Discriminating between authentic and fake emotional vocalisations - insights from electroencephalography.

Maciej Kosilo, Mónica Costa, César Lima, Ana Pinheiro, Diana Prata

Biomedical Neuroscience Lab (Diana Prata's Lab)

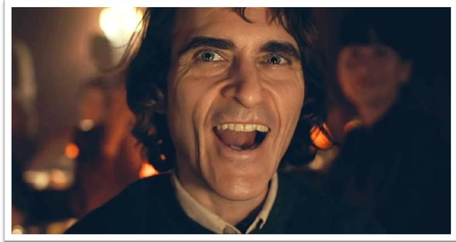
IBEB

Ciências ULisboa

FCT Fundação para a Ciência e a Tecnologia

Bial

Background



- Affective voice conveys information on the person's emotional state. However, emotional vocalisations can be both genuine, produced spontaneously (i.e. authentic), and voluntary (deliberate, controlled, "fake"). Voluntary, fake vocalisations carry different meanings and evoke different social responses.
- Ability to tell the genuine from the voluntary is important in everyday social interactions and is thought to be related to cognitive empathy (Shamay-Tsoory, 2011)
- fMRI findings suggest that different brain regions are involved in the processing of authentic and voluntary emotions, with the former being more automatic, and the latter engaging areas involved in mentalising (McGettigan et al., 2015)
- However, little is known about the time course of emotional processing in relation to auditory authenticity recognition.

Aims and hypotheses

- To explore the time course of authenticity processing using ERPs
- Can we observe a difference between authentic/fake vocalisations in early or later ERP components?
- Can we attribute the differences to individual's trait empathy scores?

Methods (EEG)

EEG preprocessing

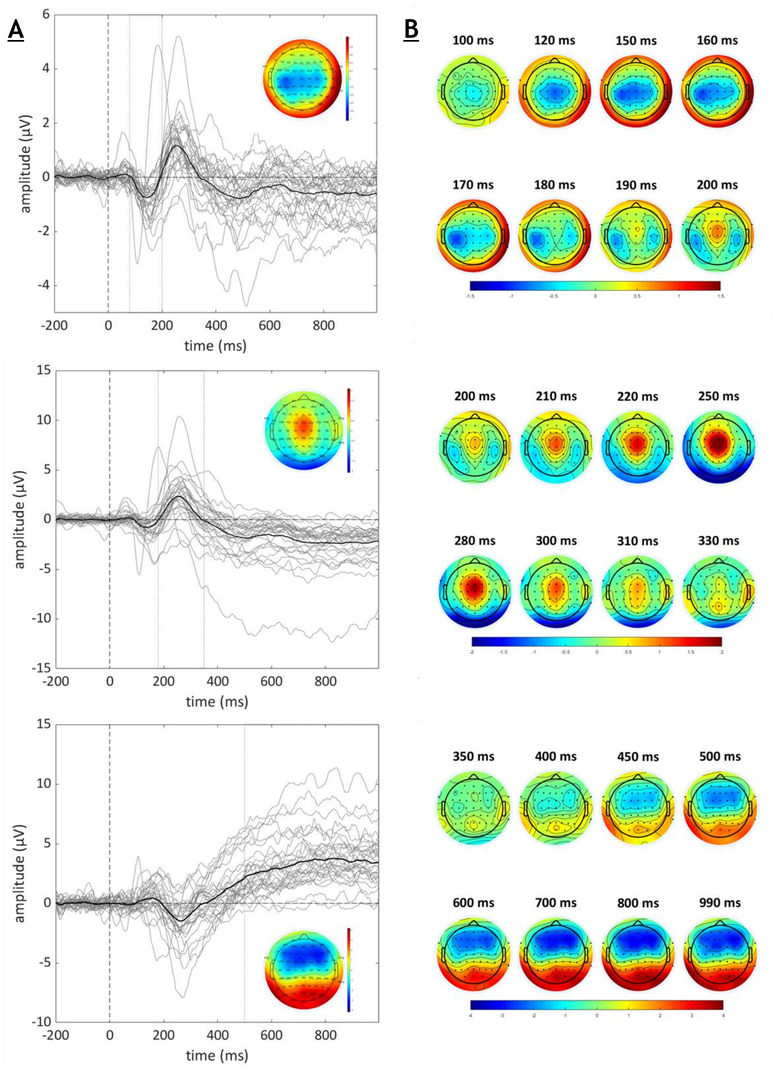
- 64-channel BrainVision actiCHamp system (Brain Products, München, Germany), sampled at the 512Hz rate.
- An offline 0.1 high and 30Hz low-pass Butterworth filter
- Re-referenced to average (after removal of noisy channels)
- Epochs: -200 before and 1000ms after stimulus onset.
- Epochs inspected visually and removed if they contained non-stereotypical artifacts (large voltage offsets, muscle and head movements).
- Ocular artifact correction using Independent Component Analysis (ICA).
- Baseline correction (baseline: 200ms prior to stimulus onset)

ERP channel and time-window justification

- ERP components in this study:
  - N100; early auditory processing; shown to be **sensitive to emotionality**
  - P200; **sensitive to changes in motivational significance of events** (e.g. higher amplitude for reward-linked stimuli, relative to neutral stimuli).
  - LPP (late positive potential); linked to higher-order processes and cognitive evaluation.

