UNIVERSIDADE DE LISBOA

Faculdade de Medicina

Interactions between spatial location and emotion in voice perception: An ERP study

Sara Temudo Emara

Orientador(es): Prof. Doutora Ana Patrícia Teixeira Pinheiro
Prof. Doutora Isabel Pavão Martins

Dissertação especialmente elaborada para obtenção do grau de Mestre em Neurociências

2021

UNIVERSIDADE DE LISBOA

Faculdade de Medicina

Interactions between spatial location and emotion in voice perception: An ERP study

Sara Temudo Emara

Orientador(es): Prof. Doutora Ana Patrícia Teixeira Pinheiro
Prof. Doutora Isabel Pavão Martins

Dissertação especialmente elaborada para obtenção do grau de Mestre em Neurociências

2021

A impressão desta dissertação foi aprovada pelo Conselho Científico da Faculdade de Medicina de Lisboa em reunião de *23 de Novembro de 2021*.

Index

Abstract

In present study we used Event-related potentials (ERP) to understand how spatial and emotional features interact during voice processing, as well as to elucidate the role of attentional requirements on these interactions. Spatialized nonverbal vocalizations differing in valence (neutral, positive [amusement] or negative [anger]) were presented from different locations around the head (front *vs*. back; left *vs*. right), while participants performed either a location discrimination (Task 1) or an emotion recognition task (Task 2). Confirming our first hypothesis, emotional voices elicited decreased N1 amplitudes (though only on the Right-Left axis), and increased P2 and LPP amplitudes, compared to neutral vocalizations. Contrary to what we expected, there were no interactions between emotion and space on P2 and LPP amplitudes, and we did not observe any main effect of task on the amplitude of the N1, P2 or the LPP component. Nevertheless, there was a significant main effect of space on P2 amplitude, indicating that voices presented on the right side elicited increased amplitudes compared to voices presented from the left; and a significant interaction between space and task on P2 amplitude revealed an increased amplitude to voices presented on the front *vs*. back space, but only in the emotion recognition task. These results further confirm the automaticity of emotional processing, as vocal emotion effects were robust to spatial manipulations and occurred even when emotional cues were task-irrelevant and attention was directed away from them (i.e., location discrimination task). On the other hand, spatial effects seem to arise from the salience conferred by voice location and are modulated by attention focus. Taken together, the present findings reinforce the automatic nature of emotional processing during voice perception and reveal a dissociation between the timing of emotional and spatial effects, as emotion effects occur earlier (i.e., N1 timeframe) than spatial effects (i.e., P2 timeframe).

Keywords: emotion, spatial location, attention focus, voice perception, even-related potentials.

Resumo

A voz humana é um dos sons mais salientes no nosso ambiente. Através da voz, é transmitida não só informação de cariz linguístico (e.g., discurso), como informação para-linguística, que nos permite inferir diversas características sobre o seu falante (e.g., identidade, idade, sexo e estado emocional). Contudo, à semelhança de outros sons, a voz também contém informação de cariz espacial que nos informa, por exemplo, da localização do falante.

Estudos de neuroimagem sugerem que o processamento de características espaciais e não-espaciais do som é sustentado por diferentes mecanismos neurais, levando à postulação de modelos de dupla via. De acordo com estes modelos, a perceção do som envolve duas vias de processamento paralelas e parcialmente independentes: informação não-espacial (e.g., identidade ou emoção) recruta áreas mais rostrais do córtex auditivo, formando a via ventral; enquanto informação espacial (e.g., localização espacial) é processada por áreas caudais, formando a via dorsal.

Apesar do processamento de características emocionais e espaciais ser assegurado por mecanismos neurais distintos, a extensão das interações entre as duas vias, durante a perceção do som, é ainda debatida. Adicionalmente, estudos comportamentais apoiam interações entre o processamento emocional e espacial durante a perceção do som. Por exemplo, comparativamente a sons apresentados à frente do participante, sons vindos de trás não só são mais rápida e corretamente localizados, como tendem a induzir emoções negativas mais fortes, sugerindo um viés atencional para o espaço atrás. Recentemente, Pinheiro e colaboradores (2019) investigaram estas interações durante a perceção da voz, manipulando o foco atencional através das instruções da tarefa. Vocalizações não-verbais espacializadas, transmitindo alegria, raiva ou um estado emocional neutro, foram apresentadas em diferentes localizações à volta do participante. As mesmas vocalizações foram utilizadas em duas tarefas distintas: tarefa de discriminação da localização espacial (Tarefa 1) e tarefa de reconhecimento emocional (Tarefa 2). Quando os participantes atendiam à localização da voz, encontrou-se uma interação entre emoção e espaço, revelando maior precisão para vozes emocionais (i.e., divertimento e raiva) quando apresentadas à direita (*vs.* esquerda) ou atrás (*vs.* frente). Quando os participantes atendiam à qualidade emocional da voz, a acurácia no reconhecimento emocional foi elevada, independentemente da emoção expressa e da sua localização espacial. Estes

resultados sugerem que a discriminação espacial é influenciada pela qualidade emocional da voz, mesmo quando pistas emocionais são irrelevantes para a tarefa. No entanto, a categorização emocional não é afetada por manipulações espaciais. Os autores concluíram que as interações entre emoção e espaço, durante a perceção da voz, são moduladas pelas instruções da tarefa. Assim, a nível comportamental, o processamento implícito de propriedades emocionais da voz é afetado por manipulações espaciais. No entanto, os mecanismos neurais subjacentes as estas interações ainda não são conhecidos. Por outro lado, a nível neural, o impacto das instruções da tarefa sobre estas interações é, também, desconhecido.

A técnica de potenciais evocados (ERP) do eletroencefalograma (EEG) permite examinar os diferentes estádios da perceção da voz, desde a apresentação do estímulo até ao momento em que uma resposta comportamental é dada. Os componentes ERP iniciais, como N1 e P2, são sensíveis às propriedades físicas da voz, enquanto os componentes mais tardios, como o *Late Positive Potential* (LPP), refletem fases posteriores do processamento, onde ocorrem operações a nível cognitivo. O componente N1 auditivo surge por volta dos 100ms após a apresentação do estímulo e reflete uma análise sensorial inicial das propriedades acústicas da voz. O componente P2, observado por volta dos 200ms, reflete a deteção automática de saliência emocional da voz. O LPP, tipicamente observado por volta dos 400ms pós estímulo, reflete a avaliação explícita do significado emocional da voz. Assim, a excelente resolução temporal da técnica de ERP é ideal para explorar o decurso das interações entre o processamento emocional e espacial durante a perceção da voz, assim como o impacto das instruções da tarefa sobre estes processos.

O presente estudo teve como objetivo averiguar as interações entre o processamento de pistas emocionais e espaciais transmitidas pela voz, através da análise de potenciais evocados (ERP). Pretendia-se, também, determinar o impacto da manipulação do foco atencional sobre diferentes características da voz (i.e., emoção ou localização espacial) nas interações entre o processamento emocional e espacial. Para esse efeito, vocalizações não-verbais espacializadas, expressando diferentes categorias emocionais (neutro, alegria ou raiva), foram apresentadas em várias localizações à volta do participante (frente *vs*. atrás; direta *vs*. esquerda). A manipulação do foco atencional foi conseguida através das instruções da tarefa. Assim, cada sujeito participou numa de duas tarefas: discriminação da localização espacial da voz (Tarefa 1) ou reconhecimento emocional (Tarefa 2). No total, analisaram-se os dados de 42 participantes, dos quais 20 realizaram a tarefa de discriminação espacial (10 homens), enquanto 22 completaram a tarefa de reconhecimento emocional (11 homens).

Com base em estudos prévios, esperava-se um efeito modulatório da emocionalidade da voz na amplitude dos componentes N1, P2 e LPP. Assim, vozes emocionais resultariam na redução da amplitude de N1 e no aumento da amplitude de P2 e LPP, comparativamente a vozes neutras. Por outro lado, esperava-se um efeito de interação entre emoção e localização espacial da voz nos componentes mais tardios (i.e., P2 e LPP), resultando num aumento da amplitude de P2 e LPP em resposta a vozes emocionais apresentadas atrás (*vs*. frente). Esta hipótese é sustentada por estudos comportamentais que reportam interações entre o processamento emocional e espacial durante a perceção de voz, e sugerem um viés atencional para o espaço atrás. Por fim, a manipulação do foco atencional, conseguida através das instruções da tarefa, deveria promover a alocação de recursos atencionais para diferentes pistas vocais, levando à implementação de estratégias de avaliação do estímulo distintas, tanto a nível percetivo como cognitivo. Assim, seriam de esperar diferenças nas amplitudes dos componentes N1, P2 e LPP, em resposta às mesmas vocalizações, em função do foco atencional (i.e., foco na localização espacial *vs*. foco no conteúdo emocional).

Em conformidade com a primeira hipótese, vozes emocionais (*vs*. neutras) resultaram numa redução da amplitude de N1, ainda que apenas no eixo direita-esquerda, e num aumento da amplitude de P2 e LPP. Por outro lado, ao contrário do esperado, não foram observados efeitos de interação entre emoção e espaço na amplitude de P2 e LPP, nem efeitos principais de tarefa sobre a amplitude de nenhum dos três componentes estudados. No entanto, obteve-se um efeito principal de espaço no componente P2, uma vez que a amplitude foi maior para vozes apresentadas à direita *vs*. esquerda. Adicionalmente, observou-se uma interação entre espaço e tarefa, também no componente P2, revelando uma maior amplitude em resposta a vocalizações apresentadas à frente (*vs*. atrás) na tarefa de reconhecimento emocional. Na tarefa de localização espacial não foram observadas diferenças entre os dois espaços.

Os resultados do presente estudo sustentam a automaticidade do processamento emocional, uma vez que os efeitos emocionais foram robustos a manipulações espaciais e surgiram mesmo quando a emocionalidade da voz não era relevante para a tarefa, nem alvo de atenção (i.e., tarefa de localização espacial). Por outro lado, a nível eletrofisiológico, os efeitos espaciais encontrados parecem resultar da saliência conferida

pela localização da voz, e são sensíveis à manipulação do foco atencional. Por último, estes resultados revelam uma dissociação entre os estádios do processamento vocal onde são observados efeitos emocionais e espaciais, uma vez que os efeitos da emocionalidade da voz surgem mais cedo (i.e., componente N1) que os efeitos da localização espacial (i.e., componente P2). Assim, os resultados do presente estudo reforçam a automaticidade do processamento emocional, e sugerem uma dissociação entre o *timing* dos efeitos emocionais e espaciais durante a perceção da voz.

Palavras-chave: emoção, localização espacial, foco atencional, perceção de voz, potenciais evocados.

1. Introduction

The human voice is one of the most salient sounds in our acoustic environment (Belin et al., 2004). Besides speech, voices also convey information regarding speaker's identity, age, sex, or even emotional state (Belin et al., 2004). The special status of the human voice is further confirmed by neuroimaging studies reporting the existence of voice-selective areas in the auditory cortex, located bilaterally in the superior temporal sulcus (e.g., Belin et al., 2000). Emotional vocal cues are particularly relevant for social interactions, as they allow us to make inferences about the speaker's intentions and current mood (Ceravolo et al., 2016a; Pinheiro et al., 2019), and shape our responses accordingly. Event-related potential (ERP) studies have confirmed a rapid differentiation of emotional and neutral cues within 100ms after voice onset (e.g., Liu et al., 2012). Emotions conveyed through the voice are also automatically detected even when task-irrelevant (Pinheiro et al., 2017). Together, these studies suggest that emotional cues are prioritized during voice perception (e.g., Pinheiro et al., 2015).

Current models of voice perception have focused on the mechanisms underlying the processing of paralinguistic information. For instance, the neurobiological model of Schirmer and Kotz (2006) describes vocal emotion comprehension as a multi-stage process involving the sensory analysis of acoustic information, followed by detection of the emotional salience, and, finally, the cognitive evaluation of the emotional significance of the voice. On the other hand, Belin et al. (2004) proposed a neurocognitive account of voice perception, similar to face perception, where the processing of speech, affective and identity vocal information rely on distinct, partially segregated, functional pathways.

In addition to emotional and identity information ('*what*' cues), voices, like other sounds, also communicate relevant spatial information ('*where*' cues), such as sound source location and proximity, and even motion cues (Middlebrooks, 2015). Auditory spatial information plays a critical role in our everyday lives, especially when visual cues are not available, guiding attention towards salient events in the environment and, in turn, influencing our behavior (Derey et al., 2017). Nonetheless, current neurobiological models of voice perception (e.g., Belin et al., 2004; Schirmer & Kotz, 2006) do not account for spatial features, neglecting the mechanisms involved in spatial processing, and the interplay between '*what'* and '*where'* vocal cues.

Dual-stream models of sound perception

Neuroimaging studies in the auditory modality report distinct brain activation patterns in response to spatial *vs.* nonspatial sound features, suggesting that different neural mechanisms underpin the processing of '*what*' and '*where*' information (Ahveninen et al., 2006). Non-spatial sound features (e.g., pitch or identity cues) typically engage rostral areas of the auditory cortex, such as the anterior portions of the superior temporal gyrus and superior temporal sulcus, planum polare, anterior-lateral Heschl's gyrus, and anterior planum temporale (Ahveninen et al., 2006). Vocal emotional cues also seem to engage rostral-lateral regions of the auditory cortex, namely the right anterior-lateral superior temporal gyrus (Kryklywy et al., 2013) and the right middle superior temporal sulcus (i.e., part of the voice-sensitive areas; Ceravolo et al., 2016a). On the other hand, spatial features recruit caudal areas of the auditory cortex, such as the posterior portions of the superior temporal gyrus and planum temporale (Ahveninen et al., 2006). These studies support the existence of parallel, partially independent, processing streams: whereas voice emotional and identity cues are thought to rely on rostral areas of the auditory cortex (ventral stream; Kryklywy et al., 2013; Rauschecker & Scott, 2009), spatial cues are thought to engage caudal areas (dorsal stream; Ahveninen et al., 2006; Rauschecker & Scott, 2009).

Interactions between emotion and space in voice perception

Although the processing of emotional and spatial features appears to rely on distinct brain pathways, the extent to which these streams interact is still a matter of debate. Behavioral evidence supports interactions between emotional and spatial information during sound perception. For instance, Asutay and Västfjäll (2015) found that sounds presented from the rear were located faster and more accurately, and induced stronger negative emotions, compared to sounds presented from the front. These findings suggest an attentional bias towards the rear space. Specifically, spatial location may ascribe emotional salience to a sound, leading to enhanced allocation of attentional resources, and ultimately resulting in facilitation effects observed in task performance (Asutay & Västfjäll, 2015).

Recently, Pinheiro et al. (2019) investigated the interactions between emotional and spatial properties of vocalizations when attention was focused on either emotion or spatial location. Spatialized nonverbal vocalizations expressing amusement, anger or a neutral state were presented in different locations around the head. When participants directed their attention to voice location, an interaction between emotion and space revealed a more accurate location of emotional voices if presented on the right (*vs.* left) side or from the back (*vs.* front). When participants directed their attention to the emotional quality of the voice, emotion recognition was overall high for all emotion types, regardless of spatial location. These results suggest that spatial discrimination is influenced by the emotional quality of the voice, even when task-irrelevant. However, spatial properties did not affect emotional recognition, suggesting that emotion categorization is unaffected by spatial manipulations. Together, these findings reveal that emotion and space interactions in voice perception are modulated by task instructions.

