Universidade do Minho
Escola de Psicologia

Marcelo Francisco Vieira Dias

The role of attention in the processing of emotional vocalizations: ERP insights

outubro de 2013

Universidade do Minho
Escola de Psicologia

Marcelo Francisco Vieira Dias

The role of attention in the processing of emotional vocalizations: ERP insights

Dissertação de Mestrado
Mestrado Integrado em Psicologia,
Área de Especialização em Psicologia Clínica e da Saúde

Trabalho realizado sob orientação da
Professora Doutora Ana P. Teixeira Pinheiro
e da
Professora Doutora Yvonne Delevoye-Turrel

outubro de 2013

DECLARAÇÃO

Nome: Marcelo Francisco Vieira Dias

Endereço electrónico: a56017@alunos.uminho.pt

Número do Bilhete de Identidade: 13709749

Título dissertação de mestrado

The role of attention in the processing of emotional vocalizations: ERP insights

Orientador(es):

Ana Patrícia Teixeira Pinheiro

Yvonne Delevoye-Turrell

Ano de conclusão: 2013

Designação do Mestrado:

Especialização em Psicologia Clínica e da Saúde

É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA DISSERTAÇÃO APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE

Universidade do Minho, 17 / 10 / 2013

Assinatura: ____

TABLE OF CONTENTS

AGRADECIMENTOS

À Professora Ana Pinheiro pela orientação científica e acompanhamento cuidado de todo este trabalho. Obrigado pela preocupação, ajuda constante e pelas oportunidades valiosas de aprendizagem e envolvimento com o trabalho de investigação ao longo destes últimos anos.

Um obrigado a toda a minha família, em especial aos meus pais, ao meu irmão e os meus avós, pelo apoio sempre presente e desmedido durante todo este tempo. Joel, um abraço especial para ti, por tudo, desde sempre.

Rocky, Nocas, Júlia e Nuno, obrigado por toda a paciência e amizade ao longo destes cinco longos anos.

Tatiana, obrigado pela amizade e preocupação. A tua ajuda e apoio foram indispensáveis para conseguir realizar este trabalho.

Ao pessoal do CCG e à malta das *biclas* pelos momentos de lazer e convívio incomparáveis!

Marrie, muito obrigado pela tua amizade. Obrigado por todos estes anos de preocupação! Fazes cá falta.

Um obrigado muito especial à Betty e aos 'gatos' pelos indispensáveis momentos de sanidade e lucidez proporcionados.

Guida, muito obrigado! Meu bem, sem a tua dedicação, apoio, compreensão e companheirismo em todos os momentos, teria sido impossível. Obrigado por estares sempre presente!

Mestrado Integrado em Psicologia da Universidade do Minho

Área de Especialização em Psicologia Clínica e da Saúde

O papel da atenção no processamento de vocalizações emocionais: *insights* electrofisiológicos

Marcelo Francisco Vieira Dias

Ana P. Teixeira Pinheiro

Yvonne Delevoye-Turrell

RESUMO

Identificar rapidamente as emoções transmitidas pela voz e pelas faces dos outros é fundamental para um funcionamento social adequado. Vários estudos comportamentais e electrofisiológicos analisaram o processamento dos sinais usados na comunicação emocional. Contudo, a maioria utilizou faces ou palavras faladas como estímulos. Assim, ainda pouco se sabe sobre os correlatos neuronais e o desenvolvimento temporal do processamento de vocalizações não-verbais. O presente estudo usou a técnica de ERP (potenciais relacionados com eventos) para estudar o processamento de vocalizações não-verbais emocionais (alegria e raiva) e neutras num estado tardio do processamento. O componente de onda P300 foi analisado. Foi descoberto um efeito modulatório da emoção sobre a amplitude deste. Vocalizações de alegria e raiva evocaram amplitudes mais positivas para o componente P300 em comparação com neutras. Mais, foi encontrado um efeito do contexto emocional no processamento de sons neutros. Vocalizações de neutralidade num contexto de raiva evocaram um componente P300 de amplitude mais positiva em comparação com vocalizações de neutralidade num contexto de alegria. Adicionalmente, foi observada uma diferença de género: verificou-se uma amplitude mais positiva do P300 para mulheres. Estes resultados sugerem um efeito da valência dos estímulos a um nível atencional e de memória imediata, indexado pelo componente P300.

Palavras-Chave: Processamento vocal emocional; Processamento atencional; Potenciais evocados; P300.

Integrated Master's in Psychology of University of Minho

Specialty in Clinical and Health Psychology

The role of attention in the processing of emotional vocalizations: ERP insights

Marcelo Francisco Vieira Dias

Ana P. Teixeira Pinheiro

Yvonne Delevoye-Turrell

ABSTRACT

Rapidly and effectively identifying the emotions conveyed by others' faces and voices is fundamental for an adequate social functioning. Several behavioral and electrophysiological studies have analyzed the processing of emotional communication signals. Nonetheless, the majority used faces or spoken words as stimuli. Yet, little is known about the neural correlates and temporal course of the processing of non-verbal vocalizations. The present study used the ERP (Event-Related Potential) methodology to study the processing of non-verbal emotional (angry and happy) versus neutral vocalizations at a later stage of processing. The P300 component was analyzed. Its amplitude was differently modulated as a function of emotion. Angry and happy vocalizations elicited more positive amplitudes for the P300 component in comparison with neutral ones. Furthermore, emotional context was found to have an effect on the processing of neutral sounds. Neutral vocalizations in an angry context elicited a more positive amplitude P300, in comparison to neutral vocalizations in a happy context. Furthermore, a gender difference was observed: the P300 amplitude was found to be more positive for female relative to male participants. Together, these findings suggest an effect of stimulus (non-verbal vocalizations) valence at an attentional and immediate memory level, as indexed by the P300 component.

Keywords: Emotional vocal processing; Attentional processing; ERP; P300.

1. INTRODUCTION

The ability to rapidly and effectively identify emotions conveyed by others' faces and voices is fundamental for an adequate social functioning (Paulmann & Kotz, 2007; Schirmer & Kotz, 2006). It is of particular importance for the understanding of attitudes, emotional states and for predicting the intentions of those who surround us (Hawk, van Kleef, Fischer & van der Schalk, 2009).

We can convey emotional information through various channels, such as facial expressions (Ekman & Friesen, 1971), semantic verbal content (Scott, O'Donnell, Leuthold & Sereno, 2009) or prosody, the latter being the suprasegmental feature of speech (Kotz & Paulmann, 2007). Yet, the information transmitted acoustically - as vocal cues - is especially effective for the communication of emotion, since it is not absolutely sight dependent (Hawk et al., 2009).

