
‘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

A PREPRINT

Isabella Czedik-Eysenberg

Department of Musicology

University of Vienna, Austria

isabella.czedik-eysenberg@univie.ac.at

Oliver Wieczorek

Department of Sociology, esp. Sociological Theory

University of Bamberg, Germany

oliver.wieczorek@uni-bamberg.de

Christoph Reuter

Department of Musicology

University of Vienna, Austria

christoph.reuter@univie.ac.at

September 20, 2021

ABSTRACT

We look into the connection between the musical and lyrical content of metal music by combining automated extraction of high-level audio features and quantitative text analysis on a corpus of 124.288 song lyrics from this genre. Based on this text corpus, a topic model was first constructed using Latent Dirichlet Allocation (LDA). For a subsample of 503 songs, scores for predicting perceived musical *hardness/heaviness* and *darkness/gloominess* were extracted using audio feature models. By combining both audio feature and text analysis, we (1) offer a comprehensive overview of the lyrical topics present within the metal genre and (2) are able to establish whether or not levels of *hardness* and other music dimensions are associated with the occurrence of particularly harsh (and other) textual topics. Twenty typical topics were identified and projected into a topic space using multidimensional scaling (MDS). After Bonferroni correction, positive correlations were found between musical *hardness* and *darkness* and textual topics dealing with ‘*brutal death*’, ‘*dystopia*’, ‘*archaisms and occultism*’, ‘*religion and satanism*’, ‘*battle*’ and ‘*(psychological) madness*’, while there is a negative associations with topics like ‘*personal life*’ and ‘*love and romance*’.

Keywords Metal Music · Topic Modeling · Latent Dirichlet Allocation · Audio Feature Extraction

1 Introduction

As audio and text features provide complementary layers of information on songs, a combination of both data types has been shown to improve the automatic classification of high-level attributes in music such as genre, mood and emotion [18, 14, 11, 12]. Multi-modal approaches interlinking these features offer insights into possible relations between lyrical and musical information (see [19, 16, 31]).

In the case of metal music, sound dimensions like loudness, distortion and particularly *hardness* (or *heaviness*)

play an essential role in defining the sound of this genre [1, 25, 17, 10]. Specific subgenres – especially doom metal, gothic metal and black metal – are further associated with a sound that is often described as *dark* or *gloomy* [21, 30].

These characteristics are typically not limited to the acoustic and musical level. In a research strand that has so far been generally treated separately from the audio dimensions, lyrics from the metal genre have come under relatively close scrutiny (cf. [9]). Topics typically ascribed to metal lyrics include *sadness*, *death*, *freedom*, *nature*, *occultism* or *unpleasant/disgusting objects* and are over-

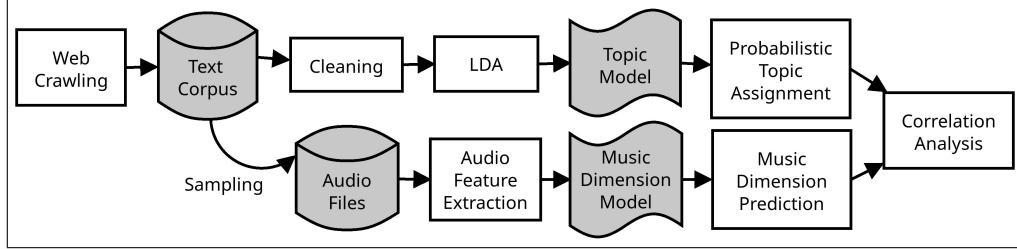


Figure 1: Processing steps of the approach illustrating the parallel analysis of text and audio features

all characterized as *harsh*, *gloomy*, *dystopian*, or *satanic* [23, 9, 28, 22, 5].

Until now, investigations on metal lyrics were limited to individual cases or relatively small corpora – with a maximum of 1,152 songs in [5]. Besides this, the relation between the musical and the textual domain has not yet been explored. Therefore, we examine a large corpus of metal song lyrics, addressing the following questions:

1. Which *topics* are present within the corpus of metal lyrics?
2. Is there a connection between characteristic musical dimensions like *hardness* and *darkness* and certain topics occurring within the textual domain?

2 Methodology

In our sequential research design, the distribution of textual topics within the corpus was analyzed using latent Dirichlet allocation (LDA). This resulted in a topic model, which was used for a probabilistic assignment of topics to each of the song documents. Additionally, for a subset of these songs, audio features were extracted using models for high-level music dimensions. The use of automatic models for the extraction of both text as well as musical features allows for scalability as it enables a large corpus to be studied without depending on the process of manual annotation for each of the songs. The resulting feature vectors were then subjected to a correlation analysis. Figure 1 outlines the sequence of the steps taken in processing the data. The individual steps are explained in the following subsections.

2.1 Text Corpus Creation and Cleaning

For gathering the data corpus, a web crawler was programmed using the Python packages Requests and BeautifulSoup. In total, 152,916 metal music lyrics were extracted from www.darklyrics.com.

Using Python’s langdetect package, all non-English texts were excluded. With the help of regular expressions, the texts were scanned for tokens indicating meta-

information, which is not part of the actual lyrics. To this end, a list of stopwords referring to musical instruments or the production process (e.g. ‘recorded’, ‘mixed’, ‘arrangement by’, ‘band photos’) was defined in addition to common stopwords. After these cleaning procedures, 124,288 texts remained in the subsample. For text normalization, stemming and lemmatization were applied as further preprocessing steps.

2.2 Topic Modelling via Latent Dirichlet Allocation

We performed a LDA [2] on the remaining subsample to construct a probabilistic topic model. The LDA models were created by using the Python library Gensim [24]. The lyrics were first converted to a bag-of-words format, and standard weighting of terms provided by the Gensim package was applied.

Log perplexity [6, p. 4] and log UMass coherence [26, p. 2] were calculated as goodness-of-fit measures evaluating topic models ranging from 10 to 100 topics. Considering these performance measures as well as qualitative interpretability of the resulting topic models, we chose a topic model including 20 topics – an approach comparable with [29]. We then examined the most salient and most typical words for each topic.

Moreover, we used the 1davis package to analyze the structure of the resulting topic space [27]. In order to do so, the Jensen-Shannon divergence between topics was calculated in a first step. In a second step, we applied multidimensional scaling (MDS) to project the inter-topic distances onto a two-dimensional plane. MDS is based on the idea of calculating dissimilarities between pairs of items of an input matrix while minimizing the strain function [4]. In this case, the closer the topics are located to one another on the two-dimensional plane, the more they share salient terms and the more likely a combination of these topics appear in a song.

2.3 High-Level Audio Feature Extraction

The high-level audio feature models used had been constructed in previous examinations [7, 8]. In those music perception studies, ratings were obtained for 212 music stimuli in an online listening experiment by 40 raters.

Table 1: Overview of the resulting topics found within the corpus of metal lyrics ($n = 124,288$) and their correlation to the dimensions *hardness* and *darkness* obtained from the audio signal (see section 3.2)

Topic	Interpretation	Salient Terms (Top 10)	Darkness ρ	Hardness ρ
1	<i>personal life</i>	know, never, time, see, way, take, life, feel, make, say	-0.195**	-0.262**
2	<i>sorrow & weltschmerz</i>	life, soul, pain, fear, mind, eye, lie, insid, lost, end	0.042	-0.002
3	<i>night</i>	dark, light, night, sky, sun, shadow, star, black, moon, cold	-0.082	-0.098*
4	<i>love & romance</i>	night, eye, love, like, heart, feel, hand, run, see, come	-0.196**	-0.273**
5	<i>religion & satanism</i>	god, hell, burn, evil, soul, lord, blood, death, satan, demon	0.11*	0.164**
6	<i>battle</i>	fight, metal, fire, stand, power, battl, steel, sword, burn, march	0.158**	0.159**
7	<i>brutal death</i>	blood, death, dead, flesh, bodi, bone, skin, cut, rot, rip	0.176**	0.267**
8	<i>vulgarity</i>	fuck, yeah, gon, like, shit, littl, head, girl, babi, hey	0.075	0.056
9	<i>archaisms & occultism</i>	shall, upon, thi, flesh, thee, behold, forth, death, serpent, thou	0.115*	0.175**
10	<i>epic tale</i>	world, time, day, new, end, life, year, live, last, earth	0.013	0.018
11	<i>landscape & journey</i>	land, wind, fli, water, came, sky, river, high, ride, mountain	-0.067	-0.125*
12	<i>struggle for freedom</i>	control, power, freedom, law, nation, rule, system, work, peopl, slave	0.095*	0.076
13	<i>metaphysics</i>	form, space, exist, beyond, within, knowledg, shape, mind, circl, sorc	0.066	0.064
14	<i>domestic violence</i>	kill, mother, children, pay, child, live, anoth, father, name, innoc	0.060	0.070
15	<i>dystopia</i>	human, race, disease, breed, destruct, machin, mass, seed, destroy, earth	0.191**	0.240**
16	<i>mourning rituals</i>	ash, word, dust, stone, speak, weep, smoke, breath, tongu, funer	0.031	0.015
17	<i>(psychological) madness</i>	mind, twist, brain, mad, self, half, mental, terror, urg, obsess	0.153*	0.107*
18	<i>royal feast</i>	king, rain, drink, fall, crown, sun, rise, bear, wine, color	-0.046	-0.031
19	<i>Rock'n'Roll lifestyle</i>	rock, roll, train, addict, explod, wreck, shock, chip, leagu, raw	0.032	0.038
20	<i>disgusting things</i>	anim, weed, ill, fed, maggot, origin, worm, incest, object, thief	0.075	0.064

Based on this ground truth, prediction models for the automatic extraction of high-level music dimensions – including the concepts of perceived *hardness/heaviness* and *darkness/gloominess* in music – had been trained using machine learning methods. In a second step, the model obtained for *hardness* had been evaluated using further listening experiments on a new unseen set of audio stimuli [8]. The model has been refined against this backdrop, resulting in an R^2 value of 0.80 for *hardness/heaviness* and 0.60 for *darkness/gloominess* using five-fold cross-validation.

The resulting models embedded features implemented in LibROSA [15], Essentia [3] as well as the timbral models developed as part of the AudioCommons project [20].

2.4 Investigating the Connection between Audio and Text Features

Finally, we drew a random sample of 503 songs and used Spearman’s ρ to identify correlations between the topics retrieved and the audio dimensions obtained by the high-level audio feature models. We opted for Spearman’s ρ since it does not assume normal distribution of the data, is less prone to outliers and zero-inflation than Pearson’s r .

Bonferroni correction was applied in order to account for multiple-testing.

3 Results

3.1 Textual Topics

Table 1 displays the twenty resulting topics found within the text corpus using LDA. The topics are numbered in descending order according to their prevalence (weight) in the text corpus. For each topic, a qualitative interpretation is given along with the 10 most salient terms¹.

The salient terms of the first topic – and in parts also the second – appear relatively generic, as terms like e.g. ‘know’, ‘never’, and ‘time’ occur in many contexts. However, the majority of the remaining topics reveal distinct lyrical themes described as being characteristic for the metal genre. ‘*Religion & satanism*’ (topic #5) and descriptions of ‘*brutal death*’ (topic #7) can be considered as being typical for black metal and death metal respectively, whereas ‘*battle*’ (topic #6), ‘*landscape & journey*’ (topic #11), ‘*struggle for freedom*’ (topic #12), and ‘*dystopia*’ (topic #15), are associated with power metal and other metal subgenres.

¹Note that the terms are presented in their stemmed form (e.g. ‘fli’ instead of ‘fly’ or ‘flying’).

