Lyrics, Topics, Musics and Their Mouse Clicks

Correlating Song Lyrics and Audio Feature Dimensions and How They Relate to the Popularity of Songs on Online Music Platforms

Isabella Czedik-Eysenberg[1], Oliver Wieczorek[2], Christoph Reuter[1]

[1]Musicological Department, University of Vienna; [2]Department of Sociology, especially Sociological Theory, University of Bamberg

Background

Do songs with a high danceability tend to deal with party topics? Does a distorted, rough sound indicate lyrics concerned with a Rock'n'Roll lifestyle and delinquency? And how does all of this relate to the popularity of music on online music platforms?

The relevance of lyrics as a complementary part of music next to audio features has repeatedly been demonstrated for tasks like musical genre classification (Neumayer and Rauber, 2007) and mood detection (Hu et al., 2009; Raschka, 2016). Lyrics features have also been shown to be able to predict hit songs, sometimes even outperforming audio features (Singhi & Brown, 2015). Research dealing with multimodal music data and metadata on a large scale helped to gain insight into the understanding of song lyrics by connecting them with user interpretations of songs (Choi, 2018; Choi et al., 2018). On the example of the metal music genre, it has been shown that the occurrence of textual topics that deal with harsh topics such as *brutal death*, *dystopia* and *satanism* are associated with audio dimensions that indicate a *hard/heavy* and *dark/gloomy* sound of music (Czedik-Eysenberg et al., 2019).

Research Questions & Aim

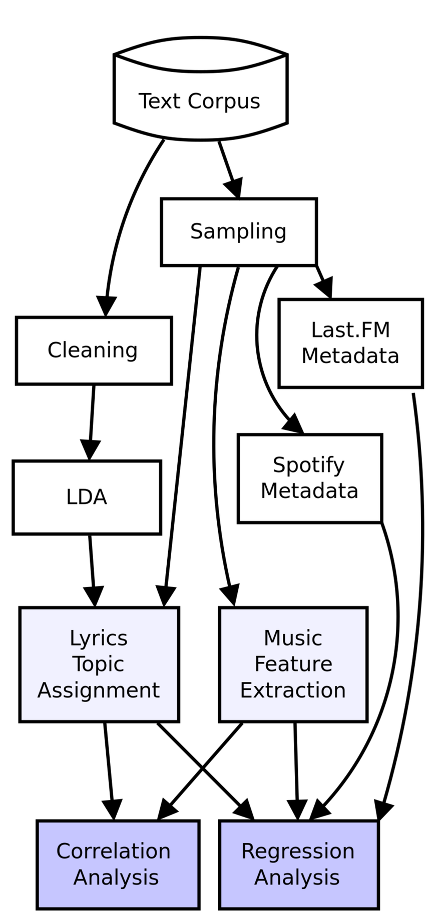
1. Offer an overview of the **lyrical topics** within popular music.
2. How are specific **music dimensions** (like e.g. *danceability* or *melodiousness*) associated with the occurrence of specific textual topics (like e.g. *love* or *politics*)?
3. Can these lyrical topics and audio dimensions play a role in predicting the **popularity** of music?

Data & Method

Topic Models

We retrieved a dataset of 944,553 song lyrics from 22 different genres via the API of *genius.com*. After **cleaning procedures** (restricting language to English, removing stopwords, stemming), the subsample included 771,663 texts.

We applied **latent Dirichlet allocation** (LDA, Blei et al., 2003) on this text corpus to construct a probabilistic **topic model**. Taking interpretability as well as Log-Perplexity and Log-U-Mass Coherence as goodness of fit-measures into consideration, a topic model including 22 topics was chosen (see below).



Metadata & Audio Features

Audio features for the songs were retrieved via the *Spotify* API (*acousticness - valence*).

Additional to these, further high-level music features were extracted using prediction models that were trained based on previous music perception experiments and available datasets (Czedik-Eysenberg et al., 2017; 2018; 2020; Aljanaki 2018). These models embedded audio features implemented in:

- *LibROSA* (McFee et al., 2015)
- *Essentia* (Bogdanov et al., 2013)
- *AudioCommons timbral models* (Pearce et al., 2017)

