Scary music mimics alarming acoustic feature of human screams

Caitlyn Trevor[1†], Luc Arnal[2] and Sascha Frühholz[1,3]
[1] Department of Psychology, University of Zurich, Zurich, Switzerland
[2] Department of Fundamental Neuroscience, University of Geneva, Geneva, Switzerland
[3] Department of Psychology, University of Oslo, Oslo, Norway
[†] Corresponding author: caitlyn.trevor@psychologie.uzh.ch

Introduction

Music used to underscore frightening scenes in movies is often described as sounding "scream-like". A well-known example is the music accompanying the infamous shower murder scene in Alfred Hitchcock's film Psycho (1960) with, "screeching, upward glissandi," from the violins (Brown, 1982; p. 46). Although 'scream-like' is a common descriptor, the question remains: do these scary film soundtrack excerpts actually sound like and are perceived similarly to human screams? Research has demonstrated that screams have a unique auditory feature called "roughness" (Arnal et al., 2015; Schwartz et al., 2019). To measure roughness, we employ the modulation power spectrum (MPS) method and parameters used by Arnal et al. (2015). The MPS is a two-dimensional Fourier transformation of a spectrogram that quantifies both temporal and spectral power modulations (Elliott & Theunissen, 2009). Previous research indicates that human screams feature higher MPS values than non-alarming vocalizations in the 30 to 150 Hz range of the temporal modulation rate dimension of the MPS (Arnal et al., 2015). To investigate whether scream-like music has the same roughness feature as, and is perceived similarly to, human screams, we conducted two studies. In the first study, we ran an acoustic analysis to test whether recorded screams and scream-like music exhibit enhanced roughness compared with control recordings. In the second study, we collected valence and arousal ratings for the audio files in order to test whether screams and scream-like music are perceived as sharing similar emotional qualities. We made the following hypotheses. First, we hypothesized that the mean power of the MPS within the roughness region (henceforth "roughness") would be similar for screams and scream-like music, and would be significantly greater for screams compared to non-screaming vocalizations and for scream-like music compared to non-scream-like music. Second, given that roughness may be a universal cue for danger (Arnal et al., 2015), we hypothesized that roughness would correlate negatively with valence ratings and positively with arousal ratings for both music and vocal stimuli. Taken together, these results would demonstrate that scream-like music both sounds like and is perceived similarly to actual human screams.

Method

The audio recordings used in the studies were deployed in a 2x2 factorial design, with one factor corresponding to the sound source (music, voice) and the other one to the scream-likeness of the sounds (scream-like, non-scream-like). Specifically, the four collections included (a) fearful scream vocalizations, (b) scream-like film music excerpts, (c) non-fearful human vocalizations (sounding similar to a held "ah" sound), and (d) non-scream-like film music excerpts as controls. All audio recordings are 800ms in duration, RMS normalized, sampled at 16kHz, and are in wav-file format. The film music excerpts were curated from recently released horror film soundtracks (to download the excerpts, see our Open Science Framework project page at https://osf.io/7d2cy/). The vocalizations were recorded at the University of Zurich. Using MATLAB, the MPS of each excerpt was measured using the same procedure and equations as used by Arnal and colleagues (Arnal et al., 2015). Specifically, the initial spectrograms were obtained using a filter-bank approach with 128 Gaussian windows whose outputs were Hilbert transformed and then log-transformed. Then the modulation power spectra were obtained by applying a two-dimensional Fourier transform to the spectrogram (24 channels/octave) and log-transforming the resulting spectral power density estimates (Arnal et al., 2015). From there, the mean amplitude in the roughness range of 30 to 150 Hz along the temporal modulation range was taken [see Fig. 1(A); see Elliott and Theunissen (2009) for more detailed information on the MPS]. In the second study, 20 healthy

participants (twelve female) listened to each of the 200 audio files and rated the valence and arousal of the conveyed emotion using two continuous sliding scales.

Results

Recall that our first hypothesis predicted that scream-like music and screams would share a similar roughness level that would be higher than their matched controls. To test our first hypothesis, we used a standard general linear regression analysis. We also tested for an interaction effect between sound types and scream-likeness. Finally, another standard general linear regression analysis was used to test for a main effect of scream-likeness on roughness for just the voice (coded as 1) for replicative comparison to the findings of Arnal et al. (2015). The results are reported in Table 1.

Table 1: Roughness as predicted by Scream-likeness and Sound Type

	Roughness (Voice Only)	**Roughness** (Main Effect)	**Roughness** (Interaction)
Intercept	**-0.88*****	**-0.70*****	**-0.39*****
	[-1.00, -0.75]	[-0.89, -0.52]	[-0.59, -0.19]
Scream-likeness (Screaming = 1)	**1.89*****	**1.27*****	**0.65*****
	[1.72, 2.06]	[1.05, 1.48]	[0.37, 0.93]
Sound Type (Voice = 1)		0.14	**-0.48****
		[-0.08, 0.35]	[-0.76, -0.21]
Scream-likeness X Sound Type			**1.24*****
			[0.85, 1.64]
N	100	200	200
R^2 / R^2 *Adjusted*	0.826 / 0.824	0.407 / 0.401	0.503 / 0.495
*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$			CI 95%

Our second hypothesis was that roughness would correlate negatively with valence ratings and positively with arousal ratings for both music and vocal stimuli, supporting its reputation as an aural cue for danger (Arnal et al., 2015). To test this hypothesis, we used emotion ratings (valence and arousal) as the predicted values for two mixed effects linear regression models. Results are reported in part in Figure 1 [for more details on the results, see (Trevor, Arnal, & Frühholz, 2020)].

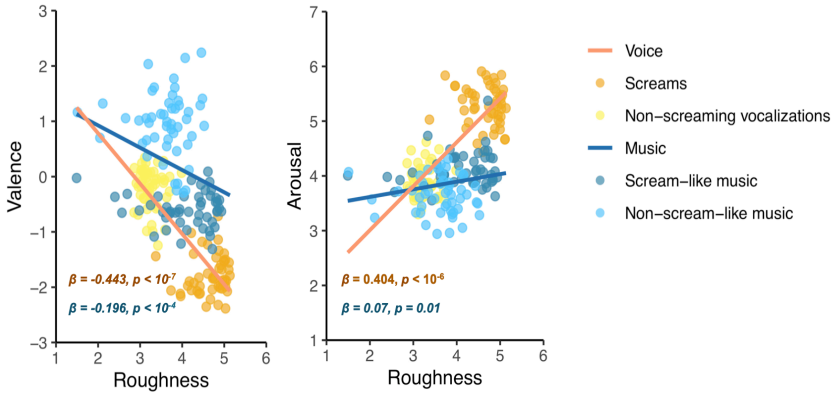

Figure 1: The left plot shows the relationship between the average valence ratings and average roughness of the four stimulus categories. There is a significant negative correlation between valence and roughness for both vocal ($\beta = -0.443$, $p < 10-7$) and musical ($\beta = -0.196$, $p < 10-4$) stimuli.

However, the correlation is significantly more pronounced for vocal stimuli as opposed to musical stimuli ($p < 10-5$). The right plot shows the relationship between the average arousal ratings and average roughness of the four stimulus categories. There is a significant positive correlation between arousal and roughness for both vocal ($\beta = 0.404$, $p < 10-6$) and musical ($\beta = 0.07$, $p = 0.01$) stimuli. However, the correlation is significantly more pronounced for vocal stimuli as opposed to musical stimuli ($p < 10-5$).

Consistent with our hypotheses, we found that both screams and scream-like music exhibited a higher level of roughness and were rated as having a more negative valence and a higher arousal level than their non-screaming counterparts. However, contrary to our hypotheses, screams had a higher roughness level than scream-like music. Overall, the results demonstrated a greater difference in roughness levels and emotion ratings between the vocal stimuli than between the musical stimuli.

Discussion

These results suggest that while scream-like music does seem to sound like and be perceived similarly to human screams, the musical rendition is still a muted version of the real thing and therefore may not provoke as potent of a reaction. Overall, the results suggest that roughness can effectively translate from a vocal cue for danger into a musical cue for danger. It is therefore reasonable to suggest that scream-like music might scare viewers in part because it is evocative of a human scream, a naturally alarming sound. For further details and analyses, please see our published paper in The Journal of the Acoustical Society of America (Trevor, Arnal, & Frühholz, 2020).

Acknowledgments

C.T. received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement (No. 835682). S.F. received funding from Swiss National Science Foundation (Grants Nos. SNSF PP00P1_157409/1 and PP00P1_183711/1). The authors thank Lawrence Feth for guidance regarding the acoustic analyses and to David Huron for valuable feedback on the project. Finally, the authors thank Arkady Konovalov for helpful input regarding the statistical analyses.

References

Arnal, L. H., Flinker, A., Kleinschmidt, A., Giraud, A.-L., & Poeppel, D. (2015). Human Screams Occupy a Privileged Niche in the Communication Soundscape. *Current Biology, 25*(15), 2051–2056. https://doi.org/10.1016/j.cub.2015.06.043

Brown, R. S. (1982). Herrmann, Hitchcock, and the music of the irrational. *Cinema Journal, 21*(2), 14-49. https://doi.org/10.2307/1225034

Elliott, T. M., & Theunissen, F. E. (2009). The modulation transfer function for speech intelligibility. *PLoS Computational Biology, 5*(3), 1–14.

Schwartz, J. W., Engelberg, J. W., & Gouzoules, H. (2019). What is a scream? Acoustic characteristics of a human call type. *The Journal of the Acoustical Society of America, 145*(3), 1776–1776. https://doi.org/10.1121/1.5101500

Trevor, C., Arnal, L. H., & Frühholz, S. (2020). Terrifying film music mimics alarming acoustic feature of human screams. *The Journal of the Acoustical Society of America, 147*(6), EL540-EL545. https://doi.org/10.1121/10.0001459