THE UNIVERSITY OF CHICAGO

Changing Semantics of Gendered Insults in Music Lyrics

By

Zhuxiao (Jasmine) Huang June 2022

A paper submitted in partial fulfillment of the requirements for the Master of Arts degree in the Master of Arts in Computational Social Science

Faculty Advisor: Jon Clindaniel

Preceptor: Pedro Alberto Arroyo

Acknowledgement

I would like to express my deepest appreciation to my thesis advisor for being an amazing mentor. I also had the pleasure of learning from and working with many professors at the University of Chicago, whose passions for teaching and deep expertise never failed to inspire me. Big thanks to my classmates and cohort members who constantly encouraged me to work harder and build my confidence. Lastly, I am extremely grateful for my family, especially my sister and mother, who believed in me since the day 1. Of course, this would not have happened without the best boy ever, my corgi Cashew. Your love and goofiness powered me through all the challenges.

Abstract

Popular culture has been heavily criticized for normalizing gendered insults and perpetuating gender stereotypes. Lyrics, powerful platforms for social change and artistic expressions, are particular salient venues to study the changing meanings of gendered insults and the cultural shifts underneath. Using neural network-based word representations, this article dynamically captures different uses of a gender slur in more than 1.6 million lyric sentences over time among different gender groups. First of its kind in both scale and approach, this analysis tracks the evolutions of the referents and connotations of gendered insults to highlight the progress, struggles, and limitations of the feminist movement in the popular culture space.

1. Introduction

It's no secret that profanity has been a major theme of popular culture - from prime time television to late-night comedy, from popular radio stations to books (Feldman et al., 2017), and from viral internet memes to presidential campaigns (Slatcher et al., 2007). Despite being an antisocial behavior in many contexts, using profanity seems to have long-standing appeals to people of different cultures, genders, social classes, and age groups (Jdetawy, 2019). When not inflicting harm, profanity can be viewed as an expression of individuality. Swearing is linked to personality traits such as extraversion, dominance, assertiveness and narcissism (Fast & Funder, 2008; Holtzman, Vazire, & Mehl, 2010; Schwartz et al., 2013), which are connected to beliefs in individualism (Triandis & Suh, 2002). Furthermore, profane language is often considered to be more authentic, honest and unfiltered (Jay, 2000; Feldman et al, 2017). Speeches that contain profane words invoke stronger emotional responses and have a stronger impact on people than speeches that do not (Jay, Caldwell-Harris, & King, 2008). Moreover, using tabooed words can be fun (Kaye & Sapolsky, 2009), cathartic (Robbins et al., 2011; Vingerhoets, Bylsma, & de Vlam, 2013), thrilling, persuasive (Scherer & Sagarin, 2006), expressive of emotional arousal (Stephens & Zile, 2017), and can mitigate more harmful antisocial behaviors (Jay, 2009).

The popularity of different types of profanity has changed over time. Scholars have noted a rise in the use of sexual and gender-based insults in recent years on social platforms and cultural products (Thelwall, 2008; Muhanovic et al., 2018; Frisby & Behm-Morawitz, 2019). Gendered slurs are derogatory words with gendered referents. These languages often encode hegemonic gender norms and are used to sanction actions that deviate from socially accepted behaviors for both sexes (James, 1998).

However, words change meanings over time (Traugott & Dasher, 2001; Wijaya & Yeniterzi, 2011). Scholars and pop culture consumers alike have noted an increase in the "reappropriation" of gendered insults (Kleinman et al., 2009; Bianchi, 2014). Through reclaiming, oppressed groups use formally pejorative words in non-derogatory manners and purposes (Bianchi, 2014). In Relevance Theory (Wilson & Sperber, 2002) terms, "reappropriation" happens when in-groups echo derogatory uses in ways that disassociate them from the offensive context. That is, women can effectively redefine the meanings of gendered insults by using those words in positive ways as neutral nouns or even compliments among in-groups. From a feminist perspective, women's reappropriation of gendered swear words - previously masculine symbols - challenges the patriarchic structure that invented those words and concepts (James, 1998). However, the reappropriation of gendered slurs might simultaneously risk reinforcing sexism (Kleinman et al., 2009). Kleinman et al. argue that the use of "bitch" for female empowerment is a "false power" and an indication that sexism is so prevalent and normalized in society that people forgot its existence (2009). They warn that using these languages in a positive light is simply a feel-good, deflective solution to gender inequality that does not threaten the status quo.

The study of semantic change takes us to the past and present societies and enables us to observe culture at different points in time (Bloomfield & Newmark, 1963). Cultural corpora - textualized cultural products such as movie scripts, lyrics and social media posts - are natural sites for analyzing these changes. One needs to look no further than popular music lyrics, to understand how the meanings of gendered insults have changed. Music is one of the most consumed cultural artifacts of modern times. In this paper, I look at the ways gendered insults are used in music lyrics between the 1970s to 2010s, with special attention paid to the word "bitch". Using neural network-based word representations, I capture the social harms and benefits of "bitch" as well as

its semantic movements through time to uncover popular attitudes towards gender minorities like women and the LGBTQ+ community as well as their self-representations.

As discussed more in detail in Section 2, gendered slurs are often used to reinforce traditional gender norms, especially those related to idealized forms of sexualities (James, 1998; Jeffreys, 2005; Frisby & Behm-Morawitz, 2019; Felmlee et al., 2020). Mainstreaming and popularization of gendered slurs in cultural channels incite criticisms and applauds alike among researchers (Kleinman et al., 2009; Ashwell, 2016; Gray & Rich, 2019). Studying how the meanings of gender insults changed over time in different lyrical contexts elucidates the objectives, progress, and obstacles of feminism movements in the popular culture space.

2. Literature Review

2.1 Gendered Insults, Swearing and Profanity

It is necessary to define some aforementioned terms here to clarify their contexts in this paper. Profanity encompasses different families of obscene languages such as cultural taboos, swear words, sexual references, objects eliciting disgust, racial slurs, or offensive slang (Mabry, 1974). In this study, I use profanity as the umbrella term for all socially inappropriate languages.

Slurs are sometimes defined as derogatory words directed to human subjects where there exist neutral counterparts (correlates) (Hornsby, 2001; Camp, 2013). That is, a language should have non-offensive ways of describing the same group of people targeted by a slur. One example is a racial slur that is used to refer to a member of a particular demographic group in a demeaning, and hateful way. However, there also exists a neutral correlate that can identify such a group that is descriptive and without prejudice.

Gendered slurs are derogatory words targeting individuals based on their gender or sexual orientations (Bianchi, 2013). However, unlike racial slurs, these gendered insults usually concern behaviors and are evaluative based on social norms. Ashwell (2016) points out that gendered pejoratives sometimes do not have neutral correlates. The neutral counterpart for "slut", for example, is not simply "women". One possible candidate is "a woman who has sex with many people" (Ashwell, 2016). In most cultural contexts, such a person is often strongly frowned upon. Hence, Ashwell (2016) argues that "slut" is not purely descriptive, but has hidden normative judgments behind it – the number of sexual partners that would be considered "too many" is a social construct. Additionally, "slut" can be used when the sexual behaviors are not directly observed or known. Ashwell (2016) notes that, more often, women are labeled "sluts" based on assumptions about their behaviors, even if there is no evidence. In this case, a more accurate translation would be "woman who is inappropriately disposed toward sexual relations" (Ashwell, 2016). "Slut" is just one of many gender pejoratives that are both "evaluatively non-neutral", due to disapproval in a social context, and "covertly normative" (Ashwell, 2016), that encodes certain gendered norms. In this study, I use gendered slurs, gendered insults, and sexist slurs interchangeably to describe such terms.

Swear and taboo words are emotionally charged language that is offensive to listeners (Jay, 2009). I use swear words to describe offensive languages that do not have gendered, sexual, or racial referents. This includes generic curse words like "damn", "shit" and "fuck".

2.2. Uses of Gendered Insults for Men and Women

Gendered insults reveal different evaluative standards for male and female referents (James, 1998). James (1998) uses participant observations and survey methods to collect and analyze 549

gendered insults uttered by both male and female University of Toronto students. He finds that, on the one hand, 5 out of 7 categories of primarily female-referential terms are related to sexuality - sexual attractiveness, availability, and loyalty. On the other hand, 5 of the 7 male-referential categories involved attacking the referents as incompetent, either in character or in mental or physical abilities. These discrepancies suggest that men are evaluated based on being capable at all times, while women are judged based on confirming heterosexual males' desires and needs (James, 1998).

Furthermore, James (1998) studies the perceived offensiveness of different types of slurs for men and women. He finds that terms that liken men to women are more demeaning than those that liken women to men. In fact, the author shows that the worst things one could call a man, according to men, are insults from the category of "weak/like woman/homosexual". The worst things for women, according to both genders, are words in the "mistreats others" category, such as "bitch", or "promiscuous" category such as "slut" (James, 1998). The differences in intensity of emotional responses to certain categories of insults highlight the different priorities for ideal femininity and masculinity. The fact that "masculine/lesbian" terms are less insulting to women than "weak/like a woman/homosexual" terms are to men suggests that masculine traits are superior to feminine ones, for both sexes, which is consistent with patriarchic values (James, 1998). In other words, by displaying masculine traits, a woman lifts her status and is therefore undesirable (for heterosexual men), as she challenges her male counterparts and the gender roles established by them. By displaying feminine traits, a man lowers his status and is therefore considered inept (by heterosexual men). Gendered insults, by delivering strong evaluative signals to the referred subjects, serve to sanction behaviors and emotions that fall outside the narrow definitions of gender norms.

2.3 Semantic Origin of "Bitch"

"Bitch" is a particularly interesting case of gender insults. Unlike "slut" or "whore", the original referent for "bitch" is a non-human (female dog). In a historical analysis of the semantic development of "bitch", Colllns (1984) states that the first semantic drift of "bitch" was recorded around 1,400 when its operative referent switched from a female canine to a promiscuous woman or prostitute.

He acknowledges that such a metaphor was used to dehumanize women in exchange for highlighting the similarity in their receptiveness to sexual intercourses. Since then, "bitch" has been used as verbs, nouns, adjectives and in other forms ("bitching") and combinations ("bitch ass"). Through its repetitive metaphorical uses, "bitch" accumulates meanings that depart from its original derivation while still carrying some of its features. In modern English, "bitch" can degrade both men and women who deviate from hegemonic gender stereotypes. For example, "bitch" means being a coward or submissive (supposedly feminine traits) when used on men. However, it is used to describe women as being difficult, manipulative, annoying or vicious (masculine traits). "Bitching", as a verb, is synonymous with "complaining" or "whining", both of which invalidate and dismiss the referred subjects as trivial, and possibly coming from the subordinate gender. "Bitch" can be directed to objects that are not human ("life is a bitch" or "the test is a bitch"). However, as Kleinman et al. (2009) pointed out, the expression still implies a female referent because the feminine pejorative is seldom replaced by a masculine one ("the test is a dick") or a gender-neutral one ("the test is a jerk").