Statistical Processing & Popularity Modelling

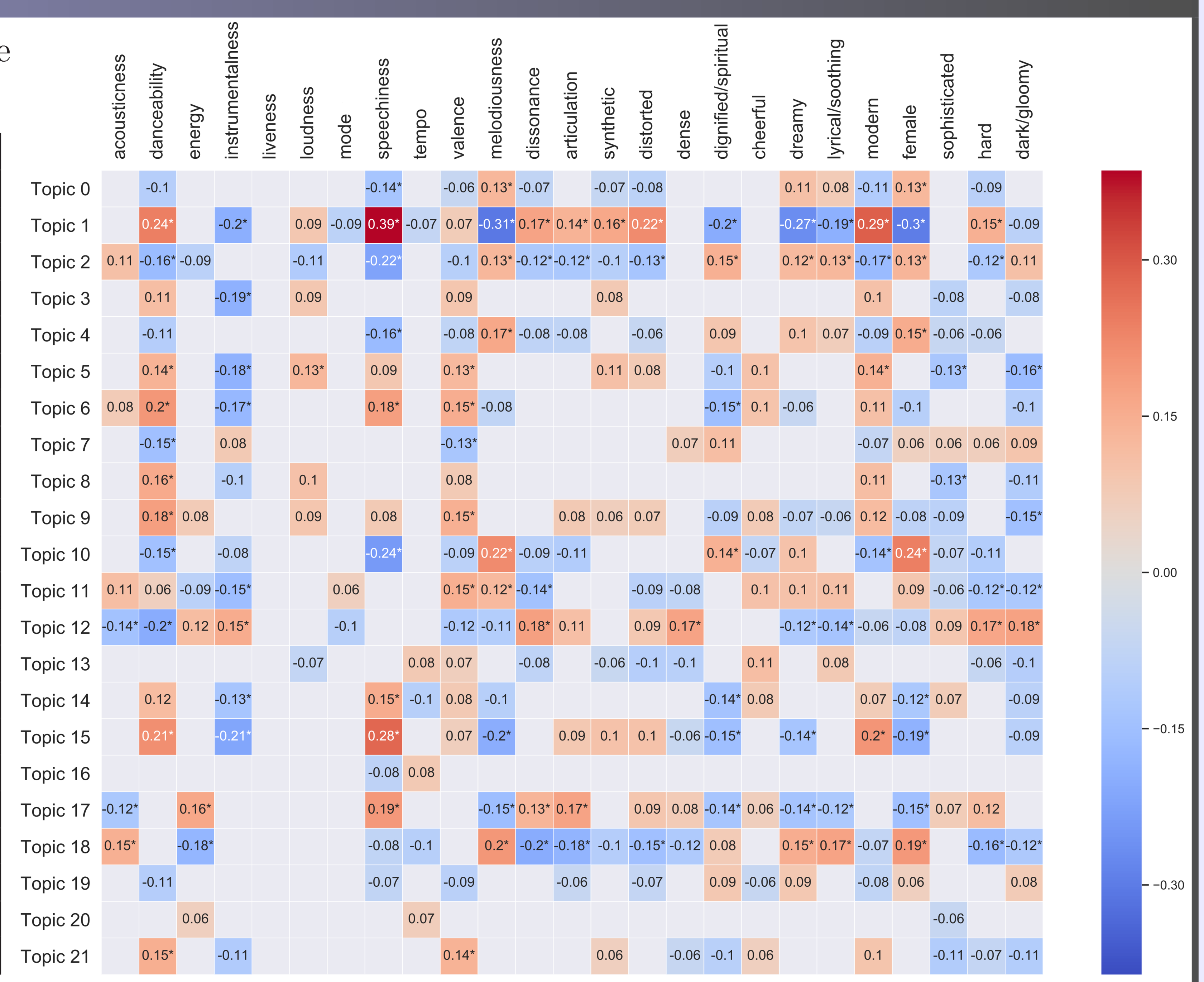
A **subsample** of 1070 songs was drawn and combined with audio features and **metadata** (e.g. *Last.FM* scrobbles, popularity on *Spotify*). We used Spearman's ρ to identify **correlations** between the topics retrieved and each of the high-level audio features. Based on the identified correlations, we further calculated the audio-textual fit for each song: $atf = \sum_{i=0}^{N} \sum_{j=0}^{M} topic_i * corr_{ij} * audio_j$

Finally, for modelling **popularity**, we chose the Spotify popularity measure to account for recent popularity and the number of Last.FM track listeners to measure long term popularity.

