns across these mediums. Such differences in rank can offer valuable insights into the distinctive linguistic characteristics of television and film contexts, highlighting the impact of media genre and scripting on language use.

Among the analysed swear words, only *fuck* exhibits adjective usage, marked by its conjugation into "fucking". In contrast, the other swear words, such as "*crap*" and "*twat,*" do not undergo this transformation. For example, "*crap*" does not change to "crapping," and

"*twat*" remains "*twat*" without any analogous alteration. This differentiation underscores the unique linguistic characteristics of *fuck* and its adaptability compared to the other swear words. Another observation emerges when comparing swear words like *crap* and *shit*. While both words function as nouns, their usage in replication differs. For instance, "fucking hell" is a common expression where *fuck* is replicated, while "shitting hell" is less prevalent. This discrepancy in replication patterns between the two nouns underscores the variations in how function words are learned and applied within different linguistic contexts.

Chapter 7 The research questions

This chapter focuses on answering the research questions from Chapter 1. Each research question will be answered as a small sub-chapter. An introduction will be given, followed by a discussion about the findings presented in Chapter 5, and lastly, an answer to the research question will be given. The results presented in Chapter 5 will determine each of the research question, and the direction that their answer will go. First, the research questions will be presented and compared to Robbie Love and his paper "Swearing in informal spoken English: 1990s-2010's". Afterward, there will be a discussion where both the findings presented in Chapter 5 and Love's findings are compared. Lastly, a conclusion will be given summarizing all five research questions.

7.1 RQ 1: How do the findings of the present master's thesis compare with Robbie Love's findings?

This paper is inspired by "Swearing in Informal Spoken English: 1990s-2010's" by Robbie Love. Love begins his paper by explaining what swearing is, then he goes on to present the list of swear words that he uses for his paper. Love (2021) states that bad language is part of everyday life. This is one of the things that he has in common with Anderson and Trudgill (1999). But the questions remain, if bad language comes naturally, when is it appropriate to use it, what kind of patterns does bad language have, and what can be counted as bad language? Four different articles and books have either written about or brushed the subject of bad language, and all four agree. According to Cambridge Dictionary, bad language is defined as "words that are considered offensive by most people" (Cambridge Dictionary, s.v. "bad language"). Cambridge Dictionary also defines swear words as "a rude or offensive word" and swearing is defined as "rude or offensive language that someone uses, especially when they are angry" (Cambridge Dictionary, s.v. "swearing").

Love mentions in his paper the frequency of a set of swear words and which ones are the most frequent. Using the frequency accrued, he continues to investigate the age and the social standing of the people who use these swear words. According to Love's (2021) research, people tend to swear more in adolescence than in any other age group. The type of the swear word is also important for the fact that the more modern the swear word is, the less likely people around thirty and above would know and use it. The fewer people use these words the

less frequency they have in that certain age group (See Chapter 2 for more detail). There are some general patterns that Robbie Love has discovered in his paper (2021). First, he points out which swear word are the ones on his list that is most frequent and gives a chart that points the swear words in the order of their frequency. He discovered that the swear word, *fuck*, is the most frequently used swear word followed by the swear word, *shit*. Second, Love argues that there is a pattern between age and swearing. Love (2021) reasons that the environment and the media exposure that the people in the age of adolescents are in is one of the reasons why it is frequently used by them. Lastly, Love (2021) points out that there is a connection between swearing and society. In a working society, it is important to keep a professional standing with co-workers and clients, which makes swearing an inconvenience. Assuming the speaker is a teacher speaking to either another teacher, a principal, their boss, or their student, the most inappropriate form of communication to use would be the usage of swearing and swear words. This is one of the reasons why the frequency is low for swearing and swear words for people in the age where they are career-focused.

Love's paper "Swearing in Informal Spoken English: 1990s-2010's" (2021) was chosen because of the topic, which is swearing and swear words the main topic in this paper. This topic is one of the topics that caught my attention from the beginning when I was choosing what to focus on in linguistics. Although other authors and linguists written about swearing in informal language, I was mostly inspired by Robbie Love's (2021) paper. Robbie Love focuses on this topic and covers most of the period that I wanted to investigate. He argues in his paper about swearing in an easy-to-understand way. His paper is a short and not overwhelming focusing on frequency. The method that he uses is the same method that is used for this paper, which is using a corpus to answer the research questions. Love uses frequency as the only way to test his theory and to answer and argue about his research questions. This paper, on the other hand, includes two topics and tools to answer the research and to test the theory, which are frequency, and collocation. One of the reasons for this is that the two topics are closely connected, and they give a wider view of the topic. The requirements given for this paper demand a much bigger scale than the one for Love's (2021) paper. This might be because Love's paper is a small article requiring a few pages with a general overview of the topic swearing in the informal English language, while this research paper is required to be specific and more detailed for academic reasons.

There are many similarities between Robbie Love (2021) and my paper, which represents the topic. First, is the method used. Both papers extract the data by using the corpus method. Both papers also use a theory that is tested using the data extracted by the corpus. The next similarity would be the usage of frequency in both papers. Love's paper focuses on finding the frequency of his set of swear words (see Table 1 in Chapter 2) in context to age and social standing, while this paper focuses on the frequency of my own set of swear words (Table 1 chapter 2) focusing on the frequency from the period 1930's to the 2010's. Aside from the similarities between the two papers, there are also a few differences. Even though this paper is inspired by Love's paper, it is not a replica of his paper. This paper uses and compares the TV corpus with the Movies corpus. Love's paper uses the BNC corpus only. The BNC corpus focuses on only the British variety of English, and the BNC corpus focuses on all different genres. This means that Love's data is gathered from for example TV, movies, interviews, recordings etc. My data, on the other hand, focuses on only two categories, TV, and movies. On top of that, by using the TV corpus and the Movies corpus, this paper can include other varieties of spoken English and analyse them.

