UNIVERSITY OF TARTU

DEPARTMENT OF ENGLISH STUDIES

Sentiment analysis of top-5 charting songs in the US and the UK from March 2020-2022

MA thesis

MEHAR ABBAS ALI HABIBI

SUPERVISOR: Assoc. Prof. JANE KLAVAN

TARTU

2023

ABSTRACT

This MA thesis aims to conduct a corpus study by using the technique of sentiment analysis to investigate the positive and negative sentiment or emotional valence of the top-5 charting songs in the US and the UK for the first two years of the pandemic from March 2020 to March 2022. The motivation for the present study is to explore the trends in popular music during times of socioeconomic crises in order to broaden the understanding of how music mirrors the emotional moods of a society. The study seeks to contribute to growing interest in the field of quantitative methods to observe patterns in popular music. By employing qualitative analysis of certain songs from the corpus, the study also seeks to provide further insights into how the emotional valence of popular song lyrics can impact individuals living in turbulent times.

The thesis is divided into the following sections: an introduction defining the key term 'popular music' and giving a brief look into the SentiStrength software which is used for sentiment analysis, the first chapter detailing the technique of sentiment analysis and the algorithm used by the software for analyzing song lyrics as short texts, a thorough review of the literature in an attempt to present a solid foundation for the study based on previous research revolving around sentiment analysis of songs, the second chapter detailing the methodological approach used in the study along with a discussion section that is further divided into sub-sections offering a brief qualitative analysis of a few songs with the highest positive and negative sentiment scores, a conclusion, and appendices that contain all the titles of the 164 unique songs present in the corpus along with their respective positive and negative sentiment scores.

Chapter 1 focuses on sentiment analysis, its definition, categories, and approaches. The chapter also looks into previous research that supports the notion of mixed emotions i.e., the ability to experience both positive and negative sentiments in parallel (Berrios et al 2015). The literature review in the first chapter is followed by a section that highlights the challenges which were faced by the researcher due to the limitations of using digital software to conduct sentiment analysis. Chapter 2 involves a deeper dive into the methodological choices made for the study, including why SentiStrength was chosen for the analysis and how the software's algorithm is designed to be dependent on a predetermined list of 2,489 positive and negative sentiment word stems that have been graded by human coders on a scale of 1 to 5. This chapter also explains the manual and time-consuming process of cleaning the individual song lyrics in the corpus sample in order to ensure consistency of language use across the corpus. The quantitative analysis section presents the statistical results for mean scores, standard deviation, range, and p-value of t-tests to confirm the significance of the results. This section is followed by a discussion to justify the confirmation or rejection of the first and the second hypothesis of the study which are based on positive and negative sentiment scores respectively. Aa qualitative analysis of select songs has been included to provide a more holistic and well-rounded analysis taking into account the context of the song lyrics and the external factors at play.

The conclusion section summarizes the key results of the study and discusses the role of sentiment analysis in investigating the emotional valence of song lyrics extracted from popular music that topped the charts during a certain time period. Overall, this thesis offers valuable insights into popular music as a means of expression of emotions during turbulent times.

Keywords: sentiment analysis, COVID-19 pandemic, popular music, US, UK, Billboard chart, top-5 songs, song lyrics, digital humanities, corpus studies

Table of Contents

LIST OF ABBREVIATIONS

ESH = Environmental Security Hypothesis

NLP = Natural Language Processing

ESM = Evaluative Space Model

INTRODUCTION

Popular music is an important tool for cultural transmission since song lyrics are used as an effective medium to communicate emotions, attitudes, opinions, and stories of people (Pettijohn and Sacco 2009). Song lyrics are a valuable cultural product that reflect changes in the emotions of individuals (DeWall et al 2011). As such, they are an excellent source of information for researchers who wish to study emotional changes occurring in a society, especially during a socioeconomic crisis. For instance, songs were used during the American Civil War as a means of "promoting patriotic fervor" when soldiers carried "collections of lyrics of popular numbers" (Leppert 2016: 29). Similarly, during the HIV/AIDS crisis in Africa, song lyrics were used to "raise awareness" and "erode the walls of silence surrounding AIDS" (Kelbessa 2013: 38). A currently ongoing crisis that has had immense emotional impact on individuals is the global spread of the COVID-19 pandemic. The pandemic resulted in many psychological repercussions, including a rise in mental illnesses, stress, anxiety, depression, breakdown of relationships, isolation, loneliness, etc. (Pedrosa et al 2020; Mucci et al 2020). The proposed study aims to investigate the sentiment of songs topping charts in two English-speaking countries, the US and the UK, during the first two years of the pandemic from March 2020 to March 2022. 11th March 2020 has been chosen as the starting date since this was when the World Health Organization officially categorized COVID-19 as a global pandemic (World Health Organization 2020).

The US and the UK have been chosen as the two primary countries for the extraction of data. Firstly, both these countries are English-speaking countries, and it follows from this that songs topping official charts in these countries would be in the English language, which circumvented the need to have translations of song lyrics in the corpus, except for one or two rare instances where lyrics of a song were in another language and their English translations were then obtained from a web-based source. It was necessary to include songs

in the corpus that were originally released in the English language to prevent another layer of mediation which would need to be considered with songs that were translated from one language to another as these translations might not carry over the exact sentiment value that the words had in their language of origin.

Another reason for choosing the US and the UK as primary countries was due to both countries having systematic official charts that explained their algorithms for selecting songs on their websites, with the UK official charts website having a much more detailed explanation than its counterpart US website. The charts and their algorithms are further explained in Section 2.1.1. Lastly, it is worth mentioning that these countries were also selected by previous studies revolving around the sentiment analysis of song lyrics such as Napier and Shamir (2018) who analyzed the sentiment of over 6,000 songs in the US Billboard Hot 100 chart spanning from 1951 to 2016, Krause and North (2019; 2020) who looked at song lyrics from the top 5 songs in the weekly singles charts of the United Kingdom between March 1960 and December 2015, North et al (2018) who also looked at the UK's weekly top 5 singles sales charts from 1962 to 2011, and lastly, Putter et al (2021) who investigated sentiment in the top-5 songs in weekly singles charts in the United States and the United Kingdom from January 1999 to August 2020. The latter research study's methodology is closely replicated by the present study and, hence, choosing the US and the UK aligned with the research aims of the present study.

For the purposes of this study, the phrase "popular music" refers to songs that have wide appeal and have topped official charts in the US and the UK. This phrase does not refer only to the genre of pop music as the official charts from which the corpus data was collected included songs from a variety of different genres, such as rap, rock, classics, R&B, etc. (Brand et al 2019; Putter et al 2021).

SentiStrength 2.3 (2017 version) was used to calculate the sentiment score of songs in the corpus. The software was created by Professor Mike Thelwall and is available for free for academic use. The software simultaneously classifies positive and negative sentiment on a scale of 1/-1 (neutral) to 5/-5 (strongest) based on a sentiment dictionary that contains a list of 2,489 positive and negative sentiment-bearing word stems or "direct affective words" (Thelwall et al 2012: 165). These words have been graded by human coders. The software and its application in the present study is explained in detail in Section 2.2.

The present study seeks to answer the following research question:

How does the sentiment of songs released in the first year of the pandemic (March 2020-March 2021) resemble or differ from the second year (March 2021-March 2022)?

Hardt-Mautner (1995) urges researchers to conduct background research and develop hypotheses before analyzing the corpus to avoid looking at the data from a naive perspective. A review of previous research studies revealed a trend towards increasingly negative song lyrics over time (Putter et al 2021) and it can be assumed that the advent of the COVID-19 pandemic in March 2020 further contributed to a spike in the negative sentiment score of popular songs. Based on this, the proposed study tests the following two hypotheses:

Hypothesis 1: The songs in the second year of the pandemic will have a lower positive sentiment score compared to the first year.

Hypothesis 2: The songs in the second year of the pandemic will have a higher negative sentiment score compared to the first year.

It is important to have two separate hypotheses concerning the positive and negative sentiment of songs since both these opposing emotions can be felt simultaneously, and thus, are measured by the SentiStrength software in parallel to each other. This notion of humans

being able to experience the simultaneity of contrasting emotions, or mixed emotions (Norris et al 2010; Berrios et al 2015), is explained in detail in Section 1.2. The first chapter of the thesis includes an overview of the technique of sentiment analysis and a description of its categories and approaches. There is also a section on the value of song lyrics and how they are a fundamental part of music, which is one of the oldest mediums of sociocultural communication, thereby elevating song lyrics to the position of vital cultural products that can be studied to observe changes in the emotions of individuals and in the psychological features of a culture or a society over time (DeWall et al 2011). Moreover, the first chapter presents a literature review discussing previous studies that used the technique of sentiment analysis and/or built a large corpus of songs spanning several years or decades to study emotional patterns in music over time. The chapter ends with a discussion of the problems and limitations of using sentiment analysis software for a quantitative corpus-based study of songs. The second chapter begins with a thorough review of the steps that were taken in the present study to build a corpus of songs spanning the initial two years of the pandemic. A section on the SentiStrength software used for analysis of corpus data is added to ensure deeper understanding of how the software works to provide a positive and a negative sentiment score for each line of a song. This section is followed by the quantitative analysis of the corpus data with results displayed in written and tabular format. The last section in the second chapter is the discussion section which aims to interpret the results of the sentiment analysis of the corpus in light of the two hypotheses of the present study.

CHAPTER 1. Sentiment Analysis and Song Lyrics

The chapter begins with a definition of sentiment analysis taken from various researchers to provide a holistic view of the technique. The first section also touches upon the different categories of sentiment and the two main approaches of sentiment analysis. The next section explains how two opposing emotions, positive and negative, can co-occur at the same time, a phenomenon known as mixed emotions. Moving on, there is a section providing an overview of previous research pertaining to sentiment analysis in the domain of song lyrics and the last section discusses the problems and challenges of using sentiment analysis software to analyze the emotional valence of short texts.

1.1. Sentiment Analysis: Definition, Categories, Approaches

Pang and Lee describe sentiment analysis as the "computational treatment of opinion, sentiment, and subjectivity in text" (2008: 5). Other terms that have been used interchangeably with sentiment analysis include opinion mining, emotion artificial intelligence, review mining, subjectivity analysis, appraisal extraction, and affective computing. The latter refers to a body of work where the aim is to enable computers to identify emotions in documents (Picard 1997). The term "sentiment" first appeared in research papers published by Das and Chen (2001) and Tong (2001) where the authors analyzed market sentiment. They used the term 'sentiment' to refer to predictive judgements present inside a text which were tracked and evaluated using the technique of sentiment analysis. In the following years, various authors (Turney 2002; Pang et al 2002; Nasukawa and Yi 2003; Yi et al 2003) used sentiment analysis as an empirical method for natural language processing (NLP) in the field of computational linguistics. The large majority of these papers utilized sentiment analysis to categorize reviews in terms of their polarity or binary classification i.e., either positive or negative.

Pang and Lee (2008) contrast sentiment analysis with traditional text categorization which is based on topics. The authors provide an in-depth explanation of how there are fewer classes in sentiment analysis as compared to topical categorization of texts. The two main types of classification in sentiment analysis are polarity classification or sentiment polarity classification and multi-class categorization. In polarity classification, a text is classified as expressing either an overall positive or negative sentiment, so there are essentially only two groups in this type of classification. In multi-class categorization, an ordinal/numerical classification is used based on a multi-point scale usually ranging from 1 to 5 such as 1 star or 4 stars, etc. so this classification has a total of five groups. On the other hand, in topic-based classification, the number of groups might be anywhere between two to one thousand or more based on complex taxonomies. Thus, Pang and Lee (2008) emphasize that "the regression-like nature of strength of feeling, degree of positivity, and so on seems rather unique to sentiment categorization" (2008: 10).

Medhat et al (2014) describe sentiment analysis as the systematic method of identifying emotions expressed in a document or text. They provide three primary levels of classification of sentiment analysis, namely sentence level, document level, and aspect level. Sentence level sentiment analysis looks at each word in a sentence and then assigns a score for the positive and negative sentiment expressed in the sentence. Document level sentiment analysis evaluates the whole document and classifies it as having positive and/or negative sentiment. While sentence level analysis is more specific than document level, there is no clear distinction between the two since one sentence can also be viewed as a short text or a document. Aspect level sentiment analysis is used when details are needed on every aspect of the text being analyzed. For the purposes of this study, a digital sentiment analysis software is used to perform a sentence level sentiment analysis that allows the software to assign a positive and a negative sentiment score to each line of a song. The positive and

negative score for each line can then be averaged using the Excel ‘Average’ formula to produce a document level positive and negative sentiment score for the whole song. This application of sentiment analysis technique is described by Rae (2020: 6) as a “text mining technique i.e., the process of using specialized software to work through large amounts of digital data”. The sentiment analysis software is dependent on sentiment lexicons or dictionaries for analysis. These lexicons or dictionaries contain a large collection of sentiment words that are present in a particular language and have emotional connotations attached to them. Liu (2015: 6) defines sentiment words or opinion words as words belonging to any language that point to either a “desirable or undesirable state”.

There are three separate categories of sentiment: type, orientation, and intensity. Sentiment type can further be classified into several other types such as linguistic-based, consumer research-based, psychology-based, etc. Sentiment intensity measures the extent to which a sentiment word is positive or negative. The intensity of a particular sentiment word can be increased or decreased by using intensifiers or diminishers respectively. For example, a sentiment word “beautiful” has a positive value and this can be increased by adding the intensifier “very” prior to it. In contrast, the positive sentiment of a word can be reduced by adding a diminisher such as “slightly” before the word (Liu 2015). Sentiment orientation, which is the focus of the present study, analyzes the orientation of sentiment words present in a text, in this case the text is lines of songs that are part of the corpus. Other words used to refer to sentiment orientation are emotional polarity, emotional valence, and semantic orientation. The emotional valence of an opinion or a sentiment word can be categorized in three different ways: neutral, positive, or negative. While the software used for this study can categorize the sentiment orientation of each word into these three categories, it comes with its own set of problems that are further discussed in Section 1.5.

There are two main approaches to conducting sentiment analysis, namely machine learning and lexicon based. The machine learning approach depends on using various techniques such as Neural Networks, Supportive Vector Machines Classifiers, Naïve Bayes Classifier, etc. (Wu et al 2021). The software that analyzes sentiment in texts with the help of machine learning technology makes use of indirect lexical indicators. These lexical indicators are automatically assigned a sentiment value based on context. However, a major issue with this technique is that the lexical words present in a sentiment dictionary might be context-dependent, thus, they cannot be employed for analysis outside of their specific domain. An example of an incorrect sentiment analysis conducted by a machine learning programme is provided by Thelwall et al (2012), who state that such a programme might attach a negative sentiment to the word 'Iraq' based on the context of news stories, hence, the same programme cannot be used when analyzing texts containing this word in another context.

The present thesis utilizes the lexicon based approach where a sentiment lexicon is used to analyze a text. This approach depends on direct sentiment indicators that have been assigned a specific semantic orientation or value by a human rater (Liu 2012). A software using the lexicon-based approach makes use of a sentiment dictionary containing a large database of terms and their corresponding emotional valence in order to classify and rate the sentiment of texts. The software matches the words present in the text to the words in the sentiment dictionary and assigns the equivalent sentiment value to each word.