Although the spatial properties of the voice only appear to affect implicit (but not explicit) emotional processing at a behavioral level, the brain mechanisms underpinning the interactions between space and emotion remain to be specified. Neuroimaging studies have reported effects of emotional salience on spatial processing (e.g., Ceravolo et al., 2016b). Additionally, there is evidence for a modulatory role of stimulus location on attentional resources devoted to emotional voices (Ceravolo et al., 2016a), as well as on the perceived arousal (Tajadura-Jiménez et al., 2010) and perceived valence of environmental sounds (Asutay & Västfjäll, 2015). However, it remains to be clarified how spatial and emotional properties of the voice interact at the brain level. On the other hand, at the behavioral level, interactions between emotional and spatial vocal cues appear to be modulated by task instructions (Pinheiro et al., 2019). Electrophysiological studies suggest that emotional processing is affected by task instructions (e.g., Ho et al., 2015; Spreckelmeyer et al., 2009; Van Strien et al., 2010), as there are amplitude differences between the ERP components elicited when participants direct their attention towards emotional cues (i.e., explicit task instructions) *vs.* when attention is directed towards other stimulus features (i.e., implicit task instructions). Nonetheless, at the brain level, the impact of task instructions on the interactions between emotional and spatial vocal cues (i.e., when the listener's attention is focused on the spatial location or emotional quality of the voice), is still unknown.

Event-Related Potentials in voice perception

The excellent temporal resolution of the electroencephalogram (EEG), especially compared to hemodynamic measures such as functional Magnetic Resonance Imaging (fMRI) or Positron Emission Tomography (PET) (Luck, 2005), allows a more precise assessment of the time-course of spatial and emotional processing, as well as of the timing of interactions between these processes (i.e., interactions between the dorsal and ventral streams). Event-related potentials (ERP) of the EEG can be used to examine the different stages of voice processing, from stimulus onset until a behavioral response is made (Kotz & Paulmann, 2011). Early ERP components, such as the N1 and P2, are sensitive to the physical features of the voice (Kotz & Paulmann, 2011; Liu et al., 2012). The auditory N1 component occurs within the first 100ms after stimulus onset and reflects the initial sensory analysis of the acoustic properties of the voice (Kotz & Paulmann, 2011; Liu et al., 2012). The P2 component, peaking approximately at 200ms after stimulus onset, is thought to reflect the automatic detection of emotional salience (Liu et al., 2012). The Late Positive Potential (LPP), typically observed 400ms after stimulus onset, reflects later cognitive operations of stimulus evaluation, including its emotional significance (Kotz & Paulmann, 2011; Pell et al., 2015).

These ERP components appear to be sensitive to the emotional properties of vocal expressions (Kotz & Paulmann, 2011; Liu et al., 2012; Paulmann et al., 2013; Pell et al., 2015). For instance, N1 amplitude tends to be less negative in response to emotional *vs*. neutral stimuli (Kotz & Paulmann, 2011; Liu et al., 2012), while the P2 and LPP components show enhanced amplitude to emotional *vs*. neutral stimuli (Kotz & Paulmann, 2011; Paulmann et al., 2013; Pell et al., 2015). Further, stimulus arousal has also been shown to modulate P2 and LPP amplitudes, as highly arousing stimuli tend to elicit increased P2 (Paulmann et al., 2013) and LPP (Cuthbert et al., 2000) amplitudes, compared to stimuli low in arousal.

Spatial manipulations were also found to affect the N1 (Getzmann & Lewald, 2012; Lewald & Getzmann, 2011; Salminen et al., 2015; Valdés-Conroy et al., 2014), P2 (Getzmann & Lewald, 2012; Lewald & Getzmann, 2011; Xie et al., 2014), and LPP components (Valdés-Conroy et al., 2014). In a sound locating task, there was an increase in N1 and P2 amplitudes to sounds presented in eccentric *vs*. central locations (Lewald & Getzmann, 2011). On the other hand, location-specific attenuations of N1 amplitude are reported in studies implementing adaptation paradigms, where adaptor-probe sound pairs are presented on either the same (e.g., both on the right) or different (e.g., adaptor on the right and probe on the left) locations (Salminen et al., 2015). Specifically, an attenuation of N1 amplitude was found in response to probes presented on the same location as adaptors, even when adaptor and probe sounds did not share the same localization cues (e.g., interaural level differences [ILD] stimuli were used as probes and interaural time differences [ITD] stimuli as adaptors or vice versa; Salminen et al., 2015). Additionally, sound motion also modulates N1 and P2 amplitude: abrupt changes on the spatial location of sounds elicit increased N1 and P2 amplitudes compared to smooth motion (Getzmann & Lewald, 2012). On the other hand, studies using visual stimuli also show increased N1 and LPP amplitude in response to objects located on the near *vs*. far space (Valdés-Conroy et al., 2014), and an enhanced P2 amplitude in response to top *vs*. bottom spatial cues (Xie et al., 2014). Of note, there is also evidence supporting interaction effects between emotional and spatial cues on the P2 (Xie et al., 2014) and LPP amplitude (Du et al., 2017) in the visual modality. Specifically, when participants keep positive words in mind, P2 amplitude is increased to visual cues presented on the top of the screen; whereas keeping negative words in mind results in increased P2 amplitudes to visual cues presented on the bottom of the screen (Xie et al., 2014). Similarly, LPP amplitude to negative stimuli is further increased when presented near *vs*. far away from the participants' hand (Du et al., 2017).

Finally, these ERP components are also modulated by task instructions, such as the implicit *vs.* explicit processing of emotional information (e.g., Ho et al., 2015). For instance, task-irrelevant vocal emotional information was associated with an increased N1 amplitude when presented in spatially attended *vs*. unattended locations (Gädeke et al., 2013). Effects of attention focus were also observed on the P2: its amplitude was decreased when vocal emotions were processed under explicit (i.e., emotion categorization task) *vs.* implicit (i.e., speaker identification task) task instructions (Spreckelmeyer et al., 2009). The LPP was also found to be increased in response to emotional faces when participants had to judge their emotional valence compared to when they were asked to perform a sex classification task (Van Strien et al., 2010). On the other hand, attention focus effects are also reported when comparing tasks involving the assessment of '*what' vs*. '*where'* sound features (Anourova et al., 2001; Leavitt et al., 2011). Anourova et al. (2001) recorded simultaneous EEG and MEG signals during location and pitch matching-to-sample tasks. The authors found an increased amplitude of the N1m (i.e., the magnetic counterpart of N1) in the match condition, and of the P2 in the nonmatch condition, when attention was focused on location *vs*. pitch discrimination (Anourova et al., 2001). Moreover, in a study presenting several animal calls on distinct spatial locations, the N1 amplitude was found to be increased when participants focused on identifying the animal *vs*. locating the sound (Leavitt et al., 2011). Taken together the evidence presented above suggests a modulatory role of attention focus on sound processing, as directing attention towards specific stimulus features leads to differences in the way those features are processed.

The current study and hypotheses

The current study examines the time course of the interactions between spatial and emotional cues in voice perception using ERPs, with a focus on the N1, P2, and LPP components. Spatialized nonverbal vocalizations portraying either positive (amusement), negative (anger) or neutral states were presented in different locations around the head (left *vs*. right; front *vs*. back). Furthermore, the role of attention focus to either emotional or spatial properties of the voice was specified: the same vocalizations was presented in two tasks, one requiring the discrimination of stimulus location, and the other requiring the identification of the emotion expressed. By manipulating the focus of attention towards distinct vocal features (i.e., emotional or spatial cues), it is possible to determine the impact of attentional requirements on the interactions between spatial and emotional vocal cues.

Based on previous ERP studies using nonverbal vocalizations, amplitude modulations were expected as a function of emotional salience (Kotz & Paulmann, 2011; Liu et al., 2012). Specifically, emotional vocalizations should elicit reduced N1 amplitude compared to neutral ones (Liu et al., 2012). On the other hand, the P2 and LPP components, should exhibit more positive amplitudes in response to emotional compared to neutral vocalizations (Liu et al., 2012; Paulmann et al., 2013; Pell et al., 2015).

We also hypothesized interactions between emotion and space on the P2 and LPP. Previous studies revealed that the spatial source of a sound may also convey salience (Asutay & Västfjäll, 2015), and specifically that sounds in the rear space are perceived as more arousing than sounds in the front space (Tajadura-Jiménez et al., 2010). Additionally, the P2 component has been proposed to reflect automatic salience detection, and both the P2 and the LPP are modulated by arousal (Kotz & Paulmann, 2011; Paulmann et al., 2013; Pell et al., 2015). On the other hand, sound location can modulate P2 amplitude (Getzmann & Lewald, 2012; Lewald & Getzmann, 2011), and studies in the visual modality report interaction effects of emotional and spatial features on both the P2 (Xie et al., 2014) and LPP components (Du et al., 2017). Considering the evidence presented above, we should observe increased P2 and LPP amplitudes in response to emotional vocalizations presented from the back (*vs.* front). If confirmed, it would provide indirect evidence for an interaction between the dorsal and ventral processing streams in voice perception from early stages of voice perception.

The manipulation of the attention focus provided by task instructions should promote the allocation of attentional resources towards distinct vocal features, resulting in the implementation of different perceptual and cognitive strategies underlying stimulus evaluation. As mentioned above, previous studies report modulatory effects of task instructions (e.g., implicit *vs*. explicit processing of emotional information) on the N1 (Ho et al., 2015), P2 (Spreckelmeyer et al., 2009), and the LPP (Van Strien et al., 2010) amplitude. Therefore, we expected differences in the amplitudes of the N1, P2, and LPP components in response to spatialized emotional vocalizations as a function of attention focus. Notwithstanding, this hypothesis is exploratory since there is no direct evidence of ERP modulations by attentional focus on emotional *vs.* spatial features of the voice.

2. Methods:

2.1. Participants:

An a priori power analysis was conducted using MorePower software (Campbell & Thompson, 2012) to estimate the sample size based on a mixed ANOVA. Considering a medium-sized effect of 0.25, α at 0.05 and power set to 0.80, the required sample size consists of 40 subjects.

Forty-nine participants were recruited at Faculty of Psychology - University of Lisbon and by word of mouth. Seven participants were excluded: two of them did not meet the inclusion criteria, and the remaining five were excluded due to excessive EEG motion artefacts. The final sample consisted of 42 participants (21 male), between 19 and 29 years of age (Mean age = 22.40, SD = 2.60). Twenty participants took part in the location discrimination task (10 male; Mean age = 22.30, SD = 2.55, age range 19–28 years), whereas 22 individuals performed the emotion recognition task (11 male; Mean age = 22.50, SD = 2.64, age range 19–29 years). Participants were tested at Faculty of Psychology - University of Lisbon. Inclusion criteria were: European Portuguese as native language; age between 18 and 31 years; right handedness (Edinburgh Handedness Inventory; Oldfield, 1971); no self-reported history of psychiatric disorder (Brief Symptoms Inventory [BSI]; Portuguese adaptation by Canavarro, 1999; Derogatis, 1993), or substance abuse (Alcohol, Smoking and Substance Involvement Screening Test [ASSIST]; validated for the Portuguese population by Mostardinha et al., 2019; WHO ASSIST Working Group, 2010); no use of medication for psychiatric disorders; normal or corrected-to-normal vision and hearing. Written informed consent was obtained from all participants, who received course credits for their participation. The study was approved by the Ethics Committee of the Faculty of Psychology - University of Lisbon.

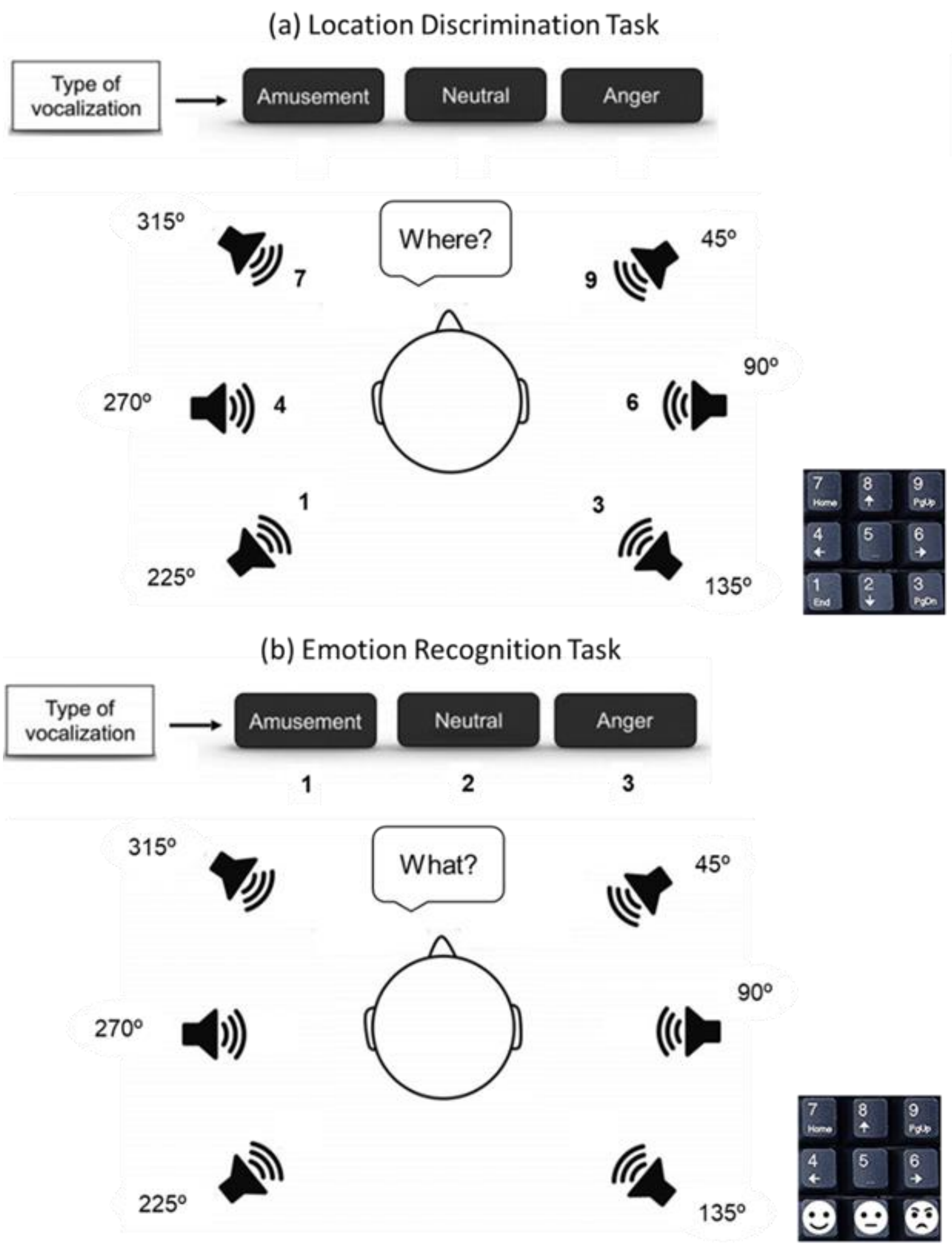
2.2. Stimuli:

Spatialized nonverbal vocalizations portraying positive (amusement), negative (anger), or neutral states were presented in six different locations around the head: 45º, 90º, 135º, 225º, 270º, 315º (see Figure 1). As in Pinheiro et al. (2019), the two emotional categories were chosen as they represent opposites of the valence continuum and share similar acoustic profiles (e.g., high intensity and variable F0; Juslin & Laukka, 2003). Emotional vocalizations were selected from the corpus of nonverbal vocalizations by Lima et al. (2013), and neutral vocalizations were selected from the Montreal Affective Voices battery (MAV; Belin et al., 2008), validated for the Portuguese population (Vasconcelos et al., 2017). The same set of stimuli was used in both tasks.

