cortex 151 (2022) 116–132

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research Report

The time course of emotional authenticity detection in nonverbal vocalizations

Tatiana Conde [a,*], *Ana I. Correia* [b], *Magda S. Roberto* [a], *Sophie K. Scott* [c], *César F. Lima* [b,c] and *Ana P. Pinheiro* [a]

[a] *CICPSI, Faculdade de Psicologia, Universidade de Lisboa, Lisboa, Portugal*

[b] *Centro de Investigação e Intervenção Social (CIS-IUL), Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal*

[c] *Institute of Cognitive Neuroscience, University College London, London, UK*

A R T I C L E I N F O

Article history:
Received 23 April 2021
Reviewed 16 July 2021
Revised 23 December 2021
Accepted 16 February 2022
Action editor Veena Kumari
Published online 19 March 2022

Keywords:
Authenticity
Emotion
Voice
Event-related potentials

A B S T R A C T

Previous research has documented perceptual and brain differences between spontaneous and volitional emotional vocalizations. However, the time course of emotional authenticity processing remains unclear. We used event-related potentials (ERPs) to address this question, and we focused on the processing of laughter and crying. We additionally tested whether the neural encoding of authenticity is influenced by attention, by manipulating task focus (authenticity versus emotional category) and visual condition (with versus without visual deprivation). ERPs were recorded from 43 participants while they listened to vocalizations and evaluated their authenticity (volitional versus spontaneous) or emotional meaning (sad versus amused). Twenty-two of the participants were blindfolded and tested in a dark room, and 21 were tested in standard visual conditions. As compared to volitional vocalizations, spontaneous ones were associated with reduced N1 amplitude in the case of laughter, and increased P2 in the case of crying. At later cognitive processing stages, more positive amplitudes were observed for spontaneous (versus volitional) laughs and cries (1000–1400 msec), with earlier effects for laughs (700–1000 msec). Visual condition affected brain responses to emotional authenticity at early (P2 range) and late processing stages (middle and late LPP ranges). Task focus did not influence neural responses to authenticity. Our findings suggest that authenticity information is encoded early and automatically during vocal emotional processing. They also point to a potentially faster encoding of authenticity in laughter compared to crying.

* *Corresponding author*. Faculdade de Psicologia, Universidade de Lisboa, Alameda da Universidade, 1649-013, Lisboa, Portugal.
E-mail address: tatianamagro@psicologia.ulisboa.pt (T. Conde).

https://doi.org/10.1016/j.cortex.2022.02.016

1. Introduction

The ability to perceive emotional information from the voice is fundamental for social interactions. Research on vocal emotions is challenged by the fact that our vocal emotional repertoire is complex and variable. It relies on both automatic and voluntary production mechanisms that might be intentionally and flexibly adjusted according to the social context and the speaker's communicative intentions (Scott et al., 2014; Sidtis & Kreiman, 2012). Consider the distinct social meanings of a laugh spontaneously produced in response to a funny situation, for instance, compared to a laugh voluntarily produced to communicate polite agreement. From the listener's perspective, differentiating authentic (*spontaneous*) from more deliberate (*volitional*) emotional expressions is an important social skill, with implications for cooperation, affiliation, and bonding (Bryant et al., 2018; Gervais & Wilson, 2005; Wood et al., 2017). There has been a growing research interest in this issue in recent years, but the neural mechanisms involved in processing emotional authenticity in nonverbal emotional vocalizations remain poorly understood.

1.1. *Differentiating spontaneous from volitional vocal emotional expressions*

Most research on auditory emotion perception relies on acted vocal portrayals (Scherer & Bänzinger, 2010). Such stimuli are typically obtained by inviting professional or nonprofessional actors to pose emotional expressions, in the absence of corresponding eliciting events. Acted portrayals are considered suitable for research on vocal emotions, allowing for more control over stimulus features (Scherer & Bänzinger, 2010). Nonetheless, recent experiments have pointed out that they differ from spontaneous emotional expressions in important ways (e.g., Anikin and Lima, 2018; Bryant & Aktipis, 2014; McGettigan et al., 2015; Neves et al., 2018). For example, differences in the acoustic features of spontaneous and volitional vocalizations may reflect distinct vocal production mechanisms (Anikin and Lima, 2018; Bryant & Aktipis, 2014; Lavan et al., 2016; McKeown et al., 2015). Moreover, spontaneous vocalizations are characterized by higher and more variable fundamental frequency (F0) and lower harmonicity than volitional expressions (Anikin and Lima, 2018). At the perceptual level, although most research has focused on laughter (e.g., Bryant & Aktipis, 2014; Lavan et al., 2016; Neves et al., 2018), there is evidence that listeners are able to differentiate spontaneous from volitional vocalizations of amusement, sadness, achievement, anger, disgust, fear, pain, and pleasure (Anikin and Lima, 2018). These studies highlight the relevance of investigating vocal emotional perception using spontaneous expressions, and of examining the differences between them and volitional ones.

At the brain level, the few existing studies on authenticity focused on laughter, and they identified distinct cortical responses to spontaneous and volitional expressions (Lavan et al., 2017; McGettigan et al., 2015). McGettigan et al. (2015) found that passively listening to spontaneous (versus volitional) laughter induced greater activity in the bilateral superior temporal gyrus, whereas volitional laughter elicited enhanced activity in anterior medial prefrontal and anterior cingulate cortices. Such stronger engagement of medial prefrontal systems highlights the potential role of mentalizing processes in the perception of acted laughs, plausibly due to their higher social-emotional ambiguity (McGettigan et al., 2015). This study also reported that improved accuracy in discriminating laughter authenticity was associated with enhanced activity in sensorimotor systems, including the presupplementary motor area and lateral somatosensory cortex, emphasizing the role of these systems in successful authenticity evaluations (McGettigan et al., 2015). Lavan et al. (2017) have further observed that the dissociable brain responses to spontaneous versus volitional laughter were related to the perceived authenticity and affective properties (i.e., valence and arousal) of the stimuli. Laughs rated as less authentic were associated with stronger activity in brain regions engaged in mentalizing (i.e., anterior medial prefrontal cortex), whereas laughs rated as more authentic and arousing were associated with stronger activity in regions engaged in voice perceptual processing (Heschl's gyrus and superior temporal gyrus).

While these findings document cortical sensitivity to laughter authenticity, it remains unclear whether this is laughter-specific or extends to other vocalizations, such as crying. Crucially, given the poor temporal resolution of fMRI, the time course of authenticity processing in emotional vocalizations remains unknown. Is authenticity processed at early sensory, or at later higher-order stages of voice perception? Event-related potentials (ERPs) are ideal to address these questions. To our knowledge, however, no previous studies have used this technique to examine authenticity processing in vocalizations.

ERP studies with acted vocal expressions have shown a differential processing of emotional and neutral vocalizations at distinct processing stages: compared to neutral vocalizations, emotional voices typically elicit reduced N1 amplitude, a component associated with sensory processing, as well as enhanced P2, a component associated with salience detection (Jessen & Kotz, 2011; Liu et al., 2012; Sauter & Eimer, 2010). These studies suggest that the N1 and P2 are sensitive to the emotional quality of the voice regardless of its valence (Jessen & Kotz, 2011; Liu et al., 2012; Sauter & Eimer, 2010; but see Pell et al., 2015 for emotion-specific effects on the P2). At later processing stages, the Late Positive Potential (LPP) component is thought to reflect sustained attention and cognitive evaluation of emotionally and motivationally significant information (Jessen & Kotz, 2011; Pell et al., 2015; Pinheiro et al., 2016a, 2017). The LPP is observed around 400–600 msec post-stimulus onset, yet it may last for up to 1 sec following stimulus offset (Hajcak & Olvet, 2008; Schupp et al., 2006). More positive LPP amplitudes for emotionally salient versus neutral stimuli have been documented for speech prosody (Paulmann et al., 2013), pictures (Cuthbert et al., 2000; Hajcak & Nieuwenhuis, 2006), and faces (Foti et al., 2010). The extent to which the LPP is sensitive to the emotional content of vocalizations is unclear. Pell et al. (2015) found that the LPP amplitude was increased for angry compared to sad and happy vocalizations, whereas Jessen and Kotz (2011) reported non-significant emotional modulations of LPP. Pinheiro et al., 2016b provided evidence for an enhanced attentional orienting (increased P3a) for emotional (happy and angry) versus neutral vocalizations, suggesting that emotionally salient voice information captures attentional resources more.

1.2. Attention focus and vocal emotional processing

The few existing studies testing the relationship between attention and vocal emotional processing have examined speech prosody and they suggest that attentional focus might affect brain responses (Grandjean et al., 2005; Sander et al., 2005). In a dichotic listening study, Sander et al. (2005) reported enhanced activity in the orbitofrontal cortex and cuneus when listening to angry prosodic stimuli that were presented in the to-be-attended versus to-be-ignored ear during a gender evaluation task. This suggests that attention affects cortical responses to emotional prosody, but this study also provided evidence for automaticity: in the right amygdala and bilateral superior temporal sulcus (STS), responses to angry prosody were similar regardless of whether the stimuli were presented in the to-be-attended or to-be-ignored ears. In a different study, Ethofer et al. (2006) found increased activity in the right posterior MTG and STS, and bilateral inferior/middle frontal gyrus, when participants' attention was directed to the emotional prosodic cues of spoken stimuli versus the content of the words. In the same vein, Frühholz et al. (2012) found distinct responses in the mid-STG, left inferior frontal gyrus, anterior cingulate cortex, and amygdala when discriminating emotions (versus gender) in prosodic stimuli, thus further documenting task effects on neural responses to vocal emotions. Nonetheless, these studies do not offer sufficient temporal resolution to elucidate when in the time course of vocal emotion perception these task effects take place.

Effects of attention on early stages of auditory information processing have been documented (e.g., Hink et al., 1978; Woldorff et al., 1993). However, in the context of emotion research specifically, visual ERP evidence indicated that the modulatory role of attention is limited to later stages reflected in the LPP (Chen et al., 2018; Ferrari et al., 2008; Schindler & Kissler, 2016), suggesting that the earlier stages of emotion information processing are automatic to an important extent. The few available ERP studies on vocal emotion (speech prosody) indicated that task focus manipulations did not significantly impact neither early (Garrido-Vásquez et al., 2013) nor late processing stages (Paulmann et al., 2013), highlighting the automaticity of these processes. Therefore, based on the existing research, it remains to be specified whether, and at which processing stages, a listener's attentional focus can affect how emotional information is decoded from voices.

One way to study the role of attention in voice perception is by manipulating the amount of visual information available in the environment, for instance by blindfolding participants (Landry et al., 2013; Lewald, 2007) or by asking them to close their eyes (Wöstmann et al., 2020). Visual information can strongly affect how information in other sensory modalities is perceived (e.g., McGurk effect, Colavita effect), and such crossmodal interactions can start at early processing stages (Koelewijn et al., 2010). Research focused on the effects of temporary visual deprivation on other sensory functions has demonstrated improvements in auditory (Fengler et al., 2015; Gibby et al., 1970; Landry et al., 2013; Lewald, 2007; Tabry et al., 2013) and tactile perceptual tasks (Facchini & Aglioti, 2003; Merabet et al., 2008). For instance, short-term visual deprivation improves the perception of harmonicity (Landry et al., 2013), loudness and pitch (Gibby et al., 1970), and it reduces inaccuracies in auditory spatial localization tasks (Lewald, 2007). Furthermore, short-term visual deprivation was found to result in enhanced activity in the visual cortex (Boroojerdi, 2000; Fierro et al., 2005), which in turn may support the processing of nonvisual information (Merabet et al., 2008). Nevertheless, it remains to be clarified if the reported advantages in the auditory modality can be generalized to the perception of voice acoustic information, since crucial differences exist in the processing of vocal versus nonvocal sounds (e.g., vocal stimuli are acoustically more complex and broadband, being processed in specialized regions of the superior temporal sulcus – Belin et al., 2011). The existing research also does not clarify when, in the course of processing, a potential effect of visual deprivation on auditory perception may take place.