This paper has a quite few patterns that were discovered through the results gathered from the data. First, we will list all the patterns from the data related to the frequency topic. The patterns found in this topic present some similarities to the patterns found in Love's thesis. The swear words from Love's (2021) paper present the swear word showing that there is a pattern between the most frequently used swear words and the age that it is used in. But there are a lot of differences present as well. For example, the timeline for the frequency of the swear words is not presented in the charts and tables shown in Love's paper. The chart for frequency that Love presents in his paper shows the swear words from most frequent to less frequent and the frequency and all the swear words and their frequency are jammed in one chart.

Next, we will list all the patterns from the data gathered from the collocation topic. The patterns for this topic are many as there are 11 swear words presented that needed data to be collected from. First, there is the usage of intensifiers through reduplications. The repetition of the swear word, used as both the swear word and the collocate, makes both the swear word and the collocation much more noticeable in a sentence. A few examples of that are the swear words *shit* and *fuck*. Both swear words had their swear word as a collocate in their list, the swear word *shit* had the collocate SHIT following it, and the swear word *fuck* had the

collocate FUCK following it. This is one of the patterns that were discovered from the data for most of the swear words (See more in Chapter 5).

Another finding that was discovered using the data collected from the corpora is that swear words are mostly used in American and Canadian varieties of spoken English. The reasoning behind that is the fact that most of the data gathered from the corpora are from either American/Canadian movies and television or the UK/Irish movies and television. I am using normalised frequencies to compensate for this. There were almost no concordance hits in the collocation for the swear words in any variety of spoken English other than those four. When finding the collocation for the swear words, the focus was on lexical words rather than function words. There are, however, quite a few exceptions. The swear words, such as *twat*, are mostly popular in the British variety of spoken English.

7.2 - RQ 2: Which swear words are the most frequent, in each corpus, over time and in which variety of spoken English?

The topic of frequency is the next on the list in this discussion. First, I will give a summary of what frequency is. Frequency allows us to investigate the number of times a specific search word or lemma is used in a corpus at a specific time or over either a long or short period, or even both (See Chapter 3 and Chapter 5 for more details). Frequency is a tool that is present in the corpus. The settings that can be controlled in the corpus used to find the frequency, may be adjusted to the necessary settings appropriate for the topic investigated at hand. For example, if there is a specific period that we wish to investigate, it is possible to enter the data in the settings so that only that specific period is shown. The same may be done for the investigation of frequency of other areas as well, such as a specific variety of spoken English, or even for a specific show or genre.

Looking at the charts mentioned in Chapter 5, the two corpora, the Movies Corpus, and the TV corpus display similar and different results. What these charts show is the frequency of the list of swear words from Table 2 from the period 1930s to the 2010s. The frequency of all 11 swear words was shown in the different charts, in both the TV corpus and the Movies corpus. Two different main patterns were discovered through the data in the investigation. First, it was clear from the charts that the Movies corpus presented more frequency for the

swear words than the TV corpus did. There are many reasons as to why the results showed this.

Next, I will be comparing my findings with the ones that Robbie Love (2021) found in his paper "Swearing in informal spoken English". According to the data found, for the majority of the swear words, there are almost no hits found in either the TV corpus or the Movies corpus for the period from the 1930s up until the 1950s and 1960s. For the swear word *fuck*, for example, the data from the TV corpus state that the data starts from the 1950s and there are only 3 hits found. This is a complete contrast to the Movie corpus where the data starts from the 1930's, and there are 43 hits found in this corpus. This is common for almost all swear words that are investigated in this thesis, as the data showing the frequency of the swear words, in the TV corpus, do not start until the 1950s, 1960's or the 1970s. Another founding is that there are many more frequency hits for the swear words in the Movies corpus than in the TV corpus. We can go back to the swear word *fuck* as an example of that, where the frequency hits in the TV corpus for the period 1960s is 1 hit while the results for the same period in the Movies corpus is 139 hits.

Love (2021) found that the three swear words, *fuck, shit,* and *bloody* are the most frequent swear words among the sixteen swear words in his research. The two swear words *shit* and *twat* show a significant increase over 10-year period, while the swear word *fuck* remained stable as the most frequently used swear word out of the set of sixteen swear words presented in Table 1 (Chapter 2). Love's theory is that the reason why *fuck* is so frequently used is because it has replaced *bloody*. He argues that despite the swear word *fuck* being one of the strongest bad language words of all, it is still the most frequent and popular swear word used among British teenagers in the ten-year gap between 1990s and the 2010s. Another discovery that Love (2021) made was that, according to the data he found, males tend to swear more than females even though the theory says that males and females have the same possibility to swear equally.

Many studies have reported and made discoveries similar and different to Love's (2021) paper. What is distinct between Love's (2021) paper and previous studies, such as Jay (2008), is that Love's research is much narrower where his entire research is based on a restricted set of words containing only 16 swear words, while other studies, such as Jay (2008). This is one of the things that this thesis and Love's (2021) paper share. Both use a restricted set of swear

words in the research to try and answer the research questions. But even so, this thesis has the smallest restricted set of swear words among all the other studies mentioned due to the limitation of this thesis. Comparing these findings with Robbie Love's 2021 research and Jay's 2008 study, several parallels and distinctions emerge. Love's research identifies *fuck, shit,* and *bloody* as the most frequent swear words among a set of sixteen. He suggests that *fuck* may have replaced *bloody* in usage. While these findings align with some aspects of the present study, there are notable differences.

The difference in swearing between movies and TV shows can be attributed to several factors, including censorship rules, time slot considerations, and the difference in content creation processes:

1. Censorship Rules: Movies and TV shows often face different censorship regulations. Movies typically go through a more rigorous rating and classification process before release, with different rating categories indicating the level of explicit content. Filmmakers may have more creative freedom in terms of language usage, as movies can be targeted at specific age groups or intended for mature audiences. However, certain countries or regions may have stricter censorship rules for movies as well.

2. Time Slot Considerations: TV shows are generally broadcasted during specific time slots, with designated hours for programming aimed at different audiences. There are regulations in place to ensure that content appropriate for children is aired during times when they are less likely to be watching, such as the "9 pm rule". This time slot restriction may limit the use of explicit language in TV shows to comply with family-friendly programming guidelines.

3. Scripted vs. Free Speech: TV shows often follow scripted narratives and storylines, which provide writers and producers with more control over the language used by the characters. The script undergoes editing and review processes to ensure adherence to network or regulatory guidelines. On the other hand, movies, particularly those targeted at mature audiences, may aim for more realism or creative expression, which can include the use of explicit language.

It's important to note that these factors can vary across countries and regions, as censorship rules and content guidelines differ. Additionally, with the rise of streaming platforms and

premium cable channels, there may be more flexibility in TV shows to include explicit language or mature content due to different subscription-based or on-demand viewing models. Overall, the difference in swearing between movies and TV can be attributed to a combination of censorship rules, time slot considerations, and the varying nature of content creation processes.

7.3 - RQ 3: Which words collocate with the swear words and on that basis what semantic prosody do the swear words have?

First, we will start by giving a summary of what collocation is and what it shows. Collocation is also the habitual juxtaposition of a particular word with another word or words with a frequency greater than chance. Using easier terms, in corpus linguistics, a collocation is a series of words or terms that co-occur more often than would be expected by chance. In the case of this thesis, the collocations investigated are for the set of swear words listed in Table 1, and presented in Chapter 2. This means that the words that are most frequently used with the swear word are called collocates for the specific swear word. For the sake of this thesis, there are 11 swear words investigated and searched the collocations for. In this investigation, there 4 main patterns that the swear words had in common. To make it clear, when the swear words are mentioned as search words, lemmas, or just swear words, they will be in *italics*, while collocates will be mentioned in CAPITAL letters.

Next, I will summarize the main patterns that were discovered through the analysis of the data. First, is the usage of intensifier through reduplication. In the collocations list, the usage of reduplication was present in most of the collocations Tables in Chapter 5. What is meant by reduplication is the usage of the swear word as both a swear word and a collocate at the same time and in the same sentence. For example, for the swear word, *shit*, the collocate SHIT was used in the same contexts and sentence, as in SHIT, SHIT, SHIT. The second pattern is that near-synonyms have similar collocates. The swear words investigated have similar meanings to each other, for example, swear words like *shit* and *crap*. Common collocates between them are almost identical, where the differences are only two sets of collocates and the order of the ranking that the collocates are in (for more detail, see Chapter 5).

The third pattern that emerged relates to the distinction between investigating collocates that are function words and those that are lexical words. It became evident that collocates containing function words primarily serve to identify the word class of the swear words and are thus less conceptually rich. The function collocates primarily indicates that the swear words in question function as nouns. Conversely, collocates consisting of lexical words, as presented in Chapter 5, contribute significantly more meaning to the results. These lexical collocates provide a deeper context and interpretation of the swear words in the corpus. The final pattern discovered through the analysis of collocations is the diverse word classes represented by the swear words. The investigation revealed that swear words can function as various parts of speech, including adjectives, nouns, intensifiers, and more. However, the predominant usage pointed towards their role as intensifiers, indicating their frequent employment to add emphasis to expressions or statements.

Next, I will compare my findings for two of the swear words with the findings of Lutzky and Kehoe in their article "Your blog is (the) shit: A corpus linguistic approach to the identification of swearing in computer-mediated communication"(Lutzky & Kehoe, 2016). A common ground between the two papers is some of the swear words used for the investigation of collocation for both studies, such as *crap* and *shit*. One of the differences between my thesis and Lutzky and Kehoe's (2016) study is the set of swear words used and the number of swear words used. Figure 1 that Lutzky and Kehoe (2016) present in their paper shows the shared collocates of all swear words investigated in their study. Two of the swear words, *shit* and *crap* show on the figure that they only partially have the same collocates. The two swear words are near-synonyms and this shows the likeness of one of the patterns discovered in my thesis; Swear words that are near-synonyms are likely to have similar collocates.

Many different collocates were presented on the lists for each Table in chapter 5. The collocates that co-occur the most for almost all Tables presenting the collocation of the swear words in chapter 5 are the collocates FUCK-3, MAN-7, FAT-4, LITTLE-4, RIGHT-5, and KNOW-3. These six collocates are present in three or more Tables, which makes them the most frequently used collocates for my set of swear words investigated in this thesis. The numbers to the left of the collocates show the number of the Tables that specific collocate has appeared in.

Another significant aspect that surfaced during the examination of collocation patterns is the concept of semantic prosody, which applies to both the collocates and the swear words employed. When swear words tend to collocate predominantly with negative terms, it imparts a negative semantic prosody. In cases where there are only a few instances of positive collocates linked to swear words, it may suggest a sense of luck or fortuity associated with the swearing. Conversely, when swear words exhibit numerous neutral collocates, such as "son" or verbs like "give" and "come," it indicates a relatively neutral semantic prosody, signifying a versatile usage context. Understanding the semantic prosody of swear words and their collocates offers valuable insights into the nuanced connotations and associations surrounding these profanities within language.

7.4 - RQ4: Using Grice's maxims of Conversation, does using swear words violate or break any of the maxims? How?