The spatialization of vocalizations was achieved using the MIT Head Related Transfer Function (HRTF) database (http://sound.media.mit.edu/resources/KEMAR.html). Binaural sounds were generated with [45°, 90°, 135°, 225°, 270°, 315°] azimuth degrees. The impulse responses of each HRTF were then convoluted with the original sound samples, originating a spatially located sound, by adding inter-aural time difference (ITD) and inter-aural level difference (ILD) cues to correspond to a new spatial location (see Pinheiro et al., 2019). The stimuli were presented via headphones (HD 202, Sennheiser). For each type of vocalization (i.e., amusement, neutral, and anger; total = 144), eight different stimuli were selected (4 female and 4 male). Sounds were repeated to achieve 35 stimuli per condition, leading to a total of 630 vocalizations presented in each task.

Figure 1.

Schematic illustration of the experimental set-up and response keys of Location Discrimination (a) and Emotion Recognition (b) tasks. Adapted from Pinheiro et al. (2019).



2.3. Procedure:

Each subject participated in one of two tasks: Task 1 - location discrimination; Task 2 - emotion recognition. At the beginning of the session, participants were asked to fill in several questionnaires: Edinburgh Handedness Inventory (Oldfield, 1971); Brief Symptoms Inventory (BSI; Portuguese adaptation by Canavarro, 1999; Derogatis, 1993) to control for psychopathological symptoms; Alcohol, Smoking and Substance Involvement Screening Test (ASSIST; Portuguese adaptation by Mostardinha et al., 2019; WHO ASSIST Working Group, 2010) to control for substance abuse; Positive and Negative Affect Scale (PANAS; Portuguese adaptation by Galinha & Pais-Ribeiro, 2012; Watson et al., 1988) to control for individual differences in mood states before the experiment. After completing the EEG task, participants rated the valence and arousal of each vocalization, using a 9-point scale (Bradley & Lang, 1994).

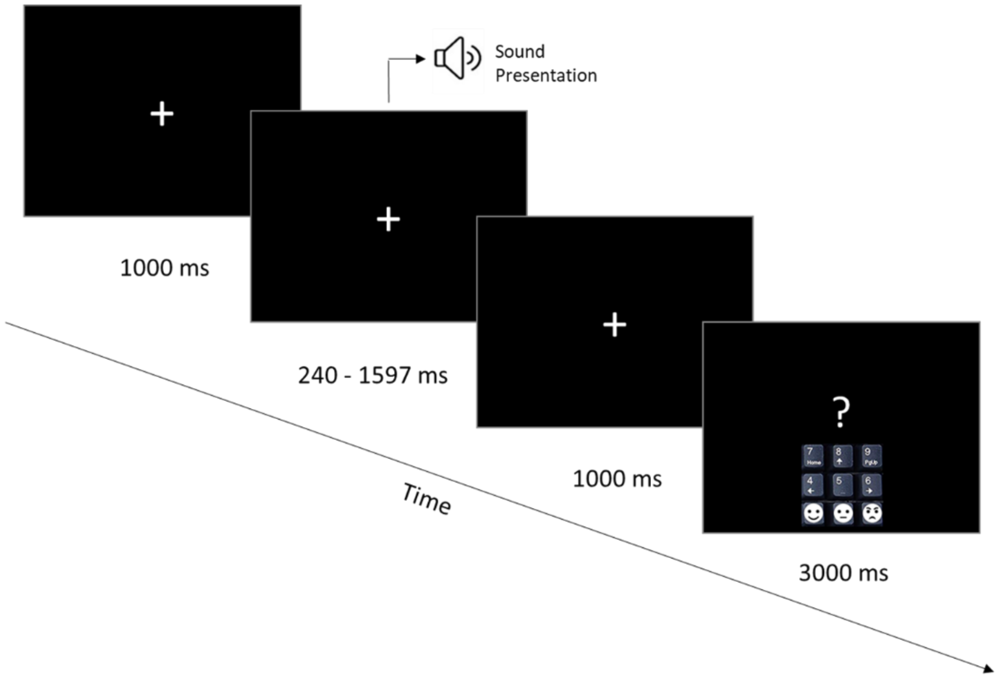
EEG sessions were conducted in an electrically shielded and acoustically isolated booth. Participants were comfortably seated 100 cm away from the computer monitor and responses were provided using the numberpad on the keyboard. Stimulus presentation, timing of events, and recording of participants' responses were controlled with Presentation software (Neurobehavioral Systems, Inc., Albany NY, USA). Both tasks involved a similar trial structure (see Figure 2). Each trial began with a fixation cross, presented 1000ms prior to the sound and which remained during and 1000ms after sound presentation. Then, a question mark, as well as a picture of the numberpad, were presented to prompt participant's response and minimize working memory demands. Participants had a maximum of 3000ms to respond before the beginning of the next trial. Each task consisted of 630 trials, distributed over four blocks. Participants were allowed short pauses of 20 seconds every 32 stimuli, as well as longer breaks between blocks. Stimuli were pseudorandomized to ensure that the same emotion or location was not presented more than three consecutive times (Bertels et al., 2013). Each task lasted approximately 60 minutes.

Location Discrimination Task: Participants were asked to discriminate the spatial location of the vocalizations (see Figure 1a). Before the EEG task, a training session with 70 trials allowed participants to get familiar with the task instructions and response keys. During training, feedback was provided. Vocalizations used in the training session were not included in the EEG task. No feedback was provided during the EEG task.

Emotion Recognition Task: Participants were asked to identify the emotional category associated with each vocalization in a forced choice emotional recognition task (see Figure 1b). Before the EEG task, a training session with six trials allowed participants to get familiar with the task instructions and response keys. No feedback was provided during training. Vocalizations presented in the training session were not included in the EEG task. Participants responded using the three bottom keys of the numberpad, which had stickers of cartoon faces portraying the three emotional categories (i.e., amusement, neutral, and angry) to minimize working memory demands (see Figure 1b; Figure 2). The keys assigned to each emotion were counterbalanced in order to avoid a bias due to an implicit association of emotion to space (Amorim & Pinheiro, 2019). No feedback was provided during the EEG task.

Figure 2.

Illustration of an experimental trial in the Emotion Recognition task.



Note. The trial structure is similar for both tasks, apart from the cartoon faces over the bottom keys on the numberpad, which were not included in the Location Discrimination Task.

Affective Ratings: After completing the EEG task, participants rated the affective properties of the stimulus in two dimensions (i.e., valence and arousal), using a 9-point scale (Bradley & Lang, 1994). Following stimulus presentation, participants indicated (a) how pleasant (1 = very unpleasant, 9 = very pleasant), and (b) how arousing they considered each sound (1 = very calm, 9 = very aroused). These ratings were collected for all 144 vocalizations and the task lasted approximately 30 minutes.

2.4. EEG Data Recording:

EEG data was recorded using a 64-channel Active Two Biosemi system (Biosemi, Amsterdam, The Netherlands) in a continuous mode at a digitization rate of 512 Hz and stored on disk for later analysis. Eye blinks and movements were monitored through electrodes placed on both temples (horizontal electrooculogram), and another one below the left eye (vertical electrooculogram). Two additional electrodes were placed on both mastoids to be later used as offline references.

2.5. EEG Data Analysis:

EEG data were analyzed using BrainVision Analyzer 2 software (Brain Products, Munich, Germany). The EEG channels were referenced offline to the average of the left and right mastoids and filtered using a band-pass filter with a 0.1 Hz and 30 Hz, low and high cutoff frequency. Individual epochs were created with a -200ms pre-stimulus baseline and 1000ms post-stimulus for each condition (i.e., Neutral 45º; Neutral 90º; Neutral 135º; Neutral 225º; Neutral 270º; Neutral 315º; Amusement 45º; Amusement 90º; Amusement 135º; Amusement 225º; Amusement 270º; Amusement 315º; Anger 45º; Anger 90º; Anger 135º; Anger 225º; Anger 270º; Anger 315º). Epochs were baseline-corrected using the -200 to 0ms pre-stimulus interval. Vertical and horizontal eye movements were corrected using the method of Gratton et al. (1983), and then the segments were semi-automatically screened for eye movements, muscle artifacts, electrode drifting and amplifier blocking. EEG epochs exceeding ±100 µV were rejected. After artifact rejection, at least 75% of segments per condition per participant entered the analyses. There were no differences between the number of segments included per condition on either task (p >.50).

Only EEG epochs associated with correct behavioral responses were included in the analysis for the emotion recognition task. In the location discrimination task, trials associated with both correct and incorrect responses were included, due to task difficulty. Data were analyzed with a focus on the N1, P2, and LPP components. Mean amplitudes for these components were analyzed in time windows selected according to previous studies (Liu et al., 2012; Paulmann et al., 2013; Pell et al., 2015; Pinheiro et al., 2016): 110-190 ms (N1), 200-280 ms (P2), and 450-700 ms (LPP). Mean amplitude was chosen over peak amplitude as it is less sensitive to high frequency noise and variability in components' latencies (Luck, 2005). Based on previous studies (Liu et al., 2012; Paulmann et al., 2013; Pell et al., 2015; Pinheiro et al., 2016), as well as careful inspection of grand average waveforms, the following regions of interest (ROI) were selected for the statistical analyses: Midline (FCz, Cz, CPz), Medial Right (FC4, C4, CP4), Medial Left (FC3, C3, CP3), Lateral Right (FC6, C6, CP6), Lateral Left (FC5, C5, CP5).

2.6. Statistical Analysis:

The SPSS statistical software package (Version 26.0, IBM Corp., Armonk, NY, USA) was used for the statistical analyses. The alpha level was set at .05.

In order to reduce the number of levels in the factor, and simplify the statistical models, locations were grouped into four spaces (Front: 45º and 315º; Back: 135º and 225º; Right: 45º, 90º and 135º; and Left: 225º, 270º and 315º), as in Pinheiro et al. (2019). The data was separately analyzed on two space axes: Front-Back and Right-Left. Since all the locations were lateralized (see Figure 1), Right and Left spaces include three locations each (e.g., Right: 45º, 90º and 135º). However, Front and Back spaces only included two locations each (e.g., Front: 45º and 315º), which are also present in Right and Left spaces (e.g., 45º is in both Front and Right spaces). Thus, analyzing all four spaces in the same statistical model would result in two main issues: The inclusion of the same trials in different conditions simultaneously (e.g., Anger 45º in both Anger_Front and Anger_Right); and the comparison of spaces that are not equivalent in the number of locations they encompass (e.g., Front space includes two locations [45º and 315º], but Right space includes three locations [45º, 90º and 225º]). To avoid these issues, the data were separately analyzed on two space axes.

2.6.1. ERP Data:

Repeated-measures analyses of variance (ANOVAs) were separately computed for N1, P2, and LPP mean amplitude, with Emotion (3 levels: Neutral, Amusement, Anger), Space (2 levels: Front and Back, or Right and Left), and ROI, (5 levels: Midline, Medial Right, Medial Left, Lateral Right, Lateral Left) as within-subject factors, and Task (2 levels: Location discrimination, Emotion recognition) as a between-subject factor. Main effects and interactions were followed up with pairwise comparisons using Bonferroni correction for multiple comparisons. The analyses were corrected for non-sphericity using the Greenhouse–Geisser method, when the Mauchly's test indicated that the assumption of sphericity had been violated.

2.6.2. Behavioral Data:

Behavioral measures include recognition accuracy data in each task (correct responses in location discrimination and emotion recognition tasks), and affective ratings (valence and arousal of each vocalization, rated on a 9-point scale; Bradley & Lang, 1994). These behavioral measures were separately analyzed using a similar model to the ERP data: a repeated-measures ANOVA with Emotion (3 levels: Neutral, Amusement, Anger) and Space (2 levels: Front and Back, or Left and Right) as within-subject factors. Main effects and interactions were followed up with pairwise comparisons using Bonferroni correction for multiple comparisons. The analyses were corrected for non-sphericity using the Greenhouse–Geisser method, when the Mauchly's test indicated that the assumption of sphericity had been violated.

3. Results

3.1. ERP results

Mean amplitudes of the N1, P2 and LPP components per condition are shown in Tables 1, 2 and 3, respectively.

3.1.1. Front-Back Axis

3.1.1.1. N1

The main effect of emotion on N1 amplitude was significant [$F(2, 80) = 4.202$; $p = .018$, *partial* $\eta^2 = .095$] (see Figure 3). However, pairwise comparisons revealed no differences between neutral and amused (*p* = .088) or angry (*p* = .058) vocalizations nor between amused and angry (*p* = 1) vocalizations (see Table 1).

Contrary to what was predicted, there was no main effect of Task [$F(1,40) < .001$, $p = .997$] (see Figure 6), nor any interaction between Emotion and Space [$F(2, 80) = .183$, $p = .833$] (see Figure 4), Emotion and Task [$F(2, 80) = .452$, $p = .638$], Space and Task [$F(1, 40) = 2.458$, $p = .125$], and Emotion, Space, and Task [$F(2, 80) = .882$, $p = .418$].

Table 1

Mean amplitude of the N1 component per condition

Emotion	Space	N1 *M (SD)*
Neutral	Front	-1.612 (0.317)
	Back	-1.421 (0.297)
	Right	-1.715 (0.318)
	Left	-1.811 (0.316)
Amusement	Front	-1.233 (0.283)
	Back	-1.132 (0.299)
	Right	-1.232 (0.289)
	Left	-1.332 (0.297)
Anger	Front	-1.200 (0.293)
	Back	-1.148 (0.308)
	Front	-1.200 (0.293)
	Back	-1.148 (0.308)

Note. M = Mean; SD = Standard Deviation.

3.1.1.2. P2

There was a significant main effect of emotion on the P2 amplitude [$F(2, 80) = 14.783$; $p<.001$, *partial* $\eta^2 = .270$] (see Figure 3), revealing an increased amplitude to emotional compared to neutral vocalizations ($ps \leq .002$), with no differences between amused and angry (*p* = .555) vocalizations (see Table 2). The interaction between space and task was also significant [$F(1, 40) = 4.669$, $p = .037$, *partial* $\eta^2 = .105$] (see Figure

5). Vocalizations presented in front locations elicited an increased P2 amplitude compared to those presented in the back (*p* = .029), but only in the emotion recognition task. In the location discrimination task, there were no differences between voices presented on the front or back locations (*p* = .410),

The main effect of Task was nonsignificant [$F(1,40) = .274$, $p = .603$] (see Figure 6), as were the interactions between emotion and space [$F(2, 80) = .152$, $p = .859$] (see Figure 4), emotion and task [$F(2, 80) = 1.146$, $p = .323$], and emotion, space and task [$F(2, 80) = .465$, $p = .630$].