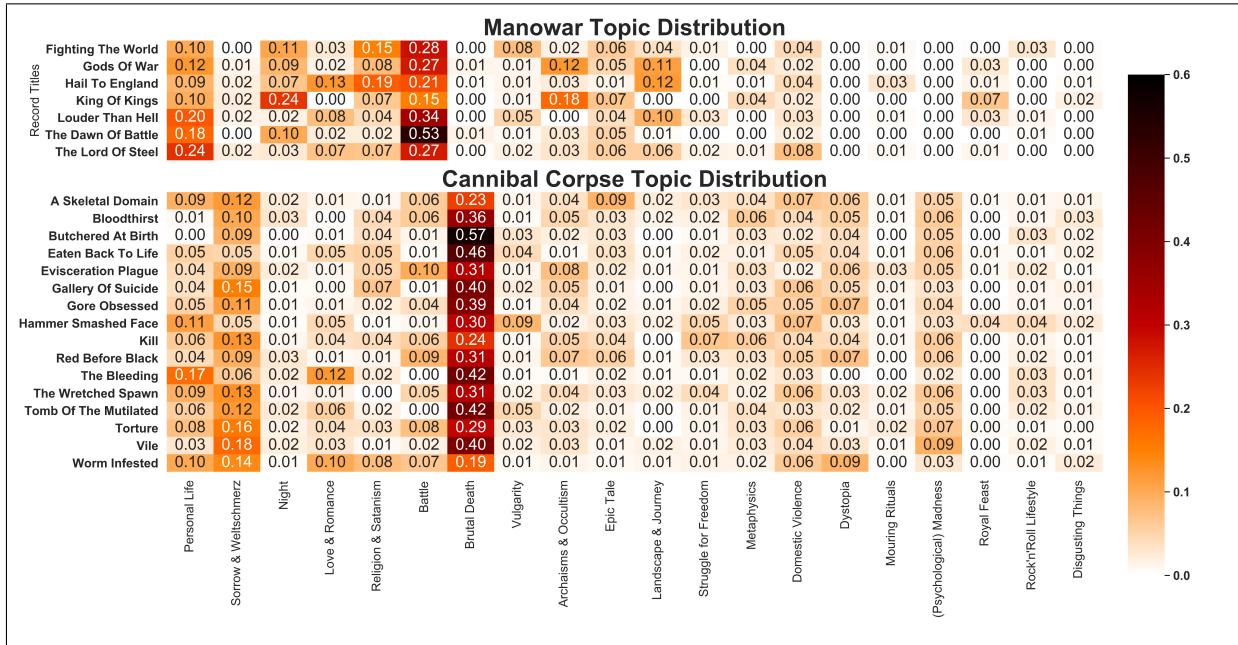


Figure 2: Comparison of the topic distributions for all included albums by the bands Manowar and Cannibal Corpse showing a prevalence of the topics 'battle' and 'brutal death' respectively

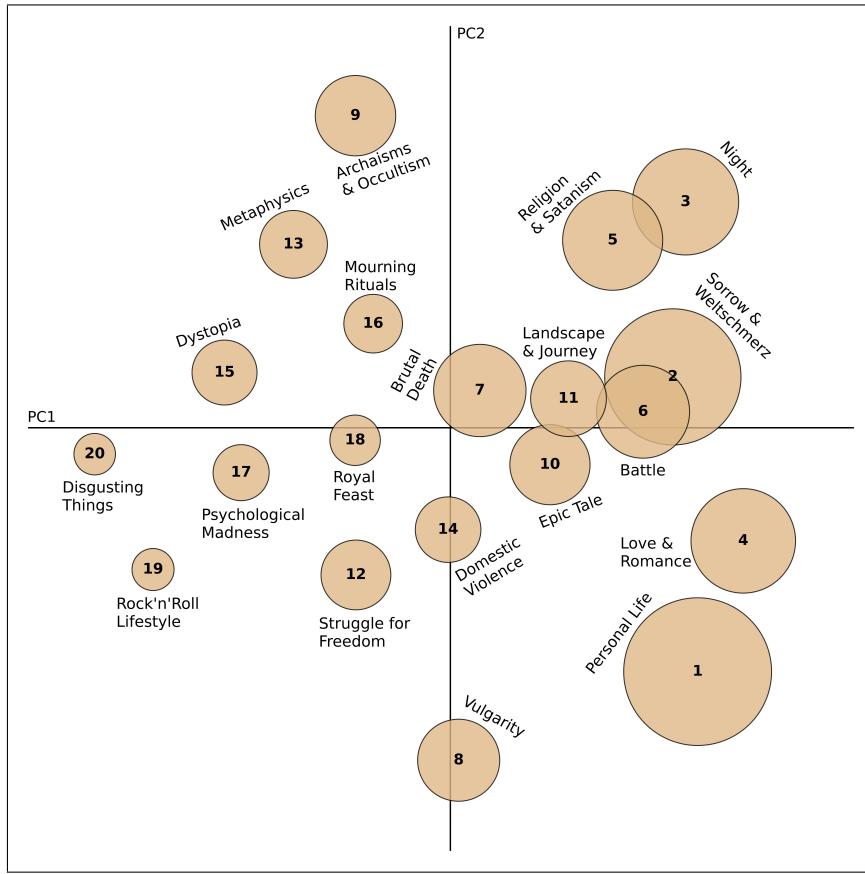


Figure 3: Topic configuration obtained via multidimensional scaling. The radius of the circles is proportional to the percentage of tokens covered by the topics (topic weight).

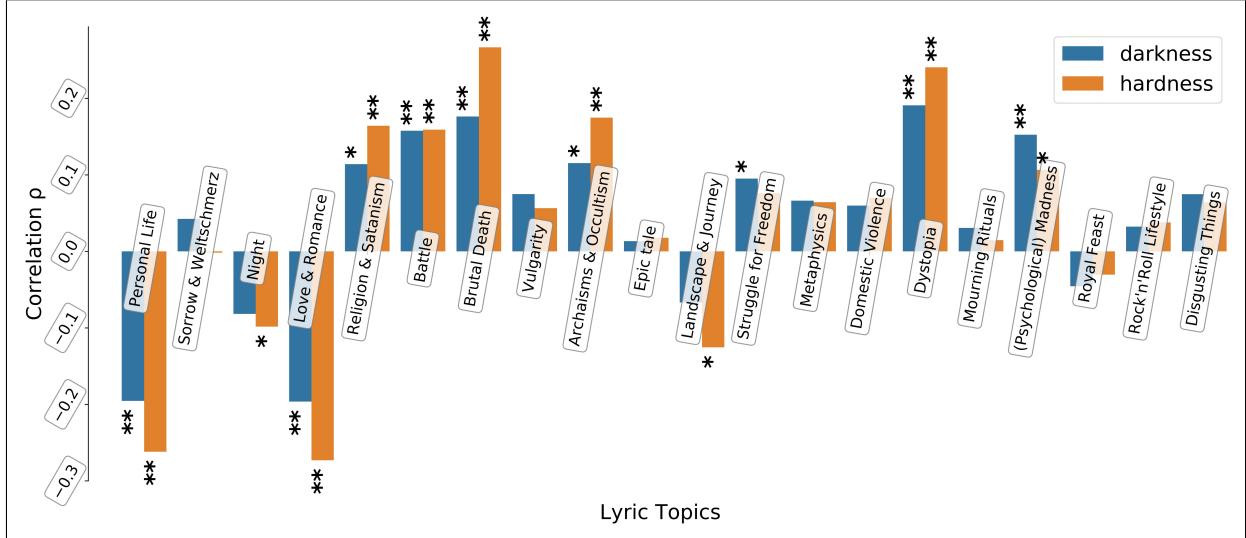


Figure 4: Correlations between lyrical topics and the musical dimensions *hardness* and *darkness*; *: $p < 0.05$, **: $p < 0.00125$ (Bonferroni-corrected significance level)

This is highlighted in detail in Figure 2. Here, the topic distributions for two exemplary bands contained within the sample are presented. For these heat maps, data has been aggregated over individual songs showing the topic distribution at the level of albums over a band's history. The examples chosen illustrate the dependence between textual topics and musical subgenres. For the band Manowar, which is associated with the genre of heavy metal, power metal or true metal, a prevalence of topic #6 ('battle') can be observed, while a distinctive prevalence of topic #7 ('brutal death') becomes apparent for Cannibal Corpse – a band belonging to the subgenre of death metal.

Within the topic configuration obtained via multidimensional scaling (see Figure 3), two latent dimensions can be identified. The first dimension (PC1) distinguishes topics with more common wordings on the right hand side from topics with less common wording on the left hand side. This also correlates with the weight of the topics within the corpus. The second dimension (PC2) is characterized by a contrast between transcendent and sinister topics dealing with *occultism*, *metaphysics*, *satanism*, *darkness*, and *mourning* (#9, #3, #5, #13, and #16) at the top and comparatively shallow content dealing with *personal life* and *Rock'n'Roll lifestyle* using a rather mundane or *vulgar* vocabulary (#1, #8, and #19) at the bottom. This contrast can be interpreted as '*otherworldliness / individual-transcending narratives*' vs. '*worldliness / personal life*'.

3.2 Correlations with Musical Dimensions

In the final step of our analysis, we calculated the association between the twenty topics discussed above and the

two high-level audio features *hardness* and *darkness* using Spearman's ρ . The results are visualized in Figure 4 and the ρ values listed in table 1.

Significant positive associations can be observed between musical *hardness* and the topics '*brutal death*', '*dystopia*', '*archaisms & occultism*', '*religion & satanism*', and '*battle*', while it is negatively linked to relatively mundane topics concerning '*personal life*' and '*love & romance*'. The situation is similar for *dark/gloomy* sounding music, which in turn is specifically related to themes such as '*dystopia*' and '*(psychological) madness*'. Overall, the strength of the associations is moderate at best, with a tendency towards higher associations for *hardness* than *darkness*. The strongest association exists between *hardness* and the topic '*brutal death*' ($\rho = 0.267$, $p < 0.01^2$).

4 Conclusion and Outlook

Applying the example of metal music, our work examined the textual topics found in song lyrics and investigated the association between these topics and high-level music features. By using LDA and MDS in order to explore prevalent topics and the topic space, typical text topics identified in qualitative analyses could be confirmed and objectified based on a large text corpus. These include e.g. *satanism*, *dystopia* or *disgusting objects*. It was shown that musical *hardness* is particularly associated with harsh topics like '*brutal death*' and '*dystopia*', while it is negatively linked to relatively mundane topics concerning personal life and love. We expect that even stronger correlations could be found for metal-specific topics when including

²The p values fall below the Bonferroni-corrected significance level.

³The previously established ground truth for the creation of the audio feature models included far more genres. Therefore, it can be assumed that our metal sample covers only a small fraction of the *darkness/hardness* range.

more genres covering a wider range of *hardness/darkness* values³.

Therefore, we suggest transferring the method to a sample including multiple genres. Moreover, an integration with metadata such as genre information would allow for the testing of associations between topics, genres and high-level audio features. This could help to better understand the role of different domains in an overall perception of genre-defining attributes such as *hardness*.

References

- [1] Berger, H. M. and Fales, C. (2005). ‘Heaviness’ in the Perception of Heavy Metal Guitar Timbres: The Match of Perceptual and Acoustic Features over Time. In: *Greene, P. and Porcello, T. (eds), Wired for Sound: Engineering and Technologies in Sonic Cultures*. Middletown: Wesleyan University Press, pp. 181-197.
- [2] Blei, D. M., Ng, A. Y. and Jordan, M. I. (2003). Latent dirichlet allocation, *Journal of machine Learning research*, 3: 993-1022.
- [3] Bogdanov, D., Wack, N., Gómez Gutiérrez, E., Guallati, S., Boyer, H., Mayor, O., Roma Trepat, G., Salamón, J., Zapata González, J.R. and Serra, X. (2013). Essentia: An audio analysis library for music information retrieval, *14th International Conference on Music Information Retrieval (ISMIR)*, Curitiba, Brazil, November 2013.
- [4] Borg, I. and Groenen, P. (2003). Modern multidimensional scaling: Theory and applications, *Journal of Educational Measurement*, 40: 277-280.
- [5] Cheung, J. O. and Feng, D. (2019). Attitudinal meaning and social struggle in heavy metal song lyrics: a corpus-based analysis, *Social Semiotics*: 1-18.
- [6] Coleman, C. A. Seaton, D. T., and Chuang, I. (2015). Probabilistic use cases: Discovering behavioral patterns for predicting certification, *Proceedings of the Second ACM Conference on Learning @ Scale*, Vancouver, Canada, March 2015.
- [7] Czedik-Eysenberg, I., Knauf, D. and Reuter, C., (2017). "Hardness" as a semantic audio descriptor for music using automatic feature extraction. In: *Eibl, M. and Gaedke, M. (Hrsg.), INFORMATIK 2017*. Gesellschaft für Informatik, Bonn, pp. 101-110. DOI: 10.18420/in2017_06
- [8] Czedik-Eysenberg, I., Reuter, C., and Knauf, D. (2018). Decoding the sound of “hardness” and “darkness” as perceptual dimensions of music. *ICMPC-ESCOM: Book of Abstracts*. Graz, pp. 112-13.
- [9] Farley, H. (2016). Demons, devils and witches: the occult in heavy metal music. In: *Bayer, G. (ed), Heavy metal music in Britain*. London: Routledge, pp. 85-100.
- [10] Herbst, J.-P. (2017). Historical development, sound aesthetics and production techniques of the distorted electric guitar in metal music, *Metal Music Studies*, 3: 23-46.
- [11] Hu, X. and Downie, J. S. (2010). Improving mood classification in music digital libraries by combining lyrics and audio, *Proceedings of the 10th annual joint conference on Digital libraries*, Surfer’s Paradise, Australia, June 2010.
- [12] Kim, Y. E., Schmidt, E. M., Migneco, R., Morton, B. G., Richardson, P., Scott, J., Speck, J.A. and Turnbull, D. (2010). Music emotion recognition: A state of the art review, *Proceedings of the 11th International Conference on Music Information Retrieval (ISMIR)*, Utrecht, Netherlands, August 2010.
- [13] Lau, J. H., Newman, D. and Baldwin, T. (2014). Machine reading tea leaves: Automatically evaluating topic coherence and topic model quality, *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, Gothenburg, Sweden, April 2014.
- [14] Laurier, C., Grivolla, J. and Herrera, P. (2008). Multimodal music mood classification using audio and lyrics, *Seventh International Conference on Machine Learning and Applications*, San Diego, CA, December 2018.
- [15] McFee, B., Raffel, C., Liang, D., Ellis, D.P.W., McVicar, M., Battenberg, E. and Nieto, O. (2015). librosa: Audio signal analysis in python, *Proceedings of the 14th python in science conference*, Austin, TX, December 2015.
- [16] McVicar, M., Freeman, T. and De Bie, T. (2011). Mining the correlation between lyrical and audio features and the emergence of mood, *Proceedings of the 12th International Conference on Music Information Retrieval (ISMIR)*, Miami, FL, October 2011.
- [17] Mynett, M. (2013). *Contemporary metal music production*. Ph.D. thesis, University of Huddersfield.
- [18] Neumayer, R. and Rauber, A. (2007). Integration of text and audio features for genre classification in music information retrieval, *European Conference on Information Retrieval*, Rome, Italy, April 2007.
- [19] Nichols, E., Morris, D., Basu, S. and Raphael, C. (2009). Relationships between lyrics and melody in popular music, *Proceedings of the 11th International Conference on Music Information Retrieval (ISMIR)*, Kobe, Japan, October 2009.
- [20] Pearce, A., Brookes, T. and Mason, R. (2017). Timbral attributes for sound effect library searching, *Audio Engineering Society Conference on Semantic Audio*, Erlangen, Germany, June 2017.
- [21] Phillips, W. and Cogan, B. (2009). *Encyclopedia of heavy metal music*. Westport: Greenwood Press.
- [22] Podoshen, J. S., Venkatesh, V. and Jin, Z. (2014). Theoretical reflections on dystopian consumer culture: Black metal, *Marketing Theory*, 14: 207-27.