Perhaps Colllns's (1984) documentation of the linguistics evolution of "bitch" was written too early to capture the new uses of the words that seem to deviate entirely from the historically

negative connotations. Kleinman et al. (2009) explain the ways feminine subjects (usually women or gay men) greet each other with a friendly "What's up, bitches". Moreover, they note that people are using "bitch" and "feminist" interchangeably in many contexts. For example, "bitch" can describe a woman (or any feminine subject) whose actions fall outside the traditional gender performances and therefore it is equivalent to a "feminist" who challenges the hegemonic gender stereotypes. Additionally, calling oneself "bitch" has a sense of "cool toughness" without the loss of sexiness (Kleinman et al., 2009). This feature can be particularly attractive to young feminists who feel empowered by sexualizing masculine traits like physical toughness and assertiveness.

The polysemy and metaphorical power of "bitch" motivate a more detailed understanding of its semantic shifts in more recent decades. With the prevalence of mass media and digital communications platforms, it is important to study the diversified producers and users of "bitch" who have gained unprecedented power and agencies in the digital age. Is there a positive change in the connotations of "bitch"? If so to what extend and what are the limitations? A temporal analysis of "bitch" might reveal interesting changes in the use cases of the slur, in light of technological and social progress.

2.4 Music, Race and Misogyny

Western popular music has been heavily criticized for perpetuating hegemonic gender stereotypes. A plethora of research has shown that popular music lyrics often depict women as submissive, inferior, nurturing, dependent, and sexual objects (Wilkinson, 1976; Freudiger & Almquist, 1978; Weitzer & Kubrin, 2009; Frisby & Behm-Morawitz, 2019; Gray & Rich, 2019).

Moreover, when female artists sing or perform, they are heavily sexualized and face significant challenges in the male-dominated industry (LeVande, 2008).

Critiques for misogyny in popular music have largely focused on hip hop and rap music (Armstrong, 2001; Weitzer& Kubrin, 2009; Wright & Rubin, 2017; Frisby & Behm-Morawitz, 2019). In a systematic examination of 403 rap songs from 130 platinum albums between 1992 and 2000, Weitzer and Kubrin (2009) conclude that misogyny is a significant theme in hip hop and that the genre was dominated by male artists at the highest level. The authors identified five high-frequency misogynistic themes that include derogatory naming of women, sexual objectification, distrust of women, legitimation of violence against women, and celebration of female promiscuity (2009).

However, the emphasis on the content of hip-hop lyrics distances the (largely) white, wealthy, male entertainment executives and audiences from such scrutiny. Furthermore, it masks the fact that these gendered insults might be equally frequent in white-dominated genres like rock, pop and country music. In fact, a recent content analysis of a random sample of pop music from top charts revealed that the genre depicts violence in lyrics as frequently as in hip hop (Frisby & Behm-Morawitz, 2019). Lastly, previous studies on gendered pejoratives in music lyrics focus on male artists as producers, while rarely unpacking the ways female artists use these vocabularies (Monk-Turner & Sylvertooth, 2008). This study complements these analyses on the gendered/racialized uses of slurs by incorporating diverse genres and artists. One of the challenges in large-scale content analysis of lyrics is the absence of gender information on the artists. I address this using a record-linkage approach to pair each lyrical data with the artist's gender.

2.5 Computational Analysis of Semantic Shifts

Word embeddings are powerful tools to study semantic shifts. Word embeddings represent a word with a vector where words with similar meanings are closer in the vector space. These methods are typically built on the 'distributional hypothesis' which states that word semantics are implicit in co-occurrence relationships (Harris, 1954; Firth, 1957). To build time into the word embedding, researchers face the "alignment problem"- an issue that occurs when embeddings are learned separately for each time slice (Yao et.al, 2018).

To make the time a factor of analysis, some approaches use a two-step method, in which embeddings are computed separately and then aligned across time. One example of this method is used by Hamilton et.al (2016), who use the orthogonal Procrustes method to align the embeddings by imposing orthogonal transformation and minimizing the distance between two adjacent time matrices. However, this method retains only the union of individual vocabulary lists of each time slice. Therefore, it is extremely disadvantageous for imbalanced corpora such as the one used in this study. Indeed, using the orthogonal Procrustes method detailed by Hamilton et.al (2016) generated a very small set of shared vocabularies for my dataset where most gendered slurs are removed due to absence in the older corpus.

Advancements in deep learning methods revolutionize the ways researchers represent and analyze textual data. Neural network-based language models beat traditional language models in almost every natural language processing task. Bidirectional Encoder Representations from Transformers (BERT) (Devlin et al., 2018) is a pre-trained deep representation of text, learned by a multi-layered Transformer architecture that considers both the left and right context of the target word. BERT is pre-trained for two tasks – Masked-Language Modeling (MLM) and Next Sentence Prediction (NSP) on BooksCorpus (800M words) and English Wikipedia (2,500M words).

BERT has a great advantage over models like Word2Vec (Mikolov et.al, 2013), where each word has a fixed representation regardless of the context within which the word appears. On the contrary, BERT produces word representations that are dynamically informed by the words around them. Effectively a word can have multiple different embeddings, depending on the context. Compares to unidirectional language models, BERT shows improved performances in a variety of Natural Language Processing tasks. What makes BERT even more attractive is that the fine-tuning process is convenient, inexpensive, and relatively fast. Without substantial task-specific architecture modifications, researchers can build state-of-the-art models for a variety of natural language understanding tasks (Devlin et al., 2018).

3. Present Study

Unpacking the stereotypes in gendered derogatory words requires researchers to take into account the producers of the messages, as well as the greater cultural shifts. The latter requires a large-scale content analysis throughout the years to evaluate changes over time. While previous quantitative content analysis has examined the representations of women and uses of gendered slurs in music lyrics, their sample sizes are usually small (under 500 songs), and the time frame is typically short (a decade or so), due to the need for human coders to process and classify lyrical references and meanings (Armstrong, 2001; Primack, 2008; Hall et al., 2012; Frisby & Behm-Morawitz, 2019). In this research, I use a scalable solution that uses a state-of-the-art language model to represent gendered insults using word embeddings. The code and data are available on Github[1].

To enrich the scholarship on the semantic shifts of "bitch" in the digital age, I analyze a large set of the lyrical corpus (~237,136 songs) between the 1970s and 2010s that contains rich

[1] https://github.com/Jasmine97Huang/gendered-insults-ma-thesis

information including lyrics, title, years, genres and the genders of the artists. I study lyrics by artists from a variety of genres and genders, in hopes of painting a more balanced picture of misogyny and gendered insults in music.

Hypothesis

I expect to see an increased frequency of gendered slurs in lyrics by all artists, which is consistent with previous studies on mainstreaming and normalization of gender-directed insults (Kleinman et al., 2009) (*H1*). The rise should be most evident among women and non-binary artists after the 2000s, when progress in gender equalities, sexual liberation, technological advancement of information, along with the rise of consumerism and neoliberalism, birth an "empowered and powerful" femininity (Lazar, 2006) (*H2*).

To observe the social change in gendered slurs, I look at the referents of "bitch" over time. A post-feminist perspective predicts that gendered slurs are becoming gender-neutral. I will see that the referents for "bitch" have become more gender-unspecific (*H3*). In this case, gendered insults may be similar to generic swear words (or even greetings) in later years. This hypothesis is supported by Ashwell's argument that the reclamation of gendered slurs requires changes in social conventions so that "there is a neutral term available" (2016). In this case, I can look for neutral counterparts that describe the behaviors of the derogatory word but without targeting specific genders.

Lastly, the uses of the slur among female and male artists are expected to be different, as observed by James (1998). However, I suspect that the discrepancies will close off over time, as usage appropriated by in-group (gender minority) will be eventually extended to out-groups (male) (Bianchi, 2013). As a result, there are fewer gender-exclusive uses over time (*H4*).

4. Data

Table 1 Data Sources

To test the hypotheses, I acquire and merge three sources of lyrical data consisting of 237,136 unique songs from Billboard Top 100 chart, Spotify, and Metro Lyrics between 1950 and 2016 (Table 1). The Billboard dataset is originally gathered by Kevin Schaich and the team, openly available on Github[2]. The Spotify[3] data is acquired from Kaggle's open-source dataset. Lastly, the Metro Lyrics dataset was originally posted on Kaggle but was removed. I acquired a subset of English songs from Github[4] where the authors used the same dataset. The three datasets include shared features such as lyrics, genres, released years and artists' names. Because there are very few 50s and 60s songs, I drop these observations to avoid severe data imbalance. I also remove columns related to audio features that are not relevant to this project. Hence, the scope of this project is songs recorded between 1970 and 2016, totaling 236,085 songs.

Data Source	Number of Songs	Time Frame
Billboard Top 100 Chart	4,024	1950-2015
Kaggle Audio Features and Lyrics of Spotify Songs	15,405	1957-2020
Kaggle Song Genre Classification Dataset (MetroLyrics)	218,209	1967-2016

4.1 Genders

[2] https://github.com/kevinschaich/billboard
[3] https://www.kaggle.com/imuhammad/audio-features-and-lyrics-of-spotify-songs
[4] https://github.com/hiteshyalamanchili/SongGenreClassification

Gender is recorded (in 3 categories: group, male and female) in the Billboard dataset but not the other two. Overall, 232,813 out of 236,085 songs have missing gender information. To determine the genders of artists in the Spotify and Metro Lyrics dataset, I use record linkage with Pronoun & Gender Dataset compiled by Chartmetric[5], an initiative that aims to assess and promote equity in the music industry. The dataset consists of detailed gender categories of 186,070 unique artists. I match the lyrics dataset with artists' names. After record linkage, there are 8,919 artists (111,201 songs) who are unidentifiable in the Pronoun & Gender Dataset and are labeled "unknown" in the "gender_matched" column. The rest of the 124,884 songs have labeled gender categories (see Table 2 for the breakdown). For parts of the analysis, I dropped songs by "unknown" and "group" artists. I then combined all categories other than "male" and "female" (such as "agender", "genderqueer") into a "non-binary" label. This information on artists' genders and genres will allow me to compare trends between styles of music and discovered gendered patterns in the use of sexist slurs.

Gender	Number of Songs
unknown	111201
group	87401
male	28900
female	8091
non-binary	317
androgynous	76
agender	36
genderfluid	32
genderqueer	12
transgender woman	12
transgender	6
gender non-conforming	1

Table 2 Gender Breakdown

4.2 Genres

Genres are recorded in different formats for three datasets. Billboard dataset has a "tag" column that records the categories of the song. For the Spotify dataset, I use the "playlist_genre" column and MetroLyrics data includes the "genre" column. I

Genre	Number of Songs
edm	9,076
latin	869
pop	40,116
r&b	6,851
rap	25,836
rock	104,946

Table 3 Gender Breakdown

[5] https://makemusicequal.chartmetric.com/pronoun-gender-database

identify the top 6 most popular genres that appeared in the three columns and use them as the universal genre categories. See Table 3 for a genre breakdown for all available songs. Note that a song could be classified into more than one genre. The dataset has an overwhelming amount of rock songs compared to others.

4.3 Unit of Time

The time unit for this analysis is 10 years or a decade. As mentioned in the data description, I remove songs in the 50s and 60s to avoid a significant data imbalance. See Table 4 for a breakdown of the dataset by decades. The processed the dataset is still very imbalanced, which might require down sampling to have meaningful comparisons across years.

Decade	Number of Songs
2000s	153,608
2010s	67,727
1990s	8,443
1980s	3,337
1970s	2,970

Table 4 Gender Breakdown

4.4 Slurs

I use the Offensive/Profane Word List[6], compiled by Luis von Ahn, a former computer science professor at Carnegie Mellon University, to get a list of profane words and phrases that include both gender- and non-gender-directed insults. Separately, I formed a list of gendered slurs based on vocabularies used by previous literature and research (Weitzer & Kubrin, 2009; Frisby, 2010; Ashwell, 2016). The resulting words are "bitch", "slut", "sissy", "whore", "ho", "hoe", "skeezer", "shawty", and "shorty". I looked up these words in the Offensive/Profane Word List to get different variations of a gender slur (see Appendix A for the comprehensive list of words/phrases). Because only "bitch" has a significant presence across time slices, my analysis of the temporal changes will only focus on this word.

5. Method

[6] http://www.cs.cmu.edu/~biglou/resources/

5.1 Prevalence Analysis (*H1*, *H2*)

I tokenize lyrics and measure the prevalence of each type of gendered insult, which is calculated as the percentage of tokens that are considered slurs or variations of slurs (see section 4.4 for the list of words) in each decade and each gender. This is calculated by dividing the frequency of a gendered insult by the total number of tokens in the decade/gender. I also calculated the proportion of each slur used in each decade, which is the frequency of a given slur divided by the sum of the frequencies for all slurs.

5.2 Word Embeddings Analysis (*H3*)

I used fine-tuned Bidirectional Encoder Representations from Transformers (BERT) (Devlin et al., 2018) model to generate dynamic word embeddings for gendered-slur "bitch". I fine-tuned BERT with the MLM task using the subset of data that only includes artists labeled "female", "male", and "non-binary".

Because lyrical corpus is very different from the book and Wikipedia corpus that BERT is pre-trained on, fine-tuning is necessary to generate appropriate embeddings. I dropped observations in my dataset where the gender information is "unknown" or "group". This results in a total of 37,483 songs - 28,900 songs by male artists, 8,091 songs by female artists, and 492 songs by non-binary artists. I segmented each song to extract sentences using SpaCy, resulting in 1,647,695 sentences (79% male, 20% female, 1% non-binary). I finetuned a "lyric-bert2" using MLM on top of the BERT base version, which has a smaller set of parameters with minimum loss of performance. BERT uses a mixed strategy of masking tokens during pre-training, where 15% of tokens are randomly selected and 80% of which are replaced with [MASK] tokens, 10% are kept the same and 10% are replaced with random tokens. Devlin et al. (2018) found that masking 100% of the selected tokens has near equivalent performances compares to an 80%-

10%-10% set up in Multi-Genre Natural Language inference and Named Entity Recognition tasks. I follow the 100% masking strategy and take a slightly different approach to select tokens. I mask all insults (both gendered- and non-gendered; from Offensive/Profane Word List) that appear in the corpus and randomly mask 15% of other tokens. I hope that by masking all profane tokens, the model will be tuned to pay attention to these words when generating predictions and adjusting parameters. Overall, 1.35% of tokens are in the profane words list and a total of 13.45 % of tokens are masked.

As suggested by Devlin et al. (2018) in their original paper, 2 to 4 epochs are sufficient for the finetuning process. I trained the model for 2 epochs on a computing cluster with one Nvidia K80 GPU and 45 cores Intel E5-2680v4 2.4GHz CPU for 8 hours. I name it "lyrics-bert2" in future references.

There are a few ways to extract word representations from BERT. Devlin et al. (2018) tested different activations from one or more layers of BERT and use them as inputs for downstream Named Entity Recognition (NER) tasks. They found that the best-performing method concatenates the output of the top four hidden layers. They also found that the second-to-last hidden layer generates comparable results as well. The latter method is additionally proposed by Han Xiao in their open-source project named bert-as-service [7], where they identify the second-to-last-layer as a reasonable sweet spot. I use the second-to-last-layer as the word embeddings for computation efficiency.

Due to the removal of close to 200,000 songs (out of 236,085) that have "unknown" or "group" as gender, only "bitch" and "whore" are found in the dataset retaining three gender classes and only the former has a meaningful presence across time slices. To visualize the differences in the

[7] https://github.com/hanxiao/bert-as-service

ways artists of different genders use slurs across time, I filtered sentences that contain “bitch” or its variations. I randomly sampled from the filtered set to reduce runtime and balance out the genders represented. I then run these subsets of tokens through the finetuned “lyrics-bert2” to retrieve embeddings. I perform dimension reduction of the embeddings and project them into the 2-D space using Principal Component Analysis and t-distributed Stochastic Neighbor Embedding (tSNE).

5.3 Clustering Analysis (*H4*)

I perform K-Means clustering on the reduce-dimensioned embeddings for each decade to localize different uses of “bitch” across time. For decades between 1970 – 1990 when the clusters are distinct and visually separable, the number of cluster K is decided by visual examination of the projection figures. For decades 2000 and 2010, I use silhouette scores in addition to visual patterns and qualitative evaluations to select the optimal number of clusters. Silhouette score is calculated as (b - a) / max(a, b) where a is the average intra-cluster distance and b is the average distance between a point and the nearest point that is not from the same cluster. The higher the score, the farther apart and well-distinguished the clusters are. Therefore the silhouette score, together with qualitative evaluations, can be used to determine the optimal number of clusters for the 2000s and 2010s.

I use the homogeneity score to measure the gender diversity of the slur. Gender diversity is defined as how exclusive a meaning/use of “bitch” is for artists of different genders. The homogeneity score falls between 0 and 1, with 1 indicating that all of the clusters contain only data points from a single class. In this context, a cluster represents one use/meaning of “bitch” and each data point (sentence) is assigned a label which is the gender of the artist. Hence, a homogeneity score of 1 means that each meaning of “bitch” is used exclusively by artists from

one gender group. A homogeneity score of 0 means each meaning of "bitch" is shared by artists of different genders equally. I also calculate the completeness score which measures whether all uses of "bitch" by a given gender category are clustered into one group. The completeness will be 1 only when "bitch" has the exact number of meanings as gender categories and each gender occupies one use case. The homogeneity could be 1 when completeness is not (for example, male artists occupy two uses of "bitch"), while the reverse does not hold. Therefore, I will mainly use a homogeneity score to determine the gender diversity of "bitch".

In addition to quantitative analysis of the diversity of producers of "bitch". I also qualitatively evaluate the most representative uses of the slur for artists of different genders in different decades. The most representative use case is simply the song that has embeddings that are closest to each cluster center. For each song, I pay particular attention to the gender and stereotypes associated with the referent(s) of the slur given the greater context of the song.

6 Results

6.1 Counts of Slurs Overtime

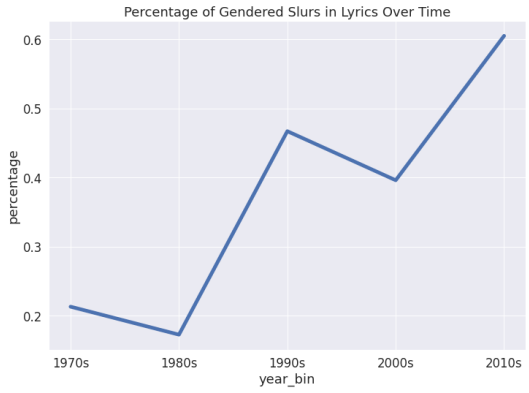


Figure 1 Increasing number of slurs used in lyrics

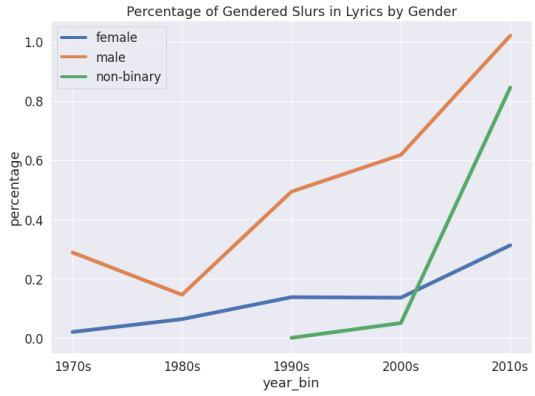


Figure 2 Gender trends in number of slurs used in lyrics

From Figure 1, I see that the frequency of gendered slurs tripled between 1970 and 2010, although the path is not linear. The biggest jump happens in the 1990s and 2010s. In the 2010s,

around 6 profane words are used for every 1000 words in lyrics. This percentage is much higher than general profanity found in books (highest at 0.3% in 2008) (Twenge et al., 2017).

Figure 2 shows the trends in the use of gender slurs for each gender category. The upward trend is evident in all genders. These patterns confirm my hypothesis that there's increasing popularity of gendered insults in lyrics across all artists (*H1*). Overall, male artists dominate the use of gendered slurs and all genders have big leaps between the 2000s and 2010s, supporting my hypothesis that the rise is most drastic post-2000 among non-binary artists. However, contrary to my expectation, the increase is steadier and milder among female artists. Non-binary artists saw a 7 fold increase in the use of gendered insults from the 2000s to the 2010s, making them the second most frequent users merely 0.1 percent behind male users in the 2010s. Due to the lack of data for binary artists in the 70s and 80s, the values are 0.

Overall, the 1980s marks a unique period where the use of gendered slurs decreased among male artists but increased among female artists. The general decline of gendered slurs could be a result of a conservative political and cultural atmosphere brought by the start of the "Reagan era". President Ronald Reagan's election in 1980 brought a dramatic change to the federal government as well as American culture and lifestyle. Several regulations were established against issues around abortion, homosexuality, and the Equal Rights Amendment (Bashevkin, 1994). Under a societal current in opposition to the Feminist Movement, female artists' increased use of gendered slurs can be viewed as a way to address their emotions and beliefs through music.

In addition to normalized frequencies, I also looked at the distributions of 6 gender insults for each gender in each decade. "Bitch" was the most common slur used by all genders. However, I see that there is a decrease in the proportion of "bitch" among male artists over time and an increase among female artists between the 2000s and 2010s. This finding suggests a potential

shift in the conventions and connotation of “bitch” as it gains popularity among female artists and loses appeal for male artists.

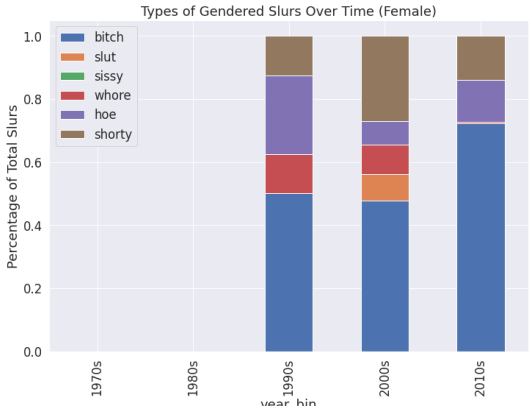


Figure 3 Proportion of slurs used by female artists

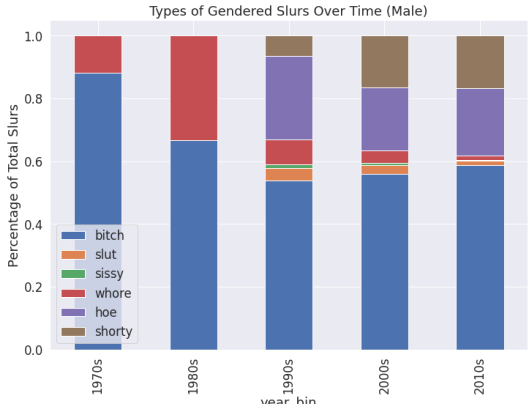


Figure 4 Proportion of slurs used by male artists

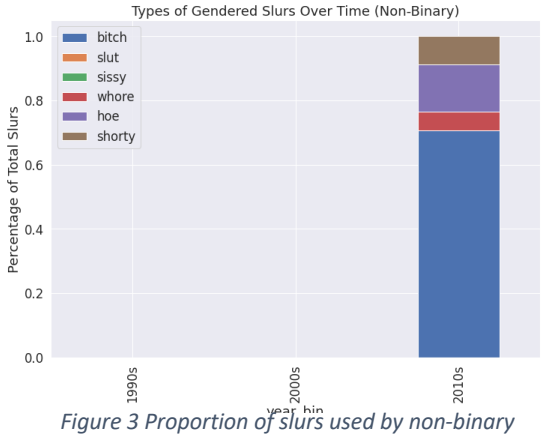


Figure 3 Proportion of slurs used by non-binary

6.2 Changing Uses of “Bitch” Overtime

6.2.1 Evolution of “Bitch” from Afar

Projecting instances of “bitch” over time, I can visualize how the uses of “bitch” evolved. Each point in Figure 4 – 8 represents a word embedding for “bitch” in each decade. The colors of the points represent the genders of the artists who used the word.

Before the 2000s, the clusters of “bitch” are loosely defined and covered more space than the niched uses afterward. “Bitch” diversified the most between the 70s and 80s and shrink sharply from 2000 and onwards.

The 80s has widely scattered dots (Figure 5), contrasting with the tight-knit cluster in the 70s (Figure 4). Note that “bitch” didn’t exist in female artists’ vocabulary until the 90s. When it did appear (Figure 6), however, it is more tight-knit like in the 70s, contrasting with the spread-out formation in the 80s. The clusters contract rapidly in the 2000s (Figure 7) when “bitch” in this decade concentrated tightly around a small region of the 90s. The 2010s saw uses of the slur among non-binary artists, who built on top of the 2000s’ use cases and defined a shape of their own (Figure 8). The macrostructures of clusters over 5 decades suggest that the semantics diversity of “bitch” has changed drastically during the 1980s and 2000s. “Bitch” in the most recent two decades is more specified and niched. In the following sections, I take a closer look at each decade and unpack the micro patterns of the embeddings for “bitch” in each decade.

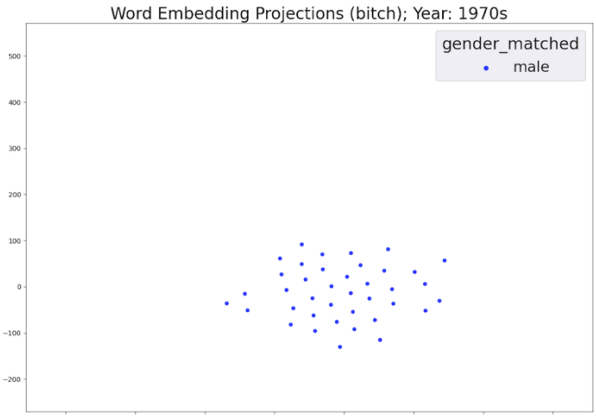


Figure 4 Word Embeddings Projections; 1970s

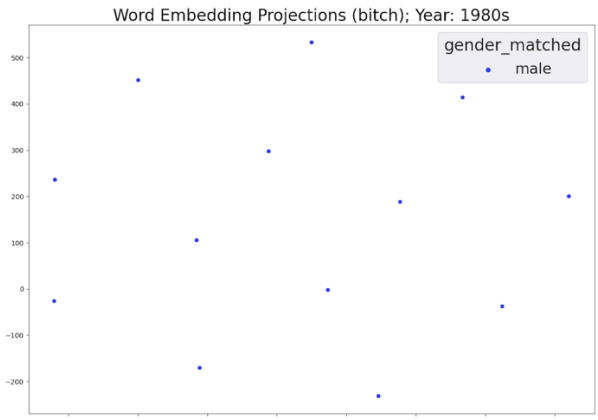


Figure 5 Word Embeddings Projections; 1980s

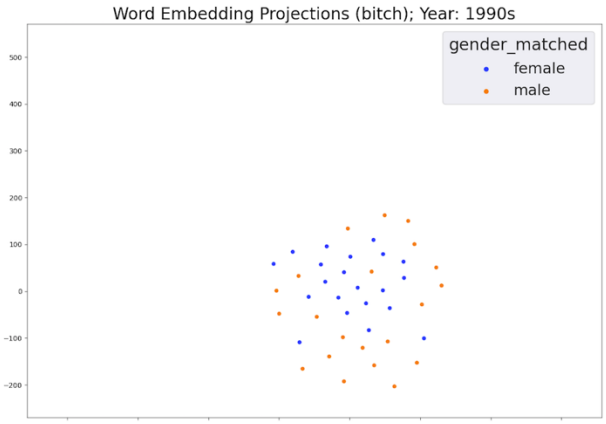


Figure 6 Word Embeddings Projections; 1990s

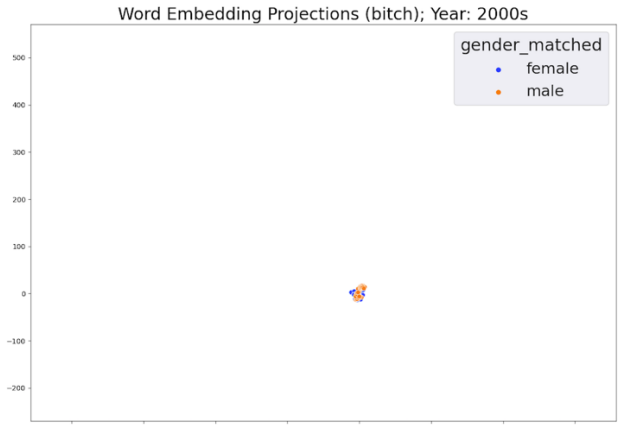


Figure 7 Word Embeddings Projections; 2020s

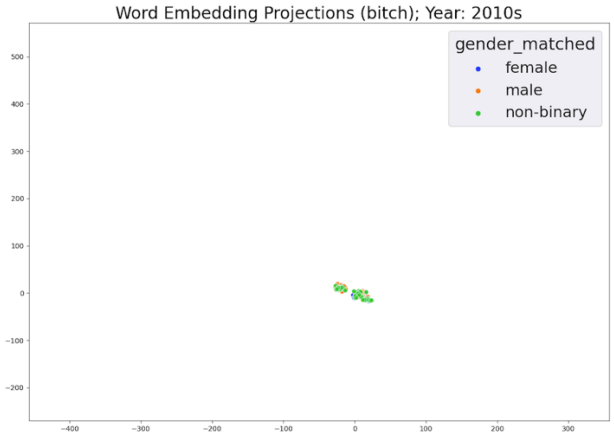


Figure 9 Word Embeddings Projections; 2010s

6.2.2 Evolution of "Bitch" under Microscope

In this section, I perform clustering analysis on lyrics (or a sample of lyrics) that used "bitch" in the dataset for each decade and qualitatively evaluate a representative song in each cluster. Silhouette scores suggest that two clusters are optimal across all decades. This is expected as the gendered slurs might be divided into two general uses- a positive and a negative one. Therefore, I report homogeneity and completeness scores based on two clusters to consistently compare how distributions of genders differ in each use case. However, for qualitative analysis, I further divide the optimal binary clusters into smaller groups to study the granular patterns. Note that I didn't report the scores for the 70s and 80s because only male artists are present in these years, making the homogeneity score 1. The scores for each measure in the 2010s exclude non-binary artists for

a fairer comparison with previous decades. A higher homogeneity score means that there are more gender-exclusive use cases (less gender diversity) of "bitch" in the time frame. The completeness score measures whether artists of the same gender use "bitch" in only one way. I see that both homogeneity and completeness scores steadily decline over decades, suggesting that a more diverse set of artists are using the insults across its opposite semantics. The drop in the homogeneity score is most evident in the 2010s, indicating that male and female artists use positive or negative "bitch" most equally in the recent decade. This supports my hypothesis that there are fewer examples of gender-exclusive uses of the slur (*H4*).

Table 2 Cluster Statistics for Two-Class (Male & Female) Clusters.
A large homogeneity score indicates a use case is more dominated by one gender. A large completeness score measures one gender specializes in one and only one use case.

Decade	Homogeneity Score	Completeness Score	Num. of Clusters	Genders
1990s	0.0747	0.865	2	Male & Female
2000s	0.0749	0.0921	2	Male & Female
2010s	0.0095	0.0108	2	Male & Female

In the following sections, I document the most representative use of "bitch" in each cluster and the referent for "bitch" in that context in Table 5 – 9. I also visualize samples of embeddings (on the left) and the optimal clustering results (on the right) in Figure 10 – 19.

1970s

The 1970s saw 2 opposite meanings of "bitch" (Figure 11). On one hand, "bitch" is used to describe a submissive, conquered figure as in David Bowie's *Moonage Daydream*, where the singer sang "I'll be a rock 'n' rollin' bitch for you". On the other hand, Elton John boasts about being a "bitch 'cause I'm better than you". It seems that a positive connotation associated with "bitch" is present in songs as early as 1974.

In general, I can see that "bitch" in the 1970s establishes a polarized landscape of negativity versus positivity which future artists continued to develop.

Table 5 Summary of the most representative use cases of "bitch" in the 1970s

Cluster	Gender	Artist	Title	Genre	Example Sentence	Referent
0	Male	David Bowie	Moonage Daydream	Rock	"I'll be a rock 'n' rollin' bitch for you"	Submissive Self
1	Male	Elton John	The Bitch Is Back	Rock	"I can bitch, I can bitch 'cause I'm better than you"	Superior, Powerful Self

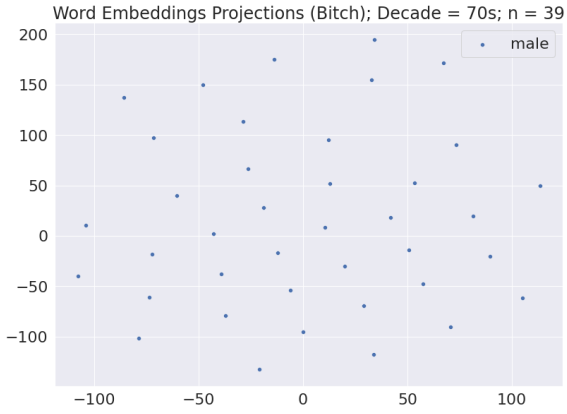


Figure 10 Word Embeddings Projections; 1970s

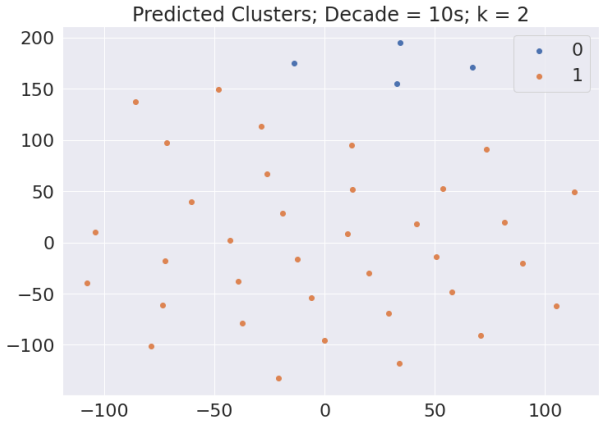


Figure 11 Clusters for the 1970s

1980s

There are only 7 songs (39 sentences) in the 1980s that contain the word "bitch". Previously, I discussed that there was an obvious diversification in the meanings of "bitch" from the 70s to the 80s by examining the cluster evolution from afar. This is confirmed by the use of "bitch" to describe a sexualized, submissive female subject in the 80s which was not typical in the 70s' lyrics.

Table 6 Summary of the most representative use cases of "bitch" in the 1980s

Cluster	Gender	Artist	Title	Genre	Example Sentence	Referent
0	Male	Daryl Hall & John Oates	Gotta Lotta Nerve (Perfect Perfect)	Rock	"Oh no stop bitching indiscretion"	Sexualized, Undeserving Female
1	Male	Scorpions	Rock You Like	Pop	"The bitch is hungry, she needs to tell. So give her inches and feed her well."	Sexualized, Dependent Female

			A Hurricane			

Specifically, in *Gotta Lotta Nerve*, "bitch" takes the form of "bitching" to describe a (female) subject who is out of her place by being overtly sexual ("…why are you getting nasty") and only deserves to be talked down ("Listen to you, your roots will be showing. Better watch what you say and do"). The *Rock You Like A Hurricane* "bitch" is also an extremely sexualized figure, consistently with the theme of the entire song.

All in all, I conclude that a new layer of the hyper-sexualization of women is evident in the "bitch" used in this era, despite an overarching social push for conservative values. Such uses deviate from the submissive or flattering uses of "bitch" in the previous decade.

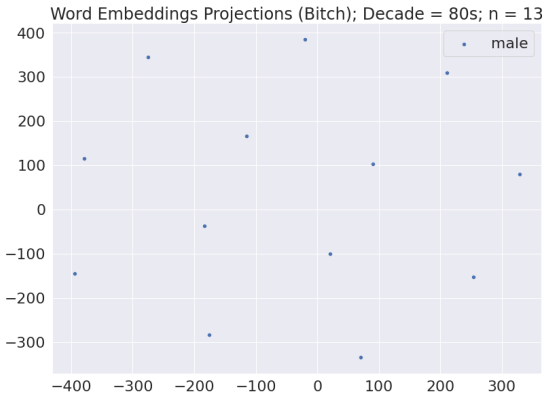


Figure 12 Word Embeddings Projections; 1980s

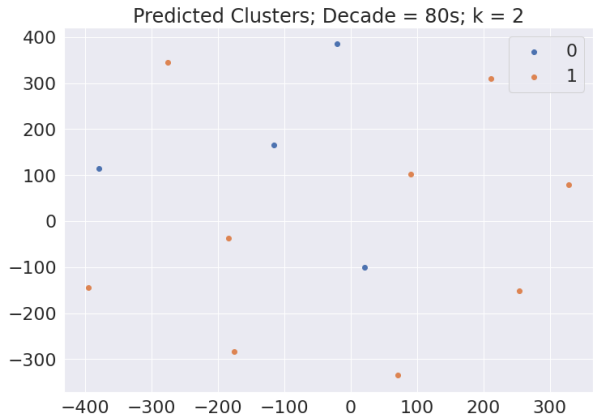


Figure 13 Clusters for 1980s

1990s

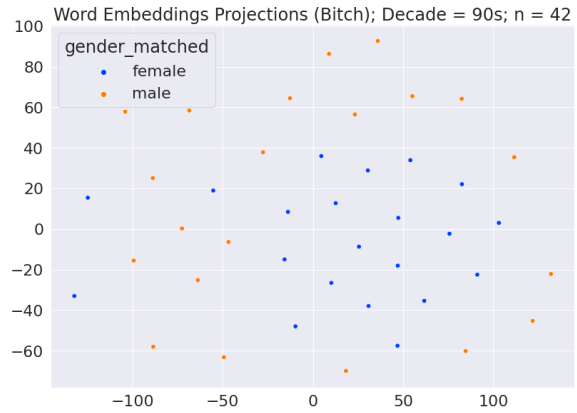


Figure 14 Word Embeddings Projections; 1990s

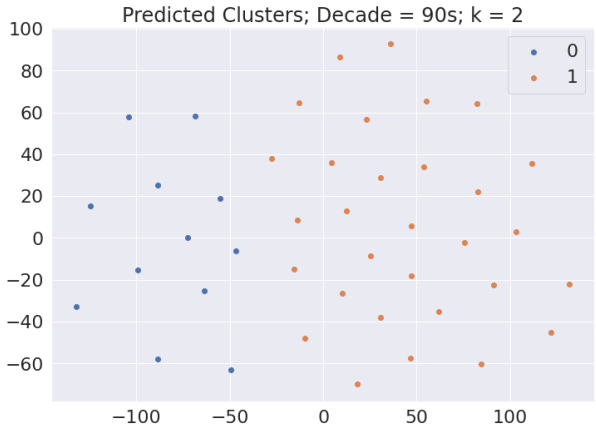


Figure 15 Clusters for 1990s

The 90s marks the entrance of female artists as the producers of "bitch". Two clusters are obvious in the 1990s. Cluster 0 mainly consists of male artists (masculine cluster) while cluster 1 has more female artists (feminine cluster).

"Bitch" in the masculine cluster (cluster 0) are used to describe opponents, enemy or anyone the singer disapproves. Lil Kim uses "bitches" to call out some of the richest, most famous, and powerful female celebrities in a demeaning way that distant herself – the start-from-nothing hustler - from the traditionally classy and feminine figures. This is a common practice in Hip Hop where MCs frequently use racial or gendered slurs to deride the other party. Similarly, Nas also uses "bitch" to refer to his opponent to whom he would not hesitate to impose violence ("box up").

The feminine cluster (cluster 1) in the 90s shows clear positivity and empowerment. Jennifer Lopez's "bad bitch" is a woman who is confident and not afraid of falling in love. 2Pac's "true down-ass bitch" is a day-one friend who's loyal.

To conclude, "bitch" in the 1990s has a clear gender divide where male-dominated clusters use it to demean opponents while female-dominated cluster embraces a positive connotation. My analysis suggests that both male (such as 2Pac) and female (such as Jennifer Lopez) artists in this

decade shook up the negativities around "bitch" and tried to convey new, positive meanings through the slur.

Table 7 Summary of the most representative use cases of "bitch" in the 1990s

Cluster	Gender	Artist	Title	Genre	Example Sentence	Referent
0	Female	Lil' Kim	No Time	Rap	"Zsa Zsa Gabor, Demi Moore, Prince Diane and all them rich bitches"	Female Opponents
0	Male	Nas	The World Is Yours	Rap	"I trip, I box up crazy bitches, aimin' guns in all my baby pictures"	Any Opponents
1	Female	Jennifer Lopez	I'm Real – (Murder Remix)	R & B	"She's a bad, bad bitch"	Powerful Female
1	Male	2Pac	I Ain't Mad At Cha	Rap	"'Cause you're a down ass bitch, and I ain't mad at cha"	Anyone Authentic, Loyal

2000s

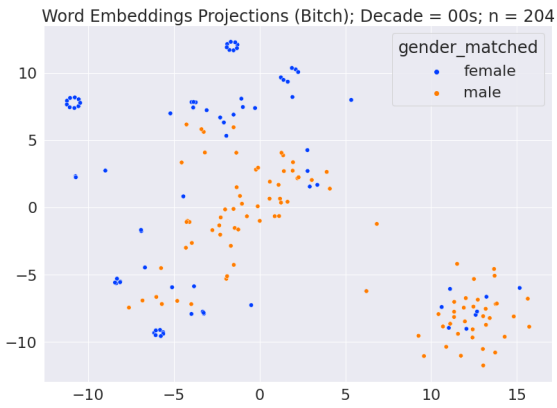


Figure 16 Word Embeddings Projections; 2000s

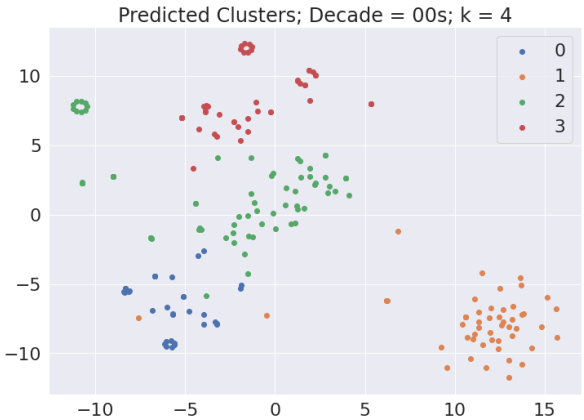


Figure 17 Clusters for the 200s

"Bitch" picks up richer sets of definitions in the 2000s. There are two main clusters in the 2000s according to the silhouette score analysis (see Appendix B for detail).

However, as I want to take a closer look at the substructures within the two main clusters, I ask K-Means algorithms to identify 4 for analysis since it does not degrade the silhouette score much (Appendix B). From Figure 17, I can identify more nuanced patterns of gender divide where female artists formed 2 small groups (clusters 0 & 3) while male artists occupy their clusters as

well (clusters 1 & 2). As shown in the following analysis, cluster 0 is a neutral use case of "bitch" with an unspecified referent. Cluster 1 is the pejorative "bitch", while clusters 2 and 3 are the positive clusters that describe powerful female subjects whose actions challenge the hegemonic gender stereotypes.

Cluster 0 is a primarily feminine cluster. Britney Spears' "It's Britney, bitch" was an instant pop-cultural holy phrase of the 2000s. It marked the artist's transition from a "girl-next-door" to a hyper-sexualized, seductive, and ultimately feminine figure. The "bitch" here does not refer to any specified person. Rather, it is asking for attention from or generally addressing someone (who might question Spears's identity). In this interpretation, one may substitute "bitch" for "y'all" - a neutral correlate (Ashwell. 2016). Such expression is common among women or non-binary people who sometimes use "Hi bitches!" as greetings (Ezzell, 2009). In this sense, "bitch" is no longer a gendered insult due to the presence of a neutral counterpart (Ashwell. 2016).

Cluster 1 is a male-dominated cluster exemplified by Eminem's untrustworthy and purposeless female "bitches" who do not deserve to be taken seriously as partners.

Cluster 2 is particularly interesting as it is a majority masculine cluster, but also includes a distinct small collection of female artists. Cluster 2 is characterized by rapper OutKast[8]'s sexually attractive, gold-digger "bitch" who has a bad attitude. This cluster reinforces the hegemonic stereotypes that women are supposed to be obedient, loyal and passive. A headstrong woman with a big personality and transactional attitude towards sex and relationship is deemed undesirable by men.

[8] "See she's the reason for the word "Bitch" (bitch)
I hope she's speedin' on the way to the club
Tryna hurry up to get to a baller or singer or somebody like that" – Roses, OutKast

Cluster 3 is a feminine cluster led by pop artist Christina Aguilera. In her song *Shut Up,* Aguilera dismissed her opponent who was constantly disrespecting her, by telling them to "shut the fuck up." Aguilera's unwavering and aggressive response to being name-called a "bitch" [9] exemplified an unapologetic attitude towards the negative association of "bitch" with powerful women. This cluster, therefore, represents a self-aware, yet defiant "bitch".

Overall, in the 2000s I can see that male-dominated clusters use "bitch" to describe an undesirable female. Primarily female clusters use "bitch" for self-references in an acceptive and proud way, carrying on the fearless "bad bitch" in the 90s. Most importantly, I find evidence for the popularity of a gender-neutral "bitch" among female artists who adopt the word as a generic reference to a group of people.

Table 8 Summary of the most representative use cases of "bitch" in the 2020s

Cluster	Gender	Artist	Title	Genre	Example Sentence	Referent
0	Female	Britney Spears	Gimme More	Pop	"It's Britney, bitch"	Unspecified
1	Male	Eminem	Superman	Rap	"Bitches they come, they go"	Untrust worthy, Undeserving Female
2	Male	OutKast	Roses	Rap	"Crazy bitch (A bitch's bitch)"	Bad attitude, materialistic Female
3	Female	Christina Aguilera	Shut Up Eminem	Pop	"Call me a bitch whenever you wish Don't give a shit, can you handle it? 'Cause I flip the script, don't seal my lips"	Negative, Unapologizing Female

2010s

[9] Call me a bitch whenever you wish
Don't give a shit, can you handle it?
'Cause I flip the script, don't seal my lips" – Shut Up, Christina Aguilera

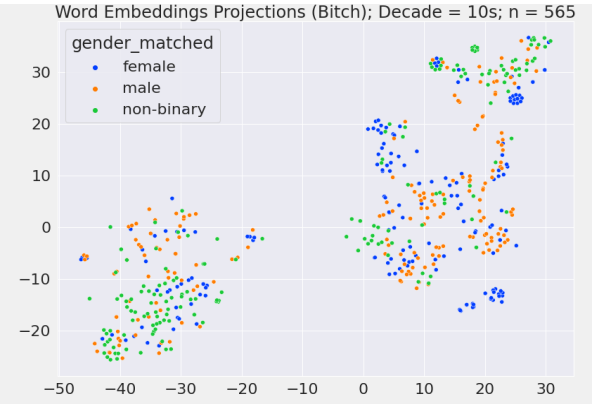


Figure 18 Word Embeddings Projections; 2010s

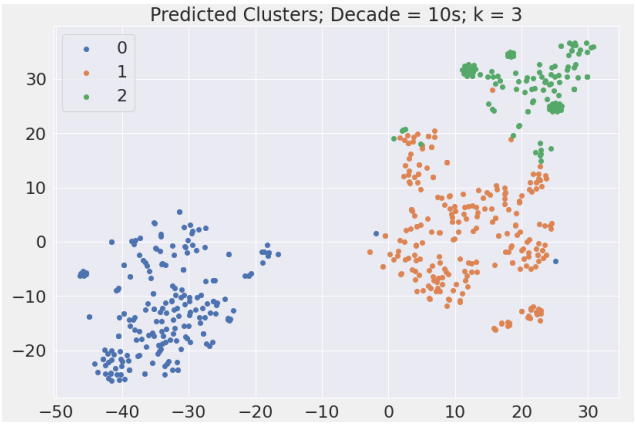


Figure 19 Clusters for the 2010s

Visually inspecting the cluster results for the 2010s, I find the genders seem more mixed across clusters (Figure 18), which is confirmed by a drop in the homogeneity score (Table 4). There are two main clusters as in the previous decade (Figure 19) – the left one (cluster 0) seems to have a higher proportion of non-binary artists while the one on the right is a mix of all. For analysis, I use K = 3 clusters to capture the female and non-binary cluster (cluster 2) at the top right corner.

Cluster 0 uses "bitch" as a negative reference to others. The "bitch" in this cluster is a materialistic woman who gives men attention for their money. Rapper Future, in his song Change Locations with Drake, brags about having so much money that money-grubbing women ("bitches") want to talk to him.

Cluster 1 is the "bad bitch". In her Grammy-nominated song Fancy, Iggy Azalea describes herself as a "bad bitch" who is authentic, attractive and someone who is not afraid of being herself and hurting feelings.

Cluster 2 is the neutral "bitch". Here, Grammy-winning singer Christina Aguilera uses "bitch" as a generic reference to the public by declaring that "I'm not cocky I just love myself, bitch!", which is much similar to "It's Brittny, Bitch!" that I discussed earlier.

To sums up, I find a continuation of the gender-neutral "bitch" into the 2010s, together with the undesirable "bitch" among male artists. I highlight a new "bad bitch" who is unfiltered, true to self and stands against gender norms.

Table 9 Summary of the most representative use cases of "bitch" in the

Cluster	Gender	Artist	Title	Genre	Example Sentence	Referent
0	Male	Drake & Future	Change Locations	Rap	"Then throw out them dollars, them bitches gon' holler"	Materialistic Female
1	Female	Iggy Azalea	Fancy	Pop	"You should want a bad bitch like this"	Authentic, True-to-self Self
2	Female	Christina Aguilera	Vanity	Latin	"I'm not cocky I just love myself, bitch!"	Unspecified

7. Discussion

Cluster homogeneity shows that the users of "bitch" have become more diverse in different uses of the insults. That is, artists of different genders participate in the production of both positive and negative connotations of bitch. This supports my hypothesis that "bitch" is indeed used by a more diverse group of artists to convey both sentiments (*H4*).

Characterizing clusters across decades illuminates the changing semantics of the slur in lyrics. The dichotomy of "bitch" was evident in the 1970s while its meanings evolved greatly in the 1980s, during which the term is synonymous with a hyper-sexualized woman. The meaning was nothing new, however. In fact, it is most similar to the origin of the human referent of "bitch" as a promiscuous woman (Colllns, 1984). In the 1990s, female artists' uses of "bitch" seemingly challenged the misogynistic definitions by associating it with positive and powerful referents, a phenomenon that scholars might call "reappropriation" (Wilson& Sperber, 2002; Bianchi, 2014). This is also a time when both male and female artists participated in the neutralization of the term. In a study in the same decade, James (1998) observes that women are leading the uses of neutralized insults in the university student population studied there, consistent with my finding

that the female-dominated cluster in the 90s demonstrated a strong positive connotation for "bitch" while male-dominated cluster continued to use the word in a demeaning way. The positivity of "bitch" continued to grow in the 2000s while gender-neutral, non-insulting uses of "bitch" were popular among female artists who adopt the word as a generic reference to a person/group. The 2010s saw a wave of "bad bitches" who openly challenged the unjust social pressure and norms placed on them by boasting their desirability and power despite not fitting a traditional definition of submissive femininity. Analyzing the patterns and referents of "bitch" throughout time, I show that referents for "bitch" have become more gender-unspecific (*H3*) and that gender-exclusive uses of "bitch" were seldom discovered (*H4*).

The restoration of "bitch" to promiscuity in the 1980s is particularly interesting. Under an overarching political conservatism influencing various aspects of the intimate and public lives of Americans (Bashevkin, 1994), music functioned as a medium for dissatisfaction and debates. The then underground rap and heavy metal music conveyed powerful messages of frustration over the racialized inequality and confusion about young adulthood and masculinity. The debut of MTV in 1981 brought the visuals of popular music to the screen and opened up new ways of music consumption. Michael Jackson, Prince, and Madonna – the pop's new faces – were both singers and performers. The emphasis on visual presentations resorted female artists to maximizing their sex appeals to reach stardom (Harrison, 2011). While social norms enforced respectability politics, female stars were pressured to be sexualized and to self-sexualize. Self-sexualization happens when a person (woman) internalizes observers' perspectives and views themselves as the object to be looked at and appreciated (Fredrickson & Roberts, 1997). Moreover, socialization theory predicts that self-sexualized individuals would come to view such obsessions with appearance as a personal choice or even natural (Costanzo, 1992). For female

artists since the 80s, self-sexualization could be a strategy to mediate the cognitive dissonance resulting from the conflicts between morality/respectability of sex and the industry's push for sexual objectification. Under this context, the promiscuous "bitch" used by male artists was a tool to sanction liberal sexual expressions so that women stay in their domestic roles as wives and mothers. The "proper women" should assume the patriarchal views that their female counterparts on TVs were "bitches" and hide their desire to follow the suit. This interpretation is consistent with observations that women are equally critical of counter-stereotypic behaviors of women in the workplace and in online discussions (Leskinen et al., 2015; Wilhelm & Joeckel, 2019). Loya et al. (2006) find that women's hostility toward women might reflect a sense of personal inadequacy and dissatisfaction. By degrading other women, women draw lines between themselves- the socially and morally superior group, and the "bitches" – women who display alternative femininity. In this case, I suspect that "proper women" preserve a favorable self-image through accepting the use of insulting language towards women whose behaviors/beliefs differ from the norm.

Such uses of gender insults to disassociate oneself and others are prevalent among female artists who use these words in negative ways on their same-sex counterparts. However, the line is sometimes blurry in lyrical narratives, where "bitch" can be used in immediate contexts to represent different sentiments towards different referents. Ciara's "a lot of these bitches they came and they went" came right after the singer declare that "yeah, I'm that bitch, like it or not".[10] Here, the "these bitches" were the dissociative uses of "bitch" to address the opponents whom Ciara disapproved. However, by calling herself "that bitch", she also openly accepted the negative associations and declared herself as the non-normative woman. This is much similar to

[10]" Yeah I'm that bitch, like it or not. Better get ready cause I'm not gon' stop, stop, stop. Let that beat drop. A lot of these bitches they came and they went." -- Jackie (B.M.F) by Ciara, 2015

the context of "bitch" in pop artist Christina Arguiler's 2008 song *Shut Up*[11]. It seems that for female artists post 2000, although misogynistic history and evaluative functions cannot be entirely removed from "bitch", the word has moved towards embodying an acceptance of a broader definition of femininity.

Indeed, as demonstrated in my analysis, the embracement of the derogatory nature is itself a threat to patriarchy. After all, one cannot penalize someone with a punishment that she does not fear. The "bitch" and "sinner" analogy in rock artist Alanis Morissette's 2008 song *I'm A Bitch I'm A Lover*[12] described her struggle living in the narrow definition of femineity and asked her male partner to "take me as I am" and "you'll have to be a stronger man". Unlike the false power described by Kleinman (2009), Morissette's self-claimed "bitch" is self-aware yet not self-deprecating. Morissette's "bitch" refused to fit into the box of a mother or a lover. Neither did she want to be only a "bitch" or "sinner". Instead, she wanted to be both and desired someone who also refuted patriarchy.

Limitation

There are a few ways future studies on the computational analysis of gendered slurs in cultural corpus could improve.

One of the biggest challenges in this study is to acquire a representative cultural sample. However, to circumvent subjectivity in the definitions of popular music and copyright issues, I used a convenience sample of lyrics available on the internet. This results in an imbalanced sample across gender, genres, and time. For example, my dataset don't have female artists who used "bitch" in the 70s and 80s, nor were non-binary artists represented in decades other than the

[11] Call me a bitch whenever you wish. Don't give a shit, can you handle it?
'Cause I flip the script, don't seal my lips" – Shut Up, Christina Aguilera

[12] "I'm a bitch, I'm a lover, I'm a child, I'm a mother. I'm a sinner, I'm a saint, I do not feel ashamed".

2010s. I suggest future researchers be more strategic about data collection, post-processing and discuss how their unique corpus might impact the conclusions. Genre-specific uses are not fully explored in this paper and I encourage future work to analyze these patterns and how different features (genre, gender, time) interact with each other. Lastly, "bitch" is only one of many gendered slurs. It would be interesting to compare if the patterns uncovered here also generalize to other gendered insults.

8. Conclusion

In this study, I present a new way of representing cultural corpus using a neutral-network-based language model to dynamically capture the various uses of a word. Clustering these word representations, I extract exemplar uses of a word and identify its semantic variations over time. This is similar to dynamic topic modeling techniques (Blei et al., 2003) but it also allows me to group corpus according to the uses of one target word, instead of the over-arching themes of the documents. In this sense, I believe that the methodology introduced here could be a flexible alternative or supplement to traditional topic models, although the validity is not fully examined in this paper. I will leave it for future work.

This study enriches the understanding of gendered slurs in popular culture. I demonstrate that the dichotomous (positive & negative) uses of "bitch" were established as early as the 1970s. Over time, I notice an increase in uses of the slur as generic references, which effectively illegitimate "bitch" as a gendered slur in such contexts. I also explore the reality of using "bitch" as a self-reference among female artists and conclude that these reappropriations, despite not taking away the insults, work as an opposition to patriarchic values. While discussions around the social harms of "bitch" are an important reality check for those dedicated to promoting gender equality, dismissing the power originating from these slurs is not beneficial to the movement. Female

resistance to hegemonic gender stereotypes might not be realistically achieved through radical changes overnight, without these clever reinterpretations and incremental appropriations of gendered slurs that encode gender norms.

9. References

Armstrong, E. G. (2001). Gangsta misogyny: A content analysis of the portrayals of against women in rap music, 1987-1993. *Journal of Criminal Justice and Popular Culture*, *8,* 96-126.

Ashwell, L. (2016). Gendered Slurs. Social Theory and Practice, 42(2), 228–239. http://www.jstor.org/stable/24871341

Bashevkin, S. (1994). Confronting Neo-conservatism: Anglo-American Women's Movements under Thatcher, Reagan and Mulroney. *International Political Science Review*, 15(3), 275–296. https://doi.org/10.1177/019251219401500304

Bianchi, C. (2014). Slurs and appropriation: An echoic account. *Journal of Pragmatics, 66, 35–44*. https://doi.org/10.1016/j.pragma.2014.02.009.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, *3*(Jan), 993-1022.

Bloomfield, M. W., & Newmark, L. (1963). A linguistic introduction to the history of English (p. 364). New York: Knopf.

Camp, E. (2013). Slurring Perspectives. *Analytic Philosophy 54 (2013): 330-49, p. 340*

Colllns, C. A. (1984). Bitch: an example of semantic development and change.

Costanzo, P. R. (1992). External socialization and the development of adaptive individuation and social connection. In D. N. Ruble, P. R. Costanzo, & M. E. Oliveri (Eds.), *The social psychology of mental health*. New York: Guilford.

Devlin, J., Chang, M.W., Lee, K. & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: *Human Language Technologies*, Volume 1 (Long and Short Papers), pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Elliot, A. J., & Devine, P. G. (1994). On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. *Journal of personality and social psychology*, *67*(3), 382.

Fast, L., & Funder, D. (2008, February). Reputational and behavioral correlates of swear word usage. Poster presented at annual meeting of the Society for Personality and Social Psychology, Albuquerque, NM.

Feldman, G., Lian, H., Kosinski, M., & Stillwell, D. (2017). Frankly, I Do Give a Damn: The Relationship Between Profanity and Honesty. *Social psychological and personality science*, *8*(7), 816–826. https://doi.org/10.1177/1948550616681055

Felmlee, D., Inara Rodis, P., & Zhang, A. (2020). Sexist slurs: Reinforcing feminine stereotypes online. *Sex roles*, 83(1), 16-28.

Firth, J.R. (1957). A synopsis of linguistic theory 1930-1955. In *Studies in Linguistic Analysis*, pp. 1-32. Oxford: Philological Society. Reprinted in F.R. Palmer (ed.), *Selected Papers of J.R. Firth 1952-1959*, London: Longman (1968).

Fischer, P., & Greitemeyer, T. (2006). Music and Aggression: The Impact of Sexual-Aggressive Song Lyrics on Aggression-Related Thoughts, Emotions, and Behavior Toward the Same and the Opposite Sex. *Personality and Social Psychology Bulletin*, *32*(9), 1165–1176. https://doi.org/10.1177/0146167206288670

Fredrickson, B. & Roberts, T. (1997). Objectification Theory: Toward Understanding Women's Lived Experiences and Mental Health Risks. *Psychology of Women Quarterly*. 21. 173-206. 10.1111/j.1471-6402.1997.tb00108.x.

Freudiger, P. & Almquist, E. M. (1978). Male and female roles in the lyrics of three genres of contemporary music. *Sex Roles*, 4, 51-65. doi:10.1007/BF00288376

Frisby, C. & Behm-Morawitz, E. (2010). " Undressing the Words: " Prevalence of Profanity, Misogyny, Violence, and Gender Role References in Popular Music.

Frisby, Cynthia. (2010). Sticks “n” stones may break my bones, but words they hurt like hell: Derogatory words in popular songs. *Media Report to Women, 12-18*.

Gray, K., & Rich, L.O. (2019). Evidence of Rape Culture in Modern Music.

Greitemeyer, T., Hollingdale, J., & Traut-Mattausch, E. (2015). Changing the track in music and misogyny: Listening to music with pro-equality lyrics improves attitudes and behavior toward women. *Psychology of Popular Media Culture, 4*(1), 56–67. https://doi-org.proxy.uchicago.edu/10.1037/a0030689

Hall, P.C., West, J. H., & Hill, S. (2012). Sexualization in lyrics of popular music from 1959 to 2009: Implications for sexuality educators. *Sexuality & Culture*, *16*(2), 103-117.

Hamilton, W. L., Leskovec, J., & Jurafsky, D. (2016b). Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change. Proceedings of the 54th Annual Meeting of the

Association for Computational Linguistics (Volume 1: Long Papers), Berlin, August 2016, 1489-1501. https://doi.org/10.18653/v1/P16-1141

Harris, Z. (1954). Distributional structure. *Word*, 10(23): 146-162.

Harrison, T. (2011). *Music of the 1980s*. Santa Barbara, Calif: Greenwood.

Holtzman, N. S., Vazire, S., & Mehl, M. R. (2010). Sounds like a narcissist: Behavioral manifestations of narcissism in everyday life. *Journal of Research in Personality,* 44, 478-484.

Hornsby, J. (2001). Meaning and Uselessness: How to Think about Derogatory Words. *Midwest Studies in Philosophy* 25 (2001): 128-41

Hyatt, C.S., Berke, D.S., Miller, J.D. and Zeichner, A. (2017), Do beliefs about gender roles moderate the relationship between exposure to misogynistic song lyrics and men's female-directed aggression?. *Aggr. Behav., 43:* 123-132. https://doi-org.proxy.uchicago.edu/10.1002/ab.21668

James, D. (1998). Gender-linked derogatory terms and their use by women and men. *American Speech*, *73*(4), 399-420.

Jay T. (2000). *Why I curse: A neuro-psycho-social theory of speech*. Amsterdam, the Netherlands: John Benjamins.

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

Jdetawy, L. F. (2019). The nature, types, motives, and functions of swear words: A sociolinguistic analysis. *International Journal of Development Research*, *9*(04), 27048-27058.

Kaye B. K., Sapolsky B. S. (2009). Taboo or not taboo? That is the question: Offensive language on prime-time broadcast and cable programming. *Journal of Broadcasting & Electronic Media*, 53, 22–37.

Kleinman, S., Ezzell, M. B., Frost, A. C. (2009). Reclaiming Critical Analysis: The Social Harms of "Bitch"

Kremin, Lena. (2017). Sexist swearing and slurs: Responses to gender-directed insults. *LingUU. 1*. 18-25.

Kutuzov, A., Øvrelid, L., Szymanski, T., & Velldal, E. (2018). Diachronic word embeddings and semantic shifts: a survey. *arXiv preprint arXiv:1806.03537*.

Leskinen, E. A., Rabelo, V. C., & Cortina, L. M. (2015). Gender stereotyping and harassment: A "catch-22" for women in the workplace. *Psychology, Public Policy, and Law, 21*(2), 192–204. https://doi.org/10.1037/law0000040.

Levande, M. (2008). Women, Pop Music, and Pornography. *Meridians: feminism, race, transnationalism, 8*, 293 - 321.

Loya, B. N., Cowan, G., & Walters, C. (2006). The role of social comparison and body consciousness in women's hostility toward women. Sex Roles, 54(7–8), 575–583. https://doi.org/10.1007/s11199-006-9024-0.

Mabry E. A. (1974). Dimensions of profanity. *Psychological Reports*, 35, 387–391.

Mast, J. F., & McAndrew, F. T. (2011). Violent lyrics in heavy metal music can increase aggression in males. *North American Journal of Psychology, 13*(1), 63–64.

Monk-Turner, E., & Sylvertooth, D. O. (2008). Rap music: Gender difference in derogatory word use. *American Communication Journal*, *10*(4).

Muhanović, M., Babić, N., & Latic, E. (2018). An analysis of gender differences in the use of swear words on Facebook. *Journal of Education and Humanities*, 1(2), Winter 2018, 89-98.

Mikolov, Tomas & Chen, Kai & Corrado, G.s & Dean, Jeffrey. (2013). Efficient Estimation of Word Representations in Vector Space. Proceedings of Workshop at ICLR. 2013.

Primack, B. A., Gold, M. A., Schwarz, E. B., & Dalton, M. A. (2008). Degrading and non-degrading sex in popular music: A content analysis. *Public Health Reports*, *123*(5), 593-600.

Robbins M. L., Focella E. S., Kasle S., López A. M., Weihs K. L., Mehl M. R. (2011). Naturalistically observed swearing, emotional support, and depressive symptoms in women coping with illness. *Health Psychology*, 30, 789–792.

Scherer, C. R., & Sagarin, B. J. (2006). Indecent influence: The positive effects of obscenity on persuasion. *Social Influence*, *1*(2), 138-146.

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., . . .Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. PLoS ONE. doi:10.1371/journal.pone.0073791

Stephens, R., & Zile, A. (2017). Does emotional arousal influence swearing fluency?. *Journal of Psycholinguistic Research*, 46(4), 983-995.

Thelwall, M. (2008). Fk yea I swear: Cursing and gender in MySpace. *Corpora, 3*, 83-107.

Traugott, E. C., & Dasher, R. B. (2001). *Regularity in semantic change* (Vol. 97). Cambridge University Press.

Triandis, H. C., & Suh, E. M. (2002). Cultural influences on personality. Annual review of psychology, 53(1), 133-160.

Trekels, J., & Eggermont, S. (2018). “I can/should look like a media figure”: The association between direct and indirect media exposure and teens’ sexualizing appearance behaviors. *The Journal of Sex Research*, *55*(3), 320-333.

Twenge, J. M., VanLandingham, H., & Keith Campbell, W. (2017). The seven words you can never say on television: Increases in the use of swear words in American books, 1950-2008. Sage Open, 7(3), 2158244017723689.

van Oosten, Johanna & Peter, Jochen & Valkenburg, Patti. (2015). The Influence of Sexual Music Videos on Adolescents' Misogynistic Beliefs: The Role of Video Content, Gender, and Affective Engagement. *Communication Research*. *42*. 10.1177/0093650214565893.

Vingerhoets A. J., Bylsma L. M., de Vlam C. (2013). Swearing: A biopsychosocial perspective. *Psihologijske Teme*, 22, 287–304.

Weitzer, R., Kubrin, C. E. (2009). Misogyny in rap music: A content analysis of prevalence and meanings. *Men and Masculinities*, 12, 3-29. doi:10.1177/1097184X08327696

Wijaya, D. T., & Yeniterzi, R. (2011, October). Understanding semantic change of words over centuries. In *Proceedings of the 2011 international workshop on DETecting and Exploiting Cultural diversiTy on the social web* (pp. 35-40).

Wilhelm, C., & Joeckel, S. (2019). Gendered morality and backlash effects in online discussions: An experimental study on how users respond to hate speech comments against women and sexual minorities. *Sex Roles*, *80*(7), 381-392.

Wilkinson, M. (1976). Romantic Love: The Great Equalizer? Sexism in Popular Music. *The Family Coordinator*, *25*(2), 161–166. https://doi.org/10.2307/582795

Wilson, D., & Sperber, D. (2002) Relevance Theory. *Handbook of Pragmatics*, Blackwell, 2002. ffijn_00000101

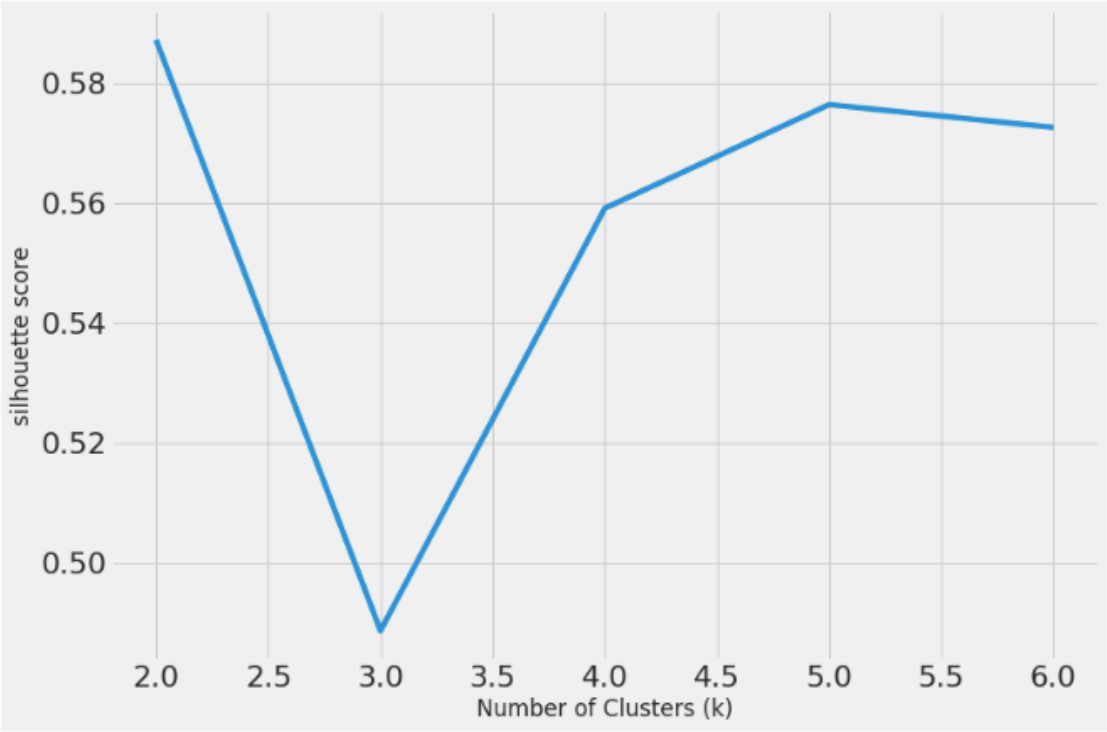
Wright, C. L., & Rubin, M. (2017). “Get lucky!” Sexual content in music lyrics, videos and social media and sexual cognitions and risk among emerging adults in the USA and Australia. *Sex Education, 17*(1), 41–56. https://doi-org.proxy.uchicago.edu/10.1080/14681811.2016.1242402

10. Appendix

A. List of Word Counted as (related-to) Gender Slurs

bitch, bitcher, bitches, bitchez, bitchin, bitching, bitchslap, bitchy, dumbbitch, fuckingbitch, nastybitch, skankbitch, skankybitch, sonofabitch, sonofbitch, easyslut, nastyslut, slut, sluts, slutt, slutting, slutty, slutwear, slutwhore, sissy, asswhore, crackwhore, crack-whore, fuckwhore, gaymuthafuckinwhore, nastywhore, sexwhore, skankwhore, skankywhore, slutwhore, twobitwhore, whore, whorefucker, whorehouse, arsehole, asshole, assholes, asshore, asswhore, bunghole, catholic, catholics, choad, chode, cornhole, crackwhore, crack-whore, cumshot, fuckwhore, gaymuthafuckinwhore, goldenshower, henhouse, ho, hobo, hodgie, hoes, hole, holestuffer, homicide, homo, homobangers, homosexual, honger, honk, honkers, honkey, honky, hook, hooker, hookers, hooters, hore, hork, horn, horney, horniest, horny, horseshit, hosejob, hoser, hostage, hotdamn, hotpussy, hottotrot, moneyshot, nastyho, nastywhore, niggerhole, peehole, peepshow, penthouse, phonesex, pocho, sexhound, sexhouse, sexwhore, shithouse, shoot, shooting, shortfuck, showtime, skankwhore, skankywhore, slutwhore, tuckahoe, twobitwhore, whop, whore, whorefucker, whorehouse, hoes, tuckahoe, shawtypimp, hoe, shorty, shawty,

B. Silhouette Scores 2000s



C. Silhouette Scores 2010s