

University of Porto
Faculty of Psychology and Educational Sciences

**BRAIN STRUCTURE CORRELATES OF EMOTIONAL SPEECH PROSODY
RECOGNITION IN CHILDREN**

Ana Isabel Costa Correia

October 2016

Dissertation presented for the Integrated Master course in Psychology, Faculty of Psychology and Educational Sciences of the University of Porto, supervised by Doctor *César Lima* (FPCEUP and University College London) and Professor *São Luís Castro* (FPCEUP).

LEGAL NOTICES

The content of the present dissertation reflects the author's perspectives, work and interpretation at the time of its delivery. This dissertation may contain inaccuracies, both conceptual and methodological, that may have been identified subsequent to the delivery time. Therefore, any use of its content should be exercised with caution.

By submitting this dissertation, the author states that it is the result of her own work, contains original contributions and all sources used are acknowledged, found properly cited in the text and identified in the references section. The author also states that she does not disclose any contents of this dissertation whose reproduction is prohibited by copyright or intellectual property rights.

Acknowledgements

First, I would like to thank my supervisor, Doctor Cesar Lima, for his continuous commitment, support, and giant patience. With his huge knowledge and incessant dedication, he provided the best guidance I could ever ask for, and for that I am truly thankful.

I am also grateful to my co-supervisor, Professor São Luís Castro, whose expertise, guidance and kindness helped me improve my knowledge on the research topic, and certainly contributed to the improvement of the final project.

I would like to express my gratitude to Marta, Leonor, and Carolina. For being always present, for their incessant concern, for providing me with materials, input, and constant motivation, I am hugely indebted to them.

I am also thankful to Helena, Paulo, and all the members of Neurocognition and Language Research Group, for having clarified my doubts, provided me with crucial support, and motivated me whenever I needed, especially when I did not realize I needed it.

Besides, I am highly grateful for having shared this experience with my colleagues from the Psychology department. Carina, Ana Mesquita, Juliana, and all of you who accompanied me throughout my years of study, thank you for all the shared moments and continuous support. You definitely made this journey a lot easier and memorable.

Finally, last but not least, I must express my deepest gratitude to my family and friends, for providing me with incessant support and encouragement throughout my entire learning path. With them, I shared my anxieties and doubts, but also my achievements. They were always there, and I would not be able to get here without them. Thank you!

Resumo

A capacidade de inferir os estados emocionais das outras pessoas a partir de variações de prosódia da fala tem valor adaptativo e é crucial para a adaptação pessoal e social. Esta capacidade emerge cedo no percurso desenvolvimental e está relacionada com a melhoria das competências socioemocionais das crianças. Contudo, as bases cerebrais do reconhecimento de prosódia emocional em crianças permanecem pouco estudadas. O principal objetivo do presente estudo foi investigar se diferenças na morfologia cerebral poderiam explicar diferenças individuais na capacidade das crianças reconhecerem emoções através de pistas prosódicas. Uma amostra de 66 crianças ($M = 8.30$ anos; $SD = 0.35$) completou uma tarefa comportamental e uma ressonância magnética estrutural. Durante a tarefa comportamental, as crianças ouviam frases neutras em termos semânticos e tinham de realizar dois julgamentos consecutivos, incluindo uma categorização de escolha-forçada da tonalidade emocional (neutro, alegre, triste, zangado, assustado) e um julgamento de intensidade, avaliando a saliência da emoção no estímulo. Os resultados revelaram que as crianças obtiveram elevadas taxas de exatidão no reconhecimento de todas as categorias emocionais e que a alegria foi a emoção melhor reconhecida. Além disso, em termos de estruturas cerebrais, houve correlações entre um melhor reconhecimento de emoções prosódicas e um maior volume de matéria cinzenta no girus fusiforme, no cerebelo e em áreas motoras/pré-motoras e pré-frontais; e um menor volume de matéria cinzenta em regiões parietais e occipitais. Do mesmo modo, descobrimos que algumas regiões cerebrais estavam correlacionadas com elevada exatidão no reconhecimento de emoções específicas, quando comparadas com as outras emoções. Os nossos resultados sugerem que as diferenças individuais na capacidade de as crianças reconhecerem emoções através de pistas prosódicas refletem diferenças na morfologia cerebral, tanto para o reconhecimento da prosódia emocional em geral como para o reconhecimento de categorias emocionais específicas.

Palavras-chave: emoção; prosódia da fala; reconhecimento de emoções; crianças; neuroimagem; *voxel-based morphometry*.

Abstract

The ability to understand others' emotional states from variations in speech prosody has an adaptive value and is crucial both for personal and social adjustment. This ability emerges early in development, and it is related with better socio-emotional skills in children. However, the neural basis of emotional prosody recognition in children remains poorly understood. The main goal of the current study was to investigate whether differences in brain morphology might explain individual differences in children's ability to recognize emotions through prosodic cues. A sample of 66 children ($M = 8.30$ years; $SD = 0.35$) completed both a behavioural task and a structural magnetic resonance imaging scan. In the behavioural task, children listened to semantically neutral sentences and had to perform two consecutive judgments for each stimulus, including a forced-choice categorization of the emotional tone (neutral, happy, sad, angry, scared) and an intensity judgment, rating the salience of the emotion in the stimulus. Results revealed that children achieved high recognition accuracy rates for all emotional categories, and happiness was the best recognized emotion. Besides, in terms of neural structures, there were correlations between higher emotional prosody recognition and increased grey matter volume in the fusiform gyrus, cerebellum, motor/premotor and prefrontal regions; and decreased grey matter volume in parietal and occipital regions. Additionally, we found that some brain regions were correlated with higher recognition accuracy of specific emotions, when directly compared to the other ones. Our findings suggest that the individual differences in children's ability to recognize emotions through prosodic cues relate to differences in brain morphology, both for the general emotional prosody recognition ability and for the recognition of specific emotional categories.

Keywords: emotion; speech prosody; emotional recognition; children; neuroimaging; voxel-based morphometry.

Résumé

La capacité à inférer les états émotionnels des autres à partir des variations de la prosodie a une valeur adaptative importante et est essentielle pour la adaptation personnelle et sociale. Cette capacité émerge très tôt dans le développement ontogénétique et est associée à l'amélioration des compétences socio-affectives des enfants. Cependant, les fondements cérébraux de la reconnaissance de la prosodie émotionnelle chez les enfants restent mal étudiés. L'objectif principal de cette étude est de déterminer si la morphologie du cerveau pourrait expliquer les différences individuelles dans les aptitudes des enfants à reconnaître les émotions par des indices prosodiques. Un échantillon de 66 enfants ($M = 30.8$ ans, $SD = 00.35$) a complété une tâche comportementale et une IRM structurelle. Au cours de la tâche comportementale, les enfants ont écouté des phrases neutres en termes sémantiques et ils ont fait deux jugements consécutifs: un classement en choix forcé de la tonalité émotionnelle (neutre, heureux, triste, en colère, peur), et un jugement d'intensité de l'émotion perçue. Les résultats ont montré que les enfants avaient des taux élevés d'exactitude dans la reconnaissance de toutes les catégories émotionnelles et que la joie était l'émotion la mieux reconnue. En outre, en termes de structures cérébrales, une meilleure corrélation a été trouvée entre la reconnaissance des émotions prosodiques et un volume plus grand de matière grise dans le gyrus fusiforme, cervelet et dans des régions motrices / pré motrices et préfrontales; et un plus petit volume de matière grise dans les régions occipitale et pariétale. De même, nous avons constaté que certaines régions du cerveau étaient associées avec une meilleure exactitude dans la reconnaissance de certaines émotions par rapport aux autres. Nos résultats suggèrent donc que les différences individuelles dans la capacité des enfants à reconnaître les émotions par des indices prosodiques reflètent des diversités dans la morphologie du cerveau, à la fois par la reconnaissance de la prosodie émotionnelle en général et par la reconnaissance de catégories émotionnelles spécifiques.

Mots-clés: émotion; prosodie de la parole; reconnaissance des émotions; enfants; neuro-imagerie; *voxel-based morphometry*.

Table of Contents

Introduction	1
1. Theoretical Perspectives on Emotions.....	1
2. Communication of Emotions.....	3
3. Emotional Prosody Recognition.....	4
4. Neurocognitive Basis of Emotional Prosody Recognition.....	5
5. Individual Differences	7
6. The Current Study	8
Method.....	10
1. Participants	10
2. Experimental Task.....	11
1. Materials	11
2. Procedure	11
3. MRI Acquisition and Data Analysis.....	14
1. Image acquisition.....	14
2. Preprocessing.....	14
3. Image Analyses	15
Results.....	16
1. Behavioural Data	16
2. Neuroanatomical Data	18
Discussion	23
Conclusion	28
References.....	30
Appendices	39
Appendix A. <i>Statistics of normality test (Shapiro-Wilk) and values of kurtosis and skewness</i>	40
Appendix B. <i>Illustration of neuroanatomical results</i>	41

List of Tables and Figures

Tables

Table 1. <i>Participants' demographic and background characteristics</i>	11
Table 2. <i>Acoustic characteristics of the stimuli</i>	12
Table 3. <i>Distribution of categorization rates for each emotion category</i>	18
Table 4. <i>Brain regions associated with emotional prosody recognition ability in a 7 to 9 years old children sample (separate-modality regressions)</i>	19
Table 5. <i>Anatomical brain regions associated with emotional prosody recognition ability in a 7 to 9 years old children sample (adjusted for variability related to the other emotional categories)</i>	21
Table 6. <i>Comparisons between anatomical brain regions associated with emotional prosody recognition ability in a 7 to 9 years old children sample</i>	22

Figures

Figure 1. <i>Schematic representation of a trial</i>	13
--	----

Introduction

Emotions have a crucial role both in social interactions and in personal adjustment to the constant changes in external environment. In the last decades, researchers have been interested in studying different aspects of emotions, considering their important adaptive role in everyday life (e.g., Lima, Alves, Scott, & Castro, 2014; Saarimäki et al., 2015). Although a consensual definition of emotion is still lacking, several authors agree that it might be considered a complex psychological state that involves at least three main components, including a subjective experience, a physiological reaction and a behavioural response (e.g., Hockenbury & Hockenbury, 2007). Even though most studies on emotions tend to include adult samples, from an ontogenetic perspective, children were shown to be able to express and recognize emotional signals from early ages (e.g., Sauter, Panattoni, & Happé, 2013). Children's knowledge about emotions might, in fact, affect their attitudes and behaviours and, consequently, the development of their social interactions. Because our knowledge of emotion recognition abilities is mostly based on studies with adult samples, further studies are required to improve our understanding of these processes in children, both at behavioural and neural levels.

1. Theoretical Perspectives on Emotions

Numerous theories and models have been proposed regarding emotion processes. The categorical and dimensional theories are the most influential ones in Psychology (Fujimura, Matsuda, Katahira, Okada, & Okanoya, 2012).

Some researchers agree that all emotional experiences are based on a limited number of emotions, which are distinct and independent from each other (e.g., happiness, sadness, anger, fear). According to this perspective, people experience emotions as discrete categories. Furthermore, these theories assume that each emotion category relies on a discrete neural system (Colibazzi et al., 2010) and has a specific adaptive value, i.e., the experience of emotions is mediated by distinct neural structures and evoke specific autonomic and behavioural response patterns, as well as facial and vocal expressions (Scherer, 2003). Besides, according to this perspective, emotions communicate relevant

information to the self and others, and influence and motivate both our own and others' actions (Kunzmann, Kappes, & Wrosch, 2014).

Most studies on emotion prosody recognition have adopted a categorical perspective: researchers often analyse the ability of different populations to identify a range of specific emotions as expressed by one or more speakers, while focusing on their tone of voice (e.g., Lima et al., 2014; Sauter et al., 2013).

