

Emotions perceived in music and speech: predicting second-by-second subjective feelings of emotion from psychoacoustic features and physiological responses

Eduardo Coutinho¹, Nicola Dibben²

¹Swiss Center for Affective Sciences, Switzerland

²University of Sheffield, United Kingdom

There is strong evidence that the structure of affective responses to auditory stimuli is largely dependent on dynamic temporal patterns in low-level music structural parameters. Previous studies have shown that spatiotemporal dynamics in a small set of psychoacoustic features can predict two fundamental psychological dimensions of emotion: valence and arousal. The study reported here aims to determine the extent to which specific physiological responses can be used in tandem with psychoacoustic cues to predict emotional responses to music and natural speech. The method consists of a behavioural study and computer simulations. The behavioural study collected two main types of data: continuous ratings of emotion perceived while listening to extracts of music and speech, using a computer interface which modelled emotion on two dimensions (arousal and valence), and physiological measures (respiration, heart rate, skin conductance, skin temperature, and blood pressure) taken while listening to each stimulus. Then we trained a neural network model to assess how well listeners' perceived emotions can be predicted from psychoacoustic cues and subjects' physiological responses. Based on previous evidence for the role of peripheral feedback in emotional response to music we expect the inclusion of physiological data in the model to contribute a small predictive value for music, and by implication, speech. Data analysis is due for completion in the next few months, and will focus on evaluating whether physiological responses are reliable predictors of emotions perceived in music and speech as well as in the comparison of both domains.

Keywords: Neural Networks, Physiology, Valence and Arousal, Psychoacoustics.