INSTITUTIONEN FÖR
SPRÅK OCH LITTERATURER

DO YOU HAVE THE TIME TO LISTEN TO ME WHINE?

First person pronoun use in English pop lyrics

Markus Berg

Title: Do you have the time to listen to me whine? First person pronoun use in English pop lyrics

Author: Markus Berg
Supervisor: Monika Mondor

Abstract

The purpose of the present study was to determine whether a pattern in first person pronoun use can be detected in the lyrics of different music genres. For this purpose, eighty songs were randomly selected by *Spotify*'s "Radio" feature, twenty from each of four different genres: hip hop, pop, punk, and club/house. The statistical data on the use of first person pronouns in these songs appears to indicate that punk lyrics use significantly fewer first person singular and plural pronouns than the other three genres. Club/house lyrics, on the other hand, tend to use slightly more first person singular pronouns than pop, both of these genres using slightly more than the combined average of the four genres. Club/house lyrics also have been found to use more first person plural pronouns than the other genres. The data gathered for the present study appears to conform to patterns observed by previous research, namely that angry people use fewer first person pronouns (Pennebaker 2011) and that the word *we* can be used to decrease social distance (Semin 2007).

Keywords: sociolinguistics, English-language music, first person pronoun frequency, quantitative study, statistical analysis, pop, hip hop, punk, club/house

Table of contents

1. Introduction

Sociolinguistics studies how people's social contexts affect the way they speak. The words someone uses and the way they pronounce those words can give a listener clues to where the speaker is from, and what kind of social background the speaker has (Chambers 2003: 4, 6-7). Chambers' book *Sociolinguistic Theory* (2003), for instance, provides examples of research on the linguistic variation in a multitude of languages. The book focuses primarily on pronunciation and dialectal word choices, but it does not discuss the use of function words, such as articles or pronouns. This is because rather few studies have been carried out on the subject of function word use, as Pennebaker points out (2011: 23).

Some studies have, however, been carried out on the topic of pronoun frequencies in creative writing. For instance, Twenge, Campbell & Gentile (2012) investigated the change of first and second personal pronoun frequencies in American literature over time. Similarly, Pennebaker, Sivertsen & Petrie (2008) examined the change in first person singular pronoun use in the lyrics of The Beatles over the course of the band's career. To my knowledge, however, no one has studied whether there are any significant differences between genres of music in terms of pronoun frequencies. According to Pennebaker (2011), the use of pronouns can tell an observer several things about the writer, such as social background, mood, and mental state. That being the case, different genres of music should exhibit differences in language use, if they are indeed written by people with differing intentions and different social backgrounds. So, the present study investigates first person pronoun frequencies in lyrics, in an attempt to determine whether there are any patterns differentiating music genres in terms of pronoun use.

1.1 Aim

The purpose of the present study is to investigate whether the findings of previous research, specifically that of Pennebaker, Mehl, and Niederhoffer (2003), Pennebaker, Sivertsen, and Petrie (2008), Semin (2007), and Gustafsson Sendén (2013) can be applied to lyrics. If music genres are thematically different in their lyrics, this should be reflected in the pronoun frequencies of the songs belonging to those genres. For instance, as punk music is by its very definition part of counter-culture, it should contain lyrics that express resentment and/or frustration with the way the world is. That being the case, my hypothesis is that punk music should exhibit a lower frequency of first person singular pronouns (FPSPs), as Pennebaker argues that angry people focus their speech on the object of their anger and utilize fewer self-

references (Pennebaker 2011: 107). In a similar vein, we should expect to see club/house music to use more instances of first person plural pronouns (FPPPs) in an attempt to shrink social distance (Semin 2007: 361). The study aims at verifying these studies against a selection of 80 songs from four genres of music, namely hip hop, pop, punk, and club/house.

1.2 Definitions
The following section contains brief definitions of the pronouns counted as well as the genres examined in the present study.

1.2.1 First person pronouns
A first person pronoun is a word used instead of a proper noun, used when referencing the self, either exclusively or as part of a group. They are categorized into three distinct types; the personal pronouns (e.g. *I, me*), the possessive pronouns (e.g. *my, mine*), and the reflexive pronouns (e.g. *myself*). These five examples are the first person singular pronouns, meaning the referent for each of these pronouns is exclusively the self. In order to include other people when referring to the self, the first person plural pronouns are used, those are: *we, us, our, ours*, and *ourselves* (Greenbaum & Nelson 2009: 46-48).

   To simplify long expressions commonly used in the present study, the following abbreviations are used:

FPSP   First Person Singular Pronoun, i.e. *I, me, my, mine*, and *myself*.
FPPP   First Person Plural Pronoun, i.e. *we, us, our, ours*, and *ourselves*.

1.2.2 Genres
The four genres under study in the present thesis are sub-genres of popular music. Popular music is somewhat difficult to define, as Tagg points out. He argues that popular music (unlike classical) is "conceived for mass distribution, [...] stored and distributed in non-written form, [...] only possible in an industrial monetary economy where it becomes a commodity and [...] should ideally sell as much as possible" (Tagg 1982: 41). In these broad terms, the phrase "pop music" can "[include] all contemporary popular forms", as Frith argues (2001: 94). However, Frith also argues that this broad definition is problematic and that pop should be considered a genre separate from, for instance, rock, hip hop, and country (Frith 2001: 94-95). For the purposes of the present study, the word *pop* and the phrase *pop music* are used to refer to this narrower definition, (further expanded on in section 1.2.4) which separates pop from, for instance, hip hop or punk. Lena and Peterson refer to this narrower definition as

"'pure' pop" (Lena & Peterson 2008, 700). On the other hand, *popular music* is used in the present thesis to refer to the broader definition described by Tagg (1982: 41). Similarly, the word *genre* is used to mean *sub-genre of popular music,* unless otherwise defined.

A genre of music refers to a collection of songs that have some characteristics in common, in terms of lyrics, composition, and the "cultural web of production, circulation and signification" (Holt 2007: 2) which they do not share with songs of other genres. As an example of this differentiation, metal music typically uses a very aggressive sound with deep, hammering drums, distorted guitars, and screamed or growled vocals, while the pop genre typically focuses on inoffensive lyrics and melodies intended to be catchy and memorable, using clean singing, as opposed to growling or screaming. It should be noted that a song can exhibit characteristics of several genres simultaneously, and some songs can be difficult to place in a genre at all. Rings refers to this as the "tension between genre [...] and the flickering, contingent identity of a given *single* [sic] song" (Rings 2013: 20).

1.2.3 Hip hop

Hip hop as a musical genre is somewhat strange, as it is used in the context of music genres as a synonym of rap music. The label hip hop actually refers to an entire culture and as such includes fashion, graffiti art, break dancing, and rap music (Morgan 2009: 51). Hip hop music as such, is not necessarily limited to rap music, but for the purposes of this essay, the two are used interchangeably.

Hip hop is characterized by one or several artists rapping over a musical track, which is often pre-recorded. Sometimes the background track is simply a drum beat from a drum machine, sometimes it is something closer to a complete song. Hip hop artists also often use so-called samples, in other words a melodic hook (or an entire melody) borrowed from another song (Boone 2013: 2). In recent years, hip hop music has, according to Rose (2008: 2-5), become reliant on sexist and homophobic lyrics, making it the target of criticism in the media, especially in the U.S.

1.2.4 Pop

Pop music is arguably the broadest genre as it encompasses a wide variety of music. It contains elements from several other genres, which makes it difficult to define, but similarly to other genres, pop music has its own characteristics that sets it apart from other genres. Lena and Peterson argue that "at its core, pop music is music found in *Billboard* magazine's Hot 100 singles chart", thus they argue that pop is a "chart [...] not a genre" (Lena and Peterson

2008: 699-700). However, Frith (2001: 94-95) argues that pop is its own genre, separate from, for instance, rock or country. He further argues that "[pop] is about giving people what they already know they want rather than pushing up against technological constraints of aesthetic conventions" (Frith, 2001: 96). Thus, pop music is typically inoffensive and lyrics tend to revolve around relatable topics, such as romantic relationships. Osborn argues that "Top-40 artists are largely dependent on conventional song forms for their commercial success" (Osborn 2011: 1). Typical pop artists include The Beatles, ABBA, and Lady Gaga.

As previously mentioned, some overlap occasionally occurs between pop music and other genres. The most relevant case for the purposes of this essay is Macklemore's "Thriftshop" (2012), which could arguably be classified as rap/hip hop, but as *Spotify's* "Radio" classifies it as a pop song, it was counted as such in this essay.

1.2.5 Punk

Punk, similar to hip hop, is a word used to describe a culture as well as a genre of music. For the purposes of the present study, the word "punk" will be used to refer specifically to the music, unless explicitly otherwise defined. The punk movement has its roots in the late sixties, when it started out primarily as an anti-establishment movement (Tschmuck 2006: 143). It should be noted that there is no unifying ideology in punk culture other than anti-establishment and anti-conformity sentiments. Typical punk bands include Bad Religion, NOFX, Rancid, the Ramones, and the Sex Pistols.

Punk music tends to be loud and relies heavily on distorted guitars, often times lyrics are screamed or growled rather than sung. Gendron describes the "themes of a punk aesthetic" to be "sheer aggressiveness and loudness, the element of physical shock" coupled with "minimalism and defiant rank amateurism" (Gendron 2002: 233, 234). This, coupled with the anti-establishment sentiments results in punk music often being angry in its tone. With this in mind, we should expect to see less reference to the self in punk lyrics compared to the lyrics of the other three genres under study, as angry people tend to focus their speech on the object of their frustration, which is seldom the self (Pennebaker 2011: 107).

1.2.6 Club/house

Club/house music, as the name implies, is music that is made to be danced to at club or house parties. House music can contain recordings of real instruments, or rearranged parts of other songs, but it is most common for songs to be created in part or entirely using a computer to synthesize sounds (Snoman [2006]: 271). It is primarily written to be played at clubs and

parties and is characterized by a simple beat (Snoman [2006]: 271), and its lyrics tend to revolve around love or having fun at a party. This type of music also tends to rely on simple rhythms so heavily that Garcia refers to electronic dance music (such as club/house) as "unapologetically repetitive" (Garcia 2005: n.p.). Since club/house songs are intended to be fun, they are typically up-beat, and it is very uncommon for club/house songs to feature dark or depressing lyrics. Typical club/house artists include David Guetta, Swedish House Mafia, and Tiesto.

2 Literature review

The following section surveys several studies on pronoun frequencies and on what these frequencies can tell the reader about the writer. This section also contains a brief discussion of how these prior studies pertain to the present study.

2.1 Previous studies on the use of pronouns in various text types

Professor James W. Pennebaker is one of the leading researchers in studying how the use of pronouns can reveal personality traits and mental health status. In his book *The Secret Life of Pronouns* (Pennebaker 2011), he summarizes several studies that he and his fellow researchers have conducted over the years on how people's use of words (especially function words, particularly, pronouns) can reveal something about their personality and motivations. Especially interesting is Pennebaker's argument that song lyrics can reveal aspects of the writer's personality (2011: 266). While the present study is not intended to psychoanalyze the mindset of songwriters, it stands to reason that if texts can reveal something about their writer, music lyrics should exhibit differences in style to reflect the different moods, attitudes, and backgrounds of their writers. Accordingly, if there is indeed a difference in content between genres, that difference should also be apparent in the style of writing, and thus in the frequency of pronoun use.

   Two of the studies summarized in *The Secret Life of Pronouns* are particularly interesting. The first of these is Pennebaker, Mehl & Niederhoffer (2003). In this study, the researchers review several methodologies of counting word use and discuss various areas where studying the way a person uses words can be of value. First, as a means of studying behavior, second, as a means to diagnose certain mental conditions. They argue that "the words people use are diagnostic of their mental, social, and even physical state" (Pennebaker et al. 2003: 548). Throughout the article they support this idea by presenting examples of other studies which

reinforce the idea that the way people use words can reveal things about them. For instance, one aspect of pronoun use is that people have a tendency to use inclusive language immediately following a shared traumatic event which has afflicted a group to which the speaker belongs. Following the September 11 terrorist attacks, for example, US students studied by Pennebaker used fewer FPSPs and more FPPPs (Pennebaker et al. 2003: 565). Another finding is about depressed people as they tend to use self-reference more frequently than the average person (Pennebaker et al. 2003: 560). This article confirms the validity of studying the use of function words as a means to glean psychological insights into the mind of the speaker. The authors conclude the article by posing suggestions for further research (Pennebaker et al. 2003: 570-571).

The second article is Pennebaker, Sivertsen & Petrie (2008) which studies the lyrics of The Beatles, and how the group's lyrics changed over time. Computerized text analysis was applied to the lyrics to see how the they changed over time and to reveal the differences between the three writers: McCartney, Lennon, and Harrison, as well as how those lyrics differed from those co-authored by McCartney and Lennon. With regards to the chronological aspect, the most notable change observed was that, as the band grew older, the frequency of FPSPs dropped significantly: from 13.6% in the period 1960-64 to 7.04% during the period 1968-70 (Pennebaker et al. 2008: 199). The study also found that Lennon had more of an impact on the style of the music of The Beatles than the other writers (2008: 201). In conclusion, the authors argue that their project "suggests a methodology by which to track groups and individuals over time to better understand their psychological dynamics" (2008: 202).

In their study of The Beatles' lyrics, Pennebaker et al. also maintain that a high frequency of first person pronouns in a song lyric is an indication of high *immediacy* in the song. Immediacy in this context is "considered to be a marker of the degree to which people are living in the moment" (Pennebaker et al. 2008: 200).

In addition to summarizing these two studies, *The Secret Life of Pronouns* describes Pennebaker's research into other aspects of pronoun and function word use. For instance, he found that as people grow older, they typically use FPSPs less frequently (Pennebaker, 2011: 63). This correlation appears to be what we can see in the lyrics of The Beatles.

Another study summarized in *The Secret Life of Pronouns* (Pennebaker, 2011: 67-69) found that there is a difference in how frequently college applicants from different social backgrounds use pronouns in their applications. Most pertinent to the present study are the findings that applicants from the lower middle class used FPSPs at a rate of 8.31%, middle

class applicants at a rate of 7.98%, and upper middle class applicants at a rate of 7.78%. This gives us an average FPSP rate of approximately 8.02% among all applicants. The differences between the social classes are not particularly relevant to the present study, but they are an interesting basis for comparison when looking at pronoun use in music lyrics, as college applications are a form of writing where one would expect writers to focus on themselves.

There are also differences between the genders in how people use function words. According to Pennebaker (2011: 58-59) women typically use more pronouns (especially pronouns with a person or group of people as the referent, i.e. *you*, *they* etc.), while men tend to use more nouns, prepositions, and articles.

In her doctoral thesis, Marie Gustafsson Sendén writes that "self-inclusive pronouns are consistently expected to be associated with more positive contexts than self-exclusive pronouns" (Gustafsson Sendén 2013: 35). At first, this might seem to contradict the findings made by Pennebaker, Mehl, and Niederhoffer (2003), who argue that there is a well documented connection between depression and a high rate of FPSP use (Pennebaker et al. 2003: 560). However, Gustafsson Sendén's thesis simply argues that self-inclusive pronouns are more likely to appear in a positive context than self-exclusive ones, not that texts that contain a high quantity of self-references are intrinsically positive in tone. Gustafsson Sendén also found that when a text was written in collaboration with multiple authors, instances of *we* occurred in similarly positive contexts as *I* (Gustafsson Sendén 2013: 55).

In 2009, Na and Choi, set out to test whether an individual's cultural background affects their assumptions about language use. The researchers showed a text in a foreign language, with pronouns underlined, to U.S. and South Korean university students. They then asked the students to guess whether each word was a first person singular pronoun (FPSP) or first person plural pronoun (FPPP). Na and Choi found that the American university students, who are (according to Na and Choi) typically more individualistic, were more inclined to guess that a foreign pronoun is a FPSP than South Korean students, who are more collectivistic (Na & Choi 2009: 1496-1497). In other words, one's cultural context affects one's assumptions regarding word use.

Similarly, Kashima and Kashima (1998) argue that languages which allow pronoun dropping, such as Japanese and Chinese, show fewer signs of individualism than languages which do not allow the elision of pronouns, such as English (471). Further, they write that language is on the one hand an expression of an individual's way of conceptualizing the world and, on the other hand, the individual's use of language affects their cognitive processes (461-

462). In simpler terms, it is not a question of whether one's use of language reflects or affects one's personality and emotions; it does both.

It has been found that a high use of FPPPs can create a sense of reduced social distance (Semin 2007: 361), which is also supported by Pennebaker who writes that a more frequent occurrence of we-words in music lyrics shows a focus on couples (Pennebaker 2011: 267). Similarly, Pennebaker, Mehl & Niederhoffer notice that "pronoun use (I, we) in marital interactions can reflect differences in the degree to which couples frame their relationship as inter- or independent" (Pennebaker et al. 2003: 567).

In his article "Power and the Language of Men" Kiesling ([2010]) set out to test how young men use language to form an identity, especially in the context of a hierarchical power structure. To achieve this, he recorded members of a U.S. university fraternity during an election meeting and analyzed their interactions, including word use. While he does not focus specifically on pronoun use, Kiesling mentions that "embedding [a statement] in 'I think'"(Kiesling [2010]: 504-505) is a hedging strategy often used by people, in this case men, in a position of inferior power, to show deference to those in the group of higher power. This agrees with the findings of Pennebaker (2011: 71) that people of higher status tend to use fewer pronouns than people of lower status.

A study conducted in 2012 by Twenge, Campbell & Gentile focuses on the pronoun use of writers of American literature. In their study, the authors argue that books are an expression of the writer's values and attitudes, and as such, indicative of contemporary culture (Twenge et al. 2012: 407). They argue that "American culture has [...] become more individualistic in recent decades" (2012: 406). The central purpose of their investigation being to determine whether the perceived shift in American culture, from collectivism towards individualism, is reflected in a change of pronoun use in American literature published from 1960 to 2008. Using *Google Books Ngram Viewer*, they confirm their hypothesis that the use of plural pronouns has decreased, while the use of singular pronouns has increased (Twenge et al. 2012: 410). Specifically, they noted that the increase in FPSPs seems to start around 1987 (2012: 411). Out of the four genres analyzed for the present study, only punk contained lyrics originally released before 1990, numbering six total. The oldest was "Beat on the Brat" by the Ramones, originally released in 1975. Assuming that the trend observed by Twenge et al. (2012) is identical in music lyrics, it would stand to reason that punk lyrics would use a higher number of FPPPs and fewer FPSPs than the lyrics of the other three genres under study. While literature and music are different genres of text and thus not directly comparable, Twenge et al. argue that the shift in pronoun use in American literature is due to a shift in American

culture. As the majority of the lyrics analyzed in the present study are written by American artists, it would stand to reason that the cultural shift would affect pronoun use in lyrics in a fashion similar to the shift observed in American literature by Twenge et al. (2012).

2.2 First person pronouns in *Google Books Ngram Viever*

In order to compare the frequency of FPSPs and FPPPs in lyrics to the frequency of FPSPs and FPPPs in English-speaking fictive literature, I conducted a case study using *Google Books Ngram Viewer* (*Google Books Ngram Viewer*, online). I entered the pronouns *I, me, my, mine, myself* in the search field, left the smoothing setting unchanged, set the corpus to "English Fiction" and the year to 2008, the most recent year available. I then used the same settings for the plural pronouns. *Google Books Ngram Viewer* is based on the work of Michel et al. (2010), spanning a corpus of over five million books. The exact size of the corpus for English fictive literature published in 2008 is 5,805,045,097 so-called 1-grams, gathered from 59,481 volumes ("Data Sets", online). A 1-gram is "a string of characters uninterrupted by a space; this includes words [...] but also numbers [...] and typos" (Michel et al. 2010: 176). The frequencies below are calculated by "dividing the number of instances of the [search term] in a given year by the total number of words in the corpus in that year" (Michel et al. 2010: 176). According to *Google Books Ngram Viewer*, the frequency (as a percentage) of FPSPs in English fictive literature published in 2008 is, if rounded off to three decimals: 1.411% (*I*) + 0.334% (*me*) + 0.327% (*my*) + 0.011% (*mine*) + 0.028% (*myself*) = 2.111%, or 2.1 instances of first person singular pronouns (FPSPs) per 100 words. If the same calculation is conducted for the FPPPs, we get: 0.204% (*we*) + 0.086% (*us*) + 0.084 (*our*) + 0.002% (*ours*) + 0.004% (*ourselves*) = 0.380%, or 0.4 instances of first person plural pronouns (FPPPs) per 100 words. The FPSP rate of 2.1 per 100 words FPPP rate of 0.4 per 100 words do not mean much in isolation, but these percentages can be a basis for comparison when looking at similar frequencies in the 80 lyrics under study.

3. Method and material

The present study is a primarily quantitative study, calculating the instances of first person singular pronouns (FPSP) and first person plural pronouns (FPPP) per lyric and per genre for the purposes of examining whether there are any noteworthy differences in first person pronoun use between the lyrics of four genres of popular music.

Function words in general, and specifically pronouns, are interesting to study as they offer a quantifiable metrics to a text as an empirical evidence of different strategies in pronoun uses. Speakers and authors do not deliberately choose the frequency of pronouns in their texts or speech (Pennebaker 2011: 19-38), but statistics can reveal strategies in those choices.

3.1 Song lyrics selection and genres

For the purposes of this study, a total of eighty songs were chosen, evenly divided between four different genres. The four genres chosen were, "hip hop", "pop", "punk", and "club/house". The selection of songs within a genre was done randomly using the music streaming software *Spotify's* "Radio" function. The *Spotify* radio is powered by software developed by a company called the Echo Nest (Richmond, 2014: n.p. online). The Echo Nest's analysis software can automatically detect aspects of a musical track and output "a complete description of all musical events, structures, and global attributes such as key, loudness, time signature, tempo, beats, sections, [and] harmony" (Jehan & DesRoches, 2014: 1). The Echo Nest's database of information about songs (so-called metadata, including e.g. track title, artist name, and genre) can then be used for free by non-profit users, or by commercial interests via a computer program provided to partners by the Echo Nest (Jehan & DesRoches, 2014: 1). Utilizing information from this database, the *Spotify* radio selects songs randomly from a genre chosen by the user. The user can then "like" or "dislike" the chosen song, in order to encourage or discourage the selection of similar songs in the future. In order to avoid affecting the algorithms selecting songs, the "like" and "dislike" features were not utilized during song selection. Using the radio feature of *Spotify* allowed for a random selection of musical tracks, and ensured that genre definitions came from a singular, consistent source. This also provided a song selection unaffected by my own arbitrary genre definitions.

The lyrics themselves were gathered from Internet resources which allow users to submit and edit lyrics, using several different websites to ensure that the lyrics were complete (for instance, that chorus repetitions were not omitted). In some cases, no one source contained complete lyrics. In those cases, the lyrics were stitched together using lyrics from multiple websites, while listening to the songs, thus completing the lyrics. In other cases, however, no lyrics could be found; in those cases, the songs were skipped entirely. Similarly, any completely instrumental songs were also skipped, as they contain no semantic information to be analyzed in the first place. Covers were also skipped, as the original author might have written the song in a style more representative of a different genre. The only such case when

selecting lyrics for the present study was the cover the Dead Kennedys made in 1987 of the 1960 song "I Fought the Law", originally written by the Crickets. It appeared during the selection of punk songs and was consequently skipped.

3.2 Counting words

When counting the total number of words in a lyric, all contractions, such as "she's" or "don't", were counted as a single word. The reason for this is that the version of Microsoft's *Word* (Standard Edition 2003) software used for this study counts such contractions as a single word. However, in order to count pronouns accurately, the word "let's" has consistently been split into "let us". When counting words, those with low semantic value, such as "u-huh" have not been counted in the totals. In some cases it was also necessary to correct the spelling of pronouns to ensure that they would be counted correctly. For a complete set of lyrics, including word counts, see the appendix (Complete lyrics and statistics of the use of first person pronouns).

When the artist of a song used their name or stage name to refer to themselves, such as Eminem rapping "Shady's back" (referring to one of his other nicknames, "Slim Shady") in "Without Me", those instances were counted as FPSPs, as they are semantically identical to the words *I* or *me* in this context. In total, 28 instances of using proper nouns in place of first person singular pronouns were found, all of them in the hip hop genre. The majority (17) of those instances were found in the song "Still D.R.E" (1999) by Dr Dre featuring Snoop Dogg.

3.3 Software

The primary pieces of software used to count words in the lyrics analyzed in the present study were the *Textalyser* (textalyser.net) and Microsoft's *Word* software (Standard Edition 2003). Testing showed that *Textalyser* would provide accurate counts for pronouns but not word totals. Accordingly, *Word* was used to count total words, while *Textalyser* was used for counting pronouns.

Textalyser allows the user to insert a text and have the program count the number of instances of each individual word. It also allows the user to define a custom stop-list, controlling which letters or symbols should be counted as punctuation marks and thus ignored. One issue that became evident with *Textalyser* was that it did not count the word *we* or total word counts accurately, so instances of the word *we* were counted manually, using *Word*'s search function. I found out later that this issue with *Textalyser* was the result of a flaw in the pre-defined English stop-list. For whatever reason, the default English stop-list was defined in

such a way that the program did not count articles or the word *we*. It is therefore recommended that anyone using the software in the future should deactivate the default English stop-list and define a custom stop-list.

3.4 Calculating frequencies

The total number of first person singular pronoun occurrences (FPSPs, i.e. *I*, *me*, *my*, *mine*, and *myself*), as well as first person plural pronoun occurrences (FPPPs, i.e. *we*, *us*, *our*, *ours*, and *ourselves*) were counted and used to calculate instances of FPSPs and FPPPs per 100 words. Additionally, instances of FPSPs and FPPPs (per 100 words) of each individual song were used to calculate an arithmetic mean (for the purposes of the present study, the word *mean* will be used interchangeably with *arithmetic mean*) of FPSP and FPPP instances per song (per 100 words) for each of the four genres. A variance calculation was also performed for the FPSP and FPPP instances per song per 100 words. Microsoft's *Excel* Software (Standard Edition 2003) was used for this calculation.

3.5 Songs analyzed

Below follows a complete list of the songs analyzed in this essay, sorted by genre, in alphabetical order by artist. When I first conducted the random generation of songs for the punk genre, I did so on my own *Spotify* account. However, since I had listened to, and rated, punk songs previously I had unintentionally weighted the algorithms used to select songs. Once I noticed this, I created a new *Spotify* account and repeated the song selection, to ensure the selection would be truly random.

3.5.1 Hip hop

The following list is a set of twenty songs included in this study. Each song was randomly selected by the hip hop channel of *Spotify's* radio feature.

2 Chainz feat. Drake - No Lie (2012)

Atmosphere - Yesterday (2008)

Childish Gambino - Heartbeat (2011)

Chris Brown feat. Lil' John & Busta Rhymes - Look at Me Now (2011)

D12 - Purple Pills (2001)

Dr. Dre feat. Snoop Dogg - Still D.R.E. (1999)

Drake - Forever (2009)

Drake - Started from the Bottom (2013)

Eminem - Cleanin' Out My Closet (2002)

Eminem - Without Me (2002)

The Game feat. 50 Cent - Hate It or Love It (2005)

Ice Cube - It Was a Good Day (1993)

Jay Z, Kanye West - Niggas in Paris (2011)

Kanye West feat. Rihanna and Kid Cudi - All of the Lights (2010)

Kanye West feat. Big Sean and Jay Z - Clique (2012)

Kid Cudi feat. Kanye West - Erase Me (2010)

Kid Cudi - Up Up & Away (2009)

OutKast - Ms. Jackson (2001)

T.I. - Whatever You Like (2008)

Wu-Tang Clan - C.R.E.A.M. (1994)

3.5.2 Pop

The following list is a set of twenty songs included in this study. Each song was randomly selected by the pop channel of *Spotify's* radio feature.

Alex Clare - Too Close (2011)

Blur - Song 2 (1997)

Calvin Harris & Ellie Goulding - I Need Your Love (2012)

Coldplay - Viva la Vida (2008)

Franz Ferdinand - Take Me Out (2004)

Fun. feat. Janelle Monáe - We are Young (2011)

Gorillaz - Feel Good Inc. (2005)

Justin Timberlake - Mirrors (2013)

The Killers - Mr: Brightside (2004)

Lorde - Royals (2012)

The Lumineers - Ho Hey (2012)

Macklemore & Ryan Lewis - Thriftshop (2012)

Marina and The Diamonds - Primadonna (2012)

Maroon 5 & Christina Aguilera - Moves Like Jagger (2011)

Muse - Uprising (2009)

Owl City & Carly Rae Jepsen - Good Time (2012)

Pitbull & Ke$ha - Timber (2013)

Rihanna feat. Calvin Harris - We Found Love (2011)

Robin Thicke feat. T.I. & Pharrell - Blurred Lines (2013)

Selena Gomez - Come & Get It (2013)

3.5.3 Punk

During the song selection for the punk genre, the songs "Basket Case" by Green Day and "Beat on the Brat" by the Ramones both appeared twice, and were consequently skipped the second time. As previously mentioned, the Dead Kennedys cover of "I Fought the Law" was skipped on account of being a cover. The following list is the final list of twenty songs included in this study. Each song was randomly selected by the punk channel of *Spotify's* radio feature.

Bad Brains - Pay to Cum (2003 Remaster) (1980/2003)

Bad Religion - 21st Century (Digital Boy) (1990)

Blink-182 - Dammit (1997)

Blink-182 - What's My Age Again? (1999)

The Distillers - City of Angels (2002)

Green Day - 21 Guns (2009)

Green Day - American Idiot (2004)

Green Day - Basket Case (1994)

Green Day - When I Come Around (1994)

Misfits - Astro Zombies (1982)

Misfits - Hybrid Moments (1997)

NOFX - Dinosaurs Will Die (2000)

Ramones - Beat on the Brat (1975)

Ramones - Blitzkrieg Bop (1976)

Rancid - Journey to the End of the East Bay (1995)

Sex Pistols - Anarchy in the UK (1976)

Sex Pistols - God Save the Queen (1977)

Sum 41 - Fat Lip (2001)

Sum 41 - In Too Deep (2001)

Sum 41 - Still Waiting (2002)

3.5.4 Club/house

During the song selection for the club/house genre, the song "UFO" by Vigiland was skipped as it is an instrumental song and thus contains no lyrics. Similarly, "Bump and Grind (2014

Remix Radio Edit)" by Waze & Odyssey featuring R. Kelly, "Sugar Man (Radio Edit)" by Yolanda Be Cool featuring Dcup, and "Tear the Roof Up (Extended Version)" by Alesso were skipped, as no reliable lyrics were available on the Internet at the time of writing the present study. Below is the final list of twenty songs included in this study. Each song was randomly selected by the club/house channel of *Spotify's* radio feature.

Alesso feat. Tove Lo - Heroes (We Could Be) (2014)

Alesso feat. Roy English - Cool (2015)

Avicii - The Nights (2014)

Axwell, Sebastian Ingrosso, Steve Angello and Laidback Luke feat. Deborah Cox - Leave the World Behind (2009)

Calvin Harris feat. Ellie Goulding - Outside (2014)

Charli XCX - Break the Rules (Tiesto Remix) (2014)

Contiez feat. Treyy G - Trumpsta (Djuro Remix) (2013)

David Guetta feat. Emeli Sandé - What I did for Love (2015)

David Guetta feat. Sam Martin - Dangerous (2014)

Ed Sheeran - Thinking Out Loud (Alex Adair Remix) (2014)

Knife Party - Bonfire (2012)

Madeon feat. Kyan - You're On (2014)

Major Lazer feat. MØ and DJ Snake - Lean On (2015)

Omi - Cheerleader (Felix Jaehn Remix Radio Edit) (2014)

Robin Schulz feat. Jasmine Thompson - Sun Goes Down (2014)

Route 94 feat. Jess Glynn - My Love (2014)

Shapeshifters - Lola's Theme (Radio Edit) (2004)

Tove Lo - Talking Body (KREAM Remix) (2015)

Years & Years - King (2015)

Zhu - Faded (2014)

4. Results and discussion

The following section contains graphs and tables reporting the data (total word counts, FPSP and FPPP instances per 100 words per genre, and FPSP and FPPP instances per song per 100 words) collected and calculated for the present study, as well as brief discussions regarding what the data could potentially imply.

4.1 Overall frequencies

This section contains tables reporting the data gathered for the present study.

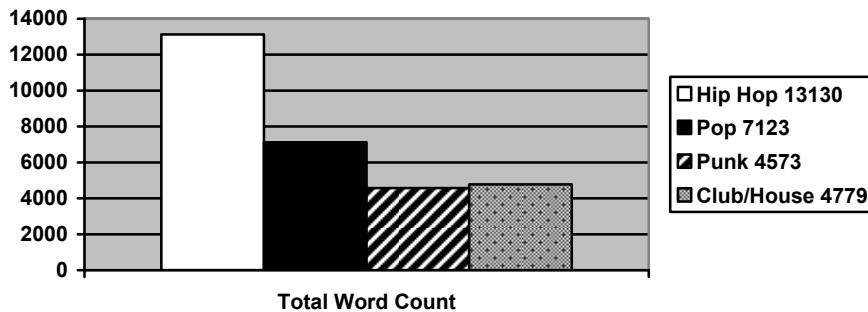The graph below represents the total word counts of the twenty songs from each genre.

Figure 1. Total word count of four genres.

Of note here is that the total number of words in the hip hop genre is nearly twice as high as that of the second wordiest genre, namely pop. The sheer volume of words seems to lend credence to the notion that hip hop focuses more on the lyrical aspect of writing music, rather than focusing on building melodies and harmonies, at least compared to the other three genres studied in this essay.

Table 1. Raw frequencies and instances of First Person Singular Pronouns (FPSPs) per 100 words, including arithmetic mean instances per 100 words per song.

Genre	Total Words	Raw Frequencies of FPSPs	Instances of FPSPs per 100 words	Mean FPSP instances per song (per 100 words)
Hip hop	13130	1209	9.2	9.1
Pop	7123	739	10.4	10.8
Punk	4432	334	7.5	6.9
Club/house	4779	512	10.7	11.3
Mean	7401.25	698.5	9.45	9.5

Out of the 80 songs under study in this essay, three had more than 20 instances of first person singular pronouns (FPSPs) per 100 words. The highest FPSP rate of all songs was "Trumpsta (Djuro Remix)" by Contiez featuring Treyy G with a FPSP rate of approximately 24.5 instances per 100 words. This exceptionally high FPSP rate is partly due to the fact that

the words "I'm Mr [sic]" are repeated several times as part of a build up at one point in the song. The second highest FPSP rate was also found in the club/house genre, specifically, "My Love" by Route 94 featuring Jess Glynn had an FPSP rate of approximately 21.8 instances per 100 words. It was closely followed by "Song 2" by Blur (from the pop genre) at 21.7.

Table 2. Raw frequencies and instances of First Person Plural Pronouns (FPPPs) per 100 words, including arithmetic mean instances per 100 words per song.

Genre	Total Words	Raw Frequencies of FPPPs	Instances of FPPPs per 100 words	Mean FPPP instances per song (per 100 words)
Hip hop	13130	122	1	1
Pop	7123	124	1.7	2
Punk	4432	71	1.6	1.6
Club/house	4779	133	2.8	2.6
Mean	7401.25	112.5	1.78	1.83

The song with the highest FPPP rate among those analyzed was "Heroes (We Could Be)" by Alesso featuring Tove Lo, belonging to the club/house genre, with 12.7 instances of FPPPs per 100 words. That it belongs to the club/house genre is not surprising, since the lyrics of the club/house genre had the highest average FPPP instances per 100 words (2.6, Table 2). The song with the second highest FPPP rate was "Uprising" by Muse (pop genre) with 9.3 instances of FPPPs per 100 words. The lyrics of "Uprising" are interesting as they contain unusually many instances of FPPPs per 100 words, but zero instances of FPSPs. This stands in stark contrast to the only other songs to contain no FPSPs, namely "Beat on the Brat" by the Ramones which contained no FPPPs.

4.2 Variance in FPSP and FPPP frequencies per song
This section contains data regarding the statistical variance in the number of FPSP and FPPP instances per 100 words per song within each of the four genres under study in the present thesis, namely hip hop, pop, punk, and club/house. There are also examples of statistical outliers in terms of FPSP and FPPP instances per song (per 100 words) from each of the four genres.

The graph below represents the variance in FPSP instances per song (per 100 words) within each of the four genres under study.
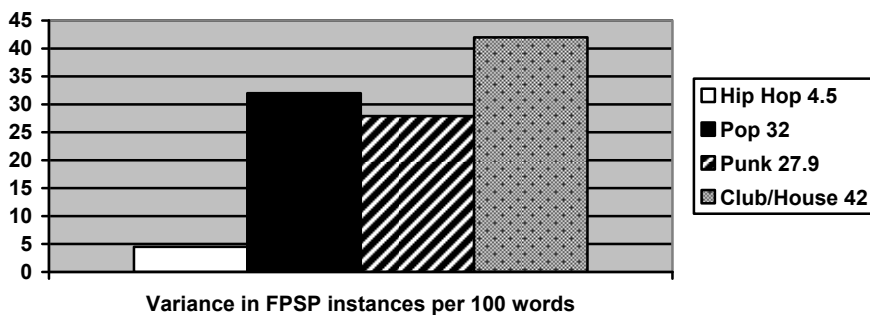

Figure 2. Variance between songs in FPSP instances per 100 words in four genres.

Out of the four genres under study, the hip hop genre had the least amount of variance in FPSP instances per 100 words (Figure 2). It contained no song with fewer instances of FPSPs per 100 words than Drake's "Started from the Bottom" (5.2), and no song with a FPSP rate higher than Eminem's "Cleanin' Out My Closet" (13.2). If a person's pronoun use can tell a reader something about the writer, as Pennebaker (2011) argues, then similarities in the way a group of people use pronouns should tell us that the group is likely homogenous. No research has been conducted with regards to the cultural and social backgrounds of the lyrics writers under study. However, given how much less variance is reported in the hip hop lyrics under study (Figure 2), it is a plausible assumption that the hip hop lyrics' writers are more homogenous as a group than the writers of the other three genres.

The pop genre contained lyrics with a significantly larger variance of FPSP instances per 100 words. Several songs from the pop genre contained lyrics with around 20 FPSPs per 100 words, e.g. "Song 2" by Blur (21.7) and "Take Me Out" by Franz Ferdinand (19.7), while others contained FPSP rates far below the pop average, e.g. "Feel Good Inc." by The Gorillaz featuring De La Soul (4). The lyrics of one song in the pop genre contained no instances of FPSPs at all, namely "Uprising" by Muse.

In the punk genre, one can also see significant variance in the FPSP rates of different lyrics (Figure 2). The only song in the punk genre with a FPSP rate approaching 20 occurrences per 100 words was "Basket Case" by Green Day, with a FPSP rate of 19. The second highest FPSP rate in the punk genre was found in the lyrics of "Astro Zombies" by The Misfits (14.2). Of all the genres under study in the present thesis, the punk genre contained the highest number of songs (fifteen) with a FPSP rate of fewer than 10 instances per 100 words e.g. "21 Guns" by Green Day (1.2) and "Blitzkrieg Bop" by The Ramones (1.1). Furthermore, the only

song other than "Uprising" by Muse to contain no occurrences of FPSPs at all belonged to the punk genre, namely "Beat on the Brat" by The Ramones.

The lyrics of the club/house genre exhibited the highest degree of variance in FPSP instances per song per 100 words (Figure 2), being the only genre to contain more than one song with over 20 occurrences of FPSPs per 100 words, namely "Trumpsta (Djuro Remix)" by Contiez featuring Treyy G. (24.5) and "My Love" by Route 94 featuring Jess Glynn (22). There were also other songs in the club/house genre with an unusually high FPSP rate per 100 words, e.g. "Lola's Theme" by The Shapeshifters (19.3), "Faded" by Zhu (17.4), and "Cheerleader (Felix Jaehn Remix)" by Omi (17.3). The lyrics of the club/house genre also contained lyrics with less than 10 FPSP occurrences per 100 words, e.g. "Sun Goes Down" by Robin Schulz featuring Jasmine Thompson (1.1) and "Lean On" by Major Lazer featuring MØ and DJ Snake (0.3). The fact that the club/house genre reports the highest degree of variance in the number of FPSP instances (per 100 words) used per song, implies that the writers of club/house lyrics, contrary to the hip hop writers, seem to be a diverse group of artists from different cultural and social backgrounds.

The graph below represents the variance in FPPP instances per song (per 100 words) within each of the four genres under study.
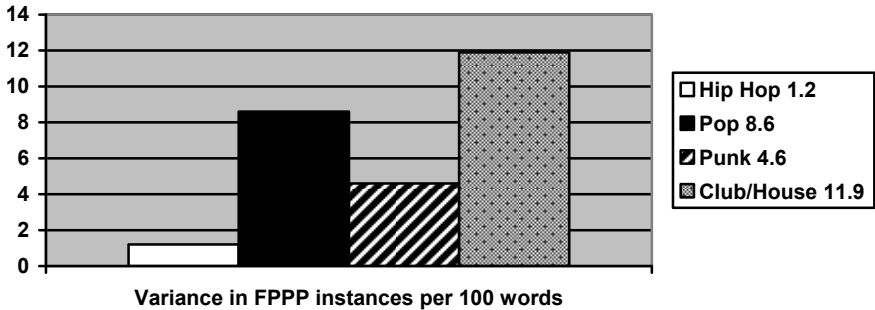
Figure 3. Variance between songs in FPPP instances per 100 words in four genres.

The data reported by Figure 3 seems to mirror the pattern in Figure 2. As previously noted, the lyrics of the hip hop genre reported the least amount of variance in FPSP instances per song per 100 words. This pattern is also consistent in the FPPP instances per song per 100 words among the hip hop lyrics under study (Figure 3). While the lyrics of all genres under study contained at least two songs with no FPPPs in the lyrics at all, the hip hop genre is the only genre that contained only two. Similarly, all of the three other genres contained at least two songs with more FPPP instances per 100 words than the song with the highest rate of FPPP instances in the hip hop genre, namely "Started from the Bottom" by Drake (5.2 instances per 100 words). Out of the hip hop lyrics under study, two lyrics contained 2 or

more FPPP instances per 100 words (the other being "Heartbeat" by Childish Gambino at 2 instances per 100 words).

The lyrics of the pop genre exhibited more variance than the hip hop lyrics in the instances of FPPPs per 100 words (Figure 3). Among the pop lyrics under study, there were four that contained zero occurrences of FPPPs. The two lyrics with the highest FPPP rates were "Uprising" by Muse (9.3) and "Royals" by Lorde (8.4).

The lyrics of the punk genre reported slightly less variance in FPPP rates than the pop lyrics (Figure 3). The song with the most instances of FPPPs per 100 words among the punk lyrics under study was "Pay to Cum (2003 Remaster)" by the Bad Brains (7.3 per 100 words).

The club/house genre reported the largest variance of FPPP instances per song per 100 words (Figure 3), with eight of the lyrics under study containing no occurrences of FPPPs. The club/house genre also contained the lyrics with the most instances of FPPPs per 100 words, namely "Heroes (We Could Be)" by Alesso featuring Tove Lo (12.7). The club/house song with the second most instances of FPPPs per 100 words was "Sun Goes Down" by Robin Schulz featuring Jasmine Thompson (8.8). Out of the club/house lyrics under study, eight contained no FPPPs at all.

4.3 First person singular pronoun frequencies

The graph below reports the instances of FPSPs per 100 words in the four genres under study.
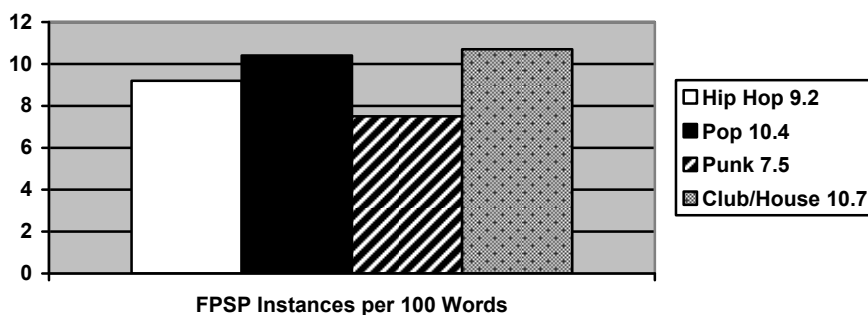
Figure 4. FPSP Instances per 100 words in four genres.

In this graph we can see what is perhaps the most interesting finding of this study: as punk music is typically angrier than the other genres studied here, we should expect punk lyrics to contain more references to the object of the writer's frustration and fewer references to the self (Pennebaker 2011: 107). This does indeed appear to be the case in music; it certainly holds true for the songs analyzed in this study, as the punk songs report an FPSP rate of 7.5 per 100 words, while the pop and club/house genres use 10.4 FPSPs per 100 words and 10.7 FPSPs per 100 words respectively (Figure 4). To illustrate the angrier lyrics in the punk genre, the

song "Still Waiting" by Sum 41 contains the lines "drop dead / a bullet to my head / your words are like a gun in hand", "Beat on the Brat" by the Ramones contains the repeated lyrics "beat on the brat with a baseball bat", City of Angels by the Distillers contains the lines "going down to the gravel head to the barrel / take this life and end this struggle", and there are several similar examples of angry or frustrated lyrics in the punk songs that were chosen for the present study.

The average FPSP rate across genres amounts to approximately 9.45 per 100 words (table 1). This is notably higher than the FPSP rate observed by Pennebaker in US college application letters, which was approximately 8.07 per 100 words (2011: 67-69). College applications are a type of writing where one might expect the writer to focus on the self to a high degree. After all, the applicant must typically describe why they want to study at the school they are applying to, as well as describe their academic achievements. However, as a college application letter is a formal document, written in an academic context, we should not expect them to contain a disproportionately high degree of pronouns, as people tend to use fewer pronouns in formal settings (Pennebaker 2011: 47). Furthermore, as lyrics and application letters are different genres of text, any comparison of pronoun usage between the two is tentative. However, it is still interesting to consider some of the differences between the text types that result in different pronoun use.

Generally, according to Pennebaker (2011: 39-72), women use FPSPs more than men, younger people more than older people, working class more than upper class, and the more dominant party in a relationship uses fewer FPSPs than the other(s). None of these patterns seem to sufficiently explain why music appears to have a higher rate of FPSP instances per 100 words than college application letters. According to the National Center for Education Statistics, a part of the U.S. Department of Education, 9.8 million female students and 7.7 male students were enrolled in undergraduate education in the United States in 2013 ("Undergraduate Enrollment", online). Assuming that the gender ratio in applications studied by Pennebaker is at least vaguely similar to the gender ratio among enrolled students, we can also assume that a majority of applicants are females. However, out of the eighty songs analyzed for the present study, only twenty were authored or co-authored by a female artist so, the gender aspect does not explain the difference in FPSP rate between college applications and music lyrics.

On the other hand, the gender aspect might explain some of the difference between the genres here, as both punk and hip hop had only one song each written or co-authored by a

woman and those two genres have the lowest FPSP rates among the songs analyzed for the present study.

Figure 5 below is the reports average FPSP rates per song from the four genres under study. The rates are given in instances per 100 words.
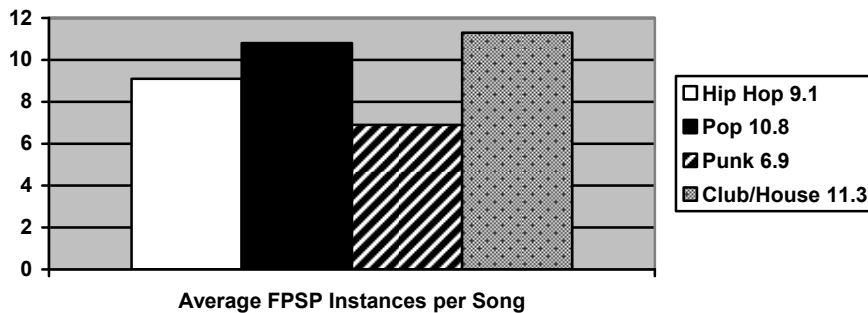

Figure 5. Average FPSP instances per song (per 100 words) in four genres.

In both Figures 4 and 5, we can see that the rate of FPSP use in music appears much higher than that used in literature. According to *Google Books Ngram Viewer,* the average FPSP rate of English fictive literature (presented in section 2.2) is approximately 2.1 instances of FPSPs per 100 words (*Google Books Ngram Viewer*, online). This means that the average FPSP frequency in music lyrics (9.5, Table 1) is approximately four and a half times higher than that in fictive literature.

This seems to imply that musicians rely on first-person narratives far more frequently than authors of fictive literature. As pop lyrics and fiction are different genres of text, one should not expect them to contain comparable frequencies of pronoun use. However, it is still interesting to discuss some of the potential explanations for pop lyrics' apparent higher immediacy and lower register of formality. Some aspects that might explain this are that older people tend to use fewer FPSPs than younger people, individuals of higher social class use fewer than lower class, and formal language tends to use fewer FPSPs (Pennebaker 2011: 63-80). While no investigations have been made into the age or social class of either authors of fiction or music lyrics for the purposes of this essay, it seems realistic to assume that writers of fiction are typically older and of higher social class than the average pop musician and as such write texts with a higher degree of formality, resulting in fewer FPSPs. However, that is a highly tentative assumption.

There are a large number of factors which affect the number of pronouns a person uses. Therefore, it is difficult to ascertain the reasons behind the differences in pronoun use between pop music and fiction, as well as the difference between punk music and the other genres. The focus on frustration, as previously discussed might be one explanation, though it

is difficult to say whether it accounts for the entire difference (approximately 3 per 100 words) between punk music and the total FPSP rate per song average across all four genres (9,5 per 100 words).

As discussed in the previous section, another possible reason for the difference between punk and the other genres could be that punk is very male-dominated (out of the songs analyzed here, the only one performed by a woman is "City of Angels" by The Distillers), and women tend to use pronouns more than men (Pennebaker 2011: 58-59). However, out of the hip hop songs analyzed in the present study, the only one to contain a female writer is "All of the Lights" by Kanye West featuring Rihanna and Kid Cudi; so, while the gender aspect might explain why hip hop and punk appear to have a lower FPSP rate than the other two genres, the gender aspect hardly explains the difference in FPSP frequency between the punk and hip hop songs.

The gender aspect also fails to explain why pop lyrics appear to have a much higher FPSP rate than English fiction (as presented in section 2.2). As previously mentioned, only 25% of songs analyzed in the present study were written or co-authored by a women. This can be compared to an investigation by the online news aggregator and commentary blog The Huffington Post, which found that approximately 32% of the books published in 2011 by four U.S. publishers (Alfred A. Knopf, Farrar, Strauss and Giroux, Crown, and Little, Brown) were written by women ("Are Book Publishers to Blame for Gender Discrimination?", online). While it is by no means certain that this number is representative of the entire publishing industry, it does indicate that the ratio of women to men in fictive literature is higher than that same ratio among lyric authors investigated in the present study. As such, the gender aspect fails to explain why authors of music lyrics appear to use FPSPs at such a high rate compared to authors of literature.

A plausible explanation for the lower rate of FPSPs in punk music is the fact that no other genre studied here contained any songs released prior to 1990, while the punk genre contained six such songs: "Anarchy in the UK", "Astro Zombies", "Beat on the Brat", "Blitzkrieg Bop", "God Save the Queen", and "Pay to Cum". As the study performed by Twenge et al. found that there is a trend in American literature, from 1960 to 2008, towards using more singular pronouns and fewer plural pronouns (Twenge et al. 2012: 410-411), we might be seeing a similar phenomenon here.

At the same time, out of the twenty punk songs analyzed for the present study, two of the six originally released before 1990 had a higher FPSP rate than the punk average of 6.9 per 100 words (Figure 5). "Anarchy in the UK" had a FPSP rate of approximately 12.8, while

"Astro Zombies" had a FPSP rate of approximately 14.2, more than twice the punk genre average. The average FPSP rate of the six punk songs released before 1990 is 5.1 per 100 words ((Anarchy:12.8+Astro:14.2+Beat:0+Blitz:1.1+God:0.5+Pay:2.1)/6≈5.1).

An FPSP rate of 5.1 per 100 words is indeed notably lower than the punk average. However, knowing that the average FPSP rate of those songs is 5.1, we can fairly easily calculate that the FPSP rate average of the remaining fourteen punk songs is approximately 8.7 per 100 words. While this is significantly higher than the punk genre average FPSP rate of 6.9 (Figure 5), it is still notably lower than the pop average of 10.8 (Figure 5) or the club/house average of 11.3 instances of FPSPs per song per 100 words.

To judge from these numbers, it appears as though the difference between the punk genre and the two others might be related to the chronological factor observed by Twenge et al. (2012: 410), and the anger/frustration aspect observed by Pennebaker (2011: 107).

Further, as the average FPSP rate of the fourteen newer punk songs is 8.7 per 100 words, we can see that excluding the six older songs we get the punk average much closer to the hip hop FPSP rate average of 9.1. So, we can see that the difference in average FPSP instances per song per 100 words between hip hop and punk shrinks if the chronological aspect is eliminated. As such, the best explanation for the fact that punk and hip hop use lower average FPSP instances per song per 100 words than pop and club/house, might be that punk and hip hop are more male-dominated, since women tend to use more pronouns than men (Pennebaker 2011: 58-59).

As previously mentioned, the data collected for the present study seems to indicate that music follows the pattern of increasing FPSP frequencies as observed by Twenge et al. in their study of the changes in pronoun use in American literature over time (2012: 410). Some caution is required, however, when attempting to draw conclusions in regards to the chronological aspect. Not only is the number of songs here older than 1990 very low, but in addition, two of the songs that were both released prior to 1990 and had lower than average FPSP rates ("Beat on the Brat" and "Blitzkrieg Bop"), were written by the Ramones, so it could simply be the case that they used very few FPSPs.

The connection described by Pennebaker et al. (2003: 560) between depression and a high FPSP rate does not appear to adequately explain the patterns of FPSP use observed in the present study. As Pennebaker (2011: 18-20) points out, the semantic content of a text is not necessarily an indicator of the mental health of the author. Accordingly, it is possible that the club/house artists analyzed here could be depressed in spite of the generally positive semantic contents of their songs, but there is nothing indicating that the club/house artists analyzed here

should have a consistently higher rate of depression than the writers in the other genres. There are, however, a few songs from various genres that appear to fit the pattern of negativity coupled with a high FPSP rate. For example, "Cleanin' Out My Closet" by Eminem is about his bad relationship with his parents, containing the lines "My whole life I was made to believe I was sick when I wasn't / 'Til I grew up, now I blew up, / it makes you sick to ya stomach / Doesn't it?"[sic], and this text has the highest FPSP rate of the hip hop genre with a FPSP rate of 13.2 instances per 100 words. In the same vein, "Erase Me" by Kid Cudi and Kanye West has a FPSP rate of 12.2 instances per 100 words and starts out with the lines: "She said I don't spend time like I really should / She said she don't know me, / anymore / I think she hates me deep down / I know she does She wants to erase me." Similarly, "Basket Case" by Green Day has the highest FPSP rate of the punk genre at 19 per 100 words and starts with the lines "Do you have the time / To listen to me whine?" The song "Ho Hey" by the Lumineers, while not altogether negative in tone, also features a high rate of FPSP instances per 100 words (18.6) coupled with lines indicating unhappiness: "All the blood that I would bleed / I don't know where I belong / I don't know where I went wrong." With a similarly negative tone, and a similarly high FPSP rate (19.7 per 100 words) "Take Me Out" by Franz Ferdinand has the lines: "And if you leave here / You leave me broken, shattered, / I lie I'm just a crosshair / I'm just a shot, then we can die. / I know I won't be leaving here with you." As previously mentioned, however, there are also several songs with a high FPSP rate that do not conform to the idea that high FPSP frequency correlates with a negative tone. One such example is "Song 2" by Blur with an FPSP rate of 21.7 per 100 words, without any clear hints of failed relationships or other negative thoughts. There are several more such examples. For instance, "Trumpsta (Djuro Remix)" by Contiez featuring Treyy G and "Cheerleader(Felix Jaehn Remix Radio Edit)" by Omi have notably high FPSP rates (25.4 per 100 words and 17.3 per 100 words respectively) without any overt signs of depression or negativity. One explanation for the higher FPSP rate in the more positive songs is a high degree of what Pennebaker, Sivertsen & Petrie (2008: 200) refer to as "immediacy".

Some additional factors that might affect the FPSP rates of the studied songs are the age, social class, and education level of the artist who wrote the songs. However, for the purposes of this study no investigation has been made into the personal lives or backgrounds of the various artists. Further qualitative research would be required to confirm exactly how the social and cultural backgrounds of the lyric writers affect the texts they produce, as well as whether the texts conform to patterns observed in prior studies.

4.4 First person plural pronoun frequencies

Figure 6 represents the instances of all first person plural pronouns (FPPP) per 100 words for the four genres under study.
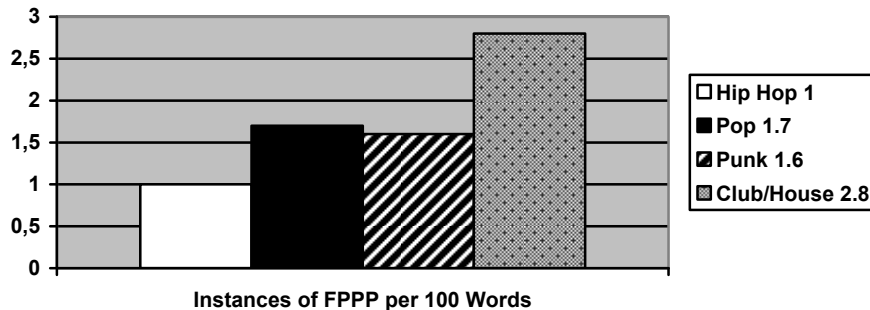
Figure 6. FPPP instances per genre.

Here we can see another interesting finding of the present study. Club/house music uses a higher rate of FPPPs per 100 words than the other genres (2.8 per 100 words, Figure 6), using 1.1 more FPPPs per 100 words than the genre with the second highest rate, namely pop (1.7 instances per 100 words, Figure 6). One possible explanation for why the club/house genre appears to use more plural pronouns than the other genres is that club/house music tends to rely more heavily on collaborations than the other three genres. Out of the twenty club/house songs analyzed for the present study, eleven were collaborations between two or more artist/groups. In contrast, while several of the hip hop tracks analyzed here were made by groups, only five of the songs were collaborations by two or more artists/groups. Another plausible reason is that club/house music tends to be created for the purpose of being played at dance clubs and parties. It would, therefore, be reasonable for those artists to write lyrics that focus on inclusiveness in a group, rather than establishing a more hostile "us versus them" dichotomy. Additionally, club/house music tends to be more positive in tone. This coupled with the higher degree of collaboration in the club/house songs analyzed here might also explain the higher FPPP rate which agrees with Gustafson Sendén's (2013: 55) findings, namely that *we* appeared in more positive contexts if the text was written in a collaborative setting. Similarly, as Semin (2007: 391) points out, the use of *we* rather than *I* can create a sense of social proximity, which might plausibly be something that club/house artists strive for, in order to facilitate social interactions at parties and clubs. It is, of course, also possible that a larger data set would even out these differences. Especially as the difference in Figure 6 between the FPPP rates of the club/house lyrics (2.8 per 100 words) and the hip hop lyrics (1 per 100 words) is only 1.8 per 100 words.

Figure 5 below represents averages of first person plural pronoun (FPPP) instances per song, presented as instances per 100 words in the four genres under study.
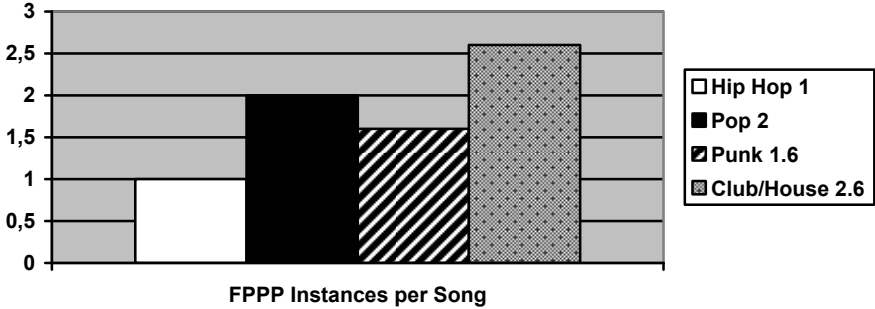
Figure 7. Average FPPP instances per song (per 100 words) across four genres.

In Figure 7 above, the general pattern remains, as the lyrics of the club/house genre have the highest average rate of FPPPs per 100 words per song (2.6), with the lyrics of the pop genre having the second highest rate (2). However, the difference between those two genres as reported in Figure 7 has decreased slightly compared to the difference reported in Figure 6; from a difference of 1.1 instances per 100 words in Figure 6, to a difference of 0.6 in Figure 7. The pattern observed in the FPSP section, namely that lyrics use significantly more FPSPs than fictive literature, also holds true for the FPPPs. According to the case study I conducted using *Google Books Ngram Viewer* (presented in section 2.2), English-language fictive literature uses approximately 0.4 instances of FPPPs per 100 words. Lyrics use significantly more. In the lyrics under study, the hip hop lyrics used on average 1 instance of FPPPs per 100 words, the pop lyrics used 2, the punk lyrics used 1.6, and the club/house lyrics used 2.6 (Figure 7). Again, lyrics and fictive literature are different genres of text and one should not expect them to be comparable at a 1 to 1 ration. However, it is worth noting that the lyrics under study consistently use first person pronouns at a much higher rate than fictive literature.

Gustafsson Sendén's study (2013: 55) found that *we* tends to appear in more positive contexts if a text is written collaboratively. Thus it would to reason, as previously mentioned, that collaborations that contain a high frequency of FPPP instances per 100 words should be positive in tone. Out of the club/house lyrics under study here, fourteen out of twenty were written either by a group or as a collaboration between multiple artists. So, it would seem reasonable to assume that they are positive in tone, since Figures 6 and 7 report that the club/house lyrics use the most FPPPs per 100 words. For example, "Heroes (We Could Be)" by Alesso featuring Tove Lo (club/house genre), contained the highest number of FPPP instances per 100 words (12.7) of any of the lyrics under study and contains the lines: "You and me we got the world in our hands" and "We can do anything". On the other hand, the

lyrics of "Lean on" by Major Lazer featuring MØ and DJ Snake express uncertainty about a relationship: "Do you recall, not long ago / [...] / All we did was care for each other / But the night was warm / We were bold and young / All around the wind blows / We would only hold on to let go" and contain a high frequency of FPPPs per 100 words (7.9). The lyrics of "What I Did for Love" by David Guetta featuring Emeli Sandé, contained 2.8 FPPPs per 100 words and is overtly negative: "Crash and we burn into flames / Stitch myself up and I'd do it again". In a similar vein, the lyrics of "Sun Goes Down" by Robin Schulz featuring Jasmine Thompson contain melancholic elements: "You are so lonely now / You are so lonely now" and a high rate of FPPPs per 100 words (8.8). However, the word *we* appears in a slightly more positive context: "Doesn't matter where we are/ [...] / If there's a moment when it's perfect / We'll carve our names". It could be the case that Gustafsson Sendén's (2013: 55) findings cannot be applied to lyrics. However, it is also possible that the lyrics of these songs were not written collaboratively. Perhaps one artist composed the music while the other wrote the lyrics, or perhaps a vocalist was brought in to perform the singing without having any other input in the production of the song. More qualitative research would be required to ascertain the details of the production of each of the songs to confirm whether Gustafsson Sendén's (2013: 55) findings apply to club/house lyrics.

5. Conclusion

The purpose of this study was to examine and compare the pronoun frequencies in music lyrics written in English. Having compiled and analyzed the data on first person singular pronoun (FPSP) and first person plural pronoun (FPPP) frequencies, the study has revealed some patterns of pronoun usage in four genres of lyrics.

The most interesting finding is a pattern differentiating the genres in terms of pronouns frequencies. Specifically, the punk genre uses significantly fewer FPSPs while the club/house genre is characterized by more instances of these pronouns. Possible explanations for this difference include the fact that punk is a very male-dominated genre, and males have been found to use less FPSPs; however, this does not explain the difference between the punk songs and hip hop songs, as hip hop is also a highly male-dominated genre. Another explanation is that punk music is typically angrier than other types of music; since punk itself is (per its definition) a counter-culture, it has to be critical of some aspect(s) of society.

At the same time, the club/house genre tends to use FPPPs significantly more frequently than the other genres. One possible reason for this is that club/house music is intended to be

played at parties and its writers want to create a sense of togetherness among the party-goers. Since the club/house songs also have an average FPSP rate that is higher than the total average FPSP rate of the four genres, it could also be the case that club/house music uses a greater amount of pronouns in general. Further research would be required to confirm whether club/house music actually does use a high frequency of pronouns. It is possible that club/house authors frequently use first person pronouns to create songs with a higher degree of what Pennebaker, Sivertsen & Petrie (2008: 200) call immediacy, although further research would still be required to determine why club/house artists would produce songs with higher immediacy than, for instance, hip hop artists.

Music lyrics appear to have significantly higher first person pronoun frequencies than English-language fictive literature. It would stand to reason that musicians rely more frequently on first person narratives than authors of fiction, though my research implies no satisfactory answers as to why authors of fiction use the first person perspective less, or why musicians use the third person perspective so sparingly. One possible explanation for the difference could be that authors of fictive literature are older, or of a higher social class, but the present study has made no investigations into the average age or social class of either group of authors.

According to the statistical data, there is significantly less variance in the frequencies of FPSPs and FPPPs per song per 100 words in the hip hop genre compared to the other three genres. This seems to indicate that writers of hip hop lyrics are more homogenous as a group than the artists of the other three genres. Further research into the cultural and social backgrounds of the artists of all four genres would be necessary in order to confirm this.

All in all, the findings of the present study appear to be in line with the findings of previous research. Thus in agreement with results of previous studies, the punk genre had the lowest FPSP rate and showed the highest degree of at least three factors that should result in a lower FPSP rate, namely the male-dominated nature (sharing the highest rate of male authors with the hip hop genre), the expression of frustration, and the fact that the punk genre contained the highest number of songs originally released prior to 1990 compared to the other three genres under study. Further, it seems possible to predict the stylistic content, specifically the pronoun frequency, of a body of song lyrics based on the genre they belong to. It appears, then, as though the four music genres analyzed here are indeed written from different mindsets or perspectives.

As a subject of further research, it would be interesting to conduct a more qualitative study, perhaps limiting its scope to a handful of artists within a genre; thus one could investigate

whether they conform to or break the patterns observed here. Another aspect worth investigating is to explore whether the patterns of first person pronoun usage observed here are recurrent or are particular to this specific selection of lyrics. By using genre classifications from a source other than *Spotify's* radio, one could also see if the patterns remain consistent, or whether a selection of lyrics based on a different classification of genres produces any significant changes in the results.

References

Primary material

AZLyrics. azlyrics.com. n.d. Web. 3 Mar. 2015
    http://www.azlyrics.com/

lololyrics. Lololyrics.com. n.d. Web. 3 Mar. 2015
    http://www.lololyrics.com/

LyricsFreak. Manifest Holdings Inc. n.d. Web. 3 Mar. 2015
    http://www.lyricsfreak.com/

PLyrics.com. PLyrics.com. n.d. Web. 3 Mar. 2015
    http://www.plyrics.com/

Rap Genius. Genius Media Group Inc. n.d. Web. 26 Aug. 2014
    http://rap.genius.com/

Sing365. SING365.com. n.d. Web. 26 Aug. 2014
    http://www.sing365.com/index.html

SongLyircs. songlyrics.com. n.d. Web. 3 Mar. 2015
    http://www.songlyrics.com/

Stretch, Owen; Hutchison, Amanda. *METROLYRICS*. CBS Interactive Inc. n.d. Web. 3
    Mar. 2015
    http://www.metrolyrics.com/

Secondary material

"Are Book Publishers to Blame for Gender Discrimination?" *HuffPost Books*.
    TheHuffingtonPost.com, Inc. Mar. 2012. Retrieved July 10 2015
    http://www.huffingtonpost.com/2012/03/13/book-gender_n_1324560.html

Boone, Christine. 2013. "Mashing: Toward a Typology of Recycled Music." In Yorgason,
    Brent (ed.) *Music Theory Online*. 19(3) n.p. Retrieved December 1 2015.
    http://www.mtosmt.org/issues/mto.13.19.3/mto.13.19.3.boone.pdf

Chambers, J.K. 2003. *Sociolinguistic Theory*. (2nd ed.) Oxford: Blackwell Publishing.

"Data Sets", *Google Books*. Google Corporation. Retrieved January 22 2016
    http://storage.googleapis.com/books/ngrams/books/datasetsv2.html

Frith, Simon. 2001. "Pop Music". In Frith, Simon; Straw, Will; Street, John (eds.). *The
    Cambridge Companion to Pop and Rock*. Cambridge: Cambridge University Press. 93-108.

Garcia, Luis-Manuel. 2005. "On and On: Repetition as Process and Pleasure in Electronic
    Dance Music." In Yorgason, Brent (ed.) *Music Theory Online*. 11(4) n.p. Retrieved
    December 1 2015

http://www.mtosmt.org/issues/mto.05.11.4/mto.05.11.4.garcia.html

Gendron, Bernard. 2002. *Between Montmartre and the Mudd Club.* Chicago and London: The
    University of Chicago Press.

Greenbaum, Sidney; Nelson, Gerald. 2009. *An Introduction to English Grammar.* (3rd ed.)
    London and New York: Pearson Education Ltd.

Gustafsson Sendén, Marie. 2013. "Personal Pronouns in Evaluative Communication." PhD
    thesis, University of Stockholm. Retrieved November 5 2014
    http://su.diva-portal.org/smash/record.jsf?pid=diva2%3A678269&dswid=-8118

"Google Books Ngram Viewer", *Google Books.* Google Corporation. Retrieved June 1 2015.
    https://books.google.com/ngrams/graph?content=I%2Cme%2Cmy%2Cmine%2Cmyself&
    year_start=1980&year_end=2008&corpus=16&smoothing=3&share=&direct_url=t1%3B
    %2CI%3B%2Cc0%3B.t1%3B%2Cme%3B%2Cc0%3B.t1%3B%2Cmy%3B%2Cc0%3B.t
    1%3B%2Cmine%3B%2Cc0%3B.t1%3B%2Cmyself%3B%2Cc0
    and:
    https://books.google.com/ngrams/graph?content=we%2Cus%2Cour%2Cours%2Courselve
    s&year_start=1980&year_end=2008&corpus=16&smoothing=3&share=&direct_url=t1%3
    B%2Cwe%3B%2Cc0%3B.t1%3B%2Cus%3B%2Cc0%3B.t1%3B%2Cour%3B%2Cc0%3
    B.t1%3B%2Cours%3B%2Cc0%3B.t1%3B%2Courselves%3B%2Cc0

Holt, Fabian. 2007. *Genre in Popular Music.* Chicago and London: The University of
    Chicago Press.

Jehan, Tristan; DesRoches, David. 2014. "Analyzer Documentation". The Echo Nest
    Corporation. Retrieved January 25 2016
    http://developer.echonest.com/docs/v4/_static/AnalyzeDocumentation.pdf

Kashima, Emiko S.; Kashima, Yoshihisa. 1998. "Culture and Language: The Case of
    Cultural Dimensions and Personal Pronoun Use." *Journal of Cross-Cultural Psychology.*
    29(3) 461-486. Retrieved October 24 2014
    http://jcc.sagepub.com.ezproxy.ub.gu.se/content/29/3/461.full.pdf+html

Kiesling, Scott Fabius. 1997 [2010]. "Power and the Language of Men." In Meyerhoff,
    Miriam; Schleef, Erik (eds.). *The Routledge Sociolinguistics Reader.* London and New
    York: Routledge. 498-513.

Lena, Jennifer C.; Peterson, Richard A. 2008. "Classification as Culture: Types and
    Trajectories of Music Genres". *American Sociological Review.* 7(5) 697-718 Retrieved
    January 29 2016
    http://www.jstor.org.ezproxy.ub.gu.se/stable/pdf/25472554.pdf

Michel, Jean-Baptiste et al. 2010. "Quantitative Analysis of Culture Using Millions of
    Digitized Books." *Science.* 331(6014) 176-182. Retrieved June 1 2015
    http://www.sciencemag.org/content/331/6014/176.full

Morgan, Marcyliena. 2009. *The Real Hiphop.* Durham and London: Duke University Press.

Na, Jinkyung; Choi, Incheol. 2009. "Culture and First Person Pronouns", *Personality and Social Psychology Bulletin.* 35(11) 1492-1499. Retrieved October 24 2014 http://psp.sagepub.com.ezproxy.ub.gu.se/content/35/11/1492.full.pdf+html

Osborn, Brad. 2011. "Understanding Through-Composition in Post-Rock, Math-Metal, and other Post-Millennial Rock Genres" In Yorgason, Brent (ed.) *Music Theory Online* 17(3) n.p. Retrieved December 9 2015 http://www.mtosmt.org/issues/mto.11.17.3/mto.11.17.3.osborn.pdf

Pennebaker, James W., *The Secret Life of Pronouns.* New York: Bloomsbury Press. 2011.

Pennebaker, James W.; Mehl, Matthias R.; Niederhoffer, Kate G. 2003. "Psychological Aspects of Natural Language Use: Our Words, Our Selves." *Annual Review of Psychology.* 54 547-577. Retrieved October 28 2014 http://www.annualreviews.org.ezproxy.ub.gu.se/doi/abs/10.1146/ annurev.psych.54.101601.145041

Pennebaker, James W.; Sivertsen, Borge; Petrie, Keith. 2008. "Things We Said Today: A Linguistic Analysis of the Beatles." *Psychology of Aesthetics, Creativity, and the Arts* 2(4) 197–202, Retrieved November 21 2013 http://homepage.psy.utexas.edu/HomePage/Faculty/Pennebaker/Reprints/Beatles.pdf

Richmond, Ben. 2014. "There Are 800 Musical Genres and Spotify Wants to Find Your Favorite". *Vice Motherboard,* 6 March 2014. Retrieved January 27, 2016 http://motherboard.vice.com/read/there-are-800-musical-genres-and-spotify-wants-to-find-your-favorite

Rings, Steven. 2013. "A Foreign Sound to Your Ear: Bob Dylan Performs 'It's Alright, Ma (I'm Only Bleeding),' 1964–2009" In Yorgason, Brent (ed.) *Music Theory Online* 19(4) n.p. Retrieved December 1 2015 http://www.mtosmt.org/issues/mto.13.19.4/mto.13.19.4.rings.pdf

Rose, Tricia. 2008 *The Hip Hop Wars: What We Talk About When We Talk About Hip Hop-And Why It Matters.* New York: Basic Civitas Books. Retrieved October 20 2014 http://aamr.alexanderstreet.com.ezproxy.ub.gu.se/View/528029

Semin, Gün R. 2007. "Linguistic Markers of Social Distance and Proximity." In Fiedler, Klaus (ed.) *Social Communication.* New York: Psychology Press. 389-408. Retrieved June 16 2015 https://books.google.se/books?hl=sv&lr=&id=zzW5dqN8NiUC&oi=fnd&pg=PA389#v=o nepage&q&f=false

Snoman, Rick. 2004 [2006]. *Dance Music Manual: Tools, Toys, and Techniques.* New York and London: Focal Press.

Tagg, Phillip. 1987. "Analyzing Popular Music: Theory, Method, and Practice". *Popular Music*. 2. 37-67 Retrieved January 29 2016 http://www.jstor.org.ezproxy.ub.gu.se/stable/pdf/852975.pdf?acceptTC=true

Textalyser. Version 1.05. textalyser.net. n.d. Retrieved June 23 2015
    http://textalyser.net/

Tschmuck, Peter. 2006. *Creativity and Innovation in the Music Industry*. Dordrecht: Springer.
    Retrieved January 29 2016
    http://link.springer.com.ezproxy.ub.gu.se/book/10.1007%2F1-4020-4275-2

Twenge, Jean M.; Campbell, Keith; Gentile, Brittany. 2012. "Changes in Pronoun Use in
    American Books and the Rise of Individualism, 1960-2008". *Journal of Cross-Cultural*
    Psychology, 44(3) 406-415, Retrieved November 22 2013
    http://web.natur.cuni.cz/~houdek3/papers/Twenge%20et%20al%202013.pdf

"Undergraduate Enrollment." *The Condition of Education.* National Center for Education
    Statistics, U.S. Department of Education. May 2015. Retrieved June 23 2015
    http://nces.ed.gov/programs/coe/indicator_cha.asp