- [23] Purcell, N. J. (2015). *Death metal music: The passion and politics of a subculture*. Jefferson: McFarland & Company.
- [24] Rehurek, R. and Sojka, P. (2010). Software framework for topic modelling with large corpora, *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, Valetta, Malta, May 2010.
- [25] Reyes, I. (2008). *Sound, Technology, and interpretation in Subcultures of Heavy Music Production*. Ph.D. thesis, Pittsburgh University.
- [26] Röder, M., Both, A. and Hinneburg, A. (2015). Exploring the space of topic coherence measures, *Proceedings of the eighth ACM international conference on Web search and data mining*, Shanghai, China, February 2015.
- [27] Sievert, C. and Shirley, K. (2014). LDAvis: A method for visualizing and interpreting topics, *Proceedings of the workshop on interactive language learning, visualization, and interfaces*. Baltimore, MD, June 2014.
- [28] Taylor, L. W. (2016). Images of human-wrought despair and destruction: Social critique in British apocalyptic and dystopian metal. In: Bayer, G. (ed), *Heavy metal music in Britain*. London: Routledge, pp. 101-22.
- [29] Viola, L. and Verheul, J. (2019). Mining ethnicity: Discourse-driven topic modelling of immigrant discourses in the USA, 1898–1920, *Digital Scholarship in the Humanities*, fqz068, DOI: 10.1093/lhc/fqz068.
- [30] Yavuz, M. S. (2017). ‘Delightfully depressing’: Death/doom metal music world and the emotional responses of the fan, *Metal Music Studies*, 3: 201-218.
- [31] Yu, Y., Tang, S., Raposo, F. and Chen, L. (2019). Deep cross-modal correlation learning for audio and lyrics in music retrieval, *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 15: 20-21.