

TITLE: A Look into Music Genres across the Years

AUTHORS: Hannah Duenke & Melinda Childs

FACULTY SPONSOR: Eric Friginal, Applied Linguistics and ESL

Introduction: Most people enjoy music and there are many different factors as to why a music genre appeals to groups of listeners. Melody/beat, message, or constant exposure to a genre (e.g., through radio or popularity in a geographic region) are often identified as factors that can influence listeners' musical preference. In this research, we focus on the analysis of song lyrics across genres by employing a quantitative, corpus-computational approach. Using the software Antconc (Anthony, 2012), a concordancer that automatically constructs an alphabetical listing of all the words used in a corpus of song lyrics, we aim to explore lexical frequencies and tokens in order to compare and contrast musical genres.

The primary purpose of this study is to compare four genres of music: Pop, Rock, R&B, and Country by examining lyrics used by the most popular artist in these genres (measured from Billboard charts) across time periods. We wanted to know if there are clear correlations between lexical distributions in each genre and the genre's popularity. Specifically, we pursued the following Research Questions:

- (1) Do listeners prefer music using higher word frequencies (e.g., more nouns and verbs) across genres or if music is something that is thought of as more an escape and not have to think too deeply about what is being presented to the audience?
- (2) Do fewer words (or shorter words) in song lyrics across genres influence popularity?
- (3) How do word frequencies, word counts, and the most frequent word types in songs compare with the popularity of their respective genres?
- (4) How have song lyrics evolved throughout the years? Does the increase/decrease of different parts of speech and filler words (e.g., 'yeah' or 'oh') help define this evolution?

Methodology: After picking 4 genres of music to compare and contrast, we looked up the top three songs from each genre from Billboard.com's Top 100 lists from each year and category using the years 2000-2010. Next, the lyrics from each of the top songs on the list were copied into text documents (.txt) and modified by eliminating unessential notes or correcting misspellings, forming our exploratory corpus of popular song lyrics within a 10-year spread. The corpus was processed by using Antconc to obtain data that can provide answers to our research questions. We tagged our listing of lexical items for part-of-speech (POS) and other linguistic categories through a spreadsheet. The frequencies of POS tags were averaged and compared in a chart to interpret results.

Results: Surprisingly, there was very little difference between word counts across genres in our exploratory corpus. This minimal difference also did not produce noteworthy distinctions across time periods. The most noticeable difference in our data can be found in adjective and verb uses in Country and R&B. Fillers, nouns, and pronouns maintained a consistently similar distribution throughout the four genres and also through the time periods. We expected pronouns (especially personal pronouns: *I, we, us, my, your*) to be more common in Country than other genres, but this

was not the case with our data. For token counts, Pop music consistently was the most popular genre and it also used more tokens. Country consistently was the least popular genre of the four, using the least amount of tokens across time periods.

Conclusion: Our goal was to analyze song lyrics across genres in terms of lexical variance as well as word counts and tokens across timeframes and to compare these data against record sales from 2000 to 2010 (from Billboard charts). We hypothesized that the higher the word count, the more popular in sales the genre as a whole would be. As briefly noted above, we found that, at least with our exploratory data, these factors did not seem to influence sales. We did notice that there was a slight difference in POS categories across genres. A more interesting difference in the number of tokens across genres per year can be further investigated in future studies (i.e., more quantitative and statistical studies).