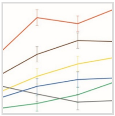

Gating Experiment Using Affect Bursts

PLOS One

Simon Schaerlaeken[1], Didier Grandjean[1]

[1]University of Geneva

dx.doi.org/10.17504/protocols.io.rr2d58e

Simon Schaerlaeken

ABSTRACT

The unfolding dynamics of the vocal expression of emotions are crucial for the decoding of the emotional state of an individual. In this study, we analyzed how much information is needed to decode a vocally expressed emotion using affect bursts, a gating paradigm, and linear mixed models. We showed that some emotions (fear, anger, disgust) were significantly better recognized at full-duration than others (joy, sadness, neutral). As predicted, recognition improved when greater proportion of the stimuli were presented. Emotions recognition curves for anger and disgust were best described by higher order polynomials (second to third), while fear, sadness, neutral, and joy were best described by linear relationships. Acoustic features were extracted for each stimulus and subjected to a principal component analysis for each emotion. The principal components were successfully used to partially predict the accuracy of recognition (i.e., for anger, a component encompassing acoustic features such as fundamental frequency (f0) and jitter; for joy, pitch and loudness range). Furthermore, the impact of the principal components on the recognition of anger, disgust, and sadness changed with longer portions being presented. These results support the importance of studying the unfolding conscious recognition of emotional vocalizations to reveal the differential contributions of specific acoustical feature sets over time. It is likely that these effects are due to the relevance of threatening information to the human mind and are related to urgent motor responses when people are exposed to potential threats as compared with emotions where no such urgent response is required (e.g., joy).

EXTERNAL LINK

https://doi.org/10.1371/journal.pone.0206216

THIS PROTOCOL ACCOMPANIES THE FOLLOWING PUBLICATION

Schaerlaeken S, Grandjean D (2018) Unfolding and dynamics of affect bursts decoding in humans. PLoS ONE 13(10): e0206216. doi: 10.1371/journal.pone.0206216

PROTOCOL STATUS

Working

Select recordings from the GEneva Multimodal Emotion Portrayals (GEMEP)

1 Five emotions were selected for this study: anger, disgust, fear, joy, and sadness. These emotions are thought to show discrete forms of expression in the face and voice. Surprise was excluded from this list because of the difficulty in simulating a realistic expression experimentally. Neutral recordings were also used.

We selected 20 recordings from the database for each emotion (except for disgust, which is represented through only 10 recordings because of the poor quality of other recordings). Each recording represented uniquely one of the emotions selected for this study. The emotion expressed in each recording did not change over the course of the recording and was unambiguous.

DATASET
GEneva Multimodal Emotion Portrayals

Normalized the recordings and remove silences

2 The volume of each recording was normalized to maintain a uniform listening experience between different stimuli (with a coefficient counterbalancing the average and maximum energy) and silences at the beginning and hlthe end of each recording were removed

SOFTWARE

Praat 5.3.51

Windows

Create the gates

3 Each complete recording were cut into smaller chunks with an incremental duration of 50 ms.

SOFTWARE

Praat 5.3.51

Windows

Generate pseudo-randomized lists of stimuli

4 We generated lists of 105 stimuli picked from the whole dataset and presented them in a pseudo-random order.

Three or four different participants evaluated each stimulus. Amongst the 8400 stimuli, we picked 72 to be presented more frequently (10 times) to the participants in order to test the reliability of emotional judgments across the different individuals.

Present stimuli to participants and record responses

5 Participants completed the experiment on a computer. The experiment itself was programmed with Windev, version 15 and ran on computers with a screen resolution of 1,280 x 1,024 pixels. Volume was set to 50% and could be adjusted by the participant. Headphones were provided. The experiment lasted 30 min. The participants had to sign an informed consent form presented in a written form and complete a demographic questionnaire before starting the experiment. The instructions were displayed on the computer screen and a set of three practice trials was presented to the participant (with stimuli not used in the main task). A different list of 105 stimuli out of the entire dataset was presented in a pseudo-random order for each participant. The participant listened to each stimulus and then rated six emotions (anger, fear, joy, sadness, disgust, surprise) and neutral with seven corresponding sliders from 0 to 100. Multiple emotion sliders could be moved but at least one of them had to be moved. The participants were also asked to rate their confidence level for the choice they made. The emotions and confidence levels were evaluated on a continuous scale. This process was repeated for the 105 stimuli used per participant. Finally, the data was anonymised based on the ethics commission's requirements.

SOFTWARE

Windev 15

Windows
by PC SOFT

Group stimuli responses by bins

6 We grouped stimuli into bins representing the percentage of the total duration presented. The different emotions were represented on the same scale independent of their global duration. We created two sets of bins: a general one at 25% intervals (25%-50%-75%-100%) and a more detailed one 10% intervals. For example, if a stimulus, created from an original recording with a duration of 100ms, lasts for 150 ms, it falls into the 25% bin and the 20% bin.

NOTE: A bin groups together multiple 'gates,' as defined when we cut the original recordings. However, for the sake of using the same terminology here as in the studies that use a gating paradigm, we also use the term 'gate' to represent the bins in the analysis.

Computation of Hu Scores

7 We decided to compute the unbiased hit rate (Hu) for every emotion at every 10% gate. In order to do so, we first defined that a trial was successfully recognized when the highest value of the continuous emotional scales corresponded to the emotion presented. With this information, we compute a confusion matrix created based on the hits and misses of all the participants for each emotion at every 10% gates. We also computed Hu score, for each participant, based on a personal confusion matrix computed for every emotion at every 25% gate. Only the cruder gates were used for this case to provide enough data for the computation of Hu. Since each participant was associated with his/her own set of Hu score, this provided us with a Hu score variability at every 25% gate allowing to compute statistical models.

Computation of acoustical features

8 We estimated a set of 42 acoustical features proposed by the Geneva Minimalist Acoustic Parameter Set, developed for emotional prosody research. Dimensionality reduction was applied through a principal component analysis (PCA) specific to each emotion. We selected four components for each emotion based on the cumulative sum of eigenvalues.

SOFTWARE
openSMILE
Windows
by audEERING

SOFTWARE
Matlab 16b
Windows
by MathWorks

Linear mixed models

9 We used these Hu score for each participant to examine the accuracy of the recognition of emotions across all trials per gate and emotion using linear mixed models (LMMs).
We used chi-square difference tests to investigate the contribution of each variable and their interaction. The fixed effects were the emotion expressed (anger, sadness, joy, disgust, fear, neutral) and the duration of the gates (25-50-75-100%). The random intercepts effects encapsulated the variability related to each participant. We used a step-up strategy while building the model to test the different combinations of fixed effects. The LMM estimated recognition curves, consisting of unbiased recognition of a specific emotions computed at every gate.
The 4 components from the Principal Component Analysis of the acoustic features were also used as fixed effect in LMM to map it to the Hu score.

SOFTWARE
lme4 1.1-17
Windows
by Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker

Recognition curves fit and polynomial contrasts

10 We computed a polynomial contrast for each recognition curve to determine the nonlinear shape of the curve over time. For the 10 gates model, we computed polynomial regression and the corresponding root mean square error (RMSE).

SOFTWARE
Phia 0.2-1
Windows
by Helios De Rosario-Martinez

SOFTWARE
Matlab 16b
Windows
by MathWorks

This is an open access protocol distributed under the terms of the **Creative Commons Attribution License** (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited