Universidade de Lisboa

Faculdade de Medicina de Lisboa

Authenticity Recognition in Laughter and Crying: An ERP Study

Mónica Gonçalves Barbeito Costa

Orientador: Prof. Doutora Diana Maria Pinto Prata

Co-orientador: Prof. Doutora Ana Patrícia Teixeira Pinheiro

Dissertação especialmente elaborada para obtenção do grau de
Mestre em Neurociências
2018

2018

Universidade de Lisboa

Faculdade de Medicina de Lisboa

Authenticity Recognition in Laughter and Crying: An ERP Study

Mónica Gonçalves Barbeito Costa

Orientador: Prof. Doutora Diana Maria Pinto Prata

Co-orientador: Prof. Doutora Ana Patrícia Teixeira Pinheiro

Dissertação especialmente elaborada para obtenção do grau de

Mestre em Neurociências

2018

A impressão desta dissertação foi aprovada pelo Conselho Científico da Faculdade de Medicina de Lisboa em reunião de 18 de setembro de 2018.

INDEX

ABSTRACT **6**

RESUMO **8**

INTRODUCTION **11**

1. Non-linguistic affective vocalizations 12
2. ERP Technique 14
3. Multi-stage approach to auditory affective processing 15
4. Authenticity 19

AIMS OF THE STUDY **26**

METHODS **27**

1. Participants 27
2. Auditory Stimuli 29
3. Task 32
4. EEG Data Acquisition and Processing 33
5. Procedure 35
6. Statistical Analyses 35
    - 6.1 Behavioural Statistical Analyses 35
    - 6.2 EEG Statistical Analyses 36
    - 6.3 Correlations 37

RESULTS **38**

1. Behavioural Results 38
    - 1.1 Mean Reaction Time 38
    - 1.2 Authenticity Ratings 39
    - 1.3 Arousal Ratings 40
    - 1.4 Emotional Contagion Ratings 41

2. EEG Results 44

2.1 Spontaneous emotional vs. neutral vocalizations 46

2.2 Voluntary emotional vs. neutral vocalizations 52

2.3 Spontaneous vs. voluntary emotional vocalizations 57

3. Correlations 60

DISCUSSION 64

1. The effects of emotion in voice processing 64

2. The effects of authenticity in voice processing 66

3. Relationship between ERPs and behavioural data 69

4. Limitations 71

CONCLUSIONS 72

ACKNOWLEDGMENTS 73

REFERENCES 74

ABSTRACT

Our ability to detect authenticity in the human affective voice, whether an emotion was evoked spontaneously (reactive, genuine) or voluntarily (deliberate, controlled), is crucial in our everyday social interactions as emotions may carry different meanings and elicit different social responses. Taking laughter as an example, while a spontaneous laughter is stimulus-driven and signals positive affect, voluntary laughter deliberately signals polite agreement or affiliation without necessarily being associated with an emotional experience. Recent functional magnetic resonance (fMRI) studies have shown brain differences between these voluntary and spontaneous laughter vocalizations. While both spontaneous and voluntary laughs engage the auditory cortex, voluntary laughter requires additional involvement of brain areas typically involved in mentalizing, possibly involved in the decoding of the intentional state behind these vocal expressions. However, how authenticity affects the temporal course of voice processing is still unclear. Previous imaging studies have shed light on the areas putatively involved in the processing of authenticity in vocal emotions. Nevertheless, fMRI lacks temporal resolution and is unable to provide information about the exact time-window on which differences in the processing between spontaneous and voluntary vocalizations in the brain may occur.

In the current study we used the event-related brain potentials (ERPs) methodology to shed light on how authenticity modulates the temporal course of vocal information processing in the brain. In particular, we investigated differences between spontaneous and voluntary non-linguistic affective vocalizations (crying and laughter) in both amplitude and latency of ERP components associated with early (N100, P200) and late stages (late positivity potential – LPP) of voice processing. We also aimed to replicate previous findings suggesting amplitude and latency differences as a function of emotionality in these three ERP components. In addition, we explored the extent to which sex differences may exist in both authenticity and emotionality modulation of these potentials. Twenty-three right-handed healthy participants (13 female) listened to spontaneous and voluntary non-linguistic affective vocalizations (happy, sad and neutral) while they rated the authenticity conveyed by the speaker, as the electroencephalogram (EEG) was recorded.

No differences in terms of amplitude or latency were found between spontaneous and voluntary vocalizations in the N100, P200 and LPP components. Emotionality effects were found at an early processing stage (N100) with happy and sad vocalizations eliciting more negative amplitude than neutral vocalizations. Happy vocalizations elicited an enhanced P200 when compared with neutral vocalizations. At later processing stages (500 – 700 ms), happy and sad vocalizations elicited a stronger late positivity (LPP) than neutral vocalizations. No differences between emotional and neutral vocalizations were detected in the latency of these components. Lastly, no sex differences were found in the amplitude or latency of N100, P200 and LPP for emotionality or authenticity effects.

Although exploratory with a small sample size and deserving further replication, all together, our results possibly suggest authenticity as unlikely to be decoded during the first 700 ms after vocalization onset. The emotional salience of the voice, on the other hand, seems to be extracted as early as 100 ms after onset. While emotional content seems to be rapidly decoded from vocal cues, authenticity may involve further elaborated processing occurring at very late stages of processing.

Key-words: Voice processing; Authenticity; Emotion; Non-linguistic vocalizations; Event-related potentials (ERP).

RESUMO

A voz humana comunica não só informação verbal, como também informação acerca da identidade e estado emocional do locutor (e.g., medo, raiva, nojo, tristeza, felicidade, surpresa) através de modulações nas propriedades acústicas (frequência, intensidade, ritmo). A autenticidade de uma expressão emocional é uma propriedade também extraída quando escutamos uma voz. Através do perfil acústico da vocalização e do seu contexto somos capazes de detetar se uma emoção foi evocada espontaneamente (ato reativo, genuíno) ou voluntariamente (ato deliberado, controlado). A capacidade de detetar a autenticidade de expressões emocionais na voz humana é crucial nas nossas interações sociais no dia-a-dia, já que estes dois tipos de expressões transmitem diferentes significados e provocam diferentes respostas sociais. Usando como exemplo a gargalhada, enquanto a gargalhada evocada espontaneamente é o resultado de um evento externo assinalando afeto positivo, a gargalhada evocada voluntariamente é deliberada indicando cortesia ou afiliação, sem necessariamente estar associada a uma experiência emocional. Estudos recentes com ressonância magnética funcional mostraram diferenças no cérebro entre gargalhadas evocadas espontaneamente e voluntariamente. Enquanto tanto a gargalhada espontânea como a gargalhada voluntária ativam áreas do córtex auditivo, a gargalhada voluntária ativa adicionalmente áreas típicas da mentalização, possivelmente envolvendo a interpretação da intenção da expressão vocal. Contudo, permanece por explorar como a autenticidade afeta o curso temporal do processamento de voz. Um modelo de múltiplos estágios de processamento da informação vocal foi proposto por Schirmer & Kotz (2006) com base em estudos com potenciais evocados e de ressonância magnética funcional. Este modelo sugere um processamento da informação vocal em três diferentes estágios: análise das propriedades acústicas (indexado pelo componente N100, ocorrendo cerca de 100 ms após o inicio da vocalização), extração da saliência emocional (indexado pelo componente P200, ocorrendo cerca de 200 ms após o inicio da vocalização) e por último, avaliação cognitiva da expressão vocal (indexado pelo *late positivity potential* – LPP, ocorre entre 500 e 700 ms após o inicio da vocalização). Diferenças no processamento da informação vocal entre estímulos emocionais e neutros, têm sido amplamente reportadas nestes três estágios de processamento. No entanto, os estudos previamente mencionados utilizaram estímulos que foram desenvolvidos instruindo atores a imitar emoções (expressões de emoção voluntária) e não expressões de emoção espontâneas. Permanece por esclarecer até que

ponto estes resultados reportados poderão ser explicados por diferenças na autenticidade da emoção. Estudos de neuroimagem prévios mostraram com elevada resolução espacial quais as áreas no cérebro putativamente envolvidas no processamento de autenticidade no processamento vocal afetivo. Não obstante, a técnica de imagem ressonância magnética funcional carece de resolução temporal, não permitindo extrair informação relativamente à janela temporal exata durante a qual estas diferenças no processamento cognitivo entre vocalizações espontâneas e voluntárias poderão ocorrer no cérebro.

No presente estudo utilizámos uma abordagem com potenciais evocados para esclarecer como a autenticidade modula o curso temporal do processamento de informação vocal afetiva no cérebro. Em particular, tivemos por objetivo investigar as diferenças entre vocalizações não-verbais espontâneas e voluntárias (gargalhada e choro) em termos da amplitude e latência dos componentes eletrofisiológicos associados a estágios iniciais (N100, P200) e a estágios mais tardios (LPP). Procurámos também replicar resultados prévios que sugerem diferenças em termos da amplitude e latência em função da emocionalidade da vocalização (emocional vs. neutro) no componente N100, P200 e LPP. Adicionalmente, como hipótese exploratória investigámos até que ponto diferenças de sexo podem existir tanto na autenticidade como na emocionalidade na modulação destes potenciais evocados. Vinte e três participantes destros saudáveis (13 mulheres) ouviram vocalizações não-verbais espontâneas e voluntárias (expressando alegria, tristeza ou tom neutro), enquanto classificavam a autenticidade expressada pelo locutor, enquanto era registado um eletroencefalograma (EEG) em simultâneo.

Em termos dos efeitos da autenticidade no processamento de informação vocal, não foram encontradas diferenças relativamente à amplitude e latência entre vocalizações espontâneas e voluntárias nos componentes N100, P200 e LPP. Efeitos de emocionalidade foram encontrados em estágios iniciais do processamento vocal (N100), com vocalizações de alegria e tristeza mostrando deflexões menos negativas quando comparadas com vocalizações neutras. Vocalizações de alegria evocaram um P200 de maior magnitude do que vocalizações neutras, não existindo diferenças significativas entre vocalizações de alegria e tristeza ou entre vocalizações de tristeza e neutras. Em estágios mais tardios do processamento (500 – 700 ms), vocalizações de alegria e tristeza evocaram uma positividade tardia (LPP) mais pronunciada do que vocalizações neutras. Os efeitos de emocionalidade reportados nos potenciais N100, P200 e LPP verificaram-se de igual modo em vocalizações espontâneas e vocalizações voluntárias. Não foram encontradas diferenças de latência entre vocalizações emocionais e neutras em nenhum

estágio de processamento vocal (N100, P200 e LPP). Por último, relativamente a diferenças de sexo no processamento da autenticidade e emocionalidade na informação vocal, não foram encontradas diferenças nestes três componentes entre homens e mulheres.

Ainda que exploratório e com necessidade de futuras replicações, os nossos resultados sugerem que a autenticidade possivelmente não será descodificada durante os primeiros 700 ms após o início da vocalização. A emocionalidade por outro lado, parece ser extraída cedo no processamento vocal, nos primeiros 100 ms (N100) após o início da vocalização irrespectivamente da valência (positiva e negativa), sendo que tanto vocalizações de alegria como de tristeza evocaram uma menor amplitude no componente N100 do que vocalizações neutras. A emocionalidade da vocalização parece deste modo ser detetada em estágios iniciais (N100), irrespectivamente da valência do estímulo. Porém no estágio seguinte, verificou-se que apenas vocalizações de alegria evocaram um P200 de maior amplitude, relativamente a vocalizações neutras. Este resultado poderá dever-se à elevada sensibilidade do componente P200 à ativação fisiológica inerente ao estímulo vocal, isto é, estímulos caracterizados por uma maior ativação fisiológica (e.g., gargalhada) são percebidos como emocionalmente mais salientes. Em estágios mais tardios do processamento verificou-se uma maior positividade do componente LPP em vocalizações emocionais (alegria e tristeza) comparativamente a vocalizações neutras. As vocalizações emocionais, independentemente da sua valência (positiva ou negativa), parecem assim promover uma elaboração cognitiva mais profunda.

Em suma, de acordo com os resultados obtidos neste estudo preliminar enquanto o conteúdo emocional parece ser rapidamente processado em pistas vocais, a autenticidade possivelmente envolve um processamento mais elaborativo que ocorre em estádios mais tardios do processamento.

Palavras-chave: Processamento de voz; Autenticidade; Emoção; Vocalizações não-verbais; Potenciais evocados.

INTRODUCTION

The human voice communicates not only verbal information, but also information about a speaker’s identity and emotional state (e.g., fear, anger, disgust, sadness, happiness, surprise) through the modulation of acoustic features such as pitch, loudness and tempo (Belin, Zatorre, Lafaille, Ahad & Pike, 2000; Karpf, 2006). Voice-selective regions can be found in the human auditory cortex, which are located bilaterally in the superior temporal sulcus (STS), when listening to human vocal sounds (speech or non-speech) compared to other sound categories (Belin et al, 2000). Three functionally distinct neural pathways of voice processing have been described: speech, identity and affect (Belin, Bestelmeyer, Latinus & Watson, 2011). Different brain areas are recruited by these voice processing pathways: speech – temporal areas (anterior and posterior STS) and inferior prefrontal regions (left hemisphere); emotion – temporo-medial regions (anterior insula and amygdala) and inferior prefrontal regions (mainly right hemisphere); and lastly, vocal identity – regions of the right temporal sulcus (Belin, Fecteau & Bédard, 2004, Belin et al, 2011).

In particular, emotional cues may be conveyed through emotional speech prosody as well as through non-linguistic affective vocalizations (Hawk, van Kleef, Fischer, van der Schalk, 2009). Emotional prosody refers to suprasegmental modulations of speech to convey emotion (speech melody). On the other hand, non-linguistic affective vocalizations represent more primitive forms of vocal expressions unconstrained by linguistic structure (e.g., laughter, sobs, screams) (Pell, Rothermich, Liu, Paulmann, Sethi & Rigoulot, 2015; Sauter, Eisner, Calder & Scott, 2010a).

1. Non-linguistic affective vocalizations

Emotion is expressed via multiple sensory channels, including the visual and auditory channels. The visual and auditory channels have been shown to convey emotional meaning on their own, transmitting enough information to allow emotion recognition without a need for complementary action from other channels (Hawk et al, 2009). A vast research has been conducted in the visual modality, showing for example that our ability to recognize ‘basic’ emotions (fear, anger, disgust, sadness, happiness and surprise) from facial expressions alone is generally accurate (Ekman, 1982; Ekman, Friesen & Hager, 2002; Calvo & Nummenmaa, 2016). Compared to research in visual emotional processing, research probing emotional processing in the auditory modality has been scarce (Hawk et al, 2009). Communication through vocal cues portrays several advantages over visual cues, as emotions may be communicated over long distance, directing our attention to relevant cues in our environment (Hawk et al, 2009).

Non-linguistic affective vocalizations (i.e., more primitive forms of vocal expressions unconstrained by linguistic structure) can be categorized according to different criteria: category (emotion valence – happiness, sadness, anger, fear, surprise, disgust) (Schröder, 2003), valence and arousal (Sauter et al, 2010a), authenticity (i.e. whether an expression is evoked spontaneously or voluntarily) (McGettigan et al, 2015) or affiliative value (emotional contagion, i.e., how much we resonate with others' emotions) (Neves et al, 2018).

Scherer (1994) was the first author to introduce the concept of ‘affect bursts’, describing them as ‘‘*very brief, discrete, nonverbal expressions of affect in both face and voice as triggered by clearly identifiable events’’* (p. 170). The concept was later updated by Schröder (2003) who described ‘affect bursts’ as emotional non-speech expressions that range in a continuum, ranging from clear non-speech vocalizations (raw affect bursts) (e.g., laughter) to interjections with a phonemic structure (e.g., ‘‘Wow!’’) (Scherer, 2003; Schröder, 2003). ‘Affect bursts’ may include expressions of happiness (laughter), anger (growling), fear (screams), sadness (sobbing), disgust and surprise (Hawk et al, 2009; Sauter et al, 2010a). Whereas speech production relies on fast, precise, coordinated actions of supra-laryngeal articulators, movement of the larynx and sub-glottal pressure (Murray & Arnot, 1993), non-linguistic vocalizations do not require these precise supra-laryngeal movements or articulations, being produced by roughly positioning pharyngeal/oral/labial constrictions (Sauter et al, 2010a). In this line, laughter has been defined as modified breathing (Provine, 2004), while crying comprises erratic inhalation accompanied sometimes by tears in adults (Vingerhoets & Bylsma, 2015).

In this work, we will focus on non-linguistic affective vocalizations, as they convey emotional meaning without the confound of concurrent semantic content or the need of a situational context (Schröder, 2003). Thus, non-linguistic vocalizations are advantageous when studying the impact of vocal cues characteristics on voice processing, such as authenticity (i.e., whether an emotion is produced spontaneously or voluntarily), as they lack potential semantic confounds which are encountered with speech prosody avoiding possible interactions between affective and semantic content (Schröder, 2003; Pell et al, 2015).

Behavioural evidence indicates that accuracy in the recognition of vocal emotions varies according to the type of vocal stimulus presented (emotional prosody – words/sentences, pseudo-words/sentences; non-linguistic affect vocalizations) (Castro & Lima, 2010; Hawk et al, 2009; Vasconcelos, Dias, Soares & Pinheiro, 2017). Concerning emotional prosody, sentences and pseudo-sentences (speech with intelligible semantic content) have similar accuracy rates (75% for sentences and 71% for pseudo-sentences) (Castro & Lima, 2010). The accuracy rates for non-linguistic affective vocalizations are higher, with a mean recognition rate of 81%, without a need of situational context (Schröder, 2003; Schreder, 2003). Hawk and colleagues (2003) further confirmed this result by comparing accuracy rates between facial, speech and non-linguistic affective expressions. Higher scores were obtained both for facial and non-linguistic affective expressions than for speech. Furthermore, non-linguistic affect vocalizations recognition is efficient even when the cognitive system is loaded with another task (Lima, Anikin, Monteiro, Scott & Castro, accepted). Similar to facial expressions, non-linguistic affective vocalizations are recognized across different cultures (Sauter, Eisner, Ekman & Scott, 2010b) due to the nonverbal evolutionary nature of the sounds (Lima, Castro & Scott, 2013). Nonetheless, an in-group effect is found with accuracy rates being higher for vocalizations produced by members of one’s own culture than for other cultures (Sauter et al, 2010b). In a recent study, Vasconcelos et al (2017) have found happy non-linguistic vocalizations (laughter) are easier to recognize (higher accuracy) when compared with other vocal emotions categories (anger, disgust, fear, sadness, pain, surprise and pleasure). Higher accuracy in the recognition happy vocalizations may be explained by the critical role laughter plays in our daily social interactions and communication, signalling affiliation or appeasement (Scott et al, 2015). Neuroimaging evidence shows voice-selective cortical regions in the human brain which respond exclusively to the human voice when compared with other sounds, such as the bilateral superior temporal sulcus (STS), not only for speech sounds but also for non-speech sounds (non-linguistic affective vocalizations) (Belin, Zatorre & Ahad, 2002; Charest et al, 2009). Moreover, evidence with near infra-red spectroscopy shows activation of voice-selective regions upon hearing emotional prosody in 7-month old infants who are still at an early stage of speech development but can already discriminate voices (Grossman, Oberecker, Koch, & Friederici, 2010). This finding further suggests these voice-selective areas are not exclusively related to speech, but also to non-verbal human displays of communication (Belin et al, 2011). Blasi and colleagues (2011) found evidence for an early specialization on processing non-verbal vocalizations and emotion in developing infants, reflected in greater activation in the anterior temporal cortex for non-linguistic vocalizations of happiness and sadness, than for non-voice environmental sounds.

Neurophysiological evidence suggests that the decoding of emotional meaning from human vocalizations in real time is species-specific and of great importance for social interactions (Pell et al, 2015; Schirmer & Kotz, 2006). Non-linguistic vocalizations have been reported to be processed faster than emotional prosody, possibly due to its primitive origins (e.g., detecting threats in the environment) (Pell et al, 2015). The event-related potentials (ERP) technique has a high temporal resolution in the order of milliseconds (ms), as opposed to fMRI which presents a temporal resolution in the order of seconds (s). Hence, it allows investigating the specific time-window in which a given cognitive process takes place and how it changes throughout time (Luck, 2005).

2. ERP Technique

The ERP (event-related potential) technique offers a non-invasive method for studying information processing in the human brain in real time due to its high temporal resolution, aiding in determining which stage or stages of processing are influenced by a specific experimental manipulation. The advantage of ERP technique compared to behavioural measures is that it allows a continuous measure of processing between a stimulus and a response (Luck, 2005). ERPs are a series of voltage deflections in the electrical brain activity evoked by a specific event (sensory, cognitive or motor), these deflections can be recorded through an electroencephalogram (EEG) (Crowley & Colrain, 2004; Luck, 2005). An ERP waveform consists of neurophysiological data averaged through multiple trials and time-locked to the presentation of a specific cognitive or sensory event, giving rise to positive or negative peaks that reflect a discrete stage of the neural processing of this event (Crowley & Colrain, 2004; Luck, 2005; Picton, Lins, & Scherg, 1995). These components can be characterized by their amplitude (magnitude measured from the maximal peak of a given component), latency (time interval between stimulus onset and maximal peak of a given component), polarity (positive or negative), scalp distribution (anatomical site of generation) or function reflected by them (Luck, 2005). Depending on whether the stimulus is visual, auditory or motor, a series of early components reflective of initial information processing will be elicited, followed by later components associated with integrative and higher-order cognitive processing (Crowley & Colrain, 2004). ERPs allow to study the temporal dynamics of neural responses elicited by dynamic stimuli (e.g., vocalizations), being fitting to probe the multi-stage model of vocal information processing. As for the neural underpinnings of emotional vocal cues processing, a multistage model has been proposed by Schirmer and Kotz (2006), integrating fMRI and electroencephalography (EEG) evidence, establishing three stages of information processing.