Table 2

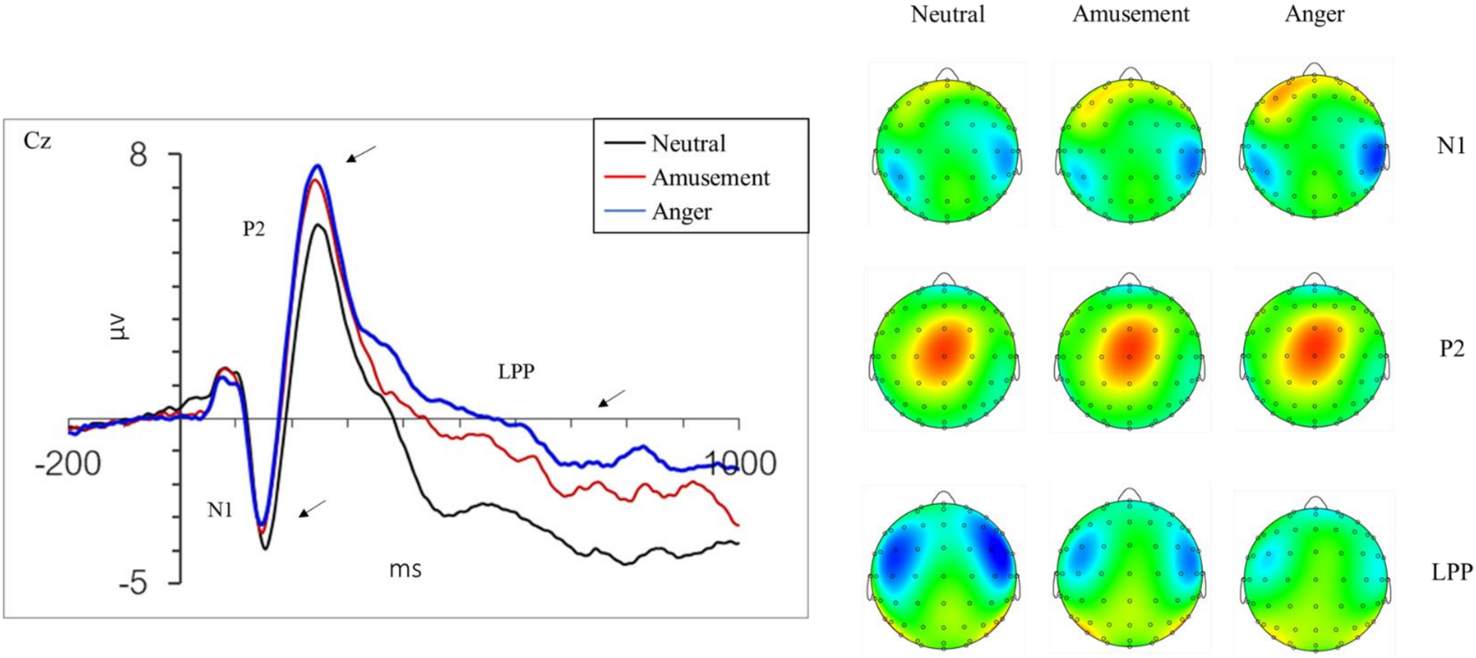
Mean amplitude of the P2 component per condition

Emotion	Space	P2 *M (SD)*
Neutral	Front	2.491 (0.348)
	Back	2.317 (0.322)
	Right	2.657 (0.359)
	Left	2.096 (0.328)
Amusement	Front	3.088 (0.320)
	Back	2.987 (0.336)
	Right	3.230 (0.335)
	Left	3.026 (0.301)
Anger	Front	3.251 (0.341)
	Back	3.242 (0.355)
	Front	3.251 (0.341)
	Back	3.242 (0.355)

Note. M = Mean; SD = Standard Deviation.

Figure 3.

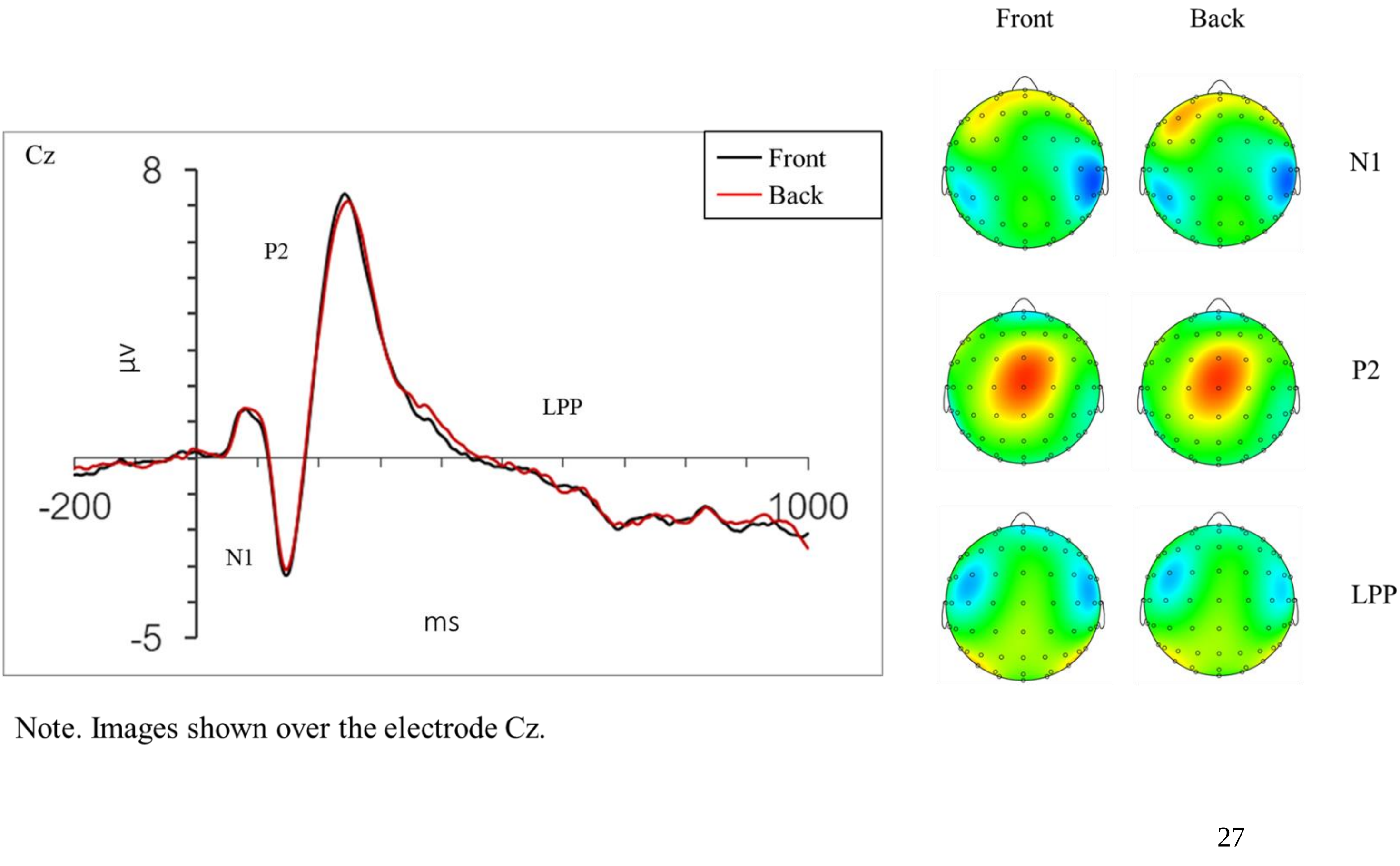
Grand average waveforms contrasting neutral, amused and angry voices. Topographic maps of the N1(110-190ms), P2 (200-280ms) and LPP (450-700ms) components.



Note. Images shown over the electrode Cz. The arrows point to statistically significant effects.

Figure 4.

Grand average waveforms contrasting emotional voices on front vs. back spaces. Topographic maps of the N1(110-190ms), P2 (200-280ms) and LPP (450-700ms) components.



Note. Images shown over the electrode Cz.

3.1.1.3. LPP

There was a significant main effect of emotion on the LPP amplitude, [$F(2, 80) = 60.927$; $p<.001$, *partial* $\eta^2 = .604$] (see Figure 3). Emotional vocalizations elicit an increased LPP amplitude compared to neutral ones (*ps*<.001), and angry vocalizations also resulted in increased LPP amplitudes compared to amused ($p = .004$) ones (see Table 3).

There was no main effect of Task [$F(1,40) =.674$, $p = .417$] (see Figure 6), nor any interactions between emotion and space [$F(2, 80) = .316$, $p = .730$] (see Figure 4), emotion and task [$F(2, 80) = .061$, $p = .940$], space and task [$F(1, 40) = .693$, $p = .410$], nor between emotion, space and task [$F(2, 80) = .722$, $p = .489$].

Table 3

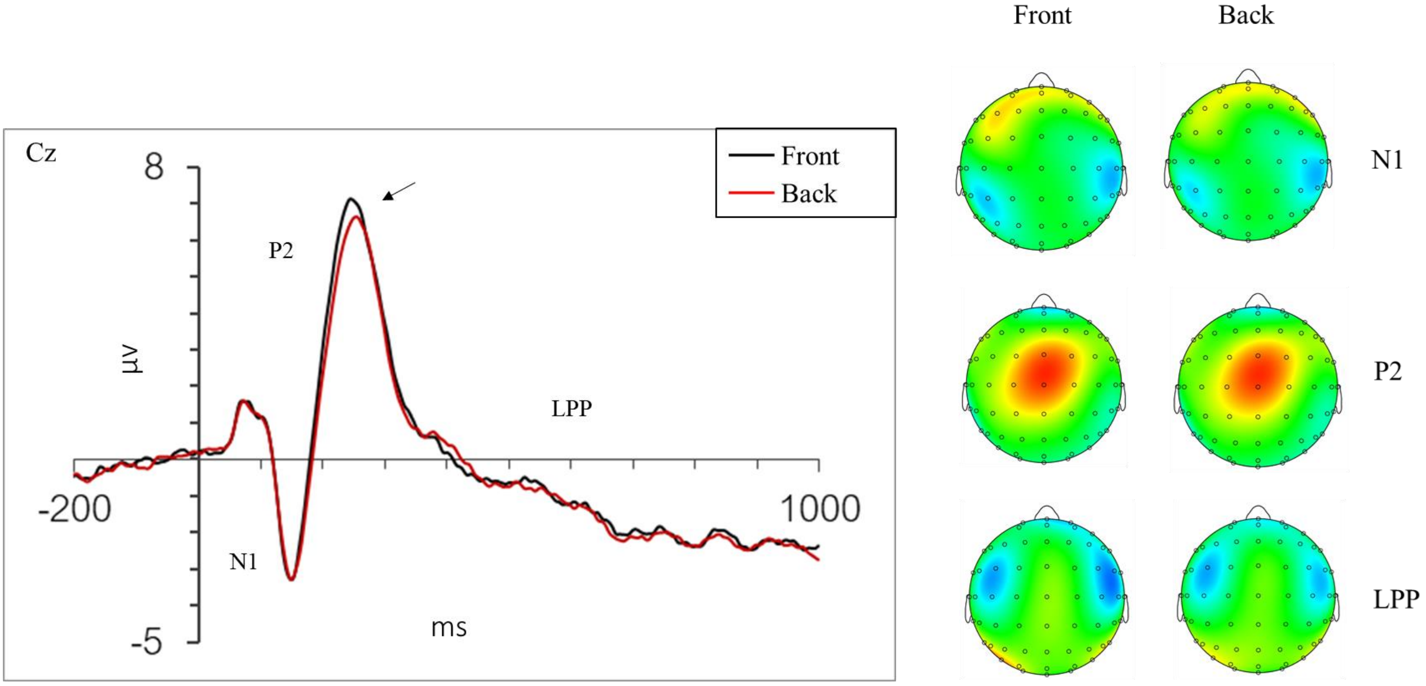
Mean amplitude of the LPP component per condition

Emotion	Space	LPP *M (SD)*
Neutral	Front	-2.496 (0.357)
	Back	-2.396 (0.280)
	Right	-2.556 (0.324)
	Left	-2.551 (0.301)
Amusement	Front	-1.190 (0.316)
	Back	-1.267 (0.322)
	Right	-1.085 (0.333)
	Left	-1.225 (0.325)
Anger	Front	-0.774 (0.380)
	Back	-0.634 (0.334)
	Front	-0.774 (0.380)
	Back	-0.634 (0.334)

Note. M = Mean; SD = Standard Deviation.

Figure 5.

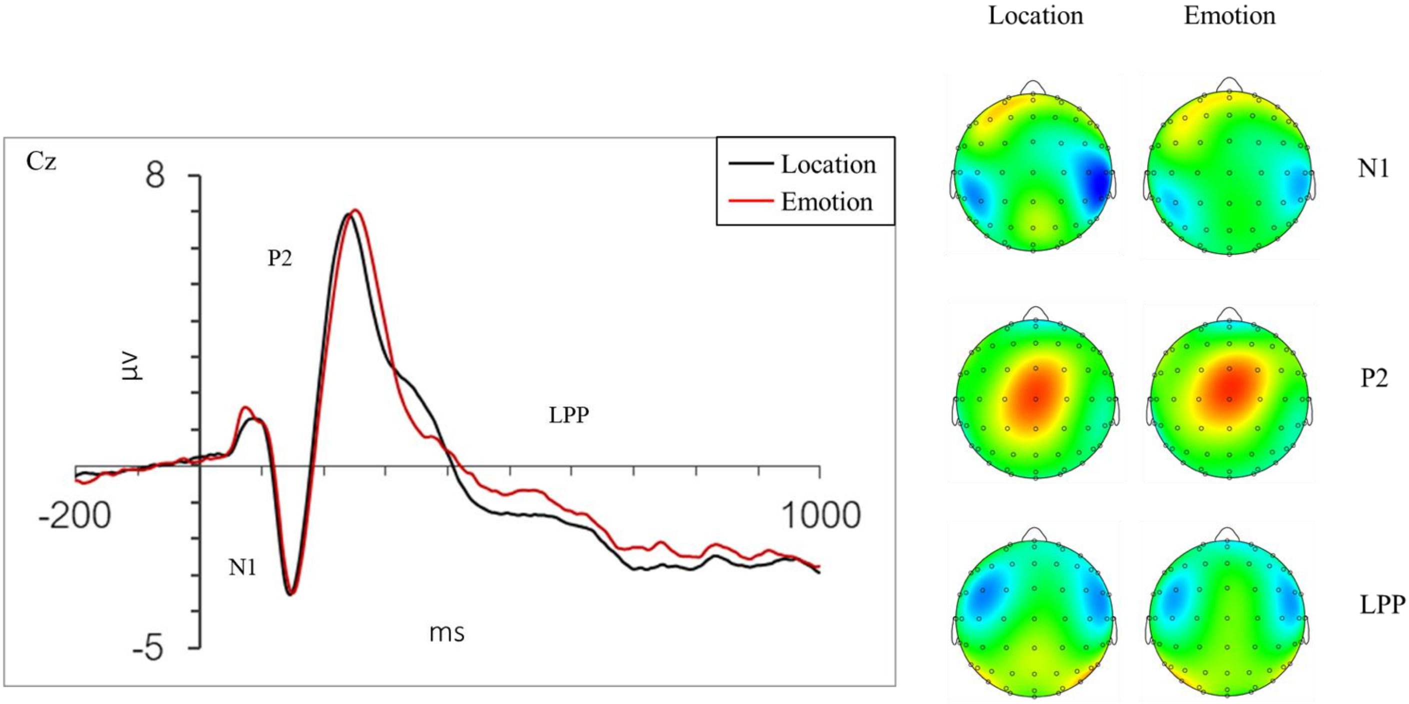
Grand average waveforms contrasting front vs. back spaces in the emotion recognition task. Topographic maps of the N1(110-190ms), P2 (200-280ms) and LPP (450-700ms) components.



Note. Images shown over the electrode Cz. The arrow points to the statistically significant effect.

Figure 6.

Grand average waveforms contrasting location discrimination and emotion recognition tasks. Topographic maps of the N1(110-190ms), P2 (200-280ms) and LPP (450-700ms) components.



Note. Images shown over the electrode Cz.

3.1.2. Right-Left Axis

3.1.2.1. N1

There was a significant main effect of emotion on N1 amplitude [$F(2, 80) = 14.783$; $p<.001$, *partial* $\eta^2 = .270$] (see Figure 3). Emotional vocalizations elicited decreased amplitudes compared to neutral ones (*ps*≤.004). There were no differences between amused and angry ($p = 1$) vocalizations (see Table 1). There main effect of Task was nonsignificant [$F(1,40) < .001$, $p = .994$] (see Figure 6).

3.1.2.2. P2

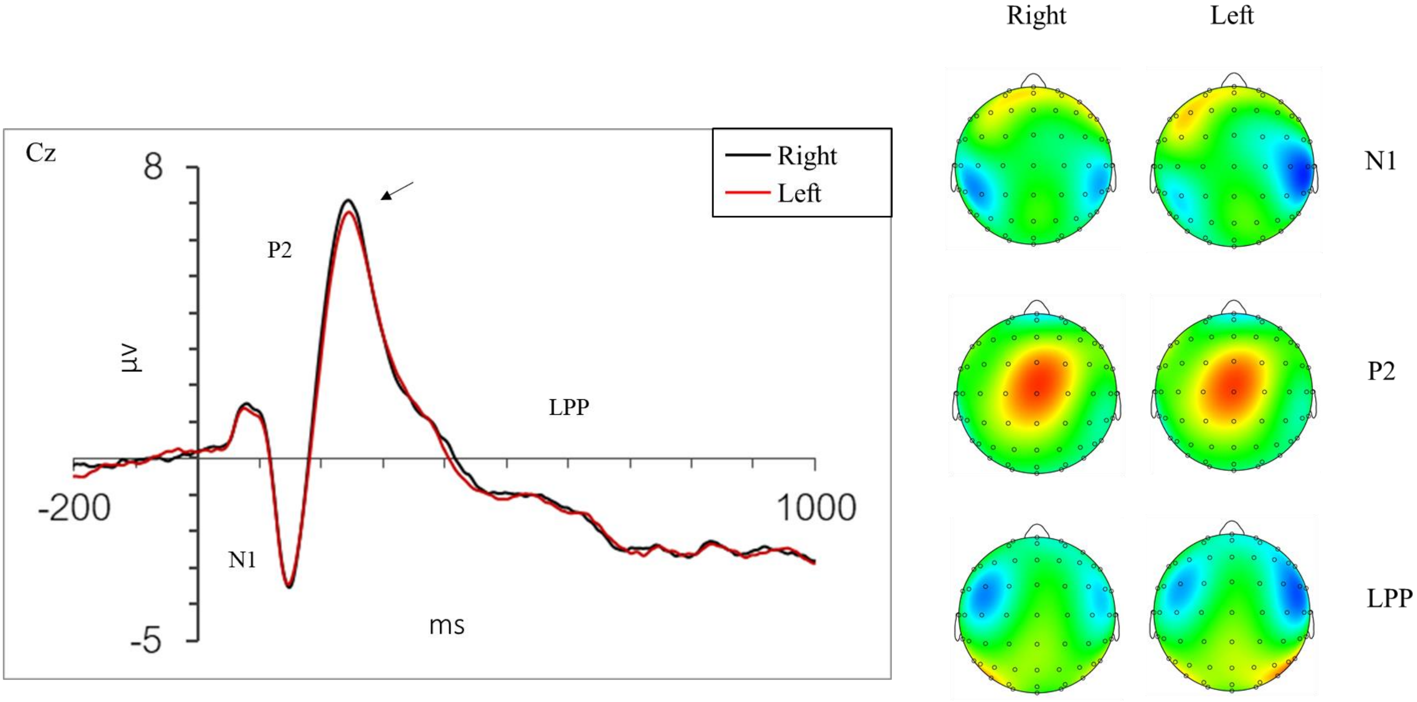
There was a significant main effect of emotion on the P2 amplitude [$F(2, 80) = 14.783$; $p<.001$, *partial* $\eta^2 = .270$] (see Figure 3). Emotional vocalizations elicited increased amplitudes compared to neutral ones (*ps*<.001), with no differences between amused and angry ($p = 1$) vocalizations (see Table 2). The main effect of space on P2 amplitude was also significant [$F(1, 40) = 8.753$; $p = .005$, *partial* $\eta^2 =.180$], indicating that voices coming from the right side elicited an increased P2 amplitude compared to voices presented on the left (see Figure 7). There was no main effect of Task [$F(1,40) = .314$; $p=.578$] (see Figure 6).

3.1.2.3. LPP

There was a significant main effect of emotion on the LPP amplitude, [$F(1.576, 63.024) = 65.612$; $p<.001$, *partial* $\eta^2 = .621$] (see Figure 3). Emotional vocalizations elicit an increased LPP amplitude compared to neutral ones (*ps*<.001), and angry vocalizations also resulted in increased LPP amplitudes compared to amused ($p = .004$) vocalizations (see Table 3). There was no main effect of Task [$F(1,40) = .418$; $p = .522$] (see Figure 6).

Figure 7.

Grand average waveforms contrasting voices presented on the right vs. left side. Topographic maps of the N1(110-190ms), P2 (200-280ms) and LPP (450-700ms) components.



Note. Images shown over the electrode Cz. The arrow points to the statistically significant effect.

3.1.3. Summary of ERP Results:

Emotional vocalizations elicited increased P2 and LPP amplitudes compared to neutral ones, and angry vocalizations elicited an increased LPP amplitude compared to voices portraying amusement. Emotional voices also elicited decreased N1 amplitudes compared to neutral ones, though only in the Right-Left axis. In the Front-Back axis, there was a significant interaction between space and task on P2 amplitude, revealing increased amplitudes to voices presented on the front (*vs*. back) space, but only in the emotion recognition task. In the Right-Left axis, there was a significant main effect of space on P2 amplitude, indicating that voices presented on the right side elicited an increased P2 amplitude compared to voices presented on the left. There were no significant main effects of task, nor any interactions between emotion and space on the amplitude of either ERP component (see Figure 8 for a schematic illustration of ERP results).

3.2. Behavioral results

Accuracy in the location discrimination and emotion recognition tasks are shown in Tables 4 and 5, respectively. Valence and arousal ratings are presented in Tables 6 and 7, respectively.

3.2.1. Front-Back Axis

3.2.1.1. Location Discrimination Task

There was a significant main effect of emotion on location accuracy [$F(2, 38) = 6.055$; $p = .005$, *partial* $\eta^2 = .242$] (see Table 4). Amused vocalizations were more accurately located than both neutral ($p = .022$) and angry ($p = .019$) vocalizations, with no differences between the latter two ($p = 1$). There was a significant main effect of space on location accuracy [$F(1, 19) = 11.687$; $p = .003$, *partial* $\eta^2 = .381$], indicating that vocalizations coming from the back were more accurately located than those presented in the front.

The interaction between emotion and space was non-significant [$F(1.452, 27.584) = 2.778$; $p = .094$, *partial* $\eta^2 = .128$].

3.2.1.2. Emotion Recognition Task

There were no significant main effects of emotion [$F(2, 42) = .606$; $p = .550$] or space [$F(1, 21) = .062$; $p = .806$] on emotion recognition (see Table 5). The interaction between emotion and space was non-significant as well [$F(2, 42) = 2.132$; $p = .131$].

Table 4

Mean percentage of correct responses in the location discrimination task

Emotion	Space	Correct Responses *M (SD)*
Neutral	Front	27.5 (4.2)
	Back	40.9 (3.9)
	Right	41.6 (2.2)
	Left	46.6 (2.5)
Amusement	Front	27.5 (3.9)
	Back	47.4 (5.1)
	Right	45.5 (2.6)
	Left	48.3 (3.3)
Anger	Front	22.5 (3.3)
	Back	43.9 (4.5)
	Right	41.0 (2.6)
	Left	44.4 (2.7)

Note. M = Mean; SD = Standard Deviation.

Table 5

Mean percentage of correct responses in the emotion recognition task

Emotion	Space	Correct Responses *M (SD)*
Neutral	Front	98.0 (0.6)
	Back	98.6 (0.5)
	Right	98.0 (0.5)
	Left	97.7 (0.6)
Amusement	Front	99.0 (0.3)
	Back	99.0 (0.4)
	Right	99.0 (0.3)
	Left	99.2 (0.2)
Anger	Front	99.0 (0.5)
	Back	98.5 (0.6)
	Right	98.3 (0.7)
	Left	99.1 (0.3)

Note. M = Mean; SD = Standard Deviation.

3.2.1.3. Valence Ratings

There was a significant main effect of emotion on valence ratings [$F(1.274, 42.039) = 441.884$; $p < .001$ *partial* $\eta^2 = .931$] (see Table 6). Amused vocalizations were considered more positive than both neutral ($p<.001$) and angry ($p<.001$) voices; and vocalizations portraying anger were considered more negative than neutral ones ($p<.001$). There was a significant interaction between emotion and space [$F(1.537, 50.732) = 8.829$; $p = .001$ *partial* $\eta^2 = .211$]. Angry vocalizations were considered more positive when presented from back compared to front locations ($p = .022$); however, no such differences were observed for amused ($p = .092$) or neutral ($p = .178$) vocalizations.

The main effect of space was not significant [$F(1, 33) = 3.894$; $p = .057$, *partial* $\eta^2 = .106$].

Table 6

Ratings of Valence per condition

Emotion	Space	Valence *M (SD)*
Neutral	Front	4.99 (0.03)
	Back	5.02 (0.02)
	Right	5.00 (0.03)
	Left	5.01 (0.02)
Amusement	Front	7.21 (0.16)
	Back	7.14 (0.16)
	Right	7.22 (0.16)
	Left	7.15 (0.16)
Anger	Front	1.91 (0.11)
	Back	2.06 (0.13)
	Right	1.98 (0.12)
	Left	1.94 (0.12)

Note. M = Mean; SD = Standard Deviation.

3.2.1.4. Arousal Ratings

There was a significant main effect of emotion on arousal ratings [F(1.520, 50.165) = 64.965; p<.001, *partial* η^2 = .663] (see Table 7). Angry vocalizations were considered more arousing than both neutral (p<.001) and amused (p = .001) ones; and voices portraying amusement were also considered more arousing than neutral ones (p<.001). There was a significant main effect of space on arousal ratings [F(1, 33) = 28.908; p<.001, *partial* η^2 = .467], indicating that vocalizations coming from the front were considered more arousing than those coming from the back.

The interaction between emotion and space was non-significant [F(2, 66) = .952; p = .391, *partial* η^2 = .028].

Table 7

Ratings of Arousal per condition

Emotion	Space	Arousal *M (SD)*
Neutral	Front	3.39 (0.26)
	Back	3.31 (0.27)
	Right	3.33 (0.26)
	Left	3.36 (0.26)
Amusement	Front	5.96 (0.31)
	Back	5.79 (0.32)
	Right	5.92 (0.32)
	Left	5.86 (0.32)
Anger	Front	7.25 (0.17)
	Back	7.14 (0.18)
	Right	7.18 (0.18)
	Left	7.27 (0.18)

Note. M = Mean; SD = Standard Deviation.

3.2.2. Right-Left Axis

3.2.2.1. Location Discrimination Task

There was a significant main effect of emotion on location accuracy [$F(2, 38) = 12.827$; $p<.001$ *partial* $\eta^2 = .403$] (see Table 4). Amused vocalizations were more accurately located than both neutral ($p = .013$) and angry ($p = .001$) vocalizations, with no differences between angry and neutral stimuli ($p = .213$). There was a significant main effect of space on location accuracy [$F(1, 19) = 6.717$; $p = .018$, *partial* $\eta^2 = .261$], indicating that vocalizations coming from the left were more accurately located than those presented on the right side.

The interaction between emotion and space was non-significant [$F(2, 38) = .651$; $p = .527$, *partial* $\eta^2 = .033$].

3.2.2.2. Emotion Recognition Task

There were no significant main effects of emotion [$F(2, 42) = 2.778$; $p = .074$] or space [$F(1, 21) = 1.456$; $p = .241$] on emotion recognition (see Table 5). The interaction between emotion and space was non-significant as well [$F(1.564, 32.850) = 1.740$; $p = .196$].

3.2.2.3. Valence Ratings

There was a significant main effect of emotion on valence ratings [$F(1.282, 42.314) = 440.917$; $p<.001$ *partial* $\eta^2 = .930$] (see Table 6). Amused vocalizations were considered more positive than both neutral ($p<.001$) and angry ($p<.001$) ones; vocalizations portraying anger were considered more negative than neutral ones ($p<.001$).

There was no main effect of space [$F(1, 33) = 2.884$; $p = .099$, *partial* $\eta^2 = .080$], nor any interaction between emotion and space [$F(1.659, 54.757) = 3.086$; $p = .063$ *partial* $\eta^2 = .086$].

3.2.2.4. Arousal Ratings

There was a significant main effect of emotion on stimulus' arousal [$F(1.521, 50.199) = 66.251$; $p<.001$, *partial* $\eta^2 = .668$] (see Table 7). Angry voices were considered more arousing than both neutral ($p<.001$), and amused ($p = .001$) voices; and voices portraying amusement were considered more arousing than neutral ones ($p<.001$). There was a significant interaction between emotion and space [$F(1.666, 54.992) = 3.849$; $p = .034$, *partial* $\eta^2 = .104$], revealing that angry voices were considered more arousing when presented on the left than on the right side ($p = .044$); however, there were no differences between right and left in arousal ratings of amused ($p = .292$) and neutral ($p = .438$) voices.

