Does Lyric Complexity Contribute to Listener Preferences?

Readability of Song Lyrics: A Trend Analysis

Alexa Walquist, Shruti Padole, Kahyun Choi

Introduction

Although music platforms such as Spotify or Apple Music examine a plethora of data points for their algorithms, there is a gap of knowledge regarding lyric complexities. Addressing this gap, **our study investigates lyric complexity using readability metrics and explores how this complexity is correlated with popularity, genre, and time.** We hope to determine what attracts listeners to certain songs through these trends and patterns.

Methods

Collection

- Gathered 4,862 songs-Billboard Year End Hot 100 chart (1965-2015)
- Genre metadata extracted from iTunes Search API
- Comments & interpretation of songs-songmeanings.com

Cleaning

- We removed foreign characters, converted all text to lower case, & removed digits/punctuation from song lyrics using Python functions
- Manually annotated lyrics & decided how many lines = a sentence

Complexity Metrics Used

- **Flesch Kincaid**- 0.39 x (words/sentences) + 11.8 x (syllables/words) – 15.59
- **SMOG**- 1.043 *√(polysyllabic words* (30/sentences)) + 3.1291
- **Gunning Fog**- 0.4 [(words/sentences) + 100 (complex words/words)]

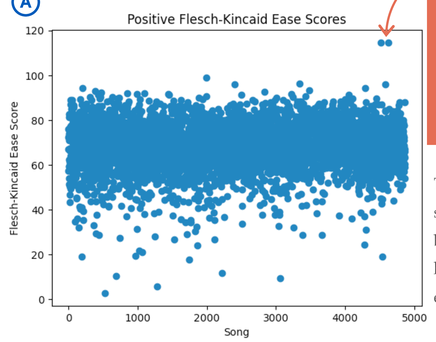
Parameter Determination

- Due to the lack of sentences in song lyrics, I tested various values of song lines to equal a sentence, and picked the one with the most similar scores to the manually annotated lyrics

Data Analysis

Score Distribution

Ⓐ



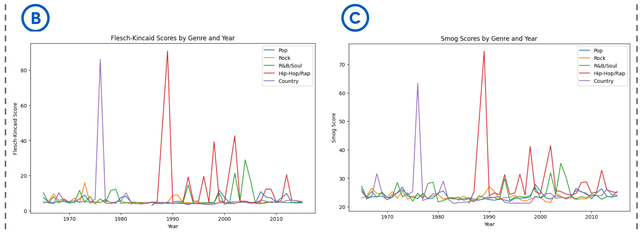
Corresponding Lyrics:

"joladadijoladadijoladadijoladadijoladadijoladadijoladadijoooohjoladadijoladadijoladadijoladadijoladadijoladadijoladadijooooh oooooooooooooh"

The range that texts should fall into is between 0-100; the higher the score, the easier it is to understand.

Genre & Time

Ⓑ Ⓒ



The majority of genres, with the exception of hip-hop/rap, fall within the ideal range for each metric consistently over time.

Ⓓ

	Average of Smog	Average of Gunning Fog	Average of Flesch Kincaid	Average of Flesch Kincaid Ease
Country	24.72	107.90	6.22	64.33
Hip-Hop/Rap	27.96	236.55	11.70	53.50
Pop	23.66	117.76	5.49	65.00
R&B/Soul	24.53	128.65	6.71	62.83
Rock	23.62	95.45	5.47	63.29

*The lower the Flesch Kincaid Ease score, the harder the text is to understand

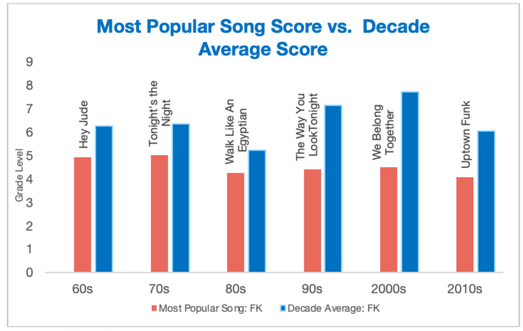
Ⓔ

	Count of Songs Charted
Country	261
Hip-Hop/Rap	479
Pop	1794
R&B/Soul	838
Rock	976

