UNIVERSIDADE DE LISBOA

Faculdade de Medicina

Can Musical Training Boost Emotional Processing Skills? An Electrophysiological Study.

Inês Filipa Estradas Martins

Orientador: Prof.ª Doutora Ana Patrícia Teixeira Pinheiro

Co-orientador: Prof.ª Doutora Maria Isabel Segurado Pavão Martins Catarino Petiz

Dissertação especialmente elaborada para obtenção do grau de Mestre em Neurociências

2020

UNIVERSIDADE DE LISBOA

Faculdade de Medicina

Can Musical Training Boost Emotional Processing Skills? An Electrophysiological Study.

Inês Filipa Estradas Martins

Orientador: Prof.ª Doutora Ana Patrícia Teixeira Pinheiro

Co-orientador: Prof.ª Doutora Maria Isabel Segurado Pavão Martins Catarino Petiz

Dissertação especialmente elaborada para obtenção do grau de Mestre em Neurociências

2020

A impressão desta dissertação foi aprovada pelo Conselho Científico da Faculdade de Medicina de Lisboa em reunião de 19 de maio de 2020.

INDEX

FIGURE INDEX

I. Methods

II. Results

TABLE INDEX

I. Methods

II. Results

I. Appendices

ACKNOWLEDGMENTS

Realizar esta dissertação foi um percurso composto por desafios, surpresas e conquistas que só foram possíveis de alcançar devido às pessoas que estiveram comigo.

Quero começar por agradecer à minha orientadora, a Professora Ana Pinheiro, por ter depositado a sua confiança em mim e neste projeto desde a primeira reunião, por me ter recebido e integrado no seu laboratório e disponibilizado todos os meios para que este projeto fosse avante. Agradecer também pela energia, motivação, conhecimento e profissionalismo que foram demonstrados todos os dias que se seguiram. Nada disto teria sido possível sem a Professora e é com gratidão que reconheço o cuidado em acompanhar todas as fases e em fazer-me alcançar cada objetivo com a maior excelência e perfecionismo possíveis.

Agradeço à Professora Isabel Pavão Martins por ter embarcado nesta aventura e permitido que este projeto acontecesse, mantendo a porta sempre aberta para dar o seu apoio.

Palavras parecerão sempre poucas para expressar a minha gratidão para com os meus colegas do VAS LAB. Durante este ano senti-me em casa principalmente por causa das pessoas incríveis que compõem este laboratório. Tive a sorte de encontrar um grupo unido, coeso, com um dos maiores sentidos de entreajuda que já vi. A todos os que contribuíram de alguma forma desde o início ao fim deste percurso, o meu sincero obrigado. Tenho de agradecer especialmente à Maria Amorim pelo apoio incondicional, pela ajuda, pelas aprendizagens que me proporcionou, pela motivação e pela amizade. Se muito do que aprendi devo a ti, muito da minha sanidade mental também! Agradeço à Tatiana Conde e ao João Sarzedas por todo o tempo que disponibilizaram e dedicaram a ajudar-me sempre que foi preciso. À Marta Pereira que foi também uma verdadeira companheira de jornada e cujo apoio me foi muito estimado. Tenho de deixar o meu agradecimento ao Ricardo Correia que desde o momento em que entrou no projeto foi um apoio indescritível. Agradecer pelos longos dias no laboratório a fazer recolhas, pelo apoio nas tarefas, por toda a disponibilidade e energia colocadas no projeto enquanto fez parte dele e por se ter tornado um amigo no processo. E claro, agradeço a confiança depositada para aprender comigo e com a minha experiência, mesmo que pouca.

Quero agradecer a todos os meus amigos e familiares que foram acompanhando este percurso e me motivaram diariamente. Obrigada pelas palavras encorajadoras, pelo calor de cada abraço, por cada sorriso, por todas as pequenas e grandes coisas que me deram sempre mais força e determinação. Agradeço especialmente aos meus pais e à minha irmã que são a base de todas as conquistas que alcanço e a principal razão de eu ser quem sou. A ousadia de sair da minha zona de conforto e de ir atrás de cada sonho que me atrevo a sonhar, devo a eles.

Agradecer ao Nelson parece sempre uma tarefa ingrata porque o que ele acrescenta a todas as minhas jornadas não se consegue simplesmente colocar em palavras. Mas tenho de agradecer pela paciência, pelo companheirismo, pela motivação e por tornar tudo mais fácil. Obrigada pela força dada, pela confiança transmitida, por seres o meu confidente, o meu psicólogo, o meu respirar fundo e o “tu consegues!”. No fundo, obrigada por teres estado sempre ao meu lado, a torcer por mim. O alcançar deste objetivo é ainda mais gratificante por ser partilhado contigo.

Por último, gostaria de agradecer a todos os participantes e a todos que ajudaram a divulgar o estudo, contribuindo para que acontecesse. Agradecer ainda à equipa da Massimo Dutti do Montijo, especialmente à Ana Lúcia, que me acolheu durante este processo, mostrando sempre disponibilidade e flexibilidade para que eu pudesse concluir com sucesso este desafio a que me propus.

A todos quantos estiveram e passaram pela minha vida no último ano e de alguma forma melhoraram esta experiência, direta ou indiretamente, o meu mais profundo obrigado.

ABSTRACT

Musical training, either vocal or instrumental, demands the integration of different sensory, motor, and cognitive abilities at the same time. The involvement of specific brain structures and functions during musicians' extensive sensorimotor training over long time periods has been a focus of interest in Neuroscience research. Evidence for experience-driven neuroplasticity due to musical practice has been documented, namely enhanced auditory perception skills in musicians compared to non-musicians. Furthermore, the beneficial effects of musical training generalize into non-musical skills, such as emotion recognition (cross-domain effect). Nevertheless, vocal and instrumental training effects in auditory emotional processing have not been directly compared, even though singers experience a different training because their instrument is contained within their body. Moreover, previous studies did not account for potential differences of the musical training effect as a function of sound category, i.e., if musical training similarly influences the processing of vocally expressed emotions vs. musically expressed emotions.

This study examined the effects of vocal and instrumental musical training in the temporal course of the neurocognitive mechanisms underlying implicit processing of emotional meaning, examining event-related potentials (ERPs) of the electroencephalogram (EEG). In particular, we investigated differences between affective vocal and musical sounds in the amplitude of ERP components associated with early (N1, P2) and late stages (late positivity potential – LPP) of auditory processing. Nineteen singers, seventeen instrumentalists and twenty-two control subjects listened to human nonverbal vocalizations and musical bursts with happy, sad and neutral valence, as well as to monkey voices and tambourine sounds which served as a response target. The N1, P2, and LPP were analyzed.

On the one hand, between-groups differences were found regarding the P2, with vocalizations eliciting a more positive amplitude than musical sounds in non-musicians only. In the case of instrumentalists and singers, no differences in the P2 amplitude elicited by vocalizations compared to musical sounds were observed. On the other hand, type of stimulus effects, independent of musical training, were observed at the three processing stages with vocalizations eliciting a reduced N1 and enhanced P2 and LPP amplitudes.

Together, these findings suggest that musical training modulates the emotional salience detection stage of the auditory emotional processing. Specifically, it seems that for musically trained individuals, musical sounds may be as salient as vocal sounds. Additionally, the current results provide evidence for differentiation of sound categories in early and late stages of auditory processing.

The current findings corroborate the activation of neuroplasticity mechanisms underlying musical training, reflected in a boosting effect of auditory skills, namely in emotion processing. Thus, there is evidence for a cross-domain effect (from music to emotion).

Key-words: Event-related potentials (ERP); Auditory Emotional Processing; Musical Training; Emotion; Neuroplasticity.

RESUMO

O treino musical tem sido considerado um exemplo de neuroplasticidade pelo facto de constituir uma tarefa de alta complexidade, que envolve a interação entre funções sensoriais, motoras e cognitivas. Especificamente, a prática de um instrumento envolve (1) a perceção e processamento do estímulo auditivo produzido pelo instrumento, (2) a perceção e processamento visual devido à leitura das pautas de música, (3) o controlo motor necessário para manusear o instrumento musical/ aparelho vocal de forma precisa, e (4) a memória – tanto a nível da sequência dos movimentos efetuados numa determinada atuação, como a nível da informação musical (sequência de notas, acordes, ritmo, métrica). Por essa razão, os músicos têm sido um foco de interesse na investigação em Neurociências. Estudos de ressonância magnética funcional (fMRI) têm reportado maior densidade de substância cinzenta em regiões visuais, auditivas e motoras em músicos quando comparados com não-músicos. Ao mesmo tempo, funções cerebrais como a perceção auditiva são fortalecidas pela prática estruturada e disciplinada da música, manifestando-se depois no desempenho de tarefas não-musicais. O aumento da capacidade de perceção auditiva foi observado em músicos (vs. não-músicos) a nível do processamento do som, em geral, e de estímulos vocais e da fala em particular. Por exemplo, quando comparados com indivíduos sem treino musical, há evidência eletrofisiológica que indica que em músicos ocorre uma deteção mais rápida de padrões sonoros, há um aumento da velocidade de respostas subcorticais a sílabas e um aumento da magnitude de resposta a alterações de frequência fundamental, tempo e timbre em vocalizações humanas. Para além disso, os efeitos do treino musical também se parecem estender ao processamento de propriedades afetivas dos estímulos sonoros. Tome-se por exemplo a maior eficiência subcortical observada durante o processamento do som do choro de um bebé: em músicos, as respostas do tronco cerebral durante a porção mais complexa do som foram de maior intensidade relativamente à porção mais periódica do som, padrão que não foi observado nos não-músicos. A nível comportamental, músicos demonstram maior acurácia no

reconhecimento de emoções expressas através da prosódia, assim como no reconhecimento de emoções expressas verbalmente em contexto de ruído. Deste modo, a evidência parece indicar que o treino musical potencia as capacidades de perceção auditiva e que estas se transferem do domínio musical para o domínio linguístico e emocional (*cross-domain effect*).

A emoção é uma propriedade bastante relevante no processamento auditivo de vozes porque transmite informação sobre o estado emocional do locutor. Estudos com potenciais evocados relacionados a eventos (*event-related potentials* – ERP) mostram que as vocalizações com conteúdo emocional são processadas mais rapidamente do que vocalizações neutras e recrutam mais recursos atencionais e cognitivos. Estes estudos, juntamente com estudos de ressonância magnética funcional, contribuíram para a conceção do modelo do processamento de vocalizações proposto por Schirmer e Kotz (2006). Este modelo sugere um processamento da informação vocal em três diferentes fases: (1) processamento sensorial, tendo por base a análise sensorial das propriedades acústicas da voz, (2) integração, envolvendo a extração da saliência emocional, e (3) cognição, em que processos cognitivos de avaliação do estímulo sonoro tomam lugar. Este modelo foi atualizado por Frühholz, Trost e Kotz (2016) que o estendeu para o processamento de emoção noutros tipos de sons, incluindo sons musicais.

As três fases diferenciadas neste modelo são indexadas por componentes eletrofisiológicos, nomeadamente o N1, que ocorre cerca de 100 ms após a apresentação do estímulo auditivo, o P2, que ocorre cerca de 200 ms após o início do som, e o potencial tardio (*late positive potential* – LPP), que ocorre numa janela temporal entre 500 e 900 ms após o início do som. Estudos anteriores mostraram que a expertise musical modula as respostas dos componentes correspondentes às três fases do modelo do processamento emocional auditivo. Contudo, as diferenças entre o treino musical vocal e o treino musical instrumental não tinham sido ainda abordadas no âmbito dos seus efeitos no processamento auditivo. Ademais, nos estudos anteriores não foram analisadas potenciais diferenças do efeito do treino musical em função da categoria do som, ou seja, se o treino musical afetaria da mesma forma o processamento de emoções veiculadas através de sons vocais vs. através de sons musicais. A verdade é que os cantores se distinguem dos restantes grupos de músicos pelo facto de conterem o seu instrumento (a voz) incorporado no corpo, o que implica a ativação de um mecanismo motor diferente. O controlo motor do aparelho vocal conta com a participação de recetores somatosensoriais que participam nos mecanismos de *feedback* envolvidos na manutenção da voz. Consequentemente, há evidência de que os cantores dependam mais do feedback somatosensorial do que os instrumentalistas.

O presente estudo teve como objetivo analisar os efeitos do treino musical vocal e instrumental no curso temporal dos mecanismos neurocognitivos subjacentes ao processamento implícito de emoções, através da análise de ERPs, com o uso da eletroencefalografia (EEG). Em particular, analisámos diferenças entre sons vocais e musicais afetivos na amplitude dos componentes de ERP associados a fases iniciais (N1, P2) e fases tardias (LPP) do processamento auditivo. Desta forma, o estudo atual permitiu comparar diretamente, pela primeira vez, os efeitos do treino musical vocal com os efeitos do treino musical instrumental no processamento auditivo de emoções, bem como analisar a relevância do tipo de estímulo (vocal ou musical). Foi ainda realizada uma tarefa comportamental com o objetivo de obter avaliações subjetivas de cada participante aos sons ouvidos durante a tarefa de EEG. As propriedades afetivas avaliadas foram a valência – quão negativo ou positivo um som é percecionado, a ativação – quão calmo ou ativo o som faz o sujeito sentir – e a categoria emocional – categorização do som enquanto alegre, triste ou neutro.

Os cantores e instrumentalistas que participaram tinham pelo menos seis anos de treino musical e treino regular do instrumento. Dezanove cantores, dezassete instrumentalistas e vinte e dois controlos ouviram vocalizações humanas não-verbais e excertos musicais expressando alegria, tristeza e tom neutro durante o registo de um eletroencefalograma. Vocalizações de primatas e sons de pandeireta serviram de sons alvo, direcionando a atenção dos participantes para esses sons durante a tarefa de EEG.

Os resultados revelaram diferenças entre grupos ao nível da P2. Foi observada uma amplitude mais positiva desse componente para vocalizações em comparação com sons musicais, apenas nos participantes controlo. Por sua vez, os dois grupos de sujeitos com treino musical (instrumentalistas e cantores) não apresentaram diferenças na amplitude de P2 gerada por sons musicais em comparação com vocalizações humanas. Além disso, foram encontradas diferentes respostas dos potenciais evocados em resposta a sons musicais comparativamente a sons vocais: as vocalizações estiveram associadas a uma amplitude mais reduzida de N1 e a amplitudes mais positivas de P2 e de LPP. A nível comportamental, foram ainda encontradas diferenças entre grupos na atribuição de classificações de valência: os cantores atribuíram valores mais positivos aos sons musicais em comparação com os participantes controlo.

Em conjunto, estes resultados sugerem que o treino musical tem influência na fase de extração de saliência emocional do processamento auditivo emocional. Especificamente, há evidência de que os sons musicais podem ser tão salientes quanto as vozes humanas. Para além disso, as diferenças encontradas em função do tipo de estímulo apontam para a ocorrência de uma diferenciação da categoria do som em fases inicias e tardias do processamento auditivo. Deste modo, estes resultados corroboram a ativação dos mecanismos de neuroplasticidade funcional subjacentes ao treino musical, que se refletem no aumento das habilidades de perceção auditiva em músicos, nomeadamente no processamento de emoções. Assim, constituem evidência de um *cross-domain effect*.

Palavras-chave: Potenciais evocados; Processamento Emocional Auditivo; Treino Musical; Neuroplasticidade.

INTRODUCTION

1. MUSICAL TRAINING: A MODEL OF EXPERIENCE-DRIVEN NEUROPLASTICITY

The human brain develops from simple cells to a complex organ that controls the human body and its experiences (Stiles & Jernigan, 2010). Brain function and structure continues to change during lifetime due to practice (Herholz & Zatorre, 2012), environmental stimuli or due to alterations that occur in the nervous system, such as molecular modifications that can affect channel configuration or gene expression (Kolb & Gibb, 2011). This process is known as neuroplasticity (Demarin et al., 2014; Kania et al., 2017). Neuroplasticity occurs at two levels: the first concerns changes in the synaptic transmission process, such as the modification in the strength between neurons; the second relates to the ability of the brain to adapt to the environment (Demarin et al., 2014) through learning and memory processes (Kania et al., 2017). The first level is called synaptic plasticity and is dominant during gestation to prepare the brain to interact with the world (Demarin et al., 2014; Lohmann & Kessels, 2014). The second level is known as functional neuroplasticity, and it plays a major role during adolescence and adult life to allow the brain to change towards injury and inputs from the environment (Demarin et al., 2014). These changes are usually related to experience or training (Herholz & Zatorre, 2012), representing endless possibilities of adaptation of the human brain (Kolb & Gibb, 2011).

Musical performance, whether via a physical instrument or the voice itself, requires a repetitive sensorimotor training over long time periods (Herholz & Zatorre, 2012; Kleber et al., 2016) putting in constant interaction several modalities and skills, such as auditory perception, motor control, visual perception, and memory (Barrett et al., 2013; Lappe et al., 2008; Schlaug, 2015). Thus, the effects of musical training rely on neural plasticity: based on the brain's plasticity mechanisms one can learn and master an instrument but that experience also produces behavioral alterations, as well as structural and functional changes in the brain (Herholz & Zatorre, 2012; Musacchia et al., 2007; Schlaug, 2015). Therefore, the musician's brain is considered a model of experience-driven neuroplasticity (Intartaglia et al., 2017; Lima, 2011; Schlaug, 2015), showing structural change (Gaser & Schlaug, 2003; Habib & Besson, 2009; Lappe et al., 2008; Schlaug, 2015).

Structural differences have been reported in the structure responsible for interhemispheric communication (Fitsiori et al., 2011), the corpus callosum. Magnetic Resonance Imaging (MRI) with voxel-based morphometry studies showed larger parts of the corpus callosum not only in adult musicians when compared to matched adult non-musicians (Oztürk et al., 2002; Schlaug et al., 1995), but also when comparing children with only 15 months of musical training with children without musical training (Hyde et al., 2009).

Other morphometrical differences have been found in various motor, auditory, and visual regions between musicians and non-musicians regarding grey matter density (Gaser & Schlaug, 2003; Schlaug, 2015; Sluming et al., 2002). For instance, an increased concentration of grey matter was found in perirolandic regions (primary motor and somatosensory areas, premotor areas, superior parietal areas, and inferior temporal gyrus bilaterally) of professional musicians when compared to non-musicians, and even with amateur musicians (Gaser & Schlaug, 2003). That is also true for the left cerebellum, left Heschl's gyrus, and left inferior frontal gyrus, specifically, in the Broca's area (Gaser & Schlaug, 2003; Sluming et al., 2002).

1.1. Vocal Musical Training – The Special Case of Singers

Singing involves the generation of a succession of musical sounds using the voice (Jones & Keough, 2008), combining language and speech properties as well as musical competences (Callan et al., 2006; Jungblut et al., 2012).

Several structures of the vocal tract coordinate the properties of singing. The fundamental frequency (F0), or perceived pitch, is determined by the way vocal folds vibrate at a particular rate after the air passes the opening of the larynx (Sundberg, 1987). Then, the laryngeal muscles are responsible for modifying the vocal folds' length to regulate and maintain a stable pitch (Sundberg, 1987). Besides that, other structures ensure the airflow and control the intensity of vocalizations such as the diaphragm, thoracic muscles and articulatory muscles (Sundberg, 1987).

When probed with fMRI, singing was found to recruit several brain regions, most of them involved also in the production of speech: pre- and post-central gyrus (Mavridis & Pyrgelis, 2016), cerebellum (Jones & Keough, 2008; Perry et al., 1999), primary auditory cortex (Brown et al., 2004; Perry et al., 1999; Zarate et al., 2010), primary motor cortex (Brown et al., 2004; Jones & Keough, 2008; Perry et al., 1999; Wilson et al., 2011), auditory association cortex (Brown et al., 2004), supplementary motor area (SMA) (Brown et al., 2004; Jones & Keough, 2008; Perry et al., 1999; Wilson et al., 2011), frontal operculum (Brown et al., 2004; Perry et al., 1999), medial frontal gyrus (MFG) (Wilson et al., 2011), and left insula (Jones & Keough, 2008; Perry et al., 1999; Zarate, 2013). Nonetheless, there are differences in auditory activation for singing vs. speaking, such as opposite hemisphere asymmetries in both motor and auditory regions with a tendency for a right-sided asymmetry during singing (Perry et al., 1999). For instance, singing activates the right temporal lobe while speech activates the left temporal lobe (Callan et al., 2006). The perisylvian cortex also shows a right-side activation that seems to be crucial for the singing activity, considering that aphasic patients with left hemispheric lesions can sing but not utter the same words (Mavridis & Pyrgelis, 2016). Furthermore, activation in some brain regions plays a central role in response to singing compared to speaking: the right ventral precentral gyrus (Perry et al., 1999), the anterior cingulate cortex (ACC) (Brown et al., 2004; Jones & Keough, 2008; Perry et al., 1999; Zarate et al., 2010), the superior temporal gyrus (STG), particularly the anterior planum polare (Brown et al., 2004; Mavridis & Pyrgelis, 2016). There is evidences for a functional connection between the auditory cortex, the ACC and the insula during singing (Zarate et al., 2010).

All of us can experience singing and keeping up with certain notes to follow a melody, but expert singers do that on a more frequent basis to achieve higher level performance. There is evidence that vocal training is associated with reduced activation of the language network for singing activities (Wilson et al., 2011), and with the use of partially different singing networks: singers showed greater bilaterally activity in the middle frontal gyrus and in the right inferior frontal gyrus while the right frontal lobe is more activated during singing in nonexperts (Henson, 1985; Wilson et al., 2011; Yamadori et al., 1977).

Singers are a special musician’s group because they use an instrument contained within their body while attaining an emotional interpretation, requiring special sensorimotor training over long periods of time to exercise finer vocal control during difficult singing tasks (Kleber et al., 2016; Zarate, 2013). Motor control of the vocal apparatus has the participation of somatosensory receptors (Kleber et al., 2017; Sundberg, 1987), present in the articulatory muscles (Sundberg, 1987), and of auditory feedback mechanisms, which is especially relevant for the maintenance of the ongoing voice production (Amir et al., 2003; Sundberg, 1987).

fMRI studies revealed increased functional activation of bilateral primary somatosensory cortex in expert classical singers along with the larynx and the articulators to send feedback information for vocal maintenance (Kleber et al., 2010). Moreover, while participants needed to ignore a masked feedback, singers displayed different patterns of brain activation, exhibiting more activity within auditory cortices (posterior STG and STS) and in the putamen than non-musicians (Zarate, 2005; Zarate & Zatorre, 2008). Those results introduced evidence for a greater reliance on internal models for F0 production in singers during singing tasks (Jones & Keough, 2008; Kleber et al., 2010, 2016). Internal model refers to a neural mechanism that simulates the response of the motor system in order to achieve a certain output (Max et al., 2004). To better understand these internal models for vocal pitch control, some researchers created a condition where some of the participants who had vocal training performed the tasks with local vocal-folds anesthesia (Kleber et al., 2013; Sundberg et al., 1995). In that condition, the sensorimotor activation decreased along with the putamen and the cerebellum (Kleber et al., 2013), leading to greater pitch deviance in singers (Kleber et al., 2013; Sundberg et al., 1995). The right anterior insula was also downregulated (Kleber et al., 2013). Surprisingly, non-singers exhibited better behavioral performance when under that anesthesia effect, which corroborates the training-related dependence of singers on sensorimotor feedback for vocal control (Kleber et al., 2013).

Likewise, with extensive vocal training and practice, the anterior insula seems to be responsible to initiate the somatosensory feedback response (Kleber et al., 2017; Zarate & Zatorre, 2008) and starts to integrate sensory feedback processing with vocal motor control (Brown et al., 2004; Wilson et al., 2011; Zarate, 2013). Evidence for an increased recruitment of the somatosensory cortex for neural motor control during singing as a function of singing expertise has been consistently found (Jones & Keough, 2008; Kleber et al., 2010; Zarate, 2013). Singers also seem to rely more on somatosensory feedback than instrumentalists (Christiner & Reiterer, 2015; Halwani et al., 2011)

2. THE INFLUENCE OF MUSICAL TRAINING ON THE AUDITORY SYSTEM

The auditory system is one of the brain systems that show more changes driven by musical training, not only structurally, but also functionally (Herholz & Zatorre, 2012). Structural changes can be linked to volume alterations or in grey matter density levels, as those mentioned above, while functional changes are related to adaptations of neuronal activity towards a specific function such as sensory representations (de Villers-Sidani & Merzenich, 2011) or rearrangement of cortical maps modulated by sensory experience, for example (Kandel et al., 2000).

The functions of the auditory system rely on the processing of basic sound features, such as pitch (Trainor, 2010). Electroencephalography (EEG) is an appropriate technique to measure responses from the auditory cortex to those sound features (Trainor, 2010) since it affords to track neurocognitive processes as they happen in real time from the moment the stimulus is presented until a response is made (Light et al., 2010). It can provide information about the oscillatory activity of different frequency bands (spectral analysis) (Kim & Im, 2018) or about electrical potentials that are time-locked to events (ERPs) (Cacioppo et al., 2007). Spectral analysis has been correlated to human behavior, cognitive state, or mental illness (Kim & Im, 2018). ERPs are voltage fluctuations that reflect the summed activity of neuronal populations. They reflect brain activity in response to distinct events and can be considered as neural manifestations of specific functional mechanisms (Cacioppo et al., 2007). EEG studies have reported differences between musicians and non-musicians in the processing of sound properties. As expected, musicians were found to exhibit increased amplitude of EEG spectrum elicited by beat and meter compared to non-musicians (Celma-Miralles & Toro, 2019), increased N1 – negative ERP component elicited in the presence of auditory stimuli (Neumann & Sanders, 1996), and P2 – positive ERP elicited by auditory stimuli that varies with selective attention and encoding of sound (Ferreira-Santos et al., 2012) – amplitudes in response to music compared to vocal expressions observed in musicians in comparison with a group without musical expertise (Rigoulot et al., 2015), and earlier (latency) and larger (amplitude) brainstem responses recorded with EEG to musical tones when compared to non-musicians (Musacchia et al., 2007).

Nonetheless, studies demonstrate the enhancement effect of musical training on auditory perception outside the music field. Faster brainstem responses to speech sounds were observed in musicians versus non-musicians (Musacchia et al., 2007). Parbery-Clark et al. (2011) observed enhanced ability to detect sound patterns in response to speech syllables in musicians in comparison to non-musicians, and Strait and colleagues (2009) found earlier brainstem responses in musicians in the processing of pitch, timing and timbre of nonverbal vocalizations, compared to non-musicians.

Pitch is a sound property that has been frequently shown to elicit different brain responses in musicians when compared to non-musicians. It is defined as the perceived repetition rate of an acoustic waveform (Oxenham, 2012). In music, the melody of a song is obtained from sequences of pitch; in speech, the intent gains different meanings because of pitch contour, either affecting prosody or word meaning (Oxenham, 2012). Firstly, musicians demonstrate reduced frequency discrimination thresholds for non-musical pure tones (Bianchi et al., 2016; Kishon-Rabin et al., 2001) and for complex tones (Micheyl et al., 2006), and exhibited faster detection of pitch change in psychoacoustic tasks (Nikjeh et al., 2008). Bianchi and colleagues (2016) also measured pupil dilation during pitch discrimination tasks and found less dilation among musicians in comparison to non-musicians, suggesting a lower processing effort. Secondly, there is also evidence for more robust encoding of linguistic pitch in brainstem recordings of musicians contrasted with non-musicians reflected in larger amplitude of frequency following responses (FFR) (Wong et al., 2007), a measure that reflects neural activity of the brainstem (Aminoff et al., 2015). One interesting study compared the FFR in response to a mandarin lexical tone and a musical pitch interval between English musicians, English non-musicians and native speakers of Mandarin Chinese (Bidelman et al., 2011). For the pitch-tracking accuracy, there were no differences between the Chinese and English musicians in both domains, whereas non-musicians exhibited lower accuracy in that measure (Bidelman et al., 2011).

A vast research performed with singers is dedicated to the understanding of the mechanisms underlying vocal pitch production and control in singing. That is not surprising considering that even though the regulation of vocal pitch is important for all human beings for speech production, it becomes crucial for singers because they need their vocalizations to match a certain output that varies from the beginning to the end of a song. One study revealed that when asked to produce a note and maintain it throughout frequency-altered feedback, singers did not present compensatory modifications in the vocal output as much as non-singers (Jones & Keough, 2008). After returning to an unaltered feedback, higher F0 values were found in comparison to baseline and control values, suggesting that singers recalibrated the connection between pitch output and the motor system that controls it. That finding was consistent with other work concerning the effects of vocal training on pitch control: non-musicians seem to be less able to ignore pitch-shifted feedback than singers (Kleber et al., 2013, 2017; Rosslau et al., 2016; Zarate & Zatorre, 2008). Similar results can be found when comparing singers with other musicians (instrumentalists). When comparing instrumentalists and singers in a task of detecting pitch deviances of 1.5%, 3%, and 6% in a mismatch negativity paradigm (MMN), singers were more accurate in the smallest pitch change condition (Nikjeh et al., 2008). Mismatch negativity (MMN) is a negative ERP component in an EEG signal when in the middle of a repetitive sound presentation, a different occasional sound is perceived (Legatt, 2015). It is an automatic (pre-attentive) response that reflects the identification of an unattended and task-irrelevant deviant stimulus (Legatt, 2015). Vocal musicians who also knew how to play another instrument exhibited earlier MMN responses (earlier latency) to all conditions, indicating a possible auditory neural advantage of vocal musicians with instrumental training over musicians with instrumental or vocal training only (Nikjeh et al., 2008).

A study focused on the structural adaptation in musicians of the arcuate fasciculus (AF), a bundle of white-matter fibers that connect regions of the brain involved in sound production and perception, feedback control, and vocal output showed that musicians (singers and instrumentalists) have more white matter density in the right and left AF (Halwani et al., 2011). The right dorsal branch of the arcuate fasciculus contains the connection between the superior temporal gyrus and the inferior frontal gyrus that has been associated with fine-grained control of pitch (Loui et al., 2009). This structural adaptation might be linked to the enhanced pitch perception of musicians described before (Bianchi et al., 2016; Bidelman et al., 2011; Kishon-Rabin et al., 2001; Kleber et al., 2013; Micheyl et al., 2006; Nikjeh et al., 2008; Zarate & Zatorre, 2008). In its turn, the singers group exhibited significant larger tract volume in the left dorsal arcuate fasciculus in comparison to both instrumentalists and non-musicians (Halwani et al., 2011). The electrophysiological differences between singers and instrumentalists reported by Nikjeh and colleagues (2008) might align with the evidence for the left AF as the structure that adapts the most to the specific demands of vocal control and integration with the motor and auditory systems while singing (Halwani et al., 2011).

Various studies have replicated the effect of musical training influence on brain responses during speech processing (Anderson & Kraus, 2011; Coffey et al., 2017; Intartaglia et al., 2017; Lima & Castro, 2011; Musacchia et al., 2007; Parbery-Clark et al., 2009, 2011, 2012; Pinheiro et al., 2015; Strait, D. L. et al., 2012; Strait et al., 2014; Zendel et al., 2015). At least three different aspects of speech elicit stronger and/or faster brain responses and higher accuracy in behavioral measures in musicians when compared to non-musicians: speech sound processing (Parbery-Clark et al., 2009, 2012; Strait et al., 2014), speech-in-noise perception (Anderson & Kraus, 2011; Coffey et al., 2017; Parbery-Clark et al., 2012; Strait, D. L. et al., 2012; Zendel et al., 2015), and emotional speech prosody decoding (Lima & Castro, 2011; Pinheiro et al., 2015). For instance, musicians have demonstrated larger amplitudes of brainstem responses to changes in speech sounds (Parbery-Clark et al., 2012; Strait, D. L. et al., 2012). A perceptual advantage for speech-in-noise perception was reflected in larger brainstem responses to shifts on speech sounds (Strait et al., 2014), similar electrical patterns of brainstem response in normal and noise conditions, suggesting that background noise does not degrade subcortical response to speech (Parbery-Clark et al., 2009), and electrophysiological evidence of enhanced lexical processing when background noise is accentuated (increased amplitude of the component indexing lexical access, the N400) (Zendel et al., 2015). Of note, a group of researchers studied the advantage of musicians in the speech domain with a foreign language task to compare neural responses to the formant frequencies (Intartaglia et al., 2017), a relevant cue in determining the phonetic content of speech sounds (Holmes et al., 1997). There were no differences between non-native musicians and native non-musicians speakers, suggesting that musical training strengthens the neural encoding of linguistic acoustic information at a native speaker level (Intartaglia et al., 2017).

These results support early and later processing changes in the brain in consequence of musical training as well as a cross-domain effect of that expertise. A cross-domain effect occurs when the skills developed in the domain of expertise (musical auditory skills enhanced in the context of musical training) transfers to other domains (auditory processing of speech sounds, which is not directly related to music) (Kraus & Chandrasekaran, 2010), opposite to within-domain effects where the expertise applies only to aspects of the same domain (musical sound processing, rhythm, beat, etc.) (Kraus & Chandrasekaran, 2010).

3. AUDITORY EMOTIONAL PROCESSING

In social interaction, the emotional state of the speaker might be expressed vocally and capture the attention of the listener (Lang & Davis, 2006). Emotional information is a key element of interpersonal relationships because it allows individuals to shape their response to others (Ding et al., 2017; Jacob et al., 2014). Emotions can be expressed with the face, where expressions change towards a specific emotional state (Ekman, 1992), and with body posture, based on the interpretation of body movements (Knapp et al., 2014). Also, emotions can be decoded from vocal cues, which is important because they can be communicated in the absence of visual contact and at larger distances (Hawk et al., 2009). Actually, vocal signals have been shown to influence the visual analysis of faces (Rigoulot & Pell, 2014). According to dimensional theories, there are two dimensions that describes an emotion: (1) how positive or negative an emotion is – valence; (2) how intense the emotional state activated is – arousal (Citron et al., 2014). Positive and negative emotions may be expressed through nonverbal vocalizations (crying, screaming, laughter) or through suprasegmental modulations of speech such as in emotional speech prosody (Hawk et al., 2009). Speech prosody is considered suprasegmental information concerning the duration, intensity and intonation (F0 contour, perceived as pitch) with which words are spoken (Koolagudi & Rao, 2012; Schirmer & Kotz, 2006). Different emotions are reflected by different combinations of these features. For example, happiness is characterized by high intensity and fast speech rate. By contrast, sad vocalizations are characterized by low intensity, slow speech rate, and steady pitch but with a high presence of spectral noise (Schirmer & Kotz, 2006). These features convey essential non-linguistic information in the context of spoken language, i.e., during verbal speech production (Pell et al., 2015). Nonverbal vocalizations such as laughter, cries, or screams are emotional vocal expressions considered 'pure' or 'primitive' because they are not contaminated by verbal information (Lang & Bradley, 2010; Schirmer & Kotz, 2006). Behavioral tasks comparing the recognition of emotion embedded in speech prosody and in nonverbal vocalizations demonstrate that affective vocalizations without verbal content are related with higher accuracy in emotion recognition tasks (Hawk et al., 2009).

Nonverbal communication may have preceded verbal communication (Burgoon et al., 2016), enabling social exchange and adaptation: nonverbal processing involves sensory input that is sent to the amygdala along with memory information from the hippocampus (Lang & Bradley, 2010). The amygdala then connects with circuits responsible for sensory and information processing and structures of the autonomic and somatic systems that are responsible for producing an action that can be defensive or appetitive (Lang & Bradley, 2010). Those systems are crucial for the survival of the individuals and have a primitive function of increasing attention, facilitating information gathering and processing, and preparing the organism for action (Lang & Bradley, 2010; Lang & Davis, 2006). This might be the reason why nonverbal cues usually have greater salience in comparison to verbal cues (Charest et al., 2009; Jacob et al., 2012, 2014; Pell et al., 2015) and why they are accurately recognized by different cultural groups, specially vocalizations that express primarily negative emotions (disgust, anger and sadness) (Sauter et al., 2010).

Schirmer and Kotz (2006) proposed a model of the time course of vocal emotional processing. Three stages were identified: (1) sensory processing – the first subprocess taking place in the auditory cortex is the analysis of acoustic features of the vocal sound, specifically in the superior temporal sulcus (STS); (2) integration – secondly, significant acoustic information is processed to detect emotional meaning in the bilateral superior temporal gyrus (STG) and STS; (3) cognition – thirdly, cognitive evaluation processing of the emotional information detected in the previous stage, such as evaluative judgments of emotion driven by the right inferior gyrus (IFG) and the orbitofrontal cortex (OFC). The model predicts influence of individual significance at any of the stages. A more recent view of the auditory emotional processing model was proposed by Frühholz, Trost and Kotz (2016), considering other types of auditory stimuli, such as musical sounds. Other brain areas were included to take part of a unified neural network of affective sound processing. For instance, along with key sensory cortical areas, such as the superior temporal cortex (STC) and the auditory cortex, the processing of nonverbal vocalizations requires the involvement of the amygdala for valence discrimination, which communicates with the geniculate nucleus (MGN) during auditory signal transmission (Frühholz et al., 2016). The auditory cortex and the amygdala are both highly involved in auditory emotional processing, but their role can be more or less relevant depending on the features of the sound. When sounds are complex *per se* the auditory cortex might intervene first; when emotion is being invoked, the amygdala might be the first to respond (Frühholz et al., 2016).

The processing of music emotions seems to recruit similar brain regions when compared to the processing of vocal emotions, namely the auditory cortex, the STC, and the MFC (Frühholz et al., 2016). The reward systems of the brain, namely the nucleus accumbens

(NAcc), and the hippocampus (HC) respond specially to musical emotions compared to vocalizations (Frühholz et al., 2016). Overall, vocalizations seem to mainly activate the brain pathways of evaluation while music preferentially recruits memory-associated emotional processes (Frühholz et al., 2016).

The multi-stage model of auditory emotional processing presented by Schirmer and Kotz (2006) is also confirmed electrophysiologically and reflected in modulations of specific ERP components. ERPs are time-locked to a specific stimulus and represent small changes in electrical potentials related to stimulus processing (Britton et al., 2016). Specifically, the N1, P2, and Late Positive Potential (LPP) constitute the ERP components that might underlie the three stages of the multi-stage model of auditory emotional processing, as it will be explored in the next section.

3.1. Effects of Musical Training on Auditory Emotional Processing

In addition to overall sound and linguistic processing differences found between musicians and non-musicians, are auditory emotions also differently processed by musicians? Individuals with musical training seem to exhibit facilitated processing and increased recognition accuracy of emotions in music when compared to non-musicians (Castro & Lima, 2014; Nolden et al., 2017; Sharp et al., 2019), but also in the human voice (Paquette et al., 2018; Pinheiro et al., 2015). At the brain level, the effects of musical training were observed in different measures: (1) through EEG, enhanced amplitude of oscillatory activity of frontal alpha band for emotional music in musicians compared to non-musicians, speech and vocalizations (Nolden et al., 2017); (2) via brainstem recordings during the presentation of an infant's unhappy cry, musicians exhibited an earlier and larger brainstem magnitude response to the spectrally complex portion of the sound (aperiodic) and smaller amplitude response for the periodic portion of the sound, demonstrating higher efficiency of emotional sound processing (Strait et al., 2009).

Behavioral findings demonstrate enhanced accuracy in recognizing emotions in speech prosody (Lima & Castro, 2011) and in speech under degraded pitch conditions (Fuller et al., 2014) was detected. Notably, Fuller and colleagues (2014) observed that musicians displayed weaker accuracy recognizing vocal emotions under poor pitch perception conditions, indicating pitch as a key auditory feature for emotional decoding.

The interaction between music and emotion has been vastly explored in neuroscience research. Until now, as described above, research on how musical training influences the brain has supported a cross-domain effect of musical training in auditory perception from sound and speech processing to emotional processing in what concerns emotions conveyed through auditory stimuli (see Appendix A for table with the literature review).

To the best of our knowledge, there are no studies including singers to examine the interactions between vocal musical training and emotion processing. Nonetheless, we consider that these interactions are highly relevant to be investigated when we take into consideration the evidence found until now about the effects of vocal training. For instance, as summarized before, singers exhibit a distinctive use of sensorimotor cortex that differentiates them from non-singers and other musicians. Sensorimotor cortex activation has been associated with improved performance on the recognition of emotional vocalizations (Banissy et al., 2010) and as part of a mechanism for the emotional interpretation of authentic sounds, such as laughter (McGettigan et al., 2015). It was also consistently observed in singers the activation of the STS that is highly involved in vocal cognition (Belin et al., 2002; Zilbovicius et al., 2006) and it is a central part of the “what” auditory pathway that integrates emotionally significant acoustic information (Schirmer & Kotz, 2006). The posterior STS, that is more engaged during singing than speaking, is recruited during recognition of unfamiliar voices and may be involved in analyzing spectral details of voices over time (Kriegstein & Giraud, 2004). Probably resulting from this practice-dependent functional reorganization, pitch discrimination accuracy and maintenance is enhanced in singers in comparison to non-singers (Banissy et al., 2010; Jones & Keough, 2008; Kleber et al., 2016), other musicians (Nikjeh et al., 2008), and even actors (Rosslau et al., 2016).

For that reason, we want to expand that knowledge by looking at the effects of musical vocal training on emotional processing in the auditory modality and compare them with the effects of musical instrumental training. Given the importance of nonverbal vocalizations and considering the neuronal substrates underpinning their processing, we the current work aimed to evaluate how different levels of musical expertise (no musical expertise, instrument expertise, and singing expertise) modulates the processing of human nonverbal vocalizations and musical bursts expressing emotional (happiness and sadness) vs. neutral information.

This research topic is relevant as it may provide knowledge of experience-dependent plasticity of the vocal system (Kleber et al., 2010) with an application to emotional processing, a mechanism that is impaired in several psychiatric disorders (Pinheiro et al., 2013; Pinheiro et al., 2017; Zilbovicius et al., 2006). The understanding of how musical training modulates the social brain (in particular, the processing of emotional cues) can open possibilities for important therapies.

4. ADVANTAGES OF ERPS TO THE STUDY OF AUDITORY EMOTIONAL PROCESSING

ERPs are voltage changes that are time locked to sensory, motor or cognitive events (Picton et al., 2000). They can be extracted from the ongoing (EEG) originated by voltage changes along time, and provide a noninvasive approach to the study of the psychophysiological correlates of mental processes (Sur & Sinha, 2009). The several ERP waveforms averaged together create a sequence of positive and negative voltage deflections that reflect sensory, motor or cognitive processes, known as components (Luck, 2005). Components can reflect physical characteristics of the event being processed (exogenous) and/or to intrinsic processing mechanisms associated to that event, such as cognitive functioning (endogenous), allowing to examine emotional processing (Picton et al., 2000). The labels given to ERP components include information of polarity and latency: for example, N1/N100 is a negative peak (polarity – N) occurring at approximately 100 ms after stimulus onset (latency – 100 ms).

Different sensory modalities elicit distinct components that differ both in waveshape and scalp distribution as a function of the sensory modality stimulated (Cacioppo et al., 2007). In the auditory modality, the N1, P2 and Late Positive Potential (LPP) have been consistently reported in ERP studies with emotional vocalizations (Jessen & Kotz, 2011; Liu et al., 2012; Pell et al., 2015; Pinheiro et al., 2013; Pinheiro et al., 2017). They reflect the three stages of emotional processing of auditory cues (Schirmer & Kotz, 2006): the N1 is thought to represent the activity of the sensory pathway, being modulated by acoustic properties of the sound; the P2 may reflect the signal transmission to central processing systems that detect emotional salience; LPP underlies the cognitive processing of the sound. Individual differences may be also reflected in ERP modulations (Cacioppo et al., 2007).

The study of emotional processing can focus on explicit emotional processing, asking participants to identify some affect properties of a stimulus (emotional or valence categorization, for example), or on implicit emotional processing, with participants paying attention to non-emotional properties of the stimuli (gender of the voice, for example) or to a different target stimulus (Cohen et al., 2016). Consequently, ERP components might be affected by attention being directed or not to the emotion. For instance, there is evidence for enhanced LPP amplitude when the task is explicit compared to an implicit task (Schupp et al., 2006).

Findings of recent studies suggest that the N1, a component generated in the supratemporal auditory cortex around 100 ms after stimulus onset (Onitsuka et al., 2013), may index the earliest point at which emotional and non-emotional stimuli can be distinguished, even though it does not provide a robust distinction between valence types (Jessen & Kotz, 2011; Liu et al., 2012). The N1 effect is more prominent over anterior electrode sites. Previous studies on emotional vocalizations processing using ERPs demonstrated valence effects on the N1 component: reduced N1 amplitude for angry and happy versus neutral vocal sounds but no differences between angry and happy vocalizations (Liu et al., 2012), and reduced N1 amplitude for fearful interjections compared to neutral ones (Jessen & Kotz, 2011). Sauter and Eimer (2010) also reported differentiation between emotional and neutral vocalizations as early as 150 ms after sound onset. The N1 occurs regardless of attention being directed to the stimulus but may be affected by it (Cacioppo et al., 2007).

The P2 is a positive deflection occurring around 200 ms post-stimulus onset that has been found to be modulated by characteristics of the stimulus and also by aspects of attention and cognitive processes underlying stimulus processing (Ferreira-Santos et al., 2012). P2 has been proposed to represent an index of the extraction of emotional salience, being the first component that indicates processing differences between types of emotions (Liu et al., 2012; Paulmann & Kotz, 2008; Pinheiro et al., 2013). Indeed, some studies reported a reduced P2 amplitude for sad vocalizations in comparison to happy and angry (Pell et al., 2015), and differences between angry and fearful vocal sounds, enhanced for angry (Jessen & Kotz, 2011). Moreover, enhanced P2 amplitude for emotional compared to neutral non-verbal vocalizations (Jessen & Kotz, 2011; Liu et al., 2012), and enhanced amplitude for neutral compared to emotional prosody (Paulmann & Kotz, 2008; Pinheiro et al., 2015) were observed.

The LPP is an ERP component that reflects sustained attention to stimuli conveying emotional information, being reportedly enhanced for emotional compared to neutral information (Dennis & Hajcak, 2009). Hence, this component is elicited during the last stage of affective auditory decoding, the cognitive processing, being larger for stimuli that are processed at a deeper level (Münte et al., 2000). Arousal modulates the LPP amplitude, being more positive for more arousing stimuli both pleasant and unpleasant

(Hajcak et al., 2010). For instance, LPP amplitude has been observed to be enhanced for angry versus sad and happy vocalizations (Pell et al., 2015), and for fearful in comparison to angry audiovisual stimuli (Jessen & Kotz, 2011). Moreover, different emotional prosodies have been shown to be distinguished from one another in this stage, with more arousing categories such as disgust eliciting a more positive amplitude of the component (Paulmann et al., 2013). In the context of speech processing, speaker identity also has a modulatory effect on LPP amplitude: emotional self-speech elicits a more positive LPP than speech produced by others (Pinheiro et al., 2016). Overall, the LPP effect seems to be driven by sustained emotional evaluation (Paulmann et al., 2013) anchored on motivational salience of the stimulus (Münte et al., 2000), with its maximal effect over centroparietal electrodes (Pell et al., 2015).

ERP studies on auditory emotional processing with musicians focusing on distinct components (N1, P2, MMN, P3) have observed enhanced amplitudes in musicians in comparison to non-musicians (Fujioka, 2006; Pinheiro et al., 2015; Rigoulot et al., 2015; Shahin et al., 2004). For example, one study investigated how musical training modulates emotional prosody processing, presenting sentences with neutral semantic content in two different conditions: intelligible semantic content (semantic content condition – SCC) and unintelligible semantic content (pure prosody condition – PPC). The sentences had neutral, happy or angry intonation. The analysis focused on the P50, N1 and P2 components. The two sentence conditions elicited a similar N1 in the musician's group only. Furthermore, musicians were also the only group exhibiting P2 amplitude differences between SCC and PPC conditions. These results suggested that musical expertise is associated with changes in vocal emotional processing mechanisms (Pinheiro et al., 2015). Another study compared how musicians and non-musicians responded to emotional auditory stimuli in the form of vocalizations and musical excerpts (Rigoulot et al., 2015). The N1 and P2 components revealed increased amplitude for musicians for musical sounds in comparison to vocal sounds. In subjects without musical training a more positive P2 was observed for vocal sounds instead (Rigoulot et al., 2015). Appendix 2 presents a table that summarizes the studies probing emotional vocal processing with EEG/ERP.

In sum, practice-related changes can be found in the early and later processing of auditory information extending from music to the speech domain. Ultimately, it extends to emotionally communicative utterances (Castro & Lima, 2014; Lima, 2011; Nolden et al., 2017; Pinheiro et al., 2015), evidencing a cross-domain effect of musical training.

5. AIMS OF THE STUDY

The present study aimed to enhance our understanding of the effects of musical training on emotional processing by investigating the temporal course of the neurocognitive mechanisms underlying implicit processing of emotional meaning conveyed through human nonverbal vocalizations and instrumental musical sounds. We chose a task where participants were not paying direct attention to the sounds of interest, an implicit task, to assess automatic responses to emotional stimuli (Liu, et al., 2012).
By comparing a group of instrumental players with a group of singers we aimed to explore the specificity of musical training on auditory emotional processing mechanisms, manipulating the emotional valence (positive, negative and neutral), to investigate: (I) if the vocal training produces enhancement of emotional processing skills exclusively for vocal cues or if there is a cross-domain effect for other type of auditory emotional stimuli; (II) if instrumental training produces enhancement of emotional processing skills exclusively for musical cues or if extends to the processing of vocalizations. The N1, P2, and LPP were the focus of the ERP analysis, considering these components as relevant indexes of the temporal processing of auditory stimuli.

Based on the electrophysiological studies reviewed on emotional auditory processing in musicians and non-musicians (Pinheiro et al., 2015; Rigoulot et al., 2015), we hypothesize a between-group effect at early and late processing stages: facilitated processing of emotional vocal information in singers and facilitated processing of musical sounds in instrumentalists when compared to the other groups, reflected in modulatory differences on N1, P2, and LPP (reduced N1 and a more positive amplitude of P2 and LPP) in response to stimuli with emotional cues compared to neutral stimuli.

Given previous behavioral results of emotion recognition with musicians (Castro & Lima, 2014; Lima & Castro, 2011), we hypothesize enhanced musical and vocal emotion recognition in instrumentalists and singers compared to non-musicians.

METHODS

PARTICIPANTS

Sixty-six adults participated in the study: twenty-one singers, twenty-two instrumental musicians, and twenty-three non-musicians (controls). Eight participants had to be excluded from the EEG analysis due to problems in data acquisition, but all participants were included in behavioral analysis. Musicians were recruited from NOVA's Master in Musical Sciences, local music schools, including the *Conservatório de Música de Lisboa* and *Escola Superior de Música de Lisboa* (ESML), local choirs, such as *Setúbal Voz*, *Coro Sinfónico Lisboa Cantat*, *Coro Gulbenkian* and *Coro da Universidade de Lisboa*. Instrumentalists played violin (n = 8), guitar (n = 6), piano (n = 3), piano and guitar (n = 1), transverse flute (n = 1), bass (n = 1), percussion instruments (n = 1) and cello (n = 1). All singers were either professional singers in activity (n = 5) or reported being active on choirs (n = 12) or on other musical projects in the present (n = 3). Both groups of musicians had at least six years of musical training and weekly practice of the instrument (Zhang et al., 2018). Instrumentalists had 11.41 (*SD* = 4.06) and singers had 15.62 (*SD* = 9.74) mean years of musical practice. Participants in the control group were college students. Before participation in the study, all participants read and signed an informed consent form. The study was approved by a local Ethics Committee (Faculdade de Psicologia, Universidade de Lisboa).

Two instrumentalists reported participating in an amateur choir for the last four years. Some singers also had instrumental training of guitar (n = 3), piano (n = 3) and oboe (n = 1). The years of musical training (χ^2 (2) = 43.19, p < .001) and the age of training onset [$F(2, 63)$ = 43.48, p < .001] only differed between controls and musicians. Instrumentalists and singers did not differ in years of musical training (p = 1.000) nor in age of training onset (p = .297). The three groups matched several criteria such education level and social economic status to try to minimize differences and the existence of factors other than the musical training that may influence the results. Musicians and non-musicians did not differ in terms of education level (χ^2 (2) = 5.03, *p* = .081) but significant differences were found regarding chronological age (χ^2 (2) = 12.03, *p* = .002). Post-hoc tests revealed that those differences are between instrumentalists (*M* = 22.86, *SD* = 4.42) and singers (*M* = 30.62, *SD* = 10.26) (*p* = .003) and between controls (*M* = 24, *SD* = 5.21) and singers (*p* = .032).

All participants were native European Portuguese speakers, and reported no auditory or visual deficit, neurological illness, history of electroconvulsive treatment or substance abuse. One male control, three male musicians, and one male singer were ambidextrous; one male singer, and one male control were left-handed as assessed by the Portuguese version of the Edinburgh Handedness Inventory (EHI) (Espírito-Santo et al., 2017). Participants completed the Positive and Negative Affect Schedule (PANAS) (Galinha & Ribeiro, 2005) and the Brief Symptom Inventory (BSI) (Canavarro, 2007). These questionnaires assessed participants' current emotional state, as well as the presence of psychological distress and psychiatric disorders. Additionally, The Goldsmiths Musical Sophistication Index (Gold-MSI) (Lima et al., 2018) was administered to assess self-reported musical skills of all participants. Table 1 presents the participants' sociodemographic and musical background characteristics.

Table 1

Sociodemographic and Musical Background Characteristics of the Participants.

	Non-musicians (*n* = 23)	**Musicians**	
		Instrumentalists (*n* = 22)	**Singers (*n* = 21)**
Age (years)	24 ± 5.21 (19-42)	22.86 ± 4.42 (18-36)	30.62 ± 10.26 (19-60)
Gender (F/M)	11/11	12/9	10/13
Education level[1]	2.04 ± 0.88 (1-3)	1.55 ± 0.48 (1-3)	2 ± 0.77 (1-3)
Gold-MSI score[2]	59.35 ± 15.13 (27-85)	95.18 ± 8.97 (80-114)	102.14 ± 25.83 (77-118)
BSI[3]	0.51 ± 0.43 (0.06 – 1.38)	0.68 ± 0.48 (0.04-1.89)	0.78 ± 0.55 (0.02 – 2.08)
Music Training (years)	-	11.41 ± 4.06 (6-19)	15.62 ± 9.74 (8-38)
Age of onset training (years)	-	10.23 ± 3.38 (5-17)	12.81 ± 4.99 (6-22)
Average practice (hours/day)	-	2.93 ± 1.36 (0.5-5)	1.93 ± 1.04 (0.5-3)

Mean ± SD (range) values are shown. [1]Education level categories were defined as the following: 1 = high school, 2 = bachelor degree, 3 = master degree, and 4 = Ph.D. [2]The Goldsmiths Musical Sophistication Index (Gold-MSI) is scored between 18 and 126. [3]The Global Symptom Index (GSI) of the BSI was considered, which is scored between 1 and 4.

STIMULI

The stimuli consisted of 24 nonverbal affective human vocalizations, 24 musical affective bursts, 10 monkey vocalizations, and 10 tambourine sounds, resulting in 68 different stimuli (Table 2). The human voices were selected from the Montreal Affective Voices (Belin et al., 2008) (8 neutral, 8 sad, and 8 happy sounds), previously validated for the Portuguese population (Vasconcelos et al., 2017). The musical sounds were selected from The Musical Emotional Bursts (Paquette et al., 2013) (8 neutral, 8 sad, and 8 happy sounds). For the target sounds, monkey voices were downloaded from the websites http://www.monkeymania.co.uk/, http://www.findsounds.com/ and http://www.soundbible.com/; the tambourine sounds were downloaded from the website https://www.pond5.com. Subsequently, they were edited into short segments and normalized at peak value (90% of maximum amplitude), using Adobe Audition 3.0 (Adobe Systems. Inc. SanJose, CA). The same procedure of peak normalization was reported for the MAV and MEB sounds (Belin et al., 2008; Paquette et al., 2013).

Table 2
List of stimuli selected from the Montreal Affective Voices (MAV) and from the Musical Emotional Bursts (MEB) and respective values of arousal, valence, and duration | Valence comprehends a range from positive/pleasant to negative/unpleasant, and arousal is defined in a range from calm to arousing, reflecting the intensity of activation (Lang & Bradley, 2010). Arousal and valence were rated from 1 (extremely calm/extremely unpleasant, respectively) to 9 (extremely aroused/extremely pleasant, respectively).

MAV	G	Arousal	Valence	Dur	MEB	Arousal	Valence	Dur
45_happiness	F	5.15	7.78	1.56	**C2_happiness**	5.92	5.79	1.75
46_happiness	F	5.53	7.95	1.01	**C3_happiness**	5.79	6.46	1.58
55_happiness	M	4.60	6.40	1.10	**C6_happiness**	5.92	6.67	1.64
58_happiness	F	5.23	7.94	1.05	**C7_happiness**	4.71	6.50	1.05
59_happiness	M	5.92	7.97	1.83	**C10_happiness**	5.92	7.13	1.63
6_happiness	M	5.25	7.78	1.74	**V3_happiness**	6.21	5.67	1.58
60_happiness	F	5.48	7.67	1.16	**V9_happiness**	6.21	5.88	1.08
61_happiness	M	5.43	7.68	2.6	**V10_happiness**	6.46	5.54	1.93
42_sadness	M	5.00	3.45	1.67	**C1_sadness**	5.08	3.75	3.03
45_sadness	F	5.10	2.32	1.78	**C2_sadness**	4.17	4.38	2.10
46_sadness	F	5.05	2.35	1.96	**C3_sadness**	4.38	5.79	2.61
53_sadness	F	4.70	2.26	2.88	**V1_sadness**	5.42	3.50	2.20

59_sadness	M	5.25	3.76	4.31	**V3_sadness**	6.04	4.21	2.94
6_sadness	M	5.00	2.21	1.64	**V4_sadness**	5.97	5.17	2.60
60_sadness	F	5.80	2.12	2.38	**V9_sadness**	5.21	5.54	2.24
61_sadness	M	4.48	2.19	2.44	**V10_sadness**	5.46	4.88	1.76
42_neutral	M	3.19	4.89	1.31	**C2_neutral**	4.04	4.83	1.59
46_neutral	F	3.30	4.95	0.24	**C3_neutral**	4.04	4.96	1.77
53_neutral	F	3.95	4.87	0.95	**C6_neutral**	3.58	5.08	1.74
55_neutral	M	3.29	4.81	1.24	**C7_neutral**	3.96	4.50	1.43
58_neutral	F	3.21	4.84	0.51	**C9_neutral**	4.63	3.96	1.70
6_neutral	M	3.15	4.90	0.90	**V1_neutral**	4.79	4.13	1.45
60_neutral	F	3.17	4.93	1.60	**V7_neutral**	4.96	4.44	0.83
61_neutral	M	3.26	4.90	1.86	**V9_neutral**	4.71	4.13	1.10

G = gender; Dur (s) = duration in seconds; In the names of MEB stimuli "C" corresponds to sounds produced by the clarinet and "V" corresponds to sounds produced by the violin.

Normative ratings of MEB stimuli did not exist for the Portuguese population until the moment. Consequently, we performed a validation of happy, sad, and neutral musical sounds to choose the sounds based on an evaluation of valence and arousal from 24 native speakers of Portuguese participants (12 female) who did not take part in the main EEG study (mean age = 27.25, *SD* = 5.50 years). For that evaluation, the presentation of the sounds was pseudo-randomized to avoid the consecutive presentation of more than two vocalizations from the same category (neutral, happy or sad). Each sound was presented only once and participants were instructed to choose a number from a scale from 1 to 9 that better represented its valence (ranging from 1—very *unpleasant* to 9—very pleasant) and arousal (ranging from 1—*very calm* to 9—*very aroused*).
Participants were also instructed to perform an emotion recognition task by choosing among "happiness", "sadness", "neutrality" or "other emotion" labels. Every time "other emotion" was selected a forced description of the emotion perceived was required. Participants were encouraged to register their responses as quickly as possible, following their first impressions.

The mean results of valence and arousal ratings of MEB sounds are presented in table 3. Analyses of variance with the factor Emotional Condition (3 levels – happiness, sadness, and neutral) were calculated for two dependent variables – valence and arousal

– followed by pairwise comparisons with Bonferroni correction for multiple comparisons, considering a *p* value of 0.05. Significant differences were found between emotional conditions regarding arousal ratings ($F(2, 21) = 14.98$, p = < .001). Follow-up comparisons revealed higher arousal ratings were attributed to happy musical sounds when compared to neutral musical sounds ($p < .001$) and between sad and neutral musical sounds ($p = .017$). No significant differences were found between happy and sad musical sounds ($p = .081$). Significant differences were also found for valence ratings between emotional conditions ($F(2, 21) = 18.04$, $p < .001$): happy musical sounds were considered more positive than neutral musical sounds ($p < .001$) and sad musical sounds ($p < .001$).

For the MAV sounds selected (please see Table 3), analyses of variance were also performed for valence and arousal ratings. Due to the normality assumption not being fulfilled for those dependent variables, non-parametric Kruskal-Wallis H Tests were performed, followed by pairwise comparisons with Bonferroni correction for multiple comparisons. Significant differences were found for arousal ratings of the MAV sounds (χ^2 (2) = 16.74, $p < .001$), but only between neutral and sad ($p = .015$) and between neutral and happy ($p < .001$) sounds. There were no differences in perceived arousal between sad and happy vocalizations ($p = .729$). Significant differences were found for valence ratings of the MAV sounds (χ^2 (2) = 20.5, $p < .001$) only between the happy and sad conditions ($p < .001$).

Table 3
Duration, valence, and arousal values of MAV and MEB stimuli per emotional condition | Mean (M) and Standard error of the mean (SE) of duration in seconds (s), valence ratings (1-9) and arousal ratings (1-9) of standard stimuli in the different emotional conditions for the two types of stimuli (human vocalizations and musical bursts).

	Duration (s)		Valence		Arousal	
	M	SE	M	SE	M	SE
Emotional Condition (MAV)						
Happiness	1.51	0.51	7.59	0.48	5.26	0.36
Sadness	2.38	0.83	2.58	0.59	5.01	0.34
Neutral	1.08	0.50	4.9	0.05	3.29	0.24
Emtotional Condition (MEB)						
Happiness	1.53	0.29	6.21	0.53	5.89	0.49
Sadness	2.44	0.41	4.65	0.78	5.21	0.61
Neutral	1.45	0.31	4.50	0.39	4.34	0.46

An additional analysis of variance with the factor Type of Stimuli (2 levels – MEB/MAV) was conducted for each emotional condition, for the same dependent variables. For the final set of standard stimuli (Table 3), mean arousal was significantly different between MAV and MEB sounds in the happiness condition ($F(1, 14) = 6.17$, $p = .026$), as mean valence ($U = 4.000$, $p = .002$). In the sadness condition, arousal ratings did not differ between MAV and MEB sounds ($F(1, 14) = .38$, $p = .548$) but differences were found concerning valence ratings ($F(1, 14) = 31.22$, $p < .001$). Lastly, in the neutral condition, differences were found for arousal ratings between MAV and MEB stimuli ($U = 1.000$, $p < .001$) but not for valence ratings ($U = 17.00$, $p = .130$).

EXPERIMENT DESIGN

The final stimulus set was composed of 68 different stimuli (24 human voice – l12 female/12 male; 24 musical bursts – 13 clarinet/11violin; 10 monkey voices; 10 tambourine sounds). In order to have enough trials to elicit a significant ERP effect (Thigpen et al., 2017), each MAV and MEB stimulus was repeated five times, and each monkey voice and tambourine sound was repeated twice, resulting in a total of 280 trials. Monkey sounds were the target stimuli in the experimental block including human vocalizations (MAV), while tambourine sounds were the target stimuli in the experimental block including musical bursts (MEB).

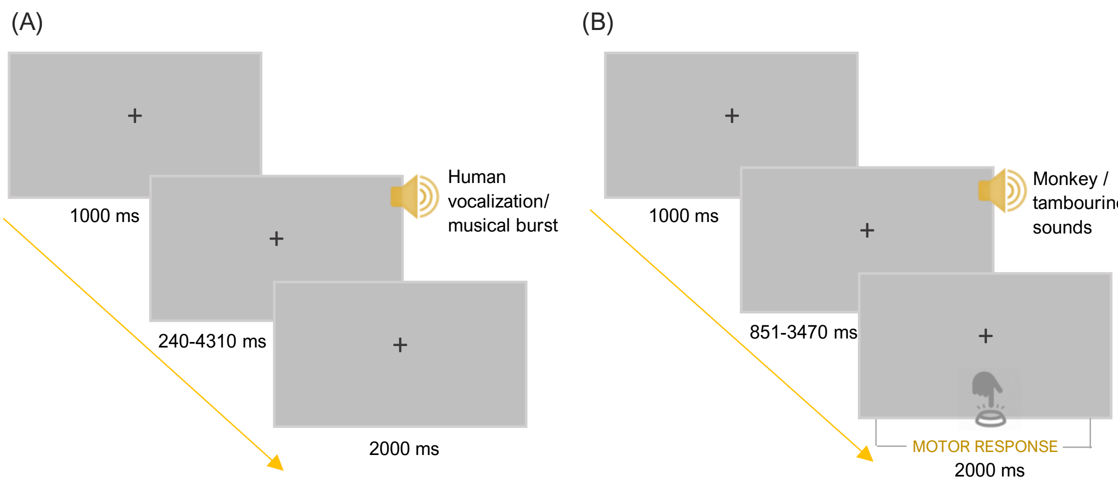
All auditory stimuli were pseudo-randomized and distributed over two blocks with 140 trials each with the following constraints: (a) the first block contained only MAV stimuli and monkey vocalizations (target stimuli); (b) the second block contained only MEB sounds and tambourine sounds (target stimuli); (c) no more than three consecutive trials of the same condition in order to avoid inducing an emotional state (Pinheiro et al., 2017; Nolden et al., 2017); (d) no consecutive target stimuli (monkey voices/tambourine sounds); (e) no more than three consecutive trials of the same type of stimuli (clarinet and violin) in the musical block.

PROCEDURE

This study was held in the facilities of Faculdade de Psicologia da Universidade de Lisboa. The EEG experiment took place in a Faraday Cage of the EEG Lab. Auditory stimuli were presented using Presentation Software (www.neurobs.com) and delivered through over-ear headphones. Participants were seated at a 100 cm distance from the computer screen and instructed to press a button using their right forefinger when they heard a monkey voice (human vocalizations condition) or when they heard a tambourine sound (musical bursts condition) (14% occurrence).

Sounds were presented while a fixation cross was shown centrally on the screen to minimize eye movements. Fig. 1 shows the design of an experimental trial. The presentation of the two experimental blocks were counterbalanced between participants. This implicit task procedure was used in a previous study concerning the neural correlates of nonverbal emotional vocalizations processing, also examining emotional versus neutral cues and with a focus on N100 and P200 ERP components (Liu et al., 2012). Rigoulot and colleagues (2015) also used an implicit task when analyzing emotional processing with musical sounds and vocalizations in an ERP paradigm. After 20 stimuli, participants had a 10 second silence break. The task had a total duration of approximately 30 min. ERPs were recorded while sounds were presented.

Figure 1
Illustration of an experiment trial for non-target (A) and target (B) stimuli.



At the end of the EEG task, participants had to complete a behavioral task concerning the evaluation of the affective properties (valence, arousal, and emotional category recognition) of the stimuli presented. The behavioral task was administered through a Qualtrics survey where each sound was presented only once followed with an emotion identification task by choosing among happiness, sadness, and neutrality labels. The valence and arousal ratings were obtained in scales from 1 to 9 ranging from 1—*very unpleasant* to 9—very pleasant (valence) and ranging from 1—*very calm'* to 9—*very* *aroused* (arousal). Vocalizations and musical bursts were presented separately in two blocks.

EEG DATA ACQUISITION AND PROCESSING

The EEG was collected using a 64-channel BioSemi system (Active 2) (BioSemi B.V., WG-Plein, Amsterdam, The Netherlands) at a digitization rate of 512 Hz. Blinks and eye movements were monitored through flat-type electrodes that were placed on the left and right temple and one below the left eye. The electrodes were arranged according to the International 10/20 System (Jasper, 1958). Electrode impedances were kept below ± 35KOhm.

Offline EEG analyses were processed using Letswave 6, an open-source Matlab toolbox (https://www.letswave.org/). EEG data was re-referenced to the mathematical average of the right and left mastoids. A bandpass 0.1-30 Hz filter was applied to event-related epochs, time-locked to sound onset. Epochs started 200 ms before each stimulus onset and ended 1000 ms after stimulus onset for each condition (happy vocalizations, sad vocalizations, neutral vocalizations, happy musical sounds, sad musical sounds and neutral musical sounds). A baseline correction was applied to epoched EEG data between -200 ms and 0 ms. Eye movements were corrected using the Gratton, Coles, and Donchin (1983) method. Then, the data were submitted to an artifact rejection procedure with amplitude criteria of ± 100 mV to exclude trials containing excessive artifacts caused by head movements and/or muscle activity. Only participants with at least 70% of trials passing the artifact rejection were included for individual ERP averaging (n = 58). Grand averages were processed for each group (non-musicians, instrumentalists, and singers) and condition (happy vocalizations, sad vocalizations, neutral vocalizations, happy musical sounds, sad musical sounds and neutral musical sounds).

The ERP waveforms were visually inspected, and three auditory components were identified for analysis: N1 (early negative peak between 120-190 ms), P2 (positive peak between 220-290 ms) and LPP (late component between 500-900 ms). These components were identified in other studies regarding auditory emotional processing (Liu et al., 2012; Pell et al., 2015; Pinheiro et al., 2015, 2017; Spreckelmeyer et al., 2009). Mean amplitude was calculated between 120-190 ms (N1), 220-290 ms (P200), 500-700 ms (LPP – early window), and 701-900 ms (LPP – late window) post-stimulus onset.

These windows were chosen based on ERP waveforms inspection and previous studies (Pinheiro et al., 2015, 2017).

STATISTICAL ANALYSIS

The IBM SPSS Statistics 22 software package (SPSS, Corp., USA) was used for statistical analyses. Analyses of variance were corrected for non-sphericity using the Greenhouse-Geisser method (Greenhouse & Geisser, 1959). Alpha level was set at .05.

For the EEG data, the mean amplitudes of N1, P2, and LPP were tested through a repeated measures of variance (ANOVA) with type of stimuli (vocalizations or musical bursts), emotional category (happiness, sadness, or neutral) and region-of-interest (ROI) (frontal, frontocentral, central, or centroparietal) as repeated-measures factor and musical expertise as between-subjects factors. Based on previous ERP studies investigating auditory emotional processing (Jessen & Kotz, 2011; Pinheiro et al., 2015) and on visual inspection of grand average waveforms, the regions of interest included the following electrodes: F3, Fz, F4 for the frontal region, FC3, FCz, FC4 for the frontocentral region, C3, Cz, C4 for the central region, and CP3, CPz and CP4 for the centroparietal region. Main effects and interactions were followed with planned comparisons with Bonferroni correction.

For the behavioral data, the ratings of valence, arousal, and emotional category recognition were submitted to a repeated measures of variance (ANOVA) with type of stimuli (vocalizations or musical bursts) and emotional category (happiness, sadness, or neutral) as repeated-measures factor and musical expertise as between-subjects factor. Main effects and interactions were followed with planned comparisons with Bonferroni correction.
The emotional category recognition responses were transformed in the unbiased hit rate or Hu, a measure that results from joining the probability that a given emotional category is correctly recognized when is presented (hit score) with the probability that when a response is used, it is correct. It was calculated based on the formula provided by Wagner (Wagner, 1993) for each participant.

Age was the only variable that was different between the three groups. Thereby, age was added as a covariate in both statistical models.

RESULTS

1. BEHAVIORAL DATA

1.1 Emotion Recognition

Recognition rates were analyzed using Hu scores (Wagner, 1993) for each participant for emotional condition (happiness, sadness, neutral) and type of stimulus (vocalizations and musical bursts). The scores go from 0 to 1, with 0 representing no correct responses, and 1 representing a perfect performance (Table 4).

Table 4

Mean Hu scores and standard deviations for vocalizations and musical sounds in each emotional condition (horizontal) as a function of musical expertise.

Group (*n*)	**Vocalizations (MAV)**			**Musical Bursts (MEB)**		
	Happy	Sad	Neutral	Happy	Sad	Neutral
Non-musicians (23)	0.96 (0.07)	0.93 (0.10)	0.98 (0.05)	0.81 (0.10)	0.68 (0.14)	0.62 (0.28)
Instrumentalists (22)	0.93 (0.08)	0.93 (0.07)	0.97 (0.06)	0.76 (0.19)	0.75 (0.19)	0.79 (0.26)
Singers (21)	0.91 (0.08)	0.92 (0.10)	0.97 (0.04)	0.73 (0.18)	0.68 (0.24)	0.72 (0.25)

A main effect of type of stimulus was identified [$F(1, 62) = 9.51$, $p = .003$, $\eta_p^2 = .133$] indicating that vocalizations were more accurately recognized than musical sounds. A significant emotional category by group interaction [$F(4, 124) = 3.67$, $p = .007$, $\eta_p^2 = .106$] showed differences in the control group that depended on emotional category: non-musicians demonstrated better performance at recognizing happy sounds in comparison to sad and neutral sounds ($p < .016$).

The main effect of group was not significant [$F(2, 62) = .576$, $p = .565$, $\eta_p^2 = .018$]. Nevertheless, a significant group by emotional category by type of stimulus interaction [$F(2, 124) = 3.36$, $\eta_p^2 = .098$] revealed that non-musicians were better at recognizing neutral vocal sounds compared to happy vocal sounds ($p = .009$), and better at recognizing happy musical sounds in comparison to neutral ($p < .001$) and sad ($p = .001$) musical sounds; singers performed better at recognizing neutral vocal sounds in comparison to sad ($p = .045$) and happy vocal sounds ($p = .010$), but no differences were found between emotional categories in the musical condition ($p > .050$) (Table 4). The performance of instrumentalists did not differ as a function of emotional category or type of stimulus ($p > .050$).The covariate did not affect the emotional recognition findings [$F(1, 62) = .002$, $p = .966$, η_p^2 .< .001].

1.2. Valence Ratings

Table 5
Mean valence ratings and standard deviations for vocalizations and musical sounds in each emotional condition (horizontal) in function of musical expertise.

Group (*n*)	**Vocalizations (MAV)**			**Musical Bursts (MEB)**		
	Happy	Sad	Neutral	Happy	Sad	Neutral
Non-musicians (23)	4.64 (0.43)	2.84 (0.80)	4.92 (0.23)	6.33 (0.72)	3.60 (0.47)	4.64 (0.43)
Instrumentalists (22)	6.78 (1.19)	2.89 (0.69)	4.79 (0.76)	6.71 (0.91)	3.91 (0.61)	4.85 (0.54)
Singers (21)	6.94 (0.86)	2.87 (0.79)	5.01 (0.47)	6.94 (0.84)	4.40 (1.17)	5.02 (0.46)

Singers rated sounds more positively in comparison to controls ($p = .019$), as shown by a main-effect of group [$F(2, 62) = 4.018$, $p = .023$, $\eta_p^2 = .115$]. A significant type of stimulus by group interaction was found [$F(2, 62) = 5.743$, $p = .005$, $\eta_p^2 = .156$]: singers attributed higher ratings of valence to musical sounds in comparison to non-musicians ($p = .001$), but not in comparison to instrumentalists ($p = .211$). Instrumentalists and controls did not differ either ($p = .109$).

A main effect of emotional category [$F(1.39, 85.84) = 35.270$, p = < .001, η_p^2 = .363] was observed, demonstrating happiness as the emotional category receiving more positive valence ratings than neutral (p < .001) and sadness (p < .001).

The covariate did not affect the emotional recognition findings [$F(1, 62) = .025$, p = .874, η_p^2 .< .001].

1.3. Arousal Ratings

Table 6
Mean arousal ratings and standard deviations for vocalizations and musical sounds in each emotional condition (horizontal) in function of musical expertise.

Group (*n*)	Vocalizations (MAV)			Musical Bursts (MEB)		
	Happy	Sad	Neutral	Happy	Sad	Neutral
Non-musicians (23)	6.18 (0.66)	5.68 (1.26)	4.60 (1.09)	6.25 (0.88)	5.50 (1.07)	4.87 (1.02)
Instrumentalists (22)	5.65 (1.65)	5.67 (1.46)	3.90 (1.62)	6.17 (1.33)	5.22 (1.21)	4.38 (1.31)
Singers (21)	5.74 (1.21)	6 (1.13)	4.03 (1.34)	6.54 (1.19)	5.62 (1.15)	4.66 (1.04)

A main effect of emotional category [$F(2, 124) = 15.842$, p < .001, η_p^2 = .664] was found for the arousal ratings showing that happy sounds were considered more arousing in comparison to sad (p < .001) and neutral sounds (p < .001).

There were no main effect of group [$F(2, 62) = .952$, p = .392, η_p^2 = .030] or interactions involving the group factor (group x type of stimulus: p = .530, group x emotional category: p = .200, group x type of stimulus x emotional category: p = .200). The covariate did not affect the emotional recognition findings [$F(1, 62) = .634$, p = .429, η_p^2 .= .010].

2. ERP RESULTS

Figure 2.

Grand average waveforms at Cz in non-musician, instrumentalists and singers for neutral (black line), happy (red line), and sad (green line) sounds in vocal and musical conditions.

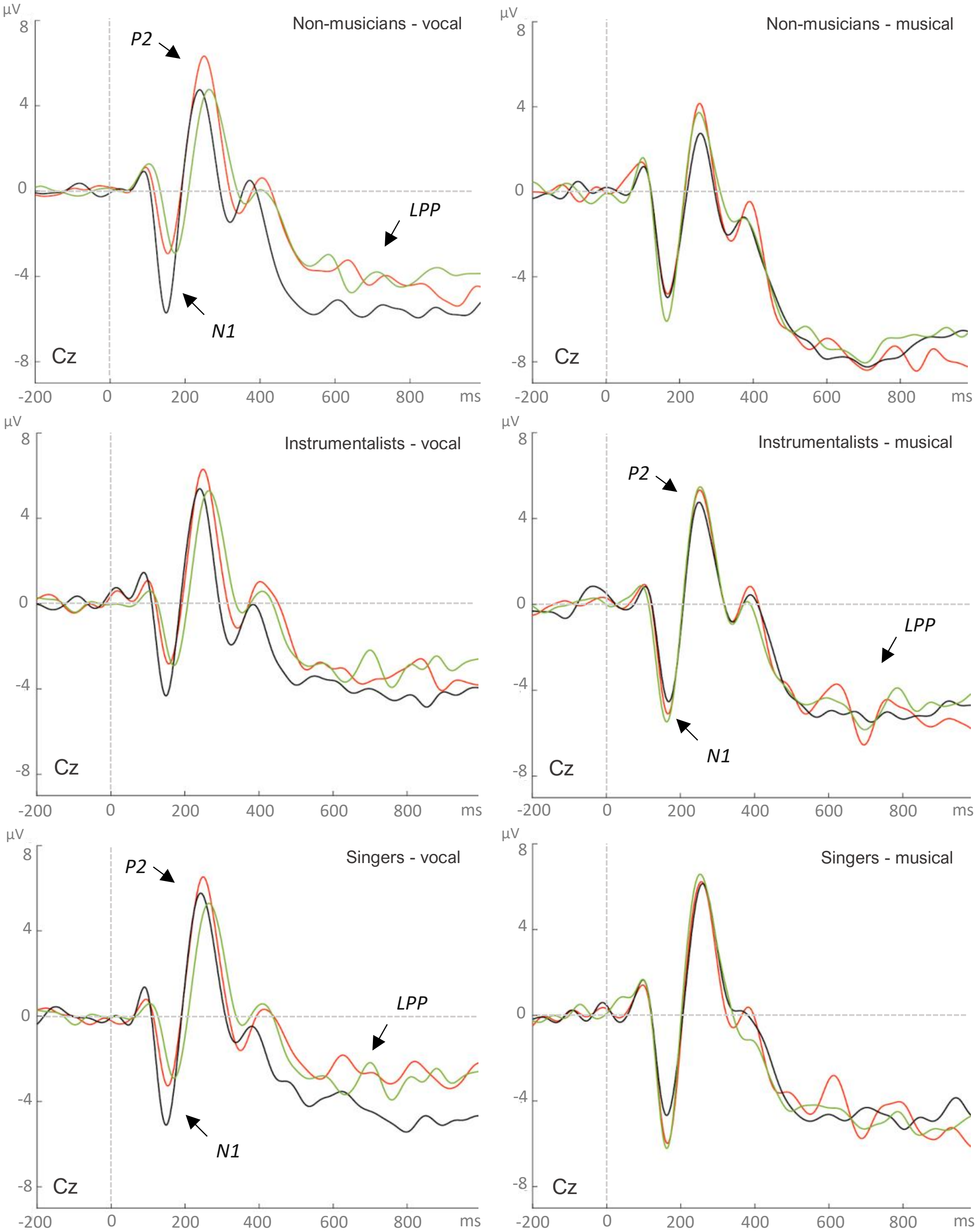
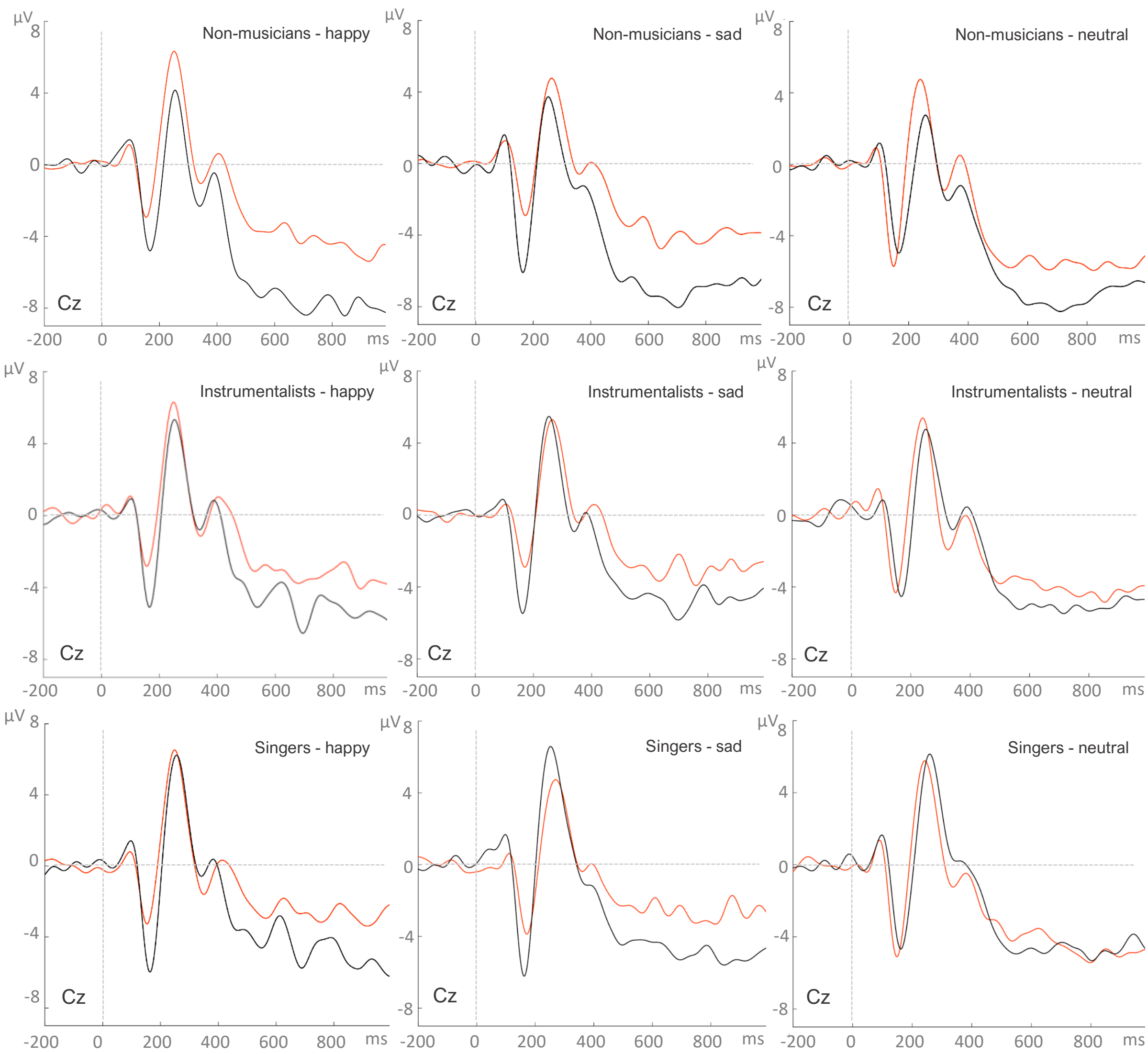


Figure 3.

Grand average waveforms at Cz in non-musician, instrumentalists and singers for musical (black line) and vocal (red line) sounds in happy, sad, and neutral conditions.



2.1. N1 Results

There was no main-effect of group [$F(2, 54) = .291$, $p = .749$, $\eta_p^2 = .011$]. No other main-effects or significant interactions were observed. A marginally significant interaction between type of stimulus and emotional category was observed [$F(2, 108) = 2.77$, $p = .067$, $\eta_p^2 = .049$], showing a tendency for musical sounds to elicit a more negative N1 in comparison to vocal sounds for both happy ($p < .001$) and sad ($p < .001$) emotional conditions (fig. 3). Also, this interaction demonstrated that for vocalizations, neutral sounds tended to elicit a more negative N1 in comparison to happy and sad sounds ($p <$

.001) while for musical bursts, sad sounds more often elicited more negative N1 amplitude only in comparison to neutral sounds (p = .003) (fig. 2). The covariate did not affect the N1 findings [$F(1, 54) = 1.27$, p = .264, η_p^2 = .023].

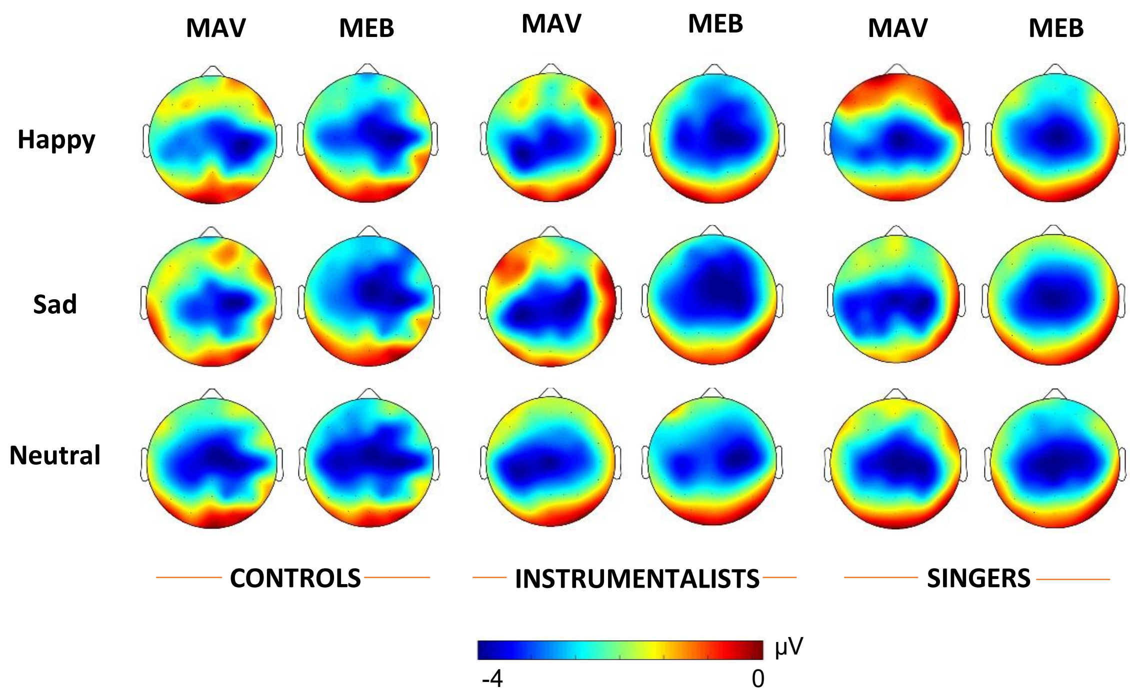
Table 7

Mean N1 amplitudes and standard deviations for vocalizations and musical sounds in each emotional condition (horizontal) as a function of musical expertise.

Group (*n*)	**Vocalizations (MAV)**			**Musical Bursts (MEB)**		
	Happy	Sad	Neutral	Happy	Sad	Neutral
Non-musicians (22)	-1.52 (0.49)	-1.23 (0.42)	-3.23 (0.49)	-2.76 (0.60)	-3.65 (0.505)	-2.94 (0.44)
Instrumentalists (17)	-1.17 (0.56)	-1.41 (0.49)	-1.89 (0.56)	-2.81 (0.69)	-3.57 (0.58)	-2.71 (0.50)
Singers (19)	-1.65 (0.56)	-1.92 (0.48)	-3.29 (0.56)	-3.44 (0.69)	-3.38 (0.58)	-2.88 (0.50)

Figure 4.

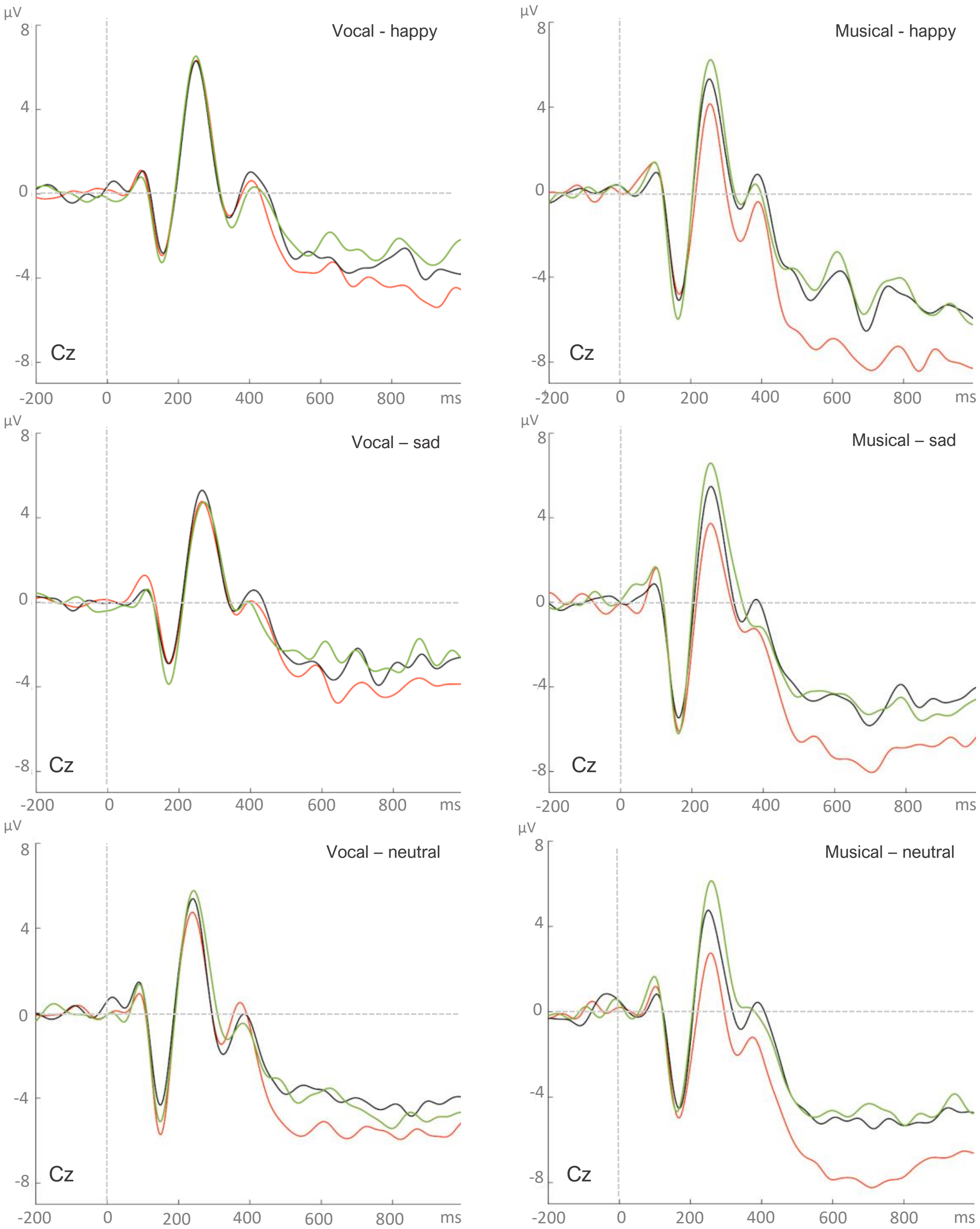
Topographic maps showing the scalp distribution of N100 voltage for happy, neutral and sad stimuli in MAV and MEB, in the three groups.



2.2. P2

Figure 5.

Grand average waveforms at Cz in response to musical and vocal sounds for controls (red), instrumentalists (black), and singers (green) in the three emotional categories.



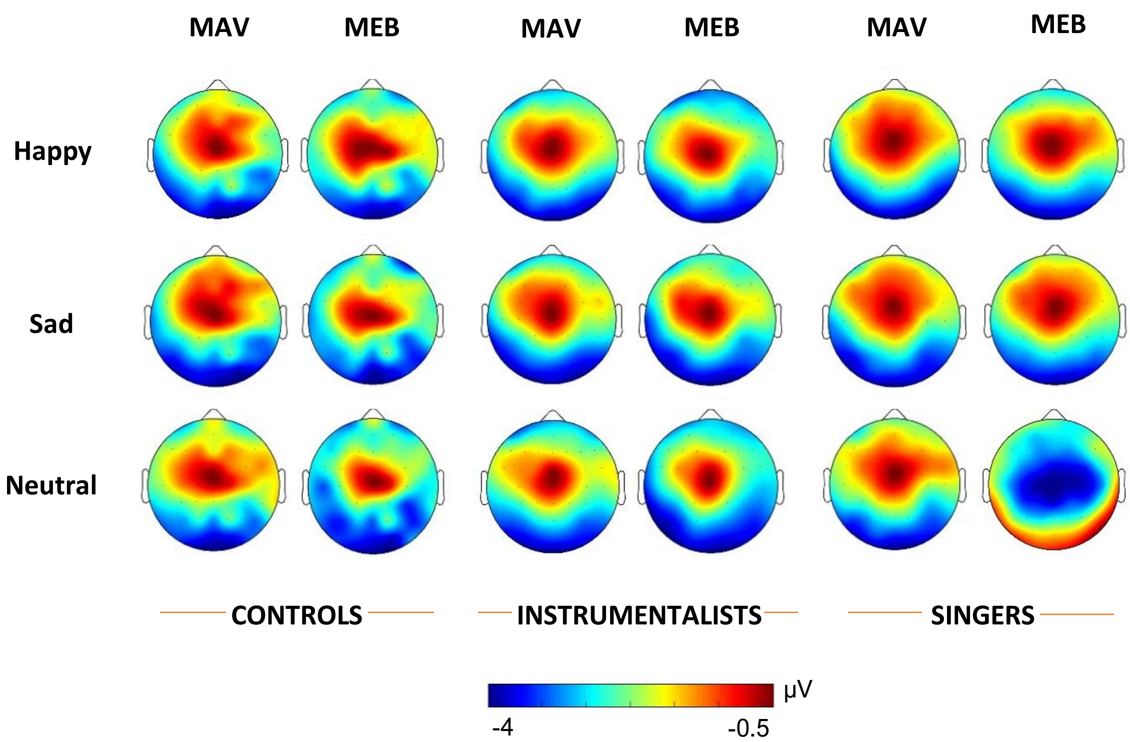
The analysis did not yield a main-effect group [F(2, 54) = .844, *p* = .436, η_p^2 = .030].

A significant interaction between group and type of stimulus was observed [F(2, 54) = 5.202, *p* = .009, η_p^2 = .162]: controls were the only group exhibiting different P2 amplitudes in response to vocalizations and musical sounds (*p* = .001) – the P2 was enhanced (i.e., more positive) for vocalizations (fig.3). In instrumentalists (*p* = .330) and singers (*p* = .220) no differences in P2 amplitude were found for the two types of stimuli (fig.3). Nonetheless, there was a tendency for an enhanced P2 amplitude in response to musical sounds in singers when compared to controls (*p* = .075) (fig.5).

The analysis also revealed main effects of emotional category [*F(*2, 108) = 4.89, *p* = .009, η_p^2 = .083] and of region of interest [*F(*1.48, 79.84) = 4.43, *p* = .024, η_p^2 = .076]. Pairwise comparisons showed that happy stimuli elicited a more positive P2 than neutral stimuli (*p* = .004), and that P2 was largest over frontocentral electrode sites compared to frontal (*p* < .001), central (*p* < .001) and centroparietal electrodes (*p* < .001). Also, both frontal and central sites registered enhanced P2 amplitudes when compared to centroparietal sites (*p* < .001).

Figure 6.

Topographic maps showing the scalp distribution of P2 voltage for happy, neutral and sad stimuli in MAV and MEB, in the three groups.



A significant type of stimulus by emotional category interaction was also found [*F(*2, 108) = 3.84, *p* = .025, η_p^2 = .066]: vocalizations elicited a more positive P2 than musical sounds in the happiness condition (*p* = .003) (fig. 5); in the case of vocalizations, happy vocal sounds were associated with an enhanced P2 in comparison to neutral (*p* = .001) and sad sounds (*p* = .001), in the case of musical sounds, the sadness condition elicited a more positive amplitude when compared to the neutral one (*p* = .008) (fig. 2).

The covariate did not affect the P2 findings [*F(*1, 54) = 3.139, *p* = .082, η_p^2 = .055].

Table 8
Mean P2 amplitudes and standard deviations for vocalizations and musical sounds in each emotional condition (horizontal) as a function of musical expertise.

Group (*n*)	**Vocalizations (MAV)**			**Musical Bursts (MEB)**		
	Happy	Sad	Neutral	Happy	Sad	Neutral
Non-musicians (22)	3.85 (0.58)	2.72 (0.56)	2.52 (0.65)	2.15 (0.64)	1.69 (0.59)	0.89 (0.60)
Instrumentalists (17)	3.78 (0.66)	2.83 (0.64)	2.92 (0.75)	2.79 (0.73)	3.09 (0.67)	2.3 (0.69)
Singers (19)	3.87 (0.66)	2.63 (0.64)	2.42 (0.75)	3.21 (0.73)	4.08 (0.67)	3.34 (0.69)

2.3. LPP (earlier window)

The effect of group did not reach statistical significance [*F(*2, 54) = .842, *p* = .436, η_p^2 = .030]. However, there was a significant main effect of stimulus type [*F(*1, 54) = 6.420, *p* = .014, η_p^2 = .106], emotional category [*F(*1.78, 96.27) = 4.15, *p* = .023, η_p^2 = .071] and region of interest [*F(*1.48, 80) = 33.40, *p* < .001, η_p^2 = .382].

The LPP was increased in response to vocalizations compared to musical sounds (*p* < .001) (fig. 5). Further, an increased LPP was observed in response to happy (*p* = .005) and sad (*p* = .023) sounds compared to neutral sounds (fig. 2). Further, the LPP was increased over centroparietal electrode sites when compared to frontal (*p* < .001), frontocentral, (*p* < .001) and central (*p* < .001), over central compared to frontal (*p* < .001) and frontocentral electrode sites (*p* < .001), and over frontocentral compared to frontal ones (*p* = .012) (fig. 7). There was a significant effect of Age on LPP amplitudes elicited [*F(*1, 54) = 13.54, p = .001, η_p^2 = .200].

Table 9
Mean LPP amplitudes (500-700ms) and standard deviations for vocalizations and musical sounds in each emotional condition (horizontal) as a function of musical expertise.

Group (*n*)	Vocalizations (MAV)			Musical Bursts (MEB)		
	Happy	Sad	Neutral	Happy	Sad	Neutral
Non-musicians (22)	-3.1 (0.58)	-3.45 (0.56)	-4.47 (0.62)	-6.25 (0.7)	-6.24 (0.63)	-6.58 (0.64)
Instrumentalists (17)	-3.02 (0.67)	-3.23 (0.65)	-3.41 (0.71)	-5.01 (0.81)	-4.65 (0.73)	-5.06 (0.74)
Singers (19)	-3.36 (0.67)	-3.42 (0.65)	-5.56 (0.71)	-5.36 (0.81)	-5.44 (0.73)	-5.91 (0.74)

2.4. LPP (later window)

We did not observe a main-effect of group [*F(*2, 54) = 1.163, p = .320, η_p^2 = .041]. A main effect of region of interest was observed [*F(*1.64, 88.51) = 49.66, *p* < .001, η_p^2 = .479]: the LPP was increased over centroparietal electrode sites compared to frontal (*p* < .001), frontocentral (*p* < .001), and central (*p* < .001), and over central sites in comparison to frontal (*p* < .001) and frontocentral (*p* < .001) (fig. 7).

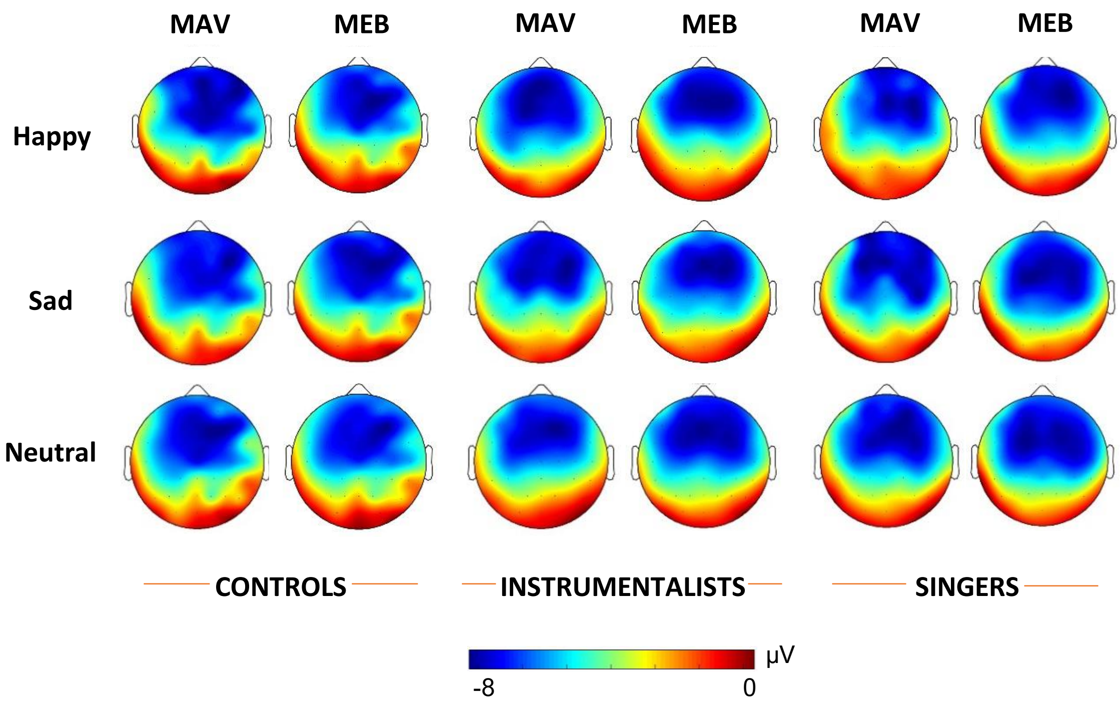
A significant type of stimulus by emotional category interaction was found [*F(*2, 108) = 3.944, *p* = .022, η_p^2 = .068]. Pairwise comparisons showed that vocalizations elicited a more positive LPP amplitude than musical sounds for emotional (*p* < .001) and neutral sounds (*p* = .013) (fig. 3). Further, in the case of vocalizations, happy and sad sounds were associated with an increased LPP compared to neutral sounds (*p* < .001). No differences between emotional categories were found for musical bursts (*p* > .050) (fig.2).

Table 10.
Mean LPP amplitudes (701-900ms) and standard deviations for vocalizations and musical sounds in each emotional condition (horizontal) as a function of musical expertise.

Group (*n*)	Vocalizations (MAV)			Musical Bursts (MEB)		
	Happy	Sad	Neutral	Happy	Sad	Neutral
Non-musicians (22)	-3.61 (0.62)	-3.59 (0.59)	-4.59 (0.58)	-6.58 (0.73)	-5.85 (0.64)	-6.58 (0.65)
Instrumentalists (17)	-3.01 (0.71)	-3.19 (0.67)	-3.76 (0.66)	-5.39 (0.84)	-4.42 (0.73)	-5.17 (0.74)
Singers (19)	-3.77 (0.71)	-3.67 (0.67)	-6.34 (0.66)	-6.17 (0.84)	-5.99 (0.73)	-5.81 (0.74)

Figure 7.

Topographic maps showing the scalp distribution of LPP voltage for happy, neutral and sad stimuli in MAV and MEB, in the three groups.



DISCUSSION

Does musical training boost emotional processing skills? This question was investigated by examining if different types of musical training (instrumental, vocal or none) influenced the processing of emotional cues conveyed by human nonverbal vocalizations and musical bursts, manipulating the emotionality of the sounds (neutral, positive or negative). For that purpose, we examined the event-related potentials elicited at the time of the presentation of those sounds during an implicit task, as well as the behavioral responses to three affective properties of the stimuli: valence, arousal and emotional categorization.

Effects of musical training on auditory emotional processing

We had hypothesized a main group effect expecting musical expertise to influence N1, P2 and LPP amplitudes: in the case of nonverbal vocalizations, we expected to observe reduced N1 and enhanced P2 and LPP amplitudes of the components for singers when compared to controls; in the case of musical bursts, we expected those differences between instrumentalists and controls. Although we did not find any main-group effects, the ERP results show evidence of musical training modulating the early processing of affective musical and vocal sounds, manifested on the P2 component. For the control group, an enhanced P2 amplitude was observed for vocalizations in comparison to musical bursts. The P2 has been associated with emotional salience detection (Liu et al., 2012; Paulmann & Kotz, 2008; A. P. Pinheiro et al., 2013), and vocalizations have been observed to have priority over other auditory stimuli, such as speech, on the emotional decoding processes at the brain level (Hawk et al., 2009; Pell et al., 2015). Hence, the absence of P2 differences in singers and instrumentalists as a functions of stimulus type (neither singers nor instrumentalists exhibited differences regarding that component), might indicate that for individuals with musical expertise, musical sounds are as salient as vocalizations. These results might be interpreted in the light of functional neuroplasticity: brain regions devoted to musical training can increase their activation for musical sounds processing. This type of neuroplasticity has been designated as “map expansion”, relying on the flexibility of local brain regions engaged in performing one type of function (Grafman, 2000), such as playing an instrument or singing.

Similar findings were observed when examining behavioral measures: controls were better at recognizing happiness among the musical sounds, while singers and instrumentalists performed similarly in emotion recognition in the musical condition, indicating an influence of musical expertise on the decoding of emotional meaning from musical sounds. In previous studies, musical expertise was linked to an enhanced emotion recognition in music (Castro & Lima, 2014). In the group of instrumentalists, we did not observe differences in emotion recognition between emotional categories of vocalizations. Even though no group differences were found, the fact that instrumentalists were the only group showing a similar performance at categorizing positive, negative and neutral valence in vocalizations aligns with the cross-domain effect of musical training on auditory emotion recognition abilities demonstrated in previous studies (Lima & Castro, 2011; Pinheiro et al., 2015; Thompson et al., 2004). Fuller et al., (2014) showed that even under degraded pitch conditions, musicians are more accurate than non-musicians at identifying emotions conveyed by the human voice. A relation between musical practice and enhanced perception of emotions through the voice was demonstrated (Strait et al., 2009) when recording the brainstem responses to human vocal sounds. Previously, the same relation had been observed not only for voice, but also for musical sounds (Musacchia et al., 2007).

Additionally, the P2 amplitude tended to be more positive in the case of singers in comparison to the control group in response to musical sounds, constituting another evidence of the influence of musical training on auditory processing. For the interpretation of this result, it is important to consider that some singers in the current study also played at least one instrument. This result might indicate that musical sounds are more salient for people with musical training, specifically with combined musical training (voice and instrument). The combination of vocal and instrumental musical training might be the reason why the differences between groups almost reached statistical significance only between singers and controls, but not between instrumentalists and controls for the musical condition. A previous study analyzing the mismatch negativity (MMN) to pitch deviances between singers and instrumentalists reported earlier responses to pitch deviances on musicians with both vocal and instrumental training (Nikjeh et al., 2008). Behaviorally, singers also rated the musical sounds as more positive than controls, indicating that valence modulates P2 amplitude as proposed in previous studies (Paulmann & Kotz, 2006; Schirmer et al., 2005).

Emotionality is processed differently in music and in the human voice

ERP effects in the three groups were observed. Emotional (vs. neutral) vocalizations elicited reduced N1 and enhanced late LPP amplitudes, whereas a valence-specific effect was found on the P2: its amplitude was enhanced in response to happy versus neutral and sad vocalizations. The observed differences are consistent with previous ERP studies with human nonverbal vocalizations that reported reduced N1 amplitudes for emotional in comparison to neutral sounds (Jessen & Kotz, 2011; Liu et al., 2012; Paulmann & Kotz, 2008; Pell et al., 2015; Pinheiro et al., 2013; Sauter et al., 2010), and enhanced P2 amplitude for happy and angry compared to sad sounds (Pinheiro et al., 2015).

All ERP components analyzed revealed differences as a function of the type of stimulus. In earlier processing stages (within 200 ms post-stimulus onset: N1 and P2), those differences were observed as an interaction with emotional category: the N1 amplitude was enhanced for musical sounds when compared to vocalizations for the emotional conditions (happiness and sadness), whilst the P2 amplitude was increased for voices in comparison to musical sounds for the positive condition (happiness). In later processing stages (after 500 ms post-stimulus onset: LPP), the type of stimulus generated a main effect, regardless of emotional category, revealing a more positive LPP for vocalizations in comparison to musical sounds. A prior study compared nonverbal human voices and violin excerpts and also reported LPP amplitude modulations – more positive for vocal than musical sounds (Proverbio et al., 2019). Another study directly compared the processing of emotions in musical and nonverbal vocalizations and reported the same result for the N1 component as observed in the current study (Rigoulot et al., 2015). Nevertheless, in their study the P2 component was modulated by stimulus type, but the P2 was still more positive for musical sounds than for vocal sounds. This may be due to the fact that they had another emotion being presented (fear), which may have affected the overall emotional salience detection from the stimuli. Moreover, the researchers of both studies did not report differences in how emotions were decoded form musical vs.

vocal sounds. In the current study, the significant interactions between type of stimulus and emotional category show, for the first time, ERP's differences between musical and vocal matched sounds. These results can provide evidence for a different treatment of sounds depending on their category, triggering distinct brain regions to respond to specific types of affective sounds, even though there is a common network for auditory emotional processing (Frühholz et al., 2016; Peretz et al., 2015). When comparing brain responses to music and voice (and their relation) with fMRI, a group of "music-preferring" neurons within the right superior temporal gyrus was identified when subtracting the response to music-preceded-by-music from the response to music-preceded-by-voice (Armony et al., 2015). When examining the N1 and P2 time-windows, the emotional content of vocalizations and musical sounds seem to be processed differently: while between vocal stimuli the N1 was decreased for emotional versus neutral vocalizations, for musical stimuli the N1 was decreased only for sad versus neutral bursts; and whereas the P2 was increased for happy vocalizations when compared to the other vocalizations, for musical bursts, the P2 was increased for sad sounds. When assessing the later time-window of the LPP component, vocal and musical sounds also elicited different ERP patterns: only vocalizations elicited different LPP amplitudes as a function of emotional category, with emotional content associated with higher amplitude than neutral content. These findings suggest different processes underlying the decoding of emotion from musical sounds and from human voices.

Musical sounds and nonverbal human vocalizations are categorized as different types of sounds. When Frühholz, Trost and Kotz (2016) proposed a typology of affective sounds, the affective nonverbal vocalizations were distinguished as "symptoms" because their origin and significance coincide – inner emotional state – and emotional musical sounds as "symbols" once they meet a conventional representation of emotion with its aesthetic expression. The observed enhanced P2 and LPP amplitudes for vocalizations versus musical sounds adds evidence for the "attentional bias" towards affective nonverbal vocalizations over other type of auditory stimuli (Jacob et al., 2014), in this case, over musical sounds. Also, it might align with the notion that as a symptom, emotional nonverbal human vocalizations consistently activate the amygdala during sound processing mechanism to generate emotional responses (Schirmer & Kotz, 2006). Amygdala activation for musical sounds is not consistently reported (Aubé et al., 2015) and if in fact emotional detection is processed by distinct brain pathways, it might be reflected in different modulations of the ERP components (Eimer & Holmes, 2007). Meanwhile, musical sounds more often activate the auditory cortex and particularly the superior temporal cortex given its importance in affective decoding over time (Frühholz et al., 2016) which can be correlated with the enhanced N1 amplitude observed for musical sounds since it has been proposed that the auditory cortex (primary and secondary) is the main neural generator of the N1 (Joos et al., 2014).

Music activates the ventral striatum of the basal ganglia, especially the nucleus accumbens, which is associated with music being perceived as a pleasurable and rewarding listening activity. In particularly, sad music is commonly perceived as pleasant, which has been explained with the "being moved" phenomenon: sad musical sounds are enjoyable when the listener feels that he/she is being moved (Vuoskoski & Eerola, 2017). Usually, that feeling is associated with empathic traits such as "compassion" and "sympathy" (Menninghaus et al., 2015). Some participants reported associating the musical bursts with movies and cartoon sounds which might indicate that the feeling of "being moved" might have been present at some level. That could explain an enhanced direction of attention for sad musical sounds reflected on the P2 component that is modulated by attention. Consequently, an enhanced P2 amplitude for sad over happy and neutral musical sounds emerges. Furthermore, the results from the behavioral emotion recognition task where vocalizations were more accurately identified than musical sounds may also be explained in the light of the "being moved" effect: sad musical sounds were more passive of being considered positive even if expressing sad meaning.

The LPP was the only ERP component whose amplitude was significantly related to age in the current study, consistent with evidence that the LPP can indicate age-related changes in emotional processes (Hajcak et al., 2010). Moreover, this component has been proposed to reflect facilitated processing of emotional stimuli compared to neutral ones. Consistently, the current results showed that LPP amplitudes were significantly different for emotional (happy and sad) compared to neutral stimuli, at earlier and later stages. Interestingly, in the later time-window, the difference was only observed for vocalizations, which again might indicate that the cognitive mechanisms used to process auditory emotion work differently (are different) according to stimulus type. These differences may be explained by the idea that emotion conveyed through music results from a self-memory or experience that is evoked in the moment of listening and generates an associated emotion, creating a very individual self-oriented experience (Grimshaw-Aagaard et al., 2019), while emotional vocalizations have an immediate functional role on adaptative behavior and survival towards others (Belin, 2006), involving more attentional and cognitive resources for its processing. Importantly, the musical bursts (MEB) were conceived as a result of an imitation of the non-verbal human vocalizations (MAV) for the purpose of being a musical counterpart of the MAV. Additionally, the selection of the sounds from the batteries of stimuli took into consideration pitch, arousal and valence properties of the sounds in order to be matched (see methods). Therefore, the differences described above regarding the emotional processing from voices and musical excerpts cannot be accounted by intrinsic features of the sounds.

The task used in this study had an implicit nature (i.e., participants allocated their attention to other target sounds than the vocalizations and the musical excerpts) indicating that the differences found between emotional and neutral cues may reflect an automatic decoding of emotion from auditory stimuli, regardless of task relevance. Previous ERP studies with implicit tasks for affective auditory emotional processing reported similar findings with voices only (Liu et al., 2012), as well as with voices and music (Proverbio et al., 2019). These results agree with the view that the attentional mechanisms receives input from several neural systems to organize the variety of sensory information (Driver, 2001). Some authors posit that a possible explanation may be the projections from amygdala to sensory cortices that enhance the processing of emotional information over non-emotional information (Vuilleumier, 2005). However, this hypothesis still needs to be tested in future studies with fMRI.

Limitations and future directions

The current study has some limitations such as the characteristics of the sample. The mean age of the participants differed between groups. Additionally, some musicians reported having depression (n = 5), anxiety (n = 2), and bipolarity (n = 1) disorders and were on medication, which may have influenced the results reported – previous studies stated neurophysiological abnormalities in the auditory processing, such as reduced P2 amplitude during the presentation of frequent tones in patients with bipolar disorder in comparison to healthy controls (Fridberg et al., 2009), and more positive P2 in response to negative relative to self-referent adjectives in depressed subjects compared to healthy subjects (Shestyuk & Deldin, 2010). The vocalizations from the MAV battery were produced by actors and even though they have different durations that account for natural properties of each emotion expression (Belin et al., 2008), one could question their ecological validity. It would be important to conduct the experiment with affective vocalizations from real social life (Anikin & Lima, 2018). Moreover, happiness and sadness in the form of laughter and crying were the only two emotional conditions used in the current study, which we can consider a poor variety of emotional stimuli when thinking about the diversity of emotions that are communicated daily. Because of that, the current study provides limited understanding of how musical training influences the auditory emotional processing overall and it should be extended to a wider variety of emotions.

One of the aims of the study was to compare two types of musical training – voice and instrument – that was compromised because seven of the eighteen singers who participated also had some level of instrumental training. The truth is that it is difficult to recruit singers who exclusively practiced singing without developing the training of some instrument. On the one hand, future research on the topic should try to include singers who only had vocal musical training to compare them with instrumental musicians and people without musical training to investigate if vocal musical training brings neural advantage on auditory processing skills per se. It could be interesting to also compare singers with other professionals who need to have a fine control of the vocal production, such as radio announcers. On the other hand, research should also aim to investigate the neurophysiological advantages of combined musical training (voice and instrument).

Considering the low spatial resolution of the EEG methodology, future studies should use both ERP and fMRI techniques to complement time resolution with spatial information underlying the affective auditory processing in musicians and non-musicians. Moreover, differences in the brain organization of musicians compared to non-musicians include multisensory systems (Peretz & Zatorre, 2005), and electrophysiological differences for the processing of speech and music have been found in both auditory and audiovisual modality (Musacchia et al., 2007). Therefore, it would be interesting to extend the current study to a cross-modal approach, investigating the emotional processing of audiovisual cues in singers, instrumentalists and non-musicians.

CONCLUSION

Music and vocalizations are two different types of sounds that engage different neural pathways of emotion processing in the brain. Our findings contributed to the understanding of how differently these two types of auditory cues are processed from stimulus onset to offset. A differential ERP response to musical relative to vocal sounds was found at early (reduced N1 and enhanced P2 amplitude in response to vocalizations compared to musical bursts) and later stages (enhanced LPP amplitude in response to vocalizations compared to musical bursts) of emotional processing. Repetitive musical training over time develops structural and functional adaptations on the brain, especially on the auditory system. The musical training effect exceeds musical skills and transfers to other domains, including the important mechanisms underlying auditory emotional processing. The present study found evidence of musical training effects at early stages of emotional processing around 200ms: both musical and vocal sounds elicited similar P2 amplitudes in singers and instrumentalists, whereas for non-musicians vocalizations elicited a more positive P2. Furthermore, a tendency for musical sounds to elicit an enhanced P2 amplitude in singers when compared to controls was observed. Thereby, according to the multi-stage model of affective sound processing, our results indicate that musical training seems to influence emotional salience detection of sounds in two ways: (1) musical sounds may be as salient as vocal sounds for musically trained individuals; (2) musicians with vocal training might have facilitated emotional detection of musical sounds when compared to musically untrained individuals. Our findings suggest that vocal training carries its own functional advantages in emotional meaning decoding from acoustic information, possibly when combined with instrumental training.

APPENDICES

Appendix A

Summary of studies reviewed examining the influence of musical training on auditory information processing in three domains: sound, speech and emotion.

INFLUENCE OF MUSICAL TRAINING ON AUDITORY INFORMATION PROCESSING			
Source	No. of Participants	Methodology	Results: Musicians vs. Non-Musicians
Sound Processing			
Pantev et al. (2001)	17 M	MEG[1]	Enhanced auditory cortical representations for tones of violin and trumpet compared to sine tones, preferentially for timbres of the instrument of training.
Kishon-Rabin et al. (2001)	30 (16 M/14 NM)	Pitch Discrimination	Enhanced frequency discrimination thresholds, especially in classical musicians.
Shahin et al. (2004)	13 children (7 M/ 6 NM)	ERP	Larger P1, N1, and P2 amplitude. P2 enhancement was specific to the instrument of practice.
Fujioka (2006)	12 (6 M/6 NM)	ERP	Larger P1 and N450 amplitude. Faster changes of N250 amplitude and latency for violin sound rather compared to noise.

Micheyl et al. (2006)	60 (30 M/30 NM)	FDTs and F0DTs[2]	Enhanced pitch discrimination performance (> FDTs).
Musacchia et al. (2007)	29 (16 M/13 NM)	ABR[3] Recording	Earlier and larger (amplitude) ABR responses to cello sound.
Parbery-Clark et al. (2009)	31 (16 M/ 15 NM)	ABR Recording	Faster ABR response, enhanced representation of speech harmonics, and less degraded response morphology in noise.
Bidelman, Gandour and Krishnan (2011)	28 (14 Chinese/ 14 English Mus)	FFR[4] Measure	Similar FFRs to Chinese for mandarin lexical tones processing and greater pitch strength for one section of the musical stimulus.
Fuller et al. (2014)	25	Behavioral Task	Enhanced melodic contour identification of piano sounds under degraded pitch conditions (cochlear implant simulation).
Rigoulot, Pell & Armony (2015)	33 (15 M/ 18 NM)	ERP	Enhanced N1 amplitude for musical sounds. Musical and vocal sounds seemed to be identically processed.
Bianchi et al. (2016)	14 (6 M/ 8 NM)	Pitch Discrimination/ Pupil Dilation Measures	Enhanced pitch discriminations; lower pupil dilations during the tasks, suggesting a lower effort in performing the tasks.
Nikjeh et al. (2017)	61 (20 VM/ 21 IM/ 20 NM)	EEG/ Pitch Discrimination	Enhanced pitch discrimination and sensory memory representations of harmonic stimuli. Faster MMN[5] response to the deviant stimulus.
Celma-Miralles & Toro (2019)	24 (12 M/12 NM)	EEG	Increased amplitude of EEG spectrum elicited by beat and meter
Speech Processing			

Musacchia et al. (2007)	29 (16M/13NM)	ABR Recording	Earlier brainstem responses.
Wong et al. (2007)	20 (10 M/ 10 NM)	FFR Measure	More faithful representation of the stimulus F0 contours and larger FFR amplitude. Behaviorally, higher identification and discriminating accuracy of the tones.
Parbery-Clark, Strait & Kraus (2011)	31 (16 M/15 NM)	ABR Recording	Enhanced perception of speech in noise and a facilitated response to F0 when speech was predictable.
Parbery-Clark et al. (2012)	50 (23 M/ 27NM)	ABR Recording	Enhanced distinct neural responses between speech syllables (syllables /ba/, /da/ and /ga/).
Strait et al. (2012)	31 (15 M/ 16 NM)	ABR Recording	Stronger neural encoding of acoustic properties for processing speech in noise in children with musical training.
Strait et al. (2014)	76 (21 preschoolers (12 M), 26 school-aged children (13 M) and 29 adults (14 M))	ABR Recording	Enhanced distinct neural responses of stop consonants in all 3 age groups. Better performance on auditory attention and working memory.
Zendel et al. (2015)	26 (13 M/ 13 NM)	ERP	More positive P1 during active listening of speech-in-noise. N4 amplitude was not influenced by the level of background noise.
Intartaglia et al. (2017)	42 (14 native NM/ 18 non-native M/ 10	ABR Recording	Native non-musicians and non-native musicians had similar neural responses to the formant frequencies of a syllable.

	non-native NM)		
Emotion Processing			
Strait et al. (2009)	30 NM	ABR Recording	Enhanced magnitude of the response for the most complex portion and decreased magnitude for more periodic portion of the stimulus.
Lima & Castro (2011)	80 (40 M/ 40 NM)	Behavioral Task (Emotion Recognition)	Increased emotion recognition accuracy (six universal emotions) in speech prosody.
Lima & Castro (2014)	80 (40 M/ 40 NM)	Behavioral Task (Emotion Recognition)	Increased emotion recognition accuracy of musical sounds.
Fuller et al. (2014)	50 (25 M/ 25 NM)	Behavioral Task (Emotion Recognition)	Enhanced vocal emotion identification (anger, sadness, joy and relief) under degraded pitch conditions (cochlear implant simulation).
Pinheiro et al. (2015)	28 (14 M/14 NM)	ERP	Reduced P50 and a more positive P200 for the pure prosody condition. Increased accuracy in recognizing angry prosody.
Nolden et al. (2017)	37 (17 M/ 20 NM)	EEG	Greater activation in frontal theta and alpha for musical sounds, speech and vocalizations and greater activation of frontal alpha when emotional content was present.
Sharp et al. (2019)	34 (17 M / 17 NM)	Behavioral Task (Emotion Recognition)	Enhanced emotion identification of fear/threatful and peaceful melodies.

M – musicians, NM – non-musicians; VM – vocal musicians (singers); IM – instrumentalists. F0 – fundamental frequency; [1] MEG – magnetoencephalography; [2] FDTs – mean frequency – and – F0DTS – F0 discrimination thresholds; [3] ABR – auditory brainstem responses; [4] FFR – frequency following responses; [5] MMN – mismatch negativity.

Appendix B

Summary of studies reviewed examining the vocal emotional processing with ERPs.

Study	N	Stimuli Type	Task	ERP Components	Results
Schirmer & Kotz (2003)	32	Verbs with positive, neutral or negative quality.	Button Press: 1) Judgement of word valence while ignoring the prosody. 2) judge emotional prosody while ignoring word valence.	Three time windows: 350-650 ms; 600-750 ms; 750-900 ms.	For all time windows: (happy and angry): incongruent > congruent word valence; 750 – 900 ms (neutral): congruent > incongruent.
Schirmer, Striano, Friderici (2005)	80	Syllabes 'dada' with anfry, neutral, happy prosody.	Passive listening.	MMN	MMN (amplitude): angry and happy > neutral. MMN (latency): neutral > happy: Arousal seems to increase MMN amplitude and thus the processing resources engaged in change detection.
Paulmann & Kotz (2008)	31	Sentences in seven different ways expressing	Button press - yes and no - about a probe word written on the screen.	P200	P2 (amplitude): emotional > neutral (irrespective of speaker voice).

		basic emotions.	after listening to a sentence.		
Spreckelmeyer et al. (2008)	16	20 syllable ‘ha’ sung with happy and sad musical tones.	Task A – emotion matching between present and previous tone; task B – identity matching – was the singer the same as before?	N1; P2; 300-400 ms; 400-1000 ms.	P2 (amplitude): emotionally congruent < incongruent happy tones; 300-400 (amplitude): congruent > incongruent.
Sauter & Eimer (2010)	10	Nonverbal vocalizations & spectrally rotated versions of the sounds.	Emotional one-back task - button press if the current sound was of the same emotional category as the one before.	Fronto-central electrodes	Fronto-central positivity (150 ms): emotional > neutral; Anterior positivity for fear/achievement/disgust but not for relief.
Jessen & Kotz (2011)	23	Emotional Interjections expressing fear, anger, or no emotion (neutral)	1. Emotion recognition; 2. Lenght of the stimulus judgement.	N1; P2; LPC.	N1 (amplitude): emotional < neutral; N1 (amplitude): fearful < angry; N1 (latency): angry < fearful; P2 (amplitude): emotional > neutral; LPC (amplitude): fearful > angry & neutral only in visual and audiovisual conditions.

Liu et al. (2012)	19	Nonverbal human vocalizations (neutral, happy and angry) and monkey voices.	Press a button when hearing a monkey voice (implicit task).	P50; N1; P2.	P50 (amplitude): angry > neutral; N1 (amplitude): happy and angry < neutral; N1 (amplitude): central and frontocentral > frontal region; N1 (latency): earlier for happy and angry relative to neutral; P2 (amplitude): happy and angry > neutral.
Liu et al., 2012	18	Human faces and human voices (angry, happy or neutral) Auditory, visual and adiovisual conditions.	Human and monkey faces discrimination.	N1; P2; N250; P3; P1; N170; P270.	N1: neutral > emotional; P2: emotional > neutral; P3: emotional > neutral; N250: neutral > emotional; P2, P3 and N250: frontal-central only.
Pinheiro, del Re, Mezin, Nestor (2012)	33	Words in SCC with short length and in	Emotion recognition of the word through button press.	P50; N1; P2.	Reduced P50 (amplitude) for happy PPC prosody was observed in schizophrenia relative to healthy controls (HC); reduced N1 amplitude in schizophrenia was found to both emotional and neutral SCC stimuli, as well as to

		the PPC (distorted)			neutral PPC stimuli; P2 (amplitude): schizophrenia > HC, for happy SCC words.
Gadeke, Focker, Roder (2013)	17	Two-syllabe pseudo-words (neutral, happy, threatening and fearful).	Identify the speaker by lifting the left or right index finger.	N1; P2.	N1: fearful > neutral; P2: fearful > neutral; neutral > happy and threatening; threatening > happy and fearful.
Pell et al. (2015)	24	Emotional speech prosody and non-linguistic vocalizations (anger, sadness or happiness).	Facial affect decision task - yes/no decision about whether the “facial expression represents an emotion”, when presented to a vocal expression followed by a static face.	N1; P2; LPP	N1: vocalizations < speech; P2: speech < vocalizations; angry and sad vocalizations > angry and sad speech; sad vocalizations < happy and angry vocalizations; LPP: speech < vocalizations; speech < vocalizations; angry > sad and happy vocalizations.
Pinheiro et al (2015)	28	Sentences – semantic content condition	Emotional valence discrimination	P50; N1; P2.	N1: PPC[1] > SCC[2] (non-musicians only); PPC > SCC (non-musicians only); happy > angry; P2: angry > happy; PPC > SCC (neutral prosody only); P2 (PPC): neutral > happy / angry > happy; P2 (Musicians): PPC > SCC;

		(SCC) and pure prosody condition (PPR).			recognition accuracy: SCC > PPC; recognition accuracy (SCC): angry > neutral; recognition accuracy (PCC): neutral/happy > angry.
Rigoulot et al. (2015)	39	Musical excerpts (violin and piano), nonlinguistic vocalizations and pseudo-utterances (fear, sadness, happiness and neutral). Pure tones as targets.	Press a button to identify the target sounds – pure tones.	N1; P2.	N1: musical excerpts > vocal sounds; P2: musical excerpts > vocal sounds.
Pinheiro et al. (2016)	16	Adjectives with different semantic valence	Voice identity recognition.	N1; P2; LPP.	N1: neutral self-speech > neutral nonself-speech; P2: positive self-speech > positive nonself-speech; LPP: positive and negative self-speech > positive and negative nonself-speech.

		(neutral, positive and negative).			
Pinheiro et al. (2017)	19	Female emotional vocalizations (happy, neutral and neutral).	Implicit task: count silently the number of target vocalizations in each block.	P3; P3a; P3b.	P3a: laughs and growls > neutral; P3b: laughs and growls > neutral. Note: the P3b response to happy targets was preceded and predicted by increased alpha desynchronization, as well as by the perceived pleasantness of the vocal cues.

‘>’ is used for representing enhanced amplitude for one group or condition in comparison to other group or condition

[1]PPC – pure prosody condition; [2]SCC – semantic content condition.

REFERENCES

Aminoff, M. J., Boller, F., & Swaab, D. F. (2015). The Human Auditory System—Fundamental Organization and Clinical Disorders. Em *Handbook of Clinical Neurology* (Vol. 129, pp. 2–704). Elsevier.

Amir, O., Amir, N., & Kishon-Rabin, L. (2003). The effect of superior auditory skills on vocal accuracy. *The Journal of the Acoustical Society of America*, *113*(2), 1102–1108. https://doi.org/10.1121/1.1536632

Anderson, S., & Kraus, N. (2011). Neural Encoding of Speech and Music: Implications for Hearing Speech in Noise. *Seminars in Hearing*, *32*(02), 129–141. https://doi.org/10.1055/s-0031-1277234

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

Armony, J. L., Aubé, W., Angulo-Perkins, A., Peretz, I., & Concha, L. (2015). The specificity of neural responses to music and their relation to voice processing: An fMRI-adaptation study. *Neuroscience Letters*, *593*, 35–39. https://doi.org/10.1016/j.neulet.2015.03.011

Aubé, W., Angulo-Perkins, A., Peretz, I., Concha, L., & Armony, J. L. (2015). Fear across the senses: Brain responses to music, vocalizations and facial expressions. *Social Cognitive and Affective Neuroscience*, *10*(3), 399–407. https://doi.org/10.1093/scan/nsu067

Banissy, M. J., Sauter, D. A., Ward, J., Warren, J. E., Walsh, V., & Scott, S. K. (2010). Suppressing Sensorimotor Activity Modulates the Discrimination of Auditory Emotions But Not Speaker Identity. *Journal of Neuroscience*, *30*(41), 13552–13557. https://doi.org/10.1523/JNEUROSCI.0786-10.2010

Barrett, K. C., Ashley, R., Strait, D. L., & Kraus, N. (2013). Art and science: How musical training shapes the brain. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00713

Belin, P. (2006). Voice processing in human and non-human primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1476), 2091–2107. https://doi.org/10.1098/rstb.2006.1933

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. https://doi.org/10.3758/BRM.40.2.531

Belin, P., Zatorre, R. J., & Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Cognitive Brain Research*, *13*(1), 17–26. https://doi.org/10.1016/S0926-6410(01)00084-2

Bianchi, F., Santurette, S., Wendt, D., & Dau, T. (2016). Pitch Discrimination in Musicians and Non-Musicians: Effects of Harmonic Resolvability and Processing Effort. *Journal of the Association for Research in Otolaryngology*, *17*(1), 69–79. https://doi.org/10.1007/s10162-015-0548-2

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Cross-domain Effects of Music and Language Experience on the Representation of Pitch in the Human Auditory Brainstem. *Journal of Cognitive Neuroscience*, *23*(2), 425–434. https://doi.org/10.1162/jocn.2009.21362

Britton, J. W., Frey, L. C., Hopp, J. L., Korb, P., Koubeissi, M. Z., Lievens, W. E., Pestana-Knight, E. M., & St. Louis, E. K. (2016). *Electroencephalography (EEG): An Introductory Text and Atlas of Normal and Abnormal Findings in Adults, Children, and Infants* (E. K. St. Louis &

L. C. Frey, Eds.). American Epilepsy Society. http://www.ncbi.nlm.nih.gov/books/NBK390354/

Brown, S., Martinez, M. J., Hodges, D. A., Fox, P. T., & Parsons, L. M. (2004). The song system of the human brain. *Cognitive Brain Research*, *20*(3), 363–375. https://doi.org/10.1016/j.cogbrainres.2004.03.016

Burgoon, J. K., Guerrero, L. K., & Floyd, K. (2016). *Nonverbal Communication*. Taylor & Francis.

Callan, D. E., Tsytsarev, V., Hanakawa, T., Callan, A. M., Katsuhara, M., Fukuyama, H., & Turner, R. (2006). Song and speech: Brain regions involved with perception and covert production. *NeuroImage*, *31*(3), 1327–1342. https://doi.org/10.1016/j.neuroimage.2006.01.036

Castro, S. L., & Lima, C. F. (2014). Age and Musical Expertise Influence Emotion Recognition in Music. *Music Perception: An Interdisciplinary Journal*, *32*(2), 125–142. https://doi.org/10.1525/mp.2014.32.2.125

Celma-Miralles, A., & Toro, J. M. (2019). Ternary meter from spatial sounds: Differences in neural entrainment between musicians and non-musicians. *Brain and Cognition*, *136*, 103594. https://doi.org/10.1016/j.bandc.2019.103594

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

Christiner, M., & Reiterer, S. M. (2015). A Mozart is not a Pavarotti: Singers outperform instrumentalists on foreign accent imitation. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00482

Citron, F. M. M., Gray, M. A., Critchley, H. D., Weekes, B. S., & Ferstl, E. C. (2014). Emotional valence and arousal affect reading in an interactive way: Neuroimaging evidence for an approach-withdrawal framework. *Neuropsychologia*, *56*(100), 79–89. https://doi.org/10.1016/j.neuropsychologia.2014.01.002

Coffey, E. B. J., Mogilever, N. B., & Zatorre, R. J. (2017). Speech-in-noise perception in musicians: A review. *Hearing Research*, *352*, 49–69. https://doi.org/10.1016/j.heares.2017.02.006

Cohen, N., Moyal, N., Lichtenstein-Vidne, L., & Henik, A. (2016). Explicit vs. implicit emotional processing: The interaction between processing type and executive control. *Cognition & Emotion*, *30*(2), 325–339. https://doi.org/10.1080/02699931.2014.1000830

de Villers-Sidani, E., & Merzenich, M. M. (2011). Lifelong plasticity in the rat auditory cortex. Em *Progress in Brain Research* (Vol. 191, pp. 119–131). Elsevier. https://doi.org/10.1016/B978-0-444-53752-2.00009-6

Demarin, V., Morovic, S., & Béné, R. (2014). Neuroplasticity. *Periodicum Biologorum*, *116*, 209–211.

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

Ding, R., Li, P., Wang, W., & Luo, W. (2017). Emotion Processing by ERP Combined with Development and Plasticity. *Neural Plasticity*, *2017*, 1–15. https://doi.org/10.1155/2017/5282670

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology (London, England: 1953)*, *92 Part 1*, 53–78.

Eimer, M., & Holmes, A. (2007). Event-related brain potential correlates of emotional face processing. *Neuropsychologia*, *45*(1), 15–31. https://doi.org/10.1016/j.neuropsychologia.2006.04.022

Ekman, P. (1992). Are there basic emotions? *Psychological Review*, *99*(3), 550–553. https://doi.org/10.1037/0033-295X.99.3.550

Espírito-Santo, H., Pires, C. F., Garcia, I. Q., Daniel, F., Silva, A. G. da, & Fazio, R. L. (2017). *Preliminary validation of the Portuguese Edinburgh Handedness Inventory in an adult sample*. *24(3)*, 275–287. https://doi.org/10.1080/23279095.2017.1290636

Ferreira-Santos, F., Silveira, C., Almeida, P. R., Palha, A., Barbosa, F., & Marques-Teixeira, J. (2012). The auditory P200 is both increased and reduced in schizophrenia? A meta-analytic dissociation of the effect for standard and target stimuli in the oddball task. *Clinical Neurophysiology*, *123*(7), 1300–1308. https://doi.org/10.1016/j.clinph.2011.11.036

Fitsiori, A., Nguyen, D., Karentzos, A., Delavelle, J., & Vargas, M. I. (2011). The corpus callosum: White matter or terra incognita. *The British Journal of Radiology*, *84*(997), 5–18. https://doi.org/10.1259/bjr/21946513

Fridberg, D. J., Hetrick, W. P., Brenner, C. A., Shekhar, A., Steffen, A. N., Malloy, F. W., & O'Donnell, B. F. (2009). Relationships between Auditory Event-Related Potentials and Mood State, Medication, and Comorbid Psychiatric Illness in Patients with Bipolar Disorder. *Bipolar disorders*, *11*(8), 857–866. https://doi.org/10.1111/j.1399-5618.2009.00758.x

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

Fujioka, T. (2006). One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain*, *129*(10), 2593–2608. https://doi.org/10.1093/brain/awl247

Fuller, C. D., Galvin, J. J. I., Maat, B., Free, R. H., & Başkent, D. (2014). The musician effect: Does it persist under degraded pitch conditions of cochlear implant simulations? *Frontiers in Neuroscience*, *8*. https://doi.org/10.3389/fnins.2014.00179

Gaser, C., & Schlaug, G. (2003). Brain Structures Differ between Musicians and Non-Musicians. *The Journal of Neuroscience*, *23*(27), 9240–9245. https://doi.org/10.1523/JNEUROSCI.23-27-09240.2003

Grafman, J. (2000). Conceptualizing functional neuroplasticity. *Journal of Communication Disorders*, *33*(4), 345–356. https://doi.org/10.1016/S0021-9924(00)00030-7

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95–112. https://doi.org/10.1007/BF02289823

Grimshaw-Aagaard, M., Walther-Hansen, M., & Knakkergaard, M. (2019). *The Oxford Handbook of Sound and Imagination, Volume 2*. Oxford University Press.

Habib, M., & Besson, M. (2009). What do Music Training and Musical Experience Teach Us About Brain Plasticity? *Music Perception: An Interdisciplinary Journal*, *26*(3), 279–285. https://doi.org/10.1525/mp.2009.26.3.279

Hajcak, G., MacNamara, A., & Olvet, D. M. (2010). Event-Related Potentials, Emotion, and Emotion Regulation: An Integrative Review. *Developmental Neuropsychology*, *35*(2), 129–155. https://doi.org/10.1080/87565640903526504

Halwani, G. F., Loui, P., Rüber, T., & Schlaug, G. (2011). Effects of Practice and Experience on the Arcuate Fasciculus: Comparing Singers, Instrumentalists, and Non-Musicians. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00156

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), 293–305. https://doi.org/10.1037/a0015178

Henson, R. (1985). Amusia. Em *Handbook of Clinical Neurology* (pp. 483–490). New York: Elsevier.

Herholz, S. C., & Zatorre, R. J. (2012). Musical Training as a Framework for Brain Plasticity: Behavior, Function, and Structure. *Neuron*, *76*(3), 486–502. https://doi.org/10.1016/j.neuron.2012.10.011

Holmes, J. N., Holmes, W. J., & Garner, P. N. (1997). *Using Formant Frequencies in Speech Recognition*. 4.

Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). Musical Training Shapes Structural Brain Development. *Journal of Neuroscience*, *29*(10), 3019–3025. https://doi.org/10.1523/JNEUROSCI.5118-08.2009

Intartaglia, B., White-Schwoch, T., Kraus, N., & Schön, D. (2017). Music training enhances the automatic neural processing of foreign speech sounds. *Scientific Reports*, *7*(1), 12631. https://doi.org/10.1038/s41598-017-12575-1

J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson. (2007). *Handbook of psychophysiology*. Cambridge Univ Pr. https://doi.org/10.13140/2.1.2871.1369

Jacob, H., Brück, C., Domin, M., Lotze, M., & Wildgruber, D. (2014). I can't keep your face and voice out of my head: Neural correlates of an attentional bias toward nonverbal emotional cues. *Cerebral Cortex (New York, N.Y.: 1991)*, *24*(6), 1460–1473. https://doi.org/10.1093/cercor/bhs417

Jacob, H., Brück, C., Kreifelts, B., Erb, M., Braun, S. R., & Wildgruber, D. (2012). *Cerebral integration of verbal and nonverbal emotional cues: Impact of individual nonverbal dominance. - PubMed—NCBI*. https://www.ncbi.nlm.nih.gov/pubmed/22516367

Jasper, H. H. (1958). The Ten-Twenty Electrode System of the International Federation. *Electroencephalography and Clinical Neurophysiology*, *10*, 370–375.

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

Jones, J. A., & Keough, D. (2008). Auditory-motor mapping for pitch control in singers and nonsingers. *Experimental Brain Research*, *190*(3), 279–287. https://doi.org/10.1007/s00221-008-1473-y

Joos, K., Gilles, A., Van de Heyning, P., De Ridder, D., & Vanneste, S. (2014). From sensation to percept: The neural signature of auditory event-related potentials. *Neuroscience & Biobehavioral Reviews*, *42*, 148–156. https://doi.org/10.1016/j.neubiorev.2014.02.009

Jungblut, M., Huber, W., Pustelniak, M., & Schnitker, R. (2012). The impact of rhythm complexity on brain activation during simple singing: An event-related fMRI study. *Restorative Neurology and Neuroscience*, *30*(1), 39–53. https://doi.org/10.3233/RNN-2011-0619

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A., & Hudspeth, A. J. (2000). *Principles of Neural Science* (Fifth Edition). McGraw-Hill QDR.

Kania, B. F., Wrońska, D., & Zięba, D. (2017). Introduction to Neural Plasticity Mechanism. *Journal of Behavioral and Brain Science*, *07*(02), 41–49. https://doi.org/10.4236/jbbs.2017.72005

Kim, D.-W., & Im, C.-H. (2018). EEG Spectral Analysis. Em C.-H. Im (Ed.), *Computational EEG Analysis: Methods and Applications* (pp. 35–53). Springer. https://doi.org/10.1007/978-981-13-0908-3_3

Kishon-Rabin, L., Amir, O., Vexler, Y., & Zaltz, Y. (2001). Pitch Discrimination: Are Professional Musicians Better than Non-Musicians? *Journal of Basic and Clinical Physiology and Pharmacology*, *12*(2). https://doi.org/10.1515/JBCPP.2001.12.2.125

Kleber, B., Friberg, A., Zeitouni, A., & Zatorre, R. (2017). Experience-dependent modulation of right anterior insula and sensorimotor regions as a function of noise-masked auditory feedback in singers and nonsingers. *NeuroImage*, *147*, 97–110. https://doi.org/10.1016/j.neuroimage.2016.11.059

Kleber, B., Veit, R., Birbaumer, N., Gruzelier, J., & Lotze, M. (2010). The Brain of Opera Singers: Experience-Dependent Changes in Functional Activation. *Cerebral Cortex*, *20*(5), 1144–1152. https://doi.org/10.1093/cercor/bhp177

Kleber, B., Veit, R., Moll, C. V., Gaser, C., Birbaumer, N., & Lotze, M. (2016). Voxel-based morphometry in opera singers: Increased gray-matter volume in right somatosensory and auditory cortices. *NeuroImage*, *133*, 477–483. https://doi.org/10.1016/j.neuroimage.2016.03.045

Kleber, B., Zeitouni, A. G., Friberg, A., & Zatorre, R. J. (2013). Experience-dependent modulation of feedback integration during singing: Role of the right anterior insula. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(14), 6070–6080. https://doi.org/10.1523/JNEUROSCI.4418-12.2013

Knapp, M. L., Hall, J. A., & Horgan, T. G. (2014). *Nonverbal Communication in Human Interaction* (8.ª ed.). Wadsworth, Cengage Learning.

Kolb, B., & Gibb, R. (2011). Brain Plasticity and Behaviour in the Developing Brain. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *20*(4), 265–276.

Koolagudi, S. G., & Rao, K. S. (2012). Emotion recognition from speech: A review. *International Journal of Speech Technology*, *15*(2), 99–117. https://doi.org/10.1007/s10772-011-9125-1

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, *11*(8), 599–605. https://doi.org/10.1038/nrn2882

Kriegstein, K. V., & Giraud, A.-L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *NeuroImage*, *22*(2), 948–955. https://doi.org/10.1016/j.neuroimage.2004.02.020

Lang, P. J., & Bradley, M. M. (2010). Emotion and the motivational brain. *Biological psychology*, *84*(3), 437–450. https://doi.org/10.1016/j.biopsycho.2009.10.007

Lang, P. J., & Davis, M. (2006). Emotion, motivation, and the brain: Reflex foundations in animal and human research. *Progress in Brain Research*, *156*, 3–29. https://doi.org/10.1016/S0079-6123(06)56001-7

Lappe, C., Herholz, S. C., Trainor, L. J., & Pantev, C. (2008). Cortical Plasticity Induced by Short-Term Unimodal and Multimodal Musical Training. *Journal of Neuroscience*, *28*(39), 9632–9639. https://doi.org/10.1523/JNEUROSCI.2254-08.2008

Legatt, A. D. (2015). Chapter 17—Electrophysiologic auditory tests. Em M. J. Aminoff, F. Boller, & D. F. Swaab (Eds.), *Handbook of Clinical Neurology* (Vol. 129, pp. 289–311). Elsevier. https://doi.org/10.1016/B978-0-444-62630-1.00017-2

Lima, C. F. (2011). *Perceiving emotions in speech and music: Modulations related to ageing, musical training, and Parkinson's disease*. https://sigarra.up.pt/fdup/pt/pub_geral.pub_view?pi_pub_base_id=201363&pi_pub_r1_id=

Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion*, *11*(5), 1021–1031. https://doi.org/10.1037/a0024521

Lima, C. F., Correia, A., Müllensiefen, D., & Castro, S. L. (2018). Goldsmiths Musical Sophistication Index (Gold-MSI): Portuguese version and associations with socio-demographic factors, personality and music preferences. *Psychology of Music*, 030573561880199. https://doi.org/10.1177/0305735618801997

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

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. https://doi.org/10.1371/journal.pone.0031001

Lohmann, C., & Kessels, H. W. (2014). The developmental stages of synaptic plasticity. *The Journal of Physiology*, *592*(Pt 1), 13–31. https://doi.org/10.1113/jphysiol.2012.235119

Loui, P., Alsop, D., & Schlaug, G. (2009). Tone Deafness: A New Disconnection Syndrome? *The Journal of Neuroscience*, *29*(33), 10215–10220. https://doi.org/10.1523/JNEUROSCI.1701-09.2009

Luck, S. (2005). An introduction to event-related potentials and their neural origins. *An Introduction to the Event-Related Potential Technique*, *107*.

Mavridis, I. N., & Pyrgelis, E.-S. (2016). Brain Activation During Singing: “Clef de Sol Activation” Is the “Concert” of the Human Brain. *Medical Problems of Performing Artists*, *31*(1), 45–50. https://doi.org/10.21091/mppa.2016.1008

Max, L., Guenther, F. H., Gracco, V. L., Ghosh, S. S., & Wallace, M. E. (2004). Unstable or Insufficiently Activated Internal Models and Feedback-Biased Motor Control as Sources of Dysfluency: A Theoretical Model of Stuttering. *Contemporary Issues in Communication Science and Disorders*, *31*(Spring), 105–122. https://doi.org/10.1044/cicsd_31_S_105

McGettigan, C., Walsh, E., Jessop, R., Agnew, Z. K., Sauter, D. A., Warren, J. E., & Scott, S. K. (2015). Individual Differences in Laughter Perception Reveal Roles for Mentalizing and Sensorimotor Systems in the Evaluation of Emotional Authenticity. *Cerebral Cortex*, *25*(1), 246–257. https://doi.org/10.1093/cercor/bht227

Menninghaus, W., Wagner, V., Hanich, J., Wassiliwizky, E., Kuehnast, M., & Jacobsen, T. (2015). Towards a psychological construct of being moved. *PloS One*, *10*(6), e0128451. https://doi.org/10.1371/journal.pone.0128451

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, *219*(1–2), 36–47. https://doi.org/10.1016/j.heares.2006.05.004

Münte, T., Urbach, T., Duzel, E., & Kutas, M. (2000). Event-related brain potentials in the study of human cognition and neuropsychology. *Handbook of Neuropsychology*, *1*.

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences*, *104*(40), 15894–15898. https://doi.org/10.1073/pnas.0701498104

Neumann, O., & Sanders, A. F. (1996). *Handbook of Perception and Action: Attention*. Academic Press.

Nikjeh, D. A., Lister, J. J., & Frisch, S. A. (2008). Hearing of note: An electrophysiologic and psychoacoustic comparison of pitch discrimination between vocal and instrumental musicians. *Psychophysiology*, *45*(6), 994–1007. https://doi.org/10.1111/j.1469-8986.2008.00689.x

Nolden, S., Rigoulot, S., Jolicoeur, P., & Armony, J. L. (2017). Effects of musical expertise on oscillatory brain activity in response to emotional sounds. *Neuropsychologia*, *103*, 96–105. https://doi.org/10.1016/j.neuropsychologia.2017.07.014

Onitsuka, T., Oribe, N., & Kanba, S. (2013). Chapter 13—Neurophysiological findings in patients with bipolar disorder. Em E. Başar, C. Başar-Eroĝlu, A. Özerdem, P. M. Rossini, & G. G. Yener (Eds.), *Supplements to Clinical Neurophysiology* (Vol. 62, pp. 197–206). Elsevier. https://doi.org/10.1016/B978-0-7020-5307-8.00013-2

Oxenham, A. J. (2012). Pitch Perception. *The Journal of Neuroscience*, *32*(39), 13335–13338. https://doi.org/10.1523/JNEUROSCI.3815-12.2012

Oztürk, A. H., Tascioglu, B., Aktekin, M., Kurtoğlu, Z., & Erden, I. (2002). Morphometric comparison of the human corpus callosum in professional musicians and non-musicians by using in vivo magnetic resonance imaging. *Journal of neuroradiology. Journal de neuroradiologie*, *29*(1), 29–34.

Paquette, S., Peretz, I., & Belin, P. (2013). The «Musical Emotional Bursts»: A validated set of musical affect bursts to investigate auditory affective processing. *Frontiers in Psychology*, *4*, 509. https://doi.org/10.3389/fpsyg.2013.00509

Paquette, S., Takerkart, S., Saget, S., Peretz, I., & Belin, P. (2018). Cross-classification of musical and vocal emotions in the auditory cortex: Cross-classification of musical and vocal emotions. *Annals of the New York Academy of Sciences*, *1423*(1), 329–337. https://doi.org/10.1111/nyas.13666

Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical Experience Limits the Degradative Effects of Background Noise on the Neural Processing of Sound. *Journal of Neuroscience*, *29*(45), 14100–14107. https://doi.org/10.1523/JNEUROSCI.3256-09.2009

Parbery-Clark, A., Strait, D. L., & Kraus, N. (2011). Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia*, *49*(12), 3338–3345. https://doi.org/10.1016/j.neuropsychologia.2011.08.007

Parbery-Clark, A., Tierney, A., Strait, D. L., & Kraus, N. (2012). Musicians have fine-tuned neural distinction of speech syllables. *Neuroscience*, *219*, 111–119. https://doi.org/10.1016/j.neuroscience.2012.05.042

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

Paulmann, S., & Kotz, S. A. (2006). *Valence, arousal, and task effects on the P200 in emotional prosody processing*.

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

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

Peretz, I., Vuvan, D., Lagrois, M.-É., & Armony, J. L. (2015). Neural overlap in processing music and speech. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1664). https://doi.org/10.1098/rstb.2014.0090

Peretz, I., & Zatorre, R. J. (2005). Brain Organization for Music Processing. *Annual Review of Psychology*, *56*(1), 89–114. https://doi.org/10.1146/annurev.psych.56.091103.070225

Perry, D. W., Zatorre, R. J., Petrides, M., Alivisatos, B., Meyer, E., & Evans, A. C. (1999). Localization of cerebral activity during simple singing: *NeuroReport*, *10*(16), 3453–3458. https://doi.org/10.1097/00001756-199911080-00035

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D., & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*(2), 127–152. https://doi.org/10.1111/1469-8986.3720127

Pinheiro, A. P., del Re, E., Mezin, J., Nestor, P. G., Rauber, A., McCarley, R. W., Gonçalves, Ó. F., & Niznikiewicz, M. A. (2013). Sensory-based and higher-order operations contribute to abnormal emotional prosody processing in schizophrenia: An electrophysiological investigation. *Psychological Medicine*, *43*(03), 603–618. https://doi.org/10.1017/S003329171200133X

Pinheiro, A. P., Rezaii, N., Nestor, P. G., Rauber, A., Spencer, K. M., & Niznikiewicz, M. (2016). Did you or I say pretty, rude or brief? An ERP study of the effects of speaker's identity on emotional word processing. *Brain and Language*, *153–154*, 38–49. https://doi.org/10.1016/j.bandl.2015.12.003

Pinheiro, A. P., Rezaii, N., Rauber, A., Nestor, P. G., Spencer, K. M., & Niznikiewicz, M. (2017). Emotional self-other voice processing in schizophrenia and its relationship with hallucinations: ERP evidence. *Psychophysiology*, *54*(9), 1252–1265. https://doi.org/10.1111/psyp.12880

Pinheiro, A. P., Vasconcelos, M., Dias, M., Arrais, N., & Gonçalves, Ó. F. (2015). The music of language: An ERP investigation of the effects of musical training on emotional prosody processing. *Brain and Language*, *140*, 24–34. https://doi.org/10.1016/j.bandl.2014.10.009

Proverbio, A. M., Benedetto, F., & Guazzone, M. (2019). Shared neural mechanisms for processing emotions in music and vocalizations. *European Journal of Neuroscience*, ejn.14650. https://doi.org/10.1111/ejn.14650

Rigoulot, S., & Pell, M. D. (2014). Emotion in the voice influences the way we scan emotional faces. *Speech Communication*, *65*, 36–49. https://doi.org/10.1016/j.specom.2014.05.006

Rigoulot, S., Pell, M. D., & Armony, J. L. (2015). Time course of the influence of musical expertise on the processing of vocal and musical sounds. *Neuroscience*, *290*, 175–184. https://doi.org/10.1016/j.neuroscience.2015.01.033

Rosslau, K., Herholz, S. C., Knief, A., Ortmann, M., Deuster, D., Schmidt, C.-M., Zehnhoff-Dinnesen, A., Pantev, C., & Dobel, C. (2016). Song Perception by Professional Singers and Actors: An MEG Study. *PLOS ONE*, *11*(2), e0147986. https://doi.org/10.1371/journal.pone.0147986

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences*, *107*(6), 2408–2412. https://doi.org/10.1073/pnas.0908239106

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, *10*(1), 24–30. https://doi.org/10.1016/j.tics.2005.11.009

Schirmer, A., Striano, T., & Friederici, A. D. (2005). *Sex differences in the preattentive processing of vocal emotional expressions*. 5.

Schlaug, G. (2015). Musicians and music making as a model for the study of brain plasticity. Em *Progress in Brain Research* (Vol. 217, pp. 37–55). Elsevier. https://doi.org/10.1016/bs.pbr.2014.11.020

Schlaug, G., Jäncke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologia*, *33*(8), 1047–1055. https://doi.org/10.1016/0028-3932(95)00045-5

Schupp, H. T., Flaisch, T., Stockburger, J., & Junghöfer, M. (2006). Emotion and attention: Event-related brain potential studies. Em *Progress in Brain Research* (Vol. 156, pp. 31–51). Elsevier. https://doi.org/10.1016/S0079-6123(06)56002-9

Shahin, A., Roberts, L., & Trainor, L. (2004). Enhancement of auditory cortical development by musical experience in children. *Neuroreport*, *15*, 1917–1921. https://doi.org/10.1097/00001756-200408260-00017

Sharp, A., Houde, M.-S., Bacon, B.-A., & Champoux, F. (2019). Musicians Show Better Auditory and Tactile Identification of Emotions in Music. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01976

Shestyuk, A. Y., & Deldin, P. J. (2010). Automatic and strategic representation of the self in major depression: Trait and state abnormalities. *The American Journal of Psychiatry*, *167*(5), 536–544. https://doi.org/10.1176/appi.ajp.2009.06091444

Sluming, V., Barrick, T., Howard, M., Cezayirli, E., Mayes, A., & Roberts, N. (2002). Voxel-Based Morphometry Reveals Increased Gray Matter Density in Broca's Area in Male Symphony Orchestra Musicians. *NeuroImage*, *17*(3), 1613–1622. https://doi.org/10.1006/nimg.2002.1288

Spreckelmeyer, K. N., Kutas, M., Urbach, T., Altenmüller, E., & Münte, T. F. (2009). Neural processing of vocal emotion and identity. *Brain and Cognition*, *69*(1), 121–126. https://doi.org/10.1016/j.bandc.2008.06.003

Stiles, J., & Jernigan, T. L. (2010). The Basics of Brain Development. *Neuropsychology Review*, *20*(4), 327–348. https://doi.org/10.1007/s11065-010-9148-4

Strait, D. L., Kraus, N., Skoe, E., & Ashley, R. (2009). Musical experience and neural efficiency—Effects of training on subcortical processing of vocal expressions of emotion. *European Journal of Neuroscience*, *29*(3), 661–668. https://doi.org/10.1111/j.1460-9568.2009.06617.x

Strait, D. L., O'Connell, S., Parbery-Clark, A., & Kraus, N. (2014). Musicians' Enhanced Neural Differentiation of Speech Sounds Arises Early in Life: Developmental Evidence from Ages 3 to 30. *Cerebral Cortex*, *24*(9), 2512–2521. https://doi.org/10.1093/cercor/bht103

Strait, D. L., Parbery-Clark, A., Hittner, E., & Kraus, N. (2012). Musical training during early childhood enhances the neural encoding of speech in noise. *Brain and Language*, *123*(3), 191–201. https://doi.org/10.1016/j.bandl.2012.09.001

Sundberg, J. (1987). *The Science of the Singing Voice.* DeKalb, IL: Nothern Illinois University Press.

Sundberg, J., Iwarsson, J., & Billström, A. H. (1995). Significance of mechanoreceptors in the subglottal mucosa for subglottal pressure control in singers. *Journal of Voice: Official Journal of the Voice Foundation*, *9*(1), 20–26. https://doi.org/10.1016/s0892-1997(05)80219-x

Sur, S., & Sinha, V. K. (2009). Event-related potential: An overview. *Industrial Psychiatry Journal*, *18*(1), 70–73. https://doi.org/10.4103/0972-6748.57865

Thigpen, N. N., Kappenman, E. S., & Keil, A. (2017). Assessing the internal consistency of the event-related potential: An example analysis. *Psychophysiology*, *54*(1), 123–138. https://doi.org/10.1111/psyp.12629

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: Do music lessons help? *Emotion (Washington, D.C.)*, *4*(1), 46–64. https://doi.org/10.1037/1528-3542.4.1.46

Trainor, L. (2010). Using Electroencephalography (EEG) to Measure Maturation of Auditory Cortex in Infants: Processing Pitch, Duration and Sound Location. *Encyclopedia on Early Childhood Development*.

Vasconcelos, M., Dias, M., Soares, A. P., & Pinheiro, A. P. (2017). What is the Melody of That Voice? Probing Unbiased Recognition Accuracy with the Montreal Affective Voices. *Journal of Nonverbal Behavior*, *41*(3), 239–267. https://doi.org/10.1007/s10919-017-0253-4

Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*(12), 585–594. https://doi.org/10.1016/j.tics.2005.10.011

Vuoskoski, J. K., & Eerola, T. (2017). The Pleasure Evoked by Sad Music Is Mediated by Feelings of Being Moved. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00439

Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior*, *17*(1), 3–28. https://doi.org/10.1007/BF00987006

Wilson, S. J., Abbott, D. F., Lusher, D., Gentle, E. C., & Jackson, G. D. (2011). Finding your voice: A singing lesson from functional imaging. *Human Brain Mapping*, *32*(12), 2115–2130. https://doi.org/10.1002/hbm.21173

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*(4), 420–422. https://doi.org/10.1038/nn1872

Yamadori, A., Osumi, Y., Masuhara, S., & Okubo, M. (1977). Preservation of singing in Broca’s aphasia. *Journal of Neurology, Neurosurgery, and Psychiatry*, *40*(3), 221–224.

Zarate, J. M. (2005). Neural Substrates Governing Audiovocal Integration for Vocal Pitch Regulation in Singing. *Annals of the New York Academy of Sciences*, *1060*(1), 404–408. https://doi.org/10.1196/annals.1360.058

Zarate, J. M. (2013). The neural control of singing. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00237

Zarate, J. M., Delhommeau, K., Wood, S., & Zatorre, R. J. (2010). Vocal accuracy and neural plasticity following micromelody-discrimination training. *PloS One*, *5*(6), e11181. https://doi.org/10.1371/journal.pone.0011181

Zarate, J. M., & Zatorre, R. J. (2008). Experience-dependent neural substrates involved in vocal pitch regulation during singing. *NeuroImage*, *40*(4), 1871–1887. https://doi.org/10.1016/j.neuroimage.2008.01.026

Zendel, B. R., Tremblay, C.-D., Belleville, S., & Peretz, I. (2015). The Impact of Musicianship on the Cortical Mechanisms Related to Separating Speech from Background Noise. *Journal of Cognitive Neuroscience*, *27*(5), 1044–1059. https://doi.org/10.1162/jocn_a_00758

Zilbovicius, M., Meresse, I., Chabane, N., Brunelle, F., Samson, Y., & Boddaert, N. (2006). Autism, the superior temporal sulcus and social perception. *Trends in Neurosciences*, *29*(7), 359–366. https://doi.org/10.1016/j.tins.2006.06.004