- Topography for all conditions averaged together were used to select electrode locations for statistical analysis ("collapsed localisers" approach, Luck, 2017); single-subject ERPs were used to aid the selection of the analysis time window.

Grand-average ERP (black line) and single-subject ERPs (grey lines) shown in A. Activity is averaged over electrodes selected from topographies (B). The final choice of time and window and channels in C.



N100: 80-200ms; central and centro-parietal electrodes

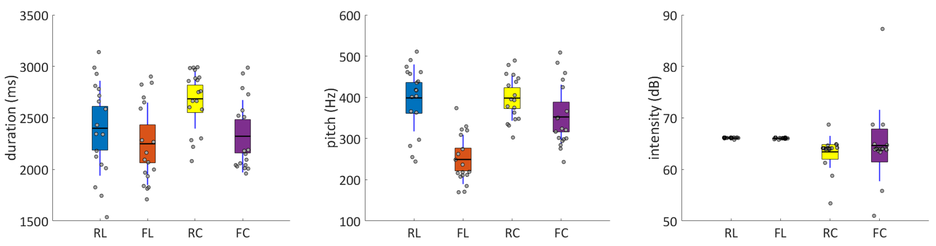
P200: 180-350ms; electrodes Cz and FCz

LPP: 500-1000ms; occipital and parieto-occipital electrodes

- Activity from the selected electrodes was collapsed; the amplitude was measured between the component's onset and offset times (onset was defined as a time point where the amplitude reached 50% of the local peak within the time window of interest and analysed using ANOVA.

Potential confounds

The stimuli used in the study differed in terms of their acoustic properties.

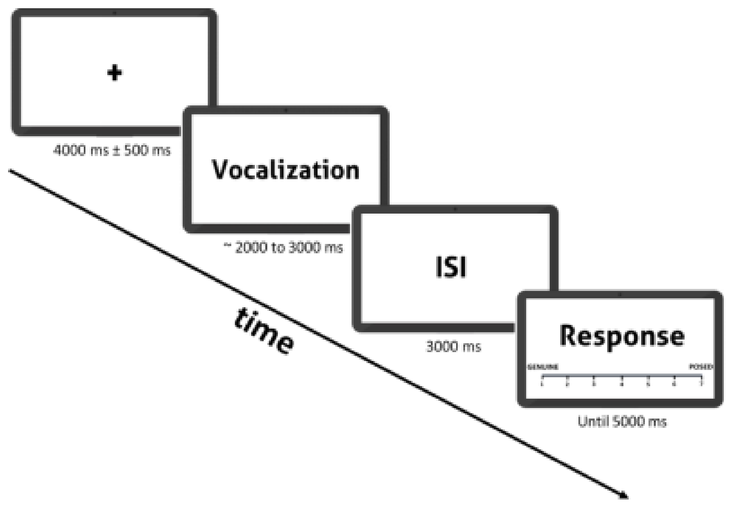


ERP amplitudes	pitch R_{rm}	pitch *p*-value	intensity R_{rm}	intensity *p*-value	duration R_{rm}	duration *p*-value
N100	.025	.055	.006	.619	.008	.518
P200	-.008	.516	.01	.444	.015	.25
LPP	.012	.345	.007	.611	-.004	.76

Repeated measure correlation coefficients (Rrm; see Bakdash & Marusich, 2017) between single-trial ERP amplitudes and acoustic stimuli properties.
This suggests that ERP effects are not explained by differences in low-level properties.

