# Audio Features vs Lyric Language in Music

Julia Stelman

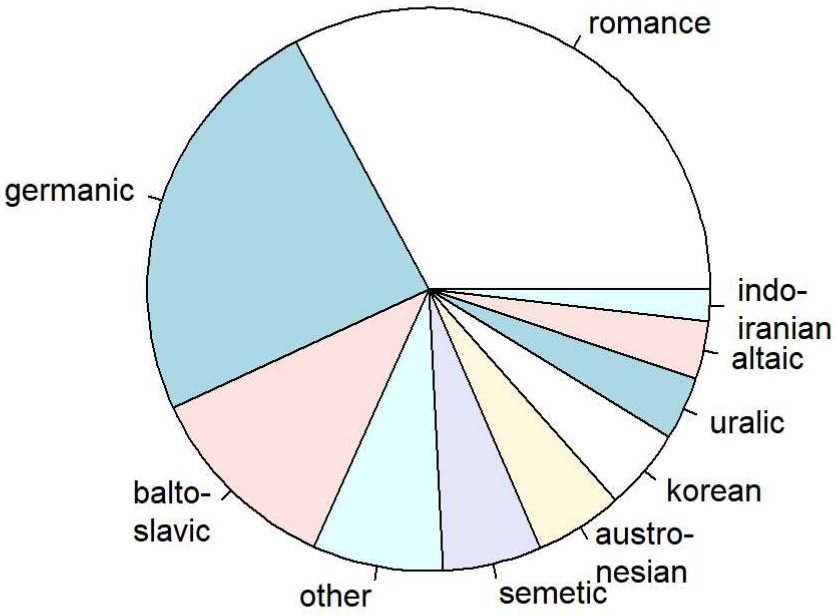
17/9/2019

Over four thousand songs in over forty different languages were used in this study. With help from the Spotify Web API (for audio feature collection), Musixmatch API (for lyric collection), and the python port of Google's language-detection library, available on Github, I was able to begin exploring the connections between lyric language and audio features in music. My list of songs came from published playlists on Spotify by the user *The Sounds of Spotify*. This user does a lot of music data mapping projects, and shares them on his site: http://everynoise.com (http://everynoise.com).

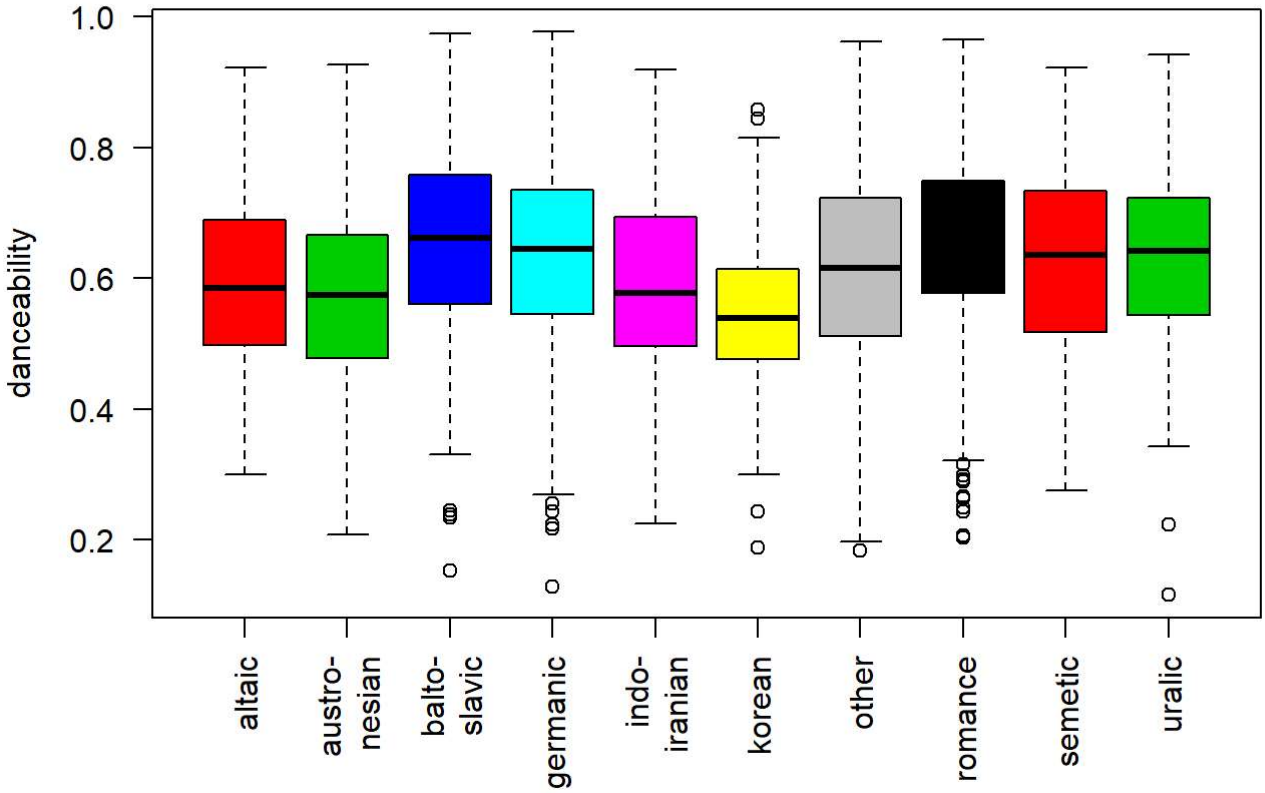
I hope to continue working on this study as time allows. It would be nice if I could expand my dataset. I definitely plan on conducting at least a few statistical tests to see how certain features follow different distributions in songs written and sung in different languages. I'm also interested in seeing if languages of the same families are any more likely to produce music with similar audio features than languages of different families.