Language is the most common tool used for communication. Naturally, most people tend to ignore language as it is part of everyday life. But language is important. What we say, how we say it, the specific words that we choose to use in a specific situation, everything is relevant when it comes to language. This paper focuses on the spoken part of language, therefore, what we choose to say is important. But what qualifies if the language is good or bad? Anderson and Trudgill (1999, p.3) argue that language can identify the characteristics of a person, whether they are good or bad, educated or not, caring, or arrogant, etc. Anderson and Trudgill (1999, p.13).

The attitude and the usage of language and words are where Grice's maxims of Conversation come in. Grice's maxims have a set of conditions that need to be followed for them to be fulfilled. The point of these maxims is not to show that this is how everyone is supposed to conduct language, but rather to show that language may not always follow these maxims and their conditions. Grice's maxims open up an opportunity for us to investigate for example how and why language tends to break these maxims and to understand that language may not always follow these theories. Bad language is one of these instances where Grice's maxims and their conditions may be broken every time (Kroeger, 2018, p. 142). But what exactly is considered bad language? Allen and Burridge (2007) argue that swearing and the usage of profanity is considered bad language. If the speaker intends to follow Grice's maxims of

conversation without breaking them, then the usage of swear words and swearing in general may not be the best option for the speaker. Swear words have meanings behind them, which makes the speaker break Grice's maxims and their conditions if they choose to use swearing as a form of conversation.

To know the speaker's true meaning, we must first know about the speaker's intent. We can separate the meaning and intentions of the speaker into two different categories for this. One is called the Pragmatic meaning, while the other is the propositional meaning. Before we can check which of the two meanings is used and when, we will first need to know what each of the meanings entails. First, is the pragmatic meaning. The Pragmatic meaning is also known as the implied meaning. The speaker in this type of meaning tends to say or speak one thing with the intention for the speaker to understand what is implied rather than speaking straightforwardly. The following example shows one way to use the pragmatic meaning.

1. Example of Implied meaning:

A: Do you know where the metro station is?

B: I'm going there now.

In the example above, A is asking B for directions, but only asks a YES or NO question. A is inquiring from B about their knowledge. A's question is implied. But so is B's answer. B only answered that he/she was going to the place that A mentioned, which implied that YES B knew where the metro station was. Both A and B used the pragmatic meaning in their communication. Pragmatic meaning may also be called for figurative meaning. With this type of meaning, we are adding additional information, which adds a meaning layer more than the literal.

The second type of meaning to explain is the propositional meaning. The propositional meaning, on the other hand, is more in the literal sense of the meaning. This type of meaning requires the hearer to understand the sentence as it is said. The speaker, in this case, is required to not have any hidden meaning when speaking and needs to understand that the hearer will take whatever the speaker says in the literal sense. The following example shows one way to make use of the propositional meaning.

2. Example of the Literal meaning:

A: You're a poor kid.

The example shows the propositional meaning in action. This example allows the speaker to point out the financial status of the hearer. In other words, the word "poor" in the example above may be taken in the literal sense if the financial status was the aim of the conversation. At the same, the word lexical "poor" in this case may have a different meaning to it. The other meaning that this example provides is the adjectives "unfortunate, unlucky, ill-fated, pitiful". The same can be said if the speaker said, "This is shit". Although the word "shit" is impolite, it still may have two different meanings, which leaves room for the hearer to interpret what the speaker said. This makes example 2 not a good example of the usages of the propositional meaning, or as it is commonly known, the literal meaning. If the speaker wishes to make use of this meaning without there being a chance for interpretation, the speaker will have to leave no room for other meanings to be interpreted. Example 3 presents this and would, therefore, be a better option.

3. Example
A: Where is this?
B: This is a school.

Speaker B in example 3 leaves no room for person A to interpret "this" as anywhere else other than "a school". Both the Pragmatic meaning and the Propositional meaning are seen to be used in the examples provided by Table 2-12 in chapter 5. The two meanings are mentioned because of their involvement and effect on the research question. The usages of either one of the two meanings, both the pragmatic meaning and the propositional meaning, have some effect on whether Grice's maxims of Conversations are being followed or broken.

7.4.1 Grice's Maxims of Conversation

Before answering Research Question 4, I will first give a small recap as to what the principle of the Maxims of Conversations are. The principle is also known as Grice's maxims, as he is the one who discovered these patterns in conversations. This principle consists of four main directions/instructions also known as maxim. These maxims present different conditions that need to be followed so that the maxims need to be fulfilled. The Four maxims of conversations are called the maxims of QUALITY, QUANTITY, RELATIONS (or

RELEVANCE), and MANNER. Kroeger (2018:141-142) gives an overview of all four maxims and their conditions.

The examples and the data provided in Chapter 5, however, show that the two most relevant maxims in this thesis are the Maxim of QUALITY and the Maxim of QUANTITY. The patterns discovered in the investigation suggest that these two maxims were broken rather often, more often than the other two, and therefore, are the most relevant when answering the research question. First, I will start with the maxim of QUALITY. As it is explained above, the maxim of QUALITY is about being truthful. Anything that is not the truth means that this maxim is broken. On top of that, this maxim has two requirements that need to be fulfilled. If even only one of the requirements of the maxim is not met, then the maxim is considered to be broken. In the data gathered for collocation shown in Chapter 5, there are many instances where the maxim of QUALITY is broken. In Chapter 5, there are many examples given out for each swear word and many of those examples, if not all, break this maxim.