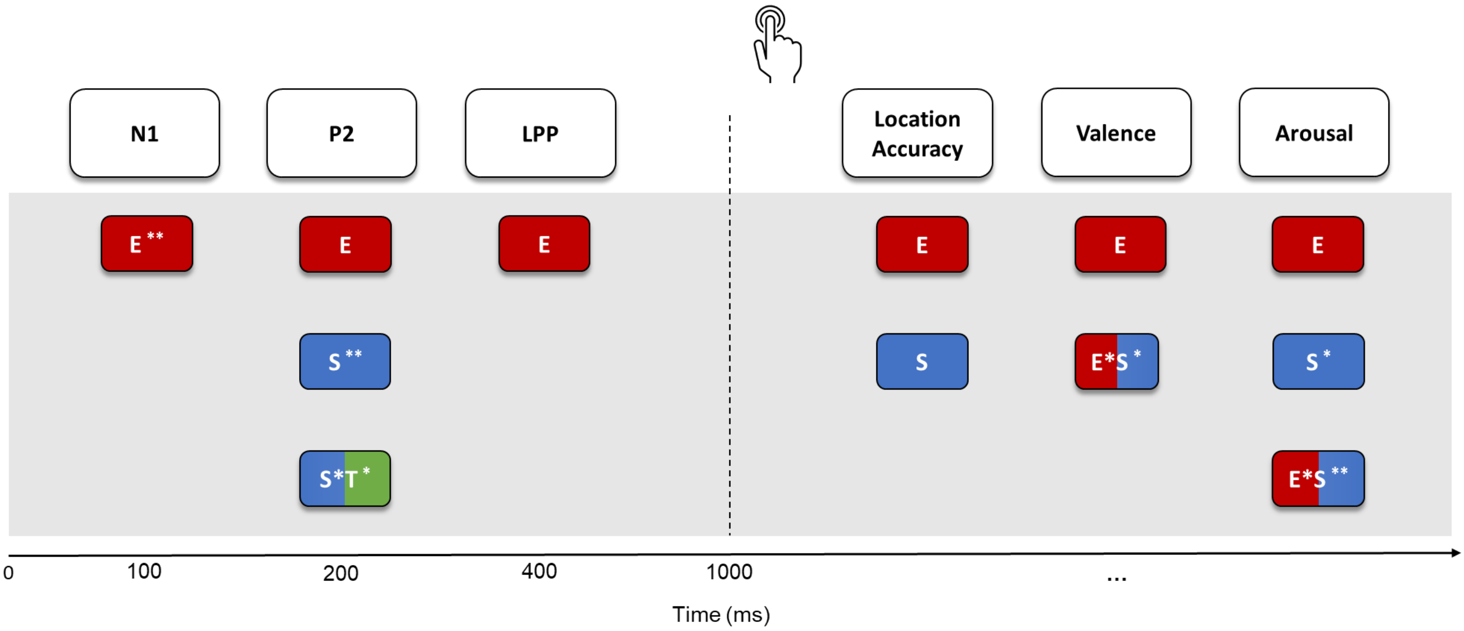
The main effect of space was not significant [$F(1, 33) = .354$; $p = .556$, *partial* $\eta^2 = .011$].

3.2.3. Summary of Behavioral Results:

Location accuracy was higher for voices portraying amusement compared to neutral and angry stimuli, and for voices presented on the back (*vs*. front) or on the left side (*vs*. right). Emotion recognition accuracy was high regardless of the emotion carried by the voice, spatial location, or the interaction between them. Valence ratings revealed that voices portraying amusement were the most pleasant, while voices portraying anger were the least pleasant. Arousal ratings were higher for emotional *vs*. neutral vocalizations, and specifically for angry *vs*. amused vocalizations. Voices coming from the front were also considered more arousing than voices presented on the back space. Finally, angry stimuli were considered less positive when presented from the front (*vs*. back), and more arousing when presented on the left (*vs*. right) side (see Figure 8 for a schematic illustration of behavioural results).

Figure 8.

Schematic illustration of the main ERP and behavioral results.



Note. Electrophysiological data were analyzed until 1000ms after stimulus onset and behavioral data, signaled by the keypress image, was collected from 1000ms onwards.

With regards to emotion effects, they first emerge in the N1 timeframe, persisting throughout the P2 and LPP timeframes and the 1000ms after stimulus onset, as indicated by the behavioral measures. Space effects emerge later, within the P2 timeframe, and are also observed after 1000ms in location accuracy and arousal ratings. Interactions between space and task also occur in the P2 timeframe. Finally, interactions between emotion and space are only observed behaviorally, from the 1000ms after stimulus onset, in affective ratings.

E = main effect of emotion; S = main effect of space; S*T = interaction between space and task; E*S = interaction between emotion and space. * Front-Back axis, ** Right-Left axis.

4. Discussion

The current study explored the time course of interactions between emotional and spatial cues during voice perception, as well as the role of attentional focus on these processes. We analysed the amplitude of three ERP components (i.e., N1, P2 and LPP) elicited in response to spatialized nonverbal emotional vocalizations, while participants performed either a location discrimination (Task 1) or an emotion recognition (Task 2) task. We expected emotional vocalizations to elicit decreased N1 and increased P2 and LPP amplitudes, compared to neutral voices. Secondly, we predicted that emotional voices would elicit increased P2 and LPP amplitudes when presented from the back (*vs.*

front) space. Lastly, we hypothesized differences in the amplitudes of the N1, P2, and LPP components in response to the same spatialized emotional vocalizations as a function of attention focus. Confirming our first hypothesis, emotion successfully modulated the amplitudes of all three components. However, contrary to our expectations, in the Front-Back axis we did not find an interaction between emotion and space on the P2 and LPP amplitudes. Finally, partially confirming our third hypothesis, we observed an interaction between space and task in the Front-Back axis, but only in the P2 timeframe.

Emotion effects

Emotional vocalizations elicited increased P2 and LPP amplitudes, compared to neutral stimuli, and angry voices resulted in an additional increase in LPP amplitude compared to amused vocalizations. Emotional voices also resulted in a reduced N1 amplitude compared to neutral stimuli. However, the effects on N1 amplitude were only significant in the Right-Left axis, suggesting an increased sensitivity to the emotional quality of the voice as a function of hemisphere. Vocal emotion effects on the N1, P2, and LPP have been extensively reported (Kotz & Paulmann, 2011; Liu et al., 2012; Paulmann et al., 2013; Pell et al., 2015). Modulations of the N1 amplitude by vocal emotion have been linked to an early processing of basic acoustic features, whereas for the P2 component these effects are usually interpreted as a result of emotional salience detection processes (Liu et al., 2012; Paulmann & Kotz, 2008). The N1 and P2 components are usually insensitive to stimulus valence (e.g., Liu et al., 2012). Later stages of stimulus evaluation occur within the LPP timeframe (Kotz & Paulmann, 2011). The increase in LPP amplitude to angry (*vs*. amused) voices is consistent with previous literature (e.g., Pell et al., 2015), revealing a differentiation between stimulus valences. Further, this finding also highlights the impact of angry vocal cues on cognitive and attentional processes, confirming that these cues engage in a more exhaustive emotional meaning evaluation, as they signal a potential threat to the individual (Pell et al., 2015). These findings reveal a dissociation between the timing of arousal and valence effects, as initial components (i.e., N1 and P2) show a general discrimination between emotional and neutral cues (i.e., arousal effects), whereas the distinction between amused and angry voices (i.e., valence effects) are only observed later, i.e. within the LPP timeframe.

Regarding the behavioral results, in the location discrimination task, amused voices were more accurately located than both neutral and angry stimuli, with no differences between the latter two. Using the same paradigm, Pinheiro et al. (2019) found no main effect of emotion on location accuracy. However, the authors report a slight advantage for positive vocalizations in the emotion recognition task (Pinheiro et al., 2019). Positive vocal cues are especially salient due to their relevance in social interactions and social bonding (Pinheiro et al., 2017; Vasconcelos et al., 2017). Thus, the relevance of this type of vocal signals might explain higher location accuracy for amused vocalizations, given their importance for social interactions. In the emotion recognition task, accuracy was high regardless of the emotional content of the voice, spatial location, or interaction between them. Overall, these results are in line with the findings of Pinheiro et al. (2019), and reinforce the automaticity of emotional processing, which is independent of attention focus and robust to spatial manipulations (Lima et al., 2019; Pinheiro et al., 2019).

Taken together, these findings provide further support to the automatic nature of vocal emotion processing, as emotion effects are evident even when attention is directed away from these vocal cues (i.e., location discrimination task).

Space effects

There was a significant main effect of space on the Right-Left axis, revealing an increased P2 amplitude to voices presented from the right *vs*. left side. The P2 component appears to be modulated by attention (Crowley & Colrain, 2004), as its amplitude tends to increase when attention is directed towards the stimulus (Liu et al., 2015, 2018). Thus, the present finding could indicate that voices coming from the right side were more salient, leading to an enhanced allocation of attentional resources to the right space. Behavioral studies have suggested that people implicitly associate positive valence with the side in each they interact with the environment more fluently with their dominant hand (Casasanto, 2009). Since our participants were right-handed, stimuli presented on their right side (i.e., participants' dominant side) might have been implicitly regarded as more pleasant and salient. Nevertheless, this interpretation is highly speculative.

Importantly, the present findings reveal a dissociation between the timing of emotional and spatial effects: whereas emotion modulations appear earlier, around 100ms after stimulus onset, spatial modulations are only seen around 200ms post stimulus onset. Thus, emotional and spatial manipulations affect different stages of voice perception.

Behaviorally, voices coming from the back were more accurately located than voices presented in the front, as in Pinheiro et al. (2019), and voices presented on the left side were also more accurately located than those presented on the right side, unlike what was found by Pinheiro et al. (2019). Increased accuracy for the back space is unlikely explained by an attentional bias towards the rear, as affective ratings revealed that stimuli presented from the front were considered more arousing than those occurring behind participants. Instead, they might reflect a tendency to respond to the back in case of uncertainty, as previously suggested (Asutay & Västfjäll, 2015). Higher location accuracy for the left *vs*. right space has been previously reported (Burke et al., 1994; Hirnstein et al., 2006), and is usually interpreted as reflecting an advantage of the right hemisphere for spatial processing. Regarding affective ratings, as previously mentioned, vocalizations were considered more arousing when presented in the front *vs*. back space. Although contrary to our expectations, this result is in line with previous reports (Pinheiro et al., 2019). The higher arousal of front locations may be tied to the relevance of the front space in our daily lives, as most of our interactions occur in the front space, where perceptual cues from all sensory modalities are available. However, this interpretation is also highly speculative.

Interactions between emotion and space

We expected emotional voices to elicit an additional increase in P2 and LPP amplitudes when presented from back (*vs*. front) locations. Contrary to our predictions, we observed no such differences, suggesting that emotional voices were processed in a similar manner whether they were presented in front or back locations. This hypothesis was mostly based on evidence from behavioral studies. For instance, Pinheiro et al. (2019) found an interaction between emotion and space on voice location performance, as location accuracy was higher for emotional voices presented on back *vs*. front locations. Other studies suggested that sound source location conveys emotional salience (Asutay & Västfjäll, 2015), and that sounds in the rear space tend to be considered more arousing (e.g., Tajadura-Jiménez et al., 2010). Contrary to what was expected, in the current study, voices presented in the front were considered more arousing than voices presented from the back. Therefore, if the back space did not confer any additional salience to the stimulus, emotional voices would not be more salient when presented from the back, thus not resulting in increased P2 and LPP amplitudes, when compared to front locations.

Concerning behavioral results, we did not replicate the previously reported interaction effects between emotion and space in the location discrimination task (Pinheiro et al., 2019). This discrepancy may arise from methodological differences between both studies. For instance, the data were collected from different samples; the statistical strategies used for data analysis were different; and we employed a between-subjects design, instead of the within-subjects design adopted in Pinheiro et al. (2019). Nonetheless, we did find significant interactions between emotion and space on affective ratings as angry vocalizations were considered more pleasant when presented from the back (*vs*. front), and more arousing when presented on the left (*vs*. right) side. As previously mentioned, in the present study, the back space did not seem to confer any additional salience to the stimulus, as vocalizations presented in front locations were considered more arousing than those coming from the back. Thus, it is plausible that angry voices might have been considered more negative when presented in highly arousing locations (i.e., front space). On the other hand, as previously mentioned, behavioral studies suggest that people implicitly associate positive valence with their dominant side (Casasanto, 2009). Since our participants were right-handed, angry stimuli may pose a greater threat when presented on their left side (i.e., nondominant side), where participants cannot act as fluently with their dominant hand and less able to defend themselves. Nevertheless, these explanations are speculative.

Notably, the findings discussed so far have revealed important discrepancies between electrophysiological and behavioral measures. For instance, electrophysiological findings revealed increased LPP amplitudes to angry *vs*. amused voices; increased P2 amplitudes to voices presented on the right *vs*. left side; and no evidence of interactions between emotion and space on the amplitude of either ERP component. However, behaviorally, we observed higher location accuracy for amused voices compared to both angry and neutral stimuli; higher location accuracy for voices presented on the left *vs*. right side; and significant interactions between emotion and space on affective ratings. Although these findings seem contradictory, they, instead, reveal an important dissociation between the phenomena captured by electrophysiological and behavioral measures: ERP components index perceptual stages related to the processing of stimulus features, whereas behavioral measures capture response processes, occurring during postperceptual stages (Luck, 2005). Therefore, the discrepancies observed between electrophysiological and behavioral measures arise from the different processes indexed by each measure.

Task effects

We expected attention focus to modulate the amplitudes of the ERP components under study, by promoting the allocation of attentional resources towards distinct vocal features (i.e., emotional or spatial cues). We did not find any main effect of task nor any significant interaction between emotion and task, suggesting that the emotional quality of the voice was processed in a similar manner whether these cues were explicit or implicitly evaluated. These results provide further support to the automatic nature of emotional processing (e.g., Lima et al., 2019; Pinheiro et al., 2017), as emotion effects emerged even when emotional cues were task-irrelevant and attention was directed away from them. Nevertheless, we found a significant interaction between space and task on P2 amplitude, suggesting that processing of spatial information was modulated by attention focus. Specifically, we observed an increased P2 amplitude to voices presented in front (*vs*. back) locations in the emotion recognition task, with no differences between front and back spaces in the location discrimination task. Given that the P2 component seems to be sensitive to stimulus salience (Liu et al., 2012), the present interaction effect suggests that, when participants explicitly attend to emotional cues, voices presented from the front were more salient than those presented from behind. However, when participants focused on voice location, front and back spaces were equally salient. Even though attention focus had a modulatory effect on spatial processing, these effects appear to depend on the salience conferred by voice location, as this interaction was only observed in the P2 timeframe. Since the N1 component reflects an early processing of basic acoustic features (Kotz & Paulmann, 2011; Liu et al., 2012) and, in the present study, spatial effects were only observed in later processing stages (i.e., P2 timeframe), it is plausible that the processes indexed by the N1 component may be more automatic in nature, and not yet sensitive to the effects of space and task. On the other hand, the LPP may be insensitive to the effects of space and task, as it reflects processes related to the evaluation of emotional significance of the stimulus (Kotz & Paulmann, 2011; Pell et al., 2015).