1.3. The current study

Using ERPs, we examined the time course of brain responses to spontaneous and volitional nonverbal vocalizations. We further asked whether these responses are affected by attention, via two orthogonal manipulations: task focus (authenticity versus emotion detection) and visual condition (standard visual condition versus visual deprivation). Vocalizations included laughter and crying, thereby allowing us to probe the generalizability of authenticity and attention effects across positive and negative expressions. As for the task focus manipulation, in one condition participants were asked to judge the emotional authenticity of vocalizations, whereas in the other they focused on the emotional category of the sounds. Approximately half the participants were assigned to a visual deprivation condition, in which they were blindfolded, whereas the other half were assigned to a standard visual condition.

Based on previous ERP experiments with actor portrayals (e.g., Jessen & Kotz, 2011; Liu et al., 2012; Pell et al., 2015), our focus was on the auditory N1, P2, and LPP components. These components are reliable online measures of the processing stages underlying vocal emotional processing: sensory, salience detection, and cognitive evaluative stages, respectively (Jessen & Kotz, 2011; Liu et al., 2012; Pell et al., 2015; Pinheiro et al., 2014). We expected authenticity to modulate early sensory (N1), salience detection (P2), and late cognitive (LPP) stages of vocal emotion perception. Effects of authenticity at early sensory and salience detection processing stages were expected because there is evidence of important acoustic and affective differences between spontaneous and volitional vocalizations (Anikin and Lima, 2018; Bryant & Aktipis, 2014; Lavan et al., 2016), and we know that the N1 and P2 are sensitive to the physical acoustic features (Näätänen & Picton, 1987; Seither-Preisler et al., 2006) and affective properties (Jessen & Kotz, 2011; Liu et al., 2012; Pell et al., 2015) of sounds. As for the direction of these effects, a tentative hypothesis was that spontaneous expressions would elicit reduced N1 and increased P2, a pattern in line with the one previously obtained for emotionally salient compared to neutral vocal expressions (Jessen & Kotz, 2011; Liu et al., 2012; Sauter & Eimer, 2010). This would occur because spontaneous expressions contain acoustic hallmarks typically associated with high emotional significance, which are absent in volitional ones (Lavan et al., 2016; Scott et al., 2014). Authenticity modulations at late processing stages, reflected in enhanced LPP for spontaneous expressions, would be consistent with evidence showing that authentic vocalizations tend to be perceived as more arousing than volitional ones (Anikin and Lima, 2018; Bryant & Aktipis, 2014; Lavan et al., 2016), and we know that this component is sensitive to arousal (Cuthbert et al., 2000; de Rover et al., 2012). An authenticity effect at later evaluative stages would be also consistent with behavioral evidence that listeners are able to reliably discriminate spontaneous from volitional vocalizations in explicit evaluation tasks (e.g., Anikin and Lima, 2018; Bryant & Aktipis, 2014; Lavan et al., 2016; Neves et al., 2018).

As for potential effects of attention, our approach was primarily exploratory. Since the N1 is enhanced by attention (e.g., Hink et al., 1978; Woldorff et al., 1993), it seems plausible that attentional manipulations, both in terms of task instructions (focus on authenticity versus emotion) and visual conditions (visual deprivation versus standard visual), could modulate early processing stages. If attentional focus critically modulates distinct stages of vocal emotional processing, then facilitated processing of authenticity versus emotion information should occur when explicitly attending to authenticity. This would translate into more prominent differences between volitional and spontaneous vocalizations in early and late ERP components, when resources were allocated to authenticity as compared to emotion, whereas emotion effects would be stronger in the emotion focus condition. The presence of such effects exclusively at late processing stages would highlight the role of more deliberate processes in authenticity perception. By contrast, if the processing of vocal emotional information is highly automatic and does not require conscious deliberation to be efficiently processed, comparable neural responses to authenticity and emotion manipulations would be obtained in the two task focus conditions, especially at early processing stages. Furthermore, if vocal emotional processing is affected by visual condition, stronger authenticity and emotion effects should be observed in the visual deprivation condition (i.e., when attentional load is reduced and information is processed in the auditory modality only) as compared to the standard visual condition (i.e., when processing resources are divided amongst the input of distinct sensory modalities). Conversely, if vocal emotional perception is largely automatic and is not influenced by the simultaneous processing of visual information, similar neural responses could be observed in the two visual conditions.

2. Methods

2.1. Participants

We report how we determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

Forty-three undergraduate students participated in this study for course credit. The size of our sample was determined based on convenience/opportunity. Participants were randomly assigned to one of two conditions: visual deprivation, VD ($n = 22$; 7 males, $M_{age} = 21.14$, $SD = 4.09$); or standard visual condition, SV ($n = 21$; 6 males, $M_{age} = 20.33$, $SD = 3.65$). In the VD condition, participants were blindfolded in a dark and electrically shielded room and performed the auditory tasks with no concurrent visual information. In the SV condition, participants performed the auditory tasks inside an electrically shielded room with typical concurrent visual information signaling the beginning of each trial and the questions. For both experimental conditions, the inclusion criteria were: being right-handed (Oldfield, 1971); normal or corrected-to-normal visual acuity; normal hearing; no history of electroconvulsive treatment, neurological illness, or DSM-IV diagnosis of drug or alcohol abuse; and no current medication with potential impact on the electroencephalogram (EEG), or on neurological and/or cognitive functioning. The inclusion/exclusion criteria were established prior to data analyses. Participants provided their informed consent, previously assessed by the local Ethics committee from the Faculty of Psychology at the University of Lisbon (Lisbon, Portugal).[1]

2.2. Stimuli

The experimental stimuli consisted of 80 nonverbal vocalizations portraying amusement (40 laughs) or sadness (40 cries). Each emotional category comprised 20 volitional and 20 spontaneous stimuli, recorded by six speakers (three women) within an anechoic chamber at the University College London. To record volitional laughter and crying, speakers were asked to voluntarily produce these expressions, without a corresponding emotional eliciting event, and to make them sound as natural as possible. Spontaneous laughter was induced through an amusement induction condition, whereby speakers watched self-selected amusing video clips, which they considered funny and would easily make them laugh aloud (for a similar procedure, see McGettigan et al., 2015). Spontaneous crying was evoked by using an emotion induction procedure: speakers were asked to recall difficult (upsetting) past episodes and/or to initially pose crying to promote a shift into spontaneous crying linked with genuine experienced sadness (see Lavan et al., 2016). Of note, feelings of amusement and sadness throughout and after recording the corresponding genuine vocalizations were reported by the six speakers. These vocalizations have been previously used

[1] The conditions of our ethics approval do not permit sharing of the data with any individual outside the author team under any circumstances.

in behavioral and neuroimaging experiments (Lavan et al., 2015, 2016; Lima et al., 2016; O'Nions et al., 2017; Neves et al., 2018). These vocalizations are available at https://osf.io/hysf3/. The acoustic and affective features of the stimuli are summarized in Table 1.

2.3. Procedure

Participants were tested in individual sessions lasting approximately 40 min (breaks included). They were seated in a comfortable chair at a distance of 100 cm from the computer monitor, in an electrically shielded and sound attenuated room. Voice stimuli were binaurally presented through headphones. Presentation® software (Version 20.1, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com/) was used to control stimulus timing and presentation, as well as to register participants' responses (the code is available here: https://osf.io/hysf3/

Each participant completed two tasks, one involving an authenticity evaluation (volitional versus spontaneous) and the other an emotion evaluation (sadness versus amusement). The order of tasks was counterbalanced across participants. In each of them, the 80 vocalizations were randomly presented twice, originating a total of 160 trials per task. Each trial was presented as follows (see Fig. 1 for details): 1) a warning sound (VD condition) or attention mark (SV condition) signaled the beginning of each trial (100 msec); 2) after a varying inter-stimulus interval (ISI; 500–1500 msec), a vocalization (<3000 msec) was presented; 3) after a 1000 msec interval, a warning sound (VD condition; 150 msec) or the written question (SV condition; 150 msec) was presented to signal the beginning of the response time (<3000 msec). The inter-trial interval (ITI) lasted 1000 msec. In the SV condition, a fixation cross (presented centrally on the screen) remained until the end of the trial to minimize eye movements. No part of the study procedures was pre-registered prior to the research being conducted.

2.4. EEG data acquisition

EEG data were recorded using a 64-channel BioSemi Active Two System (http://www.biosemi.com/products.htm) at a digitization rate of 512 Hz. Electrodes placed at left and right temples (horizontal electrooculogram – EOG) and one below the left eye (vertical EOG) were used to monitor eye movements. Electrodes were also placed at left and right mastoids for offline referencing.

Brain Vision Analyzer 2.0.4 software (www.brainproducts.com/) was used for offline analysis of EEG data. A .01 Hz high-pass filter was applied. EEG data were referenced offline to the average of the left and right mastoids. Individual ERP epochs of 1700 ms, time-locked to voice onset, were created and included a −200 msec pre-stimulus baseline. Ocular artifacts were corrected based on Gratton et al. (1983). Individual epochs containing excessive eye blinks or movement artifacts (±100 μV criterion) were excluded from the analyses. After artifact rejection, individual ERP averages were based on a minimum of 75% of segments per condition for each subject (visually deprived condition [authenticity focus condition: volitional crying = 37.23 ± 1.82; volitional laughter = 37.86 ± 2.73; spontaneous crying = 37.68 ± 2.34; spontaneous laughter = 38.05 ± 2.19; emotion focus condition: volitional crying = 37.36 ± 3.80; volitional laughter = 37.18 ± 3.03; spontaneous crying = 37.55 ± 2.89; spontaneous laughter = 37.68 ± 2.77]; standard visual condition [volitional crying = 38.52 ± 2.64; volitional laughter = 38.33 ± 2.56; spontaneous crying = 38.000 ± 3.15; spontaneous laughter = 38.33 ± 2.31; emotion focus condition: volitional crying = 38.19 ± 2.79; volitional laughter = 37.31 ± 2.83; spontaneous crying = 37.52 ± 2.71; spontaneous laughter = 33.19 ± 1.89]). The number of epochs included in the averages did not differ per condition ($P > .05$).

Based on previous auditory ERP studies and on visual inspection of grand-averaged waveforms, the following time windows were selected for analysis of mean amplitudes of the N1, P2, and LPP: 130–170 msec (N1 – Lu, Ho, Liu, Wu, & Thompson, 2015; Pinheiro et al., 2014), 220–280 msec (P2 – Masuda et al., 2018; Pinheiro et al., 2014), 450–700 msec (early LPP – Jessen & Kotz, 2011; Masuda et al., 2018; Pell et al., 2015), 700–1000 msec (middle LPP – Brown & Cavanagh, 2017; Masuda et al., 2018), and 1000–1400 msec (late LPP – Brown & Cavanagh, 2017; Brown et al., 2012; Spreckelmeyer et al., 2006) after vocalization onset. For the N1 and P2 components, the analysis included fronto-central (FC1, FCz, FC2) and central (C1, Cz, C2) channels, consistent with previous studies (e.g., Liu et al., 2012; Pell et al., 2015; Pinheiro et al., 2017; Rigoulot,

Table 1 – Acoustic and affective properties of the vocalizations.

Acoustic properties	Authenticity			
	Spontaneous		Volitional	
	Amusement	Sadness	Amusement	Sadness
*f*0 (Hz)	270.78	287.53	228.78	260.90
*f*0 min (Hz)	171.96	180.51	115.15	125.23
*f*0 max (Hz)	370.69	385.06	319.48	392.65
Duration (ms)	2402	2689	2270	2520
Intensity (dB)	66.09	65.86	66.03	66.09
Affective properties				
Valence	6.48 *(1.18)*	3.37 *(.67)*	5.63 *(.98)*	3.29 *(.67)*
Arousal	6.52 *(1.18)*	5.74 *(1.08)*	5.04 *(1.02)*	5.18 *(1.09)*
Authenticity	5.90 *(1.12)*	5.15 *(1.05)*	4.12 *(.98)*	3.90 *(.98)*

Note. SDs are given in italic.

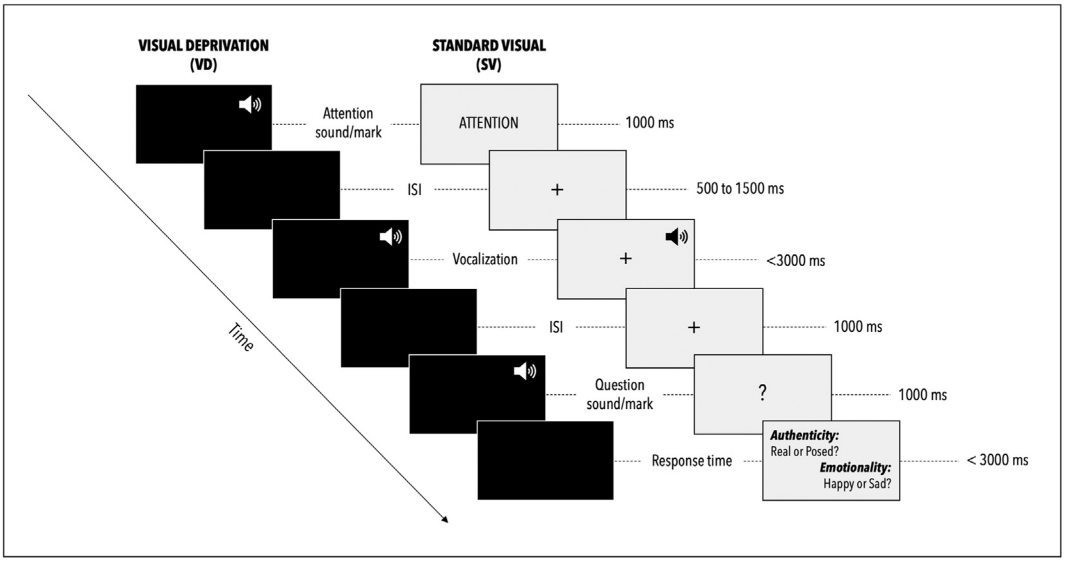

Fig. 1 – Illustration of an experimental trial.

Pell, & Armony, 2015). For the LPP, the analysis considered parietal (P1, Pz, P2) and parieto-occipital (PO3, POz, PO4) electrodes (e.g., Brown & Cavanagh, 2017; Masuda et al., 2018; Pell et al., 2015; Spreckelmeyer et al., 2006).

2.5. *Statistical analyses*

Behavioral and ERP data were analyzed with linear mixed-effects models, with lme4 (Bates et al., 2014) and lmerTest (Kuznetsova et al., 2016) packages in R-Studio (Version 1.4.1717; RStudio Team, 2021). When compared with traditional statistical analyses (e.g., repeated-measures analysis of variance), the linear mixed-effects models have the advantage of considering the variance associated with both fixed and random factors (e.g., random effects for participants) in ERP measures (Jaeger, 2008). The default variance-covariance structure (unstructured matrix) was used (Bates et al., 2015). Significant interactions were followed up by comparisons between theoretically relevant conditions, with Bonferroni-corrected pairwise comparisons using the "emmeans" R package (Lenth et al., 2018). Confidence intervals were computed using "boot" R package (Canty & Ripley, 2021). No part of the study analyses was pre-registered prior to the research being conducted.

2.5.1. *Behavioral data*

The accuracy of authenticity and emotion judgements was separately analyzed with two distinct linear mixed-effects models. In these models, we tested the interactive effects of authenticity and emotion on the accuracy of authenticity and emotion discrimination. Trial-by-trial accuracy was included as outcome, participants were included as random effects, whereas authenticity (volitional, spontaneous), emotion (sadness, amusement), and visual condition (VP, VD) were included as fixed effects.

Furthermore, we also probed whether visual condition impacted accuracy of authenticity and emotion discrimination by testing the interaction between authenticity, emotion, and visual condition (models tested: Variable ~ as.factor(Authenticity) * as.factor(EmotionType) * as.factor(VisualCond) + (1 | Participant); analysis code is available here: https://osf.io/hysf3/

2.5.2. *ERP data*

We used a multistep approach to test our hypotheses in an attempt to reduce the complexity of the statistical models and to make them as parsimonious as possible. First, we probed whether emotional authenticity modulated the distinct processing stages of vocal emotion, by testing the interactive effects of authenticity and emotion on the mean amplitudes of N1, P2, and LPPs (early, middle, and late LPP time windows). We computed distinct linear mixed-effects models for the mean amplitude of each component. The mean amplitude for each component was included as outcome, participants were included as random effects, whereas authenticity (volitional, spontaneous), emotion (sadness, amusement), task focus (authenticity, emotion), and visual condition (VP, VD) were included as fixed effects.

Additionally, to explore the potential effects of attention on the neural processing of authenticity, we first tested the interaction between authenticity, emotion, and task focus (model tested: Variable ~ as.factor(Authenticity) * as.factor(EmotionType) * as.factor(TaskFocus) + as.factor(VisualCond) + (1 | Participant)), and then the interactive effects of authenticity, emotion, and visual condition (model tested: Variable ~ as.factor(Authenticity) * as.factor(EmotionType) * as.factor(VisualCond) * as.factor(TaskFocus) + (1 | Participant)).

3. Results

3.1. *Behavioral data*

3.1.1. *Authenticity detection*

The mean accuracy of authenticity detection per emotion type is presented in Table 2.

3.1.1.1. Interactive effects of authenticity and emotion. Authenticity discrimination was modulated by an interaction between authenticity and emotion (β = .61, SE = .022, t(6837) = 13.021, p < .001, 95% CI: [.238, .326]). Follow-up analyses indicated that accuracy in authenticity discrimination was higher for crying (versus laughter) in the case of volitional expressions (p < .001), and for laughter (versus crying) in the case of spontaneous expressions (p < .001). The effect of visual condition was significant (β = .10, SE = .022, t(43) = 2.452, P = .018, 95% CI: [.006, .084]), indicating reduced authenticity detection accuracy in the visually deprived versus standard visual condition.

3.1.1.2. Interactive effects of authenticity and visual condition. The interaction between authenticity and visual condition (β = −.33, SE = .03, t(6837) = −5.025, p < .001, 95% CI: [−.220, −.096]) was significant (see supplemental material). Follow-up analyses demonstrated differences in how participants judged vocal authenticity as a function of the presence versus absence of concurrent visual stimulation, but which were specific of volitional expressions: participants who performed the experiment in the visually deprived condition were less accurate than those in the SV condition at detecting the authenticity of volitional expressions (p < .001), but no group effects were found for spontaneous vocal sounds (P = .136).

3.1.2. *Emotion detection*

The accuracy of emotion decoding per emotion type is presented in Table 3.

3.1.2.1. Interactive effects of authenticity and emotion. Emotion discrimination was modulated by an interaction between authenticity and emotion type (β = .50, SE = .01, t(6837) = 10.789, p < .001, 95% CI: [.115, .169]). Follow-up pairwise comparisons indicated increased accuracy in emotion decoding from volitional (versus spontaneous) cries (p < .001); however, participants were similarly accurate in decoding the emotional meaning of volitional and spontaneous laughs (P > .999).

3.1.2.2. Interactive effects of authenticity and visual condition. Emotion discrimination was modulated by an interaction between authenticity and visual condition (β = −.23, SE = .02, t(6837) = −3.495, p < .001, 95% CI: [−.102, −.026]), and by an interaction between authenticity, emotion and visual condition (β = .29, SE = .03, t(6837) = −3.142, p = .002, 95% CI: [.034, .130]). Follow-up pairwise comparisons indicated that participants were more accurate at decoding emotion from volitional (versus spontaneous) cries in both VD (p < .001) and SV conditions (p < .001); however, no significant differences were found in the case of laughs (all p's > .999).

3.2. *ERP data*

3.2.1. *N1 component*

3.2.1.1. Interactive effects of authenticity and emotion. N1 amplitude was modulated by an interaction between authenticity and emotion (β = .19, SE = .160, t(4085) = 4.319, p < .001, 95% CI: [.384, 1.069]). Post-hoc pairwise comparisons revealed that the N1 was reduced (i.e., less negative) for spontaneous (versus volitional) laughs (p < .001) but no differences were found in the case of crying (P = .448) (see Figs. 2 and 3).

3.2.1.2. Interactive effects of authenticity and task focus. The model revealed that authenticity interacted with emotion and task focus (β = .17, SE = .319, t(4085) = 1.962, p = .0499, 95% CI: [−.010, 1.246]; see supplemental material). Post-hoc pairwise comparisons confirmed the specificity of authenticity effects for laughter when the focus was both on authenticity (p < .001) and emotion (p < .001); however, for crying similar authenticity effects were found irrespective of task instructions (lowest P = .349).

3.2.1.3. Interactive effects of authenticity and visual condition. Authenticity interacted with visual condition (β = −.18, SE = .225, t(4085) = −2.951, p = .003, 95% CI: [−1.068, −.254]. Nevertheless, post-hoc pairwise comparisons revealed a similar N1 response in both visual conditions: N1 was overall more negative to volitional compared to spontaneous expressions both in the VD (p < .001) and SV conditions (p < .001).

3.2.2. *P2 component*

3.2.2.1. Interactive effects of authenticity and emotion. Authenticity and emotion interactively modulated P2 amplitude (β = −.58, SE = .180, t(4085) = −11.662, p < .001, 95% CI: [−2.472, −1.734]) (see Figs. 2 and 3). Follow-up analyses showed that the P2 was increased in response to spontaneous compared to volitional cries (p < .001), whereas for laughter the reverse pattern was found (P = .001).