1.1 Non-verbal Vocalizations

Non-verbal vocalizations (Schröder, 2003) seem to be especially effective for the purpose of emotional communication. These types of vocal stimuli do not have semantic content, neither the segmental structure of speech (Sauter & Eimer, 2010) and, therefore are considered to be a relatively pure vocal expression of emotion (Scott et al., 1997 *in* Sauter & Eimer, 2010).

In a review paper, Scherer (2003) reported an average of 55 to 65% of accuracy in decoding emotion from non-verbal emotional vocalizations, with these values varying between five to six times above what would be expected by chance. Also, Schröder (2003) reported an overall mean recognition rate of 81.1% in a forced choice task with ten different emotional non-verbal vocalizations categories. And more recently, Hawk and colleagues (Hawk et al., 2009) found that similarly to facial displays of emotion, non-linguistic affective vocalizations are especially effective in conveying emotional information. Participants were able to significantly decode more accurately emotion from non-linguistic affective vocalizations and facial cues than from speech-embedded vocal prosody. Therefore, non-verbal emotional vocalizations seem to be very powerful in the communication of this type of information, especially when the person who intends to communicate is not in sight.

Considering the fact that social interactions between human individuals rely vastly on auditory communication signals, it seems fairly plausible that the ability to identify emotions conveyed by the human voice has a fundamental adaptive value (Scherer, 2003). There are obvious advantages in rapidly recognizing if someone is screaming to warn us if an imminent danger is close or if he/she is in need of help. Hence, it is not surprising that for centuries, voice has attracted a great deal of attention and still does, representing a central medium to carry affective information. From Aristotle's *Rhetoric* on the importance of voice and the emotional information conveyed through it, to the 1800's with Darwin's *The Expression of the Emotions in Man and Animals* (Darwin, 1872, 1998), the vocal expression of emotion has been the focus of interest of several philosophers and scientists.

1.2 Emotion and Attention

If we take into account the seemingly endless stream of information humans are constantly subjected to, one can ask how these vital and fast decoding processes occur. The fact that our sensory systems possess a limited processing capacity allows only a fraction of the information to become the focus of sustained attention and consequent processing (Johnson & Proctor, 2004, p. 57). Therefore, a selection of the to-be-processed information has to be made. The "motivated attention" view of emotion (Lang, Greenwald, Bradley & Hamm, 1993) posits that processing of emotional stimuli is dependent on their valence (pleasant/unpleasant) and arousal (intensity) levels. Stimuli are evaluated according to their emotional/motivational significance and are either categorized in a continuum as less pleasant (i.e., aversive - negative valence) or more pleasant (i.e., appetitive - positive valence). Furthermore, according to this evaluation, cognitive and physiological resources are mobilized in order to enable an adequate behavioral response as a function of emotional significance. It is theorized that one possible implication of such mechanism is that an intense emotional stimulus, irrespective of valence, activates motivational circuits engaging sustained attention towards the relevant emotional stimulus (Lang, Bradley, & Cuthbert, 1997).

Various studies support this view of emotional priority in the allocation of attentional resources. For example, Meinhardt and Pekrun (2003) used a dual-task paradigm with simple tones in which participants had to count rare deviant stimuli from a stream of frequent standard stimuli (oddball task) as they watched emotional (pleasant/positive and unpleasant/negative) and neutral pictures. They found that emotional pictures, irrespective of valence, had a disruptive effect on the P300 ERP component. In other words, the amplitude of the P300 was smaller when an emotional picture was presented, in comparison with a neutral one. Thus, the authors proposed that there is a priority of resources allocation to emotional stimuli, even when these are not task-relevant, thus competing for task-related processing resources. The authors found the same pattern of results in a second experiment, in which participants had to perform a mental imagery task with emotional (positive and negative) and neutral content and no imagery, prior to the same oddball discrimination task. Emotional imagery (irrespective of valence) had a decreasing effect on the P300 amplitude compared to neutral or no imagery conditions.

1.3 Emotional Auditory Processing

Bearing in mind that emotional stimuli have a preferential access to attentional resources, how do we extract meaning from emotional auditory stimuli present in our everyday life? What are the cortical and subcortical processes that modulate the way we perceive other's vocal emotional information? How do these processes unfold over time? These questions gave rise to a number of models supported by imaging and electrophysiological methodologies, and based both on lesion and normal subjects' data (for an overview refer to Demaree, Everhart, Youngstrom & Harrison, 2005). More recently in an effort to integrate the conflicting experimental results in the field of vocal emotional processing, Schirmer and Kotz (2006) proposed a multi-stage model arguing for a differential brain representation of each sub-process. The proposed model comprises three sub-processes: first a) a sensory processing of vocal cues; secondly b) the integration of previously extracted emotional cues and detection of its emotional salience, and finally c) the cognitive evaluation of the emotional significance of the vocal stimulus (Kotz & Paulmann, 2011; Schirmer & Kotz, 2006).

According to Schirmer and Kotz's model, first the bilateral auditory cortices mediate the analysis of acoustic information within the first 100ms after hearing a stimulus. It has been found that differences in the frequency (Crottaz-Herbette & Ragot, 2000) and sound intensity modulate the amplitude of a negative ERP component that peaks around 100ms (N100). Secondly, these cues are integrated and emotional significance is extracted. The auditory 'what' pathway, which projects from the superior temporal gyrus (STG) to the anterior superior temporal sulcus, integrates emotionally significant acoustic information (Schirmer & Kotz, 2006). It is believed that some emotions have specific combinations of the above-cited cues (Schirmer & Kotz, 2006). For example, happiness is characterized by a fast speech rate, high intensity, mean F0 and F0 variability sounding both melodic and energetic. By contrast, sad vocalizations are characterized by a slower speech rate, low intensity, mean F0 and F0 variability with high spectral noise (Schirmer & Kotz, 2006). This level of processing occurs at around 200ms and is indexed by the P200 ERP component (Liu, Pinheiro, Deng et al., 2012; Paulmann & Kotz, 2008; Pinheiro et al., 2013). Finally, at the third stage, the emotional significance is made available to high order cognitive processes such as evaluative judgments or integration of this meaning with the semantic content (Schirmer & Kotz, 2006).

1.4 Event-Related Potentials

An excellent way of disentangling the temporal unfolding of these processes in the brain is by using the ERP methodology. An ERP is an electrical potential time-locked to a mental or physical event, measurable via electroencephalography (EEG) (Duncan et al., 2009). It is a non-invasive way of studying cognitive processes in the human brain and, as opposed to measures that rely on hemodynamic responses, it has an excellent time resolution (Luck, 2005 – page 25). Therefore, in order to effectively study the influence of emotional vocalizations on attentional processes, the use of this methodology presents itself as the most adequate.