Hip-Hop/Rap is shown to have the **most complex lyrics** across the board, as each score is significantly higher than all genres. However, R&B/Soul and Country show slightly higher complexity scores than Pop and Rock in Smog and Flesch Kincaid. Pop and Rock songs are shown in Figure E to chart the most, indicating that **more complex texts are less popular.**

Popularity

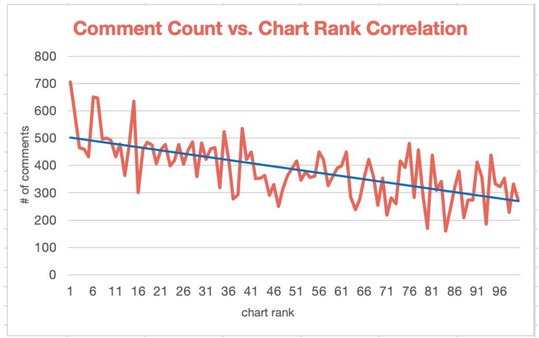
Ⓕ



*based on Flesch Kincaid metric

Each decade's average Flesch Kincaid score is consistently higher than each decade's most popular song, emphasizing that **easier readability could contribute to widespread popularity of a song.** This may be due to record labels/artists producing songs they think will reach the most audiences.

Ⓖ



Since all songs in the Billboard dataset are popular, it is unsurprising to see that they all have a relatively large number of comments on SongMeanings.com, however there still is a **strong correlation** observed between **increasing individual rank and decreasing number of comments.**

Ⓗ SPEARMAN CORRELATION : GUNNING FOG SCORE VS COMMENT COUNTS

```
[ ]  print("SCORE>>>>\n")
     corr, p = spearmanr(unique_df['gunning_fog_1'], unique_df['comment_count'])
     print(corr)
     # interpret the significance
     alpha = 0.05
     if p > alpha:
      print('Samples are uncorrelated (seem uncorrelated) p=%.3f' % p)
     else:
      print('Samples are correlated (reject H0) p=%.3f' % p)

     SCORE>>>>

     0.014354419897331699
     Samples are uncorrelated (seem uncorrelated) p=0.888
```

All metrics had similar conclusions; this suggests there are other factors contributing to song popularity and listener's preferences.

Limitations

Formulas dismiss unquantifiable complexity factors. They measure only *difficulty*, not content, organization, imagery, or tone, etc.[3]

Readability formulas aren't meant for song format. Songs are typically shorter in length than most written texts, and do not follow conventional grammatical rules or structure. They often contain a significant amount of figurative language and cultural references, which can make them more difficult to understand for readers who are not familiar with the context. [3]

Formulas do not distinguish audiences. They fail to take into account the differing purposes, maturity, or intelligence of readers. [3]

Conclusion

We observed that songs fell within the ideal range of understandability consistently over time, the most popular songs were simpler than the average charted songs, and Hip-Hop/Rap was the genre with the most complex lyrics. However, there is no meaningful correlation found between complexity score and popularity. These formulas identified polysyllabic and 'complex' words, but "I wave my hand" and "I waive my rights" have the same number of syllables with completely different levels of meaning, indicating that lyric meaning and its emotional appeal is more accurate and important to examine when listener preferences. In future research, I would like to compare charted songs and general songs to determine if popular songs are more complex than general songs, and use more advanced metrics to evaluate trends in regards to tone, style, and meaning of song lyrics.

Related Work

1. Kahyun Choi and J. Stephen Downie, "A Trend Analysis on Concreteness of Popular Song Lyrics," in Proceedings of the 6th International Digital Libraries for Musicology workshop, 2019.
2. Parmer, T., & Ahn, Y. Y. (2019). Evolution of the informational complexity of contemporary western music. arXiv preprint arXiv:1907.04292.
3. Radish, J. (n.d.). *Readability formulas have even more limitations than Klare discusses*. Retrieved March 11, 2023, from https://dl.acm.org/doi/pdf/10.1145/344599.344637