3. Multi-stage approach to auditory affective processing

A multi-stage model of vocal emotional processing has been proposed by Schirmer & Kotz (2006) establishing three different stages: 1) Sensory processing – the analysis of the acoustic properties of a vocal stimulus (extraction of pitch, loudness and tempo) that takes place bilaterally in the STS (superior temporal sulcus); 2) Integration – general detection of salience from acoustic cues that takes place in the bilateral superior temporal gyrus (STG) and STS; 3) Cognition – cognitive evaluation of the emotional significance of the voice that occurs in the inferior frontal gyrus (IFG) and orbitofrontal cortex (Schirmer & Kotz, 2006). Importantly, these stages have been also indexed by distinct event-related potential (ERP) signatures reflected: N100, P200 and Late Positivity Potential (LPP). The multi-stage model of vocal emotional processing was later updated by Frühholz, Trost and Kotz (2016) based on fMRI evidence, advocating for a unifying neural network that underlies the processing of all types of affective sounds, which included areas that were not previously considered integrative for emotion decoding (cerebellum, basal ganglia and insula). Moreover, a bilateral engagement of the inferior frontal cortex when listening to emotional cues in the human voice has been described when higher-order cognitive processes are taking place (evaluating an emotional vocalization) (Frühholz & Grandjean, 2013). According to the multi-stage model of vocal emotional processing proposed by Schirmer and Kotz (2006), the amplitude of the N100, P200 and LPP components is modulated by the acoustical properties, salience and cognitive appraisal of the vocal stimulus, respectively.

N100

The N100 is a negative ERP elicited during the sensory processing stage, generated in the bilateral secondary auditory cortex (peaking approximately at 100 ms after stimulus onset) (Rosburg, Boutros & Ford, 2008). A differentiation between human sounds (i.e., human produced vocalizations) and non-human sounds (i.e., environmental sounds) is already established at this early stage (Murray, Camen, Gonzalez-Andino, Bovet & Clark, 2006; Charest et al, 2009). Most neurophysiological studies suggest no response to the emotional quality of stimuli at this early stage (Garrido-Vázquez et al, 2013; Paulmann & Kotz, 2008; Paulmann et al., 2010; Schirmer et al., 2013). However, some exceptions indicate an early differentiation between emotional and neutral stimuli reflected in the N100, both in non-linguistic vocalizations (Liu et al, 2012; Pell et al, 2015; Wang, Pan, Liu & Chen, 2015) and emotional prosody (Iredale et al, 2013; Pinheiro et al, 2012). Liu and colleagues (2012) reported a more negative amplitude for neutral vocalizations ("mmhm") compared to happy (laughter) and angry ("humph") vocalizations. An early ERP component, the P50 (occurring around 50 ms after stimulus onset), was also elicited by angry vocalizations in this study, indicating increased automatic attention to acoustic cues signalling threat and danger (Liu et al, 2012). Pre-attentive processing (i.e., processing without conscious awareness) differences between emotional and neutral vocalizations were not only shown in terms of amplitude, but also latency, as emotional vocalizations were associated with reduced latency compared to neutral vocalizations, which may indicate they are processed faster (Liu et al, 2012; Pinheiro et al, 2012). Likewise, in emotional prosody studies reduced latency of the N100 was also reported for emotional compared to neutral words and sentences (Iredale et al, 2013; Pinheiro et al, 2012). The processing differences due to the emotional quality of the stimuli in the N100 mentioned above might be accounted for by variations in acoustic parameters (e.g., mean fundamental frequency/ pitch) of the experiment stimuli, as a majority of these studies did not control for this parameter.

A valence-tagging process (pleasant – positive or unpleasant – negative) has been established to take place only at a secondary stage of auditory information processing (P200) (Paulmann, Bleichner & Kotz, 2013). Nevertheless, two studies have reported amplitude differences with non-linguistic affective vocalizations, with happy and angry vocalizations presenting a decreased N100 amplitude when compared to sad and fearful vocalizations, respectively (Pell et al, 2015; Jessen & Kotz, 2011). An initial tagging of emotional relevant cues may be already occurring at this early stage, directing our attention to motivationally salient cues in our environment for the next processing stage of information integration, the P200 (Liu et al, 2012).

P200

The P200 is a positive ERP elicited during the integration stage and it is generated in the auditory cortex in the temporal region (peaking approximately at 200 ms after stimulus onset) (Paulmann, Pell & Kotz, 2008; Schirmer & Kotz, 2006). The P200 is a primary emotional salience detector, aiding in the classificatory process of distinguishing between emotional and neutral vocalizations (Crowley & Colrain, 2004; Schirmer & Kotz, 2006). An enhanced frontocentral positivity distribution is reported for the P200 in affective non-linguistic vocalizations (fear, achievement and disgust) (Sauter & Eimer, 2010) and emotional prosody (Paulmann et al, 2013) when compared to neutral vocalizations and prosody. While some studies report an increased P200 for emotional vocalizations when compared to neutral (Iredale et al, 2013; Liu et al, 2012; Paulmann, Seifert & Kotz, 2010; Pinheiro et al, 2012; 2014; Schirmer et al, 2013), others present a decreased amplitude for emotional vocalizations (Garrido-Vázquez et al, 2013; Paulmann & Kotz, 2008). Nonetheless, emotional and neutral vocalizations are evidenced to be distinguished at this stage regardless of emotion valence (anger, fear, disgust, happiness, surprise and sadness), type of vocal expression (non-linguistic vocalization or speech) (Pell et al., 2015), task-relevance (Kotz & Paulmann, 2007; Paulmann et al, 2013) or acoustic properties of the stimulus (Schirmer et al, 2013).

Valence seems to have a modulatory role in the P200, but whether a valence-tagging process may be initiated at this stage is not consensual across the literature. Some studies argue that the P200 is not sensitive to stimulus valence, with no distinction between a wide range of vocal emotions in emotional prosody (Paulmann & Kotz, 2008;

Paulman et al, 2010; Pinheiro et al, 2014; Iredale et al, 2013) or in non-linguistic affective vocalizations (Sauter & Eimer, 2010). Conversely, increased amplitude has been found for happy in comparison to angry prosody (Pinheiro et al, 2012) and between the six basic emotions (anger, fear, disgust, happiness, pleasant surprise and sadness) (Paulmann et al, 2013), pointing to a valence-tagging process that is indexed by the P200. Increased (i.e., more positive) P200 amplitudes were reported for anger, happiness and pleasant surprise, followed by disgust, sadness and fear, more specifically in frontal and central regions (Paulmann et al, 2013). Peak latencies have also been found to differ depending on emotion valence, with earlier peaks for happiness (laughter), followed by anger (growls) and sadness (sobs) only in non-linguistic vocalizations (Pell et al, 2015). In what concerns arousal, the P200 component seems to be partially responsive to arousal features of vocal expressions, with more arousing stimuli eliciting a more positive amplitude, irrespective of their valence (Sauter & Eimer, 2010; trends in Paulmann et al, 2013). Emotions such as anger and happiness are considered more arousing than sadness based on behavioural results, as such an enhanced sensitivity to arousal features can be observed. This subtle categorization in valence and arousal possibly predicts differences in stimulus appraisal in the following component, the Late Positivity Potential (LPP).

LPP

The LPP is a positive ERP component elicited during the higher-order cognitive evaluation stage (450–700 ms after stimulus onset), and its maximal over centroparietal electrodes (Pell et al, 2015). This late component indexes a more elaborate stage of emotional processing of affective input (i.e., evaluation of the emotional significance of a stimulus) (Kotz & Paulmann, 2012). The LPP has been widely described in visual neurophysiological studies, with greater amplitude when individuals attend to emotional compared to neutral pictures (Dennis & Hajcak, 2009; Brown et al, 2012). Similar findings have been found with auditory stimuli, with LPP being strongly modulated by the emotional category of a vocal stimulus with significant differences being described between the six basic emotions (happiness > pleasant > surprise > anger > fear > sadness > disgust) (Paulmann et al., 2013).

The LPP amplitude is increased for emotional expressions high in arousal compared to expressions low in arousal (Paulmann et al., 2013). A modulatory effect of valence is observed in this component, with a more positive sustained wave for non-linguistic vocalizations which signal cues of threat (anger) than for non-threat cues (happiness and sadness) (Paulmann et al, 2013; Pell et al, 2015). As growls (angry vocalizations) are representative of aggression and threat, the increased positivity may be the result of deeper processing for the promotion of an immediate adaptive response (Pell et al, 2015). Importantly, no interaction between valence and arousal has been reported (Paulmann et al, 2013). Cues signalling social pressure and control have also been found to lead to a preferential and more comprehensive processing, illustrated by a significantly more positive-going LPP waveform for controlling speech (i.e., order to act) than for autonomy-supportive speech (i.e., presents the choice to act) and neutral speech (Zougkou et al, 2015).

A recent study probing the processing of insults and how this processing can be modulated by the presence of a laughing crowd, found that the LPP was more positive when the insults and compliments were accompanied by a laughing crowd, compared to when they were presented without it. Even without laughter, verbal insults presented an increased LPP compared to compliments showing that insults elicit deeper emotional processing (Otten et al, 2017). This finding highlights the importance of the social context in which the communication is taking place. Not only insults elicited increase elaborative processing in the presence of a laughing crowd, but also compliments which were possibly perceived as a sarcastic comment (Otten et al, 2017).

The ERP studies aforementioned used sets of stimuli produced by actors posing emotions (voluntary expressions of emotion) instead of spontaneous expressions of the same emotions which may be more representative of real-life social interactions. It remains unclear to which extent may the previous results be explained by differences in emotion's authenticity. Additionally, considering authenticity is present in our daily social interactions and that our ability to encode and decode voluntary expressions accurately presents an evolutionary advantage, it would be an important dimension of voice processing to explore in more detail.

4. Authenticity

Discriminating whether an emotion was spontaneously or voluntarily expressed in everyday social interactions is an important and advantageous social skill in order to avoid deception (Gervais & Wilson, 2005; Lavan, Rankin, Lorking, Scott & McGettigan, 2017). Equally important is our ability to produce these same voluntary expressions in our social interactions, for example for group affiliation (Lavan, Short, Wilding & McGettigan, 2018). The ability of humans to imitate the emotional expressions of others, even if their current emotional state is different, may constitute a social skill prompting affiliation and cooperation (e.g., politely agreeing with someone even though you do not share a similar view on a given subject) (Scott, Lavan, Chen & McGettigan, 2014).

In research, three approaches to the development of sets of emotional expressions can be described (Scherer & Bänziger, 2010). The most common approach is to request professional actors or amateurs to portray an emotion with a set of guidelines (voluntary portrayal of emotion). The second approach is to induce a genuine emotional state in participants through presentation of external triggers (e.g., presenting a funny video to induce laughter) or by requesting them to recall a personal memory (e.g. feelings of loss over a close person passing away to induce crying) (spontaneous portrayal of emotion). Lastly, a more ecologically valid approach includes recording spontaneous expressions of emotion through field observation (spontaneous portrayal of emotion) (Anikin & Lima, 2017; Scherer & Bänziger, 2010). This last approach is methodologically challenging as confounding factors can influence the recording, as well as time-consuming. However, the current state of social media (facebook, youtube, twitter) allows researchers to access audio and video of individuals thoroughly engaged in highly emotional activities, serving as a tool for future studies (Anikin & Lima, 2017). Authenticity as a modulating aspect of emotion recognition has garnered more attention recently, mainly with studies concerning laughter and discrimination of smiles authenticity (Lavan & McGettigan, 2016). Laughter is a non-verbal positive expression of emotion recognized across distinct cultures that has a critical role in promoting and maintaining social bonding (Scott et al, 2015). There are two main types of laughter: spontaneous laughter (stimulus-driven and genuine) and voluntary laughter (deliberate and associated with an intentional communicative act by signalling affiliation or polite agreement) (Bryant & Atkipis, 2014; Scott et al, 2015). Smiling is a positive non-verbal signal (facial expression) denoting enjoyment, also widely recognized across cultures and signalling affiliation (Ekman, 1982). Smiling can also be categorized regarding its authenticity: spontaneous (Duchenne smile) and voluntary (non-Duchenne smile) smiles (Gunnery & Ruben, 2013). While the spontaneous smile engages the corner of the eyes (contraction of the orbicularis oculi muscle) and lifts the corners of the mouth (zygomatic major muscle), signalling genuine enjoyment, the voluntary smile lacks engagement of eye muscle contractions, signalling politeness or masking negative emotions (Ekman, Friesen & Hager, 2002; Gunnery &

Ruben, 2013). Individuals can distinguish between spontaneous and voluntary smiles due to morphological differences in facial expressions (Krumhuber & Manstead, 2009) and in laughter based on acoustic differences (Bryant & Aktipis, 2014; Lavan, Scott, & McGettigan, 2016).

Studies with laughter indicate that listeners perceive spontaneous and voluntary laughter distinctively in terms of authenticity, arousal and valence (Lavan et al, 2016). This differentiation has been encountered more robustly with non-linguistic affective vocalizations, as in most studies with emotional prosody stimuli listeners do not accurately detect authenticity (Jürgens, Drolet, Pirow, Scheiner, & Fischer, 2013; Jürgens, Grass, Drolet, & Fischer, 2015; Scherer, 2013). Nonetheless, a study by Drolet, Schubotz, and Fischer (2012) reported accurate detection of authenticity in speech above chance levels (Anikin & Lima, 2017). As for authenticity detection in non-linguistic affect vocalizations, behavioural evidence indicates an accuracy of 61% (Bryan & Atkipis, 2014) and 72% (Lavan et al, 2015). The higher accuracy in Lavan and colleagues (2015) study may be related to the pre-selection of the stimuli in which the selection criteria optimized authenticity detection (Anikin & Lima, 2017). Regarding valence and arousal perception, spontaneous laughter is rated by listeners as higher in arousal and more positively valenced than voluntary laughter (Lavan et al, 2016).

Neuroimaging studies have suggested spontaneous and voluntary laughter may engage different neural systems (McGettigan et al, 2015; Lavan et al, 2017). Whereas spontaneous laughs (designated as “evoked”) elicit greater activation in bilateral primary auditory cortex (Heschl’s gyrus) and superior temporal gyrus (STG), voluntary laughs (designated as “emitted”) elicit increased activation in mentalizing areas (anterior medial prefrontal cortex (amPFC) and anterior cingulate gyrus) (McGettigan et al, 2015). “Mentalizing” refers to the action of inferring emotions, beliefs and intentions of others and using them as cues in social interactions (Baron-Cohen, 2004). Mentalizing plays a role in discriminating the authenticity of a non-linguistic affective expression by attending to its acoustic properties and interpreting its meaning (McGettigan et al, 2015). In laughter, results reflect a differential level of engagement of the auditory cortex: spontaneous laughter presents is more rapidly and automatically processed than voluntary laughter, which requires the interpretation of the expression and motivation behind the laugh (mentalizing) (McGettigan et al, 2015). Hence, while spontaneous and voluntary laughter are both processed in the auditory cortex, voluntary laughter requires further processing in mentalizing areas for its understanding (McGettigan et al, 2015). Further support for these results comes from a study that found similar activations in the anterior medial pre-frontal cortex (amPFC) when participants listened to more socially complex laughter (joy, taunting laughter), while tickling laughter elicited a response in the right STG, similarly to McGettigan and colleagues (2015) (Szameitat et al, 2010). A recent study suggested that the increased response in the amPFC to voluntary laughter corresponds not only to the engagement of mentalizing processes by attributing a mental state/ motivation to the voluntary laughter, but also to a reflection of the social ambiguity of the vocalizations (Lavan et al, 2017). A linear decrease was observed in amPFC activation when authenticity detection increased, further corroborating that when the listener hears a clearly voluntary laughter, the amPFC is engaged to resolve the cause of the ambiguity. According to this evidence, the amPFC may have a role in higher order resolution of the meaning of emotional cues and their underlying cause and not specifically in categorical properties such as authenticity discrimination (Lavan et al, 2017).

In everyday life, emotional signals reach us in multiple modalities, with several studies showing that congruent information from multiple modalities improves emotion decoding accuracy (Lavan & McGettigan, 2016). Laughter authenticity discrimination is increased in multimodal (audio-visual) compared to unimodal contexts (visual or auditory channel), with a major influence of the auditory channel (Lavan & McGettigan, 2016). This auditory dominance may be explained by the auditory nature of laughter, prioritizing our extraction of affective information from this channel despite visual displays. Conversely, with sadness the main influence for decoding may be the visual channel, with the presence of tears in adulthood (Provine, Krosnowski, & Brocato, 2009), as spontaneous auditory crying has been found to be ambiguously categorized and confused with spontaneous laughter (Lavan, Lima, Harvey, Scott, & McGettigan, 2014).

A relationship between authenticity and emotional contagion ("propensity to resonate with others' emotions" (Neves et al, 2018, p. 3) has been recently described. Higher trait levels of empathy (experiencing a similar emotion displayed by another person) and emotional contagion (self-report questionnaires) and perceived emotional contagion (subjective ratings) seem to enhance laughter authenticity discrimination (computed an index of authenticity to determine individual ability to discriminate spontaneous and voluntary laughter), without any sex differences being reported (Neves et al, 2018). Both high mentalizing abilities (Dawel, Palermo, O'Kearney & McKone, 2015) and high trait empathy (Neves et al, 2018; Dawel et al, 2015) have been positively associated with authenticity discrimination in vocal cues. Mentalizing processes may promote cognitive strategies to infer authenticity in order to understand the meaning of the vocalization. On the other hand, high trait empathy may facilitate authenticity discrimination between spontaneous and voluntary expressions by simulation (emotional contagion), as individuals with a high trait empathy may respond with a stronger emotional response to spontaneous expressions (Dawel et al, 2015).

The portrayal of emotions by simulation (e.g., professional actor/ researcher is asked to pose a given emotion) has been thoroughly used in most studies for stimulus development, not only in the case non-linguistic affective vocalizations but also of facial expressions for behavioural, neuroimaging and ERP studies. The accumulating neuroimaging and behavioural evidence signal emotion decoding differences in the processing of spontaneous and voluntary laughter, in terms of its social and affective properties. It remains to be determined whether the authenticity of emotional vocal expressions may have influenced the results of the findings previously described. Previous imaging findings have shown which brain areas are involved in the processing of affective vocal authenticity (McGettigan et al, 2015). Nonetheless, while fMRI studies have elucidated the areas putatively involved in the processing of affective vocal authenticity, they fail to provide insight into the time course underlying authenticity recognition in vocal cues. On the other hand, the ERP technique is characterized by a high temporal resolution in the order of the milliseconds (ms), which is critical to determine the exact time-window on which differences in the processing between spontaneous and voluntary vocalizations in the brain may occur. When we study auditory stimuli such as non-linguistic vocalizations which are not static but dynamic stimuli, the ERP technique presents advantages in understanding how an auditory cue is processed in real time from stimulus onset to offset. An EEG study of the effects of authenticity in the temporal course of voice processing is thus of great importance to better understand emotional voice processing in the brain.

5. Sex differences

Evidence for sex differences in emotion recognition has been mixed. On the one hand, some studies found that women perform more accurately than men in recognizing nonverbal signs in different sensory modalities (visual, auditory, audio-visual)

(Collignon, Girard, Gosselin, Saint-Amour, Lepore & Lassonde, 2010), while other studies highlight that sex differences may be due to uncontrolled factors (sensory modality, sex of the actor and emotion) in the experimental design (Thompson & Voyer, 2014). Sensory modality functions as a moderator possibly driving sex differences as marginally larger effect sizes are found for non-verbal emotion recognition in the audio-visual modality than for the visual modality. The sex of the actor also seems to moderate these effects, as larger effect sizes for a female advantage in emotion recognition are reported when listening to male actors, compared to female or mixed actors. Lastly, larger effect sizes are obtained for negative emotions than for positive emotions which may be due to a general increased variability in negative emotions recognition (Thompson & Voyer, 2014).

In vocal emotional processing, sex differences indicating a female advantage were only found in emotional prosody studies revealing increased mismatch negativity (MMN) (elicited by the detection of auditory changes) and N400 (elicited by memory retrieval and integration of a word in a context) amplitudes, when compared to male listeners (Schirmer & Kotz, 2003; Schirmer, Kotz & Friederici, 2005; Schirmer, Striano & Friederici, 2005). However, no studies to date have indicated sex differences in the N100, P200 and LPP components during emotional vocal processing. Female listeners were reported to show a larger MMN to emotional (angry and happy) than to neutral vocalizations outside their attentional focus, in a German and Asian sample (Hung & Cheng, 2014; Schirmer, Striano & Friederici, 2005). An interaction between the semantic and the emotional meaning of a word (positive and negative words) was also found in women when asked to judge emotional congruence between the word presented and the emotion portrayed. This interaction was reflected in an increased N400 amplitude for emotional prosody than for neutral speech (Schirmer & Kotz, 2003; Schirmer, Lui, Maess & Escoffier, 2006). Later, a study by the same authors reported no sex differences between men and women in N400 amplitude, as both sexes showed a decreased N400 amplitude to emotional congruent words compared to incongruent words (Schirmer, Kotz & Friederici, 2005). The result is justified by the authors as a consequence of task design differences, as the latter study had a shorter interval between prime and target. Additionally, in a sample comparing men and women performance in a semantic task, N400 differences were found to be related to inter-individual variability (empathy) and not to sex differences (Van der Brink et al, 2012). According to the studies afore mentioned, women seem to be better at remembering the emotional tone, using emotional prosody more automatically for affective language processing than men. These findings suggest that both men and women are sensitive to the emotions conveyed in the vocalizations, but encode the information differently (Schirmer & Kotz, 2003; Schirmer, Striano & Friederici, 2005).