Currently, linguists are actually debating over which language family contains Korean. Some are saying it belongs in the Austronesian language family, while others say its belongs in the Altaic language family with Japanese. Maybe, with what I find in future analysis, I will be able to weigh in. In the boxplots below, Korean is colored yellow.

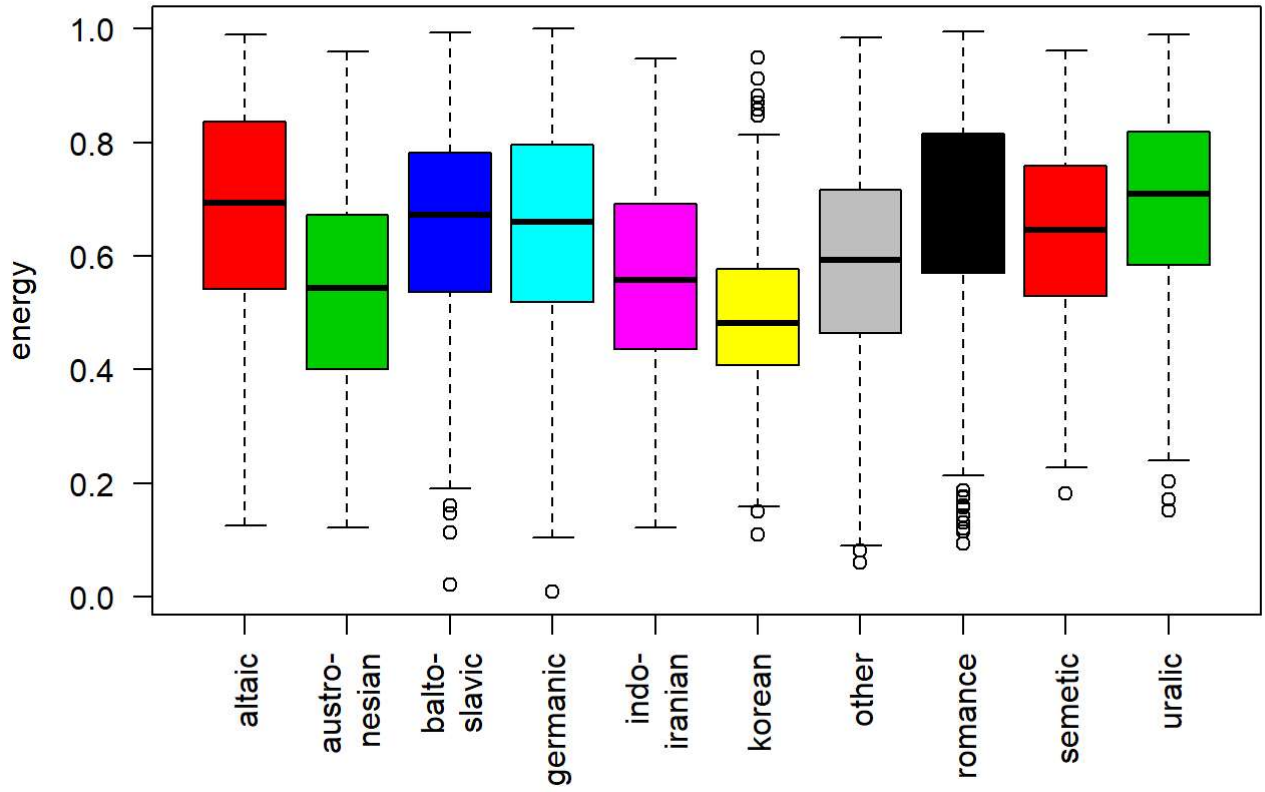
## Pie Chart of Language Family of Song Lyrics in the Sample



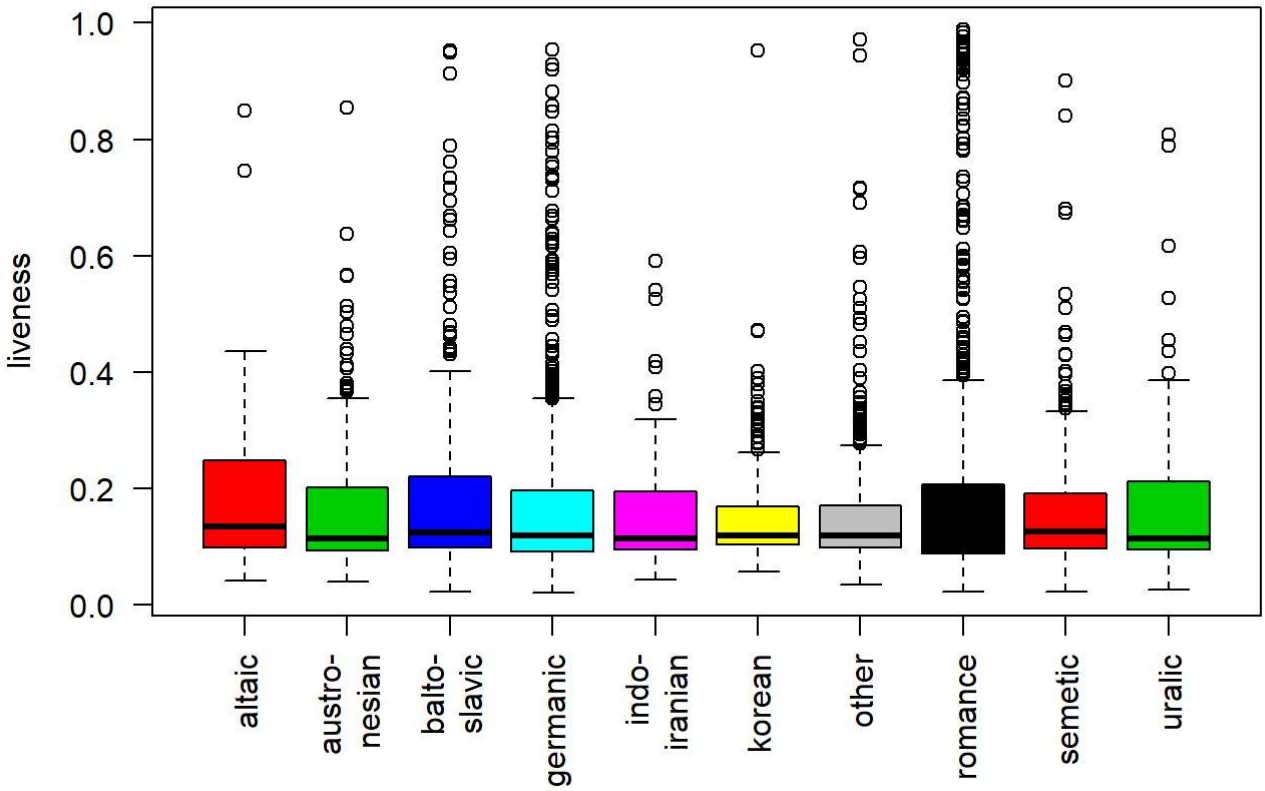
## Spotify Danceability Score vs Lyric Language Family



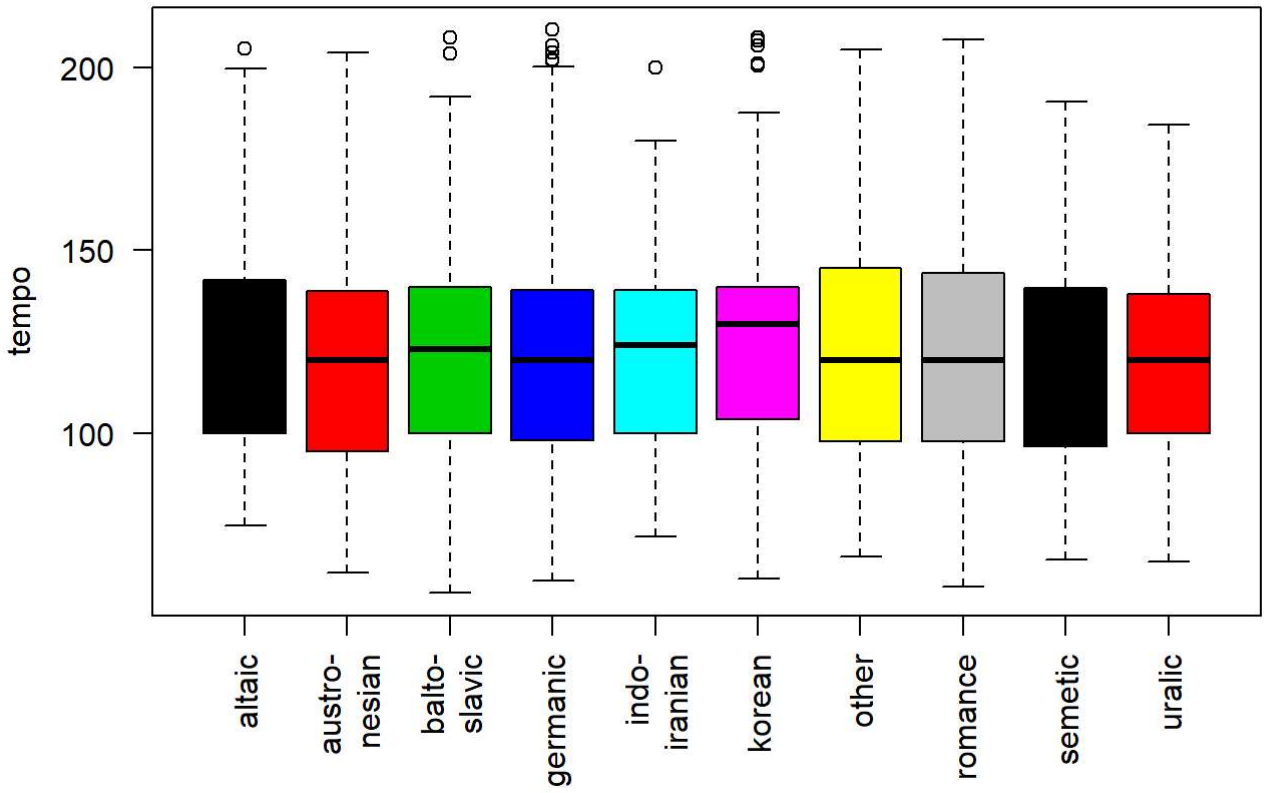

## Spotify Energy Score vs Lyric Language Family



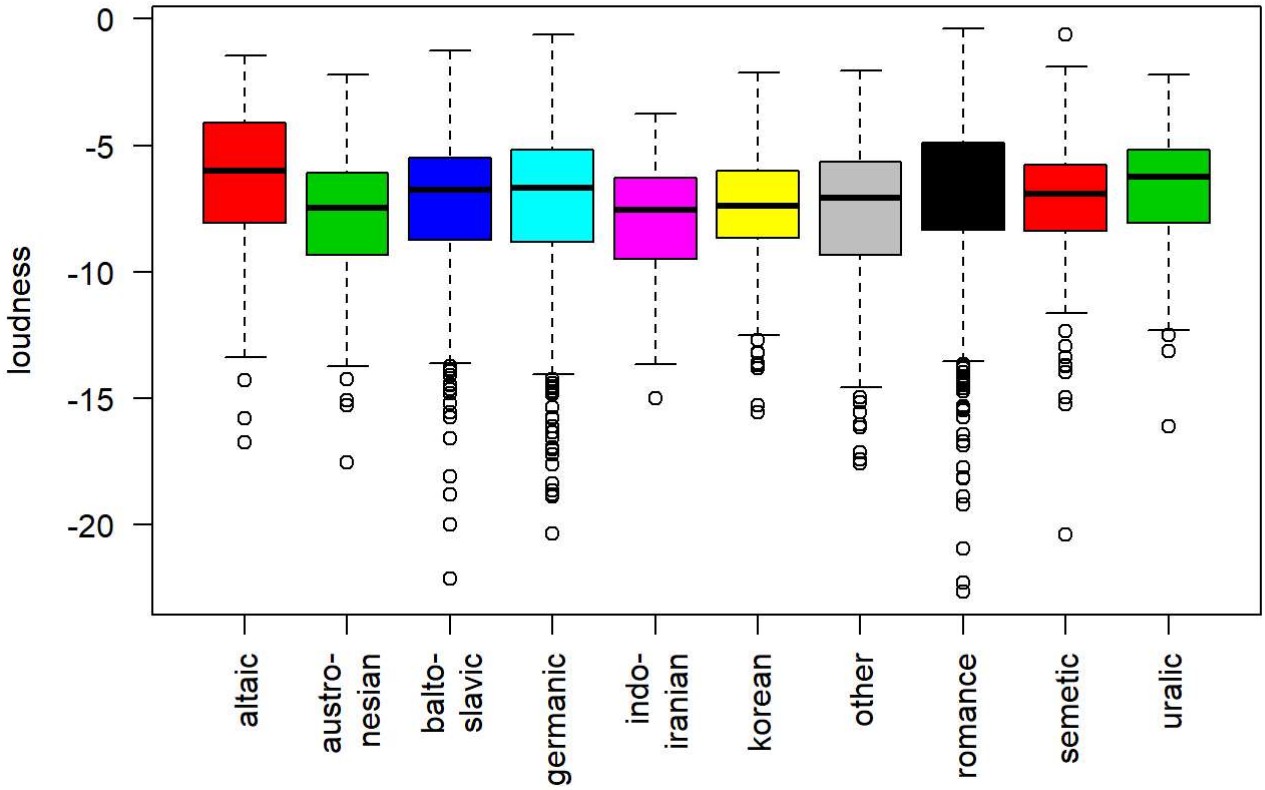
## Spotify Liveness Score vs Lyric Language Family