Task

Trial outline



Authenticity task

Two classes of emotional vocalisations produced by actors: "real" (i.e. authentic-sounding) and "fake" (see Lavan et al, 2014; McGettigan et al, 2015) ; **132** nonverbal vocalizations in total

- Real laughing (RL), Fake laughing (FL) (18 each)
- Real Crying (RC), Fake Crying (FC)(18 each)
- Neutral vocalisation (30)
- Spectrally-rotated stimuli (30 – not analysed, as the sounds were deemed too unnatural and more salient)

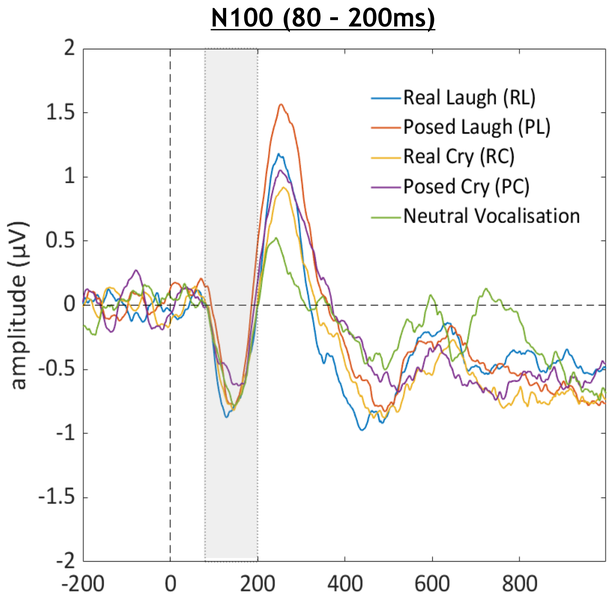
Participants passively listen to one vocalisation in a trial
Each vocalisation was subsequently rated on a scale from 1 (real/authentic) to 7 (fake).

Other measures:

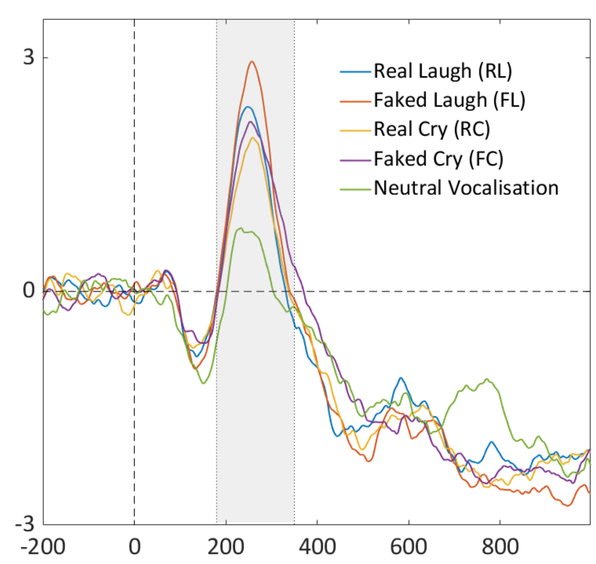
- Empathy Quotient (EQ) with subfactors
- Reading Mind from the Eyes task (RMET)
- Authenticity detection index (authentic stimuli rtings – fake stimuli ratings; Neves et al., 2018).

Results

N100 (80 – 200ms)



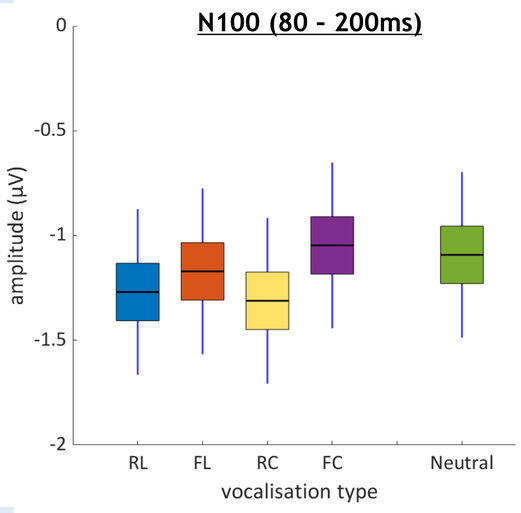
P200 (180 – 350ms)



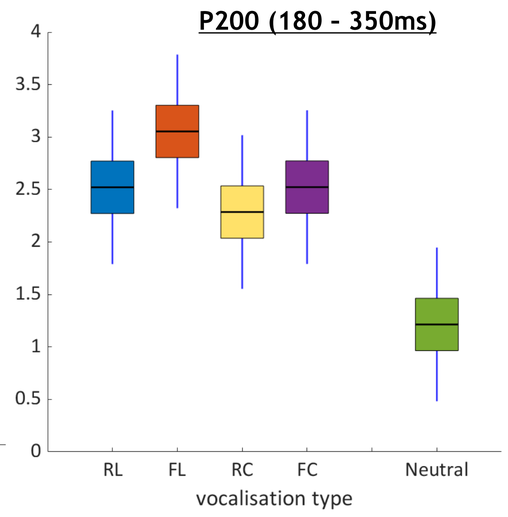
N100 amplitude modulated by authenticity ($F(1, 31)=5.67$, $p=.024$, $\eta^2=.155$), but not emotion (laugh vs crying; $F(1, 31)=.32$, $p=.576$, $\eta^2=.010$).