Similarly to other perspectives, discrete emotion theories have some limitations, such as the fact that most studies focus on facial expressions, which may not be a reliable measure of emotionality, since the same facial expression is not always associated with one specific emotion (e.g., the smile might indicate happiness, but also fear or discomfort; Colibazzi et al., 2010). Also, some researchers have argued that emotion categories are not confined to one brain region, but rather are associated to multiple inter-connected brain systems (Wager, Kang, Johnson, Nichols, Satpute, & Barrett, 2015). Thus, even though some level of specificity has been shown regarding behavioural and expressive patterns, linking discrete emotions to specific neural systems remains challenging. Yet, categorical approaches are highly prevalent in emotion research, having considerably contributed to our knowledge of emotion processes.

A distinct perspective advocates that emotions can be identified along a small number of more general affective dimensions (Briesemeistera, Kuchinke, & Jacobs, 2014). In other words, according to these theories, dimensions can be understood as distinct crossed continuous forming a sort of geometrical space (Gillioz, Fontaine, Soriano, & Scherer, 2016), where all emotions can be included. Although several dimensions have been proposed, factor analyses tend to bring out two main dimensions, namely *valence* and *arousal*. While valence ranges from extremely negative/unpleasant emotions (e.g., despair) to highly positive/pleasant ones (e.g., happiness), arousal fluctuates between deactivation and high activation (Colibazzi et al., 2010). Some studies also emphasize a third dimension, which usually represents power or control/dominance (Gillioz et al., 2016; Scherer, 2003).

Nowadays, researchers attempt to study differences of valence and arousal in the emotional states, inferred from vocal or visual stimuli. In both cases, the main focus of these theories is on the subjective experience of the individuals.

Despite its importance to advances in the field, these theories also show some limitations. Some researches, for instance, claim that a limited number of dimensions may

not capture all the existing differences between emotional experiences (Lerner, Valdesolo, & Kassam, 2015).

The question of which perspective might better explain the experience of emotions is still debated. More recently, new theories have been developed, such as componential models of emotion (e.g., Scherer, 2003), which emphasize the variability of distinct emotional states. Research in this area has been attempting to integrate different models of emotion, in order to contribute to a better understanding of the emotional experience and to, thereby, develop a more comprehensive model.

In the current study, we focus on the recognition of discrete emotion categories, including both behavioural and structural neuroimaging techniques.

2. Communication of Emotions

The ability to express and understand emotions is crucial to human interactions, considering their influence in our thoughts, feelings and actions (e.g., Darwin, 1872). Based on others' emotional expressions, we are able to make inferences about their intentions, feelings and states of mind (Sauter et al., 2013), and adapt our actions accordingly.

Understanding the processes underlying emotional communication is thus a topic of central importance. It is known that people express emotions in several different ways, including visual cues, as facial or body expressions, and auditory cues, like semantically emotional speech, nonverbal vocalizations, prosodic variations and music.

In addition to being the carrier of speech, the human voice is also crucial for non-verbal emotional communication. Nonverbal vocal emotional cues can include purely nonverbal vocalizations, such as laughter or crying, and variations in the 'tone of voice' as we speak, i.e., emotional speech prosody. Emotional speech prosody refers to variations in several acoustic cues, namely fundamental frequency (f_0), intensity, tempo, rhythm and voice quality (Gil, Hattouti, & Laval, 2016). From a production perspective, it is known that emotional states can have an effect on these cues (Banse & Scherer, 1996). From a perceptual perspective, listeners can use this information to infer the emotional state of the speaker. Variations in this suprasegmental information are associated with changes in people emotional states. In other words, the way people say things is a relevant source of information about their emotional state (Witteman, Heuven, & Schiller, 2012).

3. Emotional Prosody Recognition

The ability to understand other people emotional states from variations in speech prosody cues (Alba-Ferrara, Hausmann, Mitchell, & Weis, 2011) has an adaptive value and is crucial for personal and social adjustment (Gil et al., 2016). Some studies have shown that the ability to recognize emotions in prosody emerges early in development. For instance, Sauter and colleagues (2013) have shown that five year old children are already able to interpret emotional information conveyed via prosodic cues, indicating that they are able to use and combine different prosodic cues to distinguish emotional states. In this study, children aged five to ten years old performed an emotional recognition task, which included ten categories of verbal and non-verbal stimuli, including positive (amusement, contentment, relief, and achievement), neutral (surprise and neutral) and negative (anger, disgust, sadness and fear) categories. Results revealed an average accuracy ranging from 78% for the younger children to 84% for the older ones, indicating that emotion recognition from vocal cues improves throughout childhood.

Emotional prosody recognition is related with other important socioemotional skills, such as the development of emotional theory of mind in children (Chakrabarti & Baron-Cohen, 2006; Mitchell & Phillips, 2015). Theory of mind is, indeed, particularly useful for children when interpreting emotional states, as it implies being able to understand that other people might have thoughts, feelings and intentions different from their own (Alba-Ferrara et al., 2011). Also, a study evaluating emotion recognition abilities of children aged 8 to 10 years old found a negative correlation between the ability to interpret other people's vocal emotional cues and several aspects of social anxiety (e.g., social avoidance and distress), indicating that social anxiety may be related with an increased difficulty to interpret social information (McClure & Nowicki, 2001).

When interacting with others, children, like adults, make use of acoustic features of other people's voices to understand their intentions and feelings (Sauter et al., 2013). Prosodic cues may also provide relevant information concerning the caregivers' interest and intentions, which affects children's experience of feeling loved and desired, from very early ages, and therefore it influences their feelings as well. Moreover, children's behaviours are guided by the interpretation they make of adults' emotional states (e.g., Gil et al., 2016).

Although some studies have examined the processing of emotional speech prosody in children, this remains a poorly investigated topic. This is so because most research on emotion recognition has emphasized visual stimuli, mainly facial expressions; and the studies that examine the processing of vocal cues are typically conducted with adults. This is the case both of behavioural and of neuroimaging studies. Therefore, the current study might contribute to a better understanding of this topic.

4. Neurocognitive Basis of Emotional Prosody Recognition

A body of studies have looked at the neural basis of emotional speech prosody processing (e.g., Alba-Ferrara et al., 2011; Baum & Pell, 1999; Wildgruber, Ackermann, Klein, Riecker, & Grodd, 2000). The neural processes underlying our ability to decode emotional states through speech prosody have arguably been shaped by evolutionary processes. Since ancient times, the ability to express and interpret emotions has impelled people to develop their communication skills and, consequently, their social interactions. As a result, basic neural processes underlying speech prosody recognition evolved and become as complex as we know today (Damasio & Carvalho, 2013).

Research on emotional prosody recognition has been influenced by two relatively similar multi-stage models, which are intended to explain the neural processes involved in this ability (Witteman et al., 2012). Both models suggest that the emotional prosody recognition occurs along multiple successive processing stages. According to Wildgruber and colleagues (Brück, Kreifelts, & Wildgruber, 2011a; Wildgruber, Ethofer, Grandjean, & Kreifelts, 2009), in the initial stage, the auditory cortex, particularly the mid-superior temporal gyrus, is involved in the detection of acoustic cues from vocal stimuli; in a second phase, acoustic information is integrated to enable a global understanding of the emotional state of the speaker, which would be supported by the posterior superior temporal cortex; finally, in a third stage, emotional prosody is explicitly understood and judged, for which inferior frontal cortex and orbitofrontal cortex play a major contribution. According to this model, emotional signals may also induce automatic emotional reactions, which are supported by specific subcortical regions, as the amygdala.

The second model (Kotz & Paulmann, 2011; Schirmer & Kotz, 2006) also proposed the existence of three main stages. It agrees with the involvement of primary and secondary auditory cortex in the extraction of the acoustic properties of vocal signals,

although not relating it directly to mid-superior temporal gyrus; it assumes that the second stage includes the integration of acoustic cues, yet, according to this model, anterior superior temporal sulcus is implicated at the end of this stage; lastly, it is suggested that, in the third stage, right inferior frontal gyrus and orbitofrontal cortex, as well as subcortical structures (e.g., amygdala and basal ganglia) are involved in the explicit judgment of emotional prosody, and left inferior frontal gyrus is implicated in its integration into language processing.

Supporting these models, several neuroimaging studies have indeed shown that emotional prosody recognition is supported by a temporofrontal neural network (Alba-Ferrara et al., 2011), particularly the bilateral superior temporal cortex, and the anterior and middle superior temporal sulcus (e.g., Belin, Fecteau, & Bédard, 2004). Subcortical structures, particularly the basal ganglia and amygdala, have also been associated with emotional prosody ability (Pell & Leonard, 2003; Wiethoff, Wildgruber, Grodd, & Ethofer, 2009; Wildgruber, Riecker, Hertrich, Grodd, Ethofer, & Ackermann, 2005). Nonetheless, these models were derived from studies based on samples of adults. Therefore, the extent to which their results can be transferred to different age groups, as older adults or children, remains unclear.

Only a few studies have looked at the brain mechanisms underlying emotional speech prosody processing in children, and the existing ones used functional magnetic resonance imaging (fMRI), that is, they focused on functional responses during speech prosody processing. Although, in the current study, we focus on structure rather than on functional responses, those studies might contribute to a better understanding of the brain morphology associated with speech prosody recognition ability. Plante, Holland, and Schmithorst (2006) used a linguistic prosody task in fMRI, where children aged 5 to 18 years old listened to low-pass filtered sentences and detected those that corresponded to a target sentence. Children showed activation in frontal regions, namely the inferior and middle frontal gyrus, in bilateral insular cortex, precentral sulcus, superior temporal gyrus and anterior occipital cortex. In a similar study with a sample of children and young adults (aged 7 to 30 years old), where participants had to indicate whether each sentence they heard was a statement or a question, participants revealed activation in several frontotemporal areas, including right superior temporal gyrus, and right inferior and middle frontal gyrus (Vannest, Karunanayaka, Schmithorst, Szaflarski, & Holland, 2009).

Several studies have also suggested that there is a major involvement of right hemisphere structures in emotional prosody recognition (e.g., Adolphs, Damasio, &

Tranel, 2002; Dara, Bang, Gottesman, & Hillis, 2014; Witteman et al., 2012), although this remains a controversial issue. Children, on the other hand, reveal decreased lateralization, in general, with the exception of some brain regions, as middle frontal gyrus, which is lateralized to the right (Plante et al., 2006). Besides, some authors suggest that right hemispheric dominance is more frequently found when prosodic stimuli are non-linguistic (e.g., vocalizations; Wartenburger, Steinbrink, Telkemeyer, Friedrich, Friederici, & Obrig, 2007).

Although emotions are frequently experienced as multiple distinct feelings, the neuroanatomical basis of specific emotions is still not well understood, and researchers still aim to know whether specific brain regions are more associated with the processing of some emotions compared to others (Kragel & LaBar, 2016). Previous fMRI studies have shown that the mid-superior temporal gyrus shows increased activation to several vocal emotions when compared to neutral prosody, which is in accordance with the initial phase of the multi-stage model described above (Brück et al., 2011a; Wildgruber et al., 2009). Yet, since this increased activation is similar for all emotions, it was not possible to identify differences between discrete emotional categories (e.g., Ethofer, Van De Ville, Scherer, & Vuilleumier, 2009). Therefore, these authors suggest that the categorization of emotions might occur at posterior phases of the speech prosody processing, even though most studies fail to support the hypothesis of the existence of brain regions specifically associated with the recognition of distinct emotional categories through prosody (e.g., Kotz, Meyer, Alter, Besson, von Cramon, & Friederici, 2003; Wildgruber et al., 2005). However, research on this topic remains scant and further investigation might come to show different results. The present study will focus both the neuroanatomical basis of emotional prosody recognition, in general, and concerning specific emotional categories as well.

5. Individual Differences

The vast majority of studies on emotional speech prosody are based on group-level analysis, as researchers are trying to infer general processes across participants. Therefore, as they are not interpreted in a meaningful way, differences between individuals tend to fade among average values, and are treated as ‘noise’ (Kanai & Rees, 2011). However, it is

now well established that differences in brain morphology are often a source of individual variability in how individuals perform several tasks (e.g., Kragel & LaBar, 2016).

Although several studies have found that inter-individual differences in several aspects may be systematically related to differences in individuals' brain morphology (e.g., Banissy, Kanai, Walsh, & Rees, 2012; Kanai, Bahrami, Roylance, & Rees, 2011; Kanai, Feilden, Firth, & Rees, 2011; Lima et al., 2015), only few studies have focused on inter-individual variations in the processing of emotional voice cues (Brück, Kreifelts, Kaza, Lotze, & Wildgruber, 2011b). Consequently, we know very little about how people might differ in their ability to process speech prosody, and about how this potentially relates to differences in brain structure and function. However, the few available studies already provide some important information concerning this topic. For instance, a study by Brück and colleagues (2011b) has shown a significant influence of neuroticism on brain activation during the performance of an emotional prosody recognition task, indicating that brain mechanisms underlying emotional speech prosody recognition may be related to differences in personality.

6. The Current Study

The goal of the current study is to examine whether differences in brain morphology might contribute to explain individual differences in the ability to recognize emotions through prosodic cues, in a sample of children aged 7 to 9 years old. Each child performed a behavioural task and a structural MRI scan. The experimental task assessed children's emotional prosody recognition ability, and five emotional categories were included: happiness, sadness, anger, fear, and neutrality. Neuroimaging data were analysed using voxel-based morphometry (VBM), a technique that can be used to compare different brains on a voxel-by-voxel basis, and identify differences in the brain morphology, particularly in the amount of grey matter. VBM is a widely used technique (e.g., Ibarretxe-Bilbao, Junque, Tolosa, Marti, Valldeoriola, Bargallo, & Zarei, 2009; Rohrer, Sauter, Scott, Rossor, & Warren, 2012; Sato, Kochiyama, Uono, Kubota, Sawada, Yoshimura, & Toichi, 2015).

Behaviourally, based on previous research, we expected these children to be able to recognize prosodic emotions with accuracy levels above chance (e.g., Sauter et al., 2013). However, there might be differences in accuracy across emotions, so we expected that

children would be better at recognizing happiness and anger (e.g., Chronaki, Hadwin, Garner, Maurage, Sonuga-Barke, 2015; Oerlemans et al., 2014). In terms of neural structures, functional studies have mainly implicated frontal and temporal areas so we were expecting that structural differences in some of those areas could potentially explain variability in children's emotional recognition accuracy. However, the existing literature is based on fMRI and samples of adult participants, and so a direct link to structural analyses and development is not warranted. Our approach was thus exploratory to a certain extent.

Method

1. Participants

Eighty-eight children (forty-seven girls), from five Portuguese elementary schools (“Agrupamento de Escolas do Concelho de Matosinhos”) were initially recruited. We then excluded those who could not complete the structural MRI scan, or who had a full-scale IQ < 70 (IQ below 70 indicates intellectual disability, according to the 10th version of the International Classification of Diseases, ICD-10; World Health Organization, 1992). One additional participant was excluded due to abnormally low performance on the emotional prosody recognition task (two standard deviations below the global mean of the sample). The final sample included sixty-six children (thirty-seven girls). They were aged between 7 and 9 years old ($M = 8.30$, $SD = 0.35$) and were all attending the second grade. Three of them were left-handed and four were ambidextrous; the remaining ones were right-handed. Based on a report completed by children’s parents, all participants had normal hearing and were native speakers of European Portuguese (see Table 1 for details).

All children were tested on IQ, using the Portuguese version of Wechsler Intelligence Scale for Children – 3rd Edition (Simões, Rocha, & Ferreira, 2003; Wechsler, 1991). This is an achievement test, which determines children’s potential for learning and their current level of cognitive functioning. In this study, it was used as an index of general intellectual functioning. The full scale IQ was administered, which includes six subtests contributing to the Verbal IQ score (Information, Similarities, Arithmetic, Vocabulary, Comprehension and Digit Span), and six subtests contributing to the Performance IQ (Picture Completion, Coding, Picture Arrangement, Block Design, Object Assembly and Symbol Search). Children were also tested on their reading ability, mathematical skills, musical aptitude, executive functioning, spatial orientation and motor performance, as part of an on-going research project¹ which has received approval from the Ethics Committee of Faculty of Psychology and Educational Sciences (University of Porto), as well as from Schools’ Administration. Local school authorities, as well as children’s parents, signed a

¹ “Project Impact”, a project funded by BIAL Foundation. The goal of this project is to examine the impact of music and sports’ training on reading and mathematical abilities, as well as the modulation of the putative transfer effects by executive functioning.

consent form before the beginning of testing. The parents have also completed a safety form to ensure that the children could be scanned.

Table 1. Participants' demographic and background characteristics.

Characteristics		Range	Maximum Possible Range
N	66	–	–
Age (years)	8.30 (0.35)	7.75 – 9.5	–
Gender	29M / 37F	–	–
Mother Qualification*	1-4; 2-28; 3-18; 4-16	1 – 4	1 – 4
Full Scale IQ	95.56 (13.26)	74 – 125	40 – 160
Verbal IQ	96.52 (12.86)	69 – 124	46 – 155
Performance IQ	97.74 (13.78)	70 – 132	46 – 155

Note. *SD* values in parentheses. M – male; F – female. *1 – Elementary school; 2 – Middle School; 3 – High School; 4 – University

2. Experimental Task

2.1. Materials

A computerized task was used to measure children's ability to identify emotions from spoken utterances while focusing on the tone of voice of the speakers. The auditory stimuli were taken from a previously validated set of stimuli (Castro & Lima, 2010). They consisted of 50 short sentences (10 stimuli per emotion; mean duration = 1472 msec, *SD* = 247) with emotionally neutral semantic content (e.g., “O futebol é um desporto”, Football is a sport), as produced by two female speakers in five emotional tones – anger, fear, happiness, sadness and neutrality. The final set of stimuli included in this study was selected so that (a) half of the stimuli were produced by each speaker; (b) the mean accuracy, perceived intensity and duration were as similar as possible across emotion categories (e.g., Lima et al., 2014); and (c) “ceiling” and “floor” effects were avoided, that is, stimuli that were either associated with very high or very low recognition accuracy were not included. The characteristics of auditory stimuli are described in Table 2.

2.2. Procedure

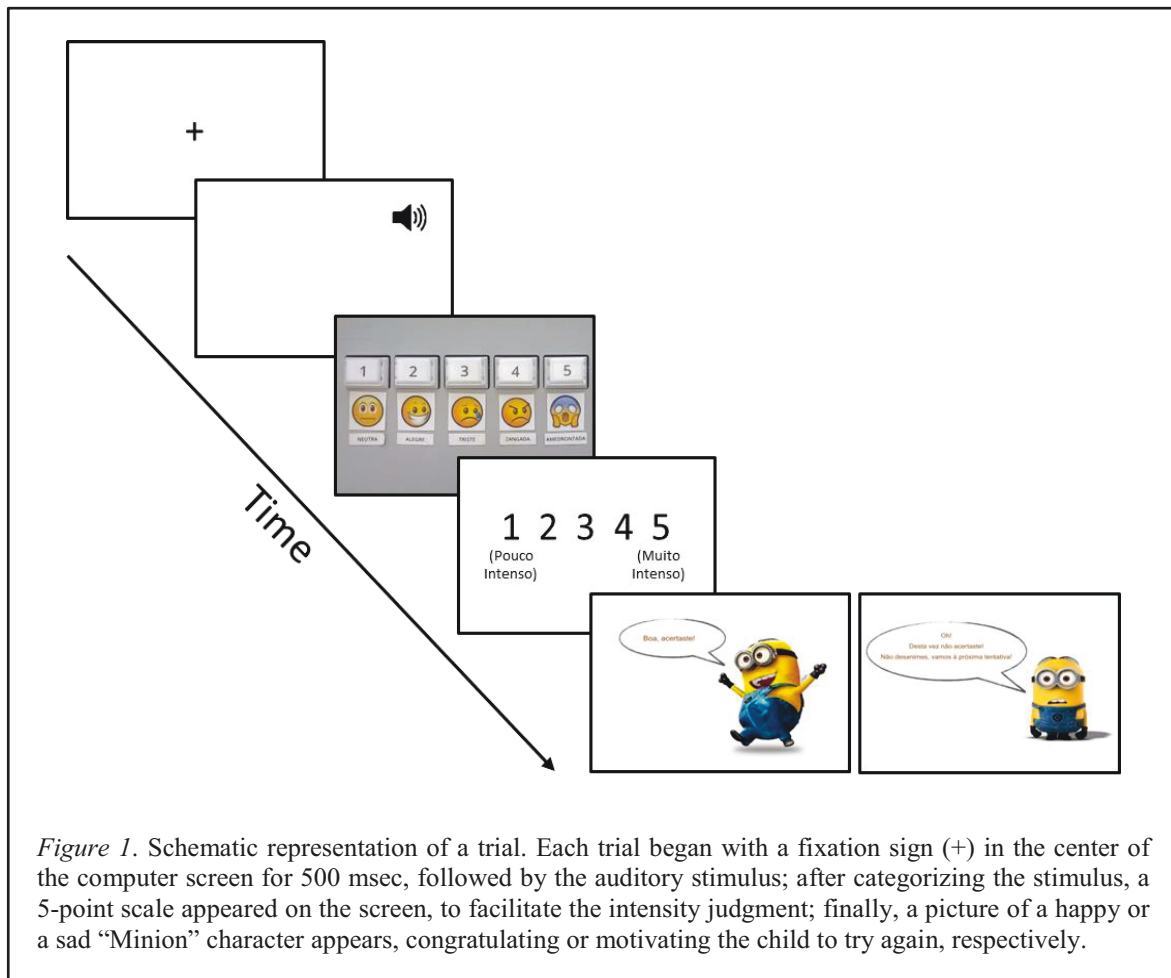
Table 2. Acoustic characteristics of the stimuli.

Emotion	f_0 (Hz)	f_0 Variation (Hz)	Duration (msec)	Accuracy (%)	Intensity (1-5)
Neutrality	213	37	1511	89	4.8
Happiness	359	86	1550	83	5.4
Sadness	200	33	1520	89	4.9
Anger	315	73	1269	84	5.0
Fear	309	41	1512	73	5.2
<i>Mean total</i>	<i>279</i>	<i>27</i>	<i>1472</i>	<i>83</i>	<i>5.1</i>

Note. Accuracy and intensity values were obtained from an adult sample, and from a task including six emotions and a neutral condition (mean age = 21.8, $SD = 6.1$; Castro & Lima, 2010).

Children were tested individually in one single session lasting about twenty minutes. The stimuli (10 trials per emotion) were randomized and divided into two blocks of 25 trials each. There was a block of 3 practice trials to allow participants to familiarize themselves with the task. This phase was also important to ensure that the volume of the stimuli was comfortable for each participant. Participants were told they would listen to short sentences that were neutral regarding semantic content and that they should focus on the tone of voice. They were instructed to perform two consecutive judgments for each sentence: a forced-choice identification of the emotional tone (neutral, happy, sad, angry, scared) and an intensity judgment rating how much the expression was present in the stimulus (Castro & Lima, 2010; Lima & Castro, 2011), on a scale from 1, low intensity, to 5, high intensity. The stimuli were presented via high quality headphones (Sennheiser HD 202), in a quiet room on children's school campus. Responses were recorded on a Macbook Air running SuperLab V5.0.3 (Heller, Matsak, Abboud, Schultz, & Zeitlin, 2014), through a seven-button response pad (model RB-740 from Cedrus Corporation). Five emoticons illustrating each of the emotional categories, as well as the corresponding labels (*neutra*, neutral; *alegre*, happy; *triste*, sad; *zangada*, angry; *amedrontada*, afraid) were placed below the buttons, while numbers for the intensity judgment (1 to 5) were placed on the buttons. Each trial began with a fixation sign (+) in the centre of the computer screen for 500 msec, preceding the auditory stimulus. Once the participant had performed the first judgement concerning the emotional tone of the speaker, the 5-point scale appeared on the screen, so the child would perform the intensity judgment. After that, in order to promote the children's motivation and attention levels

during the task, a picture of a happy or a sad “Minion” character was presented, congratulating or motivating the child to try again depending on whether the response had been correct or incorrect, respectively (see Figure 1). Between the two blocks of trials, participants were allowed to take a break, if they wished. They could also take short breaks during the experiment whenever they needed. After the experiment, children' participation was acknowledged and praised.