Limitations and future directions

One weakness of the present study was voice location difficulty, as suggested by the low accuracy in the location discrimination task (see Table 4). Analyzing only trials associated with correct responses, in this task, would result in a significant decrease in signal to noise ratio, affecting the quality of the data included in the statistical analysis. Thus, to preserve the signal to noise ratio and the quality of the data, we decided to which include trials associated with both correct and incorrect responses on the ERP analysis of the location discrimination task. On the other hand, location difficulty also resulted in discrepancies between the demands of location discrimination and emotion recognition tasks. Previous evidence suggests that location difficulty can affect auditory processing. For instance, in a study by Koiwa et al. (2010), the manipulation of sound location difficulty affected both early and late auditory processing stages. Specifically, compared to passive listening, different potentials were elicited in the early processing stages on the easy task, and in later stages of the difficult task (Koiwa et al., 2010). Importantly, the authors also found that when task difficulty was high, the pattern of results resembled more closely passive listening than easy location discrimination (Koiwa et al., 2010). Therefore, future studies should revisit the impact of attentional focus on emotional and spatial vocal processing, employing paradigms with comparable task demands, to control for the effects of task difficulty.

Another limitation of the present study is the implementation of a between-subjects design. A within-subjects design might have been more appropriate to assess the impact of attention focus on the voice perception, allowing a more direct comparison of how the same vocal stimuli are processed under different attentional requirements, accounting for interindividual differences.

On the other hand, most reports of emotional and spatial interactions on the P2 and LPP components come from studies in the visual modality (P2: Xie et al., 2014; LPP: Du et al., 2017). Therefore, more electrophysiological studies are needed to address interactions between emotional and spatial vocal features. At last, future studies should address the relationship between the salience, attention and the Right-Left axis, employing different paradigms (e.g., dichotic listening), since, as far as we are aware, there is no available literature comparing the differences between the right and left spaces using ERP.

Conclusions

The current study investigated the electrophysiological correlates of emotion and space interactions during voice perception, when attention was focused on spatial location (Task 1) or vocal emotion recognition (Task 2). Confirming our first hypothesis, emotion modulated the amplitudes of all three components. Compared to neutral stimuli, emotional voices elicited decreased N1 and increased P2 and LPP amplitudes, even though the effects on N1 amplitude were only significant in the Right-Left axis. Contrary to our expectations, there were no differences between P2 and LPP amplitudes elicited by emotional voices presented on front and back locations. Furthermore, we did not observe any main effect of task nor any interaction between emotion and task on the ERP components under study. Nevertheless, we found a significant interaction between space and task on P2 amplitude, as voices presented in the front (*vs*. back) elicited increased amplitudes, but only when emotion was explicitly processed. This finding suggests that when attention was directed towards the emotional quality of the voice, front locations were more salient. We also observed a significant main effect of space on the P2 timeframe, showing enhanced amplitudes to vocalizations presented on the right *vs*. left side. Taken together, these findings reveal dissociation in the timing of emotional and spatial effects during voice perception: whereas emotion modulations appear earlier, around 100ms after stimulus onset, spatial modulations are only seen around 200ms post stimulus onset. Moreover, the results from the present study confirm that emotional processing is robust to spatial manipulations, further highlighting the automatic nature of vocal emotion processing, as modulations of the N1, P2 and LPP amplitudes are observed even when attention is directed away from the emotional quality of the voice and those cues are irrelevant to the task. On the other hand, spatial effects during voice perception seem to arise from the salience conferred by voice location, and are modulated by attention focus.

5. References:

Ahveninen, J., Jääskeläinen, I. P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., Levänen, S., Lin, F. H., Sams, M., Shinn-Cunningham, B. G., Witzel, T., & Belliveau, J. W. (2006). Task-modulated “what” and “where” pathways in human auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.0510480103

Amorim, M., & Pinheiro, A. P. (2019). Is the sunny side up and the dark side down? Effects of stimulus type and valence on a spatial detection task. *Cognition and Emotion*. https://doi.org/10.1080/02699931.2018.1452718

Anourova, I., Nikouline, V. V., Ilmoniemi, R. J., Hotta, J., Aronen, H. J., & Carlson, S. (2001). Evidence for dissociation of spatial and nonspatial auditory information processing. *NeuroImage*. https://doi.org/10.1006/nimg.2001.0903

Asutay, E., & Västfjäll, D. (2015). Attentional and emotional prioritization of the sounds occurring outside the visual field. *Emotion*. https://doi.org/10.1037/emo0000045

Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: Neural correlates of voice perception. In *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2004.01.008

Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods*. https://doi.org/10.3758/BRM.40.2.531

Belin, P., Zatorre, R. J., Lafallie, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*. https://doi.org/10.1038/35002078

Bertels, J., Kolinsky, R., Coucke, D., & Morais, J. (2013). When a bang makes you run away: Spatial avoidance of threatening environmental sounds. *Neuroscience Letters*. https://doi.org/10.1016/j.neulet.2012.12.058

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*. https://doi.org/10.1016/0005-7916(94)90063-9

Burke, K. A., Letsos, A., & Butler, R. A. (1994). Asymmetric performances in binaural localization of sound in space. *Neuropsychologia*. https://doi.org/10.1016/0028-3932(94)00074-3

Campbell, J. I. D., & Thompson, V. A. (2012). MorePower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis. *Behavior Research Methods*. https://doi.org/10.3758/s13428-012-0186-0

Canavarro, M. C. (1999). Inventário de Sintomas Psicopatológicos [Brief Symptoms Inventory]: BSI. In R. S. Mário, S. A. Leandro, & M. G. Miguel (Eds.), *Testes e provas psicológicas em Portugal: Vol. II* (pp. 95–109). Braga: SHO/APPORT. https://doi.org/10.1590/s1413-85571999000200006

Casasanto, D. (2009). Embodiment of Abstract Concepts: Good and Bad in Right- and Left-Handers. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/a0015854

Ceravolo, L., Frühholz, S., & Grandjean, D. (2016a). Modulation of auditory spatial attention by angry prosody: An fMRI auditory dot-probe study. *Frontiers in Neuroscience*. https://doi.org/10.3389/fnins.2016.00216

Ceravolo, L., Frühholz, S., & Grandjean, D. (2016b). Proximal vocal threat recruits the right voice-sensitive auditory cortex. *Social Cognitive and Affective Neuroscience*. https://doi.org/10.1093/SCAN/NSW004

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality. In *Clinical Neurophysiology*. https://doi.org/10.1016/j.clinph.2003.11.021

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: Covariation with autonomic arousal and affective report. *Biological Psychology*. https://doi.org/10.1016/S0301-0511(99)00044-7

Derey, K., Rauschecker, J. P., Formisano, E., Valente, G., & de Gelder, B. (2017). Localization of complex sounds is modulated by behavioral relevance and sound category. *The Journal of the Acoustical Society of America*. https://doi.org/10.1121/1.5003779

Derogatis, L. R. (1993). BSI: brief symptom inventory. *Minneapolis: National Computer Systems*.

Du, F., Wang, X., Abrams, R. A., & Zhang, K. (2017). Emotional processing is enhanced in peri-hand space. *Cognition*. https://doi.org/10.1016/j.cognition.2017.04.009

Gädeke, J. C., Föcker, J., & Röder, B. (2013). Is the processing of affective prosody influenced by spatial attention? An ERP study. *BMC Neuroscience*. https://doi.org/10.1186/1471-2202-14-14

Galinha, I. C., & Pais-Ribeiro, J. L. (2012). Contribuição para o estudo da versão portuguesa da Positive and Negative Affect Schedule (PANAS): II – Estudo

Psicométrico. *Análise Psicológica*. https://doi.org/10.14417/ap.84

Getzmann, S., & Lewald, J. (2012). Cortical processing of change in sound location: Smooth motion versus discontinuous displacement. *Brain Research*. https://doi.org/10.1016/j.brainres.2012.05.033

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

Hirnstein, M., Hausmann, M., & Lewald, J. (2006). Functional cerebral asymmetry in auditory motion perception. *Laterality*. https://doi.org/10.1080/13576500600959247

Ho, H. T., Schröger, E., & Kotz, S. A. (2015). Selective attention modulates early human evoked potentials during emotional face–voice processing. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn_a_00734

Juslin, P. N., & Laukka, P. (2003). Communication of Emotions in Vocal Expression and Music Performance: Different Channels, Same Code? In *Psychological Bulletin*. https://doi.org/10.1037/0033-2909.129.5.770

Koiwa, N., Masaoka, Y., Kusumi, T., & Homma, I. (2010). Sound localization difficulty affects early and late processing of auditory spatial information: Investigation using the dipole tracing method. *Clinical Neurophysiology*. https://doi.org/10.1016/j.clinph.2010.03.016

Kotz, S. A., & Paulmann, S. (2011). Emotion, Language, and the Brain. In *Linguistics and Language Compass*. https://doi.org/10.1111/j.1749-818X.2010.00267.x

Kryklywy, J. H., Macpherson, E. A., Greening, S. G., & Mitchell, D. G. V. (2013).

Emotion modulates activity in the “what” but not “where” auditory processing pathway. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2013.05.051

Leavitt, M. V., Molholm, S., Gomez-Ramirez, M., & Foxe, J. J. (2011). “WHAT” and “WHERE” in auditory sensory processing: A high-density electrical mapping study of distinct neural processes underlying sound object recognition and sound localization. *Frontiers in Integrative Neuroscience*. https://doi.org/10.3389/fnint.2011.00023

Lewald, J., & Getzmann, S. (2011). When and where of auditory spatial processing in cortex: A novel approach using electrotomography. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0025146

Lima, C. F., Anikin, A., Monteiro, A. C., Scott, S. K., & Castro, S. L. (2019). Automaticity in the Recognition of Nonverbal Emotional Vocalizations. *Emotion*. https://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*. https://doi.org/10.3758/s13428-013-0324-3

Liu, T., Pinheiro, A. P., Deng, G., Nestor, P. G., McCarley, R. W., & Niznikiewicz, M. A. (2012). Electrophysiological insights into processing nonverbal emotional vocalizations. *NeuroReport*. https://doi.org/10.1097/WNR.0b013e32834ea757

Liu, Y., Fan, H., Li, J., Jones, J. A., Liu, P., Zhang, B., & Liu, H. (2018). Auditory-Motor Control of Vocal Production during Divided Attention: Behavioral and ERP Correlates. *Frontiers in Neuroscience*. https://doi.org/10.3389/fnins.2018.00113

Liu, Y., Hu, H., Jones, J. A., Guo, Z., Li, W., Chen, X., Liu, P., & Liu, H. (2015).

Selective and divided attention modulates auditory-vocal integration in the processing of pitch feedback errors. *European Journal of Neuroscience*. https://doi.org/10.1111/ejn.12949

Luck, S. J. (2005). An Introduction to Event-Related Potentials and Their Neural Origins. *An Introduction to the Event-Related Potential Technique*.

Middlebrooks, J. C. (2015). Sound localization. In *Handbook of Clinical Neurology*. https://doi.org/10.1016/B978-0-444-62630-1.00006-8

Mostardinha, A. R., Bártolo, A., Bonifácio, J., & Pereira, A. (2019). Validation of the alcohol, smoking and substance involvement screening test (ASSIST) among university students. *Acta Medica Portuguesa*. https://doi.org/10.20344/amp.10650

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

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

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

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

Pinheiro, A. P., Barros, C., Dias, M., & Kotz, S. A. (2017). Laughter catches attention!

Biological Psychology. https://doi.org/10.1016/j.biopsycho.2017.09.012

Pinheiro, A. P., Barros, C., & Pedrosa, J. (2015). Salience in a social landscape: Electrophysiological effects of task-irrelevant and infrequent vocal change. *Social Cognitive and Affective Neuroscience*. https://doi.org/10.1093/scan/nsv103

Pinheiro, A. P., Lima, D., Albuquerque, P. B., Anikin, A., & Lima, C. F. (2019). Spatial location and emotion modulate voice perception. *Cognition and Emotion*. https://doi.org/10.1080/02699931.2019.1586647

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*. https://doi.org/10.1016/j.bandl.2015.12.003

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. In *Nature Neuroscience*. https://doi.org/10.1038/nn.2331

Salminen, N. H., Takanen, M., Santala, O., Lamminsalo, J., Altoè, A., & Pulkki, V. (2015). Integrated processing of spatial cues in human auditory cortex. *Hearing Research*. https://doi.org/10.1016/j.heares.2015.06.006

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. In *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2005.11.009

Spreckelmeyer, K. N., Kutas, M., Urbach, T., Altenmüller, E., & Münte, T. F. (2009). Neural processing of vocal emotion and identity. *Brain and Cognition*. https://doi.org/10.1016/j.bandc.2008.06.003

Tajadura-Jiménez, A., Larsson, P., Väljamäe, A., Västfjäll, D., & Kleiner, M. (2010). When Room Size Matters: Acoustic Influences on Emotional Responses to Sounds. *Emotion*. https://doi.org/10.1037/a0018423

Valdés-Conroy, B., Sebastián, M., Hinojosa, J. A., Román, F. J., & Santaniello, G. (2014). A close look into the near/far space division: A real-distance ERP study. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2014.04.009

Van Strien, J. W., De Sonneville, L. M. J., & Franken, I. H. A. (2010). The late positive potential and explicit versus implicit processing of facial valence. *NeuroReport*. https://doi.org/10.1097/WNR.0b013e32833ab89e

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*. https://doi.org/10.1007/s10919-017-0253-4

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*. https://doi.org/10.1037//0022-3514.54.6.1063

WHO ASSIST Working Group. (2010). The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): development, reliability and feasibility. ….

Xie, J., Wang, R., & Chang, S. (2014). The mechanism of valence-space metaphors: ERP evidence for affective word processing. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0099479

6. Supplementary Material:

6.1. ERP Results involving the factor ROI

6.1.1. Front-Back Axis

6.1.1.1. N1

There was a significant main effect of ROI on N1 amplitude [$F(1.974, 78.966) = 9.060$; $p<.001$, *partial* $\eta^2 = .185$]. N1 amplitude was reduced on the Midline compared to the Medial Right region ($p=.010$), and reduced on the Medial Left region compared to all other regions ($p\leq .040$). The interaction between emotion and ROI was also significant [$F(3.825, 153) = 4.687$, $p=.002$, *partial* $\eta^2 = .105$]. For emotional vocalizations, N1 amplitude was reduced on the Medial Left region compared to both Lateral regions ($p\leq.006$), and on Midline compared to the Medial Right region ($p\leq.008$). N1 amplitude was decreased on the Medial Left *vs*. Medial Right region ($p\leq.003$), but only for neutral and amused vocalizations. The N1 amplitude elicited in response to neutral vocalizations was reduced on Lateral Right compared to Medial Right region ($p=.026$). For amused vocalizations, N1 amplitude was decreased for the Lateral Left compared to the Medial Right region ($p=.041$).

6.1.1.2. P2

There was a significant main effect of ROI on P2 amplitude [$F(2.312, 92.499)= 93.811$, *partial* $\eta^2 = .701$]. P2 amplitude was increased on the Midline compared to all other regions ($p<.001$), and on Medial regions compared to Lateral ones ($p\leq.011$). There was a significant interaction between emotion and ROI [$F(4.341, 173.623) = 20.722$; $p<.001$, *partial* $\eta^2 = .341$], indicating that P2 amplitude elicited to amused vocalizations was higher in the Lateral Left compared to Lateral Right region ($p=.010$); whereas for angry voices, P2 was increased in the Medial Right compared to Lateral Left region ($p<.001$).