3.2.2.2. Interactive effects of authenticity and task focus. No significant interactions between task focus and authenticity (β = .14, SE = .254, t(4085) = 1.950, P = .051, 95% CI: [−.49, 1.010])

Table 2 – Average accuracy scores in the authenticity detection task, considering emotion type and visual condition.

Authenticity	Emotion	Condition			
		Visual Deprivation - VD		Standard Visual - SV	
		Hits	*False alarms*	*Hits*	*False alarms*
Spontaneous	Amusement	.83 (*.19*)	.47 (*.18*)	.75 (*.12*)	.32 (*.20*)
	Sadness	.66 (*.14*)	.32 (*.20*)	.64 (*.15*)	.11 (*.09*)
Volitional	Amusement	.53 (*.15*)	.17 (*.08*)	.68 (*.20*)	.25 (*.12*)
	Sadness	.68 (*.20*)	.34 (*.14*)	.89 (*.09*)	.36 (*.15*)

Note. SDs are given in italic.

Table 3 – Average accuracy scores in the emotion detection task, considering authenticity and visual condition.

Emotion	Authenticity	Condition			
		Visual Deprivation - VD		Standard Visual – SV	
		Hits	*False alarms*	*Hits*	*False alarms*
Amusement	Spontaneous	.94 (*.06*)	.17 (*.17*)	.95 (*.05*)	.20 (*.11*)
	Volitional	.95 (*.05*)	.07 (*.12*)	.95 (*.05*)	.05 (*.04*)
Sadness	Spontaneous	.83 (*.17*)	.06 (*.06*)	.78 (*.09*)	.05 (*.05*)
	Volitional	.93 (*.12*)	.05 (*.05*)	.95 (*.04*)	.05 (*.05*)

Note. SDs are given in italic.

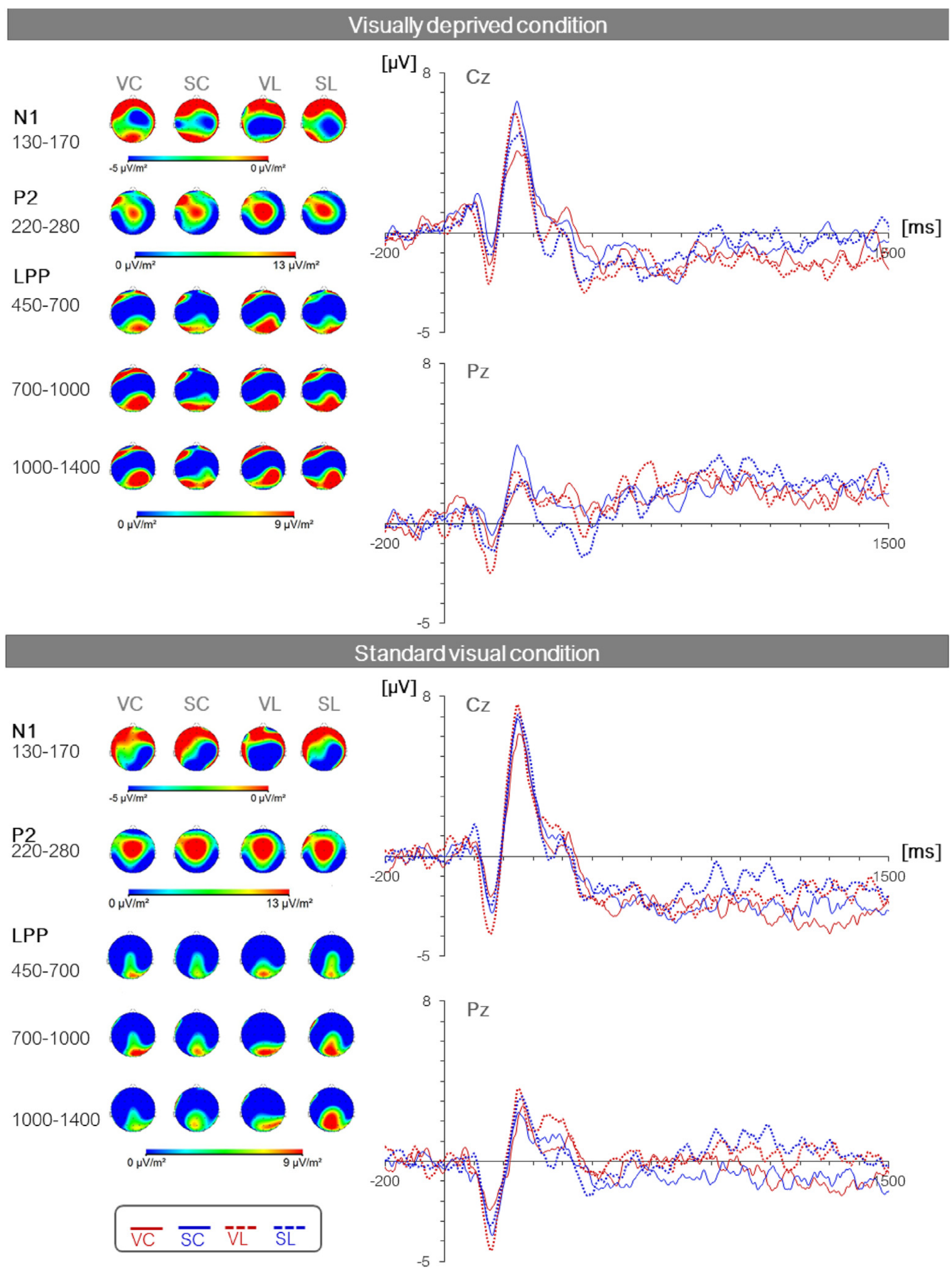

Fig. 2 – Grand average ERP waveforms for spontaneous and volitional vocalizations in the authenticity detection task, over Cz and Pz electrodes.

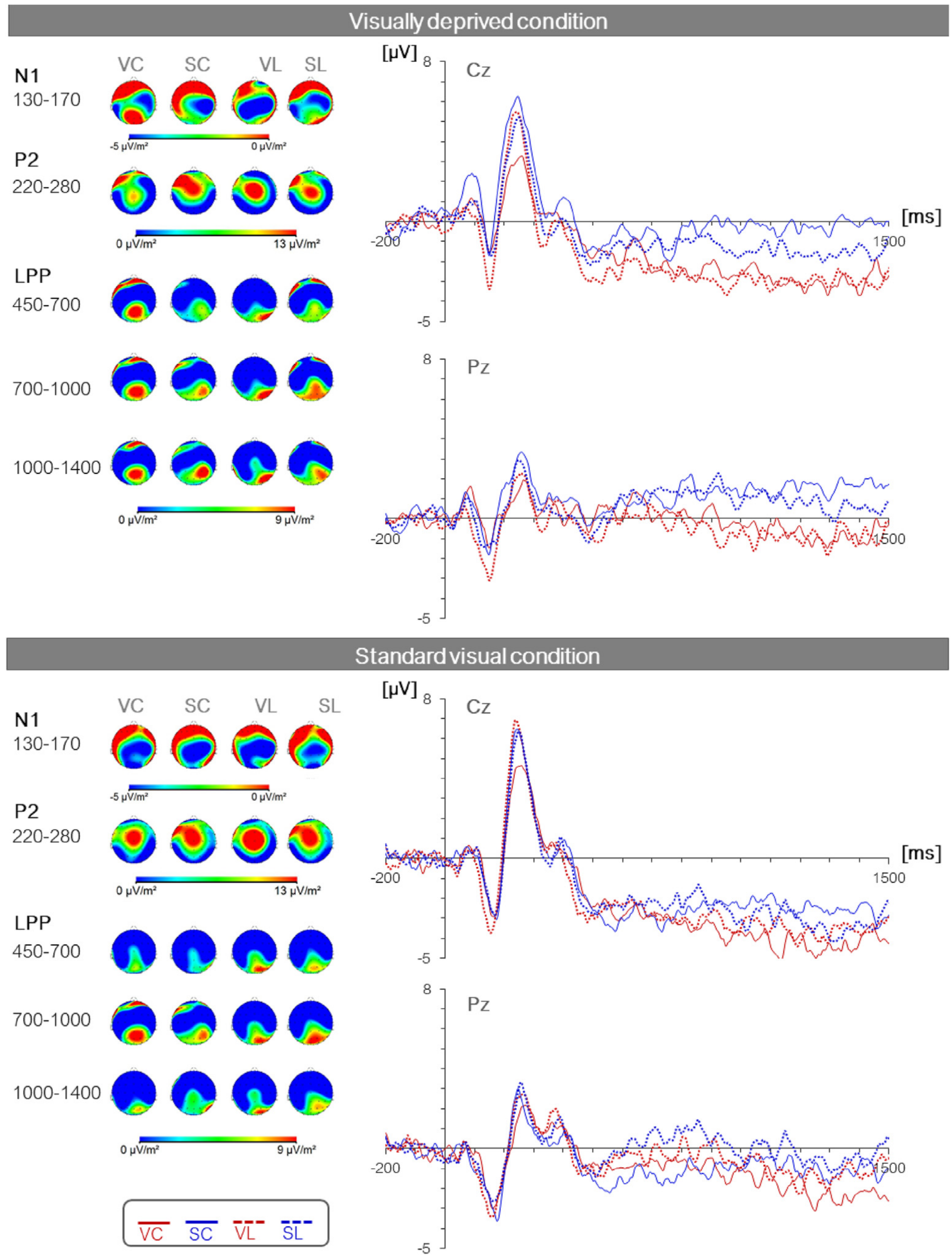

Fig. 3 – Grand average ERP waveforms for spontaneous and volitional vocalizations in the emotion detection task, over Cz and Pz electrodes.

or emotion ($\beta = .14$, SE $= .254$, $t(4085) = 1.950$, $P = .052$, 95% CI: [−.056, 1.010]) were observed, indicating that the impact of authenticity on the P2 was independent of task focus. However, task focus generally modulated P2 amplitude, as shown by a main effect of this factor ($\beta = -.22$, SE $= .1795$, $t(4085) = -4.387$, $p < .001$, 95% CI: [−1.169, −.421]): the P2 was increased when attention was focused on authenticity versus emotion discrimination.

3.2.2.3. Interactive effects of authenticity and visual condition. The P2 was modulated by an interaction between authenticity, emotion, and visual condition ($\beta = .29$, SE $= .358$, $t(4085) = 2.933$, $p = .003$, 95% CI: [.308, 1.743]). Post-hoc pairwise comparisons indicated that the P2 was enhanced in response to authentic compared to volitional cries both in the VD ($p < .001$) and SV conditions ($p < .001$). For laughs, however, the P2 was increased for volitional relative to spontaneous vocalizations in the SV condition ($P = .020$), but not in the VD condition ($P > .999$). That is, authenticity effects on the perception of laughter were specific of the standard visual condition.

3.2.3. *Early LPP (450–700 ms)*

3.2.3.1. Interactive effects of authenticity and emotion. The early LPP amplitude was not significantly predicted by authenticity ($\beta = -.04$, SE .127, $t(4085) = -1.059$, $P = .290$, 95% CI: [−.371, .144]), emotion ($\beta = -.06$, SE = .127, $t(4085) = -1.585$, $P = .113$, 95% CI: [−.450, .0665]) or interactions between authenticity and emotion ($\beta = .05$, SE = .180, $t(4085) = .887$, $P = .375$, 95% CI: [−.204, .490]) (see Figs. 2 and 3).

3.2.3.2. Interactive effects of authenticity and task focus. The amplitude was modulated by an interaction between authenticity, emotion, and task focus ($\beta = .22$, SE = .254, $t(4085) = 2.146$, $P = .032$, 95% CI: [.100, 1.407]). Nonetheless, post-hoc pairwise comparisons indicated no significant differences in the early LPP response as a function of authenticity and emotion in both tasks (lowest $P = .206$). Notwithstanding, the main effects of task focus ($\beta = -.14$, SE = .090, $t(4085) = -2.802$, $P = .005$, 95% CI: [−.838, −.174]) was significant: the LPP amplitude was generally more positive when the focus was on authenticity versus emotion discrimination.

3.2.3.3. Interactive effects of authenticity and visual condition. The amplitude was modulated by an interaction between authenticity and visual condition ($\beta = -.24$, SE = .253, $t(4085) = -3.258$, $P = .001$, 95% CI: [−1.266, −.336]). Nonetheless, post-hoc pairwise comparisons demonstrated no significant amplitude modulations as a function of authenticity in both the VD ($P = .883$) and SV conditions ($P = .109$). The interaction between emotion and visual condition was significant ($\beta = .17$, SE = .253, $t(4085) = 2.303$, $P = .021$, 95% CI: [.044, 1.128]). Post-hoc pairwise comparisons revealed that the early LPP was increased in response to crying (versus laughter) in the VD condition ($p < .001$), whereas it was enhanced in response to laughter in the SV condition ($P = .037$).

3.2.4. *Middle LPP (700–1000 msec)*

3.2.4.1. Interactive effects of authenticity and emotion. The middle LPP time window was modulated by an interaction between authenticity and emotion ($\beta = .21$, SE = .21, $t(4085) = 4.062$, $p < .001$, 95% CI: [.486, 1.262]) (see Figs. 2 and 3). Post-hoc pairwise comparisons revealed enhanced amplitude in response to authentic compared to volitional laughs ($p < .001$), but no significant authenticity effects were found for crying ($P > .999$).

3.2.4.2. Interactive effects of authenticity and task focus. No significant interactions involving task focus were identified (lowest $P = .062$), indicating that the impact of authenticity at this stage of processing is independent of directing attention to authentic versus emotional properties of the voice. However, a main effect of task focus ($\beta = -.24$, SE = .211, $t(4085) = -4.733$, $p < .001$, 95% CI: [−1.416, −.591]) revealed generally increased LPP when the task was focused on authenticity versus emotional quality of the vocalizations.

3.2.4.3. Interactive effects of authenticity and visual condition. The amplitude was modulated by an interaction between authenticity and visual condition ($\beta = -.29$, SE = .297, $t(4085) = -3.929$, $p < .001$, 95% CI: [−1.758, −.634]). Post-hoc pairwise comparisons revealed that amplitude was enhanced in response to spontaneous compared to volitional expressions in the VD condition ($p < .001$), but not in the SV condition ($P > .999$).

3.2.5. *Late LPP (1000–1400 ms)*

3.2.5.1. Interactive effects of authenticity and emotion. We observed that the late LPP was significantly predicted by authenticity ($\beta = .20$, SE = .165, $t(4085) = 5.45$, $p < .001$, 95% CI: [.566, 1.207]) (see Figs. 2 and 3): the amplitude was more positive for spontaneous versus volitional expressions, regardless of the emotional quality of the vocalizations. Further, the late LPP amplitude was not modulated by emotion ($\beta = .04$, SE = .165, $t(4085) = 1.199$, $p = .230$, 95% CI: [−.126, .512]) or interactions involving this factor ($p = .514$).

3.2.5.2. Interactive effects of authenticity and task focus. The model revealed that the late LPP was predicted by an interaction between authenticity and task focus ($\beta = .16$, SE = .329, $t(4085) = 2.126$, $P = .034$, 95% CI: [−.016, 1.357]). Nonetheless, post-hoc pairwise comparisons showed more positive amplitude for spontaneous versus volitional vocalizations both when attention was focused on authenticity ($p < .001$) and emotion ($p < .001$). That is, authenticity effects were similar across distinct task instructions.

3.2.5.3. Interactive effects of authenticity and visual condition. Visual condition interacted with authenticity ($\beta = -.18$, SE = .328, $t(4085) = -2.430$, $p = .015$, 95% CI: [−1.483, −.216]) and with emotion ($\beta = .19$, SE = .328, $t(4085) = 2.570$, $p = .010$, 95% CI: [.167, 1.450]). Post-hoc pairwise comparisons revealed that the amplitude was more positive for spontaneous compared to volitional expressions both in the VD ($P < .001$) and SV conditions ($P < .001$). This finding indicates that the impact of authenticity at this stage of processing is similar in both visual conditions. Further, they showed that the amplitude was increased for laughs (versus cries) in the standard visual condition ($p <. 001$), whereas in the VD group no significant emotion effects were observed ($P = .599$).

4. Discussion

The current study examined whether and how the authenticity of emotional vocalizations impacts ERP responses at distinct vocal processing stages. It also probed the extent to which these effects are affected by attention (i.e., task focus and visual condition).

4.1. *Early processing stages*

The auditory N1 component indexes early sensory acoustic processing (Näätänen & Picton, 1987) and is modulated by attention (Hink et al., 1978; Woldorff et al., 1993). At this processing stage, we found that authenticity modulated the N1 but only in the case of laughter: spontaneous laughs were associated with a suppression of the N1 amplitude compared to volitional expressions. Considering the existing evidence on the functional significance of the N1 (Liu et al., 2012; Meyer et al., 2007; Pell et al., 2015), this finding suggests an early preferential sensory analysis of spontaneous (versus volitional) laughs. Crucially, the authenticity effects (reduced N1 for spontaneous versus volitional laughs) were observed across distinct task instructions and visual conditions. This suggests that they are automatic to an important extent.

These findings confirm our hypothesis of authenticity effects on early sensory processing stages.

Since the N1 is sensitive to the physical acoustic features of the stimulus (Näätänen & Picton, 1987; Seither-Preisler et al., 2006), the observed N1 modulation by laughter authenticity is likely to reflect differences in the physical acoustic temporal and spectral profiles of spontaneous versus volitional laughs. For instance, spontaneous and volitional laughs were shown to differ in terms of fundamental frequency, duration, length of bursts, duration and length of inter-bursts intervals, as well as in the involvement of supralaryngeal control mechanisms (Anikin and Lima, 2018; Bryant & Aktipis, 2014; Lavan et al., 2016; Vettin & Todt, 2004). The N1 is thought to be originated from neural sources in the STG and supratemporal plane (Näätänen & Picton, 1987). The increased N1 for volitional versus spontaneous laughter is consistent with previous literature showing distinct responses for volitional and spontaneous laughs in the bilateral STG (Lavan et al., 2017; McGettigan et al., 2015), and extend these studies by showing that this dissociation initiates early in voice processing. On the other hand, this initial sensory processing stage was unaffected by the authenticity of crying. This suggests that within the first 170 msec post-stimulus onset, the acoustic differences between spontaneous and volitional crying might not be sufficiently prominent to modulate authenticity decoding, and that more time (and hence, a greater amount of acoustic information) is needed for the discrimination.

In the context of vocal emotion processing, the P2 component reflects the detection and integration of emotionally salient acoustic cues (Liu et al., 2012; Paulmann & Kotz, 2008; Pell et al., 2015; Schirmer & Kotz, 2006). The interactive effects between authenticity and emotion on the P2 revealed that emotion modulated authenticity perception in processing stages typically associated with salience detection. Specifically, our findings of enhanced P2 for spontaneous compared to volitional crying suggested facilitated detection of emotional salience from spontaneous crying. Considering previous suggestions that the P2 is enhanced for highly arousing stimuli (Paulmann et al., 2013; Sauter & Eimer, 2010), our observation of increased P2 for spontaneous crying could potentially reflect differences in the arousal properties of spontaneous versus volitional crying. This interpretation is consistent with previous suggestions that the perception of emotional authenticity relies on arousal properties (Anikin and Lima, 2018). Nevertheless, the arousal ratings obtained from an independent sample (see Pinheiro et al., 2021) revealed no significant differences between the spontaneous and volitional cries used in our study, which suggests that other properties modulated the P2 sensitivity to crying authenticity or that stimulus arousal was perceived differently by the sample tested in our study.

The fact that authenticity effects on crying were independent of task focus (on authenticity versus emotion) suggests that crying authenticity is detected in a relatively automatic manner in early processing stages. Evidence for the automatic processing of emotionally salient information from nonverbal vocalizations was previously reported (Lima et al., 2019; Pinheiro et al., 2016b). Nonetheless, laughter authenticity perception was modulated by the manipulation of visual condition: the P2 was increased for volitional compared to spontaneous laughs in the SV condition only, which indicates that laughter authenticity was processed earlier in the standard visual condition. These observations suggest heightened sensitivity of salience detection mechanisms in the condition with concurrent visual input, thus diverging from our initial hypothesis. It might be the case that the manipulation of visual condition, in which sighted participants were blindfolded, might have created a particularly unusual and potentially disadvantageous context for task performance, despite reducing attentional load.

Together, the N1 and P2 findings demonstrate a dynamic interplay between authenticity and emotion salient information during the early stages of vocal emotion processing. They suggest that the time course of authenticity processing depends on emotion, with authenticity properties being processed earlier for laughter (at N1 range) than for crying (P2).

4.2. *Late processing stages*

The LPP is believed to index sustained attentional mechanisms at higher-order late processing stages, as well as the cognitive evaluation of the emotional meaning of a stimulus (Paulmann et al., 2013; Pell et al., 2015; Pinheiro et al., 2017; Schirmer et al., 2013). Our results yield evidence for authenticity effects also at late cognitive processing stages: spontaneous vocalizations elicited increased amplitude at time windows from 700 msec onwards, earlier for laughter than crying. More positive LPP amplitudes for emotional (versus neutral) stimuli have been documented for different stimulus modalities (emotional prosody – Paulmann et al., 2013; emotional pictures – Cuthbert et al., 2000; Hajcak & Nieuwenhuis, 2006; faces – Foti et al., 2010), arguably reflecting increased sustained attention and facilitated processing of emotionally salient information. The selectivity of authenticity effects for laughter, observed in the middle LPP time window, may be linked to the important role of this type of vocalization in positive social interactions, and may indicate that authenticity is differentially processed based on emotion type. Our observation of enhanced sustained attention for spontaneous (versus volitional) vocalizations might reflect the heightened salience of spontaneous vocalizations in signaling a high arousing state of the vocalizer, as shown by previous studies (Anikin and Lima, 2018; Bryant & Aktipis, 2014; Lavan et al., 2016) and as demonstrated by the arousal ratings of the vocal stimuli used in our study Pinheiro et al., 2021). These findings converge with prior evidence from visual research (Calvo et al., 2013) showing that modulations in later cognitive processing stages (i.e., 470–720 msec) reflect sensitivity to smile authenticity. Although in our study authenticity modulations emerged from the earliest processing stages under analysis, in the study by Calvo and collaborators (2013) the effects were limited to later cognitive stages. Besides the discrepancy in the sensory modality of the stimuli (dynamic vocal versus static facial expressions), we used volitional vocalizations conveying a specific and unambiguous emotional state, whereas Calvo and colleagues compared the neural processing of genuine smiles with ambiguous facial expressions transmitting incongruous emotional information from the mouth and eyes. These differences may have contributed to the divergent findings across studies.

The fact that authenticity effects at middle and late LPP windows were observed irrespective of attentional resources being allocated to authenticity or emotion cues suggests that authenticity decoding is not facilitated by voluntarily and specifically attending to the authenticity properties of someone's voice. These findings therefore indicate that the later cognitive stages of authenticity processing are not significantly impacted by differences in task instructions (focus on authenticity versus emotion). They are in good agreement with previous ERP evidence on speech prosody showing no significant effects of task instructions on early and late ERP components (Garrido-Vásquez et al., 2013; Paulmann et al., 2013). Nevertheless, they diverge from prior evidence demonstrating dissociable neural responses to vocal emotions under highly distinct task requirements (Ethofer et al., 2006; Frühholz et al., 2012; Grandjen et al., 2005; Sander et al., 2005). In these studies, participants were asked to direct attention to (versus away from) the emotional quality of the voice. However, despite the distinct task instructions in our study, participants were still paying attention to the emotional attributes of the vocalizations, which could therefore explain why brain responses to voices did not significantly differ as a function of attention focus.

Suggesting that the later cognitive stages of emotional authenticity processing were affected by visual condition, we found some evidence for enhanced selective attention for spontaneous (versus volitional) expressions, reflected in increased LPP amplitudes (i.e., 700–1000 msec) in the visual deprivation condition. Performing the auditory task with no concurrent visual input may have enhanced the attentional and perceptual resources available (that would be otherwise directed to task-irrelevant visual information) for the processing of task-relevant auditory information, in line with previous studies (Vredeveldt et al., 2011; Wöstmann et al., 2020). This interpretation is supported by evidence demonstrating that a closed eyes (versus open) condition enhanced alpha power, a neural oscillatory measure of auditory attention, arguably reflecting processes of inhibitory control in supramodal attentional networks (Wöstmann et al., 2020). Although authenticity effects on the LPP in the visual deprivation condition seem to corroborate our hypothesis that reduced attentional load may translate into enhanced sensitivity to authenticity cues, the laughter-specific authenticity effect on the P2 in the standard visual condition pointed to the opposite direction. In fact, enhanced attention at later processing stages in the visually deprived group was not behaviorally reflected in improved discrimination of emotion or authenticity. These findings keep with Wöstmann et al., 2020 who did not find improvements in perceptual sensitivity to tone detection in an eyes closed condition, despite enhanced auditory attention. The detrimental effect of visual deprivation on authenticity discrimination of volitional expressions might be linked to our observation of enhanced sustained attention in processing stages reflecting the cognitive evaluation of voice significance. Enhanced attentional resources towards vocalizations in the VD condition might have made specific sound features of volitional expressions (otherwise neutral) abnormally salient, biasing the explicit judgement of authenticity.

Furthermore, we also found evidence for emotion effects being modulated by visual condition (between 450-700 and 1000–1400 msec post-stimulus onset). Our findings suggest enhanced cognitive processing of laughter (versus crying) in the SV condition (at 450–700 and 1000–1400 msec time windows), whereas attentional resources were enhanced in response to crying in the VD condition (at 450–700 msec time window). Even though emotion effects on the LPP have been reported for different stimulus modalities (Foti et al., 2010; Hajcak & Nieuwenhuis, 2006; Paulmann et al., 2013; Pinheiro et al., 2017), the available studies with vocalizations have produced less consistent findings (Jessen & Kotz, 2011; Pell et al., 2015). For instance, Pell et al. (2015) found non-significant differences between laughter and crying in the LPP, which may suggest that emotion effects on the LPP response are modulated by authenticity properties. Considering previous suggestions that the LPP is enhanced for highly arousing stimuli (Cuthbert et al., 2000; de Rover et al., 2012), the emotion-specific effects in the SV condition could potentially reflect the increased arousal properties of laughter (versus crying). Indeed, the subjective ratings of the stimuli (see Table 1) revealed that, when considering spontaneous and volitional vocalizations together, laughs were rated as more arousing than cries (Pinheiro et al., 2021). The selective effect of emotion in the standard visual condition, from 1000 ms onwards, could suggest heightened sensitivity of cognitive evaluative stages to the emotional quality of vocal expressions when participants are processing sounds in a more usual and standard context, i.e., with concurrent visual input. Our findings add to existing research on the effects of temporary visual deprivation on auditory perception (e.g., Landry et al., 2013; Vredeveldt et al., 2011; Wöstmann et al., 2020), and may also have implications for blindness research, which typically compares blind with blindfolded sighted participants to make the sensory conditions comparable between groups (e.g., Fairhall et al., 2017; Gamond, et al., 2017). For instance, one might claim that the approach of blindfolding sighted individuals might add a considerable bias in performance, as participants are subjected to an unusual and potentially disadvantageous context when performing a listening task without any concurrent visual input.

Our behavioral data shed light on the later stages of vocal emotion processing reflecting the evaluation and integration of stimulus emotional meaning (Schirmer & Kotz, 2006). They indicated that the authenticity and emotion of vocalizations modulated emotion recognition: emotion was better discriminated from volitional (versus authentic) cries. Previous research has reported superior emotion recognition from volitional (versus spontaneous) expressions (vocalizations – Sauter & Fisher, 2018; facial expressions – Russell, 1994; Motley & Camden, 1988). Volitional vocalizations comprise less acoustic variability than spontaneous expressions (Anikin and Lima, 2018; Lavan et al., 2015), and hence, volitional cries might exhibit a more homogeneous acoustic profile that exaggerates the acoustic distinctions between emotions. Furthermore, we found that authenticity may be differently detected based on authenticity and emotion: accuracy was increased for volitional cries (versus laughs) and for spontaneous laughs (versus cries). These results could reflect a potential bias to judge crying as less spontaneous and laughter as more ‘spontaneous’, in good agreement with earlier evidence (Anikin and Lima, 2018; Pinheiro et al., 2021). These effects of emotion in authenticity and emotion discrimination might be linked to distinct degrees of exposure to these expressions and to their highly distinct communicative functions in daily social settings. Crying is often expressed in individual (solitary) settings (Vingerhoets & Bylsma, 2016; Zeifman, 2001) and is typically a salient expression for signaling sadness (yet it might also communicate other emotions, such as intense joy), as well as for requesting empathy and prosocial behavior from the listener (Hendriks et al., 2008; Provine et al., 2009; Vingerhoets & Bylsma, 2016). Laughter, however, might express a variety of meanings and intentional states, such as joy, amusement, cheerfulness, affiliation, triumph, nervousness, dominance, taunt or schadenfreude (i.e., laughing about other's misfortune) (Szameitat et al., 2009a, 2009b; Wood et al., 2017). The more variable and complex social meaning of laughter might have increased the ambiguity of authenticity detection in volitional laughs, thus leading to reduced accuracy of judging these expressions as 'volitional' when compared to cries. Besides, laughter is typically expressed in the presence of others, promoting and reinforcing social bonds (Provine, 2001). Therefore, the tendency to be more accurate at discriminating the authenticity of spontaneous laughs (versus cries) could, therefore, be a consequence of the enhanced exposure to distinct exemplars of laughter in daily social interactions. This could explain the facilitated sensory (N1) and cognitive (middle LPP) processing of authenticity observed for laughter.

It is worth noting that our authenticity manipulation was categorical (i.e., vocalizations were assessed as either spontaneous or volitional). Nonetheless, some authors have suggested that authenticity properties are better represented as dimensional, since emotional expressions in a given social context are constrained by cultural and social norms, and therefore they represent to a greater or lesser extent 'posed' expressions (Sauter & Fischer, 2018; Scherer et al., 2011). Furthermore, given the differences in the elicitation procedure to produce the authentic vocal stimuli used in our study (i.e., amusing video clips for laughter versus mental imagery of difficult past events), one might ask to what extent the authenticity of spontaneous laughter and crying is comparable. While exposing participants to amusing videos is known to successfully induce authentic laughter, the recreation of the conditions within a laboratory setting to induce intense sadness is challenging due to important ethical constraints (Anikin and Lima, 2018; Scherer & Bänzinger, 2010). Therefore, we cannot rule out that our observations of differences in authenticity perception as a function of emotion were unrelated to the elicitation procedure of spontaneous laughter versus crying. Future studies testing authenticity effects in the context of more ecological material (e.g., Parsons et al., 2014) and including other vocalizations (e.g., fear, anger, disgust, pleasure achievement) will clarify and expand our claims. However, it should be emphasized that the rigorous experimental control over sound features of observational material might be more problematic than of spontaneous sounds induced within a laboratory setting (e.g., higher levels of background noise). Besides, the stimuli used in our study also have the advantage of involving the same number of spontaneous and volitional expressions from the same speaker, while keeping the recording conditions and sound quality equivalent between spontaneous and volitional vocalizations. In addition, future research should investigate authenticity processing in the context of distinct laughter types representing both positive (e.g., joyful laughter, tickling laughter) and negative emotions (e.g., 'schadenfreude' laughter, taunting laughter).

5. Conclusion

Together, our findings provide support for the influential role of authenticity in modulating the various stages of vocal emotion processing – early sensory (N1), salience detection (P2), and late cognitive evaluative stages (LPP), even when task-irrelevant. They additionally demonstrate that authenticity and emotion cues are simultaneously and interactively processed at early stages of voice processing. These findings add to current models of vocal emotional perception (Schirmer & Kotz, 2006) by providing evidence for an early and automatic processing of the authenticity of vocal emotions.

Credit author statement

TC: Conceptualization; Methodology; Software; Formal analysis; Investigation; Project administration; Funding acquisition; Writing – original draft.

AIC: Formal analysis.

MSR: Formal analysis.

SKS: Conceptualization; Resources.

CFL: Conceptualization; Methodology; Formal analysis; Writing - Review & Editing.

APP: Conceptualization; Methodology; Formal analysis; Project administration; Funding acquisition; Writing - Review & Editing.

Funding

This work was supported by the Bial Foundation (grant number BIAL 148/18 awarded to T.C., A.P.P., and C.F.L.) and the Portuguese Science and Technology Foundation (Fundação para a Ciência e a Tecnologia [FCT]; grant number PTDC/MHC-PCN/0101/2014 awarded to A.P.P., and a Scientific Employment Stimulus grant to T.C.).

Open practices

The study in this article earned an Open Materials badge for transparent practices. Materials and data for the study are available at https://osf.io/hysf3/?view_only=932f107e30bb47089a398320fd8c3287.

Declaration of competing interest

We have no known conflict of interest to disclose.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2022.02.016.

REFERENCES

Anikin, A., & Lima, C. F. (2018). Perceptual and acoustic differences between authentic and acted nonverbal emotional vocalizations. *The Quarterly Journal of Experimental Psychology: QJEP, 71*(3), 622–641. https://doi.org/10.1080/17470218.2016.1270976

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: linear mixed-effects models using Eigen and S4. *R package version, 1*(7), 1–23.