Next, is the maxim of QUANTITY, which is about giving out only the required information. This maxim, like the maxim of QUALITY, has two requirements, which are for the speaker to give out only what is required, and to not give out more than is required. A good example of this can be seen in Chapter 5 Table 2 and Table 3. Table 2 gives out the collocation list for words that collocate with the swear word, *fuck*, while Table 3 presents the collocation list of words collocating with the swear word, *shit*. In each of the two Tables, we can see that the swear words are repeated, for example, with the list for the swear word *fuck*, the collocate is also FUCK, and the swear word and the collocate are repeated multiple times, creating the reduplication effect. This can be seen in the example 4 under A.

4. Example:
A: Fuck, fuck, fuck, fuck!
B: Shit shit shit shit shit!

The same happens in Table 3 with the swear word, *shit*, where the collocate is also SHIT. The collocation and the swear word stand together creating a reduplication effect as it can be seen in the example above in B. What can be concluded is that the usage of reduplication is counted as intensifiers. By repeating the same word over again, the emphasis is put on the word repeated. This leads to the question of whether this has to do with habits or not. It is

extremely usual in language to use intensifiers, not just swear words, but it is also typical to use non-swear words as intensifiers, typically adverbs. The non-swear word 'very' is one of them. An example of that is when the speaker is asked how they feel, and their answer is "**very very** sad". This does not only happen in English but in other languages as well, such as Norwegian. These habits to reduplicate by using intensifier, do they mean to intensify? The answer to that is debatable. Sometimes, the reduplication is done on purpose to intensify, to emphasize the point, like in the case of the non-swear word 'very', while other times it is done out of habit, like in the examples for the swear words, *fuck,* and *shit*. Reduplication breaks the maxims of quantity. The speaker is breaking both requirements for this maxim as they are making their contribution more than what is required.

7.4.2 Politeness and impoliteness

Politeness, a fundamental aspect of human communication and social interactions, is thoroughly examined by scholars like Penelope Brown and Stephen Levinson in their seminal work "Politeness: Some Universals in Language Usage" (1987). It encompasses a wide array of behavioural aspects, ranging from the use of polite language and expressions to the adherence to cultural norms and societal expectations. The practice of politeness not only signifies respect and consideration for others but also plays a role in nurturing positive relationships while mitigating the potential for misunderstandings or conflicts (Holmes, 1988, pp. 535-555).

Nonetheless, when we go into the realm of swear words and the broader area of swearing, the application of Grice's maxims, as shown by Paul Grice, takes on a character. Swear words, often characterized by their emotional intensity, can be construed as Face Threatening Acts (FTAs) (Saeed, 2016). These linguistic expressions possess the potential to pose threats to both the negative and positive faces of the listener, making them inherently prone to inducing impoliteness, particularly in formal or polite contexts where such language is deemed inappropriate (Brown & Levinson, 1987). However, it is imperative to acknowledge the dynamic nature of the perception of politeness and impoliteness associated with swear words, which can exhibit significant variations contingent on social and cultural factors, as well as the nuanced nature of the relationship between the speaker and the listener (Holmes, 1988, pp.535-555).

Therefore, the analysis presented in this thesis transcends the conventional binary perspective of swear words as mere violations of Grice's maxims. Instead, it delves deeply into the intricate interplay of these words within the context of politeness and impoliteness. This perspective recognizes that the usage of swear words often aligns with specific social and contextual norms, thereby challenging traditional notions of politeness. Swear words, in their role as linguistic tools, serve multifaceted functions within language, communication, and society. This comprehensive approach prompts a more profound exploration of their impact, underscoring their significance as a complex and intriguing area of linguistic study (Brown & Levinson, 1987; Holmes,1988, pp.535-555).

Chapter 8 Conclusion

The research questions that formed the groundwork for this thesis are as follows:

1. How do the findings of the present master's thesis compare with Robbie Love's findings?

2. Which swear words are the most frequent, in each corpus, over time, and in which variety of spoken English?

3. Which words collocate with the swear words and on that basis what semantic prosody do the swear words have?

4. Using Grice's maxims of Conversation, does using swear words violate or break any of the maxims? How?

To answer the first question, the findings of this master's thesis both align with and extend beyond Robbie Love's research in 2021. Like Love's (2021) work, this thesis explores swear word usage and frequency. However, it diverges by analysing different corpora (TV and Movies) and focusing on diverse varieties of spoken English. Moreover, this thesis delves into the semantic prosody and examines the application of Grice's maxims of conversation, offering a more comprehensive understanding of the multifaceted world of swear words. Additionally, Jay's (2008) research has been instrumental in providing a broader context for understanding the cultural and sociolinguistic aspects of swearing, complementing, and enriching the comparative analysis with Robbie Love's (2021) findings.

The second question is answered by finding the frequency of the 11 chosen swear words in the corpora and analysing the data accordingly. The findings offer insights into the six varieties of spoken English, revealing results that differ from the initial assumptions. This study delves into swear word frequency, shedding light on language usage over time and across various English varieties. It's noteworthy that movies generally exhibit higher swear word frequencies compared to TV shows from the 1930s to the 2010s. When comparing these findings with Robbie Love's (2021) and Jay's (2008) research, both similarities and differences emerge. While Love's (2021) study includes a broader set of sixteen swear words, this thesis focuses on a smaller restricted set due to limitations. Additionally, this research explores variations across English varieties, encompassing American, Canadian, UK, and Irish English. In summary, it provides a nuanced understanding of swear word usage in diverse linguistic contexts. In terms of frequency, movies held the lead until around 1960-

1970 when TV slowly began to catch up. Figure 2 indicates that the AU/NZ variety has the highest frequency, making it the leader in this regard.

The third question is answered by finding the collocation of each of the swear words and analysing the data. this study explores collocations of 11 selected swear words, revealing insights into their linguistic usage. Four key patterns emerge: Intensifier Reduplication: Swear words are often repeated within a sentence for emphasis. Similar Collocates for Near-Synonyms: Swear words with similar meanings share common collocates. Function Words vs. Lexical Words: Collocates with function words indicate word class, while lexical collocates offer richer context. And lastly, Diverse Word Classes: Swear words can function as various parts of speech but are primarily used as intensifiers. Comparing Lutzky and Kehoe's (2016) research reveals similarities and differences in collocations. Additionally, the analysis uncovers the concept of semantic prosody, which includes negatively loaded collocates, rare instances of positively loaded collocates like "LUCKY," and numerous neutral collocates such as SON or verbs like "give" and "come." This suggests a relatively neutral semantic prosody associated with these swear words in certain linguistic contexts.

To address the fourth and final question, I apply Grice's Maxims of Conversation (Kroeger, 2018, p. 142) to assess whether the conditions of the Maxims are adhered to or breached when swear words are used. Upon this examination, it becomes evident that while the usage of swear words technically violates the Maxim of Quality and the Maxim of Quantity, the issue of swearing and the use of swear words is more closely related to matters of politeness and impoliteness. Here, the Face Threatening Act (FTA) plays a pivotal role, with the speaker potentially endangering the listener's face during swearing episodes.

8.1 Challenges

The analysis presented in this thesis faces certain challenges, one of the most prominent being the scripted nature of the data derived from the corpora. This challenge stems from the fact that the language found in movies and TV shows is meticulously crafted by screenwriters, directors, and actors, with the primary aim of serving the narrative and character development. This scripted quality means that the language used in media is less spontaneous and less reflective of natural, everyday spoken language. It is bound by the constraints of the script, often adhering to predefined dialogues and narratives. This controlled environment can

limit the accuracy and authenticity of the data, as it may not fully capture the nuances and variability inherent in spontaneous spoken discourse. While the scripted nature of the data presents a challenge, it is essential to approach the findings with an awareness of this limitation. Researchers must recognize that the language analysed here is a product of creative storytelling rather than unfiltered, real-life communication. Despite this challenge, the analysis still offers valuable insights into the usage of explicit language in media, shedding light on its role in language, communication, and society within the constraints of scripted content.

8.2 Further studies

Within the scope of this thesis, several avenues for further investigation and research have emerged, offering a myriad of opportunities for in-depth exploration:

Focus on one TV show or movie and analyse from there. concentrating on a single TV show or movie offers a promising avenue for future research, enabling a more profound exploration of the complexities surrounding the use of explicit language in scripted content and its implications for storytelling, character development, and audience reception.

Impact on Language Patterns: The usage of swear words and explicit language can influence language patterns and expressions. People may incorporate these words into their vocabulary, leading to changes in speech patterns and communication styles. Emotional Expression: Swear words often carry strong emotional connotations and are used to express frustration, anger, or emphasis. The prevalence of such language in movies and TV shows may impact the ways people express their emotions and the language they choose to use in emotional situations. Cultural and Generational Differences: The acceptance and usage of explicit language can vary across cultures and generations. Movies and TV shows can reflect and contribute to these cultural and generational differences in the way language is used and perceived. Censorship and Regulation: The prevalence of explicit language in movies and TV shows can also prompt regulatory bodies to implement stricter censorship guidelines or regulations. This can impact the creative freedom of content creators and shape the language used in media. Linguistic Creativity and Expressiveness: On the other hand, the usage of explicit language can also be seen as a form of linguistic creativity or a way to express strong emotions or attitudes. In certain contexts, it may serve as a tool for emphasis or as a means of conveying specific tones or messages.
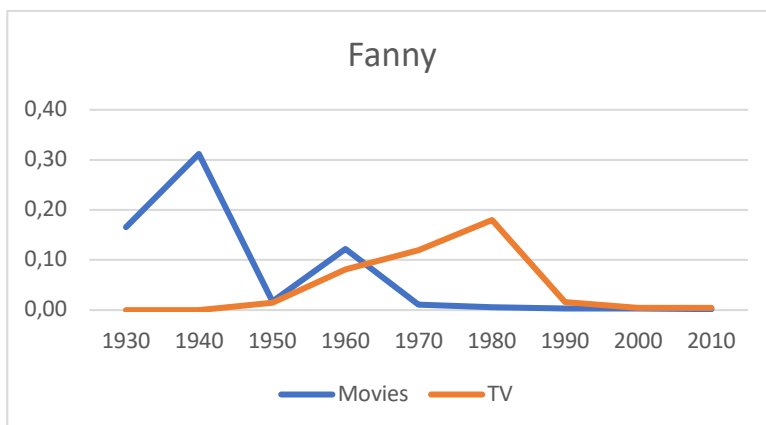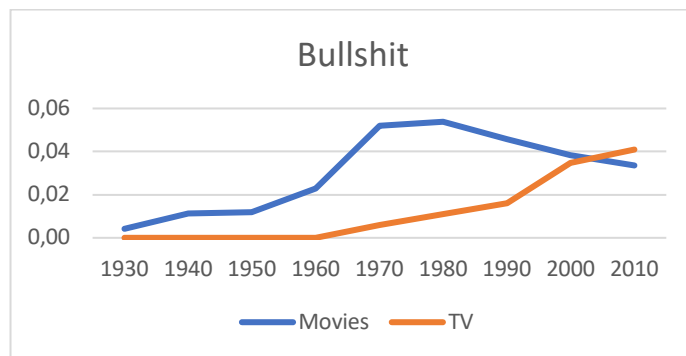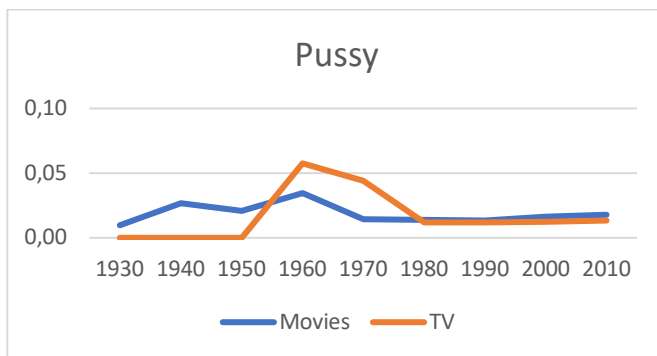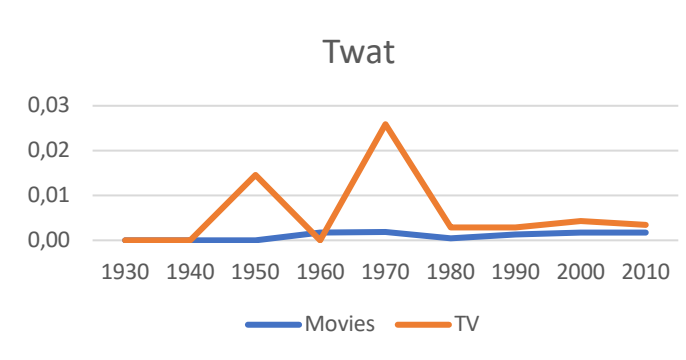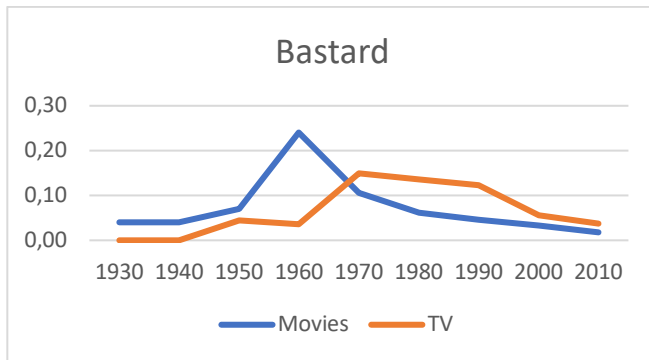
Appendix I