6.1.1.3. LPP

There was a main effect of ROI on LPP amplitude [$F(2.429, 97.151)= 9.705$, $p<.001$, *partial* $\eta^2 = .195$]. LPP amplitude was increased on the Midline compared to the

Medial Left region (*p*= .043), on Lateral Right compared to Medial Right region (*p*= .019), and on the Lateral Left compared to both Medial regions ($p<.001$). The interaction between emotion and ROI was also significant [$F(5.282, 211.266) = 14.583$; $p<.001$, *partial* $\eta^2 = .267$]. For neutral and amused vocalizations, LPP amplitude was increased in the Lateral Left region compared to the Medial Right region ($p<.001$). LPP amplitude to neutral vocalizations was increased on the Lateral Right region compared to Midline ($p=.027$) and Medial Right region ($p<.001$); and on the Lateral left compared to Midline ($p<.001$). In response to amused vocalizations, LPP amplitude was increased in the Lateral Left region compared to the Lateral Right region ($p=.011$); whereas for angry vocalizations, LPP amplitude in the Midline was increased compared to both Medial regions ($p\leq .002$).

6.1.2. Right-Left Axis

6.1.2.1. N1

There was a significant main effect of ROI on N1 amplitude [$F(2.021, 80.823) = 8.731$; $p<.001$, *partial* $\eta^2 =.179$]. N1 amplitude was decreased on the Midline compared to the Medial Right region ($p=.022$), and on the Medial Left region compared to all other ROIs ($p\leq .038$). The interaction between emotion and ROI was also significant [$F(3.561, 142.436) = 5.034$; $p=.001$, *partial* $\eta^2 =.112$]. For emotional vocalizations, N1 amplitude was reduced on the Medial Left region compared to both Lateral regions ($p\leq.013$), and on Midline compared to the Medial Right region ($p\leq.008$). N1 amplitude was decreased on the Medial Left *vs*. Medial Right region ($p\leq.002$), but only for neutral and amused vocalizations. The N1 amplitude elicited in response to neutral vocalizations was reduced on Lateral Right compared to Medial Right region ($p=.011$). For amused vocalizations, N1 amplitude was decreased for the Lateral Left compared to the Medial Right region ($p=.037$). The interaction between space and ROI was also significant [$F(1.698, 67.933) = 61.433$; $p<.001$, *partial* $\eta^2 =.606$]. In right ROIs, N1 amplitude was reduced for vocalizations presented on the right side ($p<.001$), whereas in left ROIs, N1 amplitude was decreased for vocalizations presented on the left side ($p<.001$).

6.1.2.2. P2

The main effect of ROI was significant [$F(2.312, 92.499) = 93.811$, *partial* $\eta^2 = .701$], as P2 amplitude was increased on the Midline compared to all other regions ($p<.001$), and on medial regions compared to lateral ones ($p \leq .006$). The interaction between emotion and ROI was also significant [$F(3.901, 156.052) = 26.948$; $p<.001$, *partial* $\eta^2 = .403$]. For amused vocalizations, P2 amplitude was increased in the Lateral Left compared to Lateral Right region ($p=.003$); whereas for angry vocalizations, P2 amplitude was increased in the Medial Right compared to Medial Left region ($p<.001$). The interaction between space and ROI turned out significant as well [$F(1.796, 71.833) = 45.024$; $p<.001$, *partial* $\eta^2 = .530$]. P2 amplitude was increased to stimuli presented on the right side in Midline and in both right ROIs ($p \leq .020$); and increased to left presented stimuli on the Lateral Left region ($p=.002$). There was also an interaction between emotion, space, and ROI [$F(5.119, 204.770) = 2.285$; $p=.046$, *partial* $\eta^2 = .054$]. In the Midline, P2 amplitude was increased to neutral vocalizations presented on the right (*vs*. left) side ($p=.001$). In right ROIs, P2 amplitude was increased for right (*vs*. left) presented stimuli, regardless of the emotion being expressed ($p \leq .004$). In left ROIs, P2 amplitude to angry vocalizations was increased for left (*vs*. right) presented stimuli ($p \leq .013$). Additionally, in the Lateral Left region, P2 amplitude to amused stimuli was increased for left (*vs*. right) presented stimuli ($p=.022$). The interaction between space, ROI, and Task was also significant [$F(1.796, 71.833) = 5.892$; $p=.006$, *partial* $\eta^2 = .128$]. In the Midline, P2 amplitude was increased for right (*vs*. left) presented stimuli, but only in the emotion recognition task ($p=.006$). In right ROIs, P2 amplitude was increased for right (*vs*. left) presented stimuli in both tasks ($p<.001$), whereas in left ROIs, P2 amplitude was increased for left (*vs*. right) presented stimuli on both tasks as well ($p \leq .011$). Finally, the interaction between emotion, space, ROI and Task was also statistically significant [$F(5.119, 204.770) = 3.369$; $p=.006$, *partial* $\eta^2 = .078$]. In the Midline, P2 amplitude to neutral vocalizations was increased for right (*vs*. left) presented stimuli in both tasks ($p \leq .024$). In left ROIs, P2 amplitude to emotional vocalizations was increased for left (*vs*. right) presented stimuli ($p \leq .026$) in the location discrimination task. In the Medial Right region, P2 amplitude was increased to neutral vocalizations presented on right (*vs*. left) side in both tasks ($p \leq .009$), to amused vocalizations presented on right (*vs*. left) side in the location discrimination task ($p=.023$), and to angry vocalizations presented on right (*vs*. left) side in the emotion recognition task ($p=.010$). In the Lateral Right region, P2 amplitude was increased for right (*vs*. left) presented stimuli in both tasks, regardless of the emotion being expressed by the voice ($p \leq .020$).

6.1.2.3. LPP

The main effect of ROI was significant on LPP amplitude [$F(2.362, 94.464) = 9.810$; $p<.001$, *partial* $\eta^2 =.197$]. LPP amplitude was increased on the Midline compared to Medial Left region ($p=.014$), on the Medial Right compared to Lateral Right region ($p=.046$), and on Lateral Left region compared to both medial ROIs ($p<.001$). The interaction between emotion and ROI was also significant [$F(4.966, 198.656) = 20.116$; $p<.001$, *partial* $\eta^2 =.335$]. For emotional vocalizations, LPP amplitude was increased in the Midline compared to both medial regions ($p \leq .023$). For neutral and amused vocalizations, LPP amplitude was increased in the Lateral Left region compared to the medial right region ($p<.001$). LPP amplitude to amused vocalizations was increased in the Lateral Left region compared to the Lateral Right region ($p=.012$). In response to neutral vocalizations, LPP amplitude was increased on the Lateral Right region compared to Midline ($p=.016$) and Medial Right region ($p<.001$); and on the Lateral Left compared to Midline ($p<.001$).

6.2. Supplementary Tables

Supplementary Table 1.

Participants' sociodemographic characteristics

		Total Sample	Location Discrimination Task	Emotion Recognition Task	t, p
		Mean (SD)	*Mean (SD)*	*Mean (SD)*	
N		42	20	22	n.a.
Sex		21 Male	10 Male	11 Male	n.a.
Age (years)		22.40 (2.63)	22.30 (2.62)	22.50 (2.70)	-0.243
PANAS	Positive Affect	24.45 (6.10)	23.65 (4.69)	25.18 (7.18)	-0.810
	Negative Affect	11.48 (1.47)	11.65 (1.69)	11.32 (1.25)	0.727
	Total	35.93 (6.34)	35.30 (4.95)	36.50 (7.46)	-0.608
BSI		0.59 (0.38)	0.56 (0.33)	0.62 (0.43)	-0.466
ASSIST	Alcohol	6.79 (4.88)	7.15 (4.57)	6.45 (5.23)	0.457
	Cannabis	3.62 (4.53)	4.05 (4.63)	3.23 (4.50)	0.584
	Cocaine	0.14 (0.65)	0.15 (0.67)	0.14 (0.64)	0.067
	Amphetamine-type stimulants	0.07 (0.46)	0	0.14 (0.64)	-0.952
	Inhalants	0	0	0	n.a.
	Sedatives	0.24 (0.76)	0.40 (0.99)	0.09 (0.43)	1.330
	Hallucinogens	0.14 (0.65)	0	0.27 (0.88)	-1.380
	Opioids	0.07 (0.46)	0	0.14 (0.64)	-0.952
	Total	11.07 (9.44)	11.75 (8.28)	10.45 (10.55)	0.440

Note. M = Mean; SD= Standard Deviation. n.a. = not applicable.

Supplementary Table 2.

Summary of studies investigating the effects of space and emotion-space interactions on auditory and visual modalities.

Authors	**Sensory Modality**	**Method**	**Stimuli**			**Task**	**Main Results**
			Type	**Emotion**	**Space**		
Lewald & Getzmann (2011)	Auditory	EEG	White noise	---	Left eccentric Left central Right eccentric Right central	Passive listening	N1- increased amplitude to eccentric *vs*. central sounds (**main effect of eccentricity**); shorter latencies to eccentric *vs*. central sounds (**main effect of eccentricity**) P2- increased amplitude to eccentric *vs*. central sounds (**main effect of eccentricity**)
Koiwa et al. (2010)	Auditory	EEG	White noise	---	Right: +90º, +45º, +15º Left: -90º, -45º, -15º (0º = behind)	Passive listening Active listening	N1-late (110–150 ms)- larger absolute electric potential for 15º sounds *vs*. passive listening (**main effect of task type**) SW-early (450–625 ms)- larger absolute electric potential in active *vs*. passive listening (**main effect of task type**) SW-late (625–800 ms)- larger absolute electric potential for 45º sounds *vs*. passive listening (**main effect of task type**)
Getzmann & Lewald (2012)	Auditory	EEG	White noise	---	Motion Scattered Displacement	Passive listening	N1- increased amplitude to scatter and displacement *vs*. smooth motion, with no differences between scatter and displacement (**main effect of sound condition**); shorter latencies to scatter and displacement *vs*. smooth motion, with no differences between scatter and displacement (**main effect of sound condition**) P2- increased amplitude to scatter and displacement *vs*. smooth motion, with no differences between scatter and displacement (**main effect of sound condition**)
Leavitt et al. (2011)	Auditory	EEG	Animal calls	---	Right: +90º, +60º, +30º Central: 0º Left: -90º, -60º, -30º (0º = front)	Location discrimination Sound recognition	N1- increased amplitude to sound recognition *vs*. location discrimination (**main effect of condition**)

Salminen et al. (2015)	Auditory	MEG	Bursts of broadband noise (Probe and Adaptor)	---	Right: (ITD, ILD or ITD + ILD) Left (ITD, ILD or ITD + ILD)	Passive listening	N1- attenuation of N1 amplitude to probes presented on the same side as the adaptors, across all spatial cue conditions (**Location-specific adaptation**)
Anourova et al. (2001)	Auditory	EEG MEG	Tones: 1000 Hz and 1500 Hz	---	Right Left	Location matching-to-sample Pitch matching-to-sample	N1- shorter latency for location vs. pitch task in the match condition (**interaction between task and condition**) N1m- increased amplitude and shorter latency for location vs. pitch task in the match condition (**interactions between task and condition**) P2- increased amplitude for location vs. pitch task in the nonmatch condition (**interaction between task and condition**)
Burra et al. (2019)	Auditory	EEG	Voices	Happy Neutral Aggressive	Left Right	Emotional target detection	N1- no differences between aggressive, happy and neutral voices N2ac- increased amplitude for aggressive *vs*. happy voices (**main effect of emotion**) LPCpc- increased amplitude for aggressive *vs*. happy voices (**main effect of emotion**)
Valdés-Conroy et al. (2014)	Visual	EEG	Pictures	Positive Neutral Negative	Near Far	Reaching judgment	N1- faster latencies and enhanced amplitudes to objects in the near *vs*. far space (**main effect of distance**); enhanced amplitudes to negative *vs*. neutral and positive pictures (**main effect of emotion**) LPP- enhanced amplitudes on parietal electrodes to objects in the near *vs*. far space (**interaction between distance and region**); enhanced amplitudes on occipital electrodes to emotional *vs*. neutral pictures space (**interaction between distance and region**)
Du et al. (2017)	Visual	EEG	Pictures	Neutral Unpleasant	Near Far	Passive viewing	LPP- larger amplitude for unpleasant *vs*. neutral pictures (**main effect of emotion**); larger amplitude for unpleasant pictures in hand-proximal *vs*. hand-distal condition (**interaction between proximity and emotion**);
Xie et al. (2014)	Visual	EEG	Words	Positive Negative	Top Bottom	Spatial cue detection	P2- larger amplitudes for top *vs*. bottom spatial cues (**main effect of spatial cue location**); When remembering positive words, larger P2 amplitude to top *vs*. bottom cues; when remembering negative words, larger P2 amplitude to bottom *vs*. top cues (**interaction between memory word valence and spatial cue location**)

Ceravolo et al. (2016a)	Auditory	fMRI	Voices	Neutral Aggressive	Proximal Distal	Distance evaluation	Increased activity in Superior Temporal Gyrus (STG), Thalamus and Amygdala for aggressive *vs*. neutral voices (**main effect of emotion**) Increased activity in left Insula, Inferior Parietal Lobule (IPL) and subregions of right STG for proximal *vs*. distal voices (**main effect of distance**) Increased activity in the right mid-STG to proximal *vs*. distal voices (both aggressive and neutral) and to aggressive *vs*. neutral (both proximal and distal spaces), (**interaction between distance and emotion**)
Ceravolo et al. (2016b)	Auditory	fMRI	Voices	Neutral Angry	Left Right	Auditory dot-probe task	Faster reaction times for valid vs. invalid and neutral trials Increased activation in bilateral middle and posterior parts of the Superior Temporal Sulcus (pSTS) and Medial Frontal Gyrus (MedFG), for valid vs. invalid trials
Pinheiro et al. (2019)	Auditory	Behavioral	Voices	Amused Neutral Angry	Left: front, side, back Right: front, side, back	Location Discrimination Emotional Recognition	<u>Location discrimination:</u> Higher accuracy for vocalizations coming from the back *vs*. front (**main effect of location**); Higher accuracy for emotional vocalizations when presented from the back or on the right side (**interaction between emotion and location**) <u>Emotion recognition:</u> High accuracy regardless of location
Asutay & Västfjäll (2015)	Auditory	Behavioral	Human and environmental sounds	Positive Negative	Front Back	Location discrimination	Faster reaction times for sounds occurring behind *vs*. front (**main effect of location**) Sounds were considered more negative coming from the back *vs*. front (**main effect of location**) Vocalizations were more negative and arousing *vs*. environmental scenes (**main effect of sound type**)
Bach et al. (2009)	Auditory	SCR	Full motion cue and Intensity change only sounds	---	Approaching Receding	Target detection	Faster reaction times for approaching *vs*. receding sounds (**main effect of direction**) Larger SCR magnitude for approaching *vs*. receding, with full motion cue sounds (**interaction between sound type and direction**) Approaching sounds were more unpleasant, potent, arousing, intense, salient, and threatening *vs*. receding sounds (**main effect of direction**)

Tajadura-Jiménez et al. (2010)	Auditory	EDA EMG	Natural (human and animal) and artificial (continuous and discontinuous)	Neutral Negative	Front Back	Room size estimation Sound-source distance estimation	Negative human sounds were the most unpleasant, arousing, unsafe and elicited the largest physiological responses (**interaction between valence and sound type**) Sounds presented behind were more arousing and elicited a higher EDA *vs.* front (except for artificial discontinuous) (**interaction between sound type and location**)