1.2. Simultaneity of Opposing Positive-Negative Emotions

Sentiment is a complex and multi-faceted notion that has been extensively studied by researchers in the past decade with a special focus on mixed emotions. Mixed emotions are defined as affective experiences characterized by feeling two emotions that have opposite or contrasting emotional valence at the same time, for example, happy-sad or positive-

negative (Larsen et al 2001). The notion that humans are capable of simultaneously experiencing mixed emotions has been contested by researchers who subscribe to the dimensional model of affect which theorizes that positive and negative affect lie on opposite ends of the emotional dimension, and therefore, are mutually exclusive (Russell and Carroll 1999; Watson and Tellegen 1999). Thus, mixed emotions are said to be the result of either measurement errors or reactiveness based on expectations of researchers and/or participants which can come from being confused about the reported emotions (Barrett and Bliss-Moreau 2009).

On the contrary, Cacioppo et al (1999; 2004) explained the Evaluative Space Model or ESM which postulates that the two dimensions of positive and negative affect can form a bivariate space inside of which multiple variations of both positive and negative emotions can be described, including mixed emotions where these two affects are co-activated or experienced simultaneously. The authors used the ESM model to highlight two different mechanisms through which simultaneity of opposing emotions can be experienced. The first mechanism involves the perception of both the positive and negative sentiments pertaining to a single event or stimulus, which results in co-activation of these two opposing emotions. The second mechanism features rapid alternation between the positive and negative stimulus "with sufficient speed to produce a sustained activation of both" (Norris et al 2010: 427). Thus, mixed emotions can be described as a "multifaceted emotional experience" (Berrios et al 2015: 5) involving a combination of different opposing emotions that are experienced simultaneously. According to the authors, mixed emotions are complex and marked by an affective experience where both positive and negative emotions can co-occur together.

Another model that supports the co-existence of positive and negative emotions is provided by Russell (1979) and known as the 'dimensional model of emotion' that has widely dominated research in the field of psychology of emotion. This model postulates that

sentiment can be divided into two axes or dimensions, namely the dimension of arousal or activation and the dimension of valence. A sentiment on the axis of arousal can be categorized into two levels of intensity, either low or high, while a sentiment on the axis of valence can be identified as either positive/pleasant or negative/unpleasant. Fox (2018) has provided an example of individual emotions being set along the dimensions of arousal and valence. For example, an emotion such as 'sadness' is identified on the negative or unpleasant dimension of valence while having low intensity on the arousal or activation dimension. However, an emotion such as 'fear' would be categorized as negatively valenced but having much higher intensity.

Furthermore, Watson and Tellegen (1985) summarized the findings of their research with a basic two-factor model that was slightly different from the conventional unrotated model with the two dimensions of pleasantness/unpleasantness or emotional valence and arousal. The authors' model labelled the two varimax-rotated factors as positive affect and negative affect. These terms have moved beyond their restricted use of defining two opposite moods to phrases that are now considered as two distinct dimensions in the analytical studies of affect. Watson et al (1988) defines positive affect as the extent to which a person can feel active, happy, alert, and enthusiastic. This affect is characterized by high energy, pleasurable engagement, focus, concentration, and pleasant emotions. On the other hand, negative affect is characterized by feelings of distress, unpleasable engagement, anger, disgust, fear, etc.

Numerous studies have developed research instruments and methods to study positive affect and negative affect scales. While some studies have found no correlation between these two dimensions of affect (Wills 1986; Clark & Watson 1989), other studies have found them to be significantly related (Brenner 1975; Diener & Emmons 1984). Watson et al (1988) provide a brief overview of the reasons behind the inconsistencies of the findings of these research instruments before introducing their own development of two

10-item mood scales to fill the gap for a valid and reliable positive and negative affect scale. The results of their study showed that the two scales of positive affect and negative affect were largely uncorrelated, and this finding allowed them to present sentiment as being comprised of two separately measurable components i.e., a positive and a negative component. The approach of looking at positive and negative sentiment as two distinct sentiments that can co-occur simultaneously has been adopted by the present thesis, where mixed emotions are treated as more than simply a sum of the individual emotions involved in an experience, rather the notion of mixed emotions is considered an integral and distinct affective experience.

1.3. The Value of Song Lyrics

West (2019) in his article "Introduction: The challenges of the song lyric" talks about how song lyrics have been mostly ignored in the field of literary studies and stylistics. While there is a large body of research exploring popular music using an anthropological or sociological approach (Bennett et al 2005), most researchers have avoided looking at song lyrics as pieces of text that are an essential part of the multimodal experience of popular music. West states the importance of song lyrics in the words "song lyrics are an integral part of popular music, which itself is a domain of human experience that has an enormous significance" (2019: 4). West goes on to explain that popular music emerged via the medium of revolutionary media somewhere in the beginning of the twentieth century and rose in popularity in the late 1950s when it became easily accessible and available to the masses. This allowed popular music to firmly claim its spot as a significant cultural product cataloguing the experiences of human beings. The "wide and profound appeal" enjoyed by popular music is in large part due to song lyrics which are "the most salient element and … have a crucial role to play in how people experience popular music" (ibid). Song lyrics have value in and of themselves due to the fact that they are a fundamental component of the

multimodality of music and are written to be sung by human beings with the accompaniment of tunes.

The experience popular music provides to its listeners cannot be complete without powerful song lyrics. Lyrics form the vital core of songs as they embed and ground abstract concepts and emotions into words or phrases that come together to form sentimentally charged pieces of texts. Hence, the current thesis views song lyrics as having value of their own and aims to investigate the positive and negative sentiments present in song lyrics. While popular music is indeed a multimodal experience where song lyrics, music, instrumental tunes, etc. all combine and form a complex artifact, the present work focuses only on song lyrics by considering them as the most important element in popular music. This approach allows the present study to treat song lyrics as short pieces of text that contain emotions and opinions which can be mined using a sentiment analysis software. An example of this approach is research conducted by Napier and Shamir (2018) who used automatic sentiment analysis to analyze and quantify the sentiment of songs topping the Billboard Hot 100 chart so they could measure the sentiment of these songs across several years from 1951 to 2016. This extensive study included an analysis of 6,085 songs with lyrics to measure the following sentiments: anger, disgust, fear, conscientiousness, joy, extroversion, agreeableness, sadness, and tentativeness. The results of the study showed that anger, disgust, conscientiousness, and fear steadily increased in songs across the years, while joy had a negative correlation with time. While agreeableness and extroversion remained relatively stable, the lyrics of songs also became sadder and more tentative with the passing years. There have been more attempts to explore the phenomenon of cultural evolution in relation to music and how popular songs have emerged on the scene in recent times. For instance, a review article by Savage (2019) provided the historical background of the term ‘evolution’ and considered whether culture can essentially evolve, specifically in terms of

music and musicology. Macro- and microevolution of music was also considered along with the application of this musical evolution in terms of education and copyright.

1.4. Previous Studies of Sentiment Analysis and Song Lyrics

There has been increasing interest in researching trends in popular music using quantitative methods (Mauch et al 2015; Ravignani et al 2017; Savage 2019; Youngblood 2019). The plethora of music produced in the wake of the COVID-19 pandemic encouraged researchers to investigate the effects of the pandemic on popular music. Stratton (2021) examined parodies released during the first years of the pandemic and described them as songs where the tunes of popular music from the 1970s and 1980s are taken and new lyrics are added that relate to the experience of life during a pandemic. The article encouraged readers to view these pandemic parodies as a reflection of the current socioeconomic and cultural times. The researcher further explored how introduction of new social media platforms like TikTok have changed how music is produced and consumed by the younger generation, in particular, Generation X and how the production of popular songs by a predominantly white male population calls into question the idea that music is made by and available for everyone.

For the purpose of this thesis, it is important to look at the key term ‘popular music’ and how previous research has studied it. Putter et al (2021) carried out two studies to look at song lyrics from popular songs in the top-5 position of charts in the UK and the US from March to August 2020 and March to August 2015-19. The first study considered the presence of references to interpersonal relationships, while the second study analyzed whether socioeconomic events from January 1999 to August 2020 had any influence on songs in top-5 weekly charts in the US. DICTION 7.0 content analysis software was used to analyze the sentiment of lyrics. The results of the study indicated that, during the first six months of COVID-19, song lyrics in the US charts contained significantly lower levels of words

referencing human interest and satisfaction. In addition, a significant positive association between the US Misery Index and the number of negatively valenced words was found, suggesting that song lyrics became more negative in times of hardship and economic turmoil. The study made an important contribution to previously limited research analyzing the relationship between trends in popular music and socioeconomic conditions of the time. Another study used the same DICTION software (North et al 2018) to evaluate the lyrical content of songs ranked in top-5 positions of weekly singles charts in the United Kingdom. The study investigated macroeconomic correlates in view of the environmental security hypothesis, which is explained in more detail in Section 2.4. The research was one of the first to highlight a major limitation of using digital software to analyze the sentiment of song lyrics, namely, that the software cannot detect characteristics such as metaphor, humor, slang words, irony, etc.

A significant amount of previous research has used the method of compiling a large corpus of songs to analyse sentiment in song lyrics. Brand et al (2019) built a corpus of 50 years of pop songs from 1965 to 2015 to investigate emotional valence in song lyrics and found that negatively valenced words increased, while positively valenced words decreased with time. The authors grounded their findings in the model of cultural evolution theory, which focuses on how variations in culture are transmitted from one person to another through the medium of social learning such as imitating or copying someone's habits, method of working, speech style, etc. Cultural evolutionary theorists look at how transmission biases or learning biases play a vital role in people "deciding what, when and from whom to copy" (Brand et al 2019: 2). The authors attempted to quantitively measure trends in popular music to understand what the driving force behind them is. Another study that viewed song lyrics as a tool for cultural transmission was conducted by Laland (2004), who suggested that transmission biases are behind the practice of people copying others who

are seen as successful, prestigious, or simply a part of the majority. Transmission biases fall into ‘model-based biases’ that identify who individuals want to copy from, such as success bias meaning people want to imitate others who are seen as successful. These biases are in contrast with ‘content-based biases’ which focus on the type of information people can easily learn, either due to its importance, memorability, popularity, etc. In their study, Brand et al (2019) found that unbiased transmission had an effect on positive lyrics, while content bias influenced negative lyrics. While the effects were small, the findings suggested that a random copying of song lyrics from popular songs that were topping charts led to an increase in negative words in lyrics over time, along with a general negativity bias as songs with more negative content were more popular in the official charts.

Another corpus-based research study by DeWall (2011) focused on the importance of song lyrics as a cultural product that reflects changes in the emotions of individuals. The authors tested the hypothesis that a change in individualistic traits is reflected in song lyrics by variations in word use. The Linguistic Inquiry Word Count software was used to analyze song lyrics in a corpus of popular songs from 1980 to 2007 and the results of the study revealed that words pertaining to antisocial behavior and negative sentiments increased over time, while the occurrence of words denoting positive emotions reduced in song lyrics. The study contributed to the under-researched area of analyzing song lyrics in relation to psychological emotions.

Furthermore, the notion of investigating whether socioeconomic crises with a huge emotional and psychological impact can influence song lyrics has been predominant in previous research. Pettijohn and Sacco (2009) investigated whether the lyrical content of Billboard No. 1 songs in the United States became more meaningful, expressive, and full of intergroup themes if social and economic conditions were unstable. The corpus of the study included songs from 1995 to 2003 and the General Hard Times measure was used to look at

the socioeconomic climate during these years. The authors used a questionnaire to assess whether the song lyrics were meaningful for respondents, while the Linguistic Inquiry and Word Count software was used to explore the relation between pronouns, future and person references, financial issues, etc. While the study was useful for analysing song lyrics in relation to turbulent socioeconomic times, the individual ratings provided by the respondents of the questionnaire cast doubt on the objectivity of the findings as the lyrics were provided out of context and the ratings were not obtained during the year when the songs were the most popular.

More recently, Varnum et al (2021) looked at whether the quantity of novel song choices affects the simplicity of song lyrics by examining the cross-temporal relationship between several cultural factors, including the availability of resources and an increase in the value of individualism in the US. The authors compiled a corpus of 14,661 songs released in the US from 1958 to 2016 that were listed in the Billboard Hot 100 chart. Various measures of correlation such as Kendall's test were used to look at the relationship between novel song choices and lyrical compressibility of popular songs. The cohesive and highly structured study analysed data from six decades using a range of factors in a very systematic manner. Moreover, Varnum et al published all their data and analysis codes on the Open Science Framework, following the principles of digital humanities to allow the study to be replicable and reproducible by future researchers. The main findings of the study revealed a pattern of popular music in the US becoming simpler over time in relation to song lyrics. This trend of lyrical simplicity was observed during six decades across a variety of genres, leading the researchers to suggest that an increase in lyrical simplicity might be related to a rise in new and unique songs being produced. Although one limitation of the study was its correlational nature, the researchers noticed that "in years when more novel song choices were produced, the average lyrical simplicity of the songs entering U.S. billboard charts was

greater" (Varnum et al 2021: 1). This cross-temporal association held true even after they ensured that other factors including cultural, sociological, and ecological would not influence the findings.

Lastly, the software used for the present thesis, SentiStrength, has been previously used to analyze sentiment and mine for emotional valence in Goodreads reviews (Parksepp 2019). The researcher chose 10 American novels and analysed their reviews to see which books had the highest sentiment and whether more sentiment was present in positive or negative reviews. The reviews were analysed using SentiStrength software and the results found that 5-star reviews had strong instances of both positive and negative sentiment, though positive reviews tended to be much longer than negative ones.

1.5. Limitations of Sentiment Analysis

Sentiment can be expressed more subtly in a sentence, or for the purposes of this study, in a song lyric, making it difficult to identify, categorize, or rate emotional valence and to relate the sentiment in a sentence to specific terms or keywords which express that particular sentiment. Pang and Lee (2008) provide an example of this issue by presenting a quote from Mark Twain who stated “Jane Austen’s books madden me so that I can’t conceal my frenzy from the reader. Every time I read ‘Pride and Prejudice’ I want to dig her up and beat her over the skull with her own shinbone.” The authors explain that while the second sentence in the quote contains an explicitly strong negative sentiment, it is challenging to link it with a specific phrase or keyword in the sentence. Thus, a sentence can have an overall sentiment value even without the presence of any sentiment words inside it. For example, the sentence ‘The fridge consumes too much power’ has an overall negative opinion about the fridge but there are no specific opinion words in the sentence that can be credited with carrying a negative sentiment.

On the other hand, a sentence might have a sentiment word present in it but might not express any sentiment overall. This usually occurs in sentences that are either conditional or interrogative (Liu 2015). For instance, the sentence ‘Does this taste good?’ has a positive sentiment word ‘good’ but the sentence overall does not express a positive opinion, and in fact, is asking for one from the intended recipient of the question. Similarly, a conditional sentence such as ‘If the earphones do not work very well, I will discard them’ does not have either a positive or a negative sentiment, despite the presence of a positive opinion phrase ‘very well’ and a negative opinion word ‘discard’.

Another limitation of using sentiment analysis software is that context is extremely important to sentiment mining or extraction. The software used in this study, SentiStrength, cannot detect context-dependent characteristics such as metaphor, humor, slang words, irony, etc. (North et al 2018; Rae 2020). Liu (2015) has further detailed how sarcasm and the use of metaphorical words in a sentence are an issue in sentiment analysis. The author states that sarcasm, especially, complicates the identification and categorization of words into a sentiment category since a seemingly negative word could be intended to mean something positive and vice versa.

