

Sentiment analysis of corona-musicking online reveals bifurcation of pandemic coping strategies

Niels Chr. Hansen* [1] and Rebekah Baglini [2]

[1] Aarhus Institute of Advanced Studies, Aarhus University, Denmark; Center for Music in the Brain, Aarhus University & Royal Academy of Music Aarhus/Aalborg, Denmark, [2] School of Communication and Culture, Aarhus University, Denmark; Interacting Minds Centre, Aarhus University, Denmark
nchansen@aias.au.dk

Disciplinary background A. Music Psychology. When a sweeping pandemic forced social participation into hibernation in early 2020, musical creators and consumers moved their activities online, embracing emerging technology and inventing a stylistically diverse universe of coronamusic (Hansen, 2021; Hansen et al., 2021). Interest in corona-themed music became the foremost predictor of music-aided psychological coping with a functional bifurcation between those experiencing negative and positive emotions: the former used music for self-directed emotion regulation whereas the latter used it as a proxy for social interaction (Fink et al., 2021). Indeed, approach coping has been linked with positive affect during lockdown (Eden et al., 2020), and humor, joy, and togetherness dominated anecdotal media reports about pandemic music-making (Hansen et al., 2021). Yet, quantitative investigations of positivity bias and functional bifurcation in coronamusic repertoires and in text- and video-based musical participation online are absent.

Disciplinary background B. Linguistics. In mapping psychological coping, social media data are complementary to self-report surveys in detecting broader trends in behavioral patterns at national and global levels on a more granular timescale. Such data types are multifaceted, including information about social networks, engagement (e.g., streams, likes, shares), and user-generated content (e.g., profiles, comments, posts). Natural Language Processing (NLP) offers adequate tools for collecting, analyzing, and interpreting large corpora of text-based user data from online sources in real-time as events—such as the coronavirus pandemic—unfold (Liu et al., 2021). Although NLP has been widely applied to research on digital behavior, its full potential for studying musical phenomena remains to be seen.

Abstract

To investigate if and how key findings from music-psychological self-report questionnaires manifest in participatory corona-musicking online during pandemic lockdown.

Sentiment in text corpora sourced from Twitter, Reddit, YouTube, and public news media was quantified using NLTK's Vader Analyzer (Hutto & Gilbert, 2014) and Sentiwordnet (Baccianella et al., 2010): specifically, (i) non-music-related ($n=16,619,492$) and music-related ($n=205,912$) COVID-19-themed tweets from March-May 2020 (Qazi, Imran, & Ofli, 2020); (ii) 119,926 comments posted to the "ListenToThis" and "LetsTalkMusic" subreddits during March-May 2019 and 2020; (iii) YouTube comments ($n=2*63,393$) posted in response to 329 English-language coronamusic videos matched with non-coronamusic controls; (iv) transcribed lyrics from some of these videos; and (v) coronamusic-related news coverage from the Coronavirus subset of the NOW corpus (English-Corpora.org, n.d). Valence was, moreover, obtained from the Spotify API and compared between 575,254 unique tracks from 9,486 COVID-19-themed Spotify playlists with >1 followers and a 3,706,388-track control corpus from Music Streaming Sessions Dataset (Brost et al., 2019).

Mean sentiment scores were significantly higher for music-related COVID-19 tweets—both with ($t(211553)=120.5$, $p<.0001$) and without retweets ($t(56025)=50.3$, $p<.0001$). Reddit comments from 2020 either scored higher (“LetsTalkMusic”, $U=342337948.5$, $p=.0096$) or lower (“ListenToThis”, $U=536186459$, $p<.0001$) compared to 2019, depending on topical focus. COVID-19-themed Spotify playlists, moreover, exhibited significantly higher valence than controls ($t(8826.3)=21.5$, $p<.0001$), and k-means clustering on audio features suggested two distinct categories of low-energy/low-valence “chill” and high-energy/high-valence “party” playlists. Sentiment in target YouTube comments ($p=.073$) and news articles ($p=.058$) only differed marginally non-significantly from controls. Lyrics analysis is underway.

Interdisciplinary implications. Functional bifurcation manifested in some corpora along with tendencies towards positivity bias in online corona-musicking. Consequently, consumers may have adopted short-term, mental-health benefits (Nabi & Krcmar, 2004) by regulating mood and stress levels, thus facilitating subsequent problem-oriented, long-term coping (Halfmann & Reinecke, 2021). As musical participation becomes increasingly digitalized, text-based communication and meta-data provide information-rich resources that NLP tools can readily exploit.

References

- Baccianella, S. et al. (2010). SentiWordNet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In *Proceedings of LREC*.
- Brost, B. et al. (2019). The Music Streaming Sessions Dataset. In *Proceedings of the 2019 World Wide Web Conference*, San Francisco, CA.
- Eden, A.L. et al. (2020). Media for coping during COVID-19 social distancing: stress, anxiety, and psychological well-being. *Front Psychol*.
- English-Corpora.org. (n.d.). NOW Corpus. Retrieved on June 14, 2021, from <https://www.english-corpora.org/now/>
- Fink, L. et al. (2021). Viral tunes: changes in musical behaviours and interest in coronamusic predict socio-emotional coping during COVID-19 lockdown. *Hum Soc Sci Commun*.
- Halfmann, A., & Reinecke, L. (2021). Binge-watching as case of escapist entertainment use. In *The Oxford handbook of entertainment theory*. OUP.
- Hansen, N.C. (2021). Music for hedonia and eudaimonia during pandemic social isolation. In *Arts and Mindfulness Education for Human Flourishing*. Routledge. doi:10.31234/osf.io/s9jf6
- Hansen, N.C. et al. (2021). A crowd-sourced database of coronamusic: documenting online making and sharing of music during the COVID-19 pandemic. *Front Psychol*.
- Lamsal, R. (2020). Design and analysis of a large-scale COVID-19 tweets dataset. *Appl Intell*.
- Liu, Y. et al. (2021). Monitoring COVID-19 pandemic through the lens of social media using NLP and machine learning. *Health Inf Sci Syst*.
- Nabi, R.L., & Krcmar, M. (2004). Conceptualizing media enjoyment as attitude: Implications for mass media effects research. *Commun Theory*.
- Qazi, U. et al. (2020). GeoCoV19: a dataset of hundreds of millions of multilingual COVID-19 tweets with location information. *ACM SIGSPATIAL Special*.