The stimuli and procedure used in the experimental task were generally similar to those used in previous studies with adult samples (Castro & Lima, 2010; Lima & Castro 2011). However, considering that the age of participants in the current study could increase the difficulty of the task, some changes were made. First, we included only five emotional categories (three negative, one positive and neutrality), instead of the seven categories used before, since a forced-choice paradigm with seven alternatives could be too difficult and generate floor effects. Second, the number of stimuli was reduced to fifty (ten stimuli per emotion), to avoid weariness. Third, in order to maintain participants’ interest during the

task, a feedback character was introduced after each stimulus, which was well received for all children. Although this might be responsible for a learning effect, accuracy was not substantially better for the second as compared to the first set of stimuli (70% and 63%, respectively). The accuracy on emotional prosody was quantified through the total number of correct trials and the number of correct trials per emotion. Intensity rates for the correctly identified stimulus were also included, both the average mean assessment and the mean assessment per emotion.

3. MRI Acquisition and Data Analysis

3.1. Image acquisition

High-resolution whole-brain images were obtained using a T1-weighted magnetization prepared rapid acquisition gradient echo sequence for each participant on a 1.5T Siemens Magnetom Sonata Maestro Class (repetition time = 1680 ms; echo time = 4.12 ms; flip angle = 8°; slice thickness = 1 mm; 160 sagittal slices; acquisition matrix = 250 x 250 x 160 mm, voxel size = 1 mm³).

3.2. Preprocessing

The structural images were preprocessed and analysed using SPM12 (Wellcome Trust Centre for Neuroimaging, UK; <http://www.fil.ion.ucl.ac.uk/spm/software/spm12/>), running within the Matlab R2015b (Corporate Headquarters, Natick, USA), and using the CAT12 (Computational Anatomy Toolbox) toolbox (<http://dbm.neuro.uni-jena.de/cat/>). This toolbox is an extension to SPM12 for computational anatomy. It was developed by Christian Gaser and Robert Dahnke (Jena University Hospital, Departments of Psychiatry and Neurology) in 2016, and it covers morphometric methods, including VBM. Customized pediatric tissue probability maps were generated using the TOM8 (Template-O-Matic) toolbox (Wilke, Holland, Altaye, & Gaser, 2008; <http://dbm.neuro.uni-jena.de/software/tom/>). This was important given that our sample consisted of children, and there are differences in brain structure between children and adults, including in cortical thickness, cortical surface area and cortical folding; this toolbox is thus particularly useful to minimize the potential confounds that might be introduced when using an adult template (Beal, Gracco, Brettschneider, Kroll, & De Nil, 2013; Yoon, Fonov, Perusse, & Evans, 2009). Therefore, a custom template was created according to the average age of

the sample under study. Afterwards, T1 images were spatially normalized into the same template and then segmented into grey matter, white matter, and cerebrospinal fluid. Diffeomorphic Anatomical Registration was carried out using Exponentiated Lie algebra (DARTEL method), meaning that the images were matched to a template generated from their own mean, resulting in sample-specific grey and white matter templates (Ashburner, 2007). All scans were submitted to a quality assessment, involving a visual inspection for major artefacts. Noise values were confirmed to be within acceptable boundaries for all images (< 3) and sample quality and homogeneity was also checked through CAT12 toolbox. Finally, images were smoothed with an 8 mm full-width-at-half-maximum filter (FWHM).

3.3. Image Analyses

Multiple regressions were used to analyse voxel-wise associations between grey matter volume and emotion recognition accuracy. These associations were assessed both in separate design matrices for each emotional category and for the mean total accuracy (separate-analyses), and in a combined regression matrix, i.e., including all emotional categories in a single model (combined-analysis). While the first type of analyses allows us to examine associations between emotional prosody recognition and grey matter volume for each emotion category separately (and for global emotion recognition accuracy), the second type of analyses allows us to make direct comparisons between emotion categories, and to examine unique associations between each emotion and grey matter, i.e. after controlling for variance related to the remaining emotions (for a similar approach, see Omar et al., 2011). Full scale IQ, age, sex, mother education level and total intracranial volume (TIV) were used as covariates of no interest in all the regression models, to regress out any potential confounding effects related to them. TIV is an integrated measure that combines the total volume of grey matter, white matter, and cerebrospinal fluid. It was estimated using CAT12 toolbox. The education level of children's mother was taken as an index of socioeconomic status (e.g., Erola, Jalonen, & Lehti, 2016). To avoid potential edge effects on the border areas between grey and white matter, an absolute masking threshold was used, whereby voxels with grey matter intensity below 0.05 were excluded (e.g., Lima et al., 2015).

Once the statistical model had been estimated, T-contrasts were used to test for positive and negative associations between the accuracy on emotional prosody recognition and the grey matter volume at the whole-brain level. Positive associations indicate a link

between higher emotion recognition performance and larger grey matter volume, while negative associations indicate a link between higher emotion recognition performance and smaller grey matter volume. Statistical maps were thresholded at $p < .001$, peak-level uncorrected, with a cluster extent of 30 voxels. Therefore, only clusters with more than 30 voxels were considered, and those that also survived family-wise error (FWE) correction for multiple comparisons at whole-brain level ($p < .05$) were highlighted, as it can be seen in Table 4.

Results

1. Behavioural Data

In order to examine emotion recognition performance, we considered the number of correct trials for each emotion category (i.e., number of prosodic expressions that were correctly categorized), and the number of correct rejections (i.e., the number of times a given category was not used when the prosodic expressions conveyed a different emotion). Values in diagonal cells in Table 3 indicate the percentage of correct responses for each emotion category. The average recognition accuracy was 67% ($SD = 12$), and values ranged between 87% for happy expressions and 57% for fearful expressions. Thus, all emotion categories were recognized well above the chance-level, 20%. The last row of the table indicates the average values of correct rejection per emotion, which ranged from 87% for sad trials to 97% for anger trials. Recognition accuracy rates were arcsine transformed² and submitted to a repeated-measures analysis of variance (ANOVA), including emotional category as a within-subject factor. The Tukey HSD test was used for post-hoc comparisons, allowing us to directly compare participants' accuracy across emotion categories. A significant main effect of emotion was obtained [$F(4,260) = 18.13, p < .001, \eta^2 = 0.22$], indicating that emotions differed regarding how well they are recognized. Participants were better at recognizing happy intonations ($M = 86.52, SD = 12.46, ps < .001$) than all the other ones. Additionally, the accuracy on sad trials ($M = 68.94, SD = 25.31$) was statistically higher when compared to the fearful ones ($M = 56.52, SD = 23.17; p = .03$).

The pattern of inaccurate responses was also briefly analysed (Table 3, rows). Anger was frequently confused with happiness, possibly because they are both characterized by high f_0 variation (73 Hz and 86 Hz, respectively; see Table 2). Besides, sadness and neutrality were often confused with each other, which might be associated with the fact that they both have similarly low f_0 (200 Hz and 213 Hz, respectively) and f_0 variation (33 Hz and 37 Hz). Fear was almost equally confused with happiness, neutrality and sadness. Except for anger, all emotional categories had similar duration (above 1500

² Arcsine transformation is often used to improve data's normal distribution (e.g., Haller, Raeder, Scerif, Kadosh, Lau, 2016; Lima et al., 2014); see Table A1 for a detailed description of the values of kurtosis and skewness before and after the arcsine transformation.

msec). This might be the reason why these emotional categories are rarely confused with anger, as the mean duration of anger expressions is below 1300 msec (detailed values per emotion are shown on the right column of Table 2).

An ANOVA was also performed on intensity judgments, including emotional category as a within-subject factor. Using the Tukey HSD test for post-hoc comparisons, we found a significant main effect of emotion [$F(4,236) = 10.73, p < .001, \eta^2 = 0.15$], demonstrating that intensity ratings varied across emotional categories. Participants attributed higher intensity rates to happiness ($M = 3.81, SD = 1.17$), anger ($M = 3.81, SD = 1.12$), and fear ($M = 3.65, SD = 1.27$), and lower rates to sadness ($M = 3.40, SD = 1.21$) and neutrality ($M = 3.15, SD = 1.32$). It is worth emphasizing that emotion recognition accuracy and intensity ratings were not significantly related ($r = -0.67, p = .59$), suggesting that these two judgments reflect independent processes.

Forced-choice paradigms, as the one used in the current study, may be associated with biases in participants' responses: inter-individual differences in accuracy rates might represent differences in the tendency for participants to choose some categories more frequently than others. To control this methodological issue, the data was reanalysed using the unbiased hit rate H_u (Wagner, 1993; e.g., Griffiths, Jarrold, Penton-Voak, & Munafò, 2015; Lima & Castro, 2011), a measure of perceptual sensitivity that considers both the proportion of trials in which a particular emotion is properly identified (hit rate) and the amount of times that a certain emotion is incorrectly selected (false alarms)³. Unbiased hit rates are shown in the right column of Table 3. Although corrected accuracy rates were noticeably lower than the uncorrected ones, they are still well above the chance level. An ANOVA conducted on these values revealed a similarly significant main effect of emotion [$F(4,260) = 19.10, p < .001, \eta^2 = 0.23$]. However, while happiness remains the best recognized emotion category ($p < .002$), anger was better recognized than neutrality ($p < .001$) and sadness ($p = .01$), and fear showed better results than neutrality ($p = .01$).

As it can be seen in Table 3, standard deviations of recognition accuracy are large for all emotion categories, indicating that there are large individual differences in participants' emotional recognition ability. This is an important point, given that our neuroanatomical analyses ask the question of whether individual differences in behavioural performance can be related to differences in brain morphology.

³ H_u was calculated for each emotional category and participant through the formula $H_u = A^2 / (B \times C)$, where A = number of properly identified trials, B = number of trials per category, and C = number of responses where that category was selected (including both hit rates and false alarms).

Table 3. Distribution of categorization rates for each emotion category (percentages).

Intended Emotion	N	Participants responses					Hu*
		Neutral	Happy	Sad	Angry	Fearful	
Neutrality	10	61.36 (27.28)	3.94	28.18	2.12	4.39	0.38
Happiness	10	6.97	86.52 (12.46)	1.97	2.27	2.27	0.63
Sadness	10	20.61	1.52	68.94 (25.31)	4.09	4.85	0.41
Anger	10	7.27	28.03	1.97	58.94 (27.35)	3.79	0.51
Fear	10	10.15	10.91	19.40	3.03	56.52 (23.17)	0.48
<i>Correct rejection rate</i>	–	88.75	88.90	87.12	97.12	96.17	–

Note: Average values of recognition accuracy (%) are given in bold and standard deviations in parenthesis.
 * Unbiased hit rates, used to correct for potential response biases in forced-choice paradigms, since inter-individual differences in accuracy rates might represent differences in the tendency for participants to choose some categories more frequently than others.

We have also explored possible associations between emotion recognition performance and cognitive and sociodemographic variables. A significant positive correlation was found between accuracy on emotional prosody (arcsine transformed rates) and participants' IQ (full scale IQ scores; $r = 0.33$, $p = .006$). The correlations between accuracy on emotional prosody and age ($r = -0.12$), sex ($r = 0.92$), and the education level of the children's mother ($r = 0.90$) did not reach significance.

2. Neuroanatomical Data

Brain regions significantly associated with emotional prosody recognition ability are summarized in Tables 4 to 6.

We first conducted a whole-brain analysis focusing on the average emotion recognition accuracy, which allowed us to examine whether differences in the amount of grey matter volume in some brain regions might predict children's overall performance. As can be seen in Table 4 and Figure B1, the average emotional prosody recognition was positively associated with grey matter volume in right fusiform gyrus and in left cerebellum, indicating that children who recognized prosodic emotions with higher accuracy had more grey matter in these regions. Moreover, average emotional prosody

recognition was negatively associated with a cluster with peak voxel located in the left middle occipital gyrus, extending to the left superior occipital gyrus; with clusters in the superior parietal lobule; and with the bilateral precuneus.

We then focused on each emotion category separately. Higher recognition of neutral expressions correlated with larger grey matter volume in left cerebellum, and with decreased grey matter in the left middle occipital gyrus. Higher recognition of happy expressions correlated with larger grey matter volume in the left cerebellum and left superior frontal gyrus, and with decreased grey matter volume in superior parietal and superior occipital regions, as well as with bilateral precuneus. Higher recognition of sad expressions was associated with larger grey matter volume in frontal areas, particularly the superior frontal gyrus and the right middle frontal gyrus, and in motor and premotor areas. Additionally, it correlated with smaller grey matter volume in left cerebellum, right precuneus and middle orbital gyrus. For individual differences in the recognition of anger expressions, a positive correlation was found with the left cuneus, and negative correlations were found with the right precuneus and superior parietal and middle occipital regions. Finally, the ability to recognize fearful expressions was positively associated with right fusiform and inferior occipital gyri and left cerebellum, and negatively correlated with areas from occipital and parietal lobes.

Altogether, it seems like parietal and occipital regions mostly come up in negative correlations, while the cerebellum, motor, premotor and prefrontal regions come up in positive correlations.

Table 4. Brain regions associated with emotional prosody recognition ability in a 7 to 9 years old children sample (separate-modality regressions).

Emotional category	Anatomical region	Number of voxels	MNI coordinates			<i>T</i> value	<i>Z</i> score
			<i>x</i>	<i>y</i>	<i>z</i>		
Average of all categories	<i>Positive Correlations</i>						
	R Fusiform gyrus	798	36	-58	-18	4.65	4.27 *
	L Cerebellum	181	-24	-66	-54	3.82	3.59
	L Cerebellum		-26	-68	-63	3.34	3.18
	<i>Negative Correlations</i>						
	L Middle occipital gyrus	668	-28	-86	27	4.70	4.31 *
	L Superior occipital gyrus		-21	-78	22	4.06	3.79
	R Superior parietal lobule	561	14	-54	69	5.04	4.57
	R Superior parietal lobule		24	-46	78	3.43	3.26
	R Superior parietal lobule	70	27	-70	63	3.69	3.48
	R Precuneus		15	-69	60	3.44	3.26
	L Precuneus	32	-6	-52	42	3.46	3.29
	R Superior parietal lobule	31	10	-68	69	3.87	3.64

Table 4 (cont.). Brain regions associated with emotional prosody recognition ability in a 7 to 9 years old children sample (separate-modality regressions).

Emotional category	Anatomical region	Number of voxels	MNI coordinates			T value	Z score
			x	y	z		
Neutrality	<i>Positive Correlations</i>						
	L Cerebellum	132	-24	-66	-51	3.82	3.59
	<i>Negative Correlations</i>						
	L Middle occipital gyrus	205	-27	-84	30	3.93	3.69
	L Middle occipital gyrus		-27	-86	21	3.56	3.37
Happiness	<i>Positive Correlations</i>						
	L Cerebellum	42	-20	-63	-36	3.61	3.41
	L Superior frontal gyrus	41	-16	62	6	3.91	3.67
	<i>Negative Correlations</i>						
	L Precuneus	86	-4	-54	45	3.81	3.59
	L Superior parietal lobule	80	-20	-74	51	4.37	4.05
	R Precuneus	66	9	-75	63	3.89	3.65
	R Superior parietal lobule		9	-68	69	3.80	3.58
	L Superior occipital gyrus	48	-20	-80	21	3.55	3.36
Sadness	<i>Positive Correlations</i>						
	L Paracentral lobule	706	-10	-24	68	4.76	4.35
	L Superior frontal gyrus		-12	-3	78	3.97	3.72
	L Paracentral lobule		-9	-14	70	3.90	3.66
	R Precentral gyrus	71	57	4	22	3.57	3.38
	R Precentral gyrus		66	8	24	3.47	3.30
	R Superior frontal gyrus	55	18	-12	66	3.52	3.33
	R Supplementary motor area		9	-15	68	3.42	3.25
	L Precentral gyrus	37	-38	-12	58	3.44	3.27
	R Middle frontal gyrus	32	40	12	39	3.59	3.40
	<i>Negative Correlations</i>						
	L Cerebellum	643	-40	-84	-39	4.21	3.92
	L Cerebellum		-36	-90	-27	3.89	3.65
	L Cerebellum		-40	-80	-28	3.57	3.38
	R Precuneus	302	9	-56	66	4.18	3.90
	R Precuneus		4	-58	57	3.57	3.38
	R Middle orbital gyrus	67	16	44	-2	4.08	3.81
	L Cerebellum	58	-51	-70	-45	4.06	3.79
Anger	<i>Positive Correlations</i>						
	L Cuneus	33	-6	-84	28	3.57	3.38
	<i>Negative Correlations</i>						
	R Superior parietal lobule	280	16	-68	60	3.99	3.74
	R Superior parietal lobule		26	-70	64	3.98	3.72
	R Superior parietal lobule	219	15	-56	70	4.45	4.11
	L Middle occipital gyrus	98	-34	-80	24	4.18	3.89
	R Precuneus	54	4	-69	51	3.69	3.49
Fear	<i>Positive Correlations</i>						
	R Fusiform gyrus	1143	38	-54	-14	5.28	4.75
	R Fusiform gyrus		28	-63	-14	4.55	4.19
	R Inferior occipital gyrus		36	-64	-8	4.38	4.06
	L Cerebellum	127	-6	-70	-27	4.25	3.95
	<i>Negative Correlations</i>						
	L Calcarine gyrus	384	-8	-90	-9	4.41	4.08
	R Postcentral gyrus	30	14	-46	68	3.63	3.43

Note. L: left hemisphere; R: right hemisphere. *Clusters which survived familywise error correction at $p < .05$ (correction for multiple comparisons). Data were derived from separate-modality regression analyses and thresholded at $p < .001$, peak level uncorrected; all clusters with an extent of at least 30 voxels were reported. Coordinates are given for the Montreal Neurological Institute (MNI) stereotactic space.

A combined-modalities analysis, i.e., including all emotion categories in the same model, revealed associations between anatomical brain regions and recognition accuracy rates for each emotional category, while adjusting for variability related to the other emotional categories. These analyses therefore reveal associations with grey matter that are unique for a given emotion (these results are summarized in Table 5).

Table 5. Anatomical brain regions associated with emotional prosody recognition ability in a 7 to 9 years old children sample (adjusted for variability related to the other emotional categories).

Emotional category	Brain region	Number of voxels	MNI coordinates			T value	Z score
			x	y	z		
Neutrality	<i>Negative Correlations</i>						
	L Inferior occipital gyrus	73	-42	-84	-3	3.56	3.36
	R Superior orbital gyrus	42	22	56	-6	3.81	3.57
Happiness	<i>Positive Correlations</i>						
	R Cerebellum	253	16	-70	-32	4.75	4.32
	R Superior frontal gyrus	44	12	27	54	3.64	3.43
Sadness	<i>Positive Correlations</i>						
	L Paracentral lobule	611	-10	-24	69	4.66	4.25
	L Paracentral lobule		-10	-14	72	3.73	3.51
	L Paracentral lobule		-9	-26	80	3.57	3.37
	R Middle frontal gyrus	43	40	10	39	3.90	3.56
	L Superior frontal gyrus	36	-12	-3	78		
	R Precentral gyrus	33	66	8	24	3.45	3.26
	R Precentral gyrus		58	4	22	3.44	3.26
	<i>Negative Correlations</i>						
	L Cerebellum	512	-50	-72	-45	4.31	3.98
	L Cerebellum		-40	-84	-39	3.85	3.61
	L Cerebellum		-34	-75	-33	3.79	3.55
	R Superior medial gyrus	76	16	44	-2	4.05	3.77
	R Precuneus	66	9	-56	66	3.53	3.34
Anger	<i>Positive Correlations</i>						
	R Parahippocampal gyrus	98	27	2	-33	4.30	3.97
	L Temporal pole	71	-44	14	-26	3.60	3.39
	L Temporal pole	47	40	12	-26	3.75	3.52
	<i>Negative Correlations</i>						
	L Middle occipital gyrus	42	-36	-80	22	4.15	3.85
Fear	<i>Positive Correlations</i>						
	R Fusiform gyrus	395	38	-54	-14	4.03	3.75
	R Fusiform gyrus		28	-63	-15	3.50	3.30
	L Cerebellum	135	-4	-70	-26	3.92	3.66
Fear	<i>Negative Correlations</i>						
	R Temporal pole	332	34	12	-33	4.40	4.05
	L Calcarine gyrus	124	-4	-87	-12	3.80	3.56
	L Lingual gyrus		-12	-88	-6	3.51	3.32

Note. L: left hemisphere; R: right hemisphere. Data were thresholded at $p < .001$, peak level uncorrected; all clusters with an extent of at least 30 voxels were reported. Coordinates are given for the Montreal Neurological Institute (MNI) stereotactic space.

The results were generally similar to those found in the separate-modality analyses, that is, most areas were still significantly related with recognition of specific emotional

categories, when other emotions were included as covariates. However, unlike before, we did not find any suprathreshold positive correlations with neutrality, or negative ones with happiness. Also, anger was positively correlated with temporal areas and the parahippocampal gyrus, and fear was negatively correlated with temporal areas as well.

Based on the results described above, it seems that most areas are involved in the recognition of more than one emotional category (e.g., superior parietal lobule is negatively correlated with happiness and anger; and left cerebellum is positively correlated with neutrality, happiness and fear). However, there were also regions that were significantly more related to one emotion category than to the others, as indicated by direct comparisons between each emotion category and all the others (brain regions are summarised in Table 6). For instance, the left cerebellum and the right superior medial gyrus were more strongly associated with the recognition of happy expressions as compared to the others; for the recognition of sadness, occipital and motor areas were identified; the recognition of anger was associated with a larger grey matter volume in temporal and occipital areas, as well as in parahippocampal gyrus, cuneus, precuneus, cerebellum and midcingulate cortex; fear, in turn, was correlated with left cerebellum. No suprathreshold associations were found for the recognition of neutral expressions.

Table 6. Comparisons between anatomical brain regions associated with emotional prosody recognition ability in a 7 to 9 years old children sample.

Emotional category	Anatomical region	Number of voxels	MNI coordinates			T value	Z score
			x	y	z		
Happiness > Other emotions	L Cerebellum	342	16	-69	-33	4.98	4.50
	R Superior medial gyrus	73	12	27	54	3.79	3.55
Sadness > Other emotions	L Paracentral lobule	497	-10	-24	69	4.44	4.08
	L Middle occipital gyrus		-46	-84	0	3.53	3.33
	L Inferior occipital gyrus	35	-33	-87	-9	3.73	3.51
Anger > Other emotions	R Parahippocampal gyrus	104	27	2	-33	4.26	3.94
	L Temporal pole	87	-45	15	-24	3.65	3.44
	R Middle occipital gyrus	74	39	-84	12	3.73	3.50
	L Cuneus	58	-8	-86	28	3.69	3.47
	L Cuneus		-9	-93	34	3.29	3.13
	L Midcingulate cortex	55	-2	-4	45	3.74	3.51
	R Precuneus	51	6	-46	54	3.58	3.37
Fear > Other emotions	L Cerebellum	54	-3	-74	-26	3.55	3.35

Note. L: left hemisphere; R: right hemisphere. Data were thresholded at $p < .001$, peak level uncorrected; all clusters with an extent of at least 30 voxels were reported. Coordinates are given for the Montreal Neurological Institute (MNI) stereotactic space.

Discussion

The goal of the current study was to investigate whether and how individual differences in children's ability to recognize emotions through prosodic cues might reflect differences in brain morphology. All participants completed a structural MRI scan and a behavioural task designed to assess the emotional prosody recognition ability.

As predicted, children showed high recognition accuracy rates (well above chance level) for all emotional categories, indicating that they are sensitive to paralinguistic information, and are able to distinguish different emotional states based on prosodic cues. This adds to a previous study conducted with children aged between 5 and 10 years old, native speakers of English, in experimental tasks using non-verbal and emotionally inflected speech stimuli, and including ten emotional categories (Sauter et al., 2013).

Consistent with our hypothesis, analyses of behavioural data revealed significant differences in recognition accuracy across emotions. Happiness was the best recognized emotion, followed by anger (considering the unbiased hit rates, H_u), a finding consistent with results found for non-verbal vocalizations in previous studies conducted with children samples (Chronaki et al., 2015; Oerlemans et al., 2014; Sauter et al., 2013). Some studies with adult samples have shown a similar pattern (Mill, Allik, Realo, & Valk, 2009; Wildgruber et al., 2005). Considering the unbiased hit rates, neutrality and sadness were the worst recognized emotions by children, unlike adults, for whom those intonations typically yield higher accuracy rates (Castro & Lima, 2010). The fact that neutral and sad expressions have similarly low f_0 and f_0 variation values is possibly the main reason for the lower accuracy rates obtained in these categories, since they were often confused by children. Also, the intensity rates attributed to sad and neutral trials were lower than to all other emotional categories, strengthening the association between the pattern of confusions and the acoustic properties of stimuli.

We also found large individual differences in children's emotional recognition ability, i.e., children varied widely in how well they performed the task. The VBM analyses showed that these differences might be partly explained by the amount of grey matter in specific brain regions. fMRI studies with adult samples have reported activity in frontal (e.g., left precentral, middle frontal gyrus), temporal (e.g., middle temporal gyrus, superior temporal pole) and parietal areas (e.g., right precuneus), as well as subcortical regions, such as cerebellum and thalamus (e.g., Alba-Ferrara et al., 2011; Wildgruber et al.,

2005). Even though several studies have looked at brain responses to emotional speech prosody, those studies were mostly conducted with adults, and used fMRI. Therefore, it remains unclear whether the morphological features of the areas that are shown to be active in functional studies determine how accurately prosodic emotions are recognized.

The current study showed positive correlations between higher emotion recognition accuracy and larger grey matter volume in left cerebellum and in the right fusiform gyrus (occipitotemporal area). The cerebellum has been associated with learning, memory and language functions, as well as with emotional and regulatory behaviour, and with social cognition (O'Halloran, Kinsella, & Storey, 2012). These functions not only reflect general abilities implicated in the performance of several tasks, but they also reinforce the association between emotional prosody recognition and social cognition, since similar brain structures are simultaneously implicated in both processes.

Occipitoparietal areas are typically associated with emotion recognition from faces, especially the fusiform gyrus, which has been demonstrated to be specifically implicated in this ability (e.g., Duchaine & Yovel, 2015; McCarthy, Puce, Gore, & Allison, 1997). Although we emphasized the ability to recognize emotions through prosody, in our task children had to match prosodic cues extracted from semantically neutral sentences with emoticons, i.e., facial expression cartoons representing the five emotional categories (emoticons are illustrated in Figure 1). Therefore, it could be that performance levels were partly dependent on an efficient integration of auditory with visual cues. Also, some authors suggest that listening to emotional voices might automatically evoke mental images associated to that emotion, considering that visual and auditory signals similar to those experienced during the task are often associated in everyday social interactions (e.g., Brück et al., 2011a). For example, listening to our mother yelling, during a phone call, because we are late for dinner might evoke a mental image of her angry face. Such multimodal integration processes, i.e., linking emotional auditory stimuli with facial cues might be related with the activation of brain areas specifically associated with visual modality, particularly the fusiform gyrus (Kanwisher & Yovel, 2006; Kanwisher, McDermott, & Chun, 1997). Also, the mental imagery of visual stimuli, particularly faces, has been shown to activate the same brain area (O'Craven & Kanwisher, 2000). Although these hypotheses might contribute to a better understanding of the emotional prosody decoding process, they should be interpreted with caution, since they are supported by fMRI studies and, therefore, they do not directly show that the structural features of these brain areas impact on emotion recognition performance. Nonetheless, it is plausible to

argue that differences in grey matter volume in occipitoparietal areas might explain inter-individual variability in the accuracy to recognize multimodal emotional stimuli (both visual and acoustic).

We also found that higher emotion recognition accuracy is related with reduced grey matter volume in parietal regions, including bilateral superior parietal lobule and bilateral precuneus, and occipital areas, as middle occipital gyrus. The superior parietal lobule has been related with working memory processes (Koenigs, Barbey, Postle, & Grafman, 2009), which might be associated with more general attention and learning processes. Both the superior parietal lobule and precuneus have already been shown to be specifically associated with explicit processing of emotion stimuli, when compared to semantic materials, in fMRI studies (e.g., Alba-Ferrara et al., 2011; Bach, Grandjean, Sander, Herdener, Strik, & Seifritz, 2008).

A functional study with a sample of children aged 5 to 18 years old (Plante et al., 2006) has also found temporal, parietal and frontal activity (e.g., right middle frontal gyrus) during the recognition of emotional prosody. However, unlike the current study, they have found an association with specific temporal areas, such as the superior temporal gyrus, which is often referred as being involved in initial stages of emotional prosody recognition processing (Brück et al., 2011a; Wildgruber et al., 2009).

The brain regions that are typically found in functional studies of emotional prosody processing (e.g., superior temporal gyrus, inferior frontal gyrus) do not seem to fully overlap with the regions where the amount of grey matter is predictive of behavioural performance. Previous research has suggested that the age of 9 years is considered a developmental turning point, associated with changes in the way children use emotional social signals (e.g., Gil et al., 2016). Between 8 and 10 years old, children begin to rely preferentially on prosodic cues instead of the lexical content and situational signals in order to decode emotional stimuli (e.g., Gil, Aguert, Le Bigot, Lacroix, & Laval, 2014; Morton & Trehub, 2001). As a result, children's emotion recognition processing becomes similar to the one demonstrated by adults. As the age of the children that participated in the current study coincides with this 'turning point', some mechanisms they rely on to process and interpret emotional stimuli might not be well understood yet. Besides, as previously noticed, we are mainly comparing our results with those found in functional studies, particularly with adults and, therefore, it is not surprisingly that some different brain regions have come up both for positive and negative correlations.

VBM results also revealed that distinct emotion categories are partly related to the amount of grey matter in different neural structures, some of which are specific for some emotions. Recognition of happiness was positively associated with increased grey matter in the left superior frontal gyrus, an area also implicated in working memory. Sadness was positively associated with motor and premotor areas, sometimes found to be related with prosody recognition, particularly in studies using nonverbal vocalizations (e.g., McGettigan et al., 2015; Warren et al., 2006); and with superior and middle frontal gyrus, the last one consistently found in functional studies focusing on emotional prosody recognition (e.g., Mitchell, Elliott, Barry, Cruttenden, & Woodruff, 2003; Plante et al., 2006; Witteman et al., 2012). The supplementary motor area, also associated with sadness, has been proved to be related with auditory imagery, the process by which people create and process mental images in the absence of the auditory stimuli, which can act as a facilitator of cognitive processes, and was proven to be related to visual imagery as well (Lima et al., 2015). Accuracy for anger trials was mainly related with the parahippocampal gyrus, which is often activated during the attribution of emotional salience of stimuli (e.g., Almeida, Mechelli, Hassel, Versace, Kupfer, & Phillips, 2009; Surguladze et al., 2006), and the temporal pole, an area related with social and emotional processing, particularly with the theory of mind, and therefore often shown to be activated in tasks involving the inference of others' emotional states (e.g., Olson, Plotzker, & Ezzyat, 2007). Finally, the recognition of fear was strongly related with increased grey matter volume in occipital areas, especially the fusiform gyrus, so it is possibly the emotional category that elicits the most vivid image representations. As previously stated, it seems that children make use of imagery processes to represent and process the stimuli and, according to that, choose the most suitable emotional category. It is possible that the experiences of fear in children are more frequent and salient and therefore more easily evocated from episodic memory, when compared to other emotional experiences. Accordingly, some studies have demonstrated the influence of life experiences in children's ability to recognize emotions from experimental stimuli (e.g., Pollack & Sinha, 2002; see also Brück et al., 2011a).

We have found positive and negative correlations between the grey matter volume and children's accuracy in the experimental task. Generally, parietal and occipital areas revealed negative correlations, while the cerebellum, motor, premotor and prefrontal regions were mainly positively correlated with children's accuracy on recognition of emotions from prosodic cues. Therefore, while higher accuracy rates were associated with increased grey matter volume in some brain regions (e.g., left cerebellum), they are also

related with a reduced grey matter volume in other areas, such as the middle and superior occipital gyri. Smaller brain volume may reflect differences in cortical thinning (e.g., Durston & Casey, 2006), a process that contributes to brain maturation, which is often associated with improvements in the ability to perform several tasks (e.g., Squeglia, Jacobus, Sorg, Jernigan, & Tapert, 2013). For instance, in the current study, a smaller grey matter volume in middle and superior occipital gyri, right superior parietal lobule, and precuneus seem to be related with differences in children's cortical thinning and, consequently, in the emotional prosody recognition ability.

In sum, our structural neuroimaging results revealed some regions that have been previously associated with emotional prosody processing in functional studies (e.g., middle frontal gyrus and temporal pole), yet some other regions do not seem to overlap with the network typically assumed to support emotional prosody (e.g., right fusiform gyrus and left cerebellum). Also, our results suggest that several brain areas are involved in the recognition of different emotions (e.g., left cerebellum is positively correlated with neutrality, happiness and fear), which suggests the existence of an integrated neural network generally involved in the recognition of all emotions. However, except for neutrality, some brain regions were correlated with higher recognition accuracy of specific emotions, when compared to the other ones. Therefore, it might be assumed that individual differences in children's ability to recognize emotions through prosodic cues reflect differences in brain morphology, both for the general emotional prosody recognition ability and for the recognition of specific emotional categories.

Conclusion

Considering the lack of research focusing on emotion recognition of prosodic stimuli in children, the current study brings a new contribution to the field, by showing that children are able to efficiently distinguish between several discrete emotions, being particularly good at recognizing happy intonations. Moreover, analyses of inter-individual differences in children's emotional prosody recognition ability revealed that this variability might be explained by differences in brain anatomy, and can be evaluated through VBM analyses of structural MRI scans. Since this seems to be the first study examining the structural neuroanatomical basis of emotional prosody recognition in children, it might lead to a better understanding of this topic. Besides, considering that deficits in the ability to recognize emotions from prosodic stimuli have been associated with different clinical conditions in children, such as autism (e.g., Filipe, Frota, Castro, & Vicente, 2014; van Lancker, Cornelius, & Kreiman, 1989), deafness (e.g., Hopyan-Misakyan, Gordon, Dennis, & Papsin, 2009) and pediatric traumatic brain injury (e.g., Schmidt, Hanten, Li, Orsten, & Levin, 2010), the understanding of the neural substrates of emotional prosody recognition might also behold clinical implications. For instance, research on emotional prosody recognition might support the development of diagnostic instruments or intervention programs for different clinical conditions.

It is important to mention that some of the brain areas found in the current study have also been associated, in previous studies, with different domain-general cognitive functions, such as the left cerebellum (e.g., Bellebaum & Daum, 2007), and prefrontal areas, like the superior frontal gyrus (e.g., du Boisgueheneuc et al., 2006), both implicated in executive functioning. As our experimental task involved attention and decision making processes, the association with areas implicated in the executive functioning is not unexpected. Nonetheless, it remains unclear whether these areas are specifically associated with the emotional prosody recognition or if they might reflect broader cognitive functions implicated in the experimental task.

It is also relevant to mention some of the potential limitations of the current study. First, the introduction of feedback after each trial might have influenced the results, particularly by inducing an additional learning effect. This could explain the involvement of brain areas associated with learning processes, i.e., regions related with cognitive processes, such as working memory. Nonetheless, we compared results from the first and

second blocks of stimuli, and realized that the improvement was not so substantial. Also, the limited number of stimuli per category might limit the power of comparisons between emotions. However, we still found significant effects, which we could relate with previous research. An additional limitation is the relatively small sample size, which might difficult the identification of brain areas significantly implicated in the emotional recognition prosody. Finally, we have only found two clusters that survived familywise error correction for multiple comparisons. Thus, although all the statistical maps have been thresholded at $p < 0.001$ (peak level uncorrected), results might be interpreted with caution.

To our knowledge, this is the first study examining whether differences in brain morphology explain inter-individual differences in children's ability to recognize emotions from prosodic stimuli. Therefore, further work is required in order to replicate the results in different samples. It would be interesting to compare children from different age groups, in order to allow a better understanding of the neural processes involved in the recognition of emotions from prosodic cues in different developmental stages. Also, considering the neuroanatomical results, further studies should compare the emotion recognition of prosodic and visual stimuli, especially concerning the brain morphology associated, in order to allow a better understanding of the way children perceive, process and interpret emotional stimuli.

References

- Adolphs, R., Damasio, H., & Tranel, D. (2002). Neural systems for recognition of emotional prosody: a 3-D lesion study. *Emotion*, 2(1), 23–51. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12899365>
- Alba-Ferrara, L., Hausmann, M., Mitchell, R. L., & Weis, S. (2011). The neural correlates of emotional prosody comprehension: disentangling simple from complex emotion. *PloS One*, 6(12), 1–9. doi:10.1371/journal.pone.0028701
- Almeida, J. R. C., Mechelli, A., Hassel, S., Versace, A., Kupfer, D. J., & Phillips, M. L. (2009). Abnormally increased effective connectivity between parahippocampal gyrus and ventromedial prefrontal regions during emotion labeling in bipolar disorder. *Psychiatry Research*, 174(3), 195–201. doi:10.1016/j.psychresns.2009.04.015
- Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *NeuroImage*, 38(1), 95–113. doi:10.1016/j.neuroimage.2007.07.007
- Bach, D. R., Grandjean, D., Sander, D., Herdener, M., Strik, W. K., & Seifritz, E. (2008). The effect of appraisal level on processing of emotional prosody in meaningless speech. *NeuroImage*, 42(2), 919–927. doi:10.1016/j.neuroimage.2008.05.034
- Banissy, M. J., Kanai, R., Walsh, V., & Rees, G. (2012). Inter-individual differences in empathy are reflected in human brain structure. *NeuroImage*, 62(3), 2034–2039. doi:10.1016/j.neuroimage.2012.05.081
- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70(3), 614–636.
- Baum, S. R., & Pell, M. D. (1999). The neural basis of prosody: Insights from lesion studies and neuroimaging. *Aphasiology*, 13, 581–608. doi:10.1080/026870399401957
- Beal, D. S., Gracco, V. L., Brettschneider, J., Kroll, R. M., & De Nil, L. F. (2013). A voxel-based morphometry (VBM) analysis of regional grey and white matter volume abnormalities within the speech production network of children who stutter. *Cortex*, 49(8), 2151–2161. doi:10.1016/j.cortex.2012.08.013
- Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends in Cognitive Sciences*, 8(3), 129–135. doi:10.1016/j.tics.2004.01.008

- Bellebaum, C., & Daum, I. (2007). Cerebellar involvement in executive control. *Cerebellum*, 6(3), 184–192. doi:10.1080/14734220601169707
- Briesemeister, B. B., Kuchinke, L., & Jacobs, A. M. (2014). Emotion word recognition: discrete information effects first, continuous later? *Brain Research*, 1564, 62–71. doi:10.1016/j.brainres.2014.03.045
- Brück, C., Kreifelts, B., & Wildgruber, D. (2011a). Emotional voices in context: a neurobiological model of multimodal affective information processing. *Physics of Life Reviews*, 8(4), 383–403. doi:10.1016/j.plrev.2011.10.002
- Brück, C., Kreifelts, B., Kaza, E., Lotze, M., & Wildgruber, D. (2011b). Impact of personality on the cerebral processing of emotional prosody. *NeuroImage*, 58(1), 259–268. doi:10.1016/j.neuroimage.2011.06.005
- Castro, S. L., & Lima, C. F. (2010). Recognizing emotions in spoken language: a validated set of Portuguese sentences and pseudosentences for research on emotional prosody. *Behavior Research Methods*, 42(1), 74–81. doi:10.3758/BRM.42.1.74
- Chakrabarti, B., & Baron-Cohen, S. (2006). Empathizing: neurocognitive developmental mechanisms and individual differences. *Progress in Brain Research*, 156, 403–17. doi:10.1016/S0079-6123(06)56022-4
- Chronaki, G., Hadwin, J. A., Garner, M., Maurage, P., & Sonuga-Barke, E. J. S. (2015). The development of emotion recognition from facial expressions and non-linguistic vocalizations during childhood. *The British Journal of Developmental Psychology*, 33(2), 218–236. doi:10.1111/bjdp.12075
- Colibazzi, T., Posner, J., Wang, Z., Gorman, D., Gerber, A., Yu, S., ... Peterson, B. S. (2010). Neural systems subserving valence and arousal during the experience of induced emotions. *Emotion*, 10(3), 377–89. doi:10.1037/a0018484
- Damasio, A., & Carvalho, G. B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews. Neuroscience*, 14(2), 143–152. doi:10.1038/nrn3403
- Dara, C., Bang, J., Gottesman, R. F., & Hillis, A. E. (2014). Right hemisphere dysfunction is better predicted by emotional prosody impairments as compared to neglect. *Journal of Neurology & Translational Neuroscience*, 2(1), 1037–1050. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/24945016>
- Darwin, C. R. 1872. *The expression of the emotions in man and animals*. London: John Murray.

- du Boisgucheneuc, F., Levy, R., Volle, E., Seassau, M., Duffau, H., Kinkingnehun, S., ... Dubois, B. (2006). Functions of the left superior frontal gyrus in humans: a lesion study. *Brain*, *129*(12), 3315–3328. doi:10.1093/brain/awl244
- Duchaine, B., & Yovel, G. (2015). A revised neural framework for face processing. *Annual Review of Vision Science*, *1*(1), 393–416. doi:10.1146/annurev-vision-082114-035518
- Durston, S., & Casey, B. J. (2006). What have we learned about cognitive development from neuroimaging? *Neuropsychologia*, *44*(11), 2149–2157. doi:10.1016/j.neuropsychologia.2005.10.010
- Erola, J., Jalonen, S., & Lehti, H. (2016). Parental education, class and income over early life course and children's achievement. *Research in Social Stratification and Mobility*, *44*, 33–43. doi:10.1016/j.rssm.2016.01.003
- Ethofer, T., Van De Ville, D., Scherer, K., & Vuilleumier, P. (2009). Decoding of emotional information in voice-sensitive cortices. *Current Biology*, *19*(12), 1028–1033. doi:10.1016/j.cub.2009.04.054
- Filipe, M. G., Frota, S., Castro, S. L., & Vicente, S. G. (2014). Atypical prosody in Asperger syndrome: perceptual and acoustic measurements. *Journal of Autism and Developmental Disorders*, *44*(8), 1972–1981. doi:10.1007/s10803-014-2073-2
- Fujimura, T., Matsuda, Y. T., Katahira, K., Okada, M., & Okanoya, K. (2012). Categorical and dimensional perceptions in decoding emotional facial expressions. *Cognition & Emotion*, *26*(4), 587–601. doi:10.1080/02699931.2011.595391
- Gil, S., Aguert, M., Bigot, L. L., Lacroix, A., & Laval, V. (2014). Children's understanding of others' emotional states: Inferences from extralinguistic or paralinguistic cues? *International Journal of Behavioral Development*, *38*(6), 539–549. doi:10.1177/0165025414535123
- Gil, S., Hattouti, J., & Laval, V. (2016). How children use emotional prosody: Crossmodal emotional integration? *Developmental Psychology*, *52*(7), 1064–1072. doi:10.1037/dev0000121
- Gillioz, C., Fontaine, J. R. J., Soriano, C., & Scherer, K. R. (2016). Mapping Emotion Terms into Affective Space: Further Evidence for a Four-Dimensional Structure. *Swiss Journal of Psychology*, *75*(3), 141–148. doi:10.1024/1421-0185/A000180
- Griffiths, S., Jarrold, C., Penton-Voak, I. S., & Munafò, M. R. (2015). Feedback training induces a bias for detecting happiness or fear in facial expressions that generalises

- to a novel task. *Psychiatry Research*, 230(3), 951–957. doi:10.1016/j.psychres.2015.11.007
- Haller, S. P. W., Raeder, S. M., Scerif, G., Cohen Kadosh, K., & Lau, J. Y. F. (2016). Measuring online interpretations and attributions of social situations: Links with adolescent social anxiety. *Journal of Behavior Therapy and Experimental Psychiatry*, 50, 250–256. doi:10.1016/j.jbtep.2015.09.009
- Heller, K., Matsak, E., Abboud, H., Schultz, H., & Zeitlin, V. (2014). SuperLab, stimulus presentation software (Version 5.0.3) [Computer software]. San Pedro, CA: Cedrus Corporation.
- Hockenbury, D. H. & Hockenbury, S. E. (2007). *Discovering psychology*. New York: Worth Publishers.
- Hopyan-Misakyan, T. M., Gordon, K. A., Dennis, M., & Papsin, B. C. (2009). Recognition of affective speech prosody and facial affect in deaf children with unilateral right cochlear implants. *Child Neuropsychology*, 15(2), 136–146. doi:10.1080/09297040802403682
- Ibarretxe-Bilbao, N., Junque, C., Tolosa, E., Marti, M. J., Valldeoriola, F., Bargallo, N., & Zarei, M. (2009). Neuroanatomical correlates of impaired decision-making and facial emotion recognition in early Parkinson's disease. *The European Journal of Neuroscience*, 30(6), 1162–1171. doi:10.1111/j.1460-9568.2009.06892.x
- Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews. Neuroscience*, 12(4), 231–242. doi:10.1038/nrn3000
- Kanai, R., Bahrami, B., Roylance, R., & Rees, G. (2012). Online social network size is reflected in human brain structure. *Proceedings. Biological Sciences / The Royal Society*, 279(1732), 1327–1334. doi:10.1098/rspb.2011.1959
- Kanai, R., Feilden, T., Firth, C., & Rees, G. (2011). Political orientations are correlated with brain structure in young adults. *Current Biology*, 21(8), 677–680. doi:10.1016/j.cub.2011.03.017
- Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 361(1476), 2109–2128. doi:10.1098/rstb.2006.1934
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of*

- Neuroscience*, 17(11), 4302–11. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9151747>
- Koenigs, M., Barbey, A. K., Postle, B. R., & Grafman, J. (2009). Superior parietal cortex is critical for the manipulation of information in working memory. *The Journal of Neuroscience*, 29(47), 14980–14986. doi:10.1523/JNEUROSCI.3706-09.2009
- Kotz, S. A., & Paulmann, S. (2011). Emotion, Language, and the Brain. *Language and Linguistics Compass*, 5(3), 108–125. doi:10.1111/j.1749-818X.2010.00267.x
- Kotz, S. A., Meyer, M., Alter, K., Besson, M., von Cramon, D. Y., & Friederici, A. D. (2003). On the lateralization of emotional prosody: an event-related functional MR investigation. *Brain and Language*, 86(3), 366–376. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12972367>
- Kragel, P. A., & LaBar, K. S. (2016). Decoding the nature of emotion in the brain. *Trends in Cognitive Sciences*, 20(6), 444–455. doi:10.1016/j.tics.2016.03.011
- Kunzmann, U., Kappes, C., & Wrosch, C. (2014). Emotional aging: a discrete emotions perspective. *Frontiers in Psychology*, 5, 380. doi:10.3389/fpsyg.2014.00380
- Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology*, 66, 799–823. doi:10.1146/annurev-psych-010213-115043
- Lima, C. F., & Castro, S. L. (2011). Emotion recognition in music changes across the adult life span. *Cognition & Emotion*, 25(4), 585–598. doi:10.1080/02699931.2010.502449
- Lima, C. F., Alves, T., Scott, S. K., & Castro, S. L. (2014). In the ear of the beholder: how age shapes emotion processing in nonverbal vocalizations. *Emotion*, 14(1), 145–160. doi:10.1037/a0034287
- Lima, C. F., Lavan, N., Evans, S., Agnew, Z., Halpern, A. R., Shanmugalingam, P., ... Scott, S. K. (2015). Feel the noise: relating individual differences in auditory imagery to the structure and function of sensorimotor systems. *Cerebral Cortex*, 25(11), 4638–4650. doi:10.1093/cercor/bhv134
- McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, 9(5), 605–610. doi:10.1162/jocn.1997.9.5.605
- McClure, E. B., & Nowicki, Jr., S. (2001). Associations between social anxiety and nonverbal processing skill in preadolescent boys and girls. *Journal of Nonverbal Behavior*, 25(1), 3–19. doi:10.1023/A:1006753006870

- 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. doi:10.1093/cercor/bht227
- Mill, A., Allik, J., Realo, A., & Valk, R. (2009). Age-related differences in emotion recognition ability: a cross-sectional study. *Emotion*, *9*(5), 619–630. doi:10.1037/a0016562
- Mitchell, R. L. C., Elliott, R., Barry, M., Cruttenden, A., & Woodruff, P. W. R. (2003). The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia*, *41*(10), 1410–1421. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12757912>
- Mitchell, R. L. C., & Phillips, L. H. (2015). The overlapping relationship between emotion perception and theory of mind. *Neuropsychologia*, *70*, 1–10. doi:10.1016/j.neuropsychologia.2015.02.018
- Morton, J. B., & Trehub, S. E. (n.d.). Children’s understanding of emotion in speech. *Child Development*, *72*(3), 834–843. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11405585>
- O’Craven, K. M., & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience*, *12*(6), 1013–1023. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11177421>
- O’Halloran, C. J., Kinsella, G. J., & Storey, E. (2012). The cerebellum and neuropsychological functioning: a critical review. *Journal of Clinical and Experimental Neuropsychology*, *34*(1), 35–56. doi:10.1080/13803395.2011.614599
- Oerlemans, A. M., van der Meer, J. M. J., van Steijn, D. J., de Ruiter, S. W., de Bruijn, Y. G. E., de Sonnevile, L. M. J., ... Rommelse, N. N. J. (2014). Recognition of facial emotion and affective prosody in children with ASD (+ADHD) and their unaffected siblings. *European Child & Adolescent Psychiatry*, *23*(5), 257–271. doi:10.1007/s00787-013-0446-2
- Olson, I. R., Plotzker, A., & Ezzyat, Y. (2007). The Enigmatic temporal pole: a review of findings on social and emotional processing. *Brain*, *130*(7), 1718–1731. doi:10.1093/brain/awm052
- Omar, R., Henley, S. M. D., Bartlett, J. W., Hailstone, J. C., Gordon, E., Sauter, D. A., ... Warren, J. D. (2011). The structural neuroanatomy of music emotion recognition:

- evidence from frontotemporal lobar degeneration. *NeuroImage*, 56(3), 1814–1821. doi:10.1016/j.neuroimage.2011.03.002
- Pell, M. D., & Leonard, C. L. (2003). Processing emotional tone from speech in Parkinson's disease: a role for the basal ganglia. *Cognitive, Affective & Behavioral Neuroscience*, 3(4), 275–288. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15040548>
- Plante, E., Holland, S. K., & Schmithorst, V. J. (2006). Prosodic processing by children: an fMRI study. *Brain and Language*, 97(3), 332–342. doi:10.1016/j.bandl.2005.12.004
- Pollak, S.D., & Sinha, P. (2002). Enhanced perceptual sensitivity for anger among physically abused children. *Developmental Psychology*, 38, 784–791.
- Rohrer, J. D., Sauter, D., Scott, S., Rossor, M. N., & Warren, J. D. (2012). Receptive prosody in nonfluent primary progressive aphasia. *Cortex*, 48(3), 308–316. doi:10.1016/j.cortex.2010.09.004
- Saarimäki, H., Gotsopoulos, A., Jääskeläinen, I. P., Lampinen, J., Vuilleumier, P., Hari, R., ... Nummenmaa, L. (2015). Discrete neural signatures of basic emotions. *Cerebral Cortex*, 26(6), 2563–2573. doi:10.1093/cercor/bhv086
- Sato, W., Kochiyama, T., Uono, S., Kubota, Y., Sawada, R., Yoshimura, S., & Toichi, M. (2015). The structural neural substrate of subjective happiness. *Scientific Reports*, 5, 1–7. doi:10.1038/srep16891
- Sauter, D. A., Panattoni, C., & Happé, F. (2013). Children's recognition of emotions from vocal cues. *The British Journal of Developmental Psychology*, 31(1), 97–113. doi:10.1111/j.2044-835X.2012.02081.x
- Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, 40(1), 227–256. doi:10.1016/S0167-6393(02)00084-5
- Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, 10(1), 24–30. doi:10.1016/j.tics.2005.11.009
- Schmidt, A. T., Hanten, G. R., Li, X., Orsten, K. D., & Levin, H. S. (2010). Emotion recognition following pediatric traumatic brain injury: longitudinal analysis of emotional prosody and facial emotion recognition. *Neuropsychologia*, 48(10), 2869–2877. doi:10.1016/j.neuropsychologia.2010.05.029
- Simões, M., Rocha, A. M., & Ferreira, C. (2003). *WISC-III, Escala de Inteligência de Wechsler para Crianças-3ª edição*. Lisboa: CEGOC-TEA.

- Squeglia, L. M., Jacobus, J., Sorg, S. F., Jernigan, T. L., & Tapert, S. F. (2013). Early adolescent cortical thinning is related to better neuropsychological performance. *Journal of the International Neuropsychological Society*, *19*(9), 962–970. doi:10.1017/S1355617713000878
- Surguladze, S., Russell, T., Kucharska-Pietura, K., Travis, M. J., Giampietro, V., David, A. S., & Phillips, M. L. (2006). A reversal of the normal pattern of parahippocampal response to neutral and fearful faces is associated with reality distortion in schizophrenia. *Biological Psychiatry*, *60*(5), 423–431. doi:10.1016/j.biopsych.2005.11.021
- Van Lancker, D., Cornelius, C., & Kreiman, J. (1989). Recognition of emotional prosodic meanings in speech by autistic, schizophrenic, and normal children. *Developmental Neuropsychology*, *5*(2-3), 207–226. doi:10.1080/87565648909540433
- Vannest, J., Karunanayaka, P. R., Schmithorst, V. J., Szaflarski, J. P., & Holland, S. K. (2009). Language networks in children: evidence from functional MRI studies. *AJR. American Journal of Roentgenology*, *192*(5), 1190–1196. doi:10.2214/AJR.08.2246
- Wager, T. D., Kang, J., Johnson, T. D., Nichols, T. E., Satpute, A. B., & Barrett, L. F. (2015). A Bayesian model of category-specific emotional brain responses. *PLoS Computational Biology*, *11*(4), 1–27. doi:10.1371/journal.pcbi.1004066
- Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior*, *17*(1), 3–28. doi:10.1007/BF00987006
- Warren, J. E., Sauter, D. A., Eisner, F., Wiland, J., Dresner, M. A., Wise, R. J. S., ... Scott, S. K. (2006). Positive emotions preferentially engage an auditory-motor “mirror” system. *The Journal of Neuroscience*, *26*(50), 13067–13075. doi:10.1523/JNEUROSCI.3907-06.2006
- Wartenburger, I., Steinbrink, J., Telkemeyer, S., Friedrich, M., Friederici, A. D., & Obrig, H. (2007). The processing of prosody: Evidence of interhemispheric specialization at the age of four. *NeuroImage*, *34*(1), 416–425. doi:10.1016/j.neuroimage.2006.09.009
- Wechsler, D. (1991). *The Wechsler intelligence scale for children—third edition*. San Antonio, TX: The Psychological Corporation.

- Wiethoff, S., Wildgruber, D., Grodd, W., & Ethofer, T. (2009). Response and habituation of the amygdala during processing of emotional prosody. *Neuroreport*, *20*(15), 1356–1360. doi:10.1097/WNR.0b013e328330eb83
- Wildgruber, D., Ackermann, H., Klein, M., Riecker, A., & Grodd, W. (2000). Brain activation during identification of affective speech melody: Influence of emotional valence and sex. *NeuroImage*, *11*(5), 341. doi:10.1016/S1053-8119(00)91273-4
- Wildgruber, D., Ethofer, T., Grandjean, D., & Kreifelts, B. (2009). A cerebral network model of speech prosody comprehension. *International Journal of Speech-Language Pathology*, *11*(4), 277–281. doi:10.1080/17549500902943043
- Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., & Ackermann, H. (2005). Identification of emotional intonation evaluated by fMRI. *NeuroImage*, *24*(4), 1233–1241. doi:10.1016/j.neuroimage.2004.10.034
- Wilke, M., Holland, S. K., Altaye, M., & Gaser, C. (2008). Template-O-Matic: a toolbox for creating customized pediatric templates. *NeuroImage*, *41*(3), 903–913. doi:10.1016/j.neuroimage.2008.02.056
- Witteman, J., Van Heuven, V. J. P., & Schiller, N. O. (2012). Hearing feelings: a quantitative meta-analysis on the neuroimaging literature of emotional prosody perception. *Neuropsychologia*, *50*(12), 2752–2763. doi:10.1016/j.neuropsychologia.2012.07.026
- World Health Organization. (1992). *International classification of diseases, 10th revision (ICD-10)*. Geneva: WHO.
- Yoon, U., Fonov, V. S., Perusse, D., Evans, A. C., & Brain Development Cooperative Group. (2009). The effect of template choice on morphometric analysis of pediatric brain data. *NeuroImage*, *45*(3), 769–777. doi:10.1016/j.neuroimage.2008.12.046

Appendices

Appendix A

Table A1. Statistics of normality test (Shapiro-Wilk) and values of kurtosis and skewness are given for the original, transformed (with arcsine) and corrected (Hu) accuracy rates.

Emotional Category	Original Accuracy Rates			
	Shapiro-Wilk		Kurtosis	Skewness
	Statistic	P value		
Average of all categories	0.966	0.067	-1.063	-0.048
Neutrality	0.920	<0.001	-0.689	-0.593
Happiness	0.856	<0.001	-0.302	-0.775
Sadness	0.899	<0.001	0.755	-1.017
Anger	0.947	0.007	-0.564	-0.476
Fear	0.967	0.072	-0.693	-0.127

Emotional Category	Arcsine Transformed Accuracy Rates			
	Shapiro-Wilk		Kurtosis	Skewness
	Statistic	P value		
Average of all categories	0.970	0.111	-0.953	0.103
Neutrality	0.958	0.024	-0.027	-0.423
Happiness	0.865	<0.001	-1.081	0.137
Sadness	0.924	0.001	1.050	-0.579
Anger	0.958	0.026	0.471	-0.352
Fear	0.971	0.130	0.161	0.247

Emotional Category	Unbiased Hit Rates (Hu)			
	Shapiro-Wilk		Kurtosis	Skewness
	Statistic	P value		
Average of all categories	0.968	0.088	-0.934	0.174
Neutrality	0.966	0.064	-0.979	-0.148
Happiness	0.977	0.247	-0.803	-0.021
Sadness	0.940	0.003	-0.316	-0.672
Anger	0.964	0.053	-0.805	-0.259
Fear	0.975	0.215	-0.636	0.113

Appendix B