Belin, P., Bestelmeyer, P. E. G., Latinus, M., & Watson, R. (2011). Understanding voice perception. *British Journal of Psychology, 102*(4), 711–725. https://doi.org/10.1111/j.2044-8295.2011.02041.x

Boroojerdi, B., Bushara, K. O., Corwell, B., Immisch, I., Battaglia, F., Muellbacher, W., & Cohen, L. G. (2000). Enhanced excitability of the human visual cortex induced by short-term light deprivation. *Cerebral cortex, 10*(5), 529–534. https://doi.org/10.1093/cercor/10.5.529

Brown, D. R., & Cavanagh, J. F. (2017). The sound and the fury: Late positive potential is sensitive to sound affect. *Psychophysiology, 54*(12), 1812–1825. https://doi.org/10.1111/psyp.12959

Brown, S. B., van Steenbergen, H., Band, G. P., de Rover, M., & Nieuwenhuis, S. (2012). Functional significance of the emotion-related late positive potential. *Frontiers in human neuroscience, 6*, 33. https://doi.org/10.3389/fnhum.2012.00033

Bryant, G. A., & Aktipis, C. A. (2014). The animal nature of spontaneous human laughter. *Evolution and Human Behavior, 35*(4), 327–335. https://doi.org/10.1016/j.evolhumbehav.2014.03.003

Bryant, G. A., Fessler, D., Fusaroli, R., Clint, E., Amir, D., Chávez, B., Denton, K. K., Díaz, C., Duran, L. T., Fanćovićová, J., Fux, M., Ginting, E. F., Hasan, Y., Hu, A., Kamble, S. V., Kameda, T., Kuroda, K., Li, N. P., Luberti, F. R., Peyravi, R., & Zhou, Y. (2018). The perception of spontaneous and volitional laughter across 21 societies. *Psychological science, 29*(9), 1515–1525. https://doi.org/10.1177/0956797618778235

Calvo, M. G., Marrero, H., & Beltrán, D. (2013). When does the brain distinguish between genuine and ambiguous smiles? An ERP study. *Brain and cognition, 81*(2), 237–246. https://doi.org/10.1016/j.bandc.2012.10.009

Canty, A., & Ripley, B. D. (2021). boot: Bootstrap R (S-plus) functions. *R package version, 1*, 3–26.

Chen, Y., Zhang, D., & Jiang, D. (2018). Effects of directed attention on subsequent processing of emotions: Increased attention to unpleasant pictures occurs in the late positive potential. *Frontiers in psychology, 9*, 1127. https://doi.org/10.3389/fpsyg.2018.01127

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: Covariation with autonomic arousal and affective report. *Biological Psychology, 52*(2), 95–111. https://doi.org/10.1016/s0301-0511(99)00044-7

de Rover, M., Brown, S. B., Boot, N., Hajcak, G., van Noorden, M. S., van der Wee, N. J., & Nieuwenhuis, S. (2012). Beta receptor-mediated modulation of the late positive potential in humans. *Psychopharmacology, 219*(4), 971–979. https://doi.org/10.1007/s00213-011-2426-x

Ethofer, T., Anders, S., Erb, M., Herbert, C., Wiethoff, S., Kissler, J., Grodd, W., & Wildgruber, D. (2006). Cerebral pathways in processing of affective prosody: A dynamic causal modeling study. *Neuroimage, 30*(2), 580–587. https://doi.org/10.1016/j.neuroimage.2005.09.059

Facchini, S., & Aglioti, S. M. (2003). Short term light deprivation increases tactile spatial acuity in humans. *Neurology, 60*(12), 1998–1999. https://doi.org/10.1212/01.WNL.0000068026.15208.D0

Fairhall, S. L., Porter, K. B., Bellucci, C., Mazzetti, M., Cipolli, C., & Gobbini, M. I. (2017). Plastic reorganization of neural systems for perception of others in the congenitally blind. *Neuroimage, 158*, 126–135. https://doi.org/10.1016/j.neuroimage.2017.06.057

Fengler, I., Nava, E., & Röder, B. (2015). Short-term visual deprivation reduces interference effects of task-irrelevant facial expressions on affective prosody judgments. *Frontiers in integrative neuroscience, 9*, 31. https://doi.org/10.3389/fnint.2015.00031

Ferrari, V., Codispoti, M., Cardinale, R., & Bradley, M. M. (2008). Directed and motivated attention during processing of natural scenes. *Journal of cognitive neuroscience, 20*(10), 1753–1761. https://doi.org/10.1162/jocn.2008.20121

Fierro, B., Brighina, F., Vitello, G., Piazza, A., Scalia, S., Giglia, G., Daniele, O., & Pascual-Leone, A. (2005). Modulatory effects of low- and high-frequency repetitive transcranial magnetic stimulation on visual cortex of healthy subjects undergoing light deprivation. *The Journal of physiology, 565*(Pt 2), 659–665. https://doi.org/10.1113/jphysiol.2004.080184

Foti, D., Olvet, D. M., Klein, D. N., & Hajcak, G. (2010). Reduced electrocortical response to threatening faces in major depressive disorder. *Depression and Anxiety, 27*(9), 813–820. https://doi.org/10.1002/da.20712

Frühholz, S., Ceravolo, L., & Grandjean, D. (2012). Specific brain networks during explicit and implicit decoding of emotional prosody. *Cerebral Cortex, 22*(5), 1107–1117. https://doi.org/10.1093/cercor/bhr184

Gamond, L., Vecchi, T., Ferrari, C., Merabet, L. B., & Cattaneo, Z. (2017). Emotion processing in early blind and sighted individuals. *Neuropsychology, 31*(5), 516–524. https://doi.org/10.1037/neu0000360

Garrido-Vásquez, P., Pell, M. D., Paulmann, S., Strecker, K., Schwarz, J., & Kotz, S. A. (2013). An ERP study of vocal emotion processing in asymmetric Parkinson's disease. *[Social Cognitive and Affective Neuroscience Electronic Resource], 8*(8), 918–927. https://doi.org/10.1093/scan/nss094

Gervais, M., & Wilson, D. S. (2005). The evolution and functions of laughter and humor: A synthetic approach. *The Quarterly review of biology, 80*(4), 395–430. https://doi.org/10.1086/498281

Gibby, R. G., Jr., Gibby, R. G.,S., & Townsend, J. C. (1970). Short-term visual restriction in visual and auditory discrimination. *Perceptual and Motor Skills, 30*(1), 15–21. https://doi.org/10.2466/pms.1970.30.1.15

Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). The voices of wrath: Brain responses to angry prosody in meaningless speech. *Nature neuroscience, 8*(2), 145–146. https://doi.org/10.1038/nN100392

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology, 55*(4), 468–484. https://doi.org/10.1016/0013-4694(83)90135-9

Hajcak, G., & Nieuwenhuis, S. (2006). Reappraisal modulates the electrocortical response to negative pictures. *Cognitive, Affective & Behavioral Neuroscience, 6*(4), 291–297. https://doi.org/10.3758/cabn.6.4.291

Hajcak, G., & Olvet, D. M. (2008). The persistence of attention to emotion: Brain potentials during and after picture presentation. *Emotion, 8*(2), 250–255. https://doi.org/10.1037/1528-3542.8.2.250

Hendriks, M., Croon, M., & Vingerhoets, A. (2008). Social reactions to adult crying: The help-soliciting function of tears. *The Journal of social psychology, 148*(1), 22–41. https://doi.org/10.3200/SOCP.148.1.22-42

Hink, R. F., Hillyard, S. A., & Benson, P. J. (1978). Event-related brain potentials and selective attention to acoustic and phonetic cues. *Biological Psychology, 6*(1), 1–16. https://doi.org/10.1016/0301-0511(78)90002-9

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007

Jessen, S., & Kotz, S. A. (2011). The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. *Neuroimage, 58*(2), 665–674. https://doi.org/10.1016/j.neuroimage.2011.06.035

Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2010). Attention and the multiple stages of multisensory integration: A review of audiovisual studies. *Acta Psychologica, 134*(3), 372–384. https://doi.org/10.1016/j.actpsy.2010.03.010

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13). https://doi.org/10.18637/jss.v082.i13

Landry, S. P., Shiller, D. M., & Champoux, F. (2013). Short-term visual deprivation improves the perception of harmonicity. *Journal of Experimental Psychology. Human Perception and Performance, 39*(6), 1503–1507. https://doi.org/10.1037/a0034015

Lavan, N., Lima, C. F., Harvey, H., Scott, S. K., & McGettigan, C. (2015). I thought that I heard you laughing: Contextual facial expressions modulate the perception of authentic laughter and crying. *Cognition & Emotion, 29*(5), 935–944. https://doi.org/10.1080/02699931.2014.957656

Lavan, N., Rankin, G., Lorking, N., Scott, S., & McGettigan, C. (2017). Neural correlates of the affective properties of spontaneous and volitional laughter types. *Neuropsychologia, 95*, 30–39. https://doi.org/10.1016/j.neuropsychologia.2016.12.012

Lavan, N., Scott, S. K., & McGettigan, C. (2016). Laugh like you mean it: Authenticity modulates acoustic, physiological and perceptual properties of laughter. *Journal of Nonverbal Behavior, 40*(2), 133–149. https://doi.org/10.1007/s10919-015-0222-8

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). Emmeans: Estimated marginal means, aka least-squares means. *Parasites & Vectors, 1*(1), 3.

Lewald, J. (2007). More accurate sound localization induced by short-term light deprivation. *Neuropsychologia, 45*(6), 1215–1222. https://doi.org/10.1016/j.neuropsychologia.2006.10.006

Lima, C. F., Anikin, A., Monteiro, A. C., Scott, S. K., & Castro, S. L. (2019). Automaticity in the recognition of nonverbal emotional vocalizations. *Emotion, 19*(2), 219–233. https://doi.org/10.1037/emo0000429

Lima, C. F., Brancatisano, O., Fancourt, A., Müllensiefen, D., Scott, S. K., Warren, J. D., & Stewart, L. (2016). Impaired socio-emotional processing in a developmental music disorder. *Scientific Reports, 6*, 34911. https://doi.org/10.1038/srep34911

Liu, T., Pinheiro, A. P., Deng, G., Nestor, P. G., Mccarley, R. W., & Niznikiewicz, M. A. (2012). Electrophysiological insights into processing nonverbal emotional vocalizations. *Neuroreport, 23*(2), 108–112. https://doi.org/10.1097/WNR.0b013e32834ea757

Lu, X., Ho, H. T., Liu, F., Wu, D., & Thompson, W. F. (2015). Intonation processing deficits of emotional words among Mandarin Chinese speakers with congenital amusia: An ERP study. *Frontiers in Psychology, 6*, Article 385. https://doi.org/10.3389/fpsyg.2015.00385

Masuda, F., Sumi, Y., Takahashi, M., Kadotani, H., Yamada, N., & Matsuo, M. (2018). Association of different neural processes during different emotional perceptions of white noise and pure tone auditory stimuli. *Neuroscience Letters, 665*, 99–103. https://doi.org/10.1016/j.neulet.2017.11.046

McGettigan, C., Walsh, E., Jessop, R., Agnew, Z. K., Sauter, D. A., Warren, J. E., & Scott, S. K. (2015). Individual differences in laughter perception reveal roles for mentalizing and sensorimotor systems in the evaluation of emotional authenticity. *Cerebral Cortex, 25*(1), 246–257. https://doi.org/10.1093/cercor/bht227

McKeown, G., Sneddon, I., & Curran, W. (2015). Gender differences in the perceptions of genuine and simulated laughter and amused facial expressions. *Emotion Review, 7*(1), 30–38. https://doi.org/10.1177/1754073914544475

Merabet, L. B., Hamilton, R., Schlaug, G., Swisher, J. D., Kiriakopoulos, E. T., Pitskel, N. B., Kauffman, T., & Pascual-Leone, A. (2008). Rapid and reversible recruitment of early visual cortex for touch. *Plos One, 3*(8), Article e3046. https://doi.org/10.1371/journal.pone.0003046

Meyer, M., Elmer, S., Baumann, S., & Jancke, L. (2007). Short-term plasticity in the auditory system: Differential neural responses to perception and imagery of speech and music. *Restorative Neurology and Neuroscience, 25*(3–4), 411–431.