In the present study, the SentiStrength software analyses sentiment in a text based on multi-class classification and this results in “a special type of (ordinal) regression problem because the semantics of each class may not directly correspond to a point on a scale” (Pang & Lee, 2008: 17). The issue with this ordinal rating is that each number in the class ranging from 1 to 5 may correspond to a distinct sentiment word present in the lexicon that is prebuilt or fed into the software. Hence, if the software rates a song lyric as 1 which means neutral then this might be because the positive and negative sentiments in the lyric are equally balanced and cancel each other out or it might be because the software recognized certain

keywords in a line such as ‘average’ which prompted it to classify the lyric into a neutral class.

CHAPTER 2. Analyzing the Sentiment of Songs from Two Years of COVID Era

The chapter begins with a brief overview of the decisions that were taken while compiling the song corpus for the present study and the reasoning behind those choices. The first section looks at how the song corpus was built and the steps that were taken to manually clean the lyrics of each unique song in the corpus to ensure consistency and eliminate redundancy. A detailed description of the workings of the SentiStrength software is provided in the second section followed by a report of the quantitative results that were obtained using formulas in Excel to calculate mean scores, standard deviation, range, and p-value of t-tests. The last section discusses the findings of the study in light of the first and the second hypothesis concerning the positive and negative sentiment score of songs in the corpus. A qualitative analysis of a few songs chosen from the corpus based on their extreme sentiment scores is included to offer an in-depth analysis into the trends of popular music during turbulent times.

2.1. Building the Corpus of Songs

The present study replicates and expands upon research conducted by Putter et al (2021). The authors explored the extent to which lyrics of songs topping charts during the first six months of the COVID-19 pandemic contained references to themes of social connection, interpersonal relationships, loneliness, self-interest, human satisfaction, etc. A thorough analysis of the top-5 charting songs in the official charts of the United States and the United Kingdom revealed that songs during the pandemic had higher negative valence and lower human interest and satisfaction compared to the songs topping charts from 2015-2019. The study was significant in deepening research interest and understanding of the relationship between hard socioeconomic times and trends in popular music. Following the methodological approaches of the study, the present study compiled a synchronic corpus by

adding all songs that reached top-5 positions in weekly singles charts in the UK and the US from March 2020 to March 2022, which provided the researcher with a snapshot of language use in songs during a specific period. The songs were categorized into two main groups – songs that topped the charts from March 2020 to March 2021 were labelled the first year of COVID-19, while songs from March 2021 to March 2022 were classified as the second year of the pandemic. The present thesis focused on the popularity of songs by looking at whether they topped specific charts during a particular time period, which qualified them to be added to the corpus. The dates when the songs were originally written or released was not taken into consideration for this thesis since the study focuses on the public reception of songs measured by when they topped the two specific charts of the US and the UK.

2.1.1. Choosing Songs and Compiling Data

For the United Kingdom, the song list came from the Official Singles Chart Top 100, available at https://www.officialcharts.com/. The charts on this website have been used by the British Broadcasting Corporation (BBC) in their official broadcasts, which lends credibility to the charts and allows them to be "the most widely recognized" in the UK (North et al 2018: 641). As stated on their website, the Top 100 chart is compiled by the Official Charts Company and the list is published exclusively on the website mentioned above. The chart compilation process is extremely complex and uses a sophisticated tracking service for ongoing market research to better represent the popularity of songs and display as accurately as possible a picture of consumer consumption of songs in the UK. The Official Charts Company uses Leamington Spa-based Kantar, a chart compilation contractor, to manage the vast databases containing information of the official sales of songs in order to compile "the fastest and most accurate charts in the world" (Official Charts 2023). The song lists in the chart are based on "official sales of downloads, CD, vinyl, audio streams and video streams" (ibid).

For the United States, songs were added from the Billboard Hot 100 chart found at https://www.billboard.com/. The Billboard website provides a very brief description of each of the charts available on the website, stating that the Hot 100 chart presents the most popular current songs of whichever week is selected in the calendar displayed on their webpage. These songs come from all genres of music and are ranked based on streaming activity by Luminate, which tracks data from various digital music sources, measures radio airplay and impressions of the audience, and compiles sales data of songs. It should be noted that the US Billboard website did not have a separate webpage or section on their website dedicated to explaining how they compile the list of songs for their charts and the process behind collecting and tracking official sales data of popular songs, as opposed to the UK Official Charts website which had several webpages reserved solely for explaining their program and chart compilation process.

From each country's chart, the top 5 songs were recorded in a tabular format in Microsoft Word. Both the US and the UK are English-speaking countries and songs that topped charts in these two countries were all in English, allowing the researcher to forego the need for translation, with a few exceptions. Out of a total of 164 unique songs, there were only two songs that had to be translated from another language into English. The first song was 'Peru' by Fireboy DML which contained a few lyrics in Spanish. The second song was 'Love Nwantiti' by Ckay where the chorus of the song was in Igbo. For both songs, the translations were provided by the website from where the song lyrics were copied i.e., https://genius.com/.

Following the methodological choices of previous research (Barradas & Sakka 2021), each country's corpus of songs was treated as separate to indicate that the origins of a song might not have a significant effect on how music listeners connect to song lyrics and how these lyrics influence their emotions. The authors explained that cross-cultural

differences did not significantly impact the listeners' degree of concern and connection with the song lyrics in the words "the effect of the lyrics was evident for the same set of emotions and mechanisms regardless of origin congruency, suggesting that the effects were not due to familiarity" (2021: 666). Thus, the analysis in the present study focused on comparing the sentiment of songs from the first year of COVID-19 to the second year. Although the aim of this thesis was not to compare the two countries' sentiment scores with each other, and hence, the researcher did not include any a priori hypotheses about sentiment scores being dissimilar across the two countries, the decision to keep the song corpus for each country separate proved fruitful as the results showed that there are, in fact, differences between the scores of the two countries and the discussion section of the thesis further elaborates on the similarities and differences between them.

The lyrics for the songs are freely available and accessible using web-based sources, such as http://www.azlyrics.com/ and https://genius.com/. The former website was used by Krause and North (2020) and Putter et al (2021) as an authentic source for written song lyrics, while the latter website was used by the present study in the rare cases when English translations of a full song or a few lyrics from a particular song were needed. Lyrics for all the songs in the corpus were copied from these two online sources and pasted to .txt documents using Notepad++. Each new song in the corpus was copied into a separate .txt file so it could be individually fed as a short piece of text into the SentiStrength software for analysis.

The first year of the pandemic i.e., March 2020 to 2021 in the US contained 52 weeks and each week had five songs selected from it, bringing up the total to 260 songs. Similarly, the first year in the UK had 52 weeks and five songs from each week meant a total of 260 songs. In the second year of the pandemic from March 2021 to 2022, the US had 53 weeks and five songs from each week resulted in a total of 265 songs, while the UK had 54 weeks

resulting in a total of 270 songs. The US and UK charts calculate a week differently. For example, the US Billboard chart shows the week containing the official starting date of the pandemic as 8th-14th March 2020, while the UK Officials chart shows the same week as 6th-12th March 2020, which accounts for the unequal sample sizes for these two countries. Summing up these numbers, the overall corpus contained a total of 1,055 songs. Out of these, 164 songs were unique from which 90 belonged to the US charts and 74 came from the UK charts, although many songs overlapped and were common to the charts of both countries. Tables containing the names of these songs and their average positive and negative sentiment scores are provided in the appendices.

2.1.2. Manual Cleaning of Song Lyrics

The manual cleaning of the .txt files containing the song lyrics was the most significant and time-consuming process in building and preparing the corpus for analysis. The use of word abbreviations such as ‘you’ to ‘u’, emoticons such as :), and truncated sentences in modern language has made it increasingly difficult for software to detect sentiment in online texts. Mobile phone text language has bled into daily life so quickly and pervasively that language use beyond the arena of digital communication is seen to exhibit many similarities with the informal style of language adopted in texts, messages, and DMs (direct messages) on different social platforms such as Facebook Messenger, Discord, and WhatsApp. The limited character space available for texts has also encouraged people to adopt this shortened and informal style of language in posts on social media platforms like Twitter and Instagram (Grinter & Eldridge 2003, Thurlow 2003). Baron (2003) disagrees with the view that these are instances of poor language use, instead the author views it as a natural response to the advances in technology that have led to an integration of social systems across different digital platforms. No matter the view one takes on these language changes over time, it cannot be denied that these variations make it difficult for sentiment

analysis software to effectively analyze texts with human-level accuracy since the first step in detection of sentiment in texts is speech tagging. Software can correctly tag words in sentences that follow the standard rules of spelling and grammar.

The first issue with incorrect tagging by software is due to the intentional use of non-standard spellings in texts. One example of the deliberate use of non-standard spelling in the present study's corpus can be taken from the song 'Starlight' by Dave where the frequently repeated word 'counting' is spelled as 'countin'. There are also a few instances of common words being spelled completely differently such as the word 'enough' spelled as 'nuff'. Secondly, the majority of the songs in the corpus are not concerned with using the right grammar as the focus is more on conveying the emotions present in the lyrics to the wider audience. In fact, it seems that songwriters often go out of their way to ensure they are writing in the language adopted and used by the masses to enforce the relevance of their songs to the current times. An example from the corpus is the song 'Savage' by Megan Thee Stallion with the following lines:

'I'm the hood Mona Lisa, break a nigga into pieces
Had to X some cheesy niggas out my circle like a pizza
All them lil' ass clothes only fit fake booties
Bitch, that's my trash, you the maid, so you bagged him'

From the example above, it is clear that these song lyrics do not follow the conventional rules of grammar such as the phrase 'out my circle' instead of 'out of my circle' and 'you the maid' instead of 'you are the maid'. Hence, current algorithms of both linguistic and non-linguistic machine learning software have trouble working with modern language usage. For linguistic algorithms, it might be possible to detect a range of common abbreviations that have been recorded due to large usage but human beings have an infinite ability to be creative with language use and the rapid development of newer slang words and

abbreviations means it is impossible to keep software algorithms up-to-date as there will always be novel and unique instance of words in texts which the software will not be familiar with. Similarly, while non-linguistic algorithms rely on detecting sentiment in texts based on the presence of individual words that are included in their sentiment dictionary, these types of software would also perform poorly when faced with creative ways of expressing sentiment using misspelled words, informal slang, curse words, etc (Thelwall et al 2010). In an attempt to circumvent these issues arising from using the SentiStrength software to detect sentiment in songs, all the unique songs in the corpus had to undergo a manual cleaning process.

The steps followed in the process of manually cleaning the song lyrics were similar to the two-step process explained by North et al (2018) and later adopted by Putter et al (2021) in their research. In the North et al study, the authors used two screening processes which all songs had to go through before being fed into the software for the analysis. In the first process, they identified instances of redundancy in the song lyrics, meaning all cases where redundancy had been eliminated by the online websites such as using 'You x3' in place of 'You, you, you' or adding the phrase 'Repeat chorus' or using 'x2' for lyrics that were repeated twice in the song. All these examples of redundancy were marked, and the original lyrics were reinstated with full texts added to the songs in every instance to ensure all text files of song lyrics contained the complete song representing a "verbatim account of the lyrics as recorded" (North et al 2018: 642). The present study followed a similar approach to ensure all examples of redundancy were reinstated in the song lyrics.

A few more steps were added in the current study to allow for further meticulous and thorough cleaning of the corpus. These steps included removing any words from the song lyrics that were neutral and used for sound or melodic effect such as 'Hey, hey' or 'Ooh' or 'Oh' or 'Ayy' or 'skrrt'. These words did not have any emotional valence and made no

difference to the sentiment score of the song, so they were deemed unnecessary, and therefore, removed. To continue in the same vein with reinstating redundancies, any words or phrases inside brackets, which meant they were either verbalized in the background, sung simultaneously by two people or repeated by the same singer, were removed from parentheses and added as a separate line in the text files. Abbreviations present in the song lyrics were expanded such as 'PDA' replaced with the full form 'public displays of affection'. However, some abbreviations were colloquial and not known by the researcher such as 'DWP' in the lyric 'The DWP see a number' in the song 'Seventeen going under' by Sam Fender. Abbreviations such as these would have been ignored by the software and would not have any effect on the sentiment score, so they were allowed to remain in the song lyrics. Moreover, the titles of the songs were considered separate from the song lyrics, and as such, they were not copied and included into the text files. If the lyrics had the names of any singers within the song to identify who was singing, these names were removed from the lyrics, but if the song lyrics themselves contained the names of singers or other people within the song who were mentioned or addressed directly, then these were kept within the lyrics to maintain the sentence structure and the intended meaning of each line of the song.

The second screening process used by North et al (2018) was to ensure a consistent use of language across the corpus. In order to do this, they replaced all contractions with their full forms such as 'It's' changed to 'It is' and 'Aren't' expanded to 'Are not', etc. Following the same approach, the present study also expanded all contractions and corrected the use of misspellings or informal slang. For example, words like 'wanna' and 'goin' were replaced with 'want to' and 'going' respectively. Excess repetition of letters was also removed, such as 'L-l-l-lonely' was changed to 'lonely'. The individual position of songs in the charts was not considered when they were being added to the corpus, the only criterion for song selection was that they had to be in the top 5 position in the weekly charts. There

were constant repetitions of songs in and across the charts and each song was added as a separate entry each time it appeared in the charts, meaning "each weekly appearance of a given song was counted" (Putter et al 2021: 5).

2.2. SentiStrength Software

The SentiStrength software, created by Professor Mike Thelwall from the University of Wolverhampton, United Kingdom, is a "sentiment analysis or opinion mining program" (SentiStrength. n.d.). The software was created to mine for emotions and rate the strength of positive and negative sentiment in short texts. For short and informal social web texts in the English language, the official website of the SentiStrength software confirms its human-level accuracy. The software reports two sentiment strengths: -1 (neutral or not negative) to -5 (extremely negative) and +1 (neutral or not positive) to +5 (extremely positive). The report for these two sentiments is provided by SentiStrength software simultaneously based on the findings of psychological studies which have shown that humans process both positive and negative emotions in parallel i.e., mixed emotions (Berrios et al 2015). The majority of opinion mining software group the polarity of sentiment in texts into three distinct categories: positive, negative, neutral. However, texts usually contain a combination of positive and negative emotions, and in some instances, it is beneficial to detect both types of sentiment simultaneously along with the strength of these sentiments. This is especially the case for the present study where songs are treated as short pieces of text that contain a mix of positive and negative emotions. Hence, the present study employs the use of SentiStrength software which uses "several novel methods to simultaneously extract positive and negative sentiment strength" (Thelwall et al 2010: 2545). The software analyzes texts using a sentiment lexicon with 2,489 positive and negative sentiment word stems and their corresponding strength is graded on a scale of 1 to 5.

SentiStrength software was first used by Thelwall et al (2010) to analyze a random sample of 2,600 MySpace comments. The authors chose MySpace as their source for test data since it was a popular social platform that contained a large number of short informal texts. They used subjective judgments by human coders in order to measure emotions in the texts. The comments were randomly selected to be judged on a 5-point scale for both positive and negative sentiment. The coders were provided with verbal guidelines alongside a booklet containing instructions on how to code the texts. The booklet contained a list of acronyms and emoticons that might be present in the texts. The study used three female coders to grade the sentiment words present in the texts. This was to ensure higher inter-coder agreement so that the consequent system developed from the coding process would be more homogenous and have "a consistent perspective on sentiment in the data" (2010: 2550). The mean score of the results of the three coders was calculated and rounded off to get the sentiment strength estimates. The algorithm that was developed for emotion detection by SentiStrength software used the rating provided by the coders in this pilot study.

The software's algorithm is based on a sentiment word strength list. This list contains 298 positive words and 465 negative words which have been classified as having either a positive or a negative strength value on a scale of 2 to 5. The classification of these words is based on the judgments provided by the human coders during the test study. This default manual categorization of words is further modified using a training algorithm that optimizes the sentiment strength of words. The training algorithm looks at the strengths that were allocated by humans to the words in the predefined list and then assesses whether the word requires an increase or a decrease of strength by 1 to improve the accuracy of the ratings. If the accuracy of the sentiment strength of the word increases by a minimum of 2, then the new strength for the associated word is kept. Moreover, the algorithm also has the option to correct spellings of words so a non-standard spelling such as 'hellooo' would be treated as

the word ‘hello’ and rated accordingly. However, to prevent relying solely on the algorithm, in the present study, the researcher manually corrected these non-standard spellings to ensure consistency across all songs in the corpus.

There were other lists included in the software package such as booster word list which contained words that would increase or decrease the sentiment strength of the subsequent word for both positive and negative valenced words. For example, the booster word ‘very’ would increase the positive strength of the word ‘happy’ from +3 to +4. The booster words that increased positive sentiment of words would increase it by either 1 or 2 such as ‘very’ or ‘extremely’, while booster words that would reduce the positive sentiment of a word and increase its negative sentiment would increase it by 1 such as ‘some’ or ‘little’. Moreover, a negative word list was included that contained words to invert the emotion of the following words, such as ‘very happy’ having a positive sentiment score of +4 and ‘not very happy’ having a negative sentiment of -4.

There was also an emoticon list with associated positive and negative strengths, but this was not required by the present study as no songs in the corpus had any emoticons present in the lyrics. The software also took into consideration the use of punctuation marks and their effect on sentiment strength. For example, any sentence that ended with an exclamation mark was graded with a minimum positive sentiment score of +2, while the use of repeated punctation marks would be seen as having a similar function to booster words. Lastly, negative emotion words that were present in questions were not graded so a question in a song lyric such as ‘are you mad?’ would not be graded with a negative sentiment score despite having a negative sentiment word present in the sentence. All these factors were used by the algorithm to calculate the positive and negative sentiment of each sentence in the text, meaning each lyric in the songs present in the corpus of the study.

2.3. Quantitative Analysis and Results

The first step in the analysis of the corpus was to feed the .txt files of song lyrics for each of the 164 unique songs into the software SentiStrength. The output files provided by the software contained a positive and a negative sentiment score for each line of the song and these files were downloaded to the researcher's personal laptop. It was essential to calculate an average positive and negative sentiment score for each song as some songs had 15 lines while others had over 30. Hence, an overall positive and negative sentiment score for the whole song was calculated using the 'Average' function in Excel and the sentiment score for all songs in the corpus was saved in an Excel sheet. The next step involved replacing the names of the songs with their overall positive or negative sentiment score. To maintain an organized database, several Excel sheets were used containing various sets of the data divided according to time period and emotional valence. The first two sheets corresponded to the first two years of the pandemic, so the sheet labelled 2020 contained two tables for the US 2020 scores from March 2020-2021 and the UK 2020 scores for the same time period. Two screenshots depicting the week of 11 March in the first year of the pandemic for the US 2020 and the UK 2020 tables are provided below.

	A	B	C	D	E	F
1	US (2020)	1	2	3	4	5
2	8 to 14 March	The Box by Roddy Ricch	Life is good by Future	Don't start now by Dua Lipa	Circles by Post Malone	Stupid love by Lady Gaga
3	15 to 21	The Box by Roddy Ricch	Don't start now by Dua Lipa	Life is good by Future	Blinding lights by	Circles by Post Malone
4	22 to 28	The Box by Roddy Ricch	Blinding lights by Weekend	Don't start now by Dua Lipa	Life is good by Future	Circles by Post Malone
5	29 to 4 April	Blinding lights by	The Box by Roddy Ricch	Don't start now by Dua Lipa	Heartless by Weekend	Circles by Post Malone

Figure 1. A screenshot of Excel sheet for US 2020 songs

	H	I	J	K	L	M
1	UK (2020)	1	2	3	4	5
2	6 to 12 Mar	Blinding lights by Weekend	Roses by Saint Jhn	The Box by Roddy Ricch	No time to die by Billie Eilish	Stupid love by Lady Gaga
3	13 to 19	Blinding lights by Weekend	Roses by Saint Jhn	Rain by Aitch	The Box by Roddy Ricch	Don't start now by Dua Lipa
4	20 to 26	Roses by Saint Jhn	Blinding lights by Weekend	Rain by Aitch	The Box by Roddy Ricch	Lonely by Joel Corry
5	27 to 2 Apr	Roses by Saint Jhn	Blinding lights by Weekend	The Box by Roddy Ricch	Lonely by Joel Corry	Physical by Dua Lipa

Figure 2. A screenshot of Excel sheet for UK 2020 songs

The second sheet contained data tables for the second year of the pandemic i.e., from March 2021 to 2022. Both countries' tables were included in the same Excel sheet to allow for efficient replacement of titles of songs with their sentiment scores as most songs overlapped across the two countries charts within the same year. The dataset was extended by copying these two sheets four times and labelling them as follows: 2020+, 2020-, 2021+, 2021-. The positive sign before or after a number indicates positive sentiment and the negative sign is a marker of negative sentiment. Hence, a sheet labelled 2020+ has all the positive sentiment scores for all the songs in the US and the UK for the first year of the pandemic i.e., from March 2020 to March 2021. The 'Find and Replace' function in Excel was used to replace the names of the songs with their positive and negative sentiment scores provided by the software. Two screenshots are shown below which show the first week of the US 2020+ and the UK 2020+ positive sentiment score tables.

	A	B	C	D	E	F
1	US (2020)	1	2	3	4	5
2	8 to 14 March	1.13	1.19	1.09	1.23	2.08
3	15 to 21	1.13	1.09	1.19	1.15	1.23
4	22 to 28	1.13	1.15	1.09	1.19	1.23
5	29 to 4 April	1.15	1.13	1.09	1.11	1.23

Figure 3. A screenshot of Excel sheet for US 2020+ sentiment scores

	H	I	J	K	L	M
1	UK (2020)	1	2	3	4	5
2	6 to 12 Mar	1.15	1.21	1.13	1.36	2.08
3	13 to 19	1.15	1.21	1.11	1.13	1.09
4	20 to 26	1.21	1.15	1.11	1.13	1.16
5	27 to 2 Apr	1.21	1.15	1.13	1.16	1.24

Figure 4. A screenshot of Excel sheet for UK 2020+ sentiment scores

Before the actual analysis of the corpus data using statistical tests, it should be mentioned here that a brief pilot study was carried out during the research proposal stage to test the feasibility of the study. A corpus of songs from the first four weeks of 2020-21 (labelled the first year) and 2021-22 (labelled the second year) beginning from the week of 11 March for both US and UK charts was compiled. The corpus contained 20 songs for each country, meaning a total of 40 songs in the corpus. Out of these, 16 songs were unique while others were repeated in the weekly charts and across the charts for both countries. The results showed that, for the UK, the average positive score was 1.21 for the first year and 1.18 for the second year while, for the US, the average positive score was 1.20 for the first year and 1.30 for the second year. Based simply on raw data of average positive and negative sentiment scores, the first hypothesis of the study might only be applicable to the UK where the song lyrics had a lower positive sentiment score for the second year compared to the first year. Regarding average negative scores, for the UK, the average was 1.39 for the first year and 1.21 for the second year while, for the US, the score was 1.42 for the first year and 1.38 for the second year. Thus, the second hypothesis might not be applicable for either countries'

data as the mean scores indicate that the songs topping official charts had a lower negative sentiment score in the second year compared to the first year. Since t-tests were not conducted during the pilot phase of the study, it could not be confirmed whether the differences in the mean scores were statistically significant, but the results did hint at a deviation from the findings of previous research studies which highlighted an increase in negative valence of song lyrics over time, especially due to turbulent socioeconomic conditions.

Moving on to the analysis of the present study, the ‘Average’ function in Excel was used to calculate the mean scores for the positive and negative sentiment of songs in both the US and the UK charts. Beginning with the US, the mean score for 2020+ was 1.31 and the mean for 2021+ was 1.40. The mean score for 2020- was 1.34 and the mean for 2021- was 1.27. In the case of the United Kingdom, the mean score for 2020+ was 1.29 and for 2021+ was 1.33. The mean score for 2020- was 1.30 and the score for 2021- was 1.34.

Individual songs may vary vastly in terms of their overall positive and negative sentiment scores; hence, it is important to look at standard deviation to be able to assess the validity of the average sentiment scores. To calculate the standard deviation (SD), the ‘STDEVPA’ function in Excel was used. This function finds the standard deviation of an entire population and includes logical values or text. This was important as some songs had a sentiment score of 1.00 which would not have been included in the calculations if another function such as ‘STDEV.S’ was used to calculate the standard deviation. For the United States, the SD score for 2020+ was 0.22 and for 2021+ was the same. For 2020-, the SD score was 0.22 and for 2021- was 0.24. For the United Kingdom, the SD score for 2020+ was 0.22 and for 2021+ was 0.24. In terms of negative sentiment scores, for 2020- the SD score was 0.32 and for 2021- it was 0.35. These results are presented in tabular format below.

Table 1. Mean scores and SD for US and UK

United States	**Mean scores (SD)**	**United Kingdom**	**Mean scores (SD)**
2020+	1.31 (0.22)	2020+	1.29 (0.22)
2021+	1.40 (0.22)	2021+	1.33 (0.24)
2020-	1.34 (0.22)	2020-	1.30 (0.32)
2021-	1.27 (0.24)	2021-	1.34 (0.35)

Lastly, the range for the dataset was calculated using the 'Max' and 'Min' function in Excel to find the maximum and minimum sentiment score and then subtracting the two values from each other to find the range. Looking at the range for the corpus data is important as it allows for a better understanding of the spread of the dataset by noticing the extreme values such as the highest and the lowest scores for both the positive and the negative sentiment. For the United States, in terms of positive sentiment scores for 2020, the maximum score was 2.43 and the minimum was 1.08, resulting in a range value of 1.35. In the next year, i.e., 2021+ the maximum score was 2.18 and the minimum was 1.06 so the range was 1.12. Moving on to the negative sentiment scores, in 2020, the maximum score was 2.16 and the minimum was 1 so the range was 1.16 and in 2021 the scores were exactly the same. Regarding the United Kingdom, the maximum score for 2020+ was 2.13 and the minimum was 1 so the range was 1.13. In 2021+, the maximum score was 2.77 and the minimum was 1 so the range amounted to 1.77. For 2020-, the maximum score was 5 and the minimum score was 1 so the range was 4, and for 2021- the maximum score was 4.8 and the minimum score was 1 so the range came up to 3.8. These results have been displayed in the table below.

Table 2. Max and Min scores and Range for US and UK

United States	**Max scores**	**Min scores**	**Range**	**United Kingdom**	**Max scores**	**Min scores**	**Range**
2020+	2.43	1.08	1.35	2020+	2.13	1.00	1.13
2020-	2.16	1.00	1.16	2020-	5.00	1.00	4.00
2021+	2.18	1.06	1.12	2021+	2.77	1.00	1.77
2021-	2.16	1.00	1.16	2021-	4.80	1.00	3.80

To find out whether the differences in the positive and negative sentiment scores of the first year of the pandemic compared to the second year of the pandemic were statistically significant, a t-test was applied. A T-test is a statistical measure used to determine differences between the proportions or means of two datasets that can be from the same category or can belong to entirely different groups. This type of measure is used to test a hypothesis about a certain population or sample. There are two different types of T-tests, namely one-tailed or one-sample and two-tailed or two-sample T-test. In one-sample T-test, the average score of one dataset is compared with a set or fixed value. On the other hand, in a two-sample T-test, also known as independent T-test, two datasets that are from different groups or populations are compared against each other.

In the present study, an independent two-tailed T-test was applied to compare the first year of the pandemic to the second year. There is also a further sub-division of the two types of T-tests into three categories, namely, paired sample, equal variance, and unequal variance. Paired sample T-tests are used when two datasets are from the same group or population. Equal variance T-test is applied to datasets that have the same sample size or are similar to each other in terms of variance. This test is also known as the pooled T-test. Lastly, the unequal variance T-test is used when the number of samples in the two datasets differ from each other or if there is variance in the two groups. This T-test is also known as Welch’s

test, and it was used in this study to compare the datasets from the first and the second year of the pandemic as both the datasets had an unequal sample size (Barr et al 2013).

A T-test formula was created and applied in Excel on the four different sets of corpus data i.e., US 2020+ compared to US 2021+, US 2020- compared to US 2021-, UK 2020+ compared to UK 2021+, and UK 2020- compared to UK 2021-. The T-test formula applied in the present study was as follows: =T.TEST(array1, array2, tails, type). An example of this formula as applied to the dataset is seen as follows: =T.TEST(A2:A261,B2:B266,2,3). The tails value for each of the four datasets remained fixed at 2, representing two-sample or two-tailed T-test type. Similarly, the type value for all four datasets was fixed at 3 to signify the use of the unequal variance T-test. Firstly, the dataset for US 2020+ was compared with the dataset for US 2021+ and the T-test formula was applied, resulting in a p-value of 8.55e-6 or 0.00000855. This value is lower than 0.05, and therefore, statistically significant. The next dataset to be compared was US 2020- to US 2021- and the resulting p-value was 0.0014 i.e., statistically significant. Regarding the United Kingdom datasets, a T-test applied to the two datasets of UK 2020+ and UK 2021+ resulted in a p-value of 0.05 which is equal to the standard error value of 0.05 and can be termed as statistically significant. Lastly, a comparison of the means of the datasets UK 2020- and UK 2021- using the T-test formula showed a p-value of 0.21 which is greater than 0.05, and hence, statistically non-significant. A brief summary of the p-value obtained from the T-tests applied on the four datasets is given in tabular format below.

Table 3. P-values for US and UK

United States	**P-value**	**United Kingdom**	**P-value**
2020+/2021+	$p < 0.005$	2020+/2021+	$p = 0.05$
2020-/2021-	$p = 0.0014$	2020-/2021-	$p = 0.21$

2.4. Discussion

Since the 1950s, there has been an increasing interest among researchers in studying the patterns and trends of songs and popular music. Research on the psychology of language literature shows that lyrics of songs are an important medium of communication that allow singers/songwriters to connect with their audience and tell their stories in "a manner similar to how people have conversations with each other" (Pettijohn and Sacco 2009). In the past few decades, the tone of these 'conversations' has become increasingly negative over time with "progressively lower levels of…positively valenced emotional content" (Putter et al 2021: 3). Dodds and Danforth (2010) quantified levels of happiness for large-scale texts including songs from 1960s to 1990s and found that happiness in song lyrics displayed a significant downward trend over the past three decades. Similarly, DeWall et al (2011) used the Linguistic Inquiry Word Count software to perform linguistic analyses of word use in the most popular songs in the US over the course of 1980 to 2007. The results of the study revealed an increase in words referencing antisocial behavior and self-focus, while words associated with positive emotions steadily decreased. This is further corroborated by Napier and Shamir (2018) whose findings revealed a "clear trend toward a more negative tone in pop music lyrics" (2018: 173).

In parallel, the study by Brand et al (2019) also corroborated this trend in popular music to have an increase of negatively valenced sentiment and a decrease of positively valenced sentiment over time. Their large dataset containing 160,000 songs spanned decades from 1965 to 2015 and the results of the study showed that the frequency of negative sentiment words in song lyrics increased over time. Putter et al (2021) reviewed research studies that explored the relationship between socioeconomic conditions and popular music. In general, sentiment in song lyrics was seen as both a reflector and a predictor of turbulent

times. For instance, North et al (2018) found that song lyrics which expressed less maturity were more popular during turbulence in the London Stock Exchange market and times of low economic optimism were associated with lyrics that displayed less certainty. In a similar fashion, Qiu et al (2021) found that the high unemployment rate in the US and Germany from 1980 to 2017 corresponded with high levels of anger and negative emotions in song lyrics during that time.

Looking at the evidence of the above-mentioned research studies and the sharp inclination of song lyrics to become more negative over time, especially during times of socioeconomic crises, it was evident for the researcher of the present study to formulate hypotheses suggesting an increase in negative sentiment score and a decrease in positive sentiment score in popular song lyrics from the first to the second year of the pandemic. The first two years of the COVID-19 pandemic was an unprecedented time where "limited access to necessities and/or greater risk to health and prosperity…associated with greater prevalence of psychological distress and psychiatric disorder" (Pedrosa et al 2020: 12) allowed for sufficient grounds to assume that the trend of increasingly negative song lyrics would continue, if not increase, in these two years. However, as seen from the t-test results presented in Section 2.3, the present study diverged from the results of the earlier research studies.

Firstly, the statistical results for the song lyrics of both the United States and the United Kingdom charts are discussed in relation to the first hypothesis which states that the songs from the second year of the pandemic will have a lower positive sentiment score compared to the first year. In the case of the United States, the mean of the positive sentiment scores increased from the first year to the second year of the pandemic i.e., from 1.31 for 2020 to 1.40 for 2021. The p-value of the t-test was 8.55e-6 and this identified the difference between the mean scores as statistically significant, hence, the first hypothesis is rejected as

song lyrics in the US chart became more positive moving from the first to the second year of COVID-19.

These findings are in direct contrast to the downward and increasingly negative trend that song lyrics have been displaying for the past few decades. In their study, Putter et al (2021) reported a statistically significant model for the United States when comparing song lyrics from the first six months of the pandemic to song lyrics from 2015-2019. The findings showed that songs in 2020 contained fewer words related to human satisfaction and interest as compared to songs from 2015-2019. Moreover, lyrical content in songs from the US in 2020 was also "significantly more negatively valenced relative to song lyrics reaching the top-5 positions on the US chart between 2015 and 2019" (2021: 7). The authors associated the socioeconomic turbulence that rose during the COVID-19 pandemic with the popularity of song lyrics that reflected lower human satisfaction and more social isolation. Specifically for the United States, the results of the study showed a significantly "lower incidence of positive affect (in 2020) compared to the same months in 2015-2019" (Putter et al 2021: 10).

On the other hand, Pettijohn and Sacco (2009) conducted a study to investigate lyrical content of Billboard No. 1 songs from 1955 to 2003 in relation to the changing social and economic conditions in the United States over the same period. The authors based their study on the environmental security hypothesis or ESH, which offers a theoretical model to understand the relationship between socio-economic conditions and general trend observed in popular media. The ESH (Nelson et al 2007) posits that the way individuals perceive their environmental security has an influence over their social preferences and what they gravitate towards in terms of consumption, hence, interest in popular media changes based on changing social and economic conditions. The theory further elaborates on this notion by highlighting that safety and security become more important when times are uncertain or there is greater threat to human life, health, financial stability, etc. Thus, people tend to prefer

popular media with “more meaningful themes and mature characteristics…to help reduce threat and uncertainty” (Pettijohn & Sacco 2009: 300). The authors predicted in their study that song lyrics with more meaningful content would be more popular during threatening socioeconomic conditions. Meaningful content refers to song lyrics that explore serious themes and more complex issues which have greater social and cultural significance in society. Pennebaker et al (2001) highlight how writing can be an escape for singers/songwriters who might gravitate toward songs with more mature themes in order to reduce the tension and unpleasant or aversive arousal they are experiencing due to external conflicts during turbulent times.

Furthermore, individuals in a society who are going together through the experience of globally threatening social conditions might prefer popular music with serious themes in order to alleviate their stress by listening to music that highlights their shared experience. The COVID-19 pandemic can certainly be classed as a threating and turbulent time that was experienced, perhaps in differing degrees and intensity, but still felt and shared by people across the globe. In times like these, it stands to reason that individuals would shift towards listening to music containing more positive affect words to reduce their internal tension and attain a sense of security or safety. The following paragraphs discuss examples of song lyrics taken from the corpus of songs in top-5 positions in the United States charts to explain the increase in positive sentiment from the first to the second year of the pandemic.

2.4.1. Qualitative Analysis of Love Songs in the US and the UK Corpus

The song with the highest positive score in the United States corpus was ‘I hope’ by Gabby Barrett rated +2.43. The lyrics of the song contain many references to interpersonal relationships, connection, love, and intimacy. The song starts with the line ‘I hope she makes you smile’ which was given a positive sentiment score of +3 and a negative sentiment score of -1 by the SentiStrength software, meaning this line of the song has a high positive

sentiment and a neutral or no negative sentiment. The first few verses of the song continue to paint a very positive and lighthearted picture with the following lyrics:

> ‘I hope you're both feelin' sparks by the end of the drive
> I hope you know she's the one by the end of the night
> I hope you never ever felt more free
> Tell your friends that you're so happy’

SentiStrength has graded each of these lines with a positive sentiment score of +3 with the exception of the last line which has been graded +2. It is interesting to note here that the second lyric of the verse ‘I hope you know she's the one by the end of the night’ has been awarded a higher rating of +3 even though there are no overt positive sentiment words in this line, while the last line in the verse ‘Tell your friends that you're so happy’ contains an explicit positive sentiment word ‘happy’ along with a booster word ‘so’ which should increase the rating of the overall line. Instances like these show that the algorithm of the software can glean information based on the context and the meaning of whole line, rather than just calculating an average by summing up the sentiment scores of each individual sentiment word present in a sentence. Furthermore, the software graded each of the lines mentioned above with a negative sentiment score of -1 which denotes neutrality, showcasing that there was no negative sentiment present in them.

The advantage of having the software grade both positive and negative sentiment score simultaneously becomes even more visible with this particular song. While the beginning and the ending of the song contains lyrics with a high positive sentiment in lines such as ‘I hope she's wilder than your wildest dreams, She's everything you're ever gonna need’, these lines are immediately followed by an abrupt shift in tone and mood with the subsequent lyric ‘And then I hope she cheats’ which has a strong negative sentiment with a high negative sentiment score of -3. The song ends with the lines ‘And then I hope she cheats,

Like you did on me' and these lines are what contribute to the song having an overall negative sentiment score of -1.30.

The song 'I hope' is an excellent example of humans processing mixed emotions in parallel in the form of "transient feeling states that involve two opposite affects" (Berrios et al 2015: 1). While it could be argued that the song skews more towards an overall negative sentiment rather than a positive one as the songwriter details a picturesque and romantic scene that is marred by the notion of infidelity, the SentiStrength software has graded the song with a much higher overall positive sentiment score of +2.43 compared to an overall negative sentiment score of -1.30. In this instance, I would agree with the software's sentiment score rating based on the context of the song and the feelings that are evoked by the lyrics. The emotions which the positive lyrics of the song elicit in the listener (in this case, reader) cannot be overshadowed or negated by the subsequent negative lyrics, rather both the opposing sentiments exist in parallel to each other. The picture of a young couple, wild and carefree, meeting 'the one' and falling in love with them and the emotions of happiness, joy, lust, passion, yearning, and irrepressible hope stay with the audience even after reading the next few lyrics that reveal how the narrator of the story was cheated on by the male protagonist in the song and how she now hopes the new girl he's falling for does the same to him as he did to the narrator.

The negative emotions of bitterness, loss of love, anger, a thirst for revenge, sadness, loneliness, reminiscing for the past, nostalgia, and even regret come through the new lyrics, but they are felt in conjunction to the previous positive lyrics. In fact, it seems as if there are two stories taking place in two parallel universes – the story of the young couple falling in love and the story of the past lover who was cheated on. While one universe focuses on positive affect, the other revolves around negative affect and they both combine and co-exist to form a complex, emotionally mature, and stimulating song that touches upon many of the

themes that serve to provide comfort and security to people in threatening socioeconomic conditions. In line with the ESH, Pettijohn and Sacco (2009) hypothesized in their study that, in addition to the content being meaningful and comforting, the themes of romance and love would also be more popular in media and popular music as people would prefer to hear stories based on these topics. The authors anticipated "romance, including positive and negative elements of close relationships and love, to be a meaningful topic and romantic themes to be expressed more often when times were troubling" (Pettijohn & Sacco 2009: 301).

This is certainly true for the corpus of the present study where out of the 164 unique songs that were extracted from the two years of the pandemic, there were only 8 to 10 songs that did not contain any references to love or other similar romantic notions. Dukes et al (2003) analyzed 100 songs from popular music between 1958 and 1998 and discussed the expressions of love and sex in popular songs relative to gender, race, and cultural change. Their study was vital in showcasing the pervasiveness of love songs in popular media and how songs containing romantic notions can be used to express a desire for closeness, intimacy, social cohesion, and affiliation, particularly during bad times.

Research at the Harvard Graduate School of Education found that the global COVID-19 pandemic deepened the epidemic of loneliness in the United States with over 61% of young American adults reporting increased feelings of loneliness and isolation (Weissbourd et al 2021). The need for social closeness and cohesion to nullify the feelings of loneliness and isolation resulting from the pandemic can also be seen in other songs that topped the US Billboard Hot 100 chart in 2020 and 2021. Another example of a popular song that references a romantic partnership forged in tough and unpredictable times is 'Stuck with u' by Ariana Grande and Justin Bieber. The song had a higher positive sentiment score of +1.31 compared to a negative sentiment score of -1.14. The lyrics detail the reality experienced by many

couples during the pandemic lockdowns as they were stuck indoors with a lover/partner while having nothing to do. Some lines from the song that talk about this shared experience are as follows:

‘Got all this time on our hands
Might as well cancel our plans, yeah
I could stay here for a lifetime
So, lock the door and throw out the key
Can't fight this no more, it's just you and me
And there's nothin' I, nothin' I, I can do
I'm stuck with you, stuck with you, stuck with you
I'ma get to know you better
Kinda hope we're here forever
There's nobody on these streets
If you told me that the world's endin'
Ain't no other way that I can spend it’

Many lyrics from this song bear a striking resemblance to the beginning of the pandemic when the streets were empty of people, all vacations and other types of leisure plans had to be cancelled, and it felt as if this bleak situation would continue forever since there were no official updates and no way of knowing when the pandemic would truly end. This song’s acknowledgement of a universal experience during the pandemic and the emotions of anxiety, loneliness, and fear of the unknown allows for a shift in perspective to focus on a more positive experience full of love and passion, taking the frustration of being forced to stay indoors and twisting it into an opportunity to become more physically and emotionally close with someone, for instance, turning around the hopelessness and helplessness of not being able to do anything and feeling like everything is out of control with the words ‘there’s nothing I can do’ and ‘I’m stuck with u’ and then following that up with the hopeful lyrics ‘I’ma get to know better’.

Although the subject matter of the song deals with the serious themes of lockdown, social isolation, claustrophobia, anxiety, depression, etc., the song has an overall higher positive sentiment score and leaves the audience feeling hopeful, happy, and connected. In light of this, it is no wonder that songs with love themes were more popular during the pandemic and that songs with positive affect topped charts more frequently and stayed on the charts for longer than their counterparts, resulting in the overall positive sentiment score of top 5 songs in the US Billboard Hot 100 chart to significantly increase from the first to the second year of the pandemic.

In terms of the United Kingdom, the mean score for positive sentiment also increased from +1.29 to +1.33 from one year to the next and the p-value of 0.05 proved this difference to be statistically significant. Thus, the first hypothesis is also rejected for the UK as the top-5 charting songs in the Official Singles Chart Top 100 had a higher positive sentiment score from the first to the second year of the pandemic. The highest positive sentiment score in the UK dataset across both years was for the song 'Make me feel good' by Belters Only, which had a positive sentiment score of +2.77 and a negative sentiment score of -1, meaning it had no negative sentiment and was considered neutral. The high positive sentiment score of the song can be attributed to its lyrics that are of a more sexual than romantic nature, such as the following lines which scored a high positive sentiment score of +4 and a neutral negative sentiment score of -1:

'Make me feel good
Make me feel nice
Give me your loving all through the night
Make me feel good
Make me feel nice
Give me your loving all through the night'

Christenson et al (2019) looked at lyrics from 1,040 top-40 songs in the US from 1960 to 2010 using R and their findings revealed that the most predominant themes in popular music were references to romantic and sexual relationships. Their study showed that sex-related keywords and phrases rose sharply in popular songs during the decades under study. A similarly fascinating study by Alexopoulos and Taylor (2019) performed a content analysis of 1,500 popular songs in the genres of hip-hop, country, and pop across a span of 25 years from 1991 to 2015 in the US. The results of the study found that 68.6% of the songs contained romantic themes while 48.4% discussed sexual themes. The notion of infidelity was present or was the main theme of the song in about 15% of the total songs analyzed.

A recent paper published online in the journal Psychology of Music titled ""Love lies": A content analysis of romantic attachment style in popular music" by Jorgensen-Wells et al (2023) provides further insight into how song lyrics can reveal the types of romantic attachment styles that are present in popular music. The authors used an extensive coding system to quantify the different features of songs in their corpus, and one of the features recorded about all the songs was whether they were primarily about sex, love, or both. Sexual content in song lyrics was identified by marking all instances that mentioned acts of sex including making out, talking about the sexual appeal of another person, emphasizing the physical characteristics of a body, voicing out desire for sexual interactions, and any sexually explicit language. In contrast, non-sexual or love content was recognized in lyrics that talked about emotional intimacy, acts of affection such as holding hands or kissing on the cheek/forehead, desire to spend time with another person, and other romantic references. Thus, the authors grouped songs that were high in sexual content and low in love content under the category of songs about sex, while songs that were high in love or non-sexual content and low in sexual content were marked as songs about love/romance. Songs that had relatively equal instances of both sexual and non-sexual content were tagged as both. The

results of the study revealed a positive association between songs that had a high love content and a secure attachment style where both parties feel safe and at ease with their partner in a relationship that "fulfills…emotional needs, and their fear of abandonment fades, diminishing insecurity" (Jorgensen-Wells et al 2023: 806). These positive and healthy emotions were represented in lyrics such as 'Girl, you're more than just the surface', 'My love is yours if you're willing to take it', and 'Can't we just talk?/Figure out where we're goin'/ Can't get what we want without knowin''. On the other hand, songs that had a high sexual content were positively associated with an avoidant or anxious attachment style and marked by lyrics such as 'I need somebody who can take control', 'Baby girl you just a fling', 'Spent twenty-four hours/I need more hours with you', and 'Don't stand too close'. The authors concluded that songs portraying insecure attachment style were more prevalent in popular music with over 30% of songs having characteristics of an anxious or avoidant attachment style.

Comparing these results with the present study's corpus data reveals that the same can be said for songs that were popular during the COVID-19 pandemic for the UK dataset, but not for the US dataset. Taking the example of the songs mentioned above, the song 'I hope' by Gabby Barrett with the highest positive sentiment score of +2.43 in the US charts has more references to love content than to sexual content. For example, most of the acts depicted in the song lyrics are references to emotional intimacy rather than physical closeness, such as making a person smile on the phone, thinking about them when hearing a song, feeling sparks as you drive around with them, telling friends you're happy in your relationship, spending money on the other person (buying them a ring worthy of them), and going on a first date. There are very few, if any, references to sexual content in this song except for the mention of infidelity which, as discussed earlier, forms a parallel theme displaying negative affect alongside the positive affect conveyed by the other side of the

story. It is interesting to note here the association between song lyrics having higher love content and higher positive sentiment score.

On the contrary, the opposite is true for the UK dataset. The song with the highest positive sentiment score in the UK chart across the two years is ‘Make me feel good’ by Belters Only with a positive sentiment score of +2.77. The song is centralized on the theme of sex and explicit mentions of sexual intimacy are present throughout the song lyrics beginning from the title of the song ‘Make me feel good’ to the setting of the song which is in a ‘club, somewhere near the bar’. The scene is set for a meaningless sexual encounter between two strangers who meet at a bar. The main theme of the song can be summarized in the line ‘Give me your loving all through the night’ which is repeated in full or in part at least 26 times throughout the song. The song lyrics have a positive sentiment score of +4 and a neutral negative sentiment score of -1. Other lyrics that allow the song to be categorized as a song with high sexual content include lines such as ‘make me feel good’ and ‘make me feel nice’ which refer to physical satisfaction rather than emotional closeness.

While a comparison of only two songs from the US and UK song corpus is not nearly enough to allow for any generalizations, a qualitative analysis of the songs with the highest positive sentiment scores for the two countries reveals that the US audience might have preferred songs that contained more love or romantic content with frequent depictions of emotional intimacy, whereas the UK audience gravitated towards songs that were more explicit and sexual with references to physical desire and bodily gratification. It is also interesting that the UK song had a higher overall positive sentiment score of +2.77 compared to the US song which was rated +2.43. Another minor observation is that the UK song ‘Make me feel good’ was focused on sexual content and it stands to reason then that the song was rated by the SentiStrength software as being neutral in terms of negative valence, meaning sexual content is only seen as positive sentiment without any negative sentiment present. On

the other hand, the US song ‘I hope’ had a negative sentiment score of -1.30, indicative of mixed emotions of positive and negative sentiment occurring simultaneously in a song that contained more romantic notions and a higher love content, thereby highlighting the complexity and duality of mixed emotions which are activated when thinking about love and revenge in the song.

2.4.2. Analyzing Music Nostalgia and Political Turmoil in Songs

The second hypothesis of the present study was based on the conjecture that songs in the second year of the pandemic would have a higher negative sentiment score compared to the first year. In the case of the United States, the mean score for negative sentiment decreased from -1.34 in 2020 to -1.27 in 2021. The p-value of the t-test was 0.0014, which proved this difference to be statistically significant in the opposite direction of the second hypothesis as the song lyrics in the US corpus had a lower negative sentiment score from the first to the second year of the pandemic, leading to the second hypothesis being rejected. A look at the US song corpus across the two years reveals the song ‘Save your tears’ by The Weekend as having the highest negative sentiment score of -2.16. The title of the song ‘Save your tears’ is part of the lyric ‘Save your tears for another day’ and this line is repeated numerous times throughout the song and has been rated by SentiStrength software as having a negative sentiment score of -4 and a neutral positive sentiment score of +2. The song is a nostalgic and sad love song about a man lamenting making his old lover cry.

As the previous section has already looked at love songs in detail, the current section will not focus on the negative sentiment present in love songs but will instead take another route and look at the songs with the lowest negative sentiment scores to see if they are the reason for the overall US song corpus having a reduced negative sentiment score from one year of the pandemic to the next.

In the US songs corpus, four songs had the lowest negative sentiment score of -1, which reflects neutrality as understood by looking at the coding algorithm of the SentiStrength software. The songs with no negative sentiment are ‘Watermelon sugar’ by Harry Styles, ‘Rockin around the Christmas tree’ by Brenda Lee, ‘Jingle bell rock’ by Bobby Helms, and ‘A holly jolly Christmas’ by Burl Ives. With the exception of the song ‘Watermelon sugar’ which is a love song with high sexual content, the other three songs share the common theme of Christmas. Beginning with the song ‘Rockin around the Christmas tree’ by Brenda Lee which is a famous classical song about Christmas from 1958, the lyrics are full of holiday cheer and joy such as the following lines:

‘Rockin' around the Christmas tree
Let the Christmas spirit ring
Voices singing, Let's be jolly
Deck the halls with boughs of holly!
Rockin' around the Christmas tree
Have a happy holiday
Everyone's dancing merrily’

Words that contain positive affect such as ‘happy’, ‘jolly’, and ‘merrily’ contribute to the song having an overall positive sentiment score of +1.55. The song revolves around dancing, drinking, and having a jolly good time during the Christmas season. There are no words or phrases that would infect the song with any negative affect. Songs such as these made an appearance in the top-5 positions on the weekly singles US chart in the first week of December 2021, still categorized as the first year of the pandemic, and then they appeared again in a similar fashion from the beginning of December 2022, grouped under the second year of the pandemic. In both cases, the Christmas songs stayed in the top-5 positions in the charts for about four weeks until the beginning of January. This seasonal resurgence of

Christmas songs could be attributed to the slight decrease in the overall negative sentiment score in the first two years of the pandemic as people preferred popular music that would get them into the holiday spirit and take their focus away from troubling times.

Similarly, the song ‘Jingle bell rocks’ by Bobbly Helms is also an old song that was first released in 1957, while the song ‘A holly jolly Christmas’ by Burl Ives was released as a single in 1965. A few cheerful lyrics from the two songs are as follows:

‘Jingle bell, jingle bell, jingle bell rock
Jingle bells chime in jingle bell time
Dancin' and prancin' in Jingle Bell Square
In the frosty air
What a bright time, it's the right time
To rock the night away’

‘Have a holly jolly Christmas
It's the best time of the year
I don't know if there'll be snow
But have a cup of cheer
Have a holly jolly Christmas
And when you walk down the street
Say hello to friends you know’

Both songs are classical tunes that become popular again during the Christmas season. Gibbs and Egermann (2021) conducted a study on nostalgic music, defined as music that evoked feelings of nostalgia and yearning for a certain period of life, particular places or individuals. These feelings typically occur during hard times and the researchers looked at the first COVID-19 lockdown that was put into force in the UK in 2020 to collect data about music-induced nostalgia. The authors used a pre-validated scale and recorded

responses from 570 participants and the results indicated that listening to nostalgic music as an emotion regulation strategy proved very effective for the listeners and had an overall positive effect on their mental well-being during the COVID-19 crisis. Several recent studies have found similar results of how listening to music has helped individuals regulate their emotions and better their mental health in regard to the pandemic. Music has allowed listeners to feel socially connected and helped alleviate stress, psychological tension, loneliness, depression, etc. (Granot et al 2021, Krause et al 2021), especially during long periods of social isolation which were an unfortunate common occurrence during the beginning of the pandemic.

The findings of these research studies are in alignment with the song corpus of the present study which contained many Christmas themed songs that topped charts in the US during the pandemic, allowing listeners to indulge in feelings of nostalgia about the past holiday seasons when it was easy to celebrate with a large group of friends and family and there were no restrictions on public gatherings, physical closeness, indoor events, and more. Fink et al (2021) highlight how music can help people who are physically distanced to feel emotionally and socially connected to each other, which works as a coping mechanism in times of hardship caused by external stressors such as harsh lockdown policies, risk to health and life, financial difficulties, etc. which resulted due to the pandemic. Gibbs and Egermann further elaborated on the emotions invoked by music by stating that they were a result of "appreciating the past, or acknowledgement of happy past times" (2021: 12). This is certainly the case for the classical Christmas songs in the present study's corpus which viewed Christmas as 'the best time of the year' (A holly jolly Christmas) and painted a merry picture of attending Christmas parties where people are 'rockin around the Christmas tree' and the halls are decked with 'boughs of holly' (Rockin around the Christmas tree). The authors also refer to how the participants nostalgic feelings can be associated with notions

of perseverance through times of hardship and remaining hopeful and optimistic when things seem bleak. Davis (1977) stated that while negative emotions might be unintentionally raised when a person is immersed in nostalgic trips down memory lane, these sentiments are easy to filter through a redemptive mindset where the hardships of the past remain in the past and people are able to put a positive spin on their memories in order to feel happy about them in the present.

On the other hand, in the case of the United Kingdom's song corpus, the mean score for negative sentiment increased from -1.30 in the first year of the pandemic to -1.34 in the second year. The p-value of the t-test was 0.21, which suggests that this difference is statistically non-significant, leading the researcher to reject the second hypothesis in terms of statistical significance, even though the mean score of the song lyrics in the UK corpus had a slight increase in the negative sentiment score from one year to the next. A deeper look into the song corpus for the UK chart reveals that the highest negative sentiment score of -5.00 was given by SentiStrength software to the song 'Boris Johnson is a f**king c**t' by Kunt And The Gang. The song contains only six lines with variations of the title and each of the lines has been rated -5 by the software due to the presence of the two explicit swear words at the end. The song was released by the band in July 2020 and attained a position in the top-5 weekly singles chart for the last week of December 2020. The same band released a sequel to the song titled 'Boris Johnson is still a f**king c**t' the next year around December 2021 and, while the song was not given any radio time, it made the top 5 songs list in Amazon and iTunes charts within 48 hours of its release (Murray 2021). For the purposes of the present study, the song appeared in the top 5 in the UK chart in the last week of December 2021. The sequel song had five lines and they were also different variations of the song title. All the lines were rated -5 by SentiStrength software, except for the fourth line 'He's still a c**t!' which had only one swear word and was rated -4. Hence, the overall

negative sentiment score for the sequel song was -4.80. In both instances, these songs had an overall positive sentiment score of +1.00, meaning a neutral positive valence since there were no positive sentiment words present in the song lyrics. While both songs only stayed on the UK charts for one week across both years of the pandemic, their high negative sentiment score might have skewed the results of the UK song corpus, which led to the mean negative sentiment score displaying a minor increase from one year to the next.

The slight difference between the results of the UK song corpus and the US song corpus in terms of their negative sentiment scores could be attributed to the political turmoil faced by the United Kingdom which was expressed in explicit terms in the above-mentioned two songs concerning the ex-Prime Minister, Boris Johnson. The high negative sentiment score of these songs might have contributed to the small increase in the overall negative sentiment score of the UK song corpus from the first to the second year of COVID-19. However, this increase was not statistically significant, and the overall trend shows that, in contrast to the findings of the previous studies, the present study's corpus depicts a decrease in the negative sentiment score of song lyrics over time. This divergence from previous research which prominently highlighted a trend of increasingly negative song lyrics (Napier & Shamir 2018; Brand et al 2019; Putter et al 2021) could be attributed to methodological problems in the present study. The SentiStrength software, especially the free version, which was used for this thesis, is rudimentary and limited in its capacity to process complex song lyrics with emotionally mature themes, modern language use littered with novel slang words and abbreviations, and context-dependent sentiment words and phrases. The present study also looked at only the first two years of the pandemic and it might be possible that a change in popular music to reflect the socioeconomic turbulence of the COVID-19 pandemic might take place or be visible in songs collected a few years after the beginning of the pandemic,

taking into account the long and time-consuming process of writing, producing, and releasing music.

CONCLUSION

Song lyrics are an important medium of communication and a means of expression of emotions for individuals, especially during stressful periods that threaten life or health, present risk to social security, and increase financial instability, etc. It is concerning that recent studies have found a trend of increasing negativity in song lyrics over the last few decades as songs become less positive over time (Putter et al 2021). The present study aimed to investigate whether this would remain the case in times of extreme hardship and socioeconomic turbulence, specifically in the first two years of the COVID-19 pandemic. The purpose of this MA thesis was to highlight the importance of using the technique of sentiment analysis to measure sentiment present in song lyrics as a way of understanding trends in popular music and comparing them with preceding studies to provide insight into the emotional moods of society in times of adversity.

The research question of the present study aimed at exploring the similarities and differences between the sentiment of songs released in the first year of the pandemic compared to the second year. This research question was reformed into two hypotheses focusing on the two distinct types of emotional valence or sentiment – positive sentiment and negative sentiment, which can co-activate in the human brain at the same time, a state of affect known as mixed emotions (Berrios et al 2015). In summary, the two hypotheses proposed that the sentiment score of song lyrics would display an increase in negative sentiment and a decrease in positive sentiment moving from the first to the second year, keeping in line with the findings of previous research concerning popular music becoming increasingly negative as time passes.

The methodological approach of this study consisted of employing the SentiStrength software developed by Professor Mike Thelwall as a digital tool to quantitatively measure the positive and negative sentiment in song lyrics for a period of two years of the pandemic,

beginning from 11 March 2020 when COVID-19 was first declared a global pandemic (World Health Organization 2020). Two English-speaking countries, the US and the UK, were chosen as the research field and the top-5 songs were added to the corpus from the weekly singles category of each country's official music charts. For the United Kingdom, the songs came from the Official Singles Chart Top 100, and for the US, the chart was Billboard Hot 100. A total of 1,055 songs were added to the corpus from both countries' charts, out of which 164 songs were unique and had to be manually cleaned before being fed into the software for analysis.

The manual cleaning of the song lyrics followed guidelines set by the previous research studies (North et al 2018; Putter et al 2021) in order to allow for consistency of language use across the corpus and to ensure the software had minimum trouble with repetitions, redundancy, non-standard abbreviations, etc. The SentiStrength software rated each line of the song with a positive sentiment score and a negative sentiment score on a scale of 1 to 5 based on a sentiment lexicon built into the software containing 2,489 sentiment-bearing word stems that were graded by human coders (Thelwall et al 2021). The statistical results of the study included mean scores, standard deviation scores, range, and p-value from t-test scores all conducted using the relevant formulas in Excel.

The findings of the study revealed that the song corpus for the US had a statistically significant increase in the positive sentiment score of the song lyrics from the first to the second year of the pandemic, leading to the first hypothesis being rejected. The results of the present study contrasted with previous research studies which found that threatening external factors such as a high unemployment rate can lead to songs becoming increasingly negative and displaying emotions of anger, frustration, stress, etc. (Qiu et al 2021). A qualitative analysis of the songs with the highest positive sentiment scores revealed further insights into how individuals experiencing internal conflict, stress, or adverse emotions

prefer to listen to popular music with themes of romance and love where the song lyrics are more mature, meaningful, comforting, and do not shy away from discussing serious and complex notions such as infidelity, revenge, hope, yearning for a past love, etc. (Pettijohn & Sacco 2009). The overwhelming presence of love songs in the corpus of the present study attested to the popularity of such themes, especially during socioeconomic turmoil which was prevalent during the first two years of the pandemic. The increase in the positive sentiment score of the songs could also be associated with individuals seeking comfort, social cohesion, and interconnection from songs that described a shared experience such as the loneliness which came from lockdowns imposed during the pandemic. The findings were similar for the UK songs corpus which also displayed an increase in positive sentiment score, although a deeper look into specific songs in the corpus showed a generalized trend of US songs having more romantic or love content while UK songs skewed more towards having lyrics with explicit sexual content such as references to physical desire, objectification of the body, focus on physical appearances, and more.

In regard to the second hypothesis of the present study, the results showed a difference between the US and the UK sentiment scores. While the US song corpus had a lower negative sentiment score in the second year of the pandemic compared to the first year, the results of the UK songs corpus were in line with the hypothesis that the negative sentiment in songs increases over time, although this slight increase was not statistically significant. Both the countries' official charts enjoyed a week of popular Christmas tunes resurging into the top 5 positions during the holiday season and lowering the overall negative sentiment score of the corpus with lyrics that contained words filled with positive affect, holiday cheer, and goodwill. This nostalgic music allowed people in lockdown who were alone or physically distanced from friends and family to indulge in fondly remembering the

happier times of the past, thereby helping them regulate their emotions and view the bleak present reality of their life in a more positive and hopeful light (Fink et al 2021).

The findings of the present study revealed that, in general, songs became more positive and less negative, and the discussion section attempted a brief qualitative analysis of the songs with the highest and lowest positive and negative sentiment to understand the reasons behind the present study's results contrasting with the findings of previous studies. The divergent results could be due to a weakness in methodology such as the simplicity of the SentiStrength software with only two binary categories of positive and negative sentiment which might have caused many of the context-driven nuances of song lyrics to be ignored or erroneously rated. It cannot be denied that the COVID-19 pandemic was an unprecedented global crisis and people's response to such a catastrophic event might have been different or even delayed as compared to previous generations responding to the socioeconomic crises of their time such as The Great Depression. The proposed study is significant as it enhances an understanding of how cultural products such as songs reflect emotional changes in society, especially during times of socioeconomic crises. Furthermore, it contributes to a limited, but growing, body of research analyzing sentiment in song lyrics using corpus-based quantitative methods by incorporating the use of digital software. Hence, research on song lyrics can contribute to our understanding of how cultural products reflect changes in society, especially during times of social and emotional upheaval. The thesis contributes to a limited, but growing, body of work analyzing sentiment in song lyrics of popular music.

This MA thesis reinforces the importance of popular music as a reflection of the mood of society during turbulent times by allowing song lyrics to become a global medium and a means of expressing shared emotions such as feelings of alienation, loneliness, anxiety, and struggles with mental health that left people reeling in the wake of the pandemic. Song

lyrics with meaningful and mature themes pertaining to serious topics brought to light the isolation people were struggling with and also exhibited the importance of social cohesion, connectedness, and emotional as well as physical intimacy during stressful times. The rise of hyper individualism has led to an epidemic of loneliness (Weissbourd et al 2021), especially in the United States, and it would be interesting for future studies to observe whether this societal mood influences popular music to become increasingly negative in order to mirror the emotions of individuals in the society or whether song lyrics 'fight back' to contain more positive sentiment that allows people to feel connected and provides them an escape from reality along with a measure of relief. Future research studies exploring the role of music during the COVID-19 pandemic could also expand the current study's corpus of songs from the first two years of the pandemic and collect data from more years in order to include more songs from a larger time period to help provide a deeper and more holistic look into how the positive and negative sentiment of song lyrics changes over time.

As the world becomes increasingly chaotic and political turmoil, wars and famine, economic instability, and lack of resources continue to adversely affect people's emotional and mental health, will songs become increasingly negative as a collective expression of the thoughts and emotions of a society and a generation undergoing turbulent times or will they achieve higher positive sentiment to provide hope and a safe haven for music listeners all over the world? This will remain a question for future research exploring how song lyrics and popular music can contribute to our understanding of changes in society.

REFERENCES

Alexopoulos, Cassandra and Laramie D. Taylor. 2019. Easy listening? An analysis of infidelity in top pop, hip-hop, and country song lyrics over 25 years. *Psychology of Music*, 48: 6, 795-807.

Baron, Naomi S. 2003. Language of the Internet. In Ali Farghali (ed), The Stanford Handbook for Language Engineers (pp. 59-127). Stanford: CSLI Publications.

Barr, Dale J., Roger Levy, Christoph Scheepers and Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68: 3, 255-278.

Barradas, Gonçalo T. and Laura S. Sakka. 2021. When words matter: A cross-cultural perspective on lyrics and their relationship to musical emotions. *Psychology of Music*, 50: 2, 650-669.

Barrett, Lisa F. and Eliza Bliss-Moreau. 2009. Affect as a psychological primitive. *Advances in Experimental Social Psychology*, 41, 167–218.

Bennett, Andy, Barry Shank and Jason Toynbee (eds). 2005. *The Popular Music Studies Reader*. London: Routledge.

Berrios, Raul, Peter Totterdell and Stephen Kellett. 2015. Eliciting mixed emotions: a meta-analysis comparing models, types, and measures. *Frontiers in Psychology*, 6: 428, 1-28.

Brand, Charlotte O., Alberto Acerbi and Alex Mesoudi. 2019. Cultural evolution of emotional expression in 50 years of song lyrics. *Evolutionary Human Sciences*, 1: 11, 1–14.

Brenner, Berthold. 1975. Enjoyment as a preventive of depressive affect. *Journal of Community Psychology*, 3: 4, 346-357.

Cacioppo, John T., Jeff T. Larsen, N. Kyle Smith and Gary G. Berntson. 2004. The Affect System: What Lurks below the Surface of Feelings? In Antony S. R. Manstead, Nico Frijda and Agneta Fischer (eds). *Feelings and emotions: The Amsterdam symposium*, 223-242. Cambridge University Press.

Cacioppo, John T., Wendi L. Gardner and Gary G. Berntson. 1999. The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology*, 76: 5, 839–855.

Christenson, Peter G., Silvia de Haan-Rietdijk, Donald F. Roberts and Tom F. M. ter Bogt. 2019. What has America been singing about? Trends in themes in the U.S. top-40 songs: 1960–2010. *Psychology of Music*, 47: 2, 194-212.

Clark, Lee Anna and David Watson. 1989. Diurnal variation in the Positive Affects. *Motivation and Emotion*, 13: September, 205–234.

Das, Sanjiv and Mike Chen. 2001. Yahoo! for Amazon: Extracting market sentiment from stock message boards. In Proceedings of the Asia Pacific Finance Association Annual Conference (APFA).

Davis, Fred. 1977. Nostalgia, identity and the current nostalgia wave. *The Journal of Popular Culture*, 11: 2, 414–424.

DeWall, C. Nathan, Richard S. Pond, W. Keith Campbell and Jean M. Twenge. 2011. Tuning in to psychological change: Linguistic markers of psychological traits and emotions over time in popular U.S. song lyrics. *Psychology of Aesthetics, Creativity, and the Arts*, 5: 3, 200–207.

Diener, Ed and Robert A. Emmons. 1984. The independence of positive and negative affect. *Journal of Personality and Social Psychology*, 47: 5, 1105-1117.

Dodds, Peter Sheridan and Christopher M. Danforth. 2010. Measuring the happiness of large-scale written expression: songs, blogs, and presidents. *Journal of Happiness Studies*, 11: 4, 441–456.

Dukes, Richard L., Tara M. Bisel, Karoline N. Borega, Eligio A. Lobato and Matthew D. Owens. 2003. Expressions of love, sex, and hurt in popular songs: A content analysis of all-time greatest hits. *The Social Science Journal*, 40: 4, 643-650.

Fink, Lauren, Lindsay Warrenburg, Claire Howlin, William M. Randall, Niels Chr. Hansen, Melanie Wald-Fuhrmann. 2021. Viral tunes: Changes in musical behaviours and interest in coronamusic predict socio-emotional coping during COVID-19 lockdown.. *PsyArXiv Preprint*, 1-36.

Fox, Elaine. 2008. *Emotion Science: Cognitive and Neuroscientific Approaches to Understanding Human Emotions*. Basingstoke, New York: Palgrave Macmillan.

Fox, Elaine. 2018. Perspectives from affective science on understanding the nature of emotion. *Brain and Neuroscience Advances*, 2: December, 1-8.

Gibbs, Hannah and Hauke Egermann. 2021. Music-Evoked Nostalgia and Wellbeing During the United Kingdom COVID-19 Pandemic: Content, Subjective Effects, and Function. *Frontiers in Psychology,* 12: March, 1-16.

Granot, Roni, Daniel Spitz, Boaz Cherki, Psyche Loui, Renee Timmers et al. 2021. “Help! I need somebody”: Music as a global resource for obtaining wellbeing goals in times of crisis. *PsyArXiv Preprints*, 1-70.

Grinter, Rebecca E. and Margery Eldridge. 2003. Wan2tlk?: Everyday text messaging. *CHI 2003*, April 5: 10, 441-448.

Hardt-Mautner, Gerlinde. 1995. "Only Connect". *Critical Discourse Analysis and Corpus Linguistics*. UCREL Technical Paper 6. Lancaster, UK: Lancaster University.

Hu, Minqing and Bing Liu. 2004. Mining and summarizing customer reviews. In Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD "04). New York: Association for Computing Machinery.

Jorgensen-Wells, McKell, Sarah M Coyne and Janna M Pickett. 2023. "Love lies": A content analysis of romantic attachment style in popular music. *Psychology of Music*, 51: 3, 804-819.

Kelbessa, Workineh. 2013. The Potential Contribution of Love-Sex Songs to the Spread or Prevention of HIV/AIDS : The Case of the Oromo of Ethiopia. *Africa Development*, 38: 3-4, 33-95.

Krause, Amanda E. and Adrian C. North. 2019. Pop music lyrics are related to the proportion of female recording artists: Analysis of the United Kingdom weekly top five song lyrics, 1960–2015. *Psychology of Popular Media Culture*, 8: 3, 233–242.

Krause, Amanda E. and Adrian C. North. 2020. Collaboration, cognitive effort, and self-reference in United Kingdom top 5 pop music lyrics 1960–2015. *Psychology of Aesthetics, Creativity, and the Arts*, 14: 3, 294–300.

Krause, Amanda E., James Dimmock, Amanda L. Rebar and Ben Jackson. 2021. Music listening predicted improved life satisfaction in university students during early stages of the COVID-19 pandemic. *Frontiers in Psychology*, 11: January, 1-9.

Laland, Kevin N. 2004. Social learning strategies. *Learning & Behavior*, 32: 1, 4-14.

Larsen, Jeff T., A. Peter McGraw and John T. Cacioppo. 2001. Can people feel happy and sad at the same time? *Journal of Personality and Social Psychology*, 81: 4, 684–696.

Leppert, Richard. 2016. Civil War Imagery, Song, and Poetics: The Aesthetics of Sentiment, Grief, and Remembrance. *19th-Century Music*, 40: 1, 20-46.

Liu, Bing. 2012. *Sentiment analysis and opinion mining*. Synthesis Lectures on Human Language Technologies, 5: 1, 1-167.

Liu, Bing. 2015. *Sentiment Analysis*. New York: Cambridge University Press.

Mauch, Matthias, Robert M. MacCallum, Mark Levy and Armand M. Leroi. 2015. The evolution of popular music: USA 1960–2010. *Royal Society Open Science*, 2: 5, 1-10.

Medhat, Walaa, Ahmed Hassan and Hoda Korashy. 2014. Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5: 4, 1093–1113.

Mucci, Federico, Nicola Mucci and Francesca Diolaiuti. 2020. Lockdown and isolation: psychological aspects of COVID-19 pandemic in the general population. *Clinical Neuropsychiatry*, 17: 2, 63-64.

Murray, Robin. 2021. The Kunts' 'Boris Johnson Is Still A F*cking C*nt' Gatecrashes The Charts. Available at https://www.clashmusic.com/news/the-kunts-boris-johnson-is-still-a-fcking-cnt-gatecrashes-the-charts/, accessed May 14, 2023.

Napier, Kathleen and Lior Shamir. 2018. Quantitative sentiment analysis of lyrics in popular music. *Journal of Popular Music Studies*, 30: 4, 161–176.

Nasukawa, Tetsuya and Jeonghee Yi. 2003. Sentiment analysis: Capturing favorability using natural language processing. In Proceedings of the Conference on Knowledge Capture (K-CAP).

Nelson, Leif D., Pettijohn, Terry F. And Jeff Galak. 2007. Mate preferences in social cognitive context: When environmental and personal change leads to predictable cross-cultural variation. In Viren Swami and Adrian Furnham (eds), *Body Beautiful: Evolutionary and Sociocultural Perspectives*, 185-206. New York: Palgrave.

Norris, Catherine J., Jackie Gollan, Gary G. Berntson and John T. Cacioppo. 2010. The current status of research on the structure of evaluative space. *Biological Psychology*, 84: 3, 422–436.

North, Adrian C., Amanda E. Krause, Robert Kane and Lorraine Sheridan. 2018. United Kingdom "top 5" pop music lyrics. *Psychology of Music*, 46: 5, 638–661.

Official Charts. 2023. How the Official Charts are compiled. Available at https://www.officialcharts.com/getting-into-the-charts/how-the-charts-are-compiled/, accessed March 16, 2023.

Pang, Bo and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2: 1-2, 1-90.

Pang, Bo, Lillian Lee and Shivakumar Vaithyanathan. 2002. Thumbs up? Sentiment classification using machine learning techniques. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 79–86.

Parksepp, Lotte. 2019. Sentiment Mining in Goodreads Reviews of Classic American Novels. Unpublished BA thesis. Department of English Studies, University of Tartu, Tartu, Estonia.

Pedrosa, Letícia Bitencourt, Ana Cláudia Fontoura Fróes and Maria Luíza Barreto Cazumbá et al. 2020. Emotional, Behavioral, and Psychological Impact of the COVID-19 Pandemic. *Frontiers in Psychology*, 11: October, 1-18.

Pettijohn II, Terry F. and Donald F. Sacco Jr. 2009. The Language of Lyrics: An Analysis of Popular Billboard Songs Across Conditions of Social and Economic Threat. *Journal of Language and Social Psychology*, 28: 3, 297-311.

Picard, Rosalind W. 1997. *Affective Computing*. Massachusetts: MIT Press.

Putter, Kaila C., Amanda E. Krause and Adrian C. North. 2021. Popular music lyrics and the COVID-19 pandemic. *Psychology of Music*, 1-16.

Qiu, Lin, Sarah Hian May Chan, Kenichi Ito and Joyce Yan Ting Sam. 2021. Unemployment rate predicts anger in popular music lyrics: Evidence from top 10 songs in the United States and Germany from 1980 to 2017. *Psychology of Popular Media*, 10: 2, 256–266.

Rae, Roland. 2020. *Sentiment analysis of selected works by Mark Twain with the statistical software R*. Unpublished BA thesis. Department of English Studies, University of Tartu, Tartu, Estonia.

Ravignani Andrea, Henkjan Honing and Sonja A. Kotz. 2017. Editorial: The evolution of rhythm cognition: Timing in music and speech. *Frontiers in Human Neuroscience*, 11: June, 1-8.

Reagan, Andrew J., Lewis Mitchell, Dilan P. Kiley, Christopher M. Danforth and Peter Dodds. 2016. The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science*, 5: November, 1-12.

Russell, James A. 1979. Affective space is bipolar. *Journal of Personality and Social Psychology*, 37: 3, 345-356.

Russell, James A. and James M. Carroll. 1999. On the bipolarity of positive and negative affect. *Psychological Bulletin*, 125: 1, 3–30.

Savage, Patrick E. 2019. Cultural evolution of music. *Palgrave Communications*, 5: 1, 1-12.

SentiStrength. n.d. About SentiStrength. Available at http://sentistrength.wlv.ac.uk/#About, accessed May 12, 2023.

Stratton, Jon. 2021. Parodies for a pandemic: coronavirus songs, creativity and lockdown. Cultural Studies, 35: 2-3, 412-431.

Thelwall, Mike, Kevan Buckley and Georgios Paltoglou. 2012. Sentiment strength detection for the social web. *American Society for Information Science and Technology*, 63: 1, 163–173.

Thelwall, Mike, Kevan Buckley, Georgios Paltoglou and Di Cai. 2010. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61: 12, 2544–2558.

Thurlow, Crispin. 2003. Generation Txt? The sociolinguistics of young people's text-messaging. *Discourse Analysis Online*, 1: 1, 1-27.

Tong, Richard M. 2001. An operational system for detecting and tracking opinions in on-line discussion. In Proceedings of the Workshop on Operational Text Classification (OTC).

Turney, Peter. 2002. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In Proceedings of the Association for Computational Linguistics (ACL), 417–424.

Varnum, Michael E. W., Jaimie Arona Krems, Colin Morris, Alexandra Wormley and Igor Grossmann. 2021. Why are song lyrics becoming simpler? A time series analysis of lyrical complexity in six decades of American popular music. *PLOS ONE*, 16: 1, 1-18.

Watson, David and Auke Tellegen. 1985. Toward a consensual structure of mood. *Psychological Bulletin*, 98: 2, 219-235.

Watson, David and Auke Tellegen. 1999. Issues in dimensional structure of affect—Effects of descriptors, measurement error, and response formats: Comment on Russell and Carroll (1999). *Psychological Bulletin*, 125: 5, 601–610.

Watson, David, Lee Anna Clark and Auke Tellegen. 1988. Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54: 6, 1063-1070.

Weissbourd, Richard, Milena Batanova, Virginia Lovison, and Eric Torres. 2021. Loneliness in America: How the pandemic has deepened an epidemic of loneliness and what we can do about it. Available at https://mcc.gse.harvard.edu/reports/loneliness-in-america, accessed May 13, 2023.

West, David. 2019. Introduction: The challenges of the song lyric. *Language and Literature*, 28: 1, 3-6.

Wills, Thomas A. 1986. Stress and coping in early adolescence: Relationships to substance use in urban school samples. *Health Psychology*, 5: 6, 503-529.

World Health Organization. 2020. Listings of WHO's response to COVID-19. Available at https://www.who.int/news/item/29-06-2020-covidtimeline, accessed November 24, 2021.

Wu, Chenyang, Scott Le Vine and Elizabeth Bengel et al. 2021. Sentiment analysis of popular-music references to automobiles, 1950s to 2010s. *Transportation*, 49: April, 641–678.

Yi, Jeonghee ,Tetsuya Nasukawa, Razvan Bunescu and Wayne Niblack. 2003. Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques. In Proceedings of the IEEE International Conference on Data Mining (ICDM).

Youngblood, Mason. 2019. Cultural transmission modes of music sampling traditions remain stable despite delocalization in the digital age. *PLoS ONE*, 14: 2, 1-12.

Appendix 1: US Songs with Positive and Negative Sentiment Scores

No.	Song title	Positive Sentiment Score	Negative Sentiment Score
1	Blinding lights	1.15	1.24
2	Circles	1.23	1.38
3	Don’t start now	1.09	1.41
4	Heartless	1.11	1.91
5	Life is good	1.19	1.59
6	Lonely	1.16	1.67
7	No time to die	1.36	2.03
8	Physical	1.24	1.05
9	Rain	1.11	1.27
10	Roses	1.21	1.40
11	Stupid love	2.08	1.57
12	The box	1.13	1.32
13	34+35	1.12	1.28
14	Driver’s license	1.43	1.16
15	Go crazy	1.43	1.40
16	Up	1.25	1.58
17	Save your tears	1.34	2.16
18	Toosie Slide	1.27	1.39
19	Say so	1.37	1.35
20	The Scotts	1.08	1.13
21	Savage	1.26	1.88

22	Stuck with u	1.31	1.14
23	Gooba	1.14	2.10
24	Rockstar	1.29	1.38
25	Rain on me	1.14	1.09
26	Intentions	1.53	1.18
27	Trollz	1.25	1.42
28	The bigger picture	1.19	1.65
29	What’s poppin	1.19	1.48
30	Come & go	1.96	1.10
31	Wishing well	1.27	1.98
32	Popstar	1.40	1.51
33	Cardigan	1.39	1.13
34	The 1	1.33	1.15
35	Watermelon sugar	1.11	1.00
36	WAP	1.14	1.63
37	Laugh now cry later	1.44	1.52
38	Dynamite	1.39	1.02
39	Holy	1.33	1.14
40	Mood	1.50	1.24
41	Franchise	1.29	1.17
42	Savage love	1.98	1.87
43	Positions	1.43	1.08
44	Forever after all	1.34	1.05
45	I hope by Gabby Barrett	2.43	1.30

46	Therefore I am	1.17	1.43
47	Life goes on	1.11	1.15
48	All I want for Christmas	1.32	1.09
49	Rockin around the Christmas tree	1.55	1.00
50	Dakiti	1.31	1.18
51	Jingle bell rock	1.27	1.00
52	Willow	1.18	1.38
53	A holly jolly Christmas	1.79	1.00
54	It's the most wonderful	2.10	1.23
55	Levitating	1.17	1.03
56	Whats next	1.21	1.34
57	Wants and needs	1.34	1.42
58	Lemon pepper freestyle	1.25	1.32
59	Leave the door open	1.35	1.02
60	Peaches	1.08	1.55
61	Montero	1.40	1.26
62	Rapstar	1.26	1.84
63	Kiss me more	1.57	1.24
64	Good 4 u	1.83	1.26
65	my.life	1.26	1.51
66	amarl	1.30	1.33
67	Butter	1.44	1.21
68	Déjà vu	1.29	1.21
69	Bad habits	1.23	1.44

70	Permission to dance	1.06	1.20
71	Stay	1.52	1.31
72	Industry baby	1.24	1.20
73	Rumors	1.22	1.75
74	Way 2 sexy	2.18	1.33
75	Girls want girls	1.15	1.30
76	Fair trade	1.26	1.43
77	Champagne poetry	1.41	1.48
78	Knife Talk	1.11	1.95
79	Fancy like	1.64	1.12
80	My universe	1.21	1.04
81	Easy on me	1.30	1.03
82	Shivers	1.54	1.19
83	All too well	1.27	1.31
84	Smokin out the window	1.24	1.29
85	Heat waves	1.36	1.17
86	We don’t talk about Bruno	1.24	1.22
87	Super gremlin	1.32	1.53
88	Do we have a problem?	1.32	1.41
89	abcdefu	1.29	1.83
90	oh my god	1.45	1.37

Appendix 2: UK Songs with Positive and Negative Sentiment Scores

No.	Song title	Positive Sentiment Score	Negative Sentiment Score
1	Calling my phone	1.28	1.30
2	Friday	1.04	1.03
3	Latest trends	1.28	1.57
4	Patience	1.09	1.30
5	The business	1.06	1.09
6	Wellerman	1.21	1.13
7	You'll never walk alone	1.50	1.14
8	Times like these	1.49	1.15
9	Death bed	1.21	1.11
10	Rover	1.42	1.24
11	Breaking me	1.09	1.02
12	Dinner guest	1.35	1.21
13	Head & heart	1.00	1.08
14	Lighter	1.18	1.30
15	Only you freestyle	1.32	1.18
16	West ten	1.16	1.17
17	Ain’t it different	1.38	1.22
18	Mood swings	1.36	1.38
19	Looking for me	1.19	1.00
20	Lemonade	1.28	1.58
21	You broke me first	1.17	1.93

22	What you know bout love	1.80	1.12
23	Midnight sky	1.91	1.09
24	Really love	1.86	1.29
25	See nobody	1.15	1.28
26	Sweet melody	1.51	1.14
27	Last Christmas	1.41	1.31
28	Fairytale of New York	1.27	1.20
29	Don’t stop me eatin	1.24	1.15
30	This Christmas	1.31	1.20
31	Boris Johnson	1.00	5.00
32	Rockin around JB	1.46	1.00
33	Afterglow	1.37	1.23
34	Whoopty	1.12	1.60
35	Anyone	1.65	1.13
36	Don’t play	1.01	1.23
37	Without you	1.00	1.49
38	Paradise	2.13	1.38
39	Money talks	1.23	1.25
40	Bringing it back	1.23	1.23
41	Bed	1.23	1.04
42	Body	1.27	1.22
43	Your Power	1.53	1.20
44	Traitor	1.44	1.56
45	Heartbreak anthem	1.11	2.14

46	Holiday	1.44	1.08
47	I wanna be your slave	1.30	1.52
48	3 lions	1.07	1.21
49	Clash	1.22	1.54
50	Black magic	1.03	1.08
51	Verdansk	1.34	1.41
52	Happier than ever	1.41	1.57
53	Remember	1.13	1.40
54	Visiting hours	1.35	1.23
55	Cold heart	1.10	1.28
56	Obsessed with you	1.18	1.74
57	Love Nwantiti	1.49	1.23
58	Boyz	1.91	1.64
59	Overpass Grafiti	1.41	1.24
60	Flowers	1.63	1.29
61	I drink wine	1.68	1.30
62	Merry Christmas	1.63	1.19
63	Sausage rolls for everyone	1.56	1.10
64	Boris Johnson Still	1.00	4.80
65	Seventeen going under	1.30	1.88
66	Coming for you	1.18	1.48
67	Fingers crossed	1.41	1.37
68	Peru	1.13	1.08
69	Surface pressure	1.32	1.48

70	Where are you now	1.43	1.04
71	The joker and the queen	1.28	1.10
72	Make me feel good	2.77	1.00
73	Starlight	1.64	1.65
74	Down under	1.18	1.08

RESÜMEE

TARTU ÜLIKOOL

ANGLISTIKA OSAKOND

Mehar Abbas Ali Habibi

Sentiment analysis of top-5 charting songs in the US and the UK from March 2020-2022

Ameerika Ühendriikide ja Ühendkuningriigi 2020-2022 edetabelite TOP 5 laulude meelsusanalüüs

Magistritöö

2023

Lehekülgede arv: 75

Käesoleva magistritöö eesmärgiks on läbi viia korpusuuring, kasutades meelsusanalüüsi meetodeid, et uurida positiivseid ja negatiivseid meelsuseid ehk emotsionaalset valentsi Ameerika Ühendriikide ja Ühendkuningriikide TOP 5 edetabelitesse pürginud lauludes ajavahemikus märts 2020 kuni märts 2022, mis olid esimesed kaks pandeemia-aastat. Selle uuringu motivatsiooniks on avastada levimuusika trende sotsiaalmajanduslike kriiside ajal, et arendada arusaama sellest, kuidas muusika peegeldab ühiskonna emotsionaalset meelsust. Töö sihiks on panustada kasvavasse huvisse valdkonna vastu, mis kasutab kvantitatiivseid meetodeid levimuusikas olevate mustrite uurimiseks. Kasutades kvalitatiivset analüüsi kindlate korpuses olevate laulude jaoks, üritab see uuring samuti pakkuda täiendavat teavet selle kohta, kuidas levimuusika laulusõnade emotsionaalne valentsus võib ühiskonda mõjutada.

Magistritöö on jaotatud järgnevateks osadeks: sissejuhatus, mis defineerib termini „levimuusika“ (popular music) ja annab kiire ülevaate SentiStrength tarkvara kohta, mida on analüüsiks kasutatud; esimene peatükk, mis kirjeldab meelsusanalüüsi meetodit ning seda algoritmi, mida tarkvara kasutab laulusõnade kui lühikeste tekstide analüüsiks, ja mis annab põhjaliku kirjandusülevaate selle kohta, mis on varasemalt laulude meelsusanalüüsi valdkonnas uuritud; teine peatükk, mis täpsustab uurimistöö metodoloogilist lähenemist ning mille arutlev osa on jaotatud alapeatükkideks, kus on põgusad kvalitatiivsed analüüsid mõnest laulust, millel olid kõige kõrgemad positiivset ja negatiivset meelsust kirjeldavad tulemused; kokkuvõte; ja lisad, kus võib leida kõik 164 korpuses olnud ainulaadset laulu koos oma vastavate positiivse ja negatiivse meelsuse hinnangutega.

Esimene peatükk keskendub meelsusanalüüsile, selle definitsioonile, kategooriatele ja lähenemistele. See peatükk heidab samuti pilku ka eelnevale uurimistööle, mis käsitleb vastakaid tundeid, s.t suutlikust tajuda paralleelselt nii positiivset kui negatiivset meelsust (Berrios et al 2015). Esimese peatüki kirjandusülevaatele järgneb arutelu selle kohta, mis valmistas uurimuse läbiviijale raskusi ja mis on digitaalse tarkvara kasutamise puudused meelsusanalüüsis. Teine peatükk sisaldab endas sügavamat sissevaadet uurimuse metodoloogilistesse valikutesse, kaasa arvatud ülevaadet sellest, miks on tööks valitud SentiStrength tarkvara ja kuidas selle algoritmi ülesehitus sõltub kindlaksmääratud nimekirjast, milles on 2489 positiivsete ja negatiivsete meelsustega sõnatüve, mida on hinnatud skaalal ühest viieni inimestest kodeerijate poolt. Peatükk samuti seletab, kuidas iga

unikaalse laulu sõnad manuaalselt ja aeganõudvalt korrastatakse, et tagada järjepidevus korpuse keelekasutuses. Kvantitatiivse analüüsi osa esitleb statistilisi tulemusi nagu keskmised väärtused, standardhälve, ulatus ja T-testi väärtused, et kinnitada tulemuste statistilist olulisust. Sellele osale järgneb arutelu, milles õigustatakse seda, miks uurimistöö kaks hüpoteesi kas kinnitatakse või ümber lükatakse, võttes aluseks kindlaksmääratud positiivsed ja negatiivsed meelsused. Valitud laulude kvalitatiivne analüüs on samuti välja toodud selleks, et analüüs oleks terviklikum ja mitmekülgsem, võttes arvesse laulusõnade laiemat konteksti ja väliseid tegureid, mis neid mõjutavad.

Magistritöö kokkuvõte toob välja uurimuse põhilised tulemused ja arutleb selle üle, millist rolli mängib meelsusanalüüs uurides emotsionaalset valentsi levimuusika laulusõnades, mis olid edetabelite tipus spetsiifilises ajaraamis. Üldiselt pakub see töö väärtuslikke teadmisi sellest, kuidas levimuusikat kasutatakse emotsioonide väljendamiseks keerulistel aegadel.

Märksõnad:

meelsusanalüüs, COVID-19 pandeemia, levimuusika, Ameerika Ühendriigid, Ühendkuningriik, Billboardi edetabel, TOP 5 laulud, laulusõnad, digihumanitaaria, korpusuuring

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Mehar Abbas Ali Habibi,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose

Sentiment analysis of top-5 charting songs in the US and the UK from March 2020-2022,

mille juhendaja on Jane Klavan,

reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commonsi litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.

3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.

4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Mehar Abbas Ali Habibi

16.05.2023

Autorsuse kinnitus

Kinnitan, et olen koostanud käesoleva magistritöö ise ning toonud korrektselt välja teiste autorite panuse. Töö on koostatud lähtudes Tartu Ülikooli maailma keelte ja kultuuride kolledži anglistika osakonna magistritöö nõuetest ning on kooskõlas heade akadeemiliste tavadega.

Mehar Abbas Ali Habibi (allkirjastatud digitaalselt)

16.05.2023

Lõputöö on lubatud kaitsmisele.

Jane Klavan

16.05.2023