P200 amplitude modulated by authenticity ($F(1, 31)=7.43$, $p=.010$, $\eta^2=.193$) and emotion (laugh vs crying; $F(1, 31)=4.69$, $p=.038$, $\eta^2=.131$).

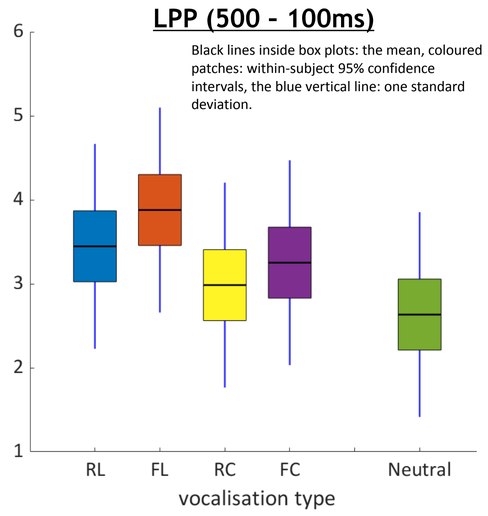
N100 (80 – 200ms)

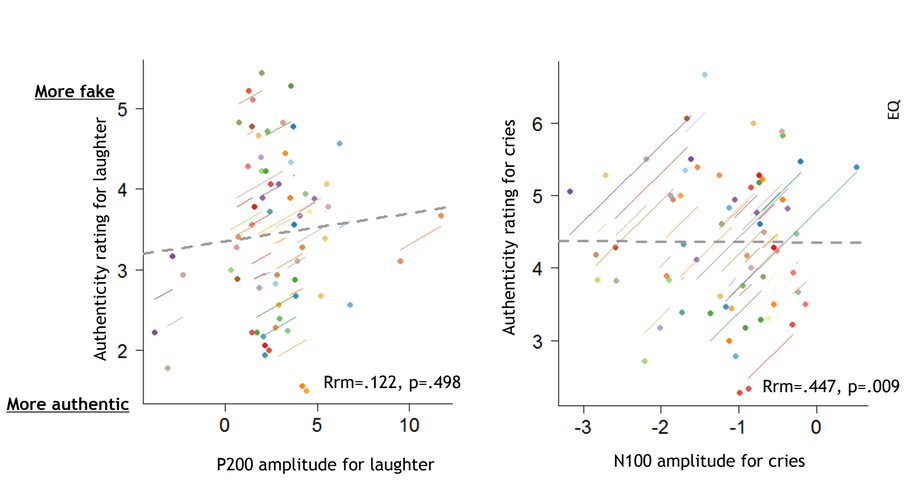


P200 (180 – 350ms)

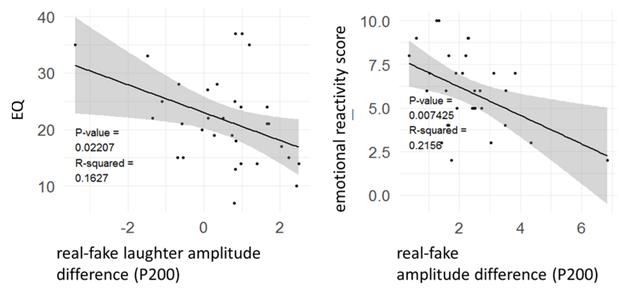


LPP (500 – 100ms)





N100 amplitude correlated with authenticity ratings



Authenticity Detection Index

	Laughter r	Laughter p-value	Cry r	Cry p-value
EQ	-.087	.640	-.146	.640
cognitive empathy	-.155	.400	-.167	.400
emotional reactivity	-.047	.800	-.066	.800
social skills	.051	.780	-.217	.780
empathic difficulties	.051	.780	-.012	.780
RMET	-.202	.270	.063	.270

P200 real-fake amplitude difference correlated with EQ measures; but the measures did not correlate with authenticity detection index (real-fake ratings)

Discussion

- Significant authentic/fake ERP differences found in the early N100 (80-100ms) and P200 (180-350ms) time windows, but not later. P200 effects for laughter might be more reliable/of greater magnitude.
- N100 amplitude highest for authentic vocalisations; P200 higher for fake.
- Lack of association between authenticity detection index, EQ and RMET suggest that a more basic processes might be responsible for the ability to tell authentic from fake in this task.
- N100 effects might be driven by emotional content of authentic vocalisations; but P200 might be driven by greater cognitive demands when processing fake stimuli, or by their incongruence. However, correlations with EQ measures suggest a role of empathy as well.

- **ERPs suggest that the real/faked distinction in the brain occurs relatively early . This early activity might impact later, more deliberate authenticity ratings.**

m.kosilo@gmail.com
@M_Kosilo
@DPrataLab