Motley, M. T., & Camden, C. T. (1988). Facial expression of emotion: A comparison of posed expressions versus spontaneous expressions in an interpersonal communication setting. *Western Journal of Speech Communication, 52*(1), 1–22. https://doi.org/10.1080/10570318809389622

Näätänen, R., & Picton, T. W. (1987). The N100 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology, 24*(4), 375–425. https://doi.org/10.1111/j.1469-8986.1987.tb00311

Neves, L., Cordeiro, C., Scott, S. K., Castro, S. L., & Lima, C. F. (2018). High emotional contagion and empathy are associated with enhanced detection of emotional authenticity in laughter. *The Quarterly Journal of Experimental Psychology: QJEP, 71*(11), 2355–2363. https://doi.org/10.1177/1747021817741800

Oldfield, R. C. (1971). The assessment and analysis of handedness: The edinburgh inventory. *Neuropsychologia, 9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

O'Nions, E., Lima, C., Scott, S., Roberts, R., McCrory, E., & Viding, E. (2017). Reduced laughter contagion in boys at risk for psychopathy. *Current Biology, 27*(19), 3049–3055.e4. https://doi.org/10.1016/j.cub.2017.08.062

Parsons, C. E., Young, K. S., Craske, M. G., Stein, A. L., & Kringelbach, M. L. (2014). Introducing the oxford vocal (OxVoc) sounds database: A validated set of non-acted affective sounds from human infants, adults, and domestic animals. *Frontiers in psychology, 5*, 562. https://doi.org/10.3389/fpsyg.2014.00562

Paulmann, S., Bleichner, M., & Kotz, S. A. (2013). Valence, arousal, and task effects in emotional prosody processing. *Frontiers in Psychology, 4*, 345. https://doi.org/10.3389/fpsyg.2013.00345

Paulmann, S., & Kotz, S. A. (2008). Early emotional prosody perception based on different speaker voices. *Neuroreport, 19*(2), 209–213. https://doi.org/10.1097/WNR.0b013e3282f454db

Pell, M. D., Rothermich, K., Liu, P., Paulmann, S., Sethi, S., & Rigoulot, S. (2015). Preferential decoding of emotion from human non-linguistic vocalizations versus speech prosody. *Biological Psychology, 111*, 14–25. https://doi.org/10.1016/j.biopsycho.2015.08.008

Pinheiro, A. P., Anikin, A., Conde, T., Sarzedas, J., Chen, S., Scott, S. K., & Lima, C. (2021). Emotional authenticity modulates affective and social trait inferences from voices.

Philosophical Transactions of the Royal Society of London: Biological Sciences, 376(1840), 20200402. https://doi.org/10.1098/rstb.2020.0402
Pinheiro, A. P., Barros, C., & Pedrosa, J. (2016b). Salience in a social landscape: Electrophysiological effects of task-irrelevant and infrequent vocal change. *Social Cognitive and Affective Neuroscience, 11*(1), 127–139. https://doi.org/10.1093/scan/nsv103
Pinheiro, A. P., Rezaii, N., Nestor, P. G., Rauber, A., Spencer, K. M., & Niznikiewicz, M. (2016a). Did you or I say pretty, rude or brief? An ERP study of the effects of speaker's identity on emotional word processing. *Brain and language, 153–154*, 38–49. https://doi.org/10.1016/j.bandl.2015.12.003
Pinheiro, A. P., Rezaii, N., Rauber, A., Liu, T., Nestor, P. G., McCarley, R. W., Gonçalves, O. F., & Niznikiewicz, M. a (2014). Abnormalities in the processing of emotional prosody from single words in schizophrenia. *Schizophrenia Research, 152*(1), 235–241. https://doi.org/10.1016/j.schres.2013.10.042
Pinheiro, A. P., Rezaii, N., Rauber, A., Nestor, P. G., Spencer, K. M., & Niznikiewicz, M. (2017). Emotional self-other voice processing in schizophrenia and its relationship with hallucinations: ERP evidence. *Psychophysiology, 54*(9), 1252–1265. https://doi.org/10.1111/psyp.12880
Provine, R. R. (2001). *Laughter: A scientific investigation*. New York: Penguin Books.
Provine, R. R., Krosnowski, K. A., & Brocato, N. W. (2009). Tearing: Breakthrough in human emotional signaling. *Evolutionary Psychology, 7*(1), 52–56. https://doi.org/10.1177/147470490900700107
Rigoulot, S., Pell, M. D., & Armony, J. L. (2015). Time course of the influence of musical expertise on the processing of vocal and musical sounds. *Neuroscience, 290*, 175–184.
RStudio Team. (2021). *RStudio: Integrated Development Environment for R. RStudio*. Boston, MA: PBC. http://www.rstudio.com/.
Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin, 115*(1), 102–141. https://doi.org/10.1037/0033-2909.115.1.102
Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). Emotion and attention interactions in social cognition: Brain regions involved in processing anger prosody. *Neuroimage, 28*(4), 848–858. https://doi.org/10.1016/j.neuroimage.2005.06.023
Sauter, D. A., & Eimer, M. (2010). Rapid detection of emotion from human vocalizations. *Journal of Cognitive Neuroscience, 22*(3), 474–481. https://doi.org/10.1162/jocn.2009.21215
Sauter, D. A., & Fischer, A. H. (2018). Can perceivers recognise emotions from spontaneous expressions? *Cognition & Emotion, 32*(3), 504–515. https://doi.org/10.1080/02699931.2017.1320978
Scherer, K. R., & Bänzinger, T. (2010). On the use of actor portrayals in research on emotional expression. In K. R Scherer, T. Bänziger, & E. Roesch (Eds.), *Blueprint for affective computing: A sourcebook and manual (affective science)* (pp. 166–178). Oxford, England: Oxford University Press.
Scherer, K. R., Clark-Polner, E., & Mortillaro, M. (2011). In the eye of the beholder? Universality and cultural specificity in the expression and perception of emotion. *International Journal of Psychology, 46*(6), 401–435. https://doi.org/10.1080/00207594.2011.626049
Schindler, S., & Kissler, J. (2016). Selective visual attention to emotional words: Early parallel frontal and visual activations followed by interactive effects in visual cortex. *Human brain mapping, 37*(10), 3575–3587. https://doi.org/10.1002/hbm.23261
Schirmer, A., Chen, C.-B., Ching, A., Tan, L., & Hong, R. Y. (2013). Vocal emotions influence verbal memory: Neural correlates and interindividual differences. *Cognitive, Affective & Behavioral Neuroscience, 13*(1), 80–93. https://doi.org/10.3758/s13415-012-0132-8
Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Science, 10*(1), 24–30. https://doi.org/10.1016/j.tics.2005.11.009
Schupp, H. T., Flaisch, T., Stockburger, J., & Junghöfer, M. (2006). Emotion and attention: Event-related brain potential studies. *Progress in Brain Research, 156*, 31–51. https://doi.org/10.1016/S0079-6123(06)56002-9
Scott, S., Lavan, N., Chen, S., & McGettigan, C. (2014). The social life of laughter. *Trends in Cognitive Sciences, 18*(12), 618–620. https://doi.org/10.1016/j.tics.2014.09.002
Seither-Preisler, A., Patterson, R., Krumbholz, K., Seither, S., & Lutkenhoner, B. (2006). Evidence of pitch processing in the N10000m component of the auditory evoked field. *Hearing Research, 213*(1–2), 88–98. https://doi.org/10.1016/j.heares.2006.01.003
Sidtis, D., & Kreiman, J. (2012). In the beginning was the familiar voice: Personally familiar voices in the evolutionary and contemporary biology of communication. *Integrative Psychological and Behavioral Science, 46*(2), 146–159. https://doi.org/10.1007/s12124-011-9177-4
Spreckelmeyer, K. N., Kutas, M., Urbach, T. P., Altenmüller, E., & Münte, T. F. (2006). Combined perception of emotion in pictures and musical sounds. *Brain Research, 1070*(1), 160–170. https://doi.org/10.1016/j.brainres.2005.11.075
Szameitat, D. P., Alter, K., Szameitat, A. J., Darwin, C. J., Wildgruber, D., Dietrich, S., & Sterr, A. (2009). Differentiation of emotions in laughter at the behavioral level. *Emotion, 9*(3), 397–405. https://doi.org/10.1037/a0015692
Szameitat, D. P., Alter, K., Szameitat, A. J., Wildgruber, D., Sterr, A., & Darwin, C. J. (2009). Acoustic profiles of distinct emotional expressions in laughter. *The Journal of the Acoustical Society of America, 126*(1), 354–366. https://doi.org/10.1121/1.3139899
Tabry, V., Zatorre, R. J., & Voss, P. (2013). The influence of vision on sound localization abilities in both the horizontal and vertical planes. *Frontiers in Psychology, 4*, 932. https://doi.org/10.3389/fpsyg.2013.00932
Vettin, J., & Todt, D. (2004). Laughter in conversation: Features of occurrence and acoustic structure. *Journal of Nonverbal Behavior, 28*(2), 93–115. https://doi.org/10.1023/B:JONB.0000023654.73558.72
Vingerhoets, A. J. J. M., & Bylsma, L. M. (2016). The riddle of human emotional crying: A challenge for emotion researchers. *Emotion Review, 8*(3), 207–217. https://doi.org/10.1177/1754073915586226
Vredeveldt, A., Hitch, G. J., & Baddeley, A. D. (2011). Eye closure helps memory by reducing cognitive load and enhancing visualisation. *Memory & cognition, 39*(7), 1253–1263. https://doi.org/10.3758/s13421-011-0098-8
Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., & Bloom, F. E. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences of the United States of America, 90*(18), 8722–8726. https://doi.org/10.1073/pnas.90.18.8722
Wood, A., Martin, J., & Niedenthal, P. (2017). Towards a social functional account of laughter: Acoustic features convey

reward, affiliation, and dominance. *Plos One, 12*(8), Article e0183811. https://doi.org/10.1371/journal.pone.0183811

Wöstmann, M., Schmitt, L. M., & Obleser, J. (2020). Does closing the eyes enhance auditory attention? Eye closure increases attentional alpha-power modulation but not listening performance. *Journal of cognitive neuroscience, 32*(2), 212–225. https://doi.org/10.1162/jocn_a_01403

Zeifman, D. M. (2001). An ethological analysis of human infant crying: Answering tinbergen's four questions. *Developmental Psychobiology, 39*(4), 265–285. https://doi.org/10.1002/dev.1005