FACULDADE DE PSICOLOGIA
E DE CIÊNCIAS DA EDUCAÇÃO
UNIVERSIDADE DO PORTO

César F. Lima

PERCEIVING EMOTIONS IN SPEECH AND MUSIC: MODULATIONS RELATED TO AGEING, MUSICAL TRAINING, AND PARKINSON'S DISEASE

TESE DE DOUTORAMENTO
PSICOLOGIA
2011

PERCEIVING EMOTIONS IN SPEECH AND MUSIC

Modulations Related to Ageing, Musical Training, and Parkinson's Disease

CÉSAR F. LIMA

PERCEIVING EMOTIONS IN SPEECH AND MUSIC

Modulations Related to Ageing, Musical Training, and Parkinson's Disease

CÉSAR F. LIMA

Thesis supervised by Professor São Luís Castro and presented at University of Porto for the Ph.D. degree in Psychology

September 2011

The studies presented in this thesis were supported by a doctoral grant from the Portuguese Foundation for Science and Technology (SFRH/BD/39306/2007) and by the Language Research Group of the Centre for Psychology at University of Porto (R&D Unit funded by FCT, 160/50; PTDC/PSI/66641/2006). They are also part of the project *Emotional processing from language and music: Comparative neurocognitive and functional neuroimaging studies,* funded by the Bial Foundation (N.º 29/08).

KEYWORDS

Emotion recognition; speech prosody; music; ageing; musical training; Parkinson's disease.

American Psychological Association (PsycINFO® Content Classification Code System):

2300 Human Experimental Psychology

2326 Auditory & Speech Perception

2343 Learning & Memory

2500 Physiological Psychology & Neuroscience

2520 Neuropsychology & Neurology

RESUMO

A fala e a música são dois poderosos meios de comunicação de emoções. A presente tese investiga como percebemos emoções em prosódia da fala – o *tom de voz* – e em música instrumental. Apresentamos uma série de estudos que visam três objectivos principais: determinar como o reconhecimento de emoções em música é modulado pela idade e pela formação musical (estudos 1 e 2); desenvolver uma base de estímulos de fala em Português para investigação em emoções prosódicas (estudo 3); e examinar a hipótese de que o reconhecimento de emoções em fala e em música depende de mecanismos neurocognitivos comuns (estudos 4 e 5). Utilizámos tarefas de escolha forçada e de julgamento de magnitude. Nos estudos 1 e 2, testámos adultos que variavam quanto à idade e formação musical no reconhecimento de alegria, serenidade, tristeza e medo em excertos musicais. Observámos que o aumento da idade está associado a uma diminuição na sensibilidade às emoções negativas, tristeza e medo, enquanto a sensibilidade às emoções positivas permanece estável. As mudanças na tristeza e medo foram significativas da meia-idade em diante. O número de anos de formação musical esteve correlacionado com uma maior sensibilidade às emoções musicais. Os efeitos de idade e formação foram independentes de diferenças em capacidades cognitivas gerais, o que sugere que aqueles efeitos têm uma origem primária. No estudo 3, duas falantes gravaram uma lista de frases e de pseudo-frases em Português variando a prosódia para comunicar neutralidade e seis emoções: alegria, fúria, medo, repulsa, surpresa e tristeza. Os perfis acústicos das diferentes emoções foram consistentes com descrições em outras línguas. As emoções foram reconhecidas com elevados níveis de exatidão, quer nas frases (75% correto), quer nas pseudo-frases (71%). Os estímulos foram incluídos numa base de dados (190 frases e 178 pseudo-frases) que disponibilizámos para as comunidades de investigação e clínica. Nos estudos 4 e 5 determinámos em que medida o reconhecimento de emoções em fala e em música recruta mecanismos partilhados. Os resultados sugerem o envolvimento de uma combinação de mecanismos partilhados e específicos ao domínio. No estudo 4, investigámos se a formação num dos domínios, música, está associada a benefícios no outro domínio, fala. Comparámos músicos com participantes sem formação musical no reconhecimento de neutralidade e de seis emoções em fala. Encontrámos um efeito robusto de transferência entre domínios: os músicos foram mais exatos do que os participantes sem formação em todas as emoções. Este resultado indica que os mecanismos são pelo menos parcialmente partilhados. No estudo 5, implementámos um design comparativo entre fala e música na doença de Parkinson (DP). Será que a DP produz um perfil similar de défices em ambos os domínios? Encontrámos uma dissociação: os doentes tiveram défices no reconhecimento de emoções positivas em música, alegria e serenidade, mas foram normais na fala; tiveram défices no reconhecimento de tristeza em fala, mas foram normais na música. Os défices na música não podem ser explicados por dificuldades percetivas e cognitivas; na fala, o reconhecimento de emoções esteve fortemente associado à disfunção executiva dos doentes. Esta dissociação neuropsicológica é evidência de que fala e música podem recrutar mecanismos específicos ao domínio. No seu conjunto, os cinco estudos aqui apresentados contribuem para o avanço da nossa compreensão sobre diferenças

individuais no reconhecimento de emoções em fala e em música, apresentam uma ferramenta para investigação em emoções vocais, e informam o debate atual sobre as relações música-linguagem. Do ponto de vista clínico, os nossos resultados constituem também um contributo para o conhecimento sobre a fenomenologia da DP.

RÉSUMÉ

La parole et la musique sont deux puissants moyens de communication émotionnelle. Cette thèse investigue la manière dont nous percevons les émotions à travers la prosodie de la parole – le *ton de la voix* – et la musique instrumentale. Nous poursuivons trois buts principaux à travers une série d'études: déterminer comment la reconnaissance de l'émotion dans la musique est modérée par l'âge et l'entraînement musical (études 1 et 2); fournir une base de données des stimuli vocale en portugais pour la recherche sur les émotions vocales (étude 3); examiner l'hypothèse selon laquelle le traitement émotionnel de la parole et de la musique fait appel à des mécanismes neurocognitifs communs à ces deux domaines (études 4 et 5). Des réponses à choix forcé et des tâches d'évaluation de magnitude ont été employées. Dans les études 1 et 2, nous avons testé des adultes variant en âge et en entraînement musical sur leur reconnaissance du joie, de la sérénité, de la tristesse et de la peur dans des extraits musicaux. Un âge supérieur était associé à une sensibilité diminuée aux émotions négatives, tristesse et peur, mais la sensibilité aux émotions positives demeurait stable. Les changements pour la tristesse et la peur étaient significatifs par rapport aux adultes d'âge moyen. Le nombre d'années d'entraînement musical corrélait avec une sensibilité accrue aux émotions musicales. L'âge et l'entraînement étaient indépendants des différences de capacités cognitives générales, ce qui suggère qu'ils sont antérieurs à ces dernières. Dans l'étude 3, deux femmes ont enregistré une liste de phrases en portugais et de pseudo-phrases, en variant la prosodie pour exprimer la neutralité ainsi que six émotions: la colère, le dégoût, la peur, le joie, la tristesse et la surprise. Les profils acoustiques spécifiques à chaque émotion étaient cohérents avec les descriptions faites pour d'autres langues. Les émotions étaient reconnues avec justesse dans les phrases (75% de réponses correctes) et dans les pseudo-phrases (71%). Ces stimuli étaient inclus dans une base de données (190 phrases et 178 pseudo-phrases) que nous avons rendue accessible aux communautés cliniques et de recherches intéressées. Dans les études 4 et 5, nous avons déterminé le degré auquel la reconnaissance des émotions dans la parole et la musique font appel à des mécanismes communs. Les résultats suggèrent qu'à la fois des mécanismes communs et des mécanismes spécifiques à chaque domaine sont engagés. Dans l'étude 4, nous avons cherché à savoir si l'entraînement dans un domaine, la musique, est associé à des bénéfices dans l'autre domaine, la parole. Des musiciens et des auditeurs non entraînés étaient comparés dans la reconnaissance de la neutralité et des six émotions à travers la parole. Nous avons trouvé un effet de transfert robuste d'un domaine à l'autre: les musiciens reconnaissaient avec plus de justesse que les auditeurs non entraînés l'ensemble des émotions. Ceci indique que des mécanismes sont en partie partagés. Dans l'étude 5, nous avons implémenté un design comparatif entre la parole et la musique dans la maladie de Parkinson. Est-ce que Parkinson produit des profils similaires de déficits dans les deux domaines? Une dissociation a été trouvée: les patients montraient un déficit de reconnaissance des émotions positives dans la musique, joie et sérénité, mais leur reconnaissance dans la musique était normale. Les déficits pour la musique ne peuvent pas être expliqués par des défauts perceptuels et cognitifs; pour la prosodie de la parole, la reconnaissance des émotions était fortement associée avec le dysfonctionnement exécutif des patients. Cette dissociation

neuropsychologique est une preuve que des mécanismes spécifiques à chaque domaine peuvent être activés par la parole et la musique. Ensemble, ces études font progresser nos connaissance sur les différences individuelles en reconnaissance des émotions dans la parole et la musique, nous fournissent un outil pour la recherche sur les émotions vocales, et participent au débat sur les relations entre la musique et le langage. D'un point de vue clinique, nos résultats contribuent à la compréhension de la phénoménologie de Parkinson.

ABSTRACT

Speech and music are two powerful means of emotional communication. This thesis investigates how we perceive emotions in speech prosody, the *tone of voice*, and in instrumental music. We pursue three main goals in a series of studies: to determine how emotion recognition in music is modulated by ageing and musical training (studies 1 and 2); to provide a database of speech stimuli in Portuguese for research in vocal emotions (study 3); and to examine the hypothesis that emotion recognition in speech and music engages shared neurocognitive mechanisms (studies 4 and 5). Forced-choice and rating tasks were employed. In studies 1 and 2, we tested adults varying in age and musical training in the recognition of happiness, peacefulness, sadness and fear in music excerpts. Advancing age was associated with decreased sensitivity to the negative emotions, sadness and fear, but sensitivity to the positive ones remained stable. Changes for sadness and fear were significant from middle-age onwards. The number of years of musical training correlated with increased sensitivity to musical emotions. Ageing and training effects were independent of differences in general cognitive abilities, suggesting that they are primary in origin. In study 3, two female speakers recorded a list Portuguese sentences and pseudo-sentences varying prosody to express neutrality and six emotions: anger, disgust, fear, happiness, sadness and surprise. The emotion-specific acoustic profiles were consistent with descriptions for other languages. Emotions were recognized accurately both in sentences (75% correct) and in pseudo-sentences (71%). These stimuli were included in a database (190 sentences and 178 pseudo-sentences), which we made available to interested research and clinical communities. In studies 4 and 5 we determined the extent to which emotion recognition in speech and music recruit shared mechanisms. Results suggest that both shared and domain-specific mechanisms are engaged. In study 4, we investigated whether training in one domain, music, is associated with benefits in the other domain, speech. Musicians and untrained listeners were compared in the recognition of neutrality and six emotions in speech. We found a robust cross-domain transfer effect: musicians had higher accuracy than untrained listeners for all emotions. This indicates that mechanisms are at least partly shared. In study 5, we implemented a comparative design between speech and music in Parkinson's disease (PD). Does PD produce a similar profile of impairments in both domains? A dissociation was found: patients had impaired recognition of positive emotions in music, happiness and peacefulness, but were normal in speech; they had impaired recognition of sadness in speech, but were normal in music. Impairments for music cannot be accounted for by perceptual and cognitive defects; for speech prosody, emotion recognition was strongly associated with the patients' executive dysfunction. This neuropsychological dissociation is evidence that domain-specific mechanisms may be engaged by speech and music. Altogether, these studies advance new knowledge on individual differences in emotion recognition in speech and music, provide a tool for research on vocal emotions, and add to the debate on the music-language relations. From a clinical standpoint, our results contribute to the understanding of the phenomenology of PD.

ACKNOWLEDGMENTS

This thesis would not exist without the help and support of many people. It is a pleasure to express my deep gratitude to all of them. Professor São Luís Castro, my supervisor, has guided, encouraged and criticized me in constructive ways over the last years. She has been determinant in shaping my development as a researcher. I thank her for the collaborative work, for the learning opportunities she provided me with, as well as for creating favourable practical conditions to make this research possible. I am grateful to the Speech Laboratory members for the innumerable and stimulating scientific discussions, for the technical and practical support throughout the journey of this work, and more than anything for their friendship and personal encouragement. Thank you so much Selene Vicente, Rui Alves, Tânia Fernandes, Lénia Carvalhais, Manuela Cameirão, Carlos Silva Pereira, Susana Silva, Manuela Santos, Teresa Limpo, Marisa Filipe, Maria Luís Almeida, Susana Carvalho, Diana Pereira, Pedro Chaves, Nuno Sousa and Paulo Branco. Maria Luís Almeida and Susana Carvalho, undergraduate research assistants, gave an important contribution concerning stimulus preparation and recording, recruitment of participants, and data collection. Mariana Schöner gave her voice to record part of the emotional speech stimuli. Professor Marc D. Pell warmly received me in the Neuropragmatics and Emotion Lab at McGill University. Our discussions on the design of the study with Parkinson's disease patients were illuminating. Professor Carolina Garrett collaborated in the selection and neurological evaluation of the patients. I acknowledge her precious help and dedication to this project. The neurologists Fernanda Simões, João Massano, Maria José Rosas, Miguel Gago and Sara Vieira also helped in selecting and evaluating patients. Professors Isabel Rocha Pinto and Patrício Costa provided valuable comments and suggestions on statistical issues. Susana Silva gave an important contribution by helping in the acoustic and structural analyses of speech and musical stimuli. Professors Anders Flykt and Arne Öhman kindly made available and gave permission for the utilization of the Karolinska Directed Emotional Faces database. Dr. Bruno Verschuere provided normative data on this database. Universidade Sénior das Antas and Universidade Sénior da Foz let us recruit their students as research participants and generously allowed us to collect data in their facilities. The studies presented here benefited very much from the critical comments and expert knowledge of many reviewers and journal editors. I am indebted to all research participants: students of different ages, musicians, and patients with Parkinson's disease and their families. Without them this thesis really would not have been possible. They made the process of data collection a greatly personally rewarding experience. I am also in debt to the many people who helped in the difficult and time-consuming process of finding and recruiting participants. I thank Angélica Relvas, Gaëtan Cousin, Manuela Cameirão and Paulo Santiago for their suggestions and comments on an earlier version of the text.

My deepest acknowledgment goes to my family and to my friends – the old friends and the ones I was lucky enough to meet during this journey. They give personal meaning to my work. I specially thank my parents and my sister. Words cannot describe my gratitude for their unconditional love and support.

Porto, summer 2011

ABBREVIATIONS

ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
AQ	The Autism-Spectrum Quotient questionnaire
dB	Decibel
e.g.	*Exempli gratia,* for example
ERP	Event-related potential
et al.	*Et alii,* and others
F_0	Fundamental frequency
fMRI	Functional Magnetic Resonance Imaging
Hu	Unbiased hit rate
i.e.	*Id est,* that is
MBEA	Montreal Battery of Evaluation of Amusia
MMSE	Mini-Mental State Examination
MOCA	Montreal Cognitive Assessment
ms	Milliseconds
ns	Non-significant
PD	Parkinson’s disease
PET	Positron Emission Tomography
RT	Reaction time
s	Seconds
SD	Standard deviation
TIPI	Ten-Item Personality Inventory
UPDRS	Unified Parkinson’s Disease Rating Scale
WAIS-III	Wechsler Adult Intelligence Scale, version III

LIST OF PAPERS

The work presented in this thesis is an expanded and updated version of five scientific articles which were submitted for publication in international peer-reviewed journals (same order as in the thesis):

1. Lima, C. F., & Castro, S. L. (2011). Emotion recognition in music changes across the adult life span. *Cognition and Emotion, 25,* 585-598. doi:10.1080/02699931.2010.502449 [Chapter VI, Study 1]

2. Castro, S. L., & Lima, C. F. (submitted). Emotion recognition in music is robust, and yet variable: The impact of age and musical expertise. [Chapter VI, Study 2]

3. Castro, S. L., & Lima, C. F. (2010). Recognizing emotions in spoken language: A validated set of Portuguese sentences and pseudo-sentences for research on emotional prosody. *Behavior Research Methods, 42*, 74-81. doi:10.3758/BRM.42.1.74 [Chapter VII]

4. Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion, 11,* 1021-1031. doi:10.1037/a002452 [Chapter VIII]

5. Lima, C. F., Garrett, C., & Castro, S. L. (submitted). Perceiving emotions in music and in speech prosody is dissociated in Parkinson's disease. [Chapter IX]

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ENERGIA E ÉTICA

Sei isto: a minha energia está canalizada
Para a palavra fazer, gosto da ideia da construção
E o que dela existe nos movimentos normais.
Agrada-me a palavra engenharia e o que ela
Representa: não saias de um sítio sem deixares algo
Atrás de ti. Dirijo-me apenas às coisas que me excitam
Positivamente e me levam a fazer outras coisas, dirijo-me
Às pessoas de que gosto, nunca às de que não gosto;
Sempre me pareceu insensato que se pare,
Nem que por um momento, de admirar, há
Sempre actos e coisas que nos ajudam
neste cálculo infernal da distância entre o dia de hoje
e a nossa morte. E qualquer pessoa dar um passo que seja
em direcção ao que não aprecia, para insultar, ou derrubar,
parece-me brutal perda de tempo, uma falha grave
no órgão de admirar o mundo
(deves combater uma ou duas vezes na vida,
se combateres duzentas vezes
é porque os combates são fracos).
Não sei pois como viver. O que li e vi
Serve-me apenas para ser mais lúcido, não
Para ser melhor pessoa. Adquiri esta regra (ou nasci com ela):
- e é talvez uma moral -
mover-me apenas em direcção ao que gosto.
Se o prédio alto, escuro, feio
me impede de ver o sol, não fico a insultá-lo, não
moverei um dedo para o deitar abaixo:
contorno sim os edifícios necessários
até chegar ao espaço de onde possa receber aquilo que
quero. Se chegar lá de noite, montarei acampamento.

Gonçalo M. Tavares, 2004

GENERAL INTRODUCTION

Humans are hard-wired to connect with the sounds of voice and music. Words, tones of voice, laughter, sobs, screams, sighs, timbres, musical notes, melodies, rhythms, beats, all that, are building blocks of our auditory environment and mental life. Our senses are perpetually inundated by these sounds. Since ancient times in our species evolution, throughout our life span, across all societies, we are adept at making sense out of voice and music. We use them as ways to establish meaningful relations with the others and with ourselves. They are multifaceted and powerful communication tools, able to convey a multitude of kinds of information. Voice and music have also a startling capacity to move us. If you think of your own experience, it will be probably easy to remember of a situation in which you felt strongly aroused by a music that you particularly enjoyed, or deeply moved while listening to someone's speech. This thesis is about a key common feature of voice and music – the communication of emotions. The focus is on how we perceive emotions communicated by speech prosody, the *tone of voice*, and by instrumental music. It comes naturally to us to tell whether someone is angry or scared by the sound of her or his voice, or whether a piece of music is happy or sad. The apparent ease of these operations, though, belies the complexity of the neurocognitive mechanisms that are involved in transforming streams of sound – pressure waves producing vibrations of the basilar membrane in the ear – into meaningful subjective representations of emotion. Despite recent advances in cognitive and neuroscientific research on this topic, many questions remain open. Herein, we set out investigate some of them in a series of studies using behavioural and neuropsychological approaches.

We pursue three main goals. The *first* one is to examine how the recognition of emotions in music changes with advancing age and as a function of musical training. The idea that ageing and musical training modulate many cognitive, neural and socio-emotional functions is firmly entrenched (e.g., Habib & Besson, 2009; Hedden & Gabrieli, 2004; Salthouse, 2009; Samanez-Larkin & Carstensen, 2011), but their impact on musical emotions is still poorly understood. The *second* goal is to devise and validate a database of speech stimuli in Portuguese for research on vocal emotions. Analogous materials have been developed for different languages (Burkhardt, Paeschke, Rolfes, Sendlmeier, & Weiss, 2005; Makarova & Petrushin, 2002; Pell, 2002; Ross, Thompson, & Yenkosky, 1997; Staroniewicz & Majewski, 2009; Wu, Yang, Wu, & Li, 2006). They may be useful for studies with native speakers of the stimuli's language, for cross-language research, as well as for the assessment of pragmatic skills in clinical settings. The *third* goal is to examine the extent to which emotion processing in speech prosody and in music engages shared neurocognitive mechanisms. This hypothesis has been highly debated in the last years (e.g., Juslin & Laukka, 2003; Juslin, Liljeström, Västfjäll, & Lundqvist, 2010; Juslin & Västfjäll, 2008; Nieminen, Istók, Brattico, Tervaniemi, & Huotilainen, 2011; Patel, 2008b; Peretz, 2010), but there is a dearth of empirical research on it. We investigate this issue in two ways. One consists of examining possible transfer effects from one domain to the other. Specifically, we determine whether expertise in music is associated with enhanced ability to recognize emotions in speech prosody. The other approach consists of using a comparative design between the two domains in a neuropsychological study. We determine whether Parkinson's disease (PD), a neurodegenerative disorder involving damage in the basal ganglia, leads to a similar profile of impairments in emotion recognition in speech and music.

The thesis is divided in two parts. Part 1 comprises a review of the literature. This review covers emotion theories (Chapter I), the acoustic and neurocognitive foundations of vocal and musical emotions (Chapters II and III), and a discussion on the parallels between both domains concerning emotion expression and processing (Chapter IV). In the end of this part the goals of the empirical studies are described (Chapter V). Part 2 comprises the empirical chapters. The first one (Chapter VI) presents two studies in which listeners of different ages and varying in musical training are examined in the recognition of emotions in short instrumental music excerpts. The second empirical chapter (Chapter VII) describes the development and recording of a set of Portuguese speech stimuli expressing different emotions, as well as the acoustic and perceptual validation of these stimuli. The third empirical chapter (Chapter VIII) presents a study in which musicians are compared with musically untrained listeners in the recognition of emotions in speech prosody. The last empirical chapter (Chapter IX) corresponds to the neuropsychological study. PD patients are compared with healthy matched controls in the recognition of emotions in speech prosody and in music. Each empirical chapter follows the typical structure of a journal article, including the sections introduction, method, results, and discussion. A brief conclusion is also provided. In the end of Part II (Chapter X), which coincides with the end of the thesis, we summarize the findings and conclusions of the empirical studies and discuss avenues for future research.

PART 1

SOUNDS WITH EMOTIONAL MEANING:
EMOTION THEORIES, EMOTIONAL SPEECH PROSODY, AND MUSICAL EMOTIONS

CHAPTER I

Emotion Theories

It is common sense that emotions play a pivotal role in human existence. They accompany almost every significant event in our lives, and shape strongly the way we think, feel and act. What is an emotion? This question was the title of one of the most famous William James' (1884) papers. Yet, more than one century afterwards a definitive answer is still lacking, even though the study of emotions is currently highly popular in cognitive neuroscience. They are notoriously hard to define and there is no general consensus regarding even what should count or not as an instance of emotion. However, experts do agree to a significant extent with respect to the general characteristics of emotions, as well as about the functions they serve (Izard, 2009). Emotions correspond to relatively brief and intense reactions to potentially important changes in our external or internal environment (e.g., stimuli or events representing subjective challenges or opportunities). They typically involve a number of sub-components – subjective feeling, cognitive changes, physiological arousal, expressive behaviour, action tendencies, and regulation – that are more or less synchronized. They are coordinated reactions intended to promote adaptive behaviours in response to a changing environment (e.g., Davidson, Scherer, & Goldsmith, 2002; Ekman, 1999; Juslin & Sloboda, 2010; Levenson, 1999; Scherer, 2005; Smith & Lazarus, 1990). Although the mechanisms underlying emotional reactions may be of several kinds, in everyday life emotions are often the result of evaluating (cognitive appraisal) an event as relevant to subjective intentions, goals, motives and concerns (Scherer, 2005; Sloboda & Juslin, 2010). The two theoretical frameworks that have most strongly influenced research on emotions in the last decades are discrete categories and dimensional emotion theories, which we outline below.

1.1. DISCRETE EMOTION THEORIES

These perspectives postulate that we experience emotions as discrete categories that are distinct from each other (e.g., happiness, sadness, anger, fear). The theory of *basic emotions* is perhaps the most prominent of categorical approaches to emotion (e.g., Ekman, 1992a, 1992b, 1999; Izard, 2009). It stems from evolutionary views and was shaped by Charles Darwin's work on biology (Darwin, 1872/2009). Emotions are conceptualized as a limited number of innate and biologically evolved functions, from which all emotional states can be derived. Each basic emotion category is thought to have its own adaptive significance and to have played an important and specific role for survival during our ancestral past. It is also presumed that basic emotions are found in all cultures, emerge early in ontogenetic development, are experienced as unique feelings, involve distinct cognitive appraisals, physiological activity and action tendencies, can be inferred in other species, and are associated with unique expressive profiles in facial and vocal expressions. There is no agreement as for the set of categories that correspond to basic emotions, but they usually include happiness, sadness, anger, fear, disgust and surprise (Power & Dalgleish, 2008). Recently, Levenson (2011) suggested that relief/contentment, interest and love are also part of the list. Support for this theory comes from studies confirming that basic emotions can be expressed and recognized universally (Ekman, 1992b, 1994; Elfenbein & Ambady, 2002; Sauter, Eisner, Ekman, & Scott, 2010; Scherer, Banse, & Wallbott, 2001), may rely on independent structural and functional brain systems (Calder et al., 2003; Levenson, 2011; Phan, Wager, Taylor, & Liberzon, 2002; Scott, Sauter, & McGettigan, 2010; Vytal & Hamann, 2010), and are perceived as discrete categories in categorical perception experiments including continua of expressions (Laukka, 2005; Sauter, Guen, & Haun, in press). A common criticism is that such a limited number of emotions cannot do justice to the complexity and variety of our emotional life (Sloboda & Juslin, 2010). More recent *component process* theories also conceive emotions as discrete categories, but instead of defining a limited list of possible categories, they emphasize the variability of emotional states, which are hypothesized to be produced by different types of appraisal patterns (e.g., Scherer, 2003, 2005). Within this framework, emotions are driven by a series of subjective appraisals for internal or external events concerning important information (e.g., novelty, coping potential, conduciveness, intrinsic pleasantness). The experienced emotional categories can be as many and fine-grained as the pattern of appraisals they result from. By highlighting that emotions derive from subjective appraisals, this approach is able to explain why the same event can elicit different reactions in different people. Alternatively, *psychological constructionist* models put forward that the experience of discrete emotions depends on learned conceptual knowledge for emotion categories, rather than on biologically determined programs. According to Barrett and colleagues (Barrett, 2006a, 2006b, 2009; Lindquist & Barrett, 2010), we experience discrete emotions because we automatically and implicitly use a conceptual system to categorize instances of more basic core affective states. Core affect would correspond to a psychological primitive with valence and arousal properties, from which specific emotions could be experienced (only) through a conceptual act of categorization. The conceptual system for emotion would consist of embodied conceptualizations for emotion categories, which would be context-dependent, learned through experience, and highly dependent on language, i.e., on emotion words. Such a perspective can be considered as an interesting compromise between discrete and dimensional approaches to emotion (see below).

1.2. DIMENSIONAL EMOTION THEORIES

Dimensional approaches consider that emotions are continuous entities defined by their placement along a small number of broad core affective dimensions. These perspectives tend to concentrate on how people subjectively represent emotional states (feelings) in terms of their internal affective space. The number of affective dimensions varies across models (e.g., Sloboda & Juslin, 2010), but the most influential is James Russell's two-dimensional *circumplex model* of affect (Posner, Russell, & Peterson, 2005; Russell, 1980, 2003). This model comprises a circular structure, referred to as the circumplex, placed on two orthogonal lines featuring the two affective dimensions: pleasure or valence, and arousal. These two dimensions are presumed to be independent and bipolar neurophysiological systems, and this has received empirical support (e.g., Anderson et al., 2003; Colibazzi et al., 2010). They are also thought to account for all differences among emotional states – different emotional states correspond to different places in the circumplex affective space. Emotions can thus be conceptualized as linear combinations of the two core dimensions, or as different degrees of valence and arousal. The valence dimension represents the subjective pleasantness of the emotional state, with states like glad, happy, pleased and serene at the pleasant extreme, and miserable, gloomy, sad and depressed at the unpleasant extreme. The arousal dimension refers to a sense of mobilization and energy of the emotion, with states like alarmed, aroused, excited and tense at the high arousal extreme (activation), and sleepy, tired, droopy and bored at the low arousal extreme (deactivation). While discrete emotion approaches conceptualize emotional states as mutually exclusive categories with clear-cut boundaries, this model conceives them as fuzzy, ambiguous and overlapping experiences, more or less related to each other depending on their distance in the affective space. For instance, calm and relaxed would be highly correlated emotional states, as they are both pleasant and with low arousal – they are placed closely in the circumplex. Happy and sad, on the other hand, may be conceived of as bipolar emotional states, as they correlate inversely both in terms of valence and arousal – they are in opposite locations of the circle.

Discrete categories and dimensional theories of emotion tend to correspond to different research traditions, with proponents of both types of approaches carrying out different kinds of studies, resorting to different methods, and arguing fiercely with each other. Notwithstanding, albeit they are theoretically distinct, these approaches could be considered, not as mutually exclusive, but rather as descriptions focusing on different dimensions or levels of the same phenomena (Sauter, 2006).

1.3. EMOTION EXPRESSION AND RECOGNITION

A crucial aspect of emotions is that they involve an expressive component, that is, they are displayed through different means, such as body postures, facial expressions, voice and language (Darwin, 1872/2009; de Gelder & Stock, 2011; Keltner & Ekman, 2002). Humans have a rich repertoire of emotion signals that also include odours, music and literature, for instance (Delplanque et al., 2009; Johnson-Laird & Oatley, 2010; Zentner, Grandjean, & Scherer, 2008). Expressive features allow to communicate emotional states to others, a function considered of paramount importance for social functioning. Emotional expressions broadcast valuable information about the person's state, behavioural intentions, relationship status, as well as about objects and events in the social environment (e.g., Matsumoto, Keltner, Shiota, O'Sullivan, & Frank, 2010). Such kind of information will exert direct influence on others' behaviour. Expressions can also regulate social life by eliciting emotional responses in others (Russell, Bachorowski, & Fernández-Dols, 2003). From the side of the perceiver, being able to accurately and quickly interpret emotion signals is also highly adaptive. This ability is a crucial part of our emotional competence (Scherer & Scherer, in press) and correlates with important indicators of personal and social functioning (e.g., mental health; social and work-related competencies; Hall, Andrzejewski, & Yopchick, 2009). Arguably, the adaptive pressures that led to the development of emotions also favoured the development of mechanisms for expression and recognition of emotions (Hauser & McDermott, 2003), in what constitutes a fundamental dimension of social communication. Most research on emotion expression and recognition has been conducted on facial expressions. Other modalities, namely in the auditory domain, remain much less well understood (e.g., de Gelder, 2010).

How emotion recognition processes are conceived vary across theoretical perspectives. For instance, from the standpoint of basic emotions theory, recognizing an emotion corresponds to a simple, automatic and undemanding process: each biologically evolved emotion has distinct and universal expressive features (e.g., specific facial muscle movements), and our perceptual systems are hard-wired to detect them (e.g., Ekman, 1994). According to more recent simulation models, recognizing an emotion in others involves the simulation of the corresponding emotional state, setting into action some of the same mechanisms that underlie the actual emotional experience. Inferring emotions in another person would thus be an empathetic process, in the sense that it would implicate feeling the emotion of the other person, at least in part (Adolphs, 2006, 2010). The constructionist model by Barrett and colleagues (Barrett, 2006a, 2006b, 2009; Barrett & Kensinger, 2010; Lindquist & Barrett, 2010) holds that emotion recognition involves bringing embodied conceptual knowledge to bear in a context sensitive manner. For example, recognizing anger in a facial expression would result from categorizing that facial expression as an instance of *anger.* This categorization act would be a function of what we know about the specific emotion category (emotion concepts) and about the context (e.g., the social situation in which the face appears). On the other hand, dimensional theories highlight that we perceive emotions in terms of affective dimensions such as valence and arousal (Russell et al., 2003).

In this thesis we adopt a discrete categorical approach to emotions, but the role of broader affective dimensions in emotion processing, notably valence, is taken into consideration as well. We focus on the recognition of emotional expressions in auditory signals, concretely in voice/speech and in music. In the next chapter we review literature about vocal emotions.

CHAPTER II

Speech Prosody and Emotion

The human voice can be considered the most important sound of our auditory environment. First and foremost, it is the carrier of speech – we can speak and understand linguistic messages, and this ability defines us as humans. In parallel to speech, though, the voice expresses a wealth of socially relevant information. It communicates information about the speaker's gender, identity, age, size, health, attractiveness, trustworthiness, dominance, feelings and mood, even when speech content is not available (for instance because it is a unknown language, or because the vocalization is nonverbal, such as a laughter or a sigh). Thus, the voice is very much like an "auditory face" (Belin, Bestelmeyer, Latinus, & Watson, in press).

We are endowed with abilities to extract paralinguistic information from voices, in what reflects a more primitive and universal form of communication than language itself – vocalizations have been used by many species for millions of years before language emerged (Belin, Fecteau, & Bédard, 2004; Bruckert et al., 2010; Latinus & Belin, 2011). The expression of emotions through vocal cues has particular relevance for social interactions. The acoustic, psychological and neural mechanisms involved in the processing of vocal emotions are outlined in the following pages. The review focuses chiefly on the type of emotional vocal cue that we examine in the empirical part of this thesis, speech prosody. Briefly put, speech prosody refers to the musical aspects of speech, the *tone of voice*, which often convey information about the speakers' emotions.

2.1. VOICE AS A MEANS OF EMOTIONAL EXPRESSION

The role of voice for emotional expression has long been recognized, as indicated by the treatises of the topic found in Greek and Roman manuals on rhetoric (Scherer, 2003), and by Darwin's work on evolutionary biology in the 19th century (Darwin, 1872/2009). An interesting issue is that vocal expression may be the most phylogenetically continuous of all forms of emotional communication. A large number of non-human species use vocalizations to communicate motivational and emotional states. Vocalizations are especially important in social mammals, whose life is based on complex and cooperative interactions between individuals. There are similarities across many species of mammals, including humans, in the neural control of voice production, in voice production mechanisms, as well as in the acoustic characteristics of the expressions that signal different states (Fichtel, Hammerschmidt, & Jürgens, 2001; Grandjean, Bänziger, & Scherer, 2006; Juslin & Laukka, 2003; Scherer, 1995; Scherer, Johnstone, & Klasmeyer, 2003). Because the expression of vocal emotions serves a highly adaptive function for survival in social environments, it is likely that is was shaped by natural selection – vocalizations usually correspond to the communication of biologically important events (e.g., Hauser & McDermott, 2003). Examples of non-human vocal expressions are calls of warning, threat, submissive or affiliative states, mating, attention-getting, desire for social contact and companionship (Juslin & Laukka, 2003). The richness of the expressive repertoire varies across species depending on the complexity and differentiation of the sound-producing apparatus, with some species producing only a few innate vocal behaviours (e.g., frogs), and others exhibiting a varied repertoire of voluntarily controlled productions (e.g., non-human primates). In humans, the formalized and abstract systems of language acquired special prominence as a form of communication, namely for the expression of emotional meanings (through semantics). Notwithstanding, more ancient paralinguistic emotional expressions are also present in our vocal behaviours under different forms. We often produce purely emotive nonverbal vocalizations, such as laughter, sighs, sobs or screams. These are perceived rapidly (Sauter & Eimer, 2009), accurately (Sauter, Eisner, Calder, et al., 2010) and across cultures (Sauter, Eisner, Ekman, et al., 2010). In addition to nonverbal expressions, we routinely produce short interjections with affective connotation, such as "wow", "Ah", "yuck" or "oh" (Belin, Fillion-Bilodeau, & Gosselin, 2008; Scherer, 1995).

Emotions are also expressed as an integrant part of the speech signal proper via speech prosody cues – suprasegmental vocal modulations that occur in the course of a spoken utterance. It is well established that emotional states are associated with changes in the way we speak: emotions involve psychophysiological changes, for instance in heart rate, blood flow and muscle tension, which affect respiration, vocal fold vibration and articulation, in such a way that all these influence the acoustic characteristics of certain voice features while we speak. For example, higher emotional arousal increases laryngeal tension and subglottal pressure, thereby increasing the intensity and changing the timbre of the voice. Thus, the suprasegmental signals that ride atop of speech provide diagnostic information concerning our states (Banse & Scherer, 1996; Scherer, 1986, 2003; Schirmer & Kotz, 2006).

2.2. ACOUSTIC CUES OF EMOTION IN VOICE

Speech prosody comprises different acoustic cues: fundamental frequency (F_0), intensity, tempo, rhythm and voice quality (e.g., Grandjean et al., 2006; Schirmer & Kotz, 2006). F_0 reflects the frequency of vibration of the vocal folds during phonation and is subjectively perceived as voice pitch. Variations in the level, range and contour of F_0 during the utterance are important cues of emotion. Intensity corresponds to the vocal energy and reflects the effort required to produce speech; it is subjectively perceived as loudness. Tempo corresponds to the number of phonemic segments per time unit, and rhythm to the structure of F_0 accents, intensity peaks and distribution of pauses in the utterance. Finally, voice quality reflects the shape of the vocal tract, which can be modified, for instance, by the muscle tension in the larynx. Voice quality is acoustically characterized by the distribution of energy in the frequency spectrum, and is subjectively perceived as voice timbre (e.g., roughness and sharpness). The relative amount of energy in the high vs. low frequency region of the spectrum is highly informative to differentiate emotions (as the amount of high-frequency energy increases, the voice sounds more sharp and less soft). Modulations in these prosodic cues are associated with the communication of different emotional states, as we discuss below.

An important notion is that changes in prosody can be spontaneous reflections of psychophysiological emotion-related states, but they can also be influenced by voluntary modulations. With regard to this issue, Scherer and colleagues have distinguished between *push effects* and *pull effects* in vocal expression (e.g., Banse & Scherer, 1996; Scherer, 1986; Scherer, 2003). Push effects refer to the direct impact of psychophysiological mechanisms involved in emotional episodes over the voice production system, which would be innately determined to a large extent. Pull effects, on the other hand, reflect the fact that prosody may incorporate sociocultural conventions and display rules, and can be strategically sculpted by our communicative intentions, independently of the presence of a real emotional state. During everyday life social interactions, the expression of emotions in prosody is often a joint product of push and pull effects.

2.3. EMOTION COMMUNICATION THROUGH SPEECH PROSODY

As a communicative phenomenon, emotional prosody can be theoretically framed in the context of a modified version of the Brunswik's functional lens model, which combines the analysis of expression and perception of affect signals (Brunswik, 1956; Grandjean et al., 2006; Juslin & Laukka, 2003; Scherer, 2003; Scherer et al., 2003). The process starts with the speaker expressing, or encoding, an emotional state through several acoustic cues, which can be objectively measured in the speech waveform. These objective parameters are called *distal cues,* as they are distant from the listener. Then they are transmitted, as part of the speech signal, and perceived by the auditory system of the listener, who forms a proximal percept. The perceived (subjective) cues are called *proximal cues*. The decoding process consists of combining and integrating the different internalized proximal cues in order to make a subjective inference regarding the speakers attitudes and emotions. Proximal cues are based on the distal ones, but they might suffer influences from the transmission channel (e.g., distance; noise) and from the characteristics of the perceptual system as well (e.g., selective enhancement of specific frequency bands). Within this framework, a functionally valid communication process occurs when the subjective judgments provided by the listeners correspond to the criteria for the state of the speaker (e.g., a specific emotional state, such as anger) with agreement rates above chance level. Over the last decades, a large number of empirical studies have been conducted to understand both expression and perception of emotional prosody.

EXPRESSION | Studies that focus on expression have attempted to determine how differentiated patterns of acoustic voice cues are reliable indicators of specific emotional states. To that end, researchers collect vocal expressions corresponding to different emotional states and then measure them for acoustic attributes in order to examine associations between cue patterning and emotions. Classically measured acoustic cues are F_0 (mean, variability and range), intensity and speech rate, although some studies include additional measures, such as jitter, high-frequency energy, formant frequency and F_0 contours (e.g., Juslin & Laukka, 2001). An important methodological question concerns how to collect expressions for posterior analyses. Three major methods have been used, each of these with advantages and drawbacks (e.g., Scherer, 2003). One involves *natural vocal expressions*, that is, materials recorded during naturally occurring emotion episodes, for instance dangerous flight situations or journalists reporting emotionally charged events. This is a highly ecological method, but it has serious disadvantages: recordings are often very brief and obtained from a single speaker, they frequently suffer from bad recording quality, and experimental control is very limited. Another method consists of *inducing emotional states* experimentally, for example through films, and then recording speech samples while the speakers are under these states. This approach facilitates experimental control over the vocal materials produced, but it is difficult to induce strong and well-differentiated emotional states, such that the resulting recordings frequently express weak affect. Furthermore, one cannot assume that the same emotion elicitor procedure will produce the same emotional responses in all speakers. By far the most frequently employed method makes use of *simulated actor portrayals*, in which actors or untrained speakers are asked to pose expressions on the basis of emotion labels and/or scenarios. In a typical situation, they read the same verbal materials (e.g., sentences with neutral semantic content) while varying prosody to express different emotions. It has often been argued that this procedure may yield unnatural and stereotypical expressions, but acoustic convergence has been found between these expressions and the so-called natural ones (Juslin & Laukka, 2001). Acted portrayals contributed greatly for what is currently known about emotion communication through prosody (for a review of pros and cons of acted portrayals, Bänziger & Scherer, 2007).

Acoustic analyses have revealed that there are different acoustic profiles for some emotion categories, thus confirming that voice inflections are informative about the speaker's state (e.g., Banse & Scherer, 1996; Hammerschmidt & Jürgens, 2007; Juslin & Laukka, 2001; Paulmann, Pell, & Kotz, 2008b; Scherer, Banse, Wallbott, & Goldbeck, 1991). In a review of the available evidence, Scherer and colleagues (Banse & Scherer, 1996; Pittam & Scherer, 1993; Scherer, 2003) observed that anger is generally associated with increased mean F_0 and voice intensity, although other features might also be found, such as increased variability and range of F_0, increased high-frequency energy, faster rate of articulation and downward-directed F_0 contours; fear is associated with increased mean and range of F_0, increased high-frequency energy and faster rate of articulation; sadness is characterized by decreased mean and range of F_0, decreased voice intensity and high-frequency energy, slower rate of articulation and downward-directed F_0 contours; and joy/happiness appears to be associated with increased mean, range and variability of F_0, increased voice intensity, and increased high-frequency energy and faster rate of articulation. These findings are summarized in Table 1.

Table 1. Summary of empirical data on patterns of acoustic cues for anger, fear, sadness and joy in vocal expression. Adapted from Scherer (2003, p. 233).

Acoustic cue	Anger	Fear	Sadness	Joy
Intensity	↗	↗	↘	↗
F_0 floor/mean	↗	↗	↘	↗
F_0 variability	↗		↘	↗
F_0 range	↗	↗(↘)	↘	↗
Sentence contours	↘		↘	
High-frequency energy	↗	↗	↘	(↗)
Speech and articulation rate	↗	↗	↘	(↗)

PERCEPTION | Research on the perception of emotional prosody, on the other hand, has examined the extent to which listeners are able to accurately infer the emotional state of a speaker from speech samples. In most studies, pre-recorded vocal stimuli – usually posed actor portrayals – expressing a number of different emotions are presented to a group of listeners. They frequently perform a forced-choice emotion recognition task, that is, they are asked to select the emotion that best describes the stimulus from a list of emotion labels. Then the percentage of stimuli correctly recognized per emotion is computed (percentage of matches between the selected emotion and the criterion for the stimulus' emotion). It has been repeatedly shown that listeners are adept at perceiving emotions in prosody, with accuracy rates about four to five times higher than what would be expected if responses were given in a random manner (e.g., Juslin & Laukka, 2003; Pell, Paulmann, Dara, Alasseri, & Kotz, 2009; Scherer, 2003; Scherer et al., 2003). For example, Banse and Scherer (1996) found a global identification accuracy of 48% for 14 different emotion categories: hot anger, cold anger, panic fear, anxiety, desperation, sadness, elation, happiness, interest, boredom, shame, pride, disgust and contempt. Speech samples in this study consisted of pseudo-sentences composed of phonemes from several Indo-European languages, which were recorded by actors to express the intended emotions. Juslin and Laukka (2001) found an accuracy of 56% for stimuli consisting of sentences with neutral semantic content, which were recorded to express five emotions: anger, disgust, fear, happiness and sadness. Adolphs, Damasio and Tranel (2002) found a recognition accuracy of 81%, also for sentences with neutral semantic content expressing five emotions: anger, fear, happiness, sadness and surprise.

Pell (2002) obtained an accuracy of 78% for pseudo-sentences that resembled English, which expressed six emotions: anger, disgust, happiness, pleasant surprise, sadness and neutrality. Altogether, these studies are strong evidence that we can perceive specific emotions from voice cues alone, i.e., even when concurrent semantic cues are not available or are emotionally neutral. Albeit recognition accuracy is generally high, significant differences are often detected between emotions. Sadness and anger are usually best recognized, followed by fear and happiness (Juslin & Laukka, 2003; Thompson & Balkwill, 2006). Disgust seems to be particularly difficult to recognize in prosody (Banse & Scherer, 1996; Scherer et al., 1991), in spite of being quite well detected in nonverbal vocal expressions (Belin et al., 2008; Sauter, Eisner, Calder, et al., 2010) and in faces (e.g., Goeleven, Raedt, Leyman, & Verschuere, 2008). The differential ease with which emotions and modalities are recognized might be related to differential evolutionary pressures (e.g., Scherer, 2003).

An important issue is whether emotion recognition in prosody is governed by universal principles or determined by cultural and language-specific factors. Cross-cultural studies are pivotal to approach this question. Thompson and Balkwill (2006) examined how English-speaking listeners recognize joy, sadness, anger and fear as expressed in English, German, Chinese, Japanese and Tagalog speech samples (sentences with neutral semantic content). Recognition accuracy was highest for English stimuli, thus revealing an in-group advantage, but listeners were able to perceive emotions in the other languages well above the chance level, including in highly dissimilar non-Western languages. Consistently, Pell, Monetta, Paulmann, and Kotz (2009) found that Spanish-speaking listeners can recognize joy, sadness, anger, fear and disgust as expressed in Spanish, English, German and Arabic pseudo-sentences, even though they are better in their native language. Further evidence comes from a study by Scherer, Banse and Wallbott (2001), who compared how listeners from nine countries across Europe, America and Asia perceive anger, sadness, fear, joy, and neutrality as expressed by German actors (materials were pseudo-sentences). All listeners recognized emotions above the chance level, but accuracy rates varied across countries: generally, performance was higher when the native language of the listeners was closer to German (e.g., Dutch), and lower when it was highly dissimilar (e.g., Malay; see also Elfenbein

& Ambady, 2002). Finding that prosodic emotions can be recognized cross-culturally suggests that the acoustic cues and the way in which listeners process prosody are universal to a significant extent. On the other hand, the in-group advantage indicates that cultural factors also exert a role. In other words, the recognition of emotional prosody appears to be determined by a combination of universal and culturally specific processes. Therefore, devising speech samples for different languages and analysing how speakers with different linguistic backgrounds process emotions is of paramount importance for better understanding emotion communication through prosody.

FROM EXPRESSION TO PERCEPTION | Some studies have approached expression and perception of emotional prosody in a comprehensive manner in order to throw light on the nature of the inference rules used by the listeners during emotion recognition. Which acoustic cues (distal cues) are attended to in voice, and how do they lead to the recognition of a certain emotion? One strategy to answer this question consists of manipulating systematically specific acoustic cues (via synthesis or resynthesis) and examining how these manipulations affect emotion perception. For example, Breitenstein, Lancker and Daum (2001) manipulated prosodic stimuli for speech rate and F_0 variability. They observed that speech rate was the most potent cue for subjective responses, with slow rate being reliably associated with sadness, and fast rate with anger, fear and neutrality; F_0 variability was less potent but still determinant of responses, with reduced variability being associated with sad or neutrality, and large variability with fear, anger or happiness. Another strategy to explore the utilization of acoustic cues in emotion inferences involves measuring the acoustic properties of speech samples, and then correlating these with the subjective responses provided by the listeners. In regression analyses, Banse and Scherer (1996) showed that a significant proportion of variance in the listeners' judgments was accounted for by acoustic attributes of the stimuli, such as mean and variability of F_0, mean intensity, duration of voiced periods, relative proportion of high-frequency energy (cut-off 1000Hz) and energy drop-off in the spectrum. The proportion of explained variance was highest for hot anger, 36%, and for the other emotions it ranged from 7% to 25%. More recently, Juslin and Laukka (2001) found that subjective judgments for anger, disgust, fear, happiness and sadness could be significantly predicted by a set of nine acoustic cues in regression analyses, namely variability and floor (lower 5% of F_0 values) of F_0, contour of F_0 (ratio of rising/falling contours), voice intensity, high-frequency energy (cut-off 500Hz), speech rate, proportion of pauses, attack (voice onsets in speech) and F_1 (mean frequency and bandwidth of the first formant). The explained variance ranged from 45% for fear and happiness to 58% for anger. From the cues included in the model, no less than three yielded significant beta weights for each emotion, suggesting that listeners relied on multiple cues when making emotion inferences. Additionally, the combination of cues that predicted responses was unique for each emotion – listeners used differentiated cue sets across specific emotions. For example, to infer happiness, listeners used primarily the cues high F_0 floor, large variability of F_0, and low proportion of pauses; to infer sadness, they used high F_0 floor, small variability of F_0, low voice intensity, low F_1, and high proportion of pauses; and to infer fear, they used high F_0 floor, small variability of F_0, low high-frequency energy, fast speech rate and high proportion of pauses. These results show that it is possible to characterize the inference rules used by the listeners in emotion judgments. They also indicate that there is significant overlap between the objective acoustic parameters of vocal expressions and the subjective cues that drive emotion inferences.

2.4. NEUROCOGNITIVE BASIS OF EMOTIONAL PROSODY

A classical notion, supported by data from neuropsychological patients, is that the processing of emotional prosody is a dominant and lateralized function of the right hemisphere (e.g., Blonder, Bowers, & Heilman, 1991; Gorelick & Ross, 1987; Ross, 1981). However, now it is clear that this is an oversimplification. More detailed examinations based on different methods (neuropsychology, neuroimaging, electrophysiology) have revealed that a widespread network is involved, with current models conceptualizing emotional prosody as a multi-stage function subserved by a bilateral distributed network of interconnected cortical and subcortical brain regions (e.g., Berckmoes & Vingerhoets, 2004; Ethofer et al., 2006; Frühholz, Ceravolo, &

Grandjean, in press; Leitman, Wolf, et al., 2010; Paulmann, Ott, & Kotz, 2011; Paulmann, Pell, & Kotz, 2008a; Pell, 1998, 2006; Schirmer & Kotz, 2006; Wildgruber, Ackermann, Kreifelts, & Ethofer, 2006; Wildgruber et al., 2005). Schirmer and Kotz (2006) proposed a working model for vocal emotional processing comprising the following sub-processes: (1) sensory processing, which involves the low-level extraction and analysis of the suprasegmental acoustic cues of voice; (2) integration of meaningful acoustic cues and detection of their emotional significance; and (3) higher order cognitive-evaluative processes, which include explicit emotion judgments, as well as the integration of prosody with verbal information. For a similar conceptualization, see Wildgruber and colleagues (2006). Note that the processing of prosody is neurally segregated from syntactic and lexico-semantic operations in language comprehension – which are bound to a left-lateralized temporo-frontal network –, even though they are interactive systems (e.g., Friederici & Alter, 2004).

SENSORY PROCESSING | The suprasegmental cues of speech that carry emotional information (e.g., F_0; intensity) run from the ear to the brainstem and thalamus, and from these to the primary and secondary auditory regions in the temporal cortex, where they are analysed. Some studies have found a rightward lateralization for the encoding of these cues. This lateralization may reflect differences between the two hemispheres concerning the temporal resolution of auditory processing areas: the left hemisphere operates on a finer temporal scale, which is more appropriate to the differentiation of speech sounds at segmental levels (e.g., phonemes; syllables), whereas the right hemisphere shows lower temporal resolution, which favours the processing of the slower variations of suprasegmental speech features, particularly pitch (Ethofer et al., 2006; Scott, Blank, Rosen, & Wise, 2000; Wildgruber et al., 2005; Zatorre, 2001). Electrophysiological studies that measure event-related potentials (ERPs) are useful to track the time window when these early prosodic processes take place. It was shown that differences in frequency and intensity information modulate the amplitude of a negative wave that peaks around 100 ms after stimulus onset, a component referred to as N100. This component is generated in secondary auditory regions and does not respond to the emotional impact of the stimulus (e.g., Paulmann et al., 2011). Differences in other prosodic cues, namely duration and temporal structure, exert similarly early effects on ERPs (Schirmer & Kotz, 2006). Thus, the low-level encoding of prosody can occur within a hundred ms.

INTEGRATIVE PROCESSES AND DETECTION OF EMOTIONAL SALIENCE | After acoustic analysis, the processing of the multiple prosodic cues becomes increasingly more integrative, and emotional significance is derived. These operations implicate the right superior temporal sulcus, concretely the anterior portion (Schirmer & Kotz, 2006), though some studies detected activations in the middle (Grandjean et al., 2005; Sander et al., 2005; Wiethoff et al., 2008) and posterior portions (Beaucousin et al., 2007; Wildgruber et al., 2006). Concurrent activity in the amygdala may also play a role – the amygdala is involved in the automatic appraisal of the emotional salience of voice cues (Bach et al., 2008; Leitman, Wolf, et al., 2010; Sander et al., 2005). ERP studies showed that emotional salience can be inferred as early as 200 ms after the beginning of the stimulus: emotional speech, as compared with neutral speech, modulate the amplitude of the components P200 and Mismatch Negativity (MMN), both peaking around 200 ms after stimulus onset (Paulmann & Kotz, 2008a; Schirmer, Striano, & Friederici, 2005). Within this time window, different emotional information is extracted, such as valence, arousal, and possibly category-specific knowledge (e.g., Paulmann et al., 2011; Paulmann & Pell, 2010a).

HIGHER ORDER COGNITION | This information is then available for higher order cognitive processes that take place in frontal regions, such as explicit evaluative judgments and integration of prosody with semantics in communicative contexts. Performing explicit evaluations of vocal expressions seems to implicate the inferior frontal gyrus and the orbitofrontal cortex bilaterally. Activity in these regions is often detected when listeners are required to judge prosodic emotions explicitly (e.g., forced-choice emotion categorization), as compared to implicit tasks (e.g., gender judgments) or to a resting baseline (Bach et al., 2008; Ethofer et al., 2006; Ethofer et al., 2008; Leitman, Wolf, et al., 2010; Wildgruber et al., 2006; Wildgruber et al., 2005). Consistently, neuropsychological evidence indicates that lesions in the orbitofrontal cortex disrupt the explicit recognition of emotions, but not the early implicit stages of prosodic processing (Paulmann, Seifert, & Kotz, 2010). Moreover, Leitman and colleagues (2010) found increased activity in the inferior frontal gyrus when prosodic cues were more ambiguous, probably reflecting the recruitment of more evaluative resources to deal with the executive difficulties in response selection (emotion categorization). The integration of prosody with semantics, on the other hand, appears to rely chiefly on the left inferior frontal gyrus, a region that has been associated with semantic selection, retrieval and integration processes (Schirmer & Kotz, 2006). Activations in the left inferior frontal gyrus are stronger when emotional prosody and semantics are incongruent (e.g., a negative word spoken with a happy prosody) than congruent, indexing effortful semantic and integrative processing (Schirmer, Zysset, Kotz, & Cramon, 2004). According to ERP studies, higher order and executive prosodic functions are reached 300 to 400 ms after stimulus onset, as indicated by the modulation of the later components N300 and N400, which have been taken to reflect emotional meaning and integration operations (Bostanov & Kotchoubey, 2004; Paulmann & Pell, 2010a; Schirmer, Kotz, & Friederici, 2002).

These different processing stages of emotional prosody form a reciprocally connected network whose dynamics is context-sensitive. That is, the specificities of the situation and the significance that prosody has in a given context for a given person can influence one or more of the sub-processes. This contextual influence can occur through top-down mechanisms: attention may be voluntarily allocated to prosody and modulate its processing (Schirmer & Kotz, 2006). For instance, Leitman and colleagues (2010) observed that increasing the ambiguity of the stimuli (by reducing the salience of the acoustic cues) is associated with increased functional connectivity between the inferior frontal gyrus and superior temporal gyrus. This indicates that, in face of an ambiguous stimulus, increased top-down resources might be recruited, such that later evaluative operations augment information processing in earlier stages. Even early perceptual processes in the auditory cortex can be modulated by attention (e.g., Petkov et al., 2004). Additionally, bottom-up mechanisms may promote later processing efforts. This can occur, for example, when out attention is automatically captured by a certain emotional expression that is not expected (e.g., abrupt change in the speaker's prosody) or is highly relevant (e.g., signal of threat; Schirmer & Kotz, 2006).

Besides the neural structures mentioned so far, some neuropsychological and neuroimaging studies have indicated that the basal ganglia are also part of the neurocognitive network that supports emotional prosody. Activations in the putamen and caudate were found in fMRI studies (e.g., Bach et al., 2008; Frühholz et al., in press; Kotz et al., 2003), and impairments in emotion recognition in prosody were observed in patients with focal lesions in the basal ganglia (Paulmann et al., 2008a), as well as in patients with PD and Huntington's disease, two neurodegenerative conditions that compromise basal ganglia structures (e.g., Dara, Monetta, & Pell, 2008; Gray & Tickle-Degnen, 2010; Pell & Leonard, 2003; Speedie, Brake, Folstein, Bowers, & Heilman, 1990). The functional role of these subcortical structures is still not clear, though. Some authors suggest that the striatum acts, at an intermediate stage of processing, as a binding mechanism for the emotionally significant acoustic cues that unfold in the speech stream (e.g., Paulmann & Pell, 2010b). Alternatively, the basal ganglia might be involved in the later higher order stages of prosodic processing, through their connections with prefrontal brain regions (e.g., Mitchell & Bouças, 2009). This issue is analysed in detail in the empirical part of this thesis.

A question that is still poorly explored is whether different vocal emotion categories (e.g., sadness; anger) have distinct neurocognitive substrates. For instance, using fMRI and multivariate pattern analysis, Ethofer, Ville, Scherer, and Vuilleumier (2009) were able to uncover distinct spatial signatures in the auditory cortex for five prosodic emotion categories (anger, joy, relief, sadness and neutrality). Wittfoth and colleagues (2010) observed that the mechanisms that support conflict resolution between prosodic and semantic information might be modulated by valence: for happily intoned sentences expressing negative semantic content, they found activations in the left inferior frontal gyrus and superior temporal gyrus, whereas for angrily intoned sentences expressing positive semantic content they found activations in the thalamus and left caudate. Using a neuropsychological approach, Rymarczyk and Grabowska (2007) found that patients with damage in the prefrontal cortex are mostly impaired for happy prosody, those with temporo-parietal damage for sad prosody, and those with subcortical damage for angry prosody. Neuropsychological studies on patients with basal ganglia damage reported particularly acute impairments for negative prosodic emotions, as compared to positive ones (Dara et al., 2008; Gray & Tickle-Degnen, 2010).

2.5. INDIVIDUAL DIFFERENCES

Research on emotional prosody has been highlighting commonalities across individuals, such that general inferences can be made. However, it is important to keep in mind that in the domain of emotion processing individual variability is the rule rather than the exception (e.g., Eugène et al., 2003; Hamann & Canli, 2004). Determining the sources of variability in emotional prosody is crucial for a comprehensive understanding of the neurocognitive underpinnings of this function. Gender, age and personality traits may play significant roles. In an ERP study, Schirmer, Striano, and Friederici (2005) found that women are more sensitive than men to emotional prosodic cues at preattentive stages of processing. Women also integrate prosodic with semantic information faster and more readily than men (Schirmer et al., 2002; Schirmer et al., 2004), even though this difference might be dependent on task instructions: it disappears when men are explicitly instructed to focus their attention on prosody (Schirmer, Kotz, & Friederici, 2005). Several studies have also revealed age-related differences in how emotions are recognized in prosody, with older adults being less accurate than younger ones at recognizing some emotion categories (e.g., Laukka & Juslin, 2007; Mill, Allik, Realo, & Valk, 2009; Mitchell, Kingston, & Bouças, 2011; Ross & Monnot, 2011; Ruffman, Henry, Livingstone, & Phillips, 2008; Ryan, Murray, & Ruffman, 2010). These differences are detected early on in middle-age, around the forties (Paulmann et al., 2008b), and cannot be explained by sensory losses or by domain-general cognitive decline (Mitchell, 2007; Orbelo, Grim, Talbott, & Ross, 2005). Ageing effects in emotion recognition have been observed for other modalities as well, such as facial expressions and body postures (Calder et al., 2003; Ruffman, Halberstadt, & Murray, 2009; Ruffman et al., 2008). They may reflect neuropsychological decline in neural structures involved in emotion processing (Ruffman et al., 2008) and motivational changes associated with adult development (e.g., Samanez-Larkin & Carstensen, 2011). Personality characteristics might modulate brain responses to prosody too. Individuals higher in social orientation, a trait reflecting interest in social exchange and care about others, have enhanced activity in the amygdala and in the orbitofrontal cortex in response to emotional vocal stimuli (Schirmer et al., 2008). Higher neuroticism is associated with enhanced activity in the right amygdala, left postcentral gyrus and medial frontal structures (Brück, Kreifelts, Kaza, Lotze, & Wildgruber, 2011). Trimmer and Cuddy (2008) uncovered that individuals with higher emotional intelligence are more sensitive to emotions in speech prosody. Training in music might be another factor explaining individual differences in prosody, but available evidence is inconclusive (Thompson, Schellenberg, & Husain, 2004; Trimmer & Cuddy, 2008). This hypothesis is empirically examined in Chapter VIII.

2.6. WHEN EMOTIONAL PROSODY FAILS

Understanding the neurocognitive foundations of emotional prosody is important for the general advance of knowledge, and for applied clinical reasons as well. Difficulties in recognizing emotions in prosody are part of the phenotype of several developmental, neuropsychological and neuropsychiatric conditions. This is the case of autism spectrum disorders (Korpilahti et al., 2007; Lindner & Rosén, 2006), Williams syndrome (Pinheiro et al., 2011), paediatric and adult traumatic brain injury (Milders, Fuchs, & Crawford, 2003; Schmidt, Hanten, Li, Orsten, & Levin, 2010), acquired focal brain damage (Adolphs et al., 2002; Ross & Monnot, 2008), PD (Gray & Tickle-Degnen, 2010), Huntington's disease (Speedie et al., 1990), Alzheimer's disease (Taler, Baum, Chertkow, & Saumier, 2008), primary progressive aphasias (Rohrer, Sauter, Scott, Rossor, & Warren, in press), schizophrenia (Bach, Buxtorf, Grandjean, & Strik, 2009; Leitman et al., 2010), depression (Uekermann, Abdel-Hamid, Lehmkämper, Vollmoeller, & Daum, 2008), alcoholism (Monnot, Nixon, Lovallo, & Ross, 2001), posttraumatic stress disorder (Freeman, Hart, Kimbrell, & Ross, 2009) and psychopathy

(Blair et al., 2002). Given that emotional prosody plays an important role for communication and social interactions, difficulties in this ability should be considered as a target in clinical settings, as they may compromise the patients' psychosocial functioning (e.g., Trauner, Ballantyne, Friedland, & Chase, 1996). Higher accuracy in emotion recognition in prosody is associated with more positive interpersonal relationships and with less depression symptoms (Carton, Kessler, & Pape, 1999). It is also a predictor of social functioning, specifically of occupational performance (Hooker & Park, 2002). Even though our focus here is primarily on perception, it is important to note that deficits in the production of emotional prosody are also a feature of different clinical conditions, such as PD (e.g., Pell, Cheang, & Leonard, 2006) and autism spectrum disorders (Peppé, Cleland, Gibbon, O'Hare, & Castilla, 2011), with negative consequences for communication and social competence (e.g., Jaywant & Pell, 2010).

After discussing how emotions are communicated in the music of speech, in the next chapter we turn to emotions in music itself.

CHAPTER III

Music and Emotion

Like language, music is a uniquely human trait. Making and appreciating music are ubiquitous features of all human societies since ancient times. It is not possible to determine when exactly musical behaviours emerged during evolution – there are no fossil records of singing –, but archaeological findings date the first unequivocal evidence of musical instruments as back as circa 36,000 years, during the Upper Palaeolithic period (d'Errico et al., 2003). Musical behaviours and abilities also emerge very early in ontogenetic development, suggesting that they are guided by innate predispositions (Zentner & Eerola, 2010; Zentner & Kagan, 1996). As recently stated by Koelsch (2011), available evidence support the view that "musicality is a natural ability of the human brain" (p. 16). Indeed, there can be few individuals for whom music does not have a profound appeal, as indicated by the massive amounts of time that we spend on musical activities. A recent report found that music listening is placed among the most important and prevalent of leisure activities, above watching TV or reading books, for instance (Rentfrow & Gosling, 2003).

Music is generally approached from a socio-cultural perspective, that is, as a cultural artefact and an exquisite form of art. The consideration of music as a biological faculty whose neurocognitive foundations can be examined is relatively recent. In the last years, though, we have witnessed a rapid expansion in biological, cognitive and neuroscientific research on music (e.g., Koelsch & Siebel, 2005; Patel, 2008a, 2010; Peretz, 2006). Remarkable advances have been made in many domains, such as in the understanding of the genetic basis of music cognition (e.g., Peretz, 2008), the development of musical abilities (e.g., Hannon & Trainor, 2007; Trehub & Hannon, 2006), the neurocognitive basis of music perception (e.g., Koelsch, 2011; Koelsch &

Siebel, 2005; Peretz & Coltheart, 2003), the disorders of music cognition (e.g., Liu, Patel, Fourcin, & Stewart, 2010; McDonald & Stewart, 2008; Peretz, Champod, & Hyde, 2003), the relations between music and language (e.g., Juslin & Laukka, 2003; Patel, 2008a; Schön & Besson, 2001; Schön et al., 2010), the effects of musical training on neurocognitive plasticity (e.g., Besson, Chobert, & Marie, 2011; Habib & Besson, 2009; Moreno et al., 2009; Pantev & Herholz, in press; Patel, 2011) and the processing of emotions in music (e.g., Juslin & Västfjäll, 2008; Koelsch, 2010).

Emotions are perhaps the most important component of music. Ever since ancient Greece, there is an awareness of the ability of music to communicate emotions that we can perceive, enjoy and be moved by (e.g., Thompson, 2009). We use music to influence moods, to release emotions, to comfort and to enjoy ourselves, to match our emotional states and to alleviate stress (e.g., Juslin & Sloboda, 2010). The expressive power of music is also presumed in a number of applications in society, for example when we use it as a source of emotion in films (Cohen, 2010), marketing (North & Hargreaves, 2010) and therapy (Thaut, 2005; Thaut & Wheeler, 2010). How do the sounds of music express emotions? How does our neurocognitive system deal with emotions in music? In the following pages we review studies on emotion communication in music and on the neurocognitive foundations of musical emotions.

3.1. COMMUNICATION OF EMOTIONS IN MUSIC

Theoretical treatises about musical emotions have a long history (e.g., Budd, 1985; Kivy, 2002; Meyer, 1956), but only recently have they become sidelined by empirical research in a systematic fashion. Musical emotions were considered too mysterious, elusive and personal to be examined rigorously in psychology and neuroscience. In the last two decades, however, this has changed and the topic of emotion processing in music is now expanding at a fast pace, paralleling the growing interest dedicated to emotion research in general (Sloboda & Juslin, 2010). A large number of studies focus on how listeners perceive emotions communicated by music. In sharp contrast with the notion that this is a highly diffuse and variable experience, it has been repeatedly demonstrated that emotion recognition in music is remarkably consistent, accurate, immediate, precocious in ontogenetic development, and governed by universal principles to a significant extent.

Bigand, Vieillard, Madurell, Marozeau, and Dacquet (2005) showed that emotional responses to music are highly consistent both within and between listeners. They presented 27 instrumental music excerpts from the Western classic repertoire to listeners with and without musical training, and asked them to group together the excerpts that conveyed similar emotions. The duration of the excerpts was manipulated, 30 s or 1 s, and listeners performed the task twice, with at least one week of interval (test-retest). Groupings were very stable across trained and untrained participants (between listeners), from the first to the second testing session (within listeners), and for both longer and shorter excerpts. There is also remarkable invariance across listeners in how they judge the emotion category expressed by a certain music excerpt. Agreement rates for emotion recognition in music tend to be high (e.g., Laukka & Juslin, 2007; Mohn, Argstatter, & Wilker, in press; Resnicow, Salovey, & Repp, 2004; Vieillard et al., 2008). For instance, Vieillard and colleagues (2008) obtained high accuracy rates for the recognition of four emotions expressed by short instrumental music excerpts: 99% for happiness, 84% for sadness, 82% for fear and 67% for peacefulness. Accordingly, in a meta-analytic review of 12 studies, Juslin and Laukka (2003) concluded that the recognition accuracy of five emotion categories in music performance (anger, fear, happiness, sadness and tenderness) is higher than what would be expected by chance alone. They found a global accuracy of 70%, which was as high as the one they found for vocal emotions. With respect to the temporal dynamics of emotion recognition, it has been shown that reliable judgments can be obtained after quite short segments of musical information, a fact that indicates that the underlying mechanisms are fast-acting and immediate. Peretz, Gagnon, and Bouchard (1998) observed that listeners need only 250 ms of music to distinguish happy from sad excerpts. In a related vein, Filipic, Tillmann, and Bigand (2010) demonstrated, using a gating task, that the distinction between moving and neutral music can be made on the basis of 250 ms excerpts. Also using a gating task, Vieillard and colleagues (2008) determined the shortest duration that can trigger accurate emotion recognition in music: 483 ms for happiness, 1,446 ms for sadness, 1,737 ms for fear and 1,261 ms for peacefulness. From a developmental standpoint, available evidence indicates that the ability to perceive emotions in music emerges during early childhood. Cunningham and Sterling (1988) found that, by the age of four, children can already recognize happiness and fear in classical orchestral compositions with accuracy rates above chance, and by the age of six they are able to also identify reliably sadness and anger (see also Hunter, Schellenberg, & Stalinski, 2011). These early abilities are linked with sensitivity to specific musical features. Dalla Bella, Peretz, Rousseau, and Gosselin (2001) observed that six to eight-year-old children rely on both tempo (fast vs. low) and mode (major vs. minor) differences to distinguish happy from sad music, as adults do. Another attribute of emotion communication in music is that it is partly governed by universal principles. Some studies have shown that listeners are able to recognize emotions even in a non-familiar music style. Balkwill and Thompson (1999) found that Western listeners are sensitive to joy, sadness and anger in Hindustani raga excerpts. More recently, Fritz and colleagues (2009) compared how Western and Mafa listeners recognize happiness, sadness and fear in instrumental music excerpts composed according to the Western tonal tradition. The Mafa people are a native African ethnic group who had never been exposed to Western music. Western listeners had higher recognition accuracy that the Mafa, denoting a same-culture advantage, but the Mafa were able to recognize the three emotions with accuracy rates above chance level. Additionally, as Western listeners, they perceived consonant music as more pleasant than dissonant music.

The fact that the recognition of musical emotions is accurate, fast, precocious and universal is evidence that the underlying mechanisms show some of the same properties of emotion stimuli important for social functioning, such as facial expressions and speech prosody (e.g., Doherty, Fitzsimons, Asenbauer, & Staunton, 1999; Elfenbein & Ambady, 2002; Grandjean et al., 2005; Pell, 2005; Pons, Harris, & Rosnay, 2004; Tracy & Robins, 2008).

3.2. MUSICAL STRUCTURE AND EXPRESSIVE CUES IN PERFORMANCE

Which musical features contribute to the communication of a certain emotion? An important distinction is between features related to the music compositional structure, and those related to expressiveness in performance. Both are determinants of the perceived musical emotions (e.g., Thompson, 2009). *Musical structure* features correspond to those that are represented in conventional musical notation, such as tempo and dynamic markings, pitch, intervals, mode, melody, rhythm, harmony and formal properties (e.g., repetition; variation). These are manipulated by the composers to convey the intended emotional expressions (Gabrielsson & Lindström, 2010). Empirical research has shown that listeners are adept at perceiving emotions communicated by structural cues (Curtis & Bharucha, 2010; Gagnon & Peretz, 2003; Vieillard et al., 2008) since early childhood (Dalla Bella et al., 2001). Balkwill and Thompson (1999) observed that the structural properties tempo, rhythmic complexity, melodic complexity and pitch range are significant predictors of the listeners' subjective emotion ratings. Several associations between specific structural cues and perceived emotion are well established. This is the case for tempo, which is one of the most important emotional cues. Fast tempo might lead to the perception of activity/excitement, happiness/joy/pleasantness, potency, surprise, anger, uneasiness and fear. Slow tempo, on the other hand, might lead to the perception of peacefulness, calmness/serenity, sadness, dignity/solemnity, tenderness, boredom and longing. Mode is also an important determinant of emotions, with major mode being usually associated with happiness/joy, serenity, gracefulness and solemnity, and minor mode with sadness, anger, tension, dignity and disgust. High pitch might be associated with expressions of happiness, serenity, excitement, potency, anger, fear and activity, and low pitch might be associated with sadness, dignity/solemnity and vigour. As for melody, a wide melodic range may suggest joy and fear, and a narrow melodic range may suggest sadness, dignity, delicacy, serenity and triumph. A simple and consonant harmony might convey happiness, dignity and serenity, whereas a complex and dissonant harmony might convey tension, anger, excitement, sadness and unpleasantness. A regular/smooth rhythm may express happiness, dignity and peacefulness, and an irregular/rough rhythm may express anger, uneasiness and amusement (for comprehensive reviews, Gabrielsson, 2009; Gabrielsson & Lindström, 2010). In typical situations, emotional expression is determined by a combination of many cues, not a single one, and they may work in additive or interactive ways (e.g., Ilie & Thompson, 2006).

In addition to the structural properties determined by the composer, the way a performer plays a particular piece is crucial in shaping its emotional meaning. *Expressive cues in performance* refer to those features controlled and manipulated by the performer for aesthetic and communicative purposes, such as timing, dynamics, vibrato, articulation, timbre and phrasing (e.g., Bhatara, Tirovolas, Duan, Levy, & Levitin, 2011; Juslin, 2000; Juslin & Timmers, 2010). Variations in these features may render different emotional connotations to the same piece of music. Juslin (2000) asked three guitar players to perform the same three short melodies so as to communicate happiness, sadness, anger and fear through performance cues alone. These recordings were then measured for tempo, sound level, frequency spectrum, articulation and articulation variability. A sample of listeners rated how much they expressed the four emotions. Both performers and listeners associated consistently different emotion categories with different patterns of expressive cues, as indicated by correlation and multiple regression analyses between the performers' intentions, acoustic cues, and listeners' subjective judgments. Happiness was associated with fast tempo, high sound level, staccato articulation, high articulation variability and intermediate amount of high-frequency energy in the spectrum; sadness with slow tempo, low sound level, legato articulation, low articulation variability and little high-frequency energy in the spectrum; anger with fast tempo, very high sound level, legato articulation, low articulation variability and a lot of high-frequency energy in the spectrum; and fear with slow tempo, very low sound level, staccato articulation, high articulation variability and little high-frequency energy in the spectrum. Importantly, expressive cues in performance were very effective in communicating specific emotions: the performers' expressive intention predicted 70% of the variance in listeners' subjective judgments (see also Juslin & Laukka, 2003; Juslin & Timmers, 2010). Other studies indicate that discrete emotion categories can be accurately perceived only on the basis of expressive cues (Juslin & Laukka, 2003;

Laukka & Juslin, 2007). As mentioned for structural cues, we extract emotional information from many expressive/performance cues when listening to music, even though tempo, sound level and timbre appear to be especially important in determining the perceived emotion (Juslin & Timmers, 2010).

Note that the distinction between compositional and performance-related emotion cues is useful for theoretical and empirical reasons, but the two types of cues are intermixed in everyday musical contexts. As pointed out by Resnicow and colleagues (2004), the structure of a musical piece constrains the aesthetically permissible range of variability in performance cues such as tempo and loudness. That is, typically, the communicative intentions of the performer have a meaningful relationship with the inherent characteristics of the musical piece. Otherwise, the interpretation would sound incoherent and unnatural. Furthermore, in some cases, compositional and performance-related cues might not be easily discernible. For instance, while harmony or mode are clearly structural cues, tempo and dynamics might both be set by the composer in the musical notation, and/or modulated by the performer for communicative purposes.

3.3. EMOTION PERCEPTION AND EXPERIENCE

So far, we have focused on the expression and recognition of emotions. According to some authors, though, perceiving emotions should be distinguished from actual emotional experiences and reactions to music (e.g., Evans & Schubert, 2008; Gabrielsson, 2002; Kivy, 2002; Konecni, 2008; Sloboda & Juslin, 2010). The reason for the distinction is that, arguably, we can recognize a certain expression in music (e.g., a happy expression) on the basis of sensory and cognitive operations, without necessarily feeling the corresponding emotion, that is, without being "moved" or emotionally engaged. This relates to the dichotomy between a *cognitivist* position, holding that music simply expresses emotions, and an *emotivist* position, holding that music also induces emotions in the listener (Kivy, 1991; Peretz, 2010). It is currently well established that music can express as well as induce emotions (e.g., Koelsch, 2010; Sloboda & Juslin, 2010), but the relation between perceptions and feelings is complex and remains to be fully understood. On one hand, the link might indeed not be obligatory: different processes may be involved in perceptions and feelings. On the other hand, they are often aligned, the difference being only a matter of degree. For instance, Salimpoor, Benovoy, Longo, Cooperstock, and Zatorre (2010) asked listeners to rate self-selected and experimenter-selected music excerpts concerning the degree of valence and arousal they (1) felt in response to the excerpts, and (2) believed the excerpts were intended to express. For self-selected music, there were no differences between felt and perceived valence and arousal. For experimenter-selected music, listeners felt more neutral and less aroused than what they thought the excerpts were intended to express. In other words, for unfamiliar music the feelings of valence and arousal were weaker than the corresponding perceptions. The authors also found that perceived valence was not related to subjective pleasure ratings, suggesting that listeners could feel as much pleasure in response to music they believe was intended to convey sadness as in response to music intended to convey happiness. Zentner, Grandjean, and Scherer (2008) asked listeners to rate how often they felt and perceived a preselected list of emotions in response to their preferred music. They found that emotions tend to be more often perceived than felt, particularly negative emotions, namely sadness and dysphoria. However, for other emotions, such as tender longing and amazement, the frequency of perceptions and corresponding feelings was similar. Hunter, Schellenberg, and Schimmarck (2010) found strong correlations (r > .8) between perceptions and feelings of happiness and sadness in response to 30 s pieces of music composed by Johann Sebastian Bach, even though perception responses tended to higher than feeling responses (see also Evans & Schubert, 2008). Neuropsychological studies indicate that emotion recognition in music is not just a "cold" cognitive operation, as it depends critically on core emotion systems, namely the amygdala (Gosselin, Peretz, Hasboun, Baulac, & Samson, 2011; Gosselin, Peretz, Johnsen, & Adolphs, 2007; Gosselin et al., 2005) and striatum (Omar et al., 2011). From a theoretical standpoint, simulation (e.g., Adolphs, 2006, 2010) and embodied emotion models (Barrett, 2006b, 2009) postulate that emotion perception and experience are inextricably linked. This is also suggested by a recent proposal on the mechanisms underlying musical emotions, which we examine below (Juslin et al., 2010; Juslin & Västfjäll, 2008). In this thesis we focus on the overt recognition of musical emotions. We have no pretence of proving or disproving whether this captures (at least partly) the same mechanisms as actual emotional experience, even though this possibility is discussed.

3.4. WHICH EMOTIONS ARE EVOKED BY MUSIC?

An important question for theories of musical emotions concerns which emotive states are most frequently evoked by music. Approaches to musical emotions have relied primarily on models and measures from non-musical domains (Zentner et al., 2008). Usually, participants are presented with music excerpts and are asked to rate them for preselected emotion terms that almost invariably reflect those of basic emotions theory, such as happiness, sadness or anger. Notwithstanding, it is intuitive that music may evoke a more nuanced and diverse range of states, and it is possible that these are distinct from the ones evoked by other domains. Some studies have attempted to identify the emotions induced by music and their prevalence. Juslin and colleagues (2010) asked 706 Swedish participants to use their own words, instead of a preselected list, to describe the feelings they had recently experienced in response to music. Positive emotional states were far more referred to (84%) than negative ones (16%). Specific emotions were much more prevalent (92%) than general unspecific affect (8%). Part of the specific emotions (11%) corresponded to mixed emotions (e.g., both anger and sadness), but the remaining 89% featured single ones. The most frequently reported emotion was happy-elated, accounting for over 50% of all episodes, followed by sad-melancholic, calm-content and nostalgic-longing. Much less frequent emotions were aroused-alert, angry-irritated, loving-tender, moved-touched, interested-expectant and proud-confident, respectively. In another study, Juslin, Liljeström, Västfjäll, Barradas, and Silva (2008) examined the prevalence of musical and non-musical emotions as they occur naturally in everyday life by using the Experience Sampling Method (participants carried a palmtop that emitted a sound signal seven times a day for two weeks and, when signalled, they completed a questionnaire). Music occurred in 37% of the sampled episodes, and in 64% of the music episodes participants indicated that music influenced the way they felt. The most frequent emotions were happiness-elation and calm-contentment, while the least frequent ones were shame-guilt and disgust-contempt. In general, musical and non-musical episodes elicited similar emotional responses, but there were differences as well: happiness-elation and nostalgia-longing were more frequent during musical than during non-musical episodes, whereas anger-irritation, boredom-indifference and anxiety-fear were more common during non-musical than during musical episodes. Additionally, positive emotions were more frequent during musical than during non-musical emotion episodes. Zentner, Grandjean, and Scherer (2008) conducted four interrelated studies to characterize, measure and classify musical emotions. They came up with a list of emotion terms that best describe music-evoked emotions, from which a model with nine factors was derived: wonder, transcendence, tenderness, nostalgia, peacefulness, power, joyful activation, tension and sadness. This model was shown to account for musical emotions better than basic emotions and dimensional emotion theories do. Differing from basic emotions theories, most categories in this model are positive. Moreover, it includes categories such as wonder, nostalgia, and transcendence, which are not generally considered by any current model of emotion. Overall, these studies highlight that listening to music evokes frequently specific emotions, which are mostly positively valenced, and may differ from the ones evoked by non-musical stimuli.

3.5. MECHANISMS UNDERLYING EMOTIONAL RESPONSES TO MUSIC

According to Juslin and colleagues (Juslin et al., 2010; Juslin & Västfjäll, 2008), not enough attention has been devoted to the mechanisms by which music evokes emotions. These authors use the term *psychological mechanisms* to refer to “any kind of information processing that may lead to the induction of emotions through listening to music" (Juslin et al., 2010, p. 619). They put forward a theoretical framework featuring seven distinct mechanisms that would underlie musical emotions: brain stem reflexes, rhythmic entrainment, evaluative conditioning, contagion, visual imagery, episodic memory and musical expectancy. *Brainstem reflexes* correspond to a mechanism whereby we react emotionally to one or more fundamental acoustical attributes of music because they are processed as signalling a potentially important and urgent event. For instance, all other things being equal, we are aroused by sounds that are sudden, loud, dissonant or have fast temporal patterns. Such responses in the brainstem would reflect how certain sound qualities are captured rapidly and automatically by our perceptual system. *Rhythmic entrainment* corresponds to a mechanism by which an emotion is evoked because the rhythm of the music interacts with internal body rhythms of the listener, such as heart rate or breathing rhythms. These internal rhythms would synchronize with musical ones, and this adjustment would spread to other components of emotion (e.g., feeling) through proprioceptive feedback, thus increasing listeners' arousal. *Evaluative conditioning* corresponds to a mechanism whereby we respond emotionally to a piece of music because it was repeatedly paired with other positive or negative stimuli. For instance, a certain music can makes us feel happy because we listened to it together in time with a specific situation that always makes us feel happy, such as meeting one of our best friends. Emotional *contagion* corresponds to a mechanism whereby an emotion is elicited because we perceive a certain emotional expression in music (e.g., happiness; sadness) and then "mimic" that expression internally. Through the possible mediation of the so-called mirror neurons, perceiving an emotion would be accompanied by the induction of the corresponding emotional experience. In other words, emotion perception would trigger emotional experience. Why would we perceive emotional expressions in music? The authors suggest that it is because music features acoustic patterns similar to those of vocal emotions; the neural mechanisms that respond to vocal emotions would thus react rapidly and automatically to voice-like aspects of music. This mechanism is of special relevance in the context of the empirical work of this thesis. The close relation between emotions in music and in speech prosody is discussed in detail in the next chapter. *Visual imagery* corresponds to a mechanism by which an emotion is induced because we conjure up visual images while listening to a piece of music. For instance, the image of a beautiful nature scene might come to our mind while we listen to a certain melodic movement. The emotion experienced would result from the interaction between the image and the music. *Episodic memory* corresponds to a mechanism whereby we react emotionally to a piece of music because it evokes a memory of a specific event in our life. These emotions may be intense, possibly because the physiological responses and the experiential contents of the original events are stored along with each other in memory. Indeed, we often use music to remind ourselves of past valued events, a fact that indicates that music may play an important nostalgic function in everyday life. Finally, *musical expectancy* corresponds to a mechanism by which an emotion is induced because a feature of the music structure delays, violates or confirms our expectations concerning the continuation of the music. The expectations are built upon the previous experience with the music style. For instance, in the Western tonal system, the progression of E-F# prompts the expectation that the music will continue with G#, such that when this does not happen the listener may become surprised.

These mechanisms would operate singularly or in a combined fashion while we listen to music. According to the authors (Juslin et al., 2010; Juslin & Västfjäll, 2008), each of them would correspond to distinct brain functions. This theoretical framework is preliminary, and so the precise characteristics of the mechanisms are to be examined. However, it is useful to derive specific hypotheses and to guide empirical research on musical emotions.

3.6. NEUROCOGNITIVE BASIS OF MUSICAL EMOTIONS

In the last years, an increasing number of studies have examined the neural foundations of musical emotions (for reviews, Koelsch, 2010; Peretz, 2010). There is evidence that processing emotions in music involves subcortical brain systems, namely limbic structures and the striatum, as well as cortical systems, namely the temporal lobe, orbitofrontal cortex, ventromedial prefrontal cortex, anterior cingulate cortex and the insula.

Seminal evidence for the involvement of subcortical systems comes from a study by Blood and Zatorre (2001). They used positron emission tomography (PET) to measure regional cerebral blood flow changes while participants listened to 90 s of their own favourite music, to which they usually had highly pleasurable experiences of “shivers-down-the-spine” or “chills”. More intense chills correlated with cerebral blood flow decrease in the amygdala and left hippocampus, and with increase in the ventral striatum, dorsomedial midbrain and thalamus (chills also modulated cortical activity, with cerebral blood flow decreasing in the ventromedial prefrontal cortex, and increasing in the insula, orbitofrontal cortex and cingulate cortex; the functional role of these cortical regions is discussed below). The subcortical structures recruited by music are also recruited for reward responses to biologically relevant stimuli, which are important for survival, such as food and sex. Thus, as highlighted by the authors (ibd.), the results of this study indicate that the highly abstract sounds of music can engage ancient emotion and reward systems in the brain. Recently, Salimpoor, Benovoy, Larcher, Dagher, and Zatorre (2011) showed that the recruitment of the striatum by music-evoked chills reflects the modulation of dopaminergic activity. They found endogenous dopamine release in the striatum during music-induced peak emotional arousal, by using the neurochemical specificity of [^{11}C]raclopride PET scanning, combined with psychophysiological measures of autonomic nervous system activity. The authors further observed, in a fMRI study with the same stimuli and listeners, a functional dissociation within the striatal system: the caudate was chiefly involved in the anticipation of reward, and the nucleus accumbens in the actual experience of intense emotional responses to music. Musical emotions recruit subcortical brain systems even when participants do not experience intense chills. For instance, Koelsch, Fritz, Cramon, Müller, and Friederici (2006) compared neural responses to consonant joyful instrumental dance-tunes (pleasant music) and to electronically manipulated dissonant music (unpleasant music) in an fMRI study. They detected subcortical activations for both types of music: pleasant consonant music activated the ventral striatum, and unpleasant dissonant music activated the amygdala, hippocampus and parahippocampal gyrus. Similarly, Mitterschiffthaler, Fu, Dalton, Andrew, and Williams (2007) found that happy music activates the ventral and dorsal striatum, whereas sad music activates the hippocampus and amygdala. Thus, positively valenced responses to music appear to engage the striatum, through the modulation of dopaminergic activity, and relatively more negative emotional responses appear to engage medial temporal structures.

Neuropsychological studies indicate that core subcortical neuroaffective systems are also critical for emotion recognition in music, not only for emotional experience. In three studies, Gosselin and colleagues (2005, 2007, 2011) found that patients with damage in the amygdala had impaired recognition of fear and sadness in instrumental music excerpts, while the recognition of positive emotions, happiness and peacefulness, was relatively preserved. This pattern could not be explained by defects in low-level music perception abilities. In another study, it was found that damage in the parahippocampal cortex correlated with abnormal perception of dissonance in music, such that patients rated dissonant music as slightly pleasant, whereas healthy controls rated it as unpleasant (Gosselin et al., 2006). The more extensive the patients' damage was, the more indifferent they were to dissonance (see also Khalfa et al., 2008). Recently, Omar and colleagues (2011) examined emotion recognition in music (happiness, sadness, anger and fear) in patients with frontotemporal lobar degeneration, and explored associations between behavioural performance and grey matter losses. Patients' performance was impaired and correlated with grey matter volumes in an extensive bilateral cerebral network that included subcortical structures, more precisely the parahippocampal gyrus, hippocampus, amygdala, nucleus accumbens and ventral tegmentum (the network included cortical structures as well, notably the insula, anterior cingulate, orbitofrontal cortex, medial prefrontal cortex, dorsal prefrontal, inferior frontal, anterior and superior temporal cortices, fusiform gyrus and posterior parietal cortices).

Cortical systems comprise neural structures that are evolutionarily more recent. It has been shown that they play different functional roles in musical emotions. Auditory regions in the superior temporal cortex, particularly in the right hemisphere, support the processing of low-level perceptual aspects of music, thus setting the stage for emotional responses and interpretative processes (Mitterschiffthaler et al., 2007; Peretz, Blood, Penhune, & Zatorre, 2001; Zatorre, Belin, & Penhune, 2002). There is solid evidence that the orbitofrontal cortex is a key region for the processing of emotions in music. It is recruited when participants experience intense chills (Blood & Zatorre, 2001), when they listen to music without feeling intense responses (Blood, Zatorre, Bermudez, & Evans, 1999; Menon & Levitin, 2005), and when they perform overt emotion recognition tasks (Omar et al., 2011). This region, which is connected with the amygdala and the striatum, appears to play a role in the analysis and integration of the reward value of emotional stimuli across sensory modalities, thereby guiding behavioural judgments (Blood & Zatorre, 2001; Menon & Levitin, 2005; Peretz, 2010). The ventromedial prefrontal cortex is also an important structure for musical emotions. It is active during chill experiences (Blood & Zatorre, 2001), during emotion recognition (Omar et al., 2011), and might be involved in triggering autonomic emotional responses to music. Johnsen, Tranel, Lutgendorf, and Adolphs (2009) found that patients with damage to the ventromedial prefrontal cortex are impaired in the generation of skin-conductance responses to music. Damasio's somatic marker hypothesis focuses on this region (ibd.). Finally, the anterior cingulate and the insular cortex were also shown to be involved in music-evoked chills (Blood & Zatorre, 2001), music listening (Koelsch et al., 2006; Mitterschiffthaler et al., 2007) and emotion recognition (Omar et al., 2011). These structures seem to be implicated in coupling autonomic responses and subjective feeling states, as well as in motor-related functions and in the synchronization of biological subsystems during emotion episodes (such as physiological arousal, motor expression, motivation, monitoring processes and cognitive appraisal; Koelsch, 2010).

3.7. INDIVIDUAL DIFFERENCES

Which factors explain differences across individuals in emotional responses to music? Personality traits might be important. For instance, Juslin and colleagues (2008) observed that the prevalence of music-evoked emotions correlates with personality traits. Specifically, they found that higher levels of neuroticism are associated with more frequent occurrence of pleasure-enjoyment in response to music. They also observed that higher levels of extraversion are associated with higher overall prevalence of musical emotions, suggesting the extraverts and introverts use music differently. Barrett and colleagues (2010) found that individuals' proneness to nostalgia predicts the intensity of music-evoked nostalgia, such that individuals with higher nostalgic tendencies are more likely to experience higher levels of nostalgia in response to music. Nostalgia proneness, in turn, correlated with neuroticism. Recently, Montag, Reuter, and Axmacher (in press) showed that persons who consider themselves as easily getting absorbed by arts and music (subscale "self-forgetfulness" of the character dimension "self-transcendence") have decreased activity in the striatum while listening to their favourite songs.

Individual factors might also modulate the ability to recognize emotional expressions in music. Mohn, Argstatter, and Wilker (in press) failed to detect significant associations with personality traits, but Vuoskoski and Eerola (2011) observed that the perception of sadness in film music excerpts correlates with higher neuroticism and with lower extraversion. Resnicow and colleagues (2004) found that individuals higher in emotional intelligence, as assessed by the Mayer-Salovey-Caruso Emotional Intelligence Test, are better in the recognition of emotions in music performance. We pointed out that ageing modulates emotion recognition in emotional prosody and in other emotional stimuli (e.g., facial expressions; body postures). Whether this is also the case for musical emotions remains poorly studied. In the only study that approached this question so far, Laukka and Juslin (2007) found that older adults are less accurate than younger ones in the recognition of negative emotions in music performance. Musical training might be another source of individual variability, but available evidence is inconclusive: while some studies reported suggestive evidence that musical training is associated with enhanced sensitivity to musical emotions (Bhatara et al., 2011; Livingstone, Muhlberger, Brown, & Thompson, 2010), other studies found no effects (Bigand et al., 2005; Resnicow et al., 2004). The impact of ageing and musical training in emotion recognition in music is examined in the empirical part of this thesis (Chapter VI).

3.8. CLINICAL APPLICATIONS

The therapeutic value of music in clinical settings has long been recognized. The scientific framework and rationales for music interventions have been altered on the basis of recent developments in the cognitive neuroscience of music. According to the model developed by the Center for Music Therapy Research in Heidelberg (Viktor Dulger Institute), there are five factors or ingredients by which music can improve the psychological and physiological health of children and adults: attention modulation, emotion modulation, cognition modulation, behaviour modulation and communication modulation (Hillecke, Nickel, & Bolay, 2005). These ingredients reflect the idea that music listening and music making engage a multitude of behaviours and brain structures related to cognitive, sensorimotor and emotional processing (e.g., Koelsch & Siebel, 2005). Thus, music could be used therapeutically to modulate activity at different levels and in different domains. For instance, it was shown that listening to relaxing music, as compared to silence, facilitates recovery from a psychologically stressful task; it leads to the reduction of salivary cortisol levels (Khalfa, Dalla Bella, Roy, Peretz, & Lupien, 2003). Recently, Särkämö and colleagues (2008) conducted a single-blind, randomized and controlled trial to determine the effects of everyday music listening in the recovery of cognitive and emotional functions after middle cerebral artery stroke. It was observed that two months of daily music listening (music was self-selected), as compared with listening to audio books or no listening material, led to significant improvements in verbal memory, focused attention, depression and confused mood. These effects are at least partly mediated by the emotional component of music.

As noted by Koelsch (2010), because music modulates the activity of core emotion systems, such as the amygdala, hippocampus and striatum, it can in principle be used to stimulate and regulate these systems when they are impaired (e.g., depression; PD; posttraumatic stress disorder). In an exploratory study, Koelsch, Offermanns, and Franzke (2010) observed that a music-therapeutic method involving music making in groups can produce positive changes in mood, as indexed by decreased depression/anxiety, decreased fatigue and increased vigour. These changes were accompanied by an increase in the experience of positive emotions (higher ratings for happiness) and by a decrease in negative emotions (lower ratings for anger, sadness and anxiety). Other studies suggest that musical-rhythmic stimuli can be used to regulate motor functions, namely gait and arm control, in patients with stroke and PD. The mechanisms behind these effects appear to be arousal and priming of the motor system, as well as timing and entrainment of the motor system (Thaut et al., 2009). Music therapeutic exercises can also be used effectively to improve the "feeling" experience of emotion, the identification of emotion, the expression of emotion, the understanding of emotional communication of others, and the synthesis, control and modulation of emotional behaviours (Thaut & Wheeler, 2010).

Despite recent advances, the knowledge on the specific mechanisms that mediate the therapeutic effects of music is still limited. Research on the neurocognitive basis of music and on how music modulates brain plasticity will be essential to inform concepts and practices in music therapy. A more complete understanding of how music relates to other cognitive functions, namely speech and language, will also be illuminating for potential therapeutic applications. In the next chapter we discuss the relations between emotions in speech prosody and in music.

CHAPTER IV

Perceiving Emotions in Speech Prosody and in Music: How Close?

The reason why music conveys specific emotions, induces mood changes and evokes emotional reactions remains puzzling. On one hand, it does not serve any obvious survival or reproduction functions. Possible adaptive roles of musical behaviours have been suggested, notably sexual selection, social cohesion and socio-cognitive development, but these are controversial (for reviews, Koelsch, 2011; Patel, 2008a). On the other hand, in everyday life, emotions are typically a response to events that are appraised as having the potential to interfere with our goal-directed actions, and music rarely either serves or blocks goals (Konecni, 2008; Scherer, 2004). In other words, while in most cases emotions are *utilitarian* (they have functions in the adaptation and adjustment of individuals), in music emotions are *aesthetic* (they are not triggered by appraisals concerning goal revelance, instead they derive from the intrinsic qualities of music; Scherer, 2004).

One hypothesis is that musical emotions are an accidental by-product of capacities originally dedicated to other purposes. Music could have co-opted the neurocognitive machinery shaped by evolution for sounds of greater biological relevance, namely vocal emotions – speech prosody and nonverbal vocal expressions (Hauser & McDermott, 2003; Juslin & Laukka, 2003; Juslin et al., 2010; Nieminen et al., 2011; Patel, 2008b; Peretz, 2010; Pinker, 1997). Other theories suggest that music and language evolved as two different specializations of a common predecessor, which consisted of a vocalization system that subserved rudimentary referential-emotive communicative functions. This primitive system would have been highly adaptive for our ancestors and some of its features would still be present in the current forms of both music and language, thus explaining the intimate connections between the two domains and the emotional power of music (e.g., Brown, 2000; Mithen, 2007).

4.1. SIMILAR CODES

The notion that the emotional dimension of music might be linked with the emotional dimension of voice has long been speculated by philosophers, theorists and natural scientists (e.g., Juslin & Laukka, 2003; Scherer, 1995). For instance, as far back as ancient Greece, Plato claimed that music is able to arouse emotions in the listeners because it imitates the way people express themselves in speech and exclamations. In late Renaissance, this hypothesis was espoused again in Florence by the *Camerata,* a group of the nobility who invented what is currently known as opera (Kivy, 2002). In the 19th century, the scholars Herbert Spencer (1857) and Hermann von Helmholtz (1863/1954) also suggested that part of musical expressiveness lies in the usage of intensified and richer versions of human voice modulations. These intuitions have been submitted to empirical inquiry in modern science. For instance, a study by Snowdon and Teie (2010) found that cotton-up tamarins respond affectively to music only when it incorporates acoustical characteristics of emotional vocalizations of their species. Tamarins were generally indifferent to human music, but they showed increased arousal for musical stimuli that included features of their own threat vocalizations, and increased calm behaviour for musical stimuli based on their affiliative vocalizations. Thus, by examining a non-human species, this study demonstrated that the emotional value of music might indeed result from the inclusion of species-specific principles of vocalizations.

In the last years, systematic empirical research provided evidence for similarities between the human voice and music with respect to the acoustic encoding of specific emotions. In a landmark meta-analysis, Juslin and Laukka (2003) reviewed 104 studies on vocal expression and 41 studies on music performance. The studies on vocal expression focused on emotions communicated through suprasegmental cues of speech samples, i.e., speech prosody. The studies on music focused on expressive features of performance, that is, variations implemented by musicians when interpreting a given musical piece. About half of the studies on music performance included singing voice, and the other half included instrumental music. The authors examined five emotion categories: anger, fear, happiness, sadness and love-tenderness. The review unveiled that the recognition accuracy of these emotions is similar in speech and music (corresponding to 70% correct in a forced-choice task with five response alternatives). Importantly, the authors analysed and compared the acoustic profiles associated with the five emotions in vocal expressions and music performance. They inspected basic dimensions of pitch, intensity, duration and voice quality/timbre. Emotion-specific profiles were found to overlap significantly across domains in three main features: speech rate/musical tempo, voice intensity/sound level and high-frequency energy. For example, anger and happiness were typically associated, in prosody and music, with increased speech rate/tempo, intensity/sound level, and high-frequency energy, whereas sadness and tenderness were typically associated with decrements in these features. There were less data available for other acoustic features, but some cross-domain similarities were also found in the variability of voice intensity/sound level, in microstructural regularity versus irregularity (concerning frequency, intensity and duration), voice onsets/tone attack, mean pitch, pitch variability and pitch contours. For instance, sadness was associated with low pitch, low pitch variability and falling contours in both prosody and musical performance, whereas anger and happiness were associated with high pitch, large pitch variability and upward contours. In Table 2, we reproduce the complete cross-modal acoustic profiles hypothesized by the authors based on the reviewed data for the five emotions. This extensive meta-analysis provided strong support for the relation between the expressive features of speech prosody and music performance in terms of decoding accuracy and code usage for discrete emotions.

Table 2. Patterns of acoustic cues for discrete emotions in vocal expression and in music performance. Reprinted from Juslin and Laukka (2003, p. 802).

Emotion	Acoustic cues (vocal expression/music performance)
Anger	Fast speech rate/tempo, high voice intensity/sound level, much voice intensity/sound level variability, much high-frequency energy, high F_0/pitch level, much F_0/pitch variability, rising F_0/pitch contour, fast voice onsets/tone attacks, and microstructural irregularity
Fear	Fast speech rate/tempo, low voice intensity/sound level (except in panic fear), much voice intensity/sound level variability, little high-frequency energy, high F_0/pitch level, little F_0/pitch variability, rising F_0/pitch contour, and a lot of microstructural irregularity
Happiness	Fast speech rate/tempo, medium–high voice intensity/sound level, medium high-frequency energy, high F_0/pitch level, much F_0/pitch variability, rising F_0/pitch contour, fast voice onsets/tone attacks, and very little microstructural regularity
Sadness	Slow speech rate/tempo, low voice intensity/sound level, little voice intensity/sound level variability, little high-frequency energy, low F_0/pitch level, little F_0/pitch variability, falling F_0/pitch contour, slow voice onsets/tone attacks, and microstructural irregularity
Tenderness	Slow speech rate/tempo, low voice intensity/sound level, little voice intensity/sound level variability, little high-frequency energy, low F_0/pitch level, little F_0/pitch variability, falling F_0/pitch contours, slow voice onsets/tone attacks, and microstructural regularity

Note: F_0 = Fundamental frequency

Recent studies show that emotion cues from musical structure (e.g., tonality; harmony) can also be intimately related to vocal expressions. This is remarkable, since musical structure is highly dependent on the composer and on the musical tradition (e.g., Western tonal music), and so in principle it would be arbitrary and variable to a significant extent. Parallels with speech prosody were found for tonality. Most of the Western music is composed in the major or in the minor mode, and these modes are usually associated with positive and negative emotional qualities, respectively (e.g., Gabrielsson & Lindström, 2010). A main distinction between modes is that in the major mode there is a major third interval between the third and first scale degrees (frequency ratio of 5:4), and in the minor mode these scale degrees form a minor third interval (frequency ratio of 6:5). Curtis and Bharucha (2010) examined whether the musical interval of a minor third is somehow similar to the modulation of pitch in emotional speech. They asked actresses to produce bi-syllabic utterances (e.g., *Let's go*) varying prosody to express four emotions: anger, happiness, pleasantness and sadness. Acoustic analyses revealed that, in sad speech, the relationship between the two salient pitches, from the first to the second syllable, approximates a descending minor third in terms of pitch distance (-300 cents, which correspond to three musical semitones). In a perceptual experiment, the authors found that this prosodic interval was the most important acoustic predictor of subjective ratings for sadness. They also confirmed the well-established association between the musical minor third interval and the perception of sadness. These results show that a musical interval that differentiates major from minor music also occurs in the pitch contour of speech, and communicates sadness similarly across domains. In a related vein, Bowling, Gill, Choi, Prinz, and Purves (2010) compared the spectral qualities of excited and subdued speech with the spectral qualities of the intervals that distinguish major and minor music. Musical intervals samples were extracted from classical and folk traditional melodies composed in major and minor keys, and speech samples consisted of single words and monologues spoken in an excited or subdued manner. To compare the spectra across domains, the authors computed measures of fundamental frequency and frequency ratios. They found that the differences between excited and subdued speech on these measures parallel the differences between thirds and sixths in major and minor music. For instance, fundamental frequency was higher in excited vs. subdued speech, as well as in thirds and sixths in major vs. minor melodies.

Thus, the codes used by composers and the characteristics of musical traditions might not be as arbitrary and dependent on learned/cultural associations as previously thought. Converging evidence comes from the cross-cultural study by Fritz and colleagues (2009), who observed that Mafa individuals (native African ethnic group), despite having never been exposed to Western music, rely on mode differences to distinguish the emotional connotations of Western music excerpts, just as Western listeners do.

4.2. SHARED NEUROCOGNITIVE MECHANISMS?

From the standpoint of cognitive neuroscience, an interesting question is whether perceiving emotions in voice and in music engages common neurocognitive mechanisms. On the basis of the acoustic similarities found between domains, Juslin and colleagues hypothesized that music engages the brain modules that also respond to vocal expressions (Juslin & Laukka, 2003; Juslin et al., 2010; Juslin & Västfjäll, 2008). According to the authors, music would be perceived by these modules as a *superexpressive voice* because it sounds like the human voice but can go beyond what voice does (e.g., in terms of speed, intensity, timbre and pitch range). For instance, while the human voice expresses anger when it has fast rate, high intensity and harsh timbre, a musical instrument, say a violin, might express much more anger because it can be faster, louder and harsher than speech. This super-stimulus would be particularly suited to trigger emotional contagion: the intensified voice-like cues of music would engage the neural mechanisms involved in vocal emotions and, by contagion, the expressed emotion would trigger the corresponding emotional experience. Available empirical evidence for the putative overlapping of neurocognitive mechanisms across domains is scarce (for a review, Peretz, 2010), and thus this hypothesis remains speculative. However, recent studies have provided some hints.

The processing of pitch, a low-level feature crucial for emotion differentiation in music and speech, appears to depend on shared mechanisms (Liu et al., 2010; Marques, Moreno, Castro, & Besson, 2007; Nan, Sun, & Peretz, 2010; Schön, Magne, & Besson, 2004). Nevertheless, emotion recognition implicates the analysis and integration of a multitude of acoustic features beyond pitch, as well as higher order interpretative processes, and so it requires direct examination. Ilie and Thompson (2006) compared how identical acoustic manipulations of intensity, rate and pitch height in music and speech influence subjective ratings on valence (pleasant-unpleasant), energy arousal (awake-tired) and tension arousal (tense-relaxed). They found similar effects across domains, but they also found domain-specific effects. For instance, loud speech and loud music were similarly perceived as more unpleasant, more energetic and more tense than their soft counterparts, suggesting that intensity manipulations produce domain-general effects. However, pitch height manipulations produced opposite effects across domains: high-pitched speech was perceived as more positive than low-pitched speech, while high-pitched music was perceived as more negative and low-pitched one. The authors further observed that, across conditions, music received higher ratings on valence and energetic arousal than speech did, a finding that indicates that the two types of stimuli may differ in the extent to which they engage emotional systems. Thus, the possibility that emotion perception in speech and music does not fully overlap cannot be discarded. A combination of domain-general and domain-specific mechanisms might be engaged. This is a plausible hypothesis, since emotional reactions to music involve mechanisms beyond the features related to speech, such as rhythmic entrainment and musical expectancies (Juslin et al., 2010; Juslin & Västfjäll, 2008). As such, musical emotions may recruit neural substrates over and above those involved in vocal emotional processing. Furthermore, emotional prosody is not an unitary entity. It involves multiple cognitive operations implemented by a distributed network of cortical and subcortical systems (Leitman, Wolf, et al., 2010; Schirmer & Kotz, 2006). We can speculate that these systems are not all similarly engaged by speech and music. For example, Omar and colleagues (2011) found that, in frontotemporal degeneration, both emotion recognition in music and in nonverbal vocal expressions (e.g., laughs; sobs) were impaired, but the anatomical pattern of brain damage associated with performance in both domains was not fully overlapping: the volume of grey matter loss in some brain structures, namely in the orbitofrontal cortex, medial prefrontal cortex and insula, showed stronger associations with behavioural performance in emotion recognition in music than in voice. Notwithstanding, this study examined nonverbal expressions, not speech prosody, and so more research is needed to determine whether the recognition of emotions in speech and music recruit common mechanisms, and to specify the degree of the overlap.

Different kinds of studies can be useful to shed light on this topic. In this thesis we focus on two possibilities. One consists of comparing musicians and untrained listeners. If mechanisms are (partly) shared, musical training can be expected to produce cross-domain benefits for emotion recognition in speech prosody. Putative shared neurocognitive mechanisms for emotion perception in both domains would become further specialized due to adaptive plastic changes induced by training in music. To our knowledge, only two studies examined this hypothesis, and they obtained different results: one found suggestive evidence of a positive effect of training (Thompson et al., 2004), but the other did not find any effect (Trimmer & Cuddy, 2008). Thus, this hypothesis awaits clarification. Another kind of study consists of focusing on individuals with a neuropathological condition that impairs the recognition of emotions in speech prosody, such as receptive aprosodia caused by stroke (e.g., Ross & Monnot, 2008) or PD (e.g., Pell & Leonard, 2003). If these individuals have a similar profile of impairments for emotion recognition in prosody and music, this would suggest that the damaged neural structures are shared by both domains. A differential profile of impairments, on the other hand, would be evidence that both domains are partly independent. Up to now, such a comparative design has never been implemented. According to Patel (2008b), this approach would be a powerful test to the resource-sharing hypothesis, provided that some methodological aspects are met: patients should be tested in other modalities, like facial expressions, to guarantee that the impairment is specific to the auditory domain; possible low-level defects in auditory processing should be excluded; and cognitive abilities (e.g., memory) should be controlled for.

These studies will contribute to a better understanding of the neurocognitive relationship between emotions in speech and music. They can also be informative for applied psychological science. If both domains set into action common mechanisms, music therapy might be a useful device in language rehabilitation programs.

CHAPTER V

Goals of this Thesis

In the preceding chapters, we have reviewed research focusing on emotion theories, emotional speech prosody, musical emotions, and on the parallels between emotion processing in speech prosody and in music. Several issues that we have referred to remain poorly understood, and they lie at the basis of the empirical part of this thesis. We highlighted that musical emotions show some of the same properties of emotion modalities of unequivocal biological and social importance, namely facial expressions and speech prosody. For instance, as for these modalities, the recognition of discrete emotions in music is accurate, quick, consistent, precocious in ontogenetic development and governed by universal principles to a significant extent. It is well established that ageing operates changes in emotion processing across different modalities (e.g., Ruffman et al., 2008), but so far it is unclear whether and how ageing affects also musical emotions. In other words, we do not know whether emotion processing in music undergoes developmental changes during the adult life, as emotion processing in other domains. It is also unclear whether training in music impacts on the processing of musical emotions, as it appears to impact on several aspects of brain and cognitive functioning (e.g., Besson et al., 2011; Habib & Besson, 2009). Concerning emotional speech prosody, we described studies showing that perceiving emotions in this modality is determined by a combination of universal and language-specific processes. Thus, it is of paramount importance to provide prosodic materials and perceptual studies in different languages. To the best of our knowledge, this has not been done for European Portuguese. Finally, we reviewed comparative research on emotional prosody and musical emotions, and concluded that there are striking cross-domain commonalities regarding the acoustic encoding of emotional qualities. We also pointed out that a currently debated hypothesis in cognitive neuroscience is that emotion processing in speech prosody and in music recruits shared neurocognitive mechanisms. However, empirical research on this hypothesis is scarce. We set out to investigate these open questions in a series of empirical studies using behavioural and neuropsychological approaches.

As we underlined in the review of the literature, emotion processing involves multiple components. It is therefore important to define at the outset what exactly we will assess. In all studies we focus in the overt recognition of discrete emotions, that is, in how individuals are able to identify and evaluate explicitly emotional qualities expressed by music excerpts and by speech samples. We resort to behavioural tasks that are widely used in emotion recognition literature, namely forced-choice and rating tasks. In conventional forced-choice tasks, participants are given a predefined list of emotions and select the one that best matches the emotion expressed by the stimulus. This task has been strongly defended by some authors (Ekman, 1994), though it may present some difficulties. When the number of options is small, results may reflect discrimination between alternatives rather than true emotion recognition (Banse & Scherer, 1996; Scherer, 2003; Scherer et al., 2003). Another criticism is that this task might inflate agreement rates because the response options are given beforehand to the participant (Sauter, 2006). This potential problem is somewhat mitigated in the studies presented here because we are chiefly interested in differences between groups, not so much in agreement rates per se. In rating tasks, participants use quantitative intensity scales to evaluate the degree to which different emotions are expressed by the stimulus. In a typical situation, participants are asked to rate how much each stimulus expresses every emotion included in the study (e.g., Adolphs & Tranel, 1999; Adolphs, Tranel, Damasio, & Damasio, 1995; Gosselin et al., 2005; Pell & Leonard, 2003; Trimmer & Cuddy, 2008). In the following paragraphs we briefly outline the empirical studies that were conducted.

Do ageing and musical expertise modulate emotion recognition in music?

Little is known about individual differences in how we respond to music. One of the goals of this thesis is to examine the roles of ageing and musical training in emotion recognition in music. For modalities like facial expressions, voice and body postures, it has been demonstrated that advancing age is associated with decreased accuracy in the recognition of some emotions, particularly negative ones (e.g., Mill et al., 2009; Ruffman et al., 2008). Studies on facial expressions have also indicated that training may improve the ability to recognize emotions (Elfenbein, 2006; Matsumoto & Hwang, 2011; Silver, Goodman, Knoll, & Isakov, 2004). What about musical emotions? In two studies, presented in Chapter VI, we determined how emotion recognition in instrumental music excerpts changes along the adult life span, and whether individuals with musical training respond differently than those without training. We used an emotion rating task in both studies, and employed stimuli from the database developed and validated by Vieillard and colleagues (2008) for research on musical emotions.

How are emotions recognized in emotional prosody in Portuguese?

In the study presented in Chapter VII, we developed and validated a database of prosodic stimuli in Portuguese. The availability of this kind of resources is valuable for research with Portuguese-speaking participants and for cross-language studies. Analogous databases have been devised for different languages (e.g., Pell, 2002; Staroniewicz & Majewski, 2009; Wu et al., 2006). The verbal materials consisted of short sentences with emotionally neutral semantic content and of pseudo-sentences (utterances that include pseudo-words). Two speakers recorded these materials varying prosody in order to express neutrality and six emotion categories: anger, disgust, fear, happiness, sadness and surprise. In order to validate the stimuli, we performed acoustic analyses and conducted a perceptual experiment using a forced-choice emotion recognition task. Reaction times and intensity judgments were also collected and analysed.

Does emotion recognition in speech prosody and in music recruit shared neurocognitive mechanisms?

As we emphasized in the review of the literature, we can use different research strategies to examine the extent to which emotion processing in speech and music engages common mechanisms. One consists of determining cross-domain transfer effects from musical training to emotional prosody. Available evidence on this issue is inconclusive (Thompson et al., 2004; Trimmer & Cuddy, 2008). In the study presented in Chapter VIII, we hypothesized that if neurocognitive mechanisms are shared across domains, then musicians should show enhanced emotion recognition in speech prosody. To investigate this hypothesis, we compared highly trained musicians with untrained participants in a forced-choice emotion recognition task using the prosodic stimuli previously validated.

Another research strategy consists of determining whether a neuropathological condition that affects speech prosody also affects musical emotions in a similar manner. In the last empirical study, Chapter IX, we adopted this approach in a neuropsychological study on patients with PD. The recognition of emotions in speech prosody may be impaired in these patients, due to their dysfunction in the basal ganglia (Gray & Tickle-Degnen, 2010; Pell & Leonard, 2003). Whether the recognition of emotions in music also depends critically on the basal ganglia is still poorly understood (Omar et al., 2011), as is whether PD patients have difficulties in dealing with musical emotions (Tricht, Smeding, Speelman, & Schmand, 2010). We tested patients and matched healthy controls in an emotion rating task for music and speech prosody, and compