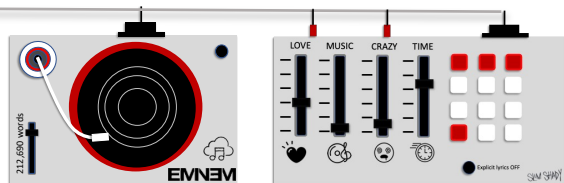


WORD UP!

A Data Analysis of Lyrics by Eminem

Eugene Garvilles

Digital Humanities, Chaminade University of Honolulu



Introduction

From its genesis in the 1970s, Hip hop music typically focused on themes of social injustice, drugs, gang crime, misogyny and race inequality¹. Eminem burst into the scene in 1996, disrupting the industry, leaving an indelible mark. This study focuses on the lyrics from his music over time to take a further look into him as an artist using words to reveal themes, creativity and change over time.

Hypothesis

There are no key themes in Eminem's work due to the wide variation, constant evolution and style in his lyrics. As a popular artist, his music will follow the trend seen in hit songs with a declining lexical density and diversity (indicating increased repetition). As seen in the trend in popular songs, the self-focused lyrics increases and other-focused decreases over time².

Methods

The dataset was taken from Kaggle, analysis done through R and visualization through Affinity Designer and Prezi.

Data Analysis & Results

The Eminem dataset is comprised of 212,690 words over 30 albums from 1996 to 2020 from Kaggle.

To create a visual representation of the data and identify the key themes, a word cloud was generated. The larger the size of the word the more frequent the word is used. After filtering non-Eminem lyrics (e.g., chorus, intro), explicit words and stop-words (i.e. is, are, the) **key themes emerged: love, time, crazy and music** (Figure 1). These words were further analyzed to review consistent mentions in his various musical generations.



Figure 1. Discography word cloud

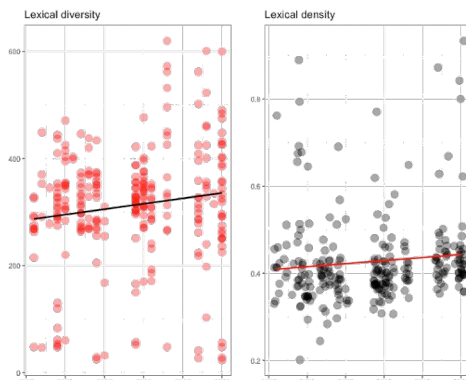


Figure 2. Lexical diversity and density

Further analyzing the lyrical complexity in terms of lexical diversity and density, there is an upward trend of unique vocabulary words used over the 14 years. Lexical diversity is a measure of different words appearing in a text while lexical density is an indicator of word repetition, a ratio between unique words and the total number of words³. This indicates that not only did he use **more words, there was also less repetition** highlighting his creativity, vocabulary and unique talent.

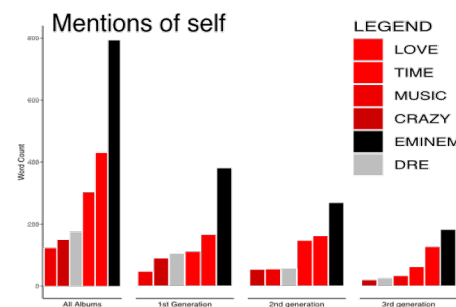


Figure 3. Frequency of mentions Eminem, Dr. Dre and key themes

Conclusion

Looking beyond explicit and non-lyrics, central themes of Eminem's discography showcase the terms love, music, crazy and time.

With each album, he increased originality in lyrical complexity avoiding repetition and offering unique words.

Over time, his mention of himself reduced while he maintained a nod to his mentor, Dr. Dre.

Limitations & Future Work

Due to availability of data, his music from 2022 was excluded, further analysis can be done to include this data. Proportion of mentions can help analyze shift from self-focus to other focus.

Credits & References

Photo Credits:

1. nounproject.com
2. <https://davedl.artstation.com/projects/VdWW9R>
3. Dafont.com

References:

1. Jukely, T. (2018, August 7). Who started rap? Retrieved from Four Over Four: <https://www.fouroverfour.com/culture/history-of-rap-hip-hop/>
2. DeWall, C., Pond, R., Campbell, W., & Twenge, J. (2011). Tuning in to psychological change: Linguistic markers of psychological traits and emotions over time in popular U.S. song lyrics. *American Psychological Association*, 200-207.
3. Liske, D. (2018, February). R NLP & Machine Learning: Lyric Analysis. Retrieved from DataCamp: <https://www.datacamp.com/tutorial/r-nlp-machine-learning>
4. Kaggle, R version, Affinity Designer version, Prezi