Various studies using ERP's have reported a modulatory effect of emotion on visual and auditory sensory and perceptual processes, as well as on attentional and immediate memory processes (e.g. Bostanov & Kotchoubey, 2004; Dolcos & Cabeza, 2002; Thierry & Roberts, 2007; Liu, Pinheiro, Deng et al., 2012; Liu, Pinheiro, Zhao et al., 2012; Pinheiro et al., 2013). In a recent study using non-verbal emotional vocalizations, Liu, Pinheiro, Deng and colleagues (2012) found a significant effect of emotion on sensory processes. The authors found a more negative N100 for emotional (happy and angry) vocalizations, compared to neutral ones. Furthermore, the authors found that the P200 component was more positive in amplitude, once again for emotional (happy and angry) vocalizations, compared to neutral ones. These results suggest a facilitated emotional processing at early processing stages as indexed by N100 and P200 findings. Concerning the later stages of processing, related to the allocation of attentional resources, Thierry and Roberts (2007) found an attentional reorientation effect of unpleasant sounds compared to neutral ones. In their study, the P3a amplitude was significantly higher for novel emotional stimuli. Therefore, it seems that unpleasant vocal emotional stimuli elicit a spontaneous reorientation effect of attention. Also, another study (Czigler, Cox, Gyimesi & Horváth, 2007) found an effect of negative valence of emotional sounds (e.g. dentist drill or vomiting) on both early (N100) and late (P3a and P3b) components. In both cases more positive amplitudes were associated with the processing of aversive sounds compared to everyday sounds (e.g. bicycle bell or door-opening). Additionally, as mentioned above, Meinhardt and Pekrun's (2003) study showed that emotional stimuli modulate the recruiting of attentional recourses, as seen by their effect on the P300 component, with emotional stimuli eliciting more positive amplitudes.

In summary, emotional stimuli and particularly human vocal cues appear to have significant modulatory effects in both early and late stages of processing (Shirmer & Kotz, 2006). Considering the models' last stage, evidence from ERP research suggests that valence has an impact on the high-order processes, as for example attention allocation (e.g. Thierry and Roberts, 2007), that take place within the time window correspondent to the third stage of the model proposed by Schirmer and Kotz (2006).

Hereupon, ERP's appear to be extremely well suited to disentangle the temporal course of these processes. The most adequate index to evaluate such effects of emotion on attention and working memory is the P300 complex, reported for the first time in 1965 by Sutton and his colleagues (Sutton, Braren, Zubin & John, 1965). The P300 is a large positive electric potential that peaks around 300ms after the onset of a rare task-relevant stimulus. Its topographic distribution is centro-pariental and it is found to be more positive in amplitude over midline electrodes (Duncan et al., 2009). Since its discovery in the mid 1960s, it probably became the most studied ERP component. In order to elicit a P300 ERP component, the "oddball" paradigm is the most frequently used method and, probably, the most used ERP paradigm ever (Picton et al., 2000). It consists of a random presentation of two different stimuli, one presented frequently (standard) and the other infrequently (deviant). The participants' task is to overtly (e.g. pressing a button) or covertly (e.g. silently count) respond to the deviant stimuli (Polish, 2007). If a given stimulus is categorized as infrequent, it will elicit an increased positivity around 300ms (P300) relative to the standard stimulus. Nonetheless, various factors affect the amplitude of this component, such as the probability of target stimuli presentation, relevance of the stimulus to the task, demands of the task (if it is more or less difficult to recognize the deviant stimuli) (Duncan et al., 2009) or the time between target stimuli (Fitzgerald & Picton, 1984; Polish, 1990; for an overview refer to Polish, 2007). Although no consensus or definitive knowledge about the functional meaning of the P300 component exists (Luck, 2005, p. 42), it is generally viewed as an index of the processes responsible for the updating of information stored in the working memory system (Donchin, 1981).

2. AIMS OF THE STUDY

We propose the use of an adapted oddball paradigm to study the effect of non-verbal emotional vocalizations on the third stage of the model proposed by Shirmer and Kotz (2006). The electrophysiological index to be used is the P300 component. Relative to prosody (e.g. Kotz & Paulmann, 2007) or emotional words (e.g. Kousta, Vinson & Vigliocco, 2009), non-verbal emotional vocalizations have the obvious advantage of being free of linguistic information. Thus, potential confounds, for example related to the processing of semantic information can be avoided.

Our goal is to assess the effect of positive and negative valence in comparison to neutral valence on the attentional and context-updating processes indexed by the P300. The combination of positive valence (happy) and negative valence (angry) with neutral vocalizations and their counterbalancing will give rise to four experimental conditions: angry, happy, neutral in the context of positive valence (neutral-1) and neutral in the context of negative valence (neutral-2) (Table 1). The use of two neutral conditions will allow us to analyze the effect of emotional context on the processing of neutral vocalizations. Thus, serving as emotional context, will be the positive and negative valence standard stimuli.

	Block 1	Block 2	Block 3	Block 4
Standard	Neutral	Happy	Neutral	Angry
Deviant	Happy	Neutral	Angry	Neutral

Table 1 - Block design and condition combination. The presentation of the four blocks was counterbalanced across subjects.

Considering the emotional sensitiveness of the P300 component that has been reported before (e.g. Liu, Pinheiro, Zhao et al., 2012; Meinhardt & Pekrun, 2003; Schupp et al., 2004; Thierry & Roberts, 2006), we expect a general effect of emotion on the amplitude of the P300. Thus, happy and angry non-verbal emotional vocalizations are expected to elicit equally increased P300 amplitude. Furthermore, neutral vocalizations presented in an emotional context (angry and happy vocalizations) are expected to elicit equally lower P300 amplitudes. Therefore, we anticipate no effect of emotional context as a function of valence and thus, no P300 differences between neutral stimuli are expected.

3. METHODOLOGY

3.1 Participants

Twenty-two students were recruited at the University of Minho, Braga, Portugal and half of them received course credit for their participation. Data from six participants were not included due to technical problems or excessive artifacts. Therefore, only sixteen (7 female; mean age: 23.25 ± 3.26) were included for analysis (see Table 2 for socio-demographic data). All the subjects were right-handed, native speakers of European Portuguese, free of medication that could affect the morphology of the EEG, had no history of psychiatric or neurological disorders, no hearing impairments and had normal or corrected vision. Also, none of them had a history of traumatic brain injury, alcohol or substance abuse. Ethics approval was obtained from the Ethics Committee on Research (School of Psychology, University of Minho). A written informed consent was provided after a detailed description of the experiment.

	Age (years)
	Mean (SD)
Male (N = 9)	24,56 (3,50)
Female (N = 7)	21,57 (2,07)
Total (N = 16)	23.25 (3.26)
	Mean (SD)
Verbal IQ[1]	121.50 (10.41)
Verbal Comprehension Index[1]	122.38 (11.07)
Working Memory Index[1]	116.06 (14.40)
BSI (Positive Symptoms Index)	1.23 (0.17)

Table 2 - Socio-demographic data from the participants' sample. [1] Verbal IQ [t (14) = 1.54; n.s.], Verbal Comprehension Index [t (14) = 1,44; n.s.] and Working Memory Index [t (12) = 2.11; n.s.] are not significantly different between women and men.