AIMS OF THE STUDY

The present study aimed to explore how authenticity (i.e., whether an emotion is evoked spontaneously or voluntarily) modulates the temporal course of vocal information. We investigated potential differences in the latency and the amplitude of the N100, P200 and LPP ERP components between spontaneous and voluntary vocalizations for happiness (laughter) and sadness (crying), considering that these ERPs may capture relevant stages of the temporal processing of different aspects of the vocalization.

Based on previous neurophysiological findings stating emotional and neutral vocalizations are processed differently, we hypothesized differences between neutral and emotional non-linguistic vocalizations at early processing stages (N100, P200 and LPP components) (Hypothesis 1). We expected to replicate previous accounts of a differential effect of neutral and emotional vocal processing, given that previous research showed that emotional stimuli elicit increased amplitude in N100, P200 and LPP components in non-linguistic vocalizations (Liu et al, 2012; Jessen & Kotz, 2011; Pell et al, 2015; Wang et al, 2015) and emotional prosody (Iredale et al, 2013; Paulmann et al, 2010; 2013; Pinheiro et al 2012; 2014; Schirmer et al, 2013; Otten et al, 2017).

In light of functional neuroimaging (McGettigan et al, 2015; Lavan et al, 2017) and behavioural evidence (Lavan et al, 2014; 2015; 2016) establishing a differential processing of spontaneous and voluntary laughs, we hypothesized that spontaneous and voluntary vocalizations would be distinguished from one another at different time points during vocal processing and reflect distinct ERP effects following a multi-stage approach.

Building on previous neurophysiological evidence suggesting the P200 and LPP amplitude and latency is affected by how motivationally salient the event is in non-linguistic affective vocalizations (Pell et al, 2015), we hypothesized that authenticity would modulate the P200 and LPP components (Hypothesis 2). We speculated that voluntary vocalizations would present increased amplitude and earlier latency in the P200 and LPP components compared to spontaneous vocalizations, due to its increased saliency.

In addition, as an exploratory hypothesis due to the mixed evidence in the literature (Schirmer & Kotz, 2003; Schirmer, Kotz & Friederici, 2005; Schirmer, Striano & Friederici, 2005), we explored sex differences in authenticity discrimination. Authenticity possibly modulates the P200 and LPP components differently in women and men, with women presenting increased amplitude in the P200 and LPP components, comparably (Hypothesis 3).

METHODS

1. Participants

A total of 38 individuals (21 men and 17 women) participated in this experiment. Fifteen participants had to be excluded due to problems in data acquisition (n = 8) or low signal-to-noise ratio (n = 7).

The final sample included 23 healthy college students (13 females) with an average age of 23.43 (*SD* = 1.67, range = 22 – 28 years). They were recruited through neurocolab.wordpress.com, a recruitment platform that was developed by the research team. The inclusion criteria for this study were: right handedness (Edinburgh Handedness Inventory) (Oldfield, 1971); European Portuguese as a first language. For female participants, an additional inclusion criterion was to be on the active weeks of contraceptive pills (Radke & Derntl, 2016). Participants provided written informed consent and were paid for their participation in the 2-hour and 30 min study.

The Positive and Negative Affect Schedule (PANAS) (Watson, Clark, & Tellegen, 1988, Portuguese version, Galinha & Ribeiro, 2005) was administered to evaluate the participants' current emotional state to ensure a low variability of mood states in the sample (Positive Affect Score – *M* = 21.65, *SD* = 5; Negative Affect Score – *M* = 15.39, *SD* = 5.50). A cognitive assessment included the Working Memory Index of the Weschler Adult Intelligence Scale – Third Edition (WAIS – III) (Weschler, 2008) to measure working memory (WM Index = 19.87, *SD* = 3.79), as deficits in working memory result in difficulties performing simple cognitive tasks and decision-making. As a control measure the Brief Symptom Inventory (Derogatis & Melisaratus, 1983; Portuguese version - Canavarro, 2007) was administered (Global Severity Index (GSI) = 1.68, SD = 0.46) as a control measure for participant's psychopathology (Table 1). No participant was excluded based on these scores.

Table 1. Socio-demographic characteristics of the sample.

		Age (years)	
Sex		**Mean ± SD**	**Range**
Male	(*N* = 10)	23.70 ± 1.95	22 – 28
Female	(*N* = 13)	23.23 ± 1.48	22 – 26

		Mean ± SD	Range
Positive Affect Score (PANAS)		21.65 ± 5	14 – 31
Negative Affect Score (PANAS)		15.39 ± 5.5	10 – 33
WM Index		19.87 ± 3.79	13 – 27
BSI (GSI)		1.68 ± 0.46	1.09 – 2.50
	Male	1.79 ± 0.49	1.13 – 2.50
	Female	1.6 ± 0.44	1.09 – 2.45

PANAS = Positive and Negative Affect Schedule, WM = Working Memory, BSI = Brief Symptom Inventory, GSI = Global Symptom Index.

Authenticity discrimination has been positively correlated with empathy (ability to share the emotional experiences of another) and mentalizing abilities (action of inferring emotions, beliefs and intentions of others) (Dawel et al, 2015; Neves et al, 2018). Empathy was assessed by administering The Empathy Quotient (EQ) (Baron-Cohen & Wheelwright, 2004; Portuguese version, Rodrigues et al, 2011): cognitive empathy (*M* = 10.83, *SD* = 2.34); emotional reactivity (*M* = 9, *SD* = 2.04); social skills (*M* = 9.65, *SD* = 2.20) and empathic difficulties (*M* = 13.30, *SD* = 2.72). Individual differences in mentalizing were assessed by applying the Reading the Mind in the Eyes Test (RMET) (Baron-Cohen, Wheelwright, Hill, Raste & Plumb, 2001; Portuguese version - Mouga & Tavares) (*M* = 26.30, *SD* = 3.15) (Table 2). Men and women did not differ significantly in empathy or mentalizing scores ($p > 0.5$).

Table 2. Empathy and mentalizing scores of the sample.

		Mean ± SD	Range
EQ		21.48 ± 7.01	8 – 36
	Male	19.6 ± 8.15	8 – 36
	Female	22.92 ± 5.93	15 – 34
RMET		26.3 ± 3.15	21 – 32
	Male	25.5 ± 3.13	22 – 30
	Female	26.92 ± 3.14	21 – 32

EQ = Empathy Quotient, RMET = Reading the Mind in the Eyes Test.

2. Auditory Stimuli

The auditory stimuli consisted of nonverbal spontaneous and voluntary vocalizations of laughter and crying (happiness and sadness) and two types of neutral vocalizations (spectral rotations, i.e. obtained by inversing the spectral characteristics of neutral or emotional vocalizations resulting in unintelligible non-emotional sounds; neutral vocalizations, i.e. vowels uttered with neutral intonation). Two types of neutral vocalizations (spectral rotations and neutral vocalizations) were included in the current study given that ERPs studies have invariably selected one or the other, but no study to date has explored which one would be methodologically advantageous. Although spectral rotations are a good match in spectro-temporal complexity to the emotional stimulus presented in those studies, they have been criticized for being anti-natural sounds not present in our environment and as such introducing noise in how we process them. As such, a secondary goal of the current study is to define which type of neutral vocalization is more adequate for a neurophysiological approach.

Spontaneous vocalizations consist of spontaneously produced vocalizations either in response to a humorous video (spontaneous laughter) or recalling of upsetting events (spontaneous crying). Voluntary vocalizations consist of acted expressions under full voluntary control. The set of stimuli used was developed at the University College of London and has been validated at both behavioural and neuroimaging levels (Lavan et al, 2014; McGettigan et al, 2015), except the neutral spectral rotations which were pre-tested in a Portuguese sample through an online questionnaire (Qualtrics), as further detailed. The vocalizations were produced by six speakers (3 women) in a soundproof anechoic chamber. For the recording of authentic laughter (spontaneous), YouTube videoclips which each speaker identified beforehand as amusing were shown, inducing them to laugh out loud. Regarding the recording of spontaneous crying, speakers were encouraged to recall personal upsetting events and/or start by pose crying in order to transition to genuine crying. Lastly, for the recording of voluntary laughter and crying, speakers were instructed to simulate laughter without experiencing any genuine feelings of amusement and to simulate crying without any genuine feelings of sadness, respectively. In order to avoid carry-over effects of genuine amusement/ sadness, the recording of the voluntary laughter/ crying always preceded the recording of spontaneous laughter/ crying. From the raw recordings separate files of laughter and crying vocalizations were sampled at a rate of 44 100 Hz to mono.wav files with 16-bit resolution. To control for the high acoustical properties variability of the raw recordings, the audio was normalised for root-mean-square (RMS) amplitude using Praat software (www.praat.org) (Lavan et al, 2014).

Regarding the spectrally rotated neutral vocalizations, an online questionnaire was created through Qualtrics platform (www.qualtrics.com) and distributed through a link. A total of 83 sounds were presented in a randomized sequence: 69 spectrally rotated sounds (34 from emotional vocalizations and 35 from baseline vocalizations) and 14 emotional sounds (3 spontaneous laughter, 4 voluntary laughter, 3 spontaneous crying, and 4 voluntary crying). The instructions were to attend to the auditory stimuli and evaluate them according to their valence in a Likert scale (1 – Negative, 4 – Neutral, 7 – Positive). A total of 27 participants (11 male) who did not take part in the main study responded to the questionnaire (mean age = 36.36, S*D* = 13.56 years). Mean, mode and standard deviation and duration was calculated for each sound. The selection criteria were the following: *Mode* = 4 (neutral) and *Mean* > 3. Two sets of thirty spectral rotations were developed in order to choose a combination similar duration to the other experimental conditions: one set with 15 baseline vocalizations which were spectrally rotated and 15 emotion vocalizations spectrally rotated and a second set, with 20 baseline spectral rotations and 10 emotion spectral rotations. The set with 20 baseline spectral rotations (which tend to have longer duration) and 10 emotions spectral rotations (tend to have shorter duration) was selected, as its average duration was more similar to the other experimental conditions. The selected spectral rotations were perceived as neutral (*M* = 3.228, *SD* = 0.194) and had a mean duration of 2.282s (*SD* = 0.689s).

A total of 132 nonverbal vocalizations were included in the experiment (18 spontaneous laughter, 18 voluntary laughter, 18 spontaneous crying, 18 voluntary crying, 30 spectrally rotated neutral, and 30 neutral vocalizations). Each of the emotional vocalizations was presented twice. The number of vocalizations produced by women and men was similar across conditions: nine produced by women and nine by men for spontaneous laughter/ crying and voluntary laughter/ crying; and fifteen by women and fifteen by men for neutral vocalizations. The acoustical properties of the stimuli (duration (ms), mean fundamental frequency – F (0), mean intensity (dB)) were obtained using Praat software (www.praat.org) (Table 3).

Table 3. Acoustic properties of the experimental stimuli.

	Stimulus Type	**Duration (ms)** *(Mean ± SD)*	**Mean F0 (Hz)** *(Mean ± SD)*	**Mean Intensity (dB)** *(Mean ± SD)*
Positive	Spontaneous Laughter	2399.94 ± 460.73	397.13 ± 90.62	66.10 ± .10
	Voluntary Laughter	2248.89 ± 400.15	257.84 ± 60.26	66.04 ± .11
Negative	Spontaneous Crying	2685.44 ± 289.36	421.38 ± 57.04	63.40 ± 3.10
	Voluntary Crying	2182.61 ± 351.48	368.62 ± 87.75	64.64 ± 6.93
Neutral	Vocalizations	2498.73 ± 292.08	182.13 ± 54.01	64.81 ± 0.04
	Spectral Rotations	2282.70 ± 689.28	235.32 ± 112.2	64.31± 4.48

Ms = milliseconds, F0 = fundamental frequency, Hz = hertz, dB = decibel.

Due to the normality assumption not being fulfilled for mean duration, mean intensity and mean fundamental frequency, non-parametric Kruskal-Wallis H Tests were performed, followed by pairwise comparisons with Bonferroni correction for multiple comparisons. Significant differences in the acoustic parameters were found for mean duration ($\chi 2(4) = 18.094$, $p = .001$), mean fundamental frequency ($\chi 2(4) = 64.499$, $p = .001$) and mean intensity ($\chi 2(5) = 59.459$, $p = .001$).

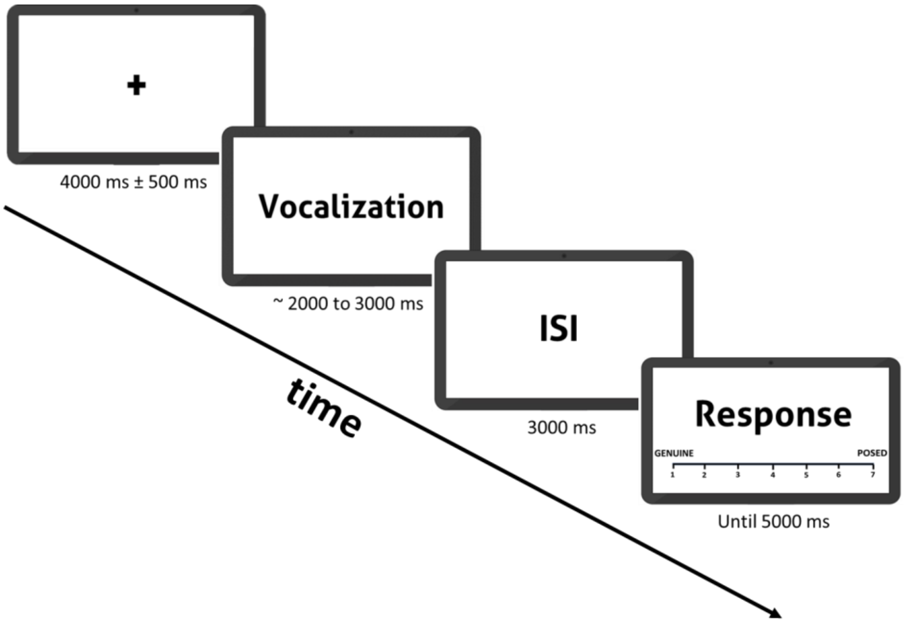
Follow-up comparisons for mean duration revealed the distributions differed significantly with spontaneous vocalizations showing increased duration of the stimuli compared to voluntary vocalizations only for crying vocalizations ($p = .001$). Concerning differences in emotionality, negative vocalizations presented a higher duration than positive vocalizations (spontaneous: $p < .001$; voluntary: $p < .001$) irrespective of authenticity. Acoustic mean duration differences were also found between neutral vocalizations and voluntary vocalizations, with increased duration for neutral than for both laughter ($p = .023$) and crying vocalizations ($p = .024$).

Pairwise comparisons showed the distributions of mean fundamental frequency differed significantly in terms of authenticity with spontaneous laughter presenting a higher fundamental mean frequency than voluntary laughter. No significant differences were found for mean fundamental frequency between spontaneous crying and voluntary crying ($p = .062$). Concerning the emotionality of the vocal expression, negative vocalizations presented an increased mean fundamental frequency compared to positive vocalizations, only for voluntary vocalizations ($p = .046$), with no differences being found for spontaneous vocalizations ($p = .448$). Neutral vocalizations presented a decreased fundamental frequency compared to all the emotional conditions, spontaneous laughter (*p* < .001), voluntary laughter (*p* < .001), spontaneous crying (*p* < .001) and voluntary crying vocalizations (*p* < .001).

Pairwise comparisons revealed the distributions of mean intensity did not differ significantly according to the authenticity of the emotional expressions (positive: *p* = .078, negative: *p* = .784). Regarding emotionality, positive vocalizations presented a higher mean intensity than negative vocalizations (spontaneous: *p* < .001; voluntary: *p* < .001), irrespective of authenticity. Neutral vocalizations presented a lower mean intensity than spontaneous laughter (*p* < .001), voluntary laughter (*p* < .001), spontaneous crying (*p* = .004) and voluntary crying vocalizations (*p* = .001).

3. Task

Before starting the task, the experimenter explained to participants that they would hear a set of sounds and would rate the emotional sounds in terms of their perceived authenticity (i.e., whether a sound is genuine or posed). Concerning neutral sounds, they were instructed only to attend to the stimulus. A trial started with a 4000 ms fixation cross with a jitter of 500 ms, followed by the presentation of the vocal expression. Before the authenticity rating, where the participant had up to 5000 ms to respond, a 3000 ms inter-stimulus interval consisting of a fixation cross was presented. Participants responded in a 7-point Likert scale their perceived authenticity of the stimuli presented, ranging from 1 (“Genuine” – spontaneous) to 7 (“Posed” – voluntary). The task had a total of 204 trials (duration = 36 min) with a fixed sequence presentation of the stimuli. The rationale for presenting a fixed sequence, instead of a pseudorandomized sequence, relates to the future applicability of this task with oxytocin intranasal intake since a variable order of the stimuli and oxytocin may interact not allowing a fair comparison between subjects. Transitions from trial to trial were taken into account in the sequence design by distributing the several types of transition from one condition to another equally. The experimental design is outlined in Figure 1.



ISI = Interstimulus interval.

Fig. 1 Illustration of an experiment trial.

After EEG data acquisition, participants were instructed to evaluate the perceived arousal and emotional contagion of the previously presented vocal stimuli in a 7-point Likert scale (Arousal: 1 – Low arousal, 7 – High arousal; Emotional Contagion: 1 – Not contagious at all, 7 – Highly contagious). Stimuli were divided across two blocks: the first block assessed the perceived arousal of each stimulus, whereas the second block assessed the perceived emotional contagion of each stimulus. Each block had a total of 72 trials (18 spontaneous laughter, 18 voluntary laughter, 18 spontaneous crying, 18 voluntary crying). A trial had the following sequence: a fixation cross presented during 1500 ms with a jitter of 500 ms, presentation of the vocalization, fixation cross during 1s and lastly, perceived arousal or emotional contagion rating depending on the block. The task was presented in a fixed sequence which accounted for transitions and had a total of 124 trials (15 min). Each vocalization was only presented once in each block.

4. EEG Data Acquisition and Processing

The electroencephalogram (EEG) was recorded at a continuous rate using a 64-channel BrainVision actiCHamp system (Brain Products, München, Germany) at a 512 Hz sampling rate. For offline reference, two flat-type electrodes were placed on the left and right mastoids. Bipolar horizontal and vertical electro-oculograms were acquired through 4 flat-type facial electrodes: two electrodes were placed at the outer corner of each eye (horizontal electro-oculogram) and two electrodes were placed below and above the left eye (vertical electro-oculogram). Electrode impedance was kept under 10kΩ for all electrodes. Offline EEG analyses was processed using BrainVision Analyser software (Brain Products GmbH, Munich, Germany) using as reference the average of the left and right mastoids. A delay between the stimulus presentation computer presenting the stimuli and the sound reaching the participant through the earphones was detected. As such, we applied a delay correction of 464.4 ms to the whole sample with the exception of one participant whose EEG data were collected posteriorly (delay of 116.09 ms).

A low-pass filter of 30 Hz and a high-pass filter of 0.1 Hz were applied, as well as a notch filter of 50 Hz in order to reduce electrical noise present in the experimental room. Individual event-related epochs, time-locked to the onset of the auditory stimuli, were defined starting 200ms before each vocalization onset and ending 1000ms after stimulus onset for each stimulus type (spontaneous laughter, voluntary laughter, spontaneous crying, voluntary crying, neutral spectral rotations, neutral vocalizations). The EEG data were baseline corrected from -200ms to 0ms before the stimulus presentation, followed by the correction of eye blink and movement artifacts using Gratton, Coles, and Donchin (1983) method. Automatic artifact rejection was performed before averaging, in order to exclude trials containing excessive blinks, eye movements and/or muscle activity (criteria: exceeding ± 100 mV). Participants were excluded if 75% of the trials for each condition did not pass artifact rejection. After visual inspection of the averages for each condition, neutral spectral rotations were excluded from further analyses due to an abnormal EEG morphology. This condition presented a high variation regarding the stimulus sound start.

The channels F3/Fz/F4, FC3/FCz/FC4, C3/Cz/C4, CP3/CPz/CP4 and P3/Pz/P4 electrodes were selected for statistical analyses based on inspection of grand average waveforms and previous voice processing ERP studies probing emotionality effects (Jessen & Kotz, 2011; Pinheiro, Barros, Dias & Niznikiewicz, 2017). Mean amplitudes were computed for the components N100 (time window = 100 to 200 ms after stimulus onset), P200 (time window = 200 to 300 ms after stimulus onset) and LPP (time window = 500 to 700 ms after stimulus onset) for each participant and condition. The time windows for each component were defined through visual inspection of the averages per condition and previous neurophysiological voice processing studies with similar stimulus duration (Pinheiro et al, 2012).