Results

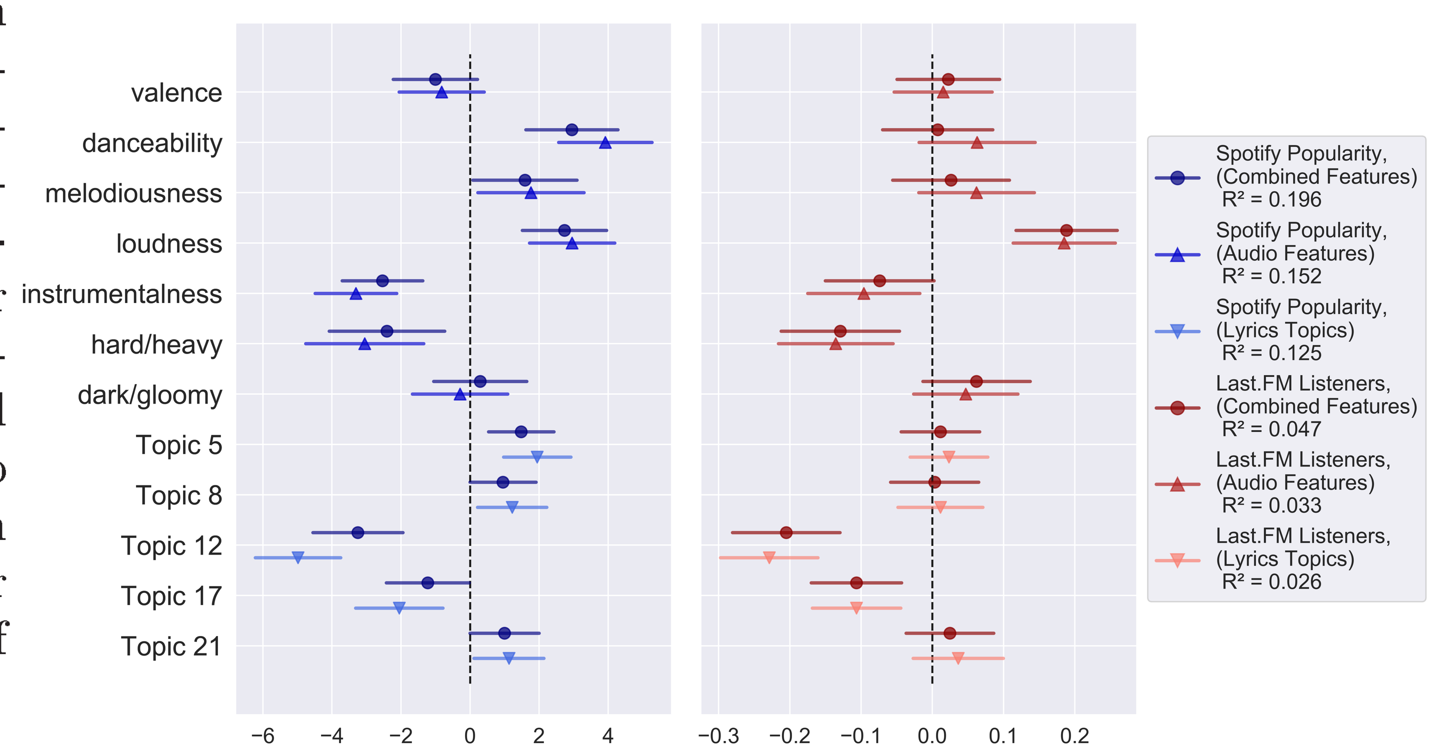
The resulting **topics** are shown below and their **correlations** with each audio feature on the right (displayed correlations: p < 0.5; *: significant after Bonferroni correction).

Topic	Interpretation	Salient Tokens/Stems (Top 10)
Topic 0	Places & Motion	away, way, run, home, gone, long, alon, walk, come, stay
Topic 1	Vulgar Language	nigga, fuck, bitch, shit, like, know, ass, caus, real, hit
Topic 2	Landscape/Visual Beauty	rain, sun, eye, sky, blue, wind, water, see, cold, fall
Topic 3	Smalltalk	like, hey, feel, okay, hello, bit, cool, say, look, act
Topic 4	Night (& Day)	time, night, day, light, wait, dream, tonight, last, shine, sleep
Topic 5	Exclamation	yeah, ayi, woah, huh, nobodi, woo, mmm, hoo, whoo, hmm
Topic 6	Abbreviated Language	goin, bout, keep, nothin, lookin, gettin, comin, wit, caus, talkin
Topic 7	Murder & Gory Topics	dead, kill, lie, blood, eye, breath, face, pain, head, hate
Topic 8	Contentment	know, babi, gon, wan, need, make, right, give, caus, tell
Topic 9	Party	danc, stop, rock, move, new, play, roll, parti, beat, bodi
Topic 10	Longing	never, know, see, life, feel, time, tri, live, could, noth
Topic 11	Talking/Dialogue	say, know, think, said, well, thing, friend, tell, would, could
Topic 12	Death/Sinister Topics	death, dark, world, rise, soul, fear, fight, end, life, burn
Topic 13	Regional & Seasonal	black, white, wild, town, countri, west, ohh, south, citi, sunday
Topic 14	Freestyle/Language Games	like, make, caus, kid, rap, rhyme, big, dog, see, kick
Topic 15	Money & Wealth	money, work, big, like, diamond, top, buy, bag, pull, car
Topic 16	Body/Physical Sensation	feel, keep, around, hold, turn, hand, hear, head, round, sound
Topic 17	Politics & Society	state, world, human, new, time, gener, right, law, self, natur
Topic 18	Love	love, heart, yuh, lover, true, nuh, nah, gyal, say, inna
Topic 19	Religion & Worship	god, lord, sing, heaven, name, jesu, soul, king, save, pray
Topic 20	Action	take, high, fire, ride, burn, fli, slow, low, side, drive
Topic 21	Flirt & Nicknames	girl, boy, pleas, bad, look, kiss, sweet, woman, crazi, ladi



Popularity Prediction

After controlling the distribution of our outcome variables, we applied **linear regression** to predict the Spotify popularity measure and **negative binomial regression** for the Last.FM listener count. Three models were calculated per outcome variable: a full model, a model comprising audio features only, and one focusing on topics. Overall, predictive power was limited, but better in case of Spotify popularity.



Furthermore, a weak negative correlation could be observed between *audio-textual fit* and *(log) Last.FM playcount*: $r = -0.143, p < 0.001$ (but not in case of the Spotify popularity: $r = -0.04, p = 0.19$)

Conclusion

In the joint appearance of textual topics and musical characteristics some significant characteristic patterns could be observed, e.g.:

- *Vulgar* language and songs dealing with *money* or *politics* tend to be characterised by a high *speechiness*.
- Songs dealing with *party* topics tend to be *danceable* and have a *bright, happy* sound (high *valence*).
- Lyrics dealing with *love* and *longing* tend to be accompanied by *melodious* and *calm* music.
- A *dissonant*, *hard* and *gloomy* sound of the music points to *sinister* and *gory* topics such as *death*.

We measured positive associations between popularity and *danceability*, *melodiousness*, *loudness*, and *party-* and *love-related* topics. Negative associations occurred with texts dealing with *death* and *politics* and a *hard/heavy*, *instrumental* sound.

References

Aljanaki, A. (2018). *Mid-level perceptual musical features.* Retrieved from osf.io/5aupt | **Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003).** Latent dirichlet allocation. *Journal of machine Learning research*, 3: 993-1022. | **Bogdanov, D. et al. (2013).** Essentia: An audio analysis library for music information retrieval. In *ISMIR'13*, pp. 493-98. | **Choi, K. (2018).** *Computational lyricology: quantitative approaches to understanding song lyrics and their interpretations.* PhD thesis, University of Illinois at Urbana-Champaign. | **Choi, K., Lee, J. H., Hu, X., & Downie, J. S. (2016).** Music subject classification based on lyrics and user interpretations. In *Proceedings of the Association for Information Science and Technology*, 53(1):1-10. | **Czedik-Eysenberg, I., Knauf, D., & Reuter, C. (2017).** "Hardness" as a semantic audio descriptor for music using automatic feature extraction. *INFORMATIK 2017.* | **Czedik-Eysenberg, I., Reuter, C., & Knauf, D. (2018).** Decoding the sound of "hardness" and "darkness" as perceptual dimensions of music. *ICMPC-ESCOM: Book of Abstracts.* Graz, pp. 112-13. | **Czedik-Eysenberg, I., Wieczorek, O., & Reuter, C. (2019).** "Warriors of the Word" - Deciphering Lyrical Topics in Music and Their Connection to Audio Feature Dimensions Based on a Corpus of Over 100,000 Metal Songs. *arXiv* preprint arXiv:1911.04952. | **Czedik-Eysenberg, I.** (2020). *Music dimension ratings and factors dataset.* https://doi.org/10.17605/OSF.IO/3VG4E | **Hu, X., Downie, J. S., & Ehmann, A. F. (2009).** Lyric text mining in music mood classification. In *ISMIR'09*, pp. 159-68. | **Neumayer, R. & Rauber, A. (2007).** Integration of text and audio features for genre classification in music information retrieval. In *European Conference on Information Retrieval*, pp. 724-727. | **Raschka, S. (2016).** Musicmood: Predicting the mood of music from song lyrics using machine learning. *arXiv* preprint arXiv:1611.00138. | **Singhi, A. & Brown, D. G. (2015).** Can song lyrics predict hits. In *Proceedings of the 11th International Symposium on Computer Music Multidisciplinary Research*, pp. 457-471.