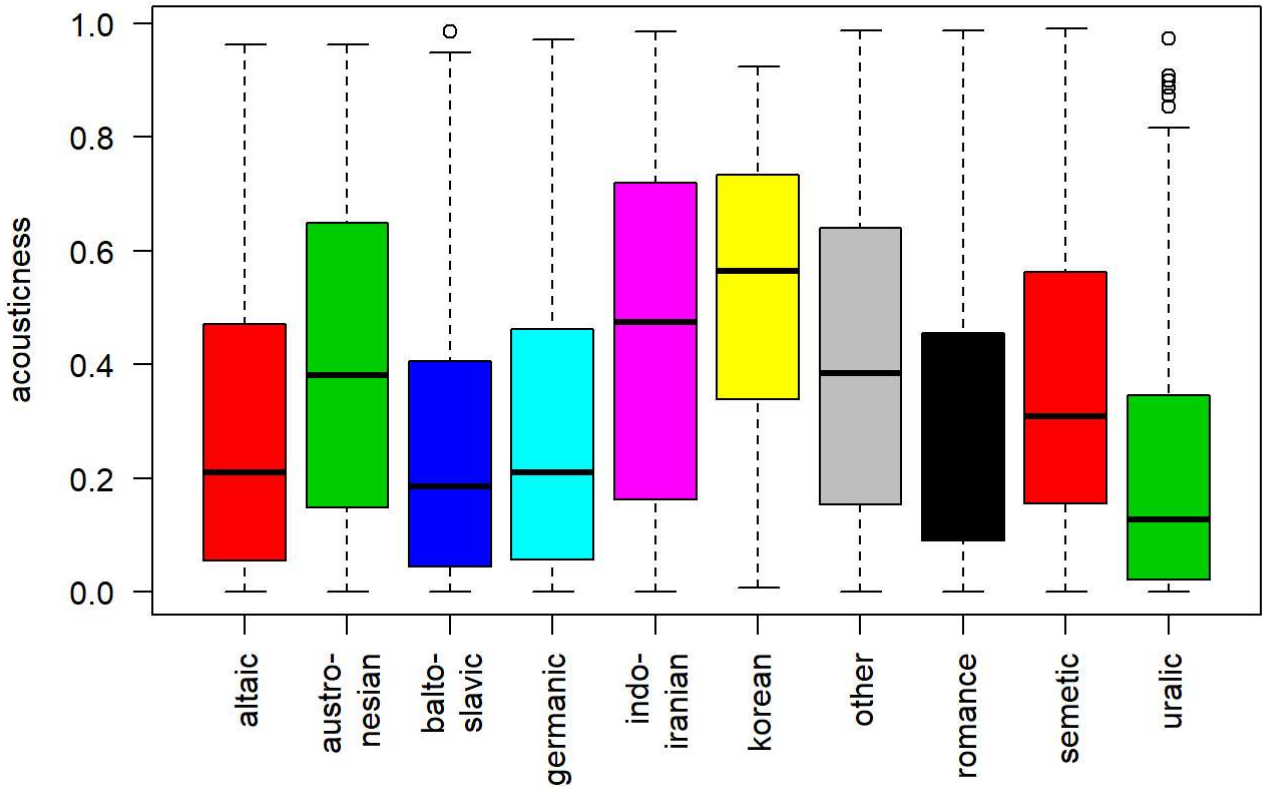

## Song Tempo vs Lyric Language Family



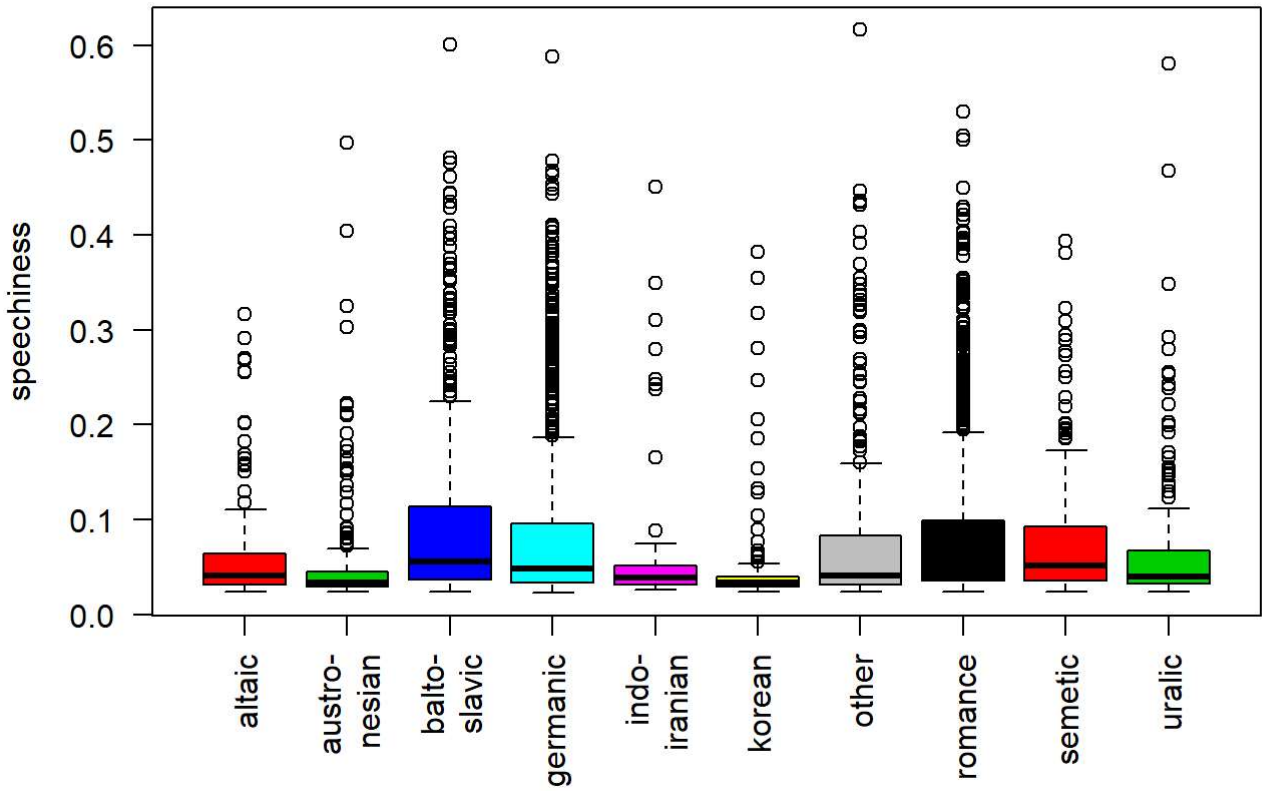
## Spotify Loudness Score vs Lyric Language Family




## Spotify Acousticness Score vs Lyric Language Family



## Spotify Speechiness Score vs Lyric Language Family




## Song Duration in ms vs Lyric Language Family