Table 12: *Frequency charts of the swear words TV corpus vs Movies Corpus*
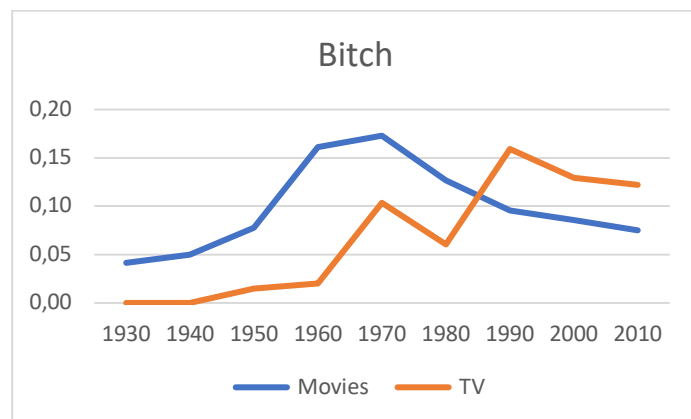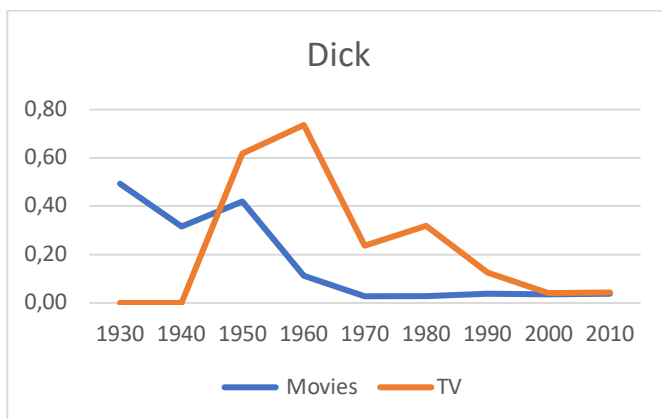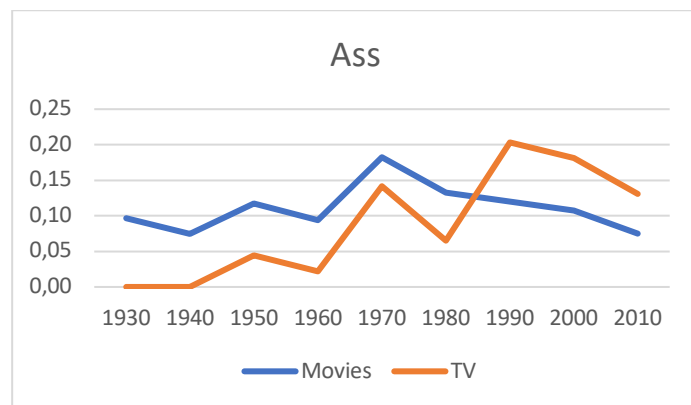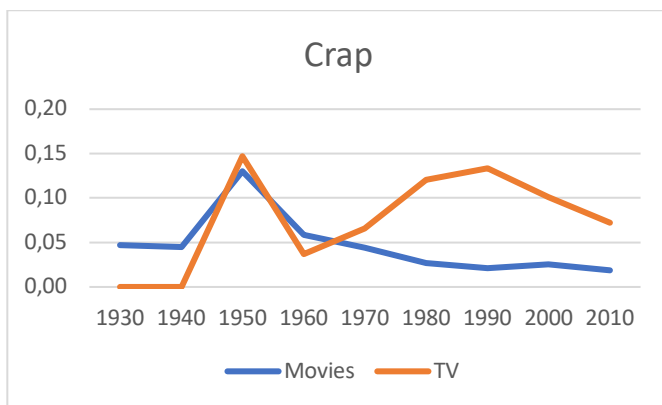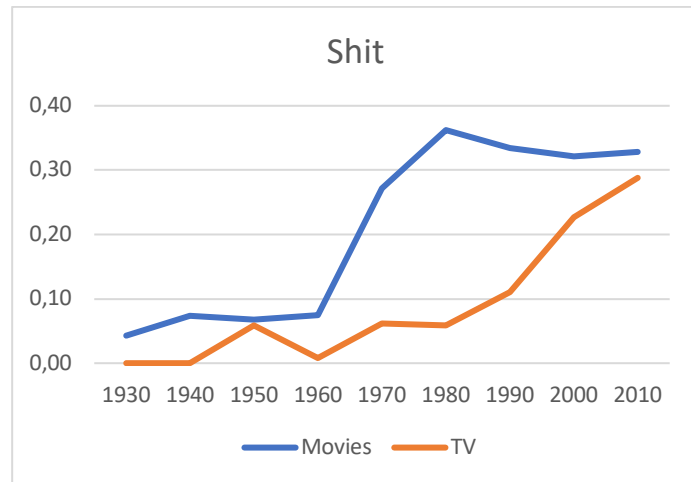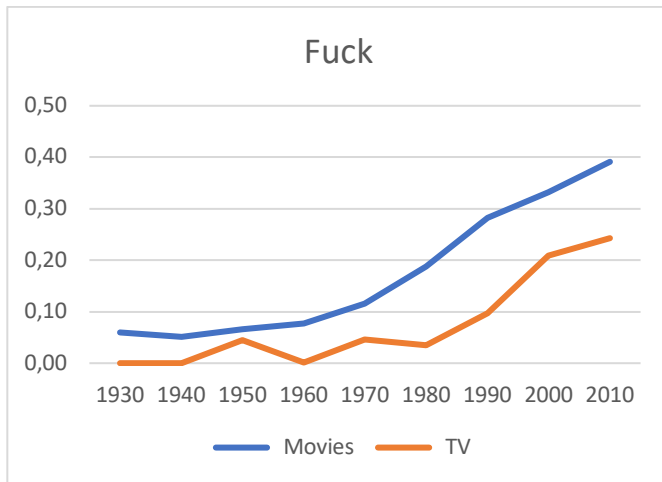
Appendix II

Table 14: *Frequency charts of the swear words TV corpus vs Movies Corpus*

Bastard

Twat

Pussy

Bullshit

Fanny

Reference list

Adams, M. (2016). *In praise of profanity*. Oxford University Press.

Allen, K., & Burridge, K. (2007). *Forbidden Words*: *Taboo and the Censoring of Language*. Cambridge University Press.

Andersson, L. G., & Trudgill, P. (1990). *Bad language*. Blackwell by arrangement with Penguin Books.

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

Brown, P., & Levinson, S. C. (2016). *Politeness: Some universals in language usage* (25th printing 2016). Cambridge University Press.

Byrne, E. (2017). *Swear!ng is g*od f*r you: The amaz!ng sc!ence of bad language*. Profile Books.

Cambridge Dictionary. (n.d.). Citation. In *Cambridge Dictionary.* Retrieved December 15, 2022, from https://dictionary.cambridge.org/

Cameron, D. (1995). *Verbal hygiene*. Routledge.

Cole, P., & Morgan, J. L. (Eds.). (1975). *Speech Acts*. BRILL.

Coulson, S., & Oakley, T. (2005). Blending and coded meaning: Literal and figurative meaning in cognitive semantics. *Journal of Pragmatics*, *37*(10), 1510–1536. https://doi.org/10.1016/j.pragma.2004.09.010

Fairclough, N. (1989). *Language and power*. Longman.

Grice, H. P. (1975). Logic and Conversation. In P. Cole & J. L. Morgan (Eds.), *Speech Acts* (pp. 41–58). BRILL. https://doi.org/10.1163/9789004368811_003

Holmes, J. (1993). New Zealand women are good to talk to: An analysis of politeness strategies in interaction. *Journal of pragmatics*, *20*(2), 91-116.

Jay, T., & Janschewitz, K. (2008). The pragmatics of swearing. *Journal of Politeness Research. Language, Behaviour, Culture*, *4*(2). https://doi.org/10.1515/JPLR.2008.013

Kroeger, P. (2018). *Analyzing meaning: An introduction to semantics and pragmatics*. Language Science Press.

Love, R. (2021). Swearing in informal spoken English: 1990s–2010s. *Text & Talk*, *41*(5–6), 739–762. https://doi.org/10.1515/text-2020-0051

Lutzky, U., & Kehoe, A. (2016). *Your blog is (the) shit*: A corpus linguistic approach to the identification of swearing in computer mediated communication. *International Journal of*

Corpus Linguistics, *21*(2), 165–191. https://doi.org/10.1075/ijcl.21.2.02lut

McEnery, T. (2006). *Swearing in English: Bad language, purity and power from 1586 to the present* (1. publ). Routledge.

Oxford Learners Dictionary. (n.d.). Citation. In *Oxford Learners Dictionary.* Retrieved

December 15, 2022, from https://www.oxfordlearnersdictionaries.com

Pinker, S. (2007). *The stuff of thought: Language as a window into human nature.*

Saeed, J. I. (2016). *Semantics* (Fourth edition). Wiley Blackwell.

Whelehan, A. 2023 May. Fanny. The Bump. https://www.thebump.com/b/fanny-baby-name

ISO 690