3.2 Materials and Stimuli

The WAIS-III (Wechsler, 2008) was used in order to assess the cognitive profile of the participants. Three cognitive measures were acquired: Verbal IQ (Intelligence Quotient) (121.50 ± 10.41), Verbal Comprehension Index (122.38 ± 11.07), and Working Memory

Index (116.06 ± 14.40). The Brief Symptom Inventory (Canavarro, 2007) was administered to participants in order to assess the presence of psychological or psychiatric symptoms, of which no evidence was found (1.23 ± 0.17). Finally, the Edinburgh Handedness Inventory (Oldfield, 1971) was used to assess handedness.

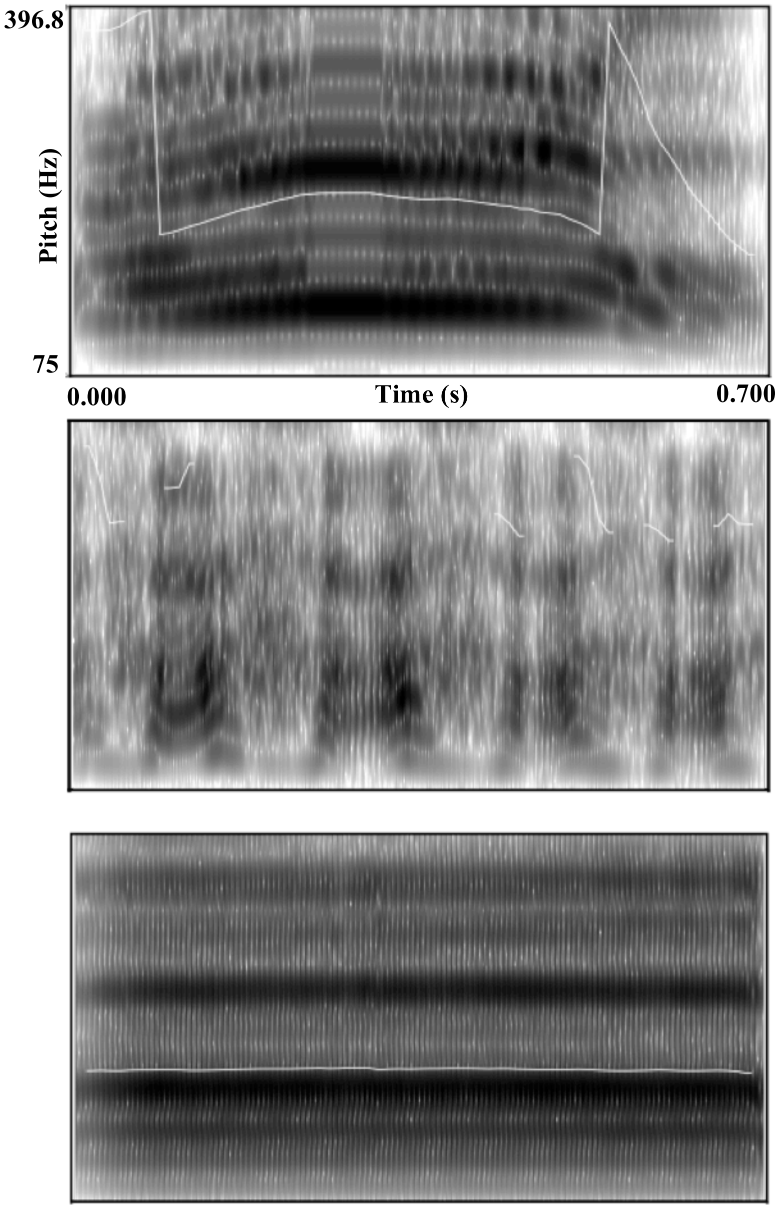

Fig. 1 - Spectrogram of each vocalization. From top to bottom: angry, happy and neutral.

The stimuli used were selected from the *Montreal Affective Voices* database (Belin, Fillion-Bilodeau & Gosselin, 2008). An exemplar of a neutral (valence: 4.67 ± 0.90; arousal: 3.32 ± 2.29; dominance: 6.02 ± 2.35), a happy (valence: 5.32 ± 1.08; arousal: 5.33 ±2.24; dominance: 6.77 ± 1.78) and an angry (valence: 3.40 ± 1.66; arousal: 5.72 ± 2.18; dominance: 5.02 ± 2.27) vocalization was selected based on the results from the adaptation of the database to the Portuguese population. Acoustic properties as well as spectrograms can be seen in Table 3 and Figure 1 respectively. The selection of the vocalizations to be used was preceded by the validation of the *Montreal Affective Voices* database (Belin et al., 2008) to the Portuguese population. The sample used was composed of 60 university students from the University of Minho, Braga, Portugal (30 females; 21.97 ± 3.382 years).

	Angry	**Happy**	**Neutral**
Duration (ms)	700	700	700
Mean Pitch (Hz)	248.32	322.02	190.57
Mean Intensity (dB)	70.56	74.89	78.04

Table 3 - Acoustic properties of the stimuli used.

3.3 Procedure

The experimental procedure was divided into three sessions. During the first session, the informed consent was provided; handedness, psychiatric symptomatology and cognitive functioning were assessed. The Edinburgh Handedness Inventory (Oldfield, 1971), the Portuguese version of the Brief Symptom Inventory (Canavarro, 2007) and all the subtests comprised in the verbal IQ (VIQ) domain from the WAIS-III (Wechsler, 2008) were administered (Vocabulary, Similarities, Information, Comprehension, Arithmetic, Digit Span, Letter-Number Sequencing). The use of these subtests allowed the additional calculation of verbal comprehension (VCI) and working memory indices (WMI).

The next two sessions involved EEG data collection. A modified oddball paradigm was used. Stimuli used were angry, happy and neutral vocalizations. Stimuli within each block were presented in a pseudo-randomized order with a minimum of one and a maximum of six standards occurring between each deviant. The order of the sounds' sequence was maintained equal across participants.

Each trial (Figure 2) began with a 1000ms fixation cross at the center of the screen. At 300ms from the onset of the cross, a vocalization was presented, which lasted for 700ms and ended simultaneously with the fixation cross. After this period of time, a 500ms interstimulus interval consisting of a blank screen was presented. The experiment was divided in four blocks that counterbalanced the combination of happy, angry and neutral vocalizations. All three vocalizations were presented as deviant and standard stimuli. Thus, resulting in four different conditions: angry, happy, neutral in the context of happy vocalizations (neutral-1) and neutral in the context of angry vocalizations (neutral-2). Each block was composed by

210 (84%) standard and 40 (16%) deviant vocalizations. The experiment was conducted in a dimly lit, sound-attenuated, and electrically shielded room. Participants were seated in a comfortable chair at 100cm distance from an LCD screen. Presentation software (version 16.3; Neurobehavioral Systems, Inc., Albany, NY, USA) was used to control the presentation and timing of stimuli. The auditory stimuli were presented through a set of Sennheiser CX 300-II ear-canal phones with 19 to 21,000 Hz of frequency response and 113 dB of sound pressure. Participants were instructed to silently count the number of deviants in each block. At the end of each block, the number of deviant stimuli counted was registered.

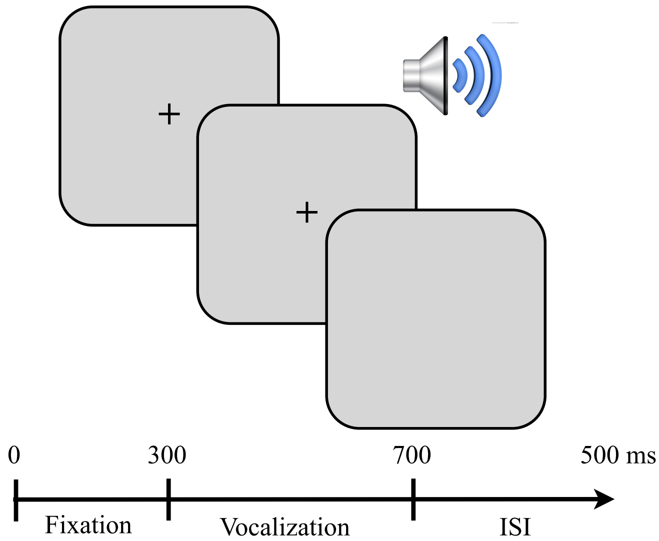

Fig. 2 - Trial design: the fixation cross lasts for 1000ms, the vocalization starts at 300ms from the beginning of the fixation cross and lasts for 700ms. ISI (blank screen) has a duration of 500ms.

Finally a 9-point-SAM-scale (Self-Assessment Manikin) (Bradley & Lang, 1994) was used at the end of the third session in order to obtain the participant's subjective evaluation of the sound they heard. The parameters were valence, arousal, dominance and the identification of the portrayed emotion in a forced-choice-task.

3.4 EEG Data Acquisition

The Electroencephalogram (EEG) was recorded using custom designed electrode caps from a 64-channel ActiveTwo BioSemi system (BioSemi, Amsterdam, The Netherlands). Electrode location was based on the modified expanded international 10/20 system (American Electroencephalographic Society, 1991). EEG was acquired in a continuous mode at a digitization rate of 512Hz, with a bandpass of 0.01–100 Hz, and stored for later analysis. Additionally, 2 flat-type electrodes placed on the left and right mastoids were recorded for offline referencing. Eye movements and blinks were monitored by placing a flat-type electrode on each temple for the vertical electrooculogram and one electrode below the left eye for the vertical electrooculogram.

3.5 EEG Processing

EEG data were analyzed using BrainVision Analyser software (Brain Products GmbH, Munich, Germany). The EEG data were referenced offline to the algebraic sum of the left and right mastoids. Individual, time-locked to the onset of the stimuli, epochs were extracted from the datasets. These started 100ms before stimulus onset and finished 1000s after stimulus onset. After this, a baseline correction was performed subtracting the 100ms prestimulus period. Eye blinks and movements were corrected using the method developed by Gratton, Coles & Donchin (1983). Segments showing excessive eye movements, blinks, muscle activity, or amplifier blocking were rejected offline before averaging (epochs with voltage levels exceeding ± 100 mV were rejected from further analysis). Individual averages were only considered if at least 50% of the segments in a given condition passed the artifact rejection (as standard: 156.50 ± 28.49 for angry; 162.19 ± 28.53 for happy; 162.19 25.53 for neutral-1; 161.63 ± 29.58 for neutral-2; as deviant: 30.44 ± 6.38 for angry; 31.88 ± 4.99 for happy; 29.06 ± 6.59 for neutral-1; 28.50 ± 7.11 for neutral-2). As stated above happy, angry and neutral stimuli were both presented as standards and deviants in different blocks. This procedure was used in order to avoid physical differences of the stimuli to have confounding effects on the P300 amplitude. Therefore, the activity of a given stimulus, as a standard was subtracted to the activity of the same stimulus as a deviant (Johnson, 1993). P300 mean amplitude and peak latency were measured from a time window ranging from 350 to 450ms, post-stimulus onset.

3.6 Statistical Analysis

Mean amplitude and peak latency of the P300 were individually subjected to repeated measures analyses of variance (ANOVA's) with emotion (happy, angry, neutral-1 and neutral-2) and electrode (Cz and Pz) as within-subject factors. Additionally, a repeated measures ANOVA was performed to assess the effects of hemisphere, with the following within-subject factors: emotion (angry, happy, neutral-1 and neutral-2), electrode (C1, C2, P1 and P2) and hemisphere (left and right). As a preliminary analysis, gender was posteriorly included as between-subjects factor. Main effects and interactions were followed with planned comparisons with *Bonferroni* correction.

As exploratory analyses, paired samples T-tests were performed to assess potential differences between women and men for the WAIS-II (VIQ, VCI and WMI) and BSI results. More, paired sample T-tests were performed to assess possible differences in the subjective ratings of arousal, valence and dominance. Additionally, a repeated measures ANOVA with emotion (happy, angry and neutral) and dimension (valence, arousal and dominance) as within-factors, and gender as between-factor was computed to investigate possible differences in ratings between women and men. Finally, the association between ERP amplitudes (Cz and Pz) and clinical/neurocognitive/behavioral measures was tested using Pearson's Correlation Coefficient. Only significant results ($p \leq 0.05$) will be reported.

4. Results

4.1 Counting of Deviant Stimuli

The number of deviant stimuli counted by the participants was very high (happy: 39.44 ± 1.21; angry: 39.19 ± 1.22; neutral-1: 39.63 ± 0.96; neutral-2: 39.50 ± 0.82). This indicated that participants were focused on counting the deviant stimuli during the EEG recording.

4.2 P300

4.2.1 P300 Amplitude

The ANOVA revealed a main effect of emotion [F (3,45) = 10.19, $p < .001$]: happy vocalizations elicited more positive amplitude compared to neutral ones in a happy context (p = .015); angry stimuli elicited a more positive P300 amplitude than neutral ones in a happy context (p = .003) and finally, neutral stimuli in an angry context evoked a more positive P300 amplitude relative to neutral stimuli presented in a happy context (p = .025).

A main effect of gender was observed [F (1,14) = 7.89, p = .014] in a preliminary analysis testing differences between men and women: more positive P300 was found in women relative to men, irrespective of condition.

Average amplitudes can be seen in Table 4. Grand average waveforms ilustrating the four conditions at C1, C2, Cz, P1, P2 and Pz can be seen in Figure 3.

	Amplitude µV
	Mean (SD)
Angry	5.99 (.68)
Happy	6.65 (.75)
Neutral in happy context (neutral-1)	3.30 (.49)
Neutral in angry context (neutral-2)	4.75 (.48)

Table 4 - Average voltage (µV) extracted from a time window ranging from 350 to 450ms, post-stimulus onset.

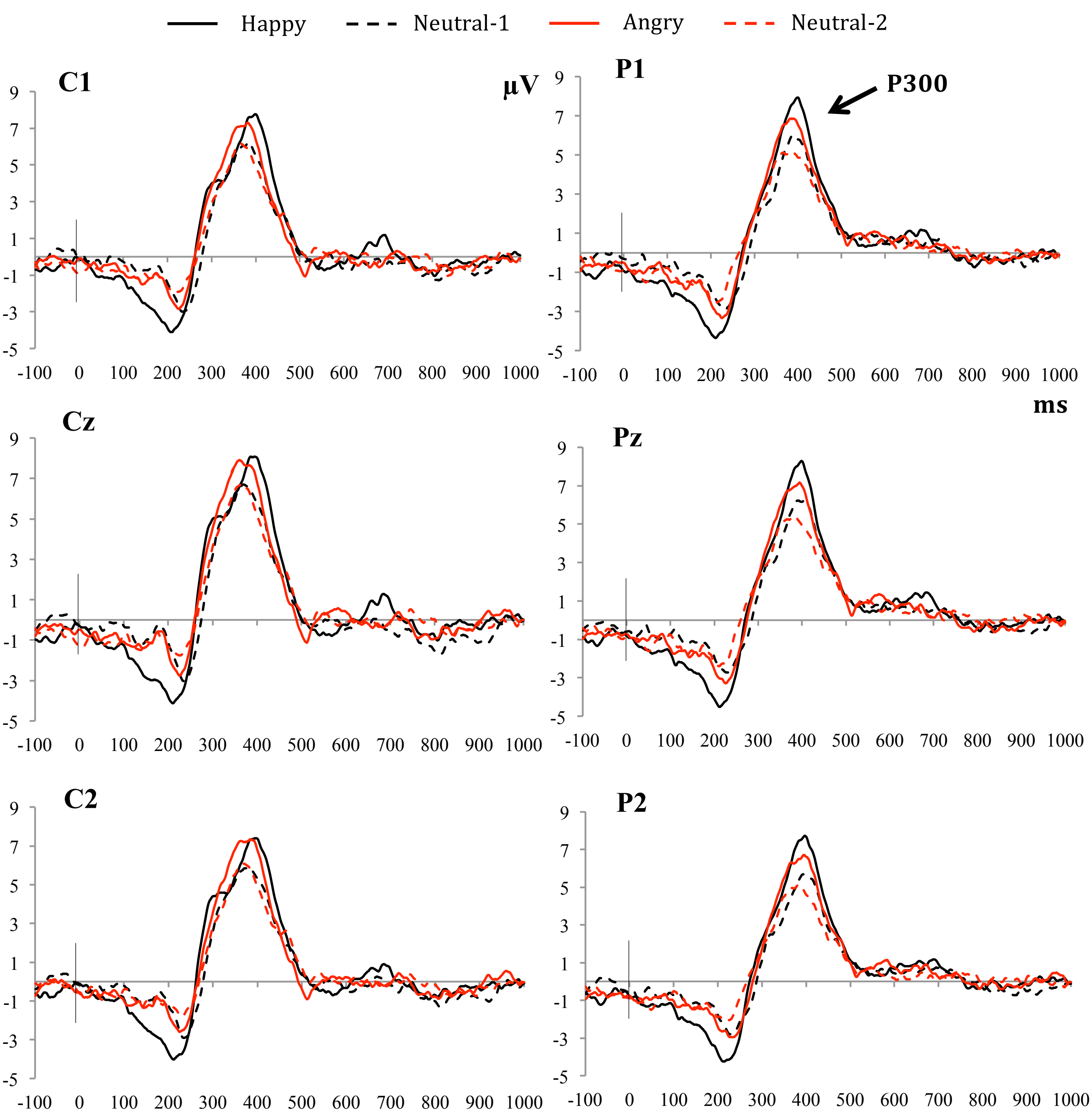

Fig. 3 - Event-related potentials time locked to the onset of the vocalizations. Grand Average waveforms of each condition (angry, happy neutral-1 and neutral-2) at C1, Cz, C2, P1, Pz and P2 sites. The waveforms above illustrated are the result of the subtraction method applied.

4.2.2 P300 Latency

Analyses of latency did not reveal any main effect or interaction, neither at midline sites (Cz and Pz) nor between hemispheres.

4.3 Ratings of Sounds' Affective Dimensions

Paired samples T-tests revealed significant differences between emotions for the three dimensions rated (Table 5). Considering the arousal dimension, neutral vocalizations differed significantly from angry [t (15) = -10.20, p < .001] and happy [t (15) = -7.16, p < .001] ones, with both emotional stimuli being rated as more arousing then neutral stimuli. No differences between angry and happy vocalizations concerning arousal ratings were found. Considering valence, all three stimuli were differently rated. The ratings of neutral vocalizations differed significantly from angry [t (15) = 4.68, p < .001] and happy [t (15) = -10.75, p < .001] ones and happy differed from angry ones [t (15) = -7.60, p < .001]. Angry vocalizations were rated as more unpleasant (rated with lower values on the 9-point-SAM-scale), followed by neutral ones and finally by happy ones. Dominance rating differed significantly between neutral and angry vocalizations [t (15) = 7.00, p < .001] and between angry and happy vocalizations [t (15) = -8.82, p < .001]. No differences in dominance ratings were found between neutral and happy vocalizations ratings.

	Valence	Arousal	Dominance
Angry	2.69 (1.89)	6.75 (1.34)	3.88 (1.78)
Happy	7.69 (0.95)	6.19 (1.42)	6.94 (1.30)
Neutral	4.88 (0.34)	3.06 (1.18)	7.31 (1.78)

Table 5 - Average 9-poit-SAM-scale ratings for valence, arousal and dominance for each sound.

4.4 Correlation Analysis

No significant correlations were found between the results of the WAIS-III indices, the BSI, the valence, arousal and dominance ratings of each sound and the P300 amplitudes. No effect or interaction of gender was found concerning the sounds' ratings.

5. Discussion

This study examined the electrophysiological correlates of the processing of emotional (angry and happy) and neutral non-verbal vocalizations at a higher-order processing stage, such as indexed by the P300 component. In order to do so, non-verbal vocalizations, devoid of linguistic confounds (e.g. semantics), were presented in a modified oddball paradigm. The comparison of the amplitude of the four conditions (happy, angry, neutral-1 and neutral-2) revealed a significant effect of emotion. Our findings indicate a differential electrophysiological response at central and parietal sites according to stimulus emotional significance, as indexed by the P300.

Thus as initially hypothesized, the processing of emotional stimuli, of either positive or negative valence led to increased amplitude of the P300 compared to neutral valence (neutral in happy context). It seems so that happy and angry auditory cues elicit a preferential access to attentional resources (Meinhardt & Pekrun, 2003) in a later stage of processing (Schirmer & Kotz, 2006), when compared with neutral stimuli. These findings. Together with subjective ratings' results point to a potential effect of arousal on attentive auditory processing. Emotional stimuli were rated as equally arousing and thus, this dimension can potentially account for the differences between the P300 amplitudes of emotional (angry and happy) and neutral (neutral-1) conditions. Still, no significant association between arousal ratings and amplitude was found.

Moreover, our results seem to support the assumption that intense emotional stimuli, irrespective of valence, engage equal processing resources (Lang et al., 1997). We found no differences between the P300 amplitudes elicited for both angry and happy conditions. More, the results obtained in the subjective rating task further support the above-cited assumption. Even though equally arousing, angry and happy vocalizations were perceived as being opposed in valence (angry vocalizations: more negative/aversive and happy vocalizations: more positive/appetitive). Still, no differences in amplitude between both were found. These results are in line with previous reports of P300 modulation by emotional significance in studies using pictures (Olofsson, Nordin, Sequeira & Polich, 2008; Schupp, Flaisch, Stockburger & Junghofer, 2006) and sounds (Czigler et al., 2007; Thierry & Roberts, 2007).

Additional effects of emotion were found when comparing neutral conditions (neutral-1 and neutral-2). In this particular case, contrary to our hypothesis, when comparing the processing of both neutral vocalizations in a positive valence context (neutral-1) and in a negative valence context (neutral-2), the second (neutral-2) elicited a more positive P300 amplitude relative to the former (neutral-1). In this case, arousal cannot account for the observed differences, as both emotional (angry and happy) stimuli used as context, were rated as equally arousing. Therefore, it seems that neutral stimuli are differently processed according to the valence of the context in which they are presented (angry vs. happy sounds). This differential processing of neutral stimuli as a function of emotional context is in line with previous findings, using ERP (Domínguez-Borràs, Garcia-Garcia & Escera, 2008) as well as fMRI (functional Magnetic Resonance Imaging) (Domínguez-Borràs et al., 2009) techniques.

If we consider the differences found for the perception of control (dominance ratings significantly lower for angry than neutral and happy), one can argue that participants perceived the angry (negative/aversive) context as more threatening than the happy context. Thus, as a consequence a heightened state of alert (Domínguez-Borràs et al., 2008) might have evoked the increased amplitude P300 of neutral stimuli in a negative valence (angry) context, as opposed to neutral stimuli in a positive valence (happy) context. In this line of thought, the 'negativity bias' framework (Cacioppo & Gardener, 1999) posits that aversive and threatening stimuli are particularly relevant and have facilitated access to processing resources. Attention is more readily oriented towards potentially dangerous or threatening events, thus facilitating a rapid processing of such stimuli. Therefore, in conformity with an evolutionary adaptive view (Cacioppo & Gardener, 1999), in a negative/aversive context, the neutral valence stimuli (irrelevant in non-threatening conditions) would be readily processed due to a gating mechanism. Thus, potentially menacing stimuli could be readily processed and an adequate response provided. In sum, even when not task relevant (not to be attended), angry vocalizations, as opposed to happy ones, modulate the processing of neutral stimuli, as seen by the increase of the P300 amplitude in this condition (neutral-2).

Finally, our preliminary analyses on gender, suggest a differential processing of emotional non-verbal vocalizations as a function of gender. Our experiment demonstrated overall more positive amplitude of the P300 component in women compared to men. This pattern of results might indicate a higher resource allocation for the processing of emotional stimuli conveyed by voice (Schirmer, Kotz & Friederici, 2002) in women relative to men. Caution is advised in the interpretation of this result considering the small sample of participants (only 7 women). Nonetheless, a previous study by Schirmer, Striano and Friederici (2005) demonstrated that, at a pre-attentive level, women apparently recruit more resources towards the processing of syllables uttered with emotional prosody as compared with neutral, relative to men. The authors found a more negative mismatch negativity (MMN) for emotional stimuli, compared to neutral stimuli, only for female participants. Further evidence seems to indicate that these differences between genders might exist at higher order processes (Hall, 1978). Still, contrary to Schirmer and colleagues' study (2005), we found no specific effect of valence.

Concerning our findings of overall increased amplitude for the P300 in women compared to men; one could arguably attribute such differences between genders to neurocognitive or clinical factors. However, our exploratory analyses revealed no significant differences of the WAIS-III (VIQ, VCI and WMI) and BSI results between genders. Thus, these factors do not seem to be mediating the P300 amplitude evoked to the emotional (angry and) and neutral vocalizations. Similarly, no differences between genders were found concerning the subjective evaluations (arousal, valence and dominance) of the stimuli.

A general effect of gender on the attentional and context updating processes indexed by the P300 seems plausible but further research is needed to verify this possibility. So it seems that a potential difference of gender in the processing of non-verbal emotional information might exist at a later stage of processing, as indexed by the P300.

This study revealed the importance and suitability of non-verbal vocalizations in the study of vocal emotional processing as well as more complex, higher order processes occurring in the third stage of processing as proposed by Schirmer and Kotz (2006). More, as this particular type of stimuli do not have confounding linguistic effects (e.g. semantics) they seem particularly adequate for studying the above-cited mechanisms. Nonetheless, future studies are needed to further clarify the role of emotional significance in the processing of these stimuli. We suggest the use of more emotional categories, such as for example: sadness, disgust or surprise. At the same time, various exemplars in each emotional category, as opposed to one, should be employed with the same goal. Additionally, arousal, valence and dominance levels of the stimuli should be controlled and systematically manipulated. This might prove to be a demanding task, as the participants' perceived levels of arousal, valence and dominance are assessed at the end of the procedure. Thus, unwanted differences concerning those ratings might arise. Finally, gender should be taken into account when recruiting participants. Finally, the necessary number of women and men should be included in future studies using non-verbal emotional and neutral vocalizations to confirm the effect of gender.

6. Conclusion

These findings demonstrate a distinct ERP response to emotional (happy and angry) non-verbal vocalizations in comparison with neural ones, as shown by the modulation of the P300 amplitude. Furthermore, neutral non-verbal vocalizations are processed differently as a function of emotional (happy and angry) context. Finally, women seem to elicit more positive amplitude P300 than men. Our results point to a possible differential processing at later stages according to gender, as indexed by the P300 component. Additional, studies are needed to properly validate an effect of gender on the processing of non-verbal emotional vocalizations, specifically concerning the P300 ERP component.

7. References

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, 40*(2), 531-539.

Bostanov, V., & Kotchoubey, B. (2004). Recognition of affective prosody: continuous wavelet measures of event-related brain potentials to emotional exclamations. *Psychophysiology, 41*(2), 259-268. doi: 10.1111/j.1469-8986.2003.00142.x

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry, 25*(1), 49-59.

Cacioppo, J. T., & Gardner, W. L. (1999). Emotion. *Annual review of psychology, 50*(1), 191-214.

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

Crottaz-Herbette, S., & Ragot, R. (2000). Perception of complex sounds: N1 latency codes pitch and topography codes spectra. *Clinical neurophysiology, 111*(10), 1759-1766.

Czigler, I., Cox, T. J., Gyimesi, K., & Horvath, J. (2007). Event-related potential study to aversive auditory stimuli. *Neurosci Lett, 420*(3), 251-256. doi: 10.1016/j.neulet.2007.05.007

Darwin, C., 1872. The Expression of Emotions in Man and Animals. London: John Murray (third ed., P. Ekman (Ed.), 1998, London: Harper Collins).

Demaree, H. A., Everhart, D. E., Youngstrom, E. A., & Harrison, D. W. (2005). Brain lateralization of emotional processing: historical roots and a future incorporating "dominance". *Behavioral and cognitive neuroscience reviews, 4*(1), 3-20.

Dolcos, F., & Cabeza, R. (2002). Event-related potentials of emotional memory: encoding pleasant, unpleasant, and neutral pictures. *Cognitive, Affective, & Behavioral Neuroscience, 2*(3), 252-263.

Dominguez-Borras, J., Garcia-Garcia, M., & Escera, C. (2008). Emotional context enhances auditory novelty processing: behavioural and electrophysiological evidence. *Eur J Neurosci, 28*(6), 1199-1206. doi: 10.1111/j.1460-9568.2008.06411.x

Dominguez-Borras, J., Trautmann, S. A., Erhard, P., Fehr, T., Herrmann, M., & Escera, C. (2009). Emotional context enhances auditory novelty processing in superior temporal gyrus. *Cereb Cortex, 19*(7), 1521-1529. doi: 10.1093/cercor/bhn188

Donchin, E. (1981). Surprise!… surprise? *Psychophysiology, 18*(5), 493-513.

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Naatanen, R., . . . Van Petten, C. (2009). Event-related potentials in clinical research: guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clin Neurophysiol, 120*(11), 1883-1908. doi: 10.1016/j.clinph.2009.07.045

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology, 17*(2), 124.

Fitzgerald, P. G., & Picton, T. W. (1984). The Effects of Probability and Discriminability on the Evoked Potentials to Unpredictable Stimulia. *Annals of the New York Academy of Sciences, 425*(1), 199-203.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and clinical neurophysiology, 55*(4), 468-484.

Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological bulletin, 85*(4), 845.

Hawk, S. T., Van Kleef, G. A., Fischer, A. H., & Van der Schalk, J. (2009). " Worth a thousand words": Absolute and relative decoding of nonlinguistic affect vocalizations. *Emotion, 9*(3), 293.

Johnson, A., & Proctor, R. W. (2004). *Attention: Theory and practice.* California: Sage.

Johnson, R. (1993). On the neural generators of the P300 component of the event-related potential. *Psychophysiology, 30*(1), 90-97.

Kotz, S. A., & Paulmann, S. (2007). When emotional prosody and semantics dance cheek to cheek: ERP evidence. *Brain Res, 1151*, 107-118.

Kotz, S. A., & Paulmann, S. (2011). Emotion, language, and the brain. *Language and Linguistics Compass, 5*(3), 108-125.

Kousta, S. T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition, 112*(3), 473-481. doi: 10.1016/j.cognition.2009.06.007

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). Motivated attention: Affect, activation, and action. *Attention and orienting: Sensory and motivational processes*, 97-135.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30*(3), 261-273.

Liu, T., Pinheiro, A. P., Zhao, Z., Nestor, P. G., McCarley, R. W., & Niznikiewicz, M. A. (2012). Emotional cues during simultaneous face and voice processing: electrophysiological insights. *PLoS One, 7*(2), e31001. doi: 10.1371/journal.pone.0031001

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

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

Meinhardt, J., & Pekrun, R. (2003). Attentional resource allocation to emotional events: An ERP study. *Cognition & emotion, 17*(3), 477-500.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: an integrative review of ERP findings. *Biological psychology, 77*(3), 247-265.

Paulmann, S., & Kotz, S. A. (2008). Early emotional prosody perception based on different speaker voices. *Neuroreport, 19*(2), 209-213.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., . . . Rugg, M. D. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology, 37*(2), 127-152.

Pinheiro, A. P., Del Re, E., Mezin, J., Nestor, P. G., Rauber, A., McCarley, R. W., ... & Niznikiewicz, M. A. (2012). Sensory-based and higher-order operations contribute to abnormal emotional prosody processing in schizophrenia: an electrophysiological investigation. Psychological medicine, 1(1), 1-16.

Polich, J. (1990). Probability and inter-stimulus interval effects on the P300 from auditory stimuli. *International journal of psychophysiology, 10*(2), 163-170.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical neurophysiology, 118*(10), 2128-2148.

Sauter, D. A., & Eimer, M. (2010). Rapid detection of emotion from human vocalizations. *Journal of Cognitive Neuroscience, 22*(3), 474-481.

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40*(1), 227-256.

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends Cogn Sci, 10*(1), 24-30. doi: 10.1016/j.tics.2005.11.009

Schirmer, A., Kotz, S. A., & Friederici, A. D. (2002). Sex differentiates the role of emotional prosody during word processing. *Cognitive Brain Research, 14*(2), 228-233.

Schirmer, A., Striano, T., & Friederici, A. D. (2005). Sex differences in the preattentive processing of vocal emotional expressions. *Neuroreport, 16*(6), 635-639.

Schröder, M. (2003). Experimental study of affect bursts. *Speech Communication, 40*(1), 99-116.

Schupp, H., Cuthbert, B., Bradley, M., Hillman, C., Hamm, A., & Lang, P. (2004). Brain processes in emotional perception: Motivated attention. *Cognition and Emotion, 18*(5), 593-611.

Schupp, H. T., Flaisch, T., Stockburger, J., & Junghöfer, M. (2006). Emotion and attention: event-related brain potential studies. *Progress in brain research, 156*, 31-51.

Scott, G. G., O'Donnell, P. J., Leuthold, H., & Sereno, S. C. (2009). Early emotion word processing: Evidence from event-related potentials. *Biological psychology, 80*(1), 95-104.

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science, 150*(3700), 1187-1188.

Thierry, G., & Roberts, M. V. (2007). Event-related potential study of attention capture by affective sounds. *Neuroreport, 18*(3), 245-248.

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