5. Procedure

The experiment consisted of one single individual session in a quiet room, lasting a total of two hours and half. Participants were seated in a chair at a distance of 80 cm away of a desktop computer with a 64-channel EEG cap and were instructed to remain as still as possible to avoid eye and motor artefacts. They were instructed to evaluate the authenticity of the auditory stimuli on Likert scale of 7 points (1 – “Genuine” to 7 – “Posed”). The auditory stimuli were presented binaurally through a set of Sennheiser CX 3.00 ear-canal phones at a comfortable listening level that was individually adjusted at the start of the experiment. Stimulus presentation, timing of events and subject’s responses were recorded using Matlab version 8.3.0 (R2014a) with Psychtoolbox 3. Participants were encouraged to respond as intuitively as possible given their 5s time limit. Buttons of the response pad were marked with the Likert scale points to minimize memory demands. In order to facilitate a quick response, participants were asked to put three fingers of their left hand and four of their right hand on the response keys (left hand – 1, 2, 3; right hand – 4, 5, 6, 7). Given the long duration of the task (36 minutes), three pauses of 30s were distributed equally throughout the experiment for the participant to rest and minimize fatigue effects. After the EEG session, participants rated perceived arousal and emotional contagion of each of the sounds presented, as well as responded to the Empathy Quotient (EQ) and Reading the Mind in the Eyes Test (RMET).

6. Statistical Analyses

The SPSS statistical software package (Version 24, SPSS Inc., Chicago, IL, USA) was used for the statistical analyses, with an alpha level set at *.05*.

6.1 Behavioural Statistical Analyses

Effects on mean reaction times (RTs) and authenticity, arousal and emotional contagion ratings were tested through Repeated-Measures ANOVAs with emotion valence (positive, negative) and authenticity (spontaneous, voluntary) as within-subject factors and sex as a between-subject factor. Main effects were followed by multiple comparisons with Bonferroni correction. For the behavioural statistical analyses of arousal and emotional contagion, data was only obtained from 20 participants of the whole sample. An outlier (> 2 SD) was detected in arousal ratings: after the removal of the outlier a total of 19 participants were included in this analysis.

6.2 EEG Statistical Analyses

To investigate the role of emotion and relevant interacting factors, in voice processing, the effects on mean amplitudes and peak latency of N100, P200 and LPP were tested through two Repeated-Measures Analyses of Variance (ANOVA) with emotion valence (positive, negative, neutral), region of interest (ROI) (frontal, fronto-central, central, central-parietal, parietal) and electrode (3, z, 4) as within-subject factors and sex as a between-subject factor – one ANOVA for spontaneous vocalizations and a separate one for voluntary vocalizations. Bonferroni correction was applied for multiple comparisons.

In order to investigate the role of authenticity differences, and relevant interacting factors, in voice processing, the effects of mean amplitudes and peak latency of N100, P200 and LPP were tested through repeated measures ANOVAs with emotion valence (positive, negative), authenticity (spontaneous, voluntary), ROI (frontal, fronto-central, central, central-parietal and parietal) and electrode (3, z, 4) as within-subject factors and sex as a between-subject factor. Given that most research on authenticity discrimination has been conducted with laughter stimuli (Lavan et al 2015; 2016; 2017; McGettigan et al, 2015) and that voluntary crying has been ambiguously categorized as spontaneous laughter in previous studies (Lavan et al, 2014), a separate analysis was conducted only including spontaneous and voluntary laughter. Bonferroni correction was applied for multiple comparisons.

As 8 participants of the sample presented a Global Severity Score (GSI) in the Brief Symptom Inventory (BSI) above the cut-off of the Portuguese population (1.7), the previously described analyses were repeated with the BSI score as a co-variate, to ensure the results found were not influenced by these scores.

Since laterality effects have been reported in voice processing studies (Kotz & Paulmann, 2012), an exploratory analysis was performed to evaluate the effects of authenticity and a possible interaction with laterality: N100, P200 and LPP were assessed in a repeated-measures ANOVA with emotion (positive, negative), authenticity (spontaneous, voluntary), ROI (frontal, fronto-central, central, central-parietal and parietal) and hemisphere (left/ right) as within-subject factors and sex as between-subject factor. This analysis included the following electrodes: left hemisphere (comprised by F3, FC3, C3, CP3 and P3) and right hemisphere (comprised by F4, FC4, C4, CP4 and P4), excluding midline electrodes (Fz, FCz, Cz, CPz and Pz). Bonferroni correction was applied for multiple comparisons.

Comparisons with more than one degree of freedom in the numerator were corrected for non-sphericity using Greenhouse-Geisser correction (Greenhouse & Geisser, 1959). Effects sizes for significant effects (*p*-value ≤ .05) are reported using the partial n-square method (η_p^2).

6.3 Correlations

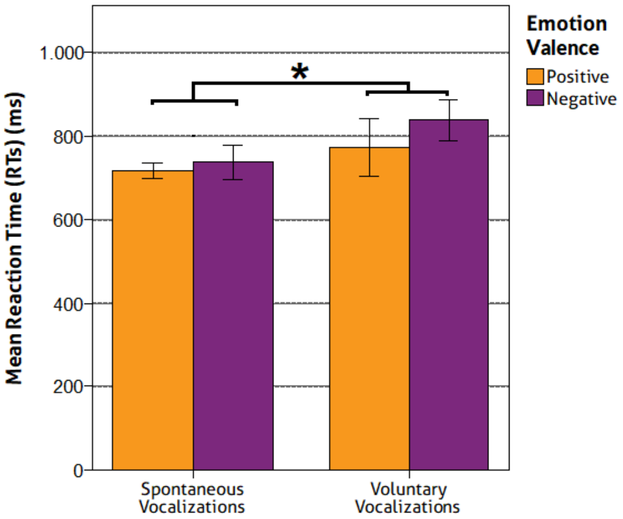
Pearson correlations (two-tailed, $p < .05$) corrected for multiple comparisons were computed to examine the relationship between behavioural measures (authenticity ratings, arousal ratings, emotional contagion ratings, EQ scores, RMET scores) and the mean amplitude of N100, P200 and LPP.

RESULTS

1. Behavioural Results

1.1 Mean Reaction Time

A significant main effect of authenticity was found on mean reaction time, showing a difference in reaction time (ms) between spontaneous vocalizations and voluntary vocalizations across subjects [*F* (1, 21) = 10.794, *p* = .004, η_p^2 = .879]: spontaneous vocalizations (*M* = 725.67, *SD* = 12.032) elicited a faster response than voluntary vocalizations (*M* = 807.668, *SD* = 22.381), irrespective of emotion valence (Figure 2, Table 4). No significant main effect of sex (*p* = .948) or interactions involving it were found (authenticity*sex: *p* = .236; emotion valence*sex: *p* = .609), except a significant interaction effect between authenticity, emotion valence and sex [*F* (1, 21) = 4.544, *p* = .045, η_p^2 = .178], follow-up comparisons were not significant.



*. Effect is significant at the 0.05 level.

Error bars represent standard deviations.

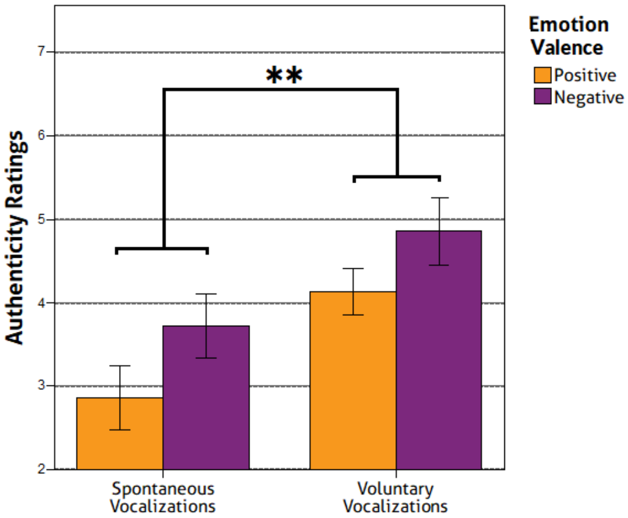
Fig. 2 Bar graph illustrating significant differences in mean reaction time (ms) between spontaneous and voluntary vocalizations.

Table 4. Mean reaction time (ms) per emotional experimental condition.

		Reaction Time (ms) ***(Mean ± SD)***
Spontaneous Laughter		718.099 ± 42.346
	Male	718.468 ± 17.858
	Female	717.816 ± 8.299
Voluntary Laughter		773. 418 ± 159.003
	Male	762.511 ± 51.920
	Female	781.809 ± 44.685
Spontaneous Crying		736.996 ± 93.626
	Male	704.127 ± 31.185
	Female	762.281 ± 23.507
Voluntary Crying		837.719 ± 111.015
	Male	885.017 ± 28.976
	Female	801.336 ± 31.621

1.2 Authenticity Ratings

A significant main effect of authenticity on authenticity ratings was found, showing a difference between spontaneous and voluntary vocalizations [*F* (1, 21) = 34.390, $p < .001$, $\eta_p^2 = .621$]: spontaneous vocalizations were reported as more authentic (*M*= 3.310, *SD* =.167) than voluntary vocalizations (*M* = 4.514, *SD* =.131) (Figure 3, Table 5). No significant interactions effects were found between the authenticity and emotion valence factor ($p = .396$). No significant main effect of sex ($p = .292$) or interactions involving the sex factor were found (authenticity*sex: $p = .955$; emotion valence*sex: $p = .393$; authenticity*emotion valence*sex: $p = .226$).



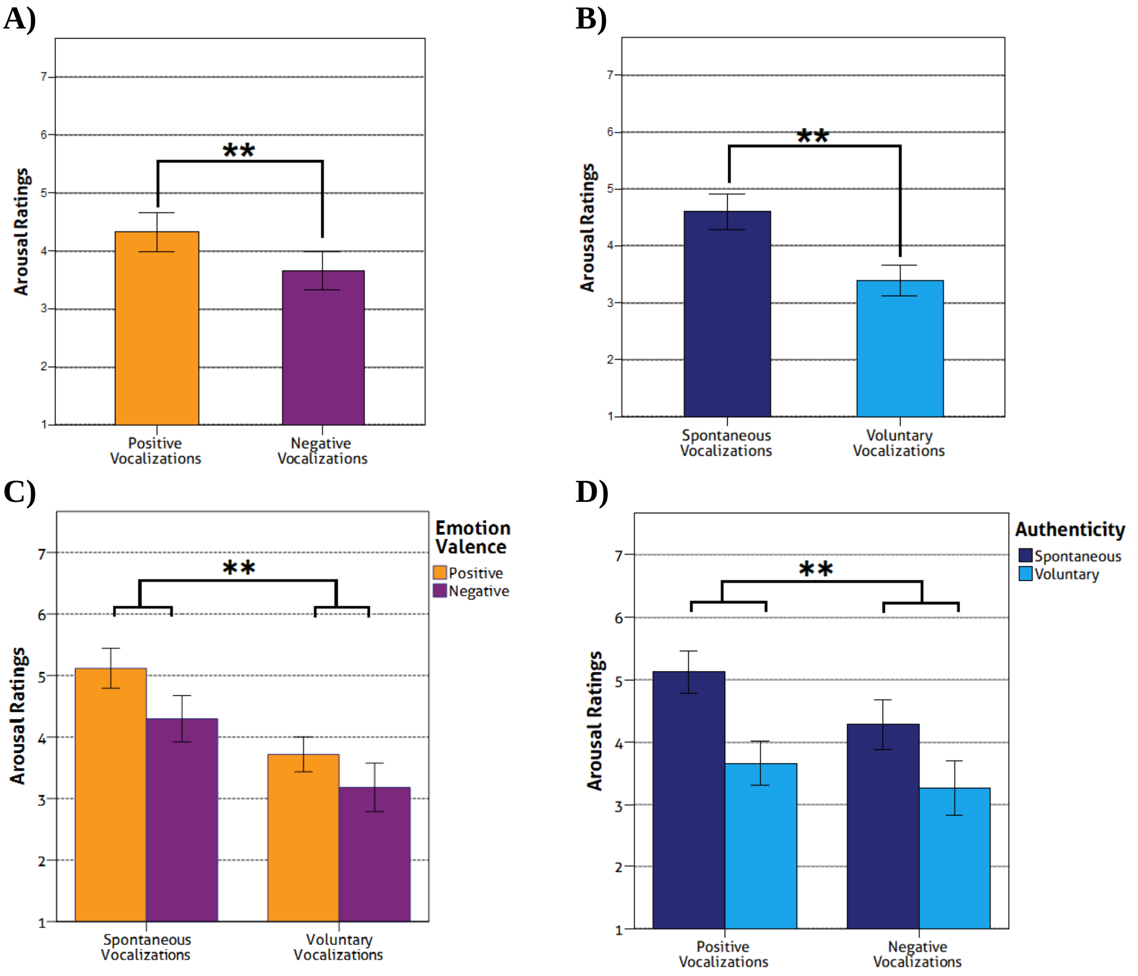
**. Effect is significant at the 0.01 level.

Error bars represent standard deviations.

Fig. 3 Bar graph illustrating significant differences in authenticity ratings between spontaneous and voluntary vocalizations, on a Likert scale from 1 "Genuine" to 7 "Posed".

1.3 Arousal Ratings

A significant main effect of emotion valence was found [*F* (1, 17) = 14.886, *p* = .001, η_p^2 = .467] on arousal ratings: positive vocalizations (*M* = 4.309, *SD* = .170) were perceived as more arousing than negative vocalizations (*M* = 3.705, *SD* = .203) (Figure 4A). A significant main effect of authenticity was found [*F* (1, 17) = 121.783, *p* < .001, η_p^2 = .878]: spontaneous vocalizations (*M*= 4.594, *SD* =.189) were reported as more arousing than voluntary vocalizations (*M*= 3.421, *SD* =.166) (Figure 4B, Table 5). No significant interaction effects were found between authenticity and emotion valence (*p* = .200) (Figure 4C, 4D). No significant main effect of sex (*p* = .473) or interactions involving the sex factor were found (emotion valence*sex: *p* = .121; authenticity*emotion valence*sex: *p* = .166) except between authenticity and sex [*F* (1, 17) = 6.726, *p* = .019, η_p^2 = .283], follow-up comparisons were not significant.



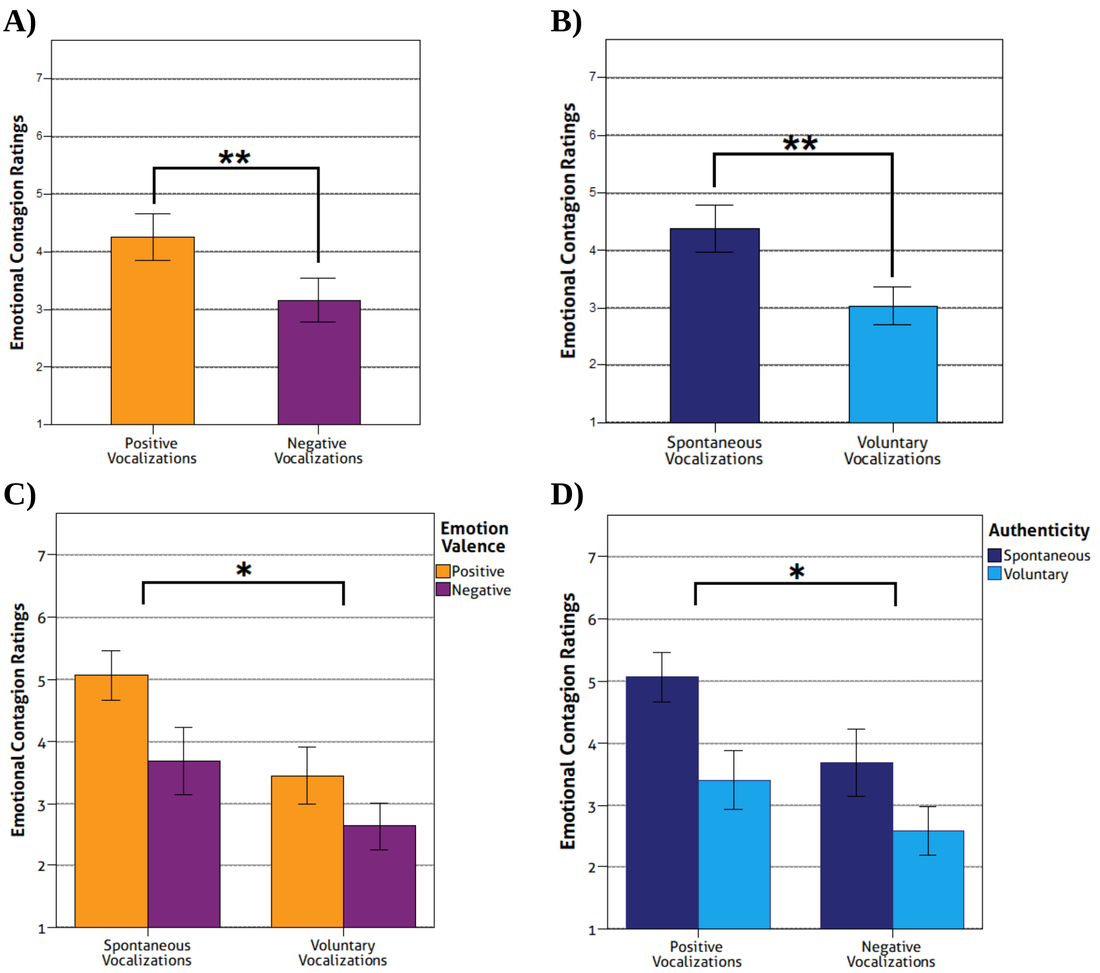
**. Effect is significant at the 0.01 level.

Error bars represent standard deviations.

Fig. 4 A) Bar graph illustrating significant differences arousal ratings between positive and negative vocalizations; B) Bar graph illustrating significant differences in arousal ratings between spontaneous and voluntary vocalizations; C) Bar graph illustrating significant differences in arousal ratings between spontaneous and voluntary vocalizations according to its emotion valence; D) Bar graph illustrating significant differences in arousal ratings between positive and negative vocalizations according to its authenticity.

1.4 Emotional Contagion Ratings

A significant main effect of emotion valence was found [F (1, 17) = 27.277, $p <$ .001, η_p^2 = .602] on emotional contagion ratings: positive vocalizations (M= 4.237, SD = .197) were reported as more contagious than negative vocalizations (M = 3.207, SD = .228) (Figure 5A). A significant main effect of authenticity was found [F (1, 17) = 73.235, p < .001, η_p^2 = .803]: spontaneous vocalizations (M = 4.379, SD =.228) were reported as more contagious than voluntary vocalizations (*M* = 3.065, *SD* =.176) (Figure 5B, Table 5). A significant interaction effect was found between authenticity and emotion valence [F (1, 17) = 5.523, *p* = .031, $\eta_p{}^{2\,=}$ .245]: both positive spontaneous (*M* = 5.117, *SD* = .181) and voluntary vocalizations (*M* = 3.545, *SD* = .215) were considered more contagious than negative spontaneous (*M* = 3.838, *SD* = .260) and voluntary vocalizations (*M* =2.749, *SD* = .185) (Figure 5C, 5D). No significant main effect of sex (*p* = .711) or interactions involving it were found (authenticity*sex: *p* = .464; emotion valence*sex: *p* = .094; authenticity*emotion valence*sex: *p* = .769).



**. Effect is significant at the 0.01 level.

Error bars represent standard deviations.

Fig. 5 A) Bar graph illustrating significant differences in emotional contagion ratings between positive and negative vocalizations; B) Bar graph illustrating significant differences in emotional contagion ratings between spontaneous and voluntary vocalizations; C) Bar graph illustrating significant differences in emotional contagion ratings between spontaneous and voluntary vocalizations according to its emotion valence; D) Bar graph illustrating significant differences in emotional contagion ratings between positive and negative vocalizations according to its authenticity.

Table 5. Perceived authenticity, arousal and emotional contagion of each emotional experimental condition.

		Authenticity *(Mean ± SD)*	**Arousal *(Mean ± SD)***	**Emotional Contagion *(Mean ± SD)***
		N = 23	***N = 19***	***N = 20***
Spontaneous Laughter		2.858 ± .897	5.144 ± .675	5,065 ± .889
	Male	3.141 ± .229	4.866 ± .318	4.853 ± .358
	Female	2.641 ± .272	5.345 ± .116	5.206 ± .234
Voluntary Laughter		4.141 ± .641	3.715 ± .583	3,451 ± 1.021
	Male	4.305 ± .179	3.835 ± .217	3.415 ± .415
	Female	4.017 ± .191	3.629 ± .173	3.475 ± .277
Spontaneous Crying		3.729 ± .886	4.267 ± .785	3,688 ± 1.222
	Male	3.727 ± .315	4.317 ± .345	3.932 ± .351
	Female	3.731 ± .231	4.231 ± .120	3.525 ± .396
Voluntary Crying		4.857 ± .929	3.242 ± .854	2,628 ± .872
	Male	4.9487 ± .309	3.616 ± .386	2.973± .276
	Female	4.788 ± .257	2.970 ± .161	2.398 ± .255

2. EEG Results

The N100, P200 and LPP components were elicited for each of the experimental conditions (voluntary laughter, spontaneous laughter, voluntary crying, spontaneous crying and neutral vocalizations), as illustrated by the grand average waveforms in Figure 6.

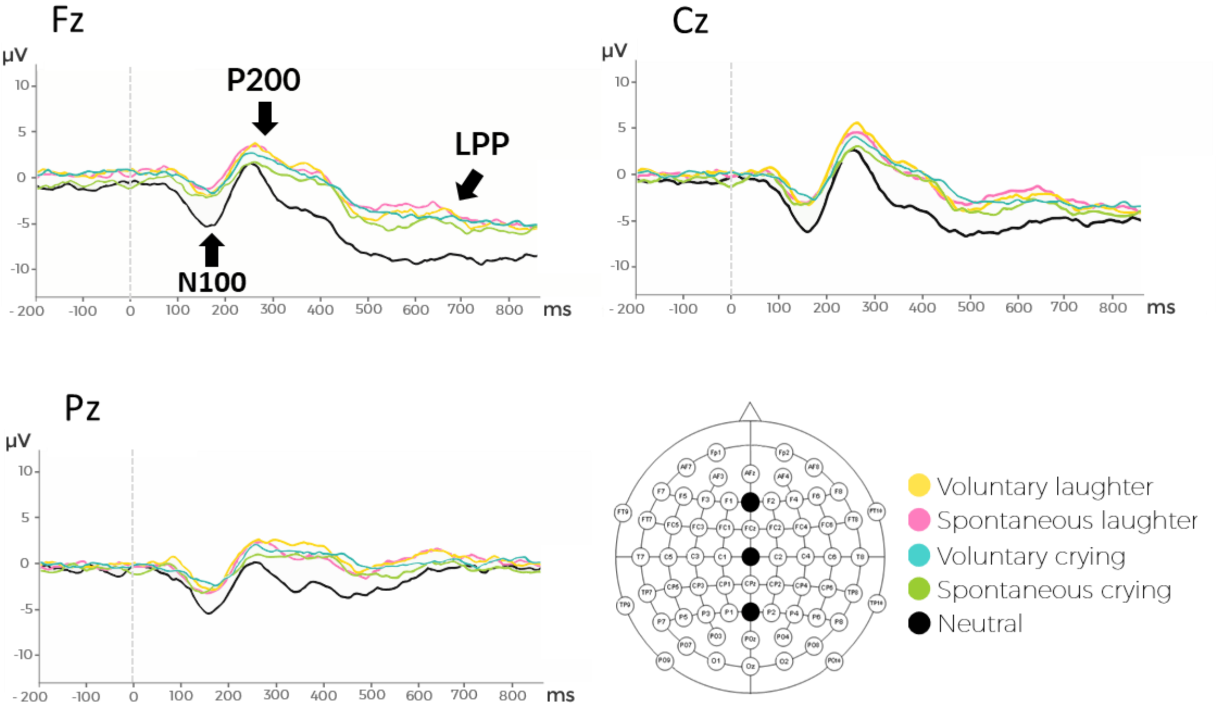


Fig. 6 Grand average waveforms for each experimental condition showing the N100, P200 and LPP components at frontal (Fz), central (Cz), parietal (Pz) midline electrodes. Positivity is plotted upwards.

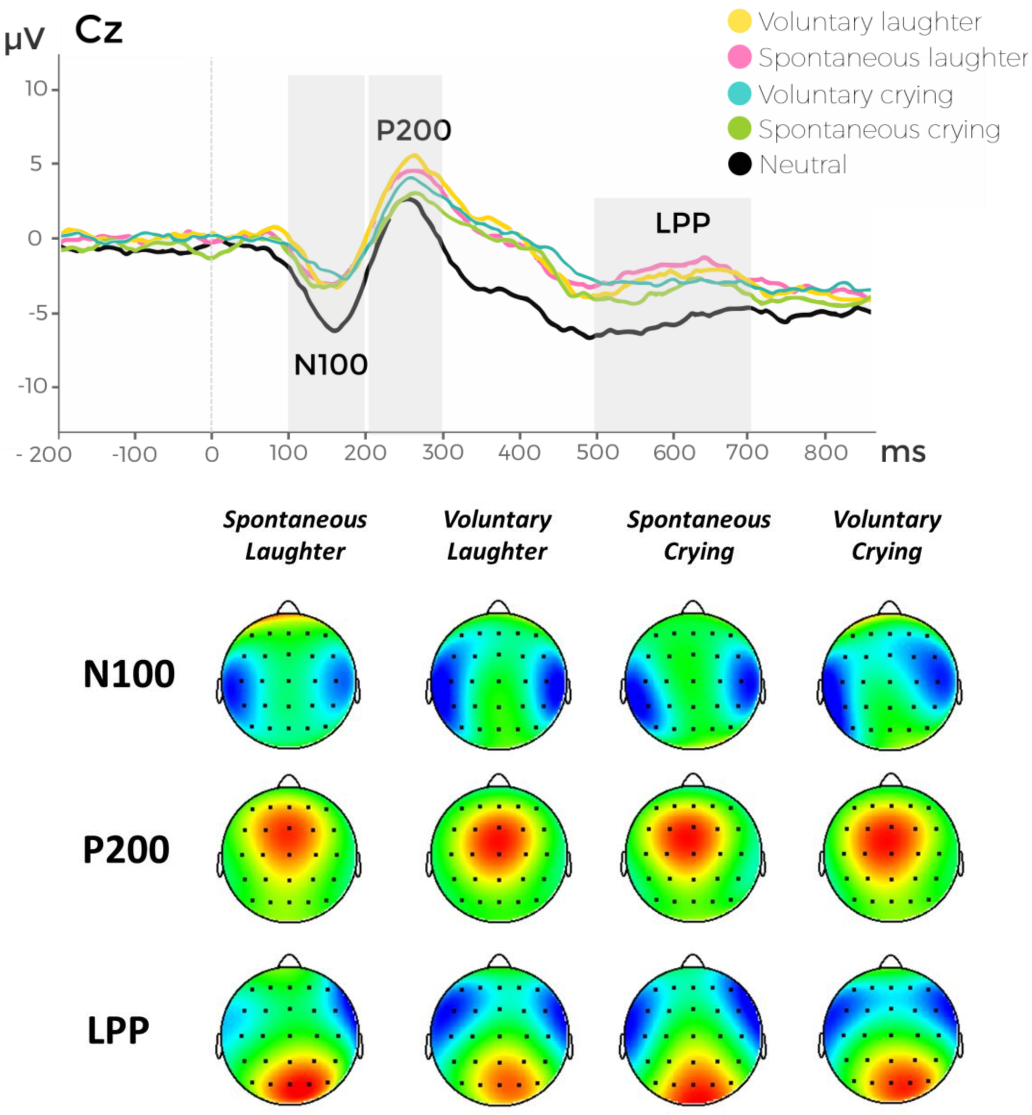


Fig. 7 Illustration of the N100 (100–2000 ms), P200 (200–300 ms) and LPP (500 – 700 ms) response topographically as a function of the experimental emotional condition.

First, the distinction between emotional (positive and negative) and neutral vocalizations was evaluated for spontaneous and voluntary vocalizations, separately. Secondly, we evaluated the effects of authenticity on emotional vocalizations by contrasting spontaneous and voluntary vocalizations.

2.1 Spontaneous emotional vs. neutral vocalizations

N100

A main effect of emotion valence was observed for N100 amplitude [*F* (2, 42) = 11.733, *p* < .001, η_p^2 = .358] on spontaneous (vs. neutral) vocalizations: N100 was significantly increased for positive (*M* = -1.718, *SD* = .500, *p* = .001) and negative spontaneous vocalizations (*M* = -2.222, *SD* = .466, *p* = .001) relative to neutral vocalizations (*M* =-4.325, *SD* = .484). No significant difference was found between positive and negative spontaneous vocalizations (*p* = 1.000) (Figure 8). Further the effect size was large (Cohen, 1988), emotion valence explaining 36% of the inter-individual variance on the N100 amplitude of spontaneous vocalizations.

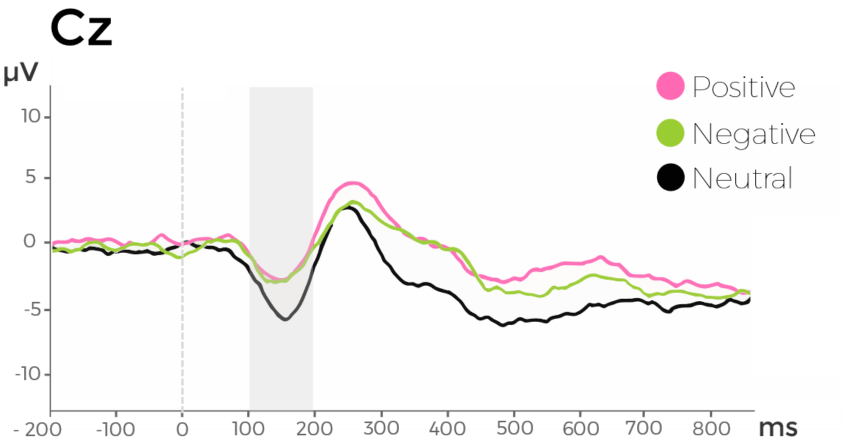


Fig. 8 Grand average waveforms at Cz illustrating main effect of emotion valence on N100 amplitude (highlighted in grey) in spontaneous (vs. neutral) vocalizations. Positivity is plotted upwards.

No significant main effect of ROI (*p* = .154) or sex were found (*p* = .058). However, a significant interaction effect between ROI and sex on N100 amplitude was found [*F* (1,088, 20) = 5.653, *p* = .024, η_p^2 = .212]: N100 amplitude was significantly more negative in frontal (*M* = -3.702, *SD* = .868, *p* = .029) and fronto-central areas (M= -3.702, *SD* = .684, *p* = .044) but not in central (*p* = .064), central-parietal (*p* = .151) and parietal areas (*p* = .601) in males, compared to females (frontal: *M* = -1.003, *SD* = .761; fronto-central: *M* = -1.759, *SD* = .600) (Table 6).

Table 6. Comparison of mean amplitude of N100 for each ROI in spontaneous vocalizations (vs. neutral) by sex.

Sex	**ROI**	**Mean ± SD**
Male	Frontal	-3.702 ± .868
	Fronto-central	-3.705 ± .684
	Central	-3.794 ± .544
	Central-parietal	-3.406 ± .411
	Parietal	-2.715 ± .315
Female	Frontal	-1.003 ± .761
	Fronto-central	-1.759 ± .600
	Central	-2.380 ± .477
	Central-parietal	-2.590 ± .361
	Parietal	-2.493 ± .276

No significant main effect of emotion valence (p = .163) or interactions involving this factor were found for N100 latency in spontaneous vocalizations (emotion valence*ROI: p = .317; emotion valence*sex: p = .143). The previous models were repeated with the BSI score as a co-variate and no significant results were found (p > .05).

P200

A main effect of emotion valence was found for P200 amplitude [F (2, 42) = 6.106, p =.005, η_p^2 = .225] on spontaneous (vs. neutral) vocalizations: P200 was increased for positive spontaneous vocalizations (M =1.821, SD = .622, p = .009), relative to neutral vocalizations (M = -.565, SD = .611). No significant difference was found between positive and negative spontaneous vocalizations (M = .449, SD = .473, p = .195), as well as between neutral (M = -.565, SD = .611) and negative spontaneous vocalizations (M = .449, SD = .473, p = .381) (Figure 9). Further the effect size was large (Cohen, 1988), emotion valence explaining 22% of the inter-individual variance on the P200 amplitude of spontaneous vocalizations.

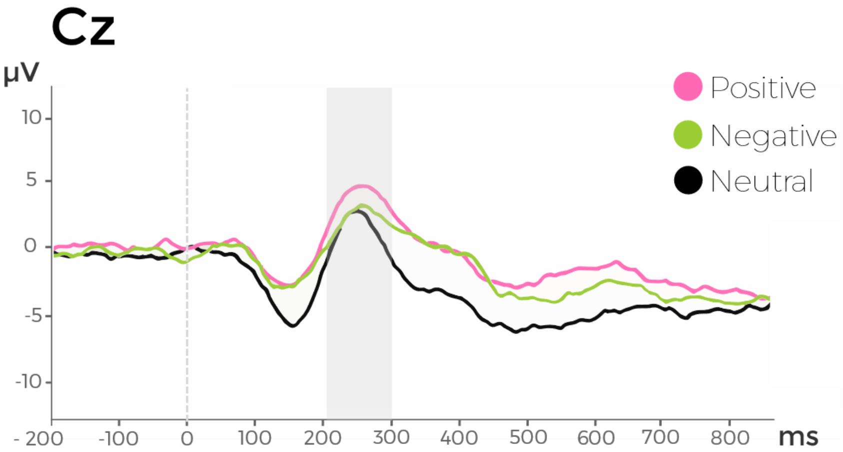


Fig. 9 Grand average waveforms at Cz illustrating main effect of emotion valence on P200 amplitude (highlighted in grey) in spontaneous (vs. neutral) vocalizations. Positivity is plotted upwards.

No significant main effect of ROI (*p* = .077) or sex were found (*p* = .063). However, a significant interaction between ROI and sex was observed [*F* (1.170, 20) = 4.516, *p* = .038, η_p^2 = .135]: P200 amplitude was increased in frontal (*M* = 1.619, *SD* = .801, *p* = .045), fronto-central (*M* = 2.194, *SD* = .702, *p* = .041) and central regions (*M* = 1.877, *SD* = .572, *p* = .038) in females, compared to males (frontal: *M* = - .963, *SD* = .913; fronto-central: *M* = -.121, *SD* = .801; central: *M* = -.045, *SD* = .652) but not in central-parietal (*p* = .140) or parietal regions (*p* = .693) (Table 7).

Table 7. Comparison of mean amplitude of P200 for each ROI in spontaneous (vs. neutral) vocalizations by sex.

Sex	**ROI**	**Mean ± SD**
Male	Frontal	-.963 ± .913
	Fronto-central	-.121 ± .801
	Central	-.045 ± .652
	Central-parietal	-.029 ± .525
	Parietal	-.073 ± .473
Female	Frontal	1.619 ± .801
	Fronto-central	2.194 ± .702
	Central	1.877 ± .572
	Central-parietal	1.041 ± .460
	Parietal	.179 ± .415

No significant main effect of emotion valence (*p* = .251) or interactions involving this factor were found for P200 latency in spontaneous vocalizations (emotion valence*ROI: *p* = .700; emotion valence*sex: *p* = .143). The previous models were repeated with the BSI score as a co-variate and no significant results were found (*p* > .05).

LPP

A main effect of emotion valence was found for LPP amplitude [*F* (2, 20) = 12.146, *p* < .001, η_p^2 = .366] on spontaneous (vs. neutral) vocalizations: LPP amplitude was increased for positive (*M* = -1.990, *SD* =.611, p = .001) and negative spontaneous vocalizations (*M* = -3.305, *SD* =.650, p = .019), compared to neutral vocalizations (*M* =-5.516, *SD* =.761). No significant difference was found between positive and negative spontaneous vocalizations (*p* = .159) (Figure 10). Further the effect size was large (Cohen, 1988), emotion valence explaining 37% of the inter-individual variance on the LPP amplitude of spontaneous vocalizations.

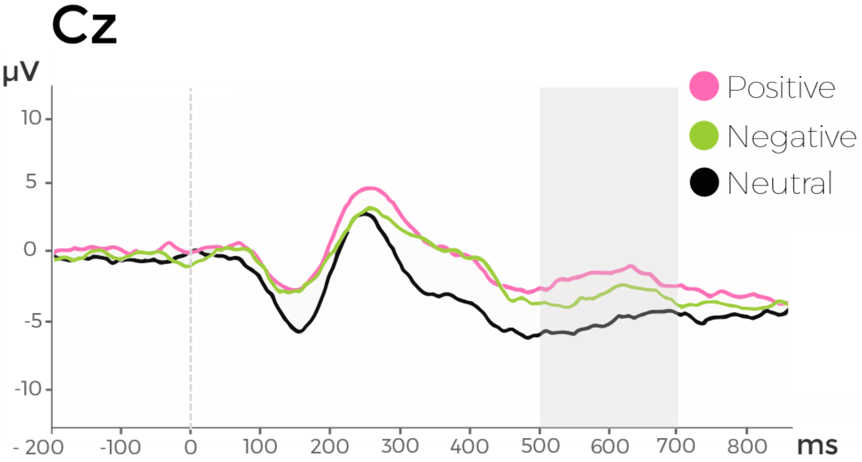


Fig. 10 Grand average waveforms at Cz illustrating main effect of emotion valence on LPP amplitude (highlighted in grey) in spontaneous (vs. neutral) vocalizations. Positivity is plotted upwards.

A main effect of ROI was observed for LPP amplitude [*F* (1.218, 25.579) = 73.486, *p* < .001, η_p^2 = .778]: LPP amplitude was significantly different in frontal, fronto-central, central, central-parietal and parietal areas. Increased amplitude was found in parietal areas (*M* = -.587, *SD* = .492), followed by central-parietal (*M* = -2.194, *SD* = .502), central (*M* = -3.943, *SD* = .558), fronto-central (*M* = -2.194, *SD* = .502) and frontal areas (*M* = -6.026, *SD* = .694), respectively (Table 8).

Table 8. Mean amplitude of LPP for each ROI in spontaneous (vs. neutral) vocalizations.

ROI	**Mean ± SD**
Frontal	-6.026 ± .694
Fronto-central	-5.268 ± .637
Central	-3.943 ± .558
Central-parietal	-2.194 ± .502
Parietal	-.587 ± .492

A significant interaction effect between emotion valence and ROI was found [F (2.410, 50.610) = 13.803, $p < .001$, $\eta_p{}^2 = .397$]: LPP amplitude was increased significantly upon hearing positive spontaneous vocalizations in frontal ($p < .001$), fronto-central ($p < .001$), central ($p = .001$), central-parietal ($p = .006$) and parietal areas (p = .032), compared to neutral vocalizations. An increased amplitude was also found for negative spontaneous vocalizations when compared to neutral vocalizations in frontal ($p = .001$), fronto-central ($p = .005$), central ($p = .027$) areas, with no significant differences in central-parietal ($p = .195$) and parietal areas ($p = .592$) (Table 9). No significant differences were found between positive and negative spontaneous vocalizations for the LPP amplitude according to ROI (frontal: $p = .239$; fronto-central: $p = .129$; central: $p =.139$; central-parietal: $p = .163$; parietal: $p = .221$)

Table 9. Comparison of mean LPP amplitude for each ROI in spontaneous (vs. neutral) vocalizations by emotion valence.

ROI	**Emotion**	**Mean ± SD**
Frontal	Positive	-3.600 ± .921
	Negative	-5.312 ± .905
	Neutral	-9.168 ± .860
Fronto-central	Positive	-3.408 ± .762
	Negative	-4.806 ± .763
	Neutral	-7.590 ± .810
Central	Positive	-2.363 ± .628
	Negative	-3.703 ± .643
	Neutral	-5.762 ± .799
Central-parietal	Positive	-.947 ± .512
	Negative	-2.119 ± .590
	Neutral	-3.517 ± .775
Parietal	Positive	.369 ± .454
	Negative	-.587 ± .563
	Neutral	-1.542 ± .786

No significant main effect of sex (*p* = .599) or interactions involving this factor were found (emotion valence*sex: *p* = .573; ROI*sex: p = .474).

No significant main effect of emotion valence (*p* = .387) or interactions involving it were found for LPP latency in spontaneous vocalizations (emotion valence*ROI: *p* = .164; emotion valence*sex: *p* = .294). The previous models were repeated with the BSI score as a co-variate and no significant results were found (*p* > .05).

2.2 Voluntary emotional vs. neutral vocalizations

N100

A main effect of emotion valence was observed for N100 amplitude [*F* (2, 42) = 8.803, *p* = .001, η_p^2 = .295] on voluntary (vs. neutral) vocalizations: N100 was increased in response to positive (*M* = -1.668, *SD* = .628, *p* = .008) and negative voluntary vocalizations (*M* = -1.370, *SD* = .671, *p* = .003) relative to neutral vocalizations (*M* = -4.325, *SD* = .484), with no significant difference between positive and negative vocalizations (*p* = 1.000) (Figure 11). Further the effect size was large (Cohen, 1988), emotion valence explaining 30% of the inter-individual variance on the N100 amplitude of voluntary vocalizations.

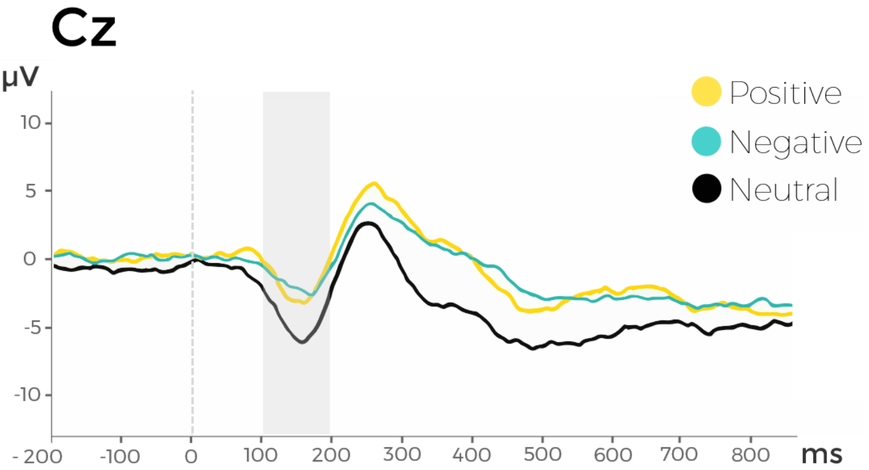


Fig. 11 Grand average waveforms at Cz illustrating main effect of emotion valence on N100 amplitude (highlighted in grey) in voluntary (vs. neutral) vocalizations. Positivity is plotted upwards.

No significant main effect of ROI (*p* = .763) and sex (*p* = .467) or interactions involving these factors were found (emotion valence*ROI: *p* = .317; emotion valence*sex: *p* = .877; ROI*sex: *p* = .180).

No significant main effect of emotion valence (*p* = .662) or interactions involving this factor were found for N100 latency in voluntary vocalizations (emotion valence*ROI: *p* = .094; emotion valence*sex: *p* = .362). The previous models were repeated with the BSI score as a co-variate and no significant results were found (*p* > .05).

P200

A main effect of emotion valence was found for P200 amplitude [*F* (2, 42) = 5.198, *p* =.010, η_p^2 = .198] on voluntary (vs. neutral) vocalizations: P200 was increased for positive voluntary vocalizations (*M* =1,971, *SD* =.637, *p* = .020) compared to neutral vocalizations (*M* = -.565, *SD* = .611) (Figure 12). No significant difference was found between positive and negative voluntary vocalizations (*M* = 1.227, *SD* = .692, *p* = 1.000) or neutral and negative voluntary vocalizations (*p* = .111). Further the effect size was large (Cohen, 1988), emotion valence explaining 20% of the inter-individual variance on the P200 amplitude of voluntary vocalizations.

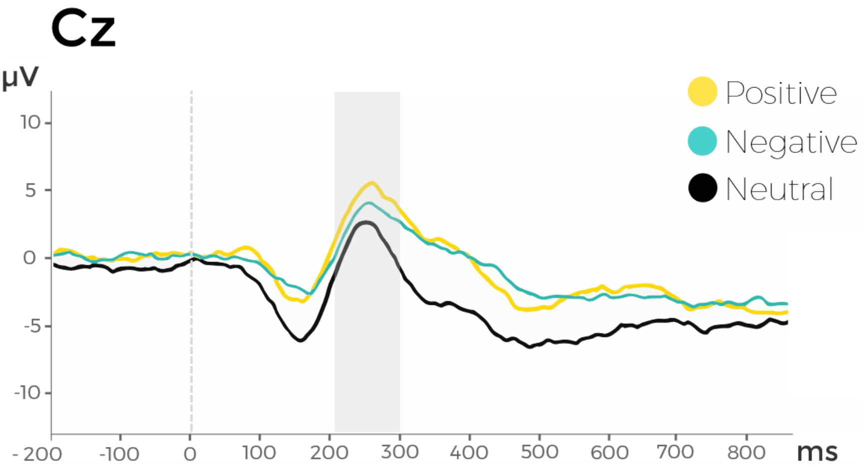


Fig. 12 Grand average waveforms at Cz illustrating main effect of emotion valence on P200 amplitude (highlighted in grey) in voluntary (vs. neutral) vocalizations. Positivity is plotted upwards.

A significant main effect of ROI was found [*F* (1.299, 27.272) = 5.791, *p* <.001, η_p^2 = .216]: P200 amplitude was significantly increased in central areas compared to central-parietal (*p* = .043) and parietal areas (*p* = .010), and in fronto-central areas compared to frontal areas (*p* = .001). No significant difference was found between frontal and central-parietal (*p* = 1.000) and parietal areas (*p* = 1.000), as well as between fronto-central and central areas (*p* = .074). No significant interaction effects involving the ROI factor were found (emotion valence*ROI: *p* = .828) (Table 10).

Table 10. Mean amplitude of P200 for each ROI in voluntary (vs. neutral) vocalizations.

ROI	Mean ± SD
Frontal	.606 ± .602
Fronto-central	1.342 ± .552
Central	1.292 ± .463
Central-parietal	.858 ± .397
Parietal	.289 ± .328

No significant main effect of sex (p = .073) or interactions involving this factor were found (emotion valence*sex: p = .648; ROI*sex: p = .433).

No significant main effect of emotion valence (p = .213) or interactions involving it were found for P200 latency in voluntary vocalizations (emotion valence*ROI: p = .394). The previous models were repeated with the BSI score as a co-variate and no significant results were found (p > .05).

LPP

A main effect of emotion valence was found for LPP amplitude [F (2, 42) = 8.426, p =.001, η_p^2 = .286] on voluntary (vs. neutral) vocalizations: LPP amplitude was increased for positive (M = -2.338, SD =.577, p = .007) and negative voluntary vocalizations (M = -2.506, SD =.697, p = .010) compared to neutral vocalizations (M = -5.516, SD =.761) (Figure 13). No significant difference was found between positive and negative voluntary vocalizations (p = 1.000). Further the effect size was large (Cohen, 1988), emotion valence explaining 29% of the inter-individual variance on the LPP amplitude of voluntary vocalizations.

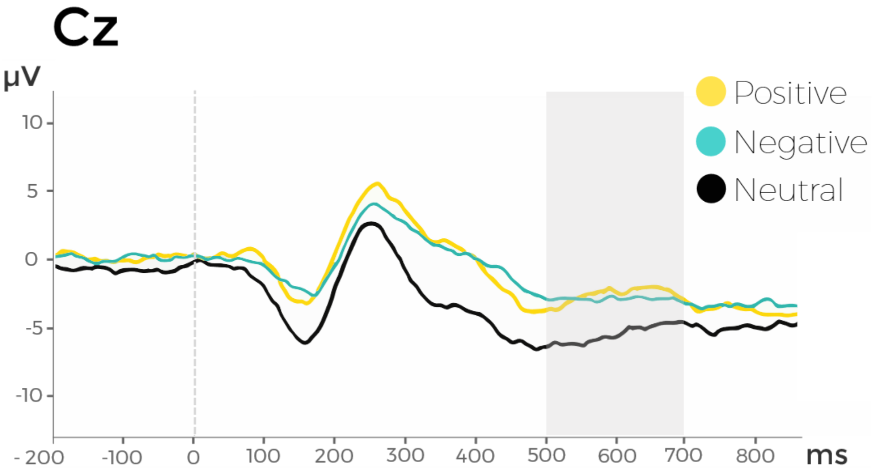


Fig. 13 Grand average waveforms at Cz illustrating main effect of emotion valence on LPP amplitude (highlighted in grey) in voluntary (vs. neutral) vocalizations. Positivity is plotted upwards.

A main effect of ROI was observed for LPP amplitude [F (1.319, 27.702) = 117.479, p <.001, η_p^2 = .848]: LPP amplitude was significantly different in frontal, fronto-central, central, central-parietal and parietal areas. Parietal areas presented increased positivity (M = -.335, SD = .448), followed by central-parietal (M = -2.021, SD = .457), central (M = -3.870, SD = .503), fronto-central (M = -5.222, SD = .542) and frontal areas (M = -5.818, SD = .536) (Table 11).

Table 11. Mean amplitude of LPP for each ROI in voluntary (vs. neutral) vocalizations.

ROI	**Mean ± SD**
Frontal	-5.818 ± .536
Fronto-central	-5.222 ± .542
Central	-3.870 ± .503
Central-parietal	-2.021 ± .457
Parietal	-.335 ± .448

A significant interaction effect between emotion valence and ROI was found [F (2.190, 45.988) = 9.875, p < .001, η_p^2 =.320]: LPP amplitude was increased significantly on frontal (positive: p = .002; negative: p = .001), fronto-central (positive: p = .005; negative: p = .004) and central areas (positive: p = .015; negative: p = .015) upon hearing positive and negative voluntary vocalizations, compared to neutral vocalizations. In central-parietal a significant difference was only found in positive voluntary vocalizations compared to neutral (*p* = .035), with no difference between negative and neutral vocalizations (*p* = .067) (Table 12). No significant difference was found for LPP amplitude between positive, negative and neutral vocalizations in parietal areas (positive and negative: *p* = 1.000; positive and neutral: *p* = .093; negative and neutral: *p* = .194).

Table 12. Comparison of mean LPP amplitude for each ROI in voluntary (vs. neutral) vocalizations by emotion valence.

ROI	**Emotion**	**Mean ± SD**
Frontal	Positive	-4.157 ± .817
	Negative	-4.129 ± .890
	Neutral	-9.168 ± .860
Fronto-central	Positive	-4.013 ± .699
	Negative	-4.062 ± .765
	Neutral	-7.590 ± .810
Central	Positive	-2.911 ± .603
	Negative	-2.938 ± .701
	Neutral	-5.762 ± .799
Central-parietal	Positive	-1.100 ± .516
	Negative	-1.448 ± .651
	Neutral	-3.517 ± .775
Parietal	Positive	.492 ± .513
	Negative	.045 ± .610
	Neutral	-1.542 ± .786

No significant main effect of sex (*p* = .298) or interactions involving the sex factor were found (emotion valence*sex: *p* = .895; ROI*sex: *p* = .482). No significant main effect of emotion valence (*p* = .107) or interactions involving it were found for LPP latency in voluntary vocalizations (emotion valence*ROI: *p* = .769; emotion valence*sex: *p* = .373). We repeated the previous models with the BSI score as a co-variate and no significant results were found (p > .05).

2.3 Spontaneous vs. voluntary emotional vocalizations

N100

No main effect of authenticity was found for N100 amplitude (p = .259, η_p^2 = .060) or latency (p = .157, η_p^2 = .093) (Figure 14). No significant main effect of the hemisphere factor was found (p = .283). However, an interaction effect was found between authenticity and hemisphere [F (1, 21) = 4.928, p = .038, η_p^2 =.190], follow-up comparisons were not significant. No significant interaction effects were found between authenticity and sex on N100 mean amplitude (p = .398, η_p^2 =.034) or latency (p = .563, η_p^2 =.016). The previous models were repeated with the BSI score as a co-variate and no significant results were found (p > .05).

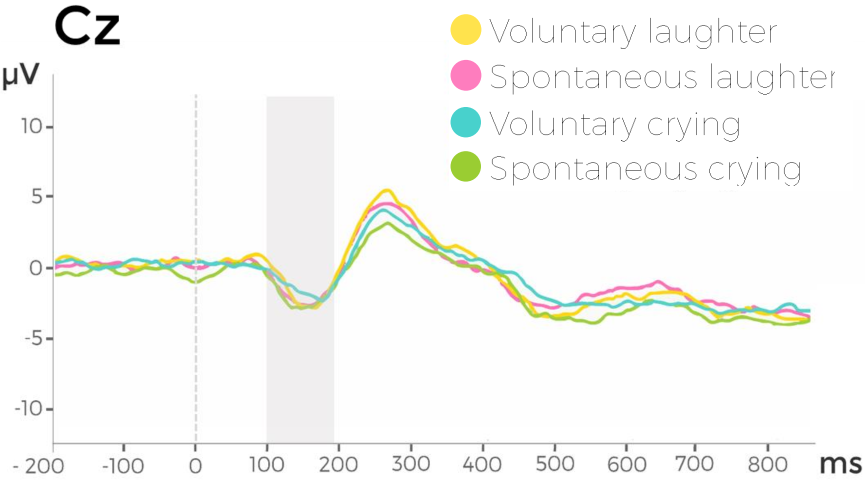


Fig. 14 Grand average waveforms at Cz illustrating N100 component (highlighted in grey) in spontaneous vs. voluntary vocalizations. Positivity is plotted upwards.

P200

No main effect of authenticity was found for P200 amplitude (p = .267, η_p^2 = .058) or latency (p = .390, η_p^2 = .035) (Figure 15). A main effect of the hemisphere factor was found [F (1) = 16.075, p = .001, η_p^2 =.434]: amplitude was increased in the left hemisphere (M = 1.266, SD =.435) compared to the right hemisphere (M = .624, SD =.419). No significant interaction effects involving the hemisphere factor were found (authenticity*hemisphere: p = .096). No significant interaction effects were found between authenticity and sex on P200 mean amplitude (p = .903, η_p^2 = .001) or latency (p = .611, η_p^2 = .013). The previous models were repeated with the BSI score as a co-variate and no significant results were found (p > .05).

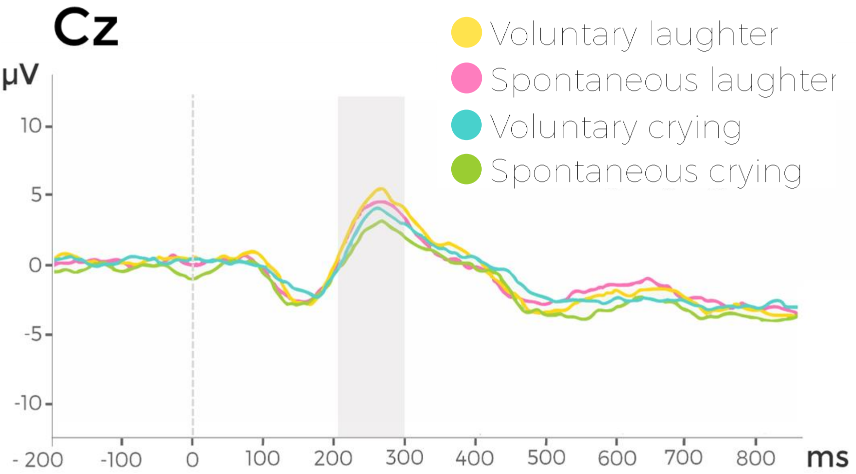


Fig. 15 Grand average waveforms at Cz illustrating P200 component (highlighted in grey) in spontaneous vs. voluntary vocalizations. Positivity is plotted upwards.

LPP

No significant main effect or interactions involving authenticity were found for LPP mean amplitude (p = .606, η_p^2 = .013) or latency (p = .649, η_p^2 = .010) (Figure 16). No significant main effect or interactions involving the hemisphere factor were found (authenticity*hemisphere: p = .340). No significant interaction effects were found between authenticity and sex on LPP mean amplitude (p = .477, η_p^2 = .024) or latency (p = .144, η_p^2 = .099). The previous models were repeated with the BSI score as a co-variate and no significant results were found (p > .05).

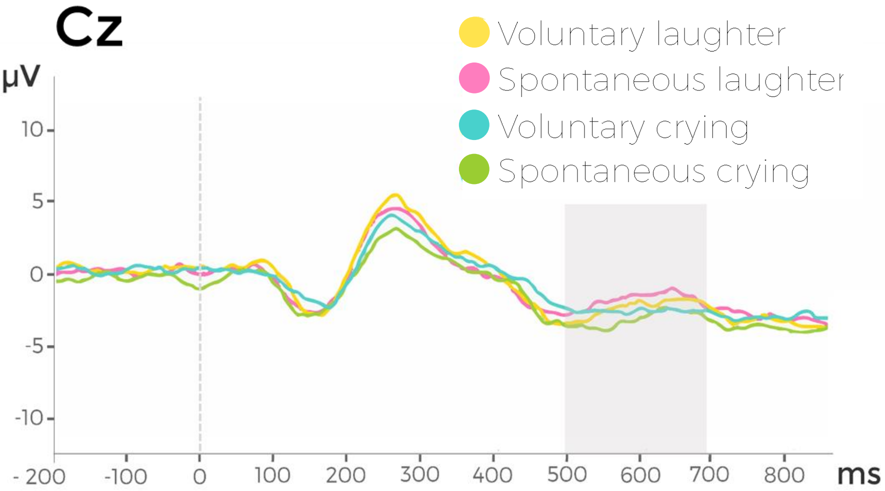


Fig. 16 Grand average waveforms at Cz illustrating LPP component (highlighted in grey) in spontaneous vs. voluntary vocalizations. Positivity is plotted upwards.

Regarding the analysis only using spontaneous and voluntary laughter vocalizations, no significant differences were found for the authenticity factor on the mean amplitude of the N100 ($p = .926$, $\eta_p^2 < .001$), P200 ($p = .814$, $\eta_p^2 = .003$) and LPP component ($p = .555$, $\eta_p^2 = .017$).

3. Correlations

No significant correlations were found between the N100, P200 and LPP mean amplitude and behavioural measures (authenticity ratings, arousal ratings, emotional contagion ratings, empathy score (EQ) or mentalizing score (RMET)) (p > .05) (Table 13, 14).

Table 13. Association between authenticity, arousal and emotional contagion ratings and N100, P200 and LPP mean amplitude.

		N100 Mean Amplitude	**P200 Mean Amplitude**	**LPP Mean Amplitude**
Authenticity Ratings	*Spontaneous Laughter*	*r* = .211, *p* = .334	*r* = .173, *p* = .430	*r* = .222, *p* = .308
	Voluntary Laughter	*r* = - .065, *p* = .768	*r* = - .350, *p* = .102	*r* = - .179, *p* = .413
	Spontaneous Crying	*r* = - .239, *p* = .272	*r* = - .179, *p* = .414	*r* = - .184, *p* = .400
	Voluntary Crying	*r* = - .038, *p* = .862	*r* = - .054, *p* = .808	*r* = .330, *p* = .124
Arousal Ratings	*Spontaneous Laughter*	*r* = .029, *p* = .934	*r* = .181, *p* = .458	*r* = .145, *p* = .553
	Voluntary Laughter	*r* = - .318, *p* = .185	*r* = - .333, *p* = .163	*r* = - .167, *p* = .495
	Spontaneous Crying	*r* = .213, *p* = .381	*r* = .057, *p* = .818	*r* = - .090, *p* = .713
	Voluntary Crying	*r* = - .230, *p* = .343	*r* = - .109, *p* = .656	*r* = .027, *p* = .914
Emotional Contagion Ratings	*Spontaneous Laughter*	*r* = - .217, *p* = .358	*r* = .027, *p* = .911	*r* = .319, *p* = .170
	Voluntary Laughter	*r* = - .528, *p* = .017	*r* = - .226, *p* = .339	*r* = .192, *p* = .417
	Spontaneous Crying	*r* = - .134, *p* = .572	*r* = .022, *p* = .927	*r* = .233, *p* = .323
	Voluntary Crying	*r* = - .390, *p* = .089	*r* = - .303, *p* = .194	*r* = - .169, *p* = .476

Table 14. Association between Empathy Quotient Score (EQ), Reading the Mind in the Eyes Test Score (RMET) and N100, P200 and LPP mean amplitude.

		EQ Score	**RMET Score**
N100 Mean Amplitude	*Spontaneous Laughter*	*r* = - .195, *p* = .372	*r* = .249, *p* = .252
	Voluntary Laughter	*r* = - .050, *p* = .820	*r* = .038, *p* = .864
	Spontaneous Crying	*r* = .132, *p* = .548	*r* = .284, *p* = .190
	Voluntary Crying	*r* = .062, *p* = .778	*r* = .368, *p* = .084
P200 Mean Amplitude	*Spontaneous Laughter*	*r* = - .143, *p* = .515	*r* = .286, *p* = .185
	Voluntary Laughter	*r* = .115, *p* = .600	*r* = - .128, *p* = .559
	Spontaneous Crying	*r* = .225, *p* = .302	*r* = .155, *p* = .479
	Voluntary Crying	*r* = .222, *p* = .308	*r* = .458, *p* = .028
LPP Mean Amplitude	*Spontaneous Laughter*	*r* = .046, *p* = .834	*r* = - .131, *p* = .552
	Voluntary Laughter	*r* = .090, *p* = .683	*r* = - .208, *p* = .340
	Spontaneous Crying	*r* = .027, *p* = .902	*r* = .208, *p* = .340
	Voluntary Crying	*r* = .034, *p* = .877	*r* = .125, *p* = .570

A significant negative correlation was found between perceived authenticity and perceived arousal for spontaneous laughter (*r* = -.462, *p* = .046) and between perceived authenticity and perceived emotional contagion for voluntary crying (*r* = -.525, *p* = .009) (Table 15, Figure 17A and 17B). No significant correlations were found between perceived authenticity and perceived arousal for voluntary laughter (*p* = .196), spontaneous crying (*p* = .509) and voluntary crying (*p* = .516). No significant correlations were found between perceived authenticity and perceived emotional contagion for spontaneous laughter (*p* = .098), voluntary laughter (*p* = .275) and spontaneous crying (*p* = .656).

Table 15. Association between authenticity, arousal and emotional contagion ratings for spontaneous and voluntary vocalizations.

Authenticity	**Arousal**	**Emotional Contagion**
Spontaneous Laughter	**$r = -.462^{*}, p = .046$**	$r = -.381$, $p = .098$
Voluntary Laughter	$r = -.310$, $p = .196$	$r = -.257$, $p = .275$
Spontaneous Crying	$r = -.161$, $p = .509$	$r = -.106$, $p = .656$
Voluntary Crying	$r = -.159$, $p = .516$	**$r = -.571^{*}, p = .009$**

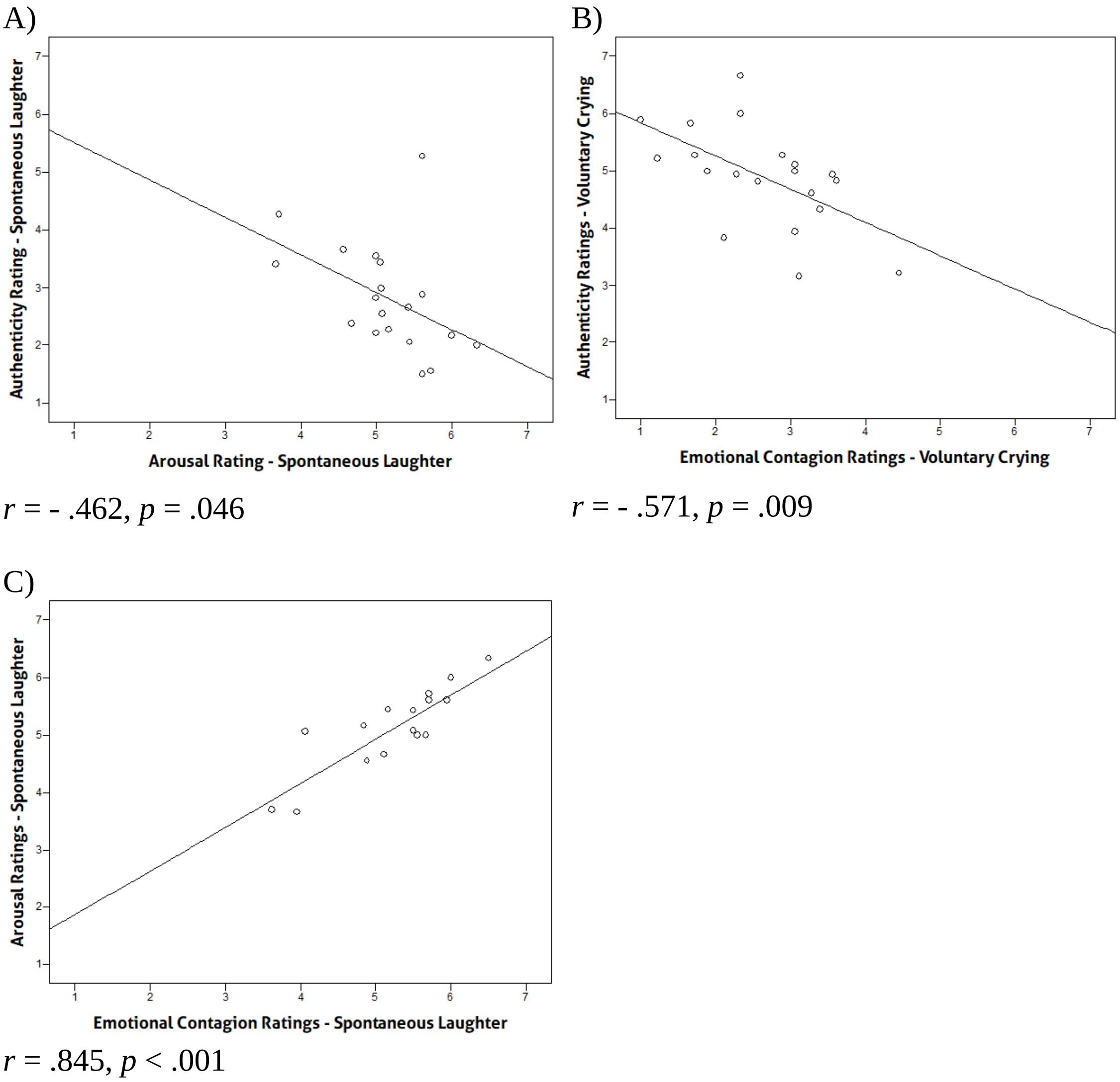
values denote Pearson's r coefficients.
**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

A significant positive correlation was found between perceived arousal and perceived emotional contagion for spontaneous laughter ($r = .845$, $p < .001$) (Table 15, Figure 17C). No significant correlations were found between perceived arousal and perceived emotional contagion for voluntary laughter ($p = .450$), spontaneous crying ($p = .431$) and voluntary crying ($p = .365$) (Table 16).

Table 16. Association between arousal and emotional contagion ratings for spontaneous and voluntary vocalizations.

Arousal	**Emotional Contagion**
Spontaneous Laughter	**$r = .845^{**}, p < .001$**
Voluntary Laughter	$r = .450$, $p = .450$
Spontaneous Crying	$r = .431$, $p = .431$
Voluntary Crying	$r = .365$, $p = .365$

values denote Pearson's r coefficients.
**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).



A) *r* = - .462, *p* = .046

B) *r* = - .571, *p* = .009

C) *r* = .845, *p* < .001

Fig. 17 Scatterplots representing A) the negative correlation between authenticity and arousal ratings of spontaneous laughter, B) the negative correlation between authenticity and emotional contagions rating of voluntary crying, and C) the positive correlation between arousal and emotional contagion ratings of spontaneous laughter.

DISCUSSION

The current study investigated how authenticity modulates the time-course of vocal affective processing. In particular, we explored how spontaneous and voluntary non-linguistic vocalizations are processed online using the ERP methodology, by focusing on the three processing stages of vocal emotional perception proposed by Schirmer and Kotz (2006): sensory processing (N100), salience detection (P200) and cognitive evaluation of the emotional significance of the voice (LPP). Our data did not reveal any amplitude or latency differences between spontaneous and voluntary vocalizations in the N100, P200 and LPP components. This finding suggests authenticity does not affect the temporal course of vocal cues processing during the first 700 ms after vocalization onset. Conversely and replicating previous findings emotional vocalizations were robustly differentiated from neutral vocalizations in terms of amplitude as early as 100 ms (N100) after listeners were exposed to the vocalization. While in the N100 and LPP components both happy and sad vocalizations elicited increased amplitudes when compared to neutral vocalizations, in the P200 component only happy vocalizations were robustly enhanced compared to neutral vocalizations. No latency differences between emotional and neutral vocalizations could be detected in these components at any phase of vocal processing. Lastly, no sex differences were found in the amplitude or latency of N100, P200 and LPP for emotionality or authenticity effects.

1. The effects of emotion in voice processing

The first differentiation between emotional and neutral vocal cues was already visible in the first stage of sensory analysis, the N100 component. Although studies have repeatedly linked emotionality effects to the P200 component (Garrido-Vázquez et al, 2013; Paulmann & Kotz, 2008; Paulmann et al., 2010; Schirmer et al., 2013), some notable exceptions report sensitivity to emotional vocalizations at this early stage of auditory processing (N100) (Liu et al, 2012; Iredale et al, 2013; Pinheiro et al, 2012; Wang et al, 2015). In light of studies reporting an early emotional salience detection in the N100, the current findings reveal an enhanced N100 in response to neutral as compared to positive and negative vocalizations. This finding reveals a rapid assessment of emotionally relevant cues as soon as 100 ms after stimulus onset, with no distinction between positive (laughter) and negative (crying) vocalizations, suggesting they were perceived to have similar emotional saliency.

The P200 has been broadly described as a salience detection marker with a frontocentral distribution (Sauter & Eimer, 2010; Paulmann et al, 2013), with a majority of the studies indicating it indexes a first distinction between emotional and neutral vocal cues (Garrido-Vázquez et al, 2013; Paulmann & Kotz, 2008; Paulmann et al, 2010; Schirmer et al, 2013; Schirmer & Kotz, 2006). Evidence with emotional prosodic stimuli is not consistent in terms of the direction of the P200 amplitude, though most of the studies showed a more pronounced P200 for emotional compared to neutral speech (Iredale et al, 2013; Paulmann et al, 2010; Pinheiro et al 2012; 2014; Schirme et al, 2013), whereas some reports presented an enhanced amplitude for neutral compared to emotional speech instead (Garrido-Vázquez et al, 2013; Paulmann & Kotz, 2008). However, when comparing emotional and neutral non-linguistic affective vocalizations a study by Liu and colleagues (2012) showed an enhancement of the P200 for emotional vocalizations irrespective of their valence, suggesting a rapid deployment of our attentional resources towards emotional nonverbal cues. Contrary to previous findings, emotionality effects were not as consistent in the P200 in the current study. We observed more pronounced P200 amplitudes for positive vocalizations (laughter) than for neutral vocalizations in both vocalizations evoked spontaneously and voluntarily, indicating no processing differences between negative and neutral stimuli or negative and positive stimuli. A similar pattern of results was obtained in studies using MMN (an ERP component that peaks 100 to 250 ms after the onset of a deviant stimulus) and P300 (a component that peaks around 300 ms after stimulus onset) experiment paradigms (Pinheiro, Barros, Dias & Kotz, 2017a; Pinheiro, Barros, Vasconcelos, Obermeier & Kotz, 2017b). MMN and P300 paradigm studies, which occur close to the P200 timeline, showed a positivity bias with positive vocalizations (laughter) presenting an enhanced P3b and MMN amplitude as compared to both negative (growls) and neutral vocalizations (Pinheiro et al, 2017a; 2017b). The finding that laughter is associated with facilitated deviance detection and enhanced attention relates well with accounts establishing a preferential processing of emotionally salient (e.g., more arousing) events (Jessen & Kotz, 2011; Paulmann et al, 2013; Pell et al, 2015), given that laughter was considered by listeners as more arousing than crying stimuli.

The LPP indexes a stage of cognitive appraisal of an event, with a centro-parietal scalp distribution (Kotz & Paulmann, 2012; Schirmer & Kotz, 2006). The LPP has been demonstrated to be sensitive to the emotional content of visual (Dennis & Hajcak, 2009; Brown et al, 2012) and auditory stimuli (Jessen & Kotz, 2011; Paulmann et al., 2013; Pell et al, 2015; Schirmer et al., 2013). The magnitude of the LPP is enhanced by both positively- and negatively-valenced stimuli compared to neutral stimuli, as amplitude is more positive for more arousing stimuli (Paulmann et al, 2013). Here, the LPP amplitude was strongly influenced by emotionality, similar to the N100 component: positive (laughter) and negative (crying) vocalizations exhibited a sustained and more positive wave than neutral vocalizations. These findings corroborate the role of LPP in more elaborative processing, being enhanced for more salient cues (visual and auditory) to allow a more sustained cognitive processing of emotional vocal cues and to promote an adaptive behavioural response (Pell et al, 2015; Otten et al, 2017). In what concerns latency, contrary to findings supporting emotional vocalizations being associated with earlier ERP responses (Liu et al, 2012; Pinheiro et al, 2012), reflected in reduced latency as compared to neutral vocalizations, we did not find emotional vocalizations to be processed in a faster manner compared to neutral vocalizations in the N100, P200 and LPP.

2. The effects of authenticity in voice processing

Building on previous work behavioural and neuroimaging evidence indicated spontaneous and voluntary vocalizations are perceived distinctively (McGettigan et al, 2015; Lavan et al, 2014; 2016; 2017). Behavioural evidence indicates listeners perceive spontaneous and voluntary vocalizations differently based on its acoustic features, accurately detecting spontaneous laughter faster (Bryant & Aktipis, 2014; Lavan et al, 2016). Neuroimaging studies confirm the differential processing of spontaneous and voluntary vocalizations in the brain, indicating that areas such as the anterior medial pre-frontal cortex (aMPFC) are activated only when listening to voluntary laughter, reflecting higher order processing for the resolution of social ambiguity (McGettigan et al 2015; Lavan et al, 2017). Event-related potential evidence shows that the degree to which an event is motivationally salient causes a shift in our attentional resources and promotes preferential processing of the salient event, reflected in an enhanced P200 and LPP amplitude (Jessen & Kotz, 2011; Paulmann et al, 2013; Pell et al, 2015; Otten et al, 2017). We hypothesized that authenticity would affect the P200 and LPP components amplitude and latency, with voluntary vocalizations being perceived as more motivationally salient and as such presenting an increased positive amplitude and being processed in a faster manner. Contrary to our hypothesis, we found that authenticity does not modulate how vocal emotional cues are processed in the first 700 ms after stimulus onset. Specifically, no significant differences were found in terms of amplitude or latency between spontaneous and voluntary non-linguistic affective vocalizations in the three processing stages: N100, P200 and LPP components.

Additionally, a separate analysis was conducted only comparing spontaneous and voluntary laughter, as previous studies with authenticity have focused on this positive emotion (Lavan et al, 2014; 2017; 2018; McGettigan et al, 2015). Moreover, spontaneous crying has also been reportedly confused with spontaneous laughter by listeners (Lavan et al, 2014), possibly acting as a confounding factor. Crying in adults is thought to be mainly decoded with the presence of tears to be perceived as authentic, not presenting an auditory dominance (Vingerhoets & Bylsma, 2015). Infants' crying, on the other hand, is a more common emotional expression presenting auditory channel dominance, as it can be accurately recognized exclusively through the auditory modality as authentic (Vingerhoets & Bylsma, 2015). Findings with laughter further confirmed the absence of an authenticity effect, showing no significant difference in the N100, P200 or LPP amplitude or latency between spontaneous and voluntary laughter. The absence of a neurophysiological effect of authenticity may arise from differences in task design, as compared with the fMRI experiments conducted in this field (Lavan et al, 2017; McGettigan et al, 2015). In the fMRI experiments, the participants listened passively to the vocalizations and only later classified them in terms of their perceived authenticity in a behavioural experiment (McGettigan et al, 2015; Lavan et al, 2017). In our study, participants listened to each vocalization and rated its authenticity immediately after listening to it, in each trial. The event-related task design used in the current study may have weakened the differences in authenticity between spontaneous and voluntary vocalizations due to increasing attention demands. Another possible explanation for these findings is that elaborative processing may have been elicited for both types of expressions, which did not occur in previous studies with passive listening of the vocalizations (and may be the case for ecological environments as well). In everyday communication we may further elaborate on voluntary vocalizations, as it is not immediately clear if they were genuinely evoked or not, decoding the intent behind the posed emotion. However, if we instruct people to make a decision for both spontaneous and voluntary vocalizations, we may be eliciting processing that may not naturally be present and thus weakening authenticity effects in voice processing.

Studies on vocal emotional cues processing, including the ones using ERP methodology, have been criticized for applying sets of stimuli with acted portrayals of emotional expressions (voluntary) (Anikin & Lima, 2017; Scherer & Banzinger, 2010), but our results further validate them. While spontaneous portrayals of emotion are ecologically more valid, as they represent accurately everyday communication, acted portrayals also present advantages due to their capacity to represent a culturally conventional signal of a given nonverbal cue and are associated with easy recognition (Scherer & Bänziger, 2010). A recent study has reported that voluntary vocalizations better express characteristics such as identity of the speaker than spontaneous vocalizations, revealing acted portrayals are preferable when investigating voice identity (Lavan et al, 2018).

While not the main focus of our study, we also explored if males and females differ in the affective processing of vocal cues, as the literature has not been consistent regarding sex-based differences in the processing of nonverbal signals (Collignon et al, 2010; Thompson & Voyer, 2014). There is some evidence indicating that females are better than males in emotion recognition irrespective of the input channel (visual, auditory and audiovisual) (Collignon et al, 2010). ERP studies probing the MMN and N400 components showed that females use social and semantic information of the auditory channel more automatically than males and outperform them in emotion recognition (Hung & Cheng, 2014; Schirmer & Kotz, 2003; Schirmer, Lui, Maess & Escoffier, 2006; Schirmer, Striano & Friederici, 2005). Conversely, other studies failed to report differences between males and females using the same N400 paradigm (Schirmer, Kotz & Friederici, 2005). To date, no studies probing the multistage model of vocal affective processing Schirmer & Kotz (2006) have reported sex differences in the N100, P200 or LPP. Also, a link between inter-individual characteristics such as trait empathy and improved emotion recognition, rather than sex-based differences has been previously pointed out by van der Brink and colleagues (2012). In agreement with prior studies, we also verified no significant differences between males and females in the processing of authenticity or emotionality in vocal cues in the three processing stages. If increased amplitude is reflective of increased processing effort (Jessen & Kotz, 2011; Paulmann et al, 2013; Pell et al, 2015; Otten et al, 2017), from this standpoint males would require more cognitive resources for the evaluation of voluntary vocalizations than females. Importantly, human social interactions seldomly entail a ratio as simple as “females are better than males at task A” or the opposite, instead it needs to be taken in to account that a more complex interaction is taking place, depending on a dynamic and constantly changing social context (McKeown, Sneddon & Curran, 2015).

3. Relationship between ERPs and behavioural data

As for our behavioural findings, we asked participants to rate in a Likert scale authenticity, arousal and emotional contagion. The use of rating scales instead of a forced-choice task allowed us to obtain a more precise measure of the continuous perceptual properties of laughter and crying vocalizations by reducing response biases due to competition/ conflict. Regarding authenticity, spontaneous vocalizations were perceived as more authentic than voluntary vocalizations for both laughter and crying, in line with previous reports (Lavan et al 2014; 2016). Anikin and Lima (2017) tested differences in authenticity recognition accuracy in a range of emotions, by exposing listeners to spontaneous and voluntary expressions of achievement, amusement, anger, disgust, fear, pleasure and sadness. Emotional expressions high in arousal, such as achievement, anger, fear and pleasure presented a higher accuracy than those low in arousal, such as amusement, disgust and sadness (Anikin & Lima, 2017). As for arousal, spontaneous vocalizations were perceived as more arousing than voluntary vocalizations, as in behavioural findings with laughter authenticity (Lavan et al, 2014; 2016). In agreement with the hypothesis that non-linguistic affective vocalizations are characterized as more salient (e.g., high arousal) than speech due to its primitive origins and are processed more readily (Pell et al, 2015), the same may apply to spontaneous vocalizations. Correspondingly, similar to speech, voluntary vocalizations demand a high voluntary control of the human voice for the production of signals which in turn may result in lower arousal (Lavan et al, 2018).

Emotional contagion and its relationship with authenticity has been recently explored in laughter. Emotional contagion seems to improve authenticity detection, with spontaneous laughter being more contagious than voluntary laughter (Neves et al, 2018). A similar conclusion was reached in our study, extending the findings from laughter to crying vocalizations as well. Laughter and crying are pervasive non-verbal expressions of emotion that not only influence social interactions but also regulate the responses of whom we interact with (Scott et al, 2015). Emotional contagion plays a role in this response regulation, as whether we resonate with others emotions modulates our communication. Though correlation analysis does not represent causality, our findings do not show a positive correlation between authenticity ratings raw score and affective empathy (Empathic Concern Scale of the IRI) or cognitive empathy (mentalizing – Basic Empathy Scale). These constructs contribute distinctively for authenticity recognition as affective empathy refers to experiencing a similar emotion, while cognitive empathy is characterized by the recognition and understanding of other's intentions and mental states, enabling the prediction of behaviours (Baron-Cohen, 2004). Neves and colleagues (2018) found a positive correlation between authenticity discrimination index (obtained by subtracting spontaneous and voluntary laughter authenticity ratings) and the IRI empathy concern scale. In the present study, no significant correlations were found between individual characteristics such as empathy (EQ) and mentalizing (RMET) and the raw authenticity ratings for both laughter and crying. Additionally, no correlation was found between the authenticity ratings and N100, P200 and LPP amplitude. A positive correlation was found between arousal and emotional contagion for spontaneous laughter only: as expected, the more arousing the emotional expression the more contagious it was rated. Additionally, with spontaneous laughter as well, a negative association between authenticity ratings and arousal was found, suggesting the more authentic a laughter is the higher the arousal it induces. This finding is in line with previous studies reports of laughs rated higher in arousal being perceived as more authentic (Lavan et al., 2016). Lastly, a negative correlation between authenticity ratings and emotional contagion was found exclusively for voluntary crying. This result states not only that voluntary expressions are less emotionally contagious but more specifically for negative voluntary emotions (crying). Laughter can be elicited by hearing another person laughing, being highly contagious, be it actively (direct interaction) or passively (via computer) (Scott et al, 2015). While the emotional contagion of positive emotions may promote bonding and group affiliation (Scott et al, 2015), the emotional contagion of negative emotions may have a strong effect on group dynamics by decreasing the chance of an effective group coping. As such, it may be the case that decreased emotional contagion for negative emotions is evolutionarily advantageous by allowing that when member of the group presents anger or sadness the rest of the group does not share the same negative emotions and promote a more effective group coping.

4. Limitations

The current study presents some limitations such as our sample size. A larger sample would be ideal, possibly potentiating authenticity correlates in neurophysiological data. As our sample was homogenous (mostly young Portuguese college students), it remains uncertain whether our conclusions extend to more diverse populations (e.g., attending to factors such as culture and age). It is important to highlight that although we used recordings of natural spontaneous vocalizations, with acoustic properties similar to those evoked in live interactions, future research should use live recordings of emotions in a context of a real social interaction for increased ecological validity (Anikin & Lima, 2017). The length and variability in duration of the auditory stimuli should also be noted. Indeed, the set of stimuli used was originally developed for fMRI with a variable duration between two and three seconds and ERP studies with non-linguistic vocalizations have a duration close to one second (Liu et al, 2012; Pell et al, 2015), which may have masked or biased the results reported in the current study. Repeating the study with a passive listening task may strengthen the effects of authenticity on the temporal course of vocal processing, as it promotes a similar experience to our daily interactions where only voluntary vocalizations are elaborated more in-depth. Furthermore, the length of the experimental session may have induced fatigue effects and influenced the results, even though we introduced three short breaks throughout the session. As spatial resolution is reduced in EEG, future studies should use both ERP and fMRI methodologies for a more accurate spatial and temporal representation of how our brain processes authenticity in non-linguistic affective cues. Extending the use of neurophysiological approaches to the study of authenticity detection in other vocalizations other than laughter and crying may help to increase our understanding of how the recognition and impact of each emotion is affected by authenticity at different phases of neural processing (Anikin & Lima, 2017). Human communication entails multisensory information, reaching us with complementary action from the visual auditory system. Authenticity accuracy has been found to be enhanced when information is conveyed through audio-visual cues (Lavan & McGettigan, 2016). As such, future studies should inspect the extent to which integration of multisensory stimuli during authenticity decoding may be captured by specific ERP correlates.

CONCLUSIONS

Our ability to decode authenticity in vocal cues, i.e., detecting whether an emotion was spontaneously or voluntarily expressed, is a relevant skill present in everyday social interactions (Lavan et al, 2017; Scott et al, 2015). Our findings shed light on how authenticity in non-linguistic affective vocalizations is processed online according to the multistage model of vocal information processing (Schirmer & Kotz, 2006). Authenticity did not affect vocal emotional early processing (in the first 700 ms after voice onset) as indicated by the absence of differences between spontaneous and voluntary vocalizations in their respective N100, P200 and LPP amplitudes or latencies. It seems a more cognitive process of authenticity appraisal possibly occurring only after 700 ms.

A differential ERP response to emotional (happy and sad) as compared with neutral non-linguistic affective vocalizations was found as early as 100 ms after vocalization onset (N100), as well as in the P200 and LPP components, irrespective of authenticity. Altogether, the present findings suggest that although emotional content of vocal cues may be rapidly decoded, authenticity according to the results obtained does not seem to be decoded during early stages of the respective neural processing.

ACKNOWLEDGMENTS

A realização desta dissertação de mestrado não seria possível sem o importante apoio e suporte que me rodeou constantemente, sempre me motivando a superar os desafios vigentes e atingir os meus objetivos.

À minha orientadora Prof. Dra. Diana Prata, por me ter acolhido no seu laboratório desde o início deste meu percurso no mestrado de neurociências, motivando sempre uma aprendizagem multidisciplinar que me permitiu desenvolver as minhas capacidades.

À minha co-orientadora Prof. Dra. Ana Pinheiro, pela sua disponibilidade, acompanhamento e expertise em EEG sem o qual esta dissertação não teria sido possível. A sua dedicação é cativante e fortaleceu ainda mais o meu interesse por esta área de estudo.

Ao Daniel Martins, pelo constante apoio, paciência e amizade, independentemente das suas ocupações profissionais e pessoais. Serás sempre um ponto de referência a seguir.

À Inês Casqueiro, Vânia Tavares, Carina Mendes e Duarte Ferreira por todo ânimo e o entusiasmo nas novas oportunidades que foram surgindo e encorajamento nos desafios profissionais e pessoais que tanto me permitiram crescer e fortaleceram esta amizade.

To Katja, my dearest post-doc for all her warm encouragement and rigorous scientific thinking.

A todos os membros do DPrata Lab e VAS lab, em especial à Beatriz Simões e Gonçalo Cosme por todo o companheirismo e preocupação.

Aos meus psis, Gonçalo Ferreira, Eduardo Mascarenhas, Beatriz Moreno, João Sarzedas, Rita Fatela e Maria Amorim, por terem sempre cultivado momentos de amizade que tornaram os piores dias memoráveis, o vosso aconselhamento fez toda a diferença.

Ao Jorge Santos, por ser o meu suporte emocional incondicionalmente, acompanhando sempre cada pedaço deste meu percurso tão pouco linear.

Aos meus pais, irmãos e irmã, por todo o carinho, dedicação e motivação para ser mais e melhor, por nunca duvidarem de mim ainda que eu fosse a primeira a fazê-lo.

REFERENCES

Anikin, A., & Lima, C. F. (2018). Perceptual and acoustic differences between authentic and acted nonverbal emotional vocalizations. *Quarterly Journal of Experimental Psychology (2006)*, *71*(3), 622–641. http://doi.org/10.1080/17470218.2016.1270976

Baron-Cohen, S. (2004). Autism and “theory of mind.” *The Applied Psychologist*, 1–20. http://doi.org/10.1007/978-1-4899-0882-7_4

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *34*(2), 163–175. https://doi.org/10.1023/B:JADD.0000022607.19833.00

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

Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends in Cognitive Sciences*, *8*(3), 129–135. http://doi.org/10.1016/j.tics.2004.01.008

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, *403*(6767), 309–312. http://doi.org/10.1038/35002078

Belin, P., Zatorre, R. J., & Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Brain Res Cogn Brain Res*, *13*. http://doi.org/10.1016/S0926-6410(01)00084-2

Blasi, A., Mercure, E., Lloyd-Fox, S., Thomson, A., Brammer, M., Sauter, D., … Murphy, D. G. M. (2011). Early specialization for voice and emotion processing in the infant brain. *Current Biology : CB*, *21*(14), 1220–1224. http://doi.org/10.1016/j.cub.2011.06.009

Brown, S. B. R. E., van Steenbergen, H., Band, G. P. H., de Rover, M., & Nieuwenhuis, S. (2012). Functional significance of the emotion-related late positive potential. *Frontiers in Human Neuroscience*, *6*, 33. http://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. http://doi.org/10.1016/j.evolhumbehav.2014.03.003

Calvo, M. G., & Nummenmaa, L. (2016). Perceptual and affective mechanisms in facial expression recognition: An integrative review. *Cognition and Emotion*, *30*(6), 1081–1106. http://doi.org/10.1080/02699931.2015.1049124

Canavarro, M. C. (2007). Inventário de Sintomas Psicopatológicos: Uma revisão crítica dos estudos realizados em Portugal. In M. Simões, C. Machado, M. Gonçalves, & L. Almeida (Eds.), *Avaliação psicológica: Instrumentos validados para a população Portuguesa* (vol. III, pp. 305-331). Coimbra: Quarteto Editora

Castro, S. L., & Lima, C. F. (2010). Recognizing emotions in spoken language: A validated set of Portuguese sentences and pseudosentences for research on emotional prosody. *Behavior Research Methods*, *42*(1), 74–81. http://doi.org/10.3758/BRM.42.1.74

Charest, I., Pernet, C. R., Rousselet, G. A., Quiñones, I., Latinus, M., Fillion-Bilodeau, S., … Belin, P. (2009). Electrophysiological evidence for an early processing of human voices. *BMC Neuroscience, 10*(1), 127. http://doi.org/10.1186/1471-2202-10-127

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Collignon, O., Girard, S., Gosselin, F., Saint-Amour, D., Lepore, F., & Lassonde, M. (2010). Women process multisensory emotion expressions more efficiently than men. *Neuropsychologia*, *48*(1), 220–225. http://doi.org/10.1016/j.neuropsychologia.2009.09.007

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *115*(4), 732–744. http://doi.org/10.1016/j.clinph.2003.11.021

Dennis, T. A., & Hajcak, G. (2009). The late positive potential: a neurophysiological marker for emotion regulation in children. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *50*(11), 1373–1383. http://doi.org/10.1111/j.1469-7610.2009.02168.x

Derogatis, L. R., & Melisaratos, N. (1983). The Brief Symptom Inventory: an introductory report. *Psychological Medicine*, *13*(3), 595–605. http://doi.org/10.1017/S0033291700048017

Drolet, M., Schubotz, R. I., & Fischer, J. (2012). Authenticity affects the recognition of emotions in speech: behavioral and fMRI evidence. *Cognitive, Affective & Behavioral Neuroscience*, *12*(1), 140–150. http://doi.org/10.3758/s13415-011-0069-3

Ekman, P. (1982). Methods for Measuring Facial Action. In Scherer, K. R. & Ekman, P. (Eds.), *Handbook of Methods in Nonverbal Behavior Research* (pp. 45-90). New York: Cambridge University Press.

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *The facial action coding system* (2nd ed.). Salt Lake City, UT: Research Nexus eBook.

Frühholz, S., & Grandjean, D. (2013). Processing of emotional vocalizations in bilateral inferior frontal cortex. *Neuroscience and Biobehavioral Reviews, 37*(10 Pt 2), 2847–2855. http://doi.org/10.1016/j.neubiorev.2013.10.007

Frühholz, S., Trost, W., & Kotz, S. A. (2016). The sound of emotions-Towards a unifying neural network perspective of affective sound processing. *Neuroscience and Biobehavioral Reviews*, *68*, 96–110. http://doi.org/10.1016/j.neubiorev.2016.05.002

Galinha, I, & Ribeiro, J. (2005). Contribuição para o estudo da versão portuguesa da Positive and Negative Affect Schedule (PANAS): I - Abordagem teórica ao conceito de afecto. *Análise Psicológica, 23* (2).

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*, *8*(8), 918–927. http://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. http://doi.org/10.1086/498281

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. http://doi.org/10.1016/0013-4694(83)90135-9

Gunnery, S. D., & Ruben, M. A. (2016). Perceptions of Duchenne and non-Duchenne smiles: A meta-analysis. *Cognition and Emotion*, *30*(3), 501–515. http://doi.org/10.1080/02699931.2015.1018817

Hawk, S. T., van Kleef, G. A., Fischer, A. H., & van der Schalk, J. (2009). "Worth a thousand words": absolute and relative decoding of nonlinguistic affect vocalizations. *Emotion (Washington, D.C.), 9*(3), 293–305. http://doi.org/10.1037/a0015178

Hung, A.Y., & Cheng, Y. (2014). Sex differences in preattentive perception of emotional voices and acoustic attributes. *Neuroreport*, *25*(7), 464–469. http://doi.org/10.1097/WNR.0000000000000115

Iredale, J. M., Rushby, J. A., McDonald, S., Dimoska-Di Marco, A., & Swift, J. (2013). Emotion in voice matters: neural correlates of emotional prosody perception. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology, 89*(3), 483–490. http://doi.org/10.1016/j.ijpsycho.2013.06.025

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

Jürgens, R., Drolet, M., Pirow, R., Scheiner, E., & Fischer, J. (2013). Encoding Conditions Affect Recognition of Vocally Expressed Emotions Across Cultures. *Frontiers in Psychology*, *4*, 111. http://doi.org/10.3389/fpsyg.2013.00111

Jürgens, R., Grass, A., Drolet, M., & Fischer, J. (2015). Effect of Acting Experience on Emotion Expression and Recognition in Voice: Non-Actors Provide Better Stimuli than Expected. *Journal of Nonverbal Behavior*, *39*(3), 195–214. http://doi.org/10.1007/s10919-015-0209-5

Karpf, A. (2006). *The Human Voice*. London: Bloomsburry.

Kotz, S. A., & Paulmann, S. (2007). When emotional prosody and semantics dance cheek to cheek: ERP evidence. *Brain Research*, *1151*, 107–118. http://doi.org/10.1016/j.brainres.2007.03.015

Kotz, S. A., & Paulmann, S. (2012). Emotion, Language, and the Brain. *Language and Linguistics Compass*, *5*(3), 108-125. http://doi.org/10.1111/j.1749-818x.2010.00267.x

Krumhuber, E. G., & Manstead, A. S. R. (2009). Can Duchenne smiles be feigned? New evidence on felt and false smiles. *Emotion (Washington, D.C.)*, *9*(6), 807–820. http://doi.org/10.1037/a0017844

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. http://doi.org/10.1080/02699931.2014.957656

Lavan, N., & McGettigan, C. (2016). Increased discriminability of authenticity from multimodal laughter is driven by auditory information. *Quarterly Journal of Experimental Psychology (2006)*, *70*(10), 2159–2168. http://doi.org/10.1080/17470218.2016.1226370

Lavan, N., Rankin, G., Lorking, N., Scott, S., & McGettigan, C. (2017). Neural correlates of the affective properties of spontaneous and voluntary laughter types. *Neuropsychologia*, *95*, 30–39. http://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. http://doi.org/10.1007/s10919-015-0222-8

Lavan, N., Short, B., Wilding, A., & McGettigan, C. (2018). Impoverished encoding of speaker identity in spontaneous laughter. *Evolution and Human Behavior*, *39*(1), 139–145. http://doi.org/10.1016/j.evolhumbehav.2017.11.002

Lima, C. F., Anikin, A., Monteiro, A.C., Scott, S. K. &, Castro, S.L. (accepted). Automaticity in the recognition of nonverbal emotional vocalizations. *Emotion*. http://doi.org/10.1037/emo0000429

Lima, C. F., Castro, S. L., & Scott, S. K. (2013). When voices get emotional: a corpus of nonverbal vocalizations for research on emotion processing. *Behavior Research Methods*, *45*(4), 1234–1245. http://doi.org/10.3758/s13428-013-0324-3.

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

Luck, S. J. (2005). *An Introduction to the event-related potential technique*. Cambridge, MA: MIT Press.

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*, (January), 246–257. http://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. http://doi.org/10.1177/1754073914544475

Mouga, S. & Tavares P. *Adult Eyes Test - Portuguese version*. Retrieved from https://www.autismresearchcentre.com/arc_tests.

Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: a review of the literature on human vocal emotion. *The Journal of the Acoustical Society of America*, *93*(2), 1097–1108. http://doi.org/10.1121/1.405558

Murray, M. M., Camen, C., Gonzalez Andino, S. L., Bovet, P., & Clarke, S. (2006). Rapid Brain Discrimination of Sounds of Objects. *J Neurosci*, *26*. http://doi.org/10.1523/JNEUROSCI.4511-05.2006

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. *Quarterly Journal of Experimental Psychology*, 1747021817741800. http://doi.org/10.1177/1747021817741800

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

Otten, M., Mann, L., van Berkum, J. J. A., & Jonas, K. J. (2017). No laughing matter: How the presence of laughing witnesses changes the perception of insults. *Social Neuroscience*, *12*(2), 182–193. http://doi.org/10.1080/17470919.2016.1162194

Paulmann, S., Bleichner, M., & Kotz, S. A. (2013). Valence, authenticity, and task effects in emotional prosody processing. *Frontiers in Psychology*, *4*, 345. http://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. http://doi.org/10.1097/WNR.0b013e3282f454db

Paulmann, S., Pell, M. D., & Kotz, S. A. (2008). Functional contributions of the basal ganglia to emotional prosody: evidence from ERPs. *Brain Research*, *1217*, 171–178. http://doi.org/10.1016/j.brainres.2008.04.032

Paulmann, S., Seifert, S., & Kotz, S. A. (2010). Orbito-frontal lesions cause impairment during late but not early emotional prosodic processing. *Social Neuroscience*, *5*(1), 59–75. http://doi.org/10.1080/17470910903135668

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*, *11*, 14–25. http://doi.org/10.1016/j.biopsycho.2015.08.008

Picton, T. W., Lins, O., & Scherg, M. (1995). The recording and analysis of event-related potentials. In F. Boller & J. Grafman (Series Eds.), & R. Johnson, Jr. (Section Ed.), *Handbook of Neuropsychology: Vol. 10, section 14. Event-related brain potentials and cognition* (pp. 3–73). Amsterdam: Elsevier

Pinheiro, A. P., Barros, C., Dias, M., & Kotz, S. A. (2017a). Laughter catches attention! *Biological Psychology*, *130*, 11–21. http://doi.org/10.1016/j.biopsycho.2017.09.012

Pinheiro, A. P., Barros, C., Vasconcelos, M., Obermeier, C., & Kotz, S. A. (2017b). Is laughter a better vocal change detector than a growl? *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *92*, 233–248. http://doi.org/10.1016/j.cortex.2017.03.018

Pinheiro, A. P., Del Re, E., Mezin, J., Nestor, P. G., Rauber, A., McCarley, R. W., … Niznikiewicz, M. A. (2012). Sensory-based and higher-order operations contribute to abnormal emotional prosody processing in schizophrenia: an electrophysiological investigation. *Psychological Medicine*, *43*(3), 603–618. http://doi.org/10.1017/S003329171200133X

Pinheiro, A. P., Rezaii, N., Nestor, P. G., Rauber, A., Spencer, K. M., & Niznikiewicz, M. (2016). 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. http://doi.org/10.1016/j.bandl.2015.12.003

Pinheiro, A. P., Rezaii, N., Rauber, A., Liu, T., Nestor, P. G., McCarley, R. W., … Niznikiewicz, M. (2014). Abnormalities in the processing of emotional prosody from single words in schizophrenia. *Schizophrenia Research*, *152*(1), 235–241. http://doi.org/10.1016/j.schres.2013.10.042

Provine, R. R. (2004). Laughing, Tickling, and the Evolution of Speech and Self. *Current Directions in Psychological Science*, *13*(6), 215–218. http://doi.org/10.1111/j.0963-7214.2004.00311.x

Provine, R. R., Krosnowski, K. A., & Brocato, N. W. (2009). Tearing: Breakthrough in Human Emotional Signaling. *Evolutionary Psychology*, *7*(1). http://doi.org/10.1177/147470490900700107

Radke, S., & Derntl, B. (2016). Affective responsiveness is influenced by intake of oral contraceptives. *European Neuropsychopharmacology : The Journal of the European College of Neuropsychopharmacology*, *26*(6), 1014–1019. http://doi.org/10.1016/j.euroneuro.2016.03.004

Rodrigues, J., Lopes, A., Giger, J.C., Gomes, A., Santos, J., & Gonçalves, G. (2011). Escalas de Medição do Quociente de Empatia/Sistematização: Um Ensaio de Validação para a População Portuguesa. *Psicologia, 25* (1), 73-89.

Rosburg, T., Boutros, N. N., & Ford, J. M. (2008). Reduced auditory evoked potential component N100 in schizophrenia--a critical review. *Psychiatry Research*, *161*(3), 259–274. http://doi.org/10.1016/j.psychres.2008.03.017

Sauter, D. A., & Eimer, M. (2009). Rapid Detection of Emotion from Human Vocalizations. *Journal of Cognitive Neuroscience*, *22*(3), 474–481. http://doi.org/10.1162/jocn.2009.21215

Sauter, D. A., Eisner, F., Calder, A. J., & Scott, S. K. (2010a). Perceptual cues in non-verbal vocal expressions of emotion. *Quarterly Journal of Experimental Psychology (2006)*, *63*(11), 2251–2272. http://doi.org/10.1080/17470211003721642

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010b). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences*, *107*(6), 2408 LP – 2412. Retrieved from http://www.pnas.org/content/107/6/2408.abstract

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. http://doi.org/10.3758/s13415-012-0132-8

Schirmer, A., & Kotz, S. A. (2003). ERP evidence for a sex-specific Stroop effect in emotional speech. *Journal of Cognitive Neuroscience*, *15*(8), 1135–1148. http://doi.org/10.1162/089892903322598102

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *TRENDS in Cognitive Sciences*, *10*(1). http://doi.org/10.1016/j.tics.2005.11.009

Schirmer, A., Kotz, S. A., & Friederici, A. D. (2005). On the role of attention for the processing of emotions in speech: sex differences revisited. *Brain Research. Cognitive Brain Research*, *24*(3), 442–452. http://doi.org/10.1016/j.cogbrainres.2005.02.022

Schirmer, A., Striano, T., & Friederici, A. D. (2005). Sex differences in the preattentive processing of vocal emotional expressions. *Neuroreport*, *16*(6), 635–639. http://doi.org/10.1097/00001756-200504250-00024

Schirmer, A., Lui, M., Maess, B., Escoffier, N., Chan, M., & Penney, T. B. (2006). Task and sex modulate the brain response to emotional incongruity in Asian listeners. *Emotion (Washington, D.C.)*, *6*(3), 406–417. http://doi.org/10.1037/1528-3542.6.3.406

Scherer, K. R. (1994) "Affect Bursts" In S. van Goozen, N.E. van de Poll & J.A. Sergeant (Eds.), *Emotions: Essays on Emotion Theory* (pp. 161-196). Hillsdale, NJ: Erlbaum.

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, *40*(1), 227–256. http://doi.org/10.1016/S0167-6393(02)00084-5

Scherer, K. R. (2013). Vocal markers of emotion: Comparing induction and acting elicitation. *Computer Speech & Language,* 27 (1), 40–58. https://doi.org/10.1016/j.csl.2011.11.003

Scherer, K. R., & Bänziger, 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* (pp. 166–176). Oxford: Oxford University Press.

Schröder, M. (2003). Experimental study of affect bursts. *Speech Communication*, *40*(1), 99–116. http://doi.org/10.1016/S0167-6393(02)00078-X

Scott, S., Lavan, N., Chen, S., & McGettigan, C. (2014). The social life of laughter. *Trends in Cognitive Sciences*, *18*(12), 618–620. http://doi.org/10.1016/j.tics.2014.09.002

Szameitat, D. P., Kreifelts, B., Alter, K., Szameitat, A. J., Sterr, A., Grodd, W., & Wildgruber, D. (2010). It is not always tickling: distinct cerebral responses during perception of different laughter types. *NeuroImage*, *53*(4), 1264–1271. http://doi.org/10.1016/j.neuroimage.2010.06.028

Thompson, A. E., & Voyer, D. (2014). Sex differences in the ability to recognise non-verbal displays of emotion: a meta-analysis. *Cognition & Emotion*, *28*(7), 1164–1195. http://doi.org/10.1080/02699931.2013.875889

Van den Brink, D., Van Berkum, J. J. A., Bastiaansen, M. C. M., Tesink, C. M. J. Y., Kos, M., Buitelaar, J. K., & Hagoort, P. (2012). Empathy matters: ERP evidence for inter-individual differences in social language processing. *Social Cognitive and Affective Neuroscience*, *7*(2), 173–183. http://doi.org/10.1093/scan/nsq094

Vasconcelos, M., Dias, M., Soares, A. P., & Pinheiro, A. P. (2017). What is the Melody of That Voice? Probing Unbiased Recognition Accuracy with the Montreal

Affective Voices. *Journal of Nonverbal Behavior*, *41*(3), 239–267. http://doi.org/10.1007/s10919-017-0253-4

Vingerhoets, A. J. J. M., & Bylsma, L. M. (2015). The Riddle of Human Emotional Crying: A Challenge for Emotion Researchers. *Emotion Review*, *8*(3), 207–217. http://doi.org/10.1177/1754073915586226

Wang, P., Pan, Z., Liu, X., & Chen, X. (2015). Emotional intensity modulates vocal emotion decoding in a late stage of processing: an event-related potential study. *NeuroReport*, *26*(17). http://doi.org/10.1097/WNR.0000000000000466

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. http://doi.org/10.1037/0022-3514.54.6.1063

Wechsler, D. (2008). *Escala de Inteligência de Wechsler para Adultos (WAIS-III).* Lisboa: Cegoc.

Zougkou, K., Weinstein, N., & Paulmann, S. (2017). ERP correlates of motivating voices: quality of motivation and time-course matters. *Social Cognitive and Affective Neuroscience*, *12*(10), 1687–1700. http://doi.org/10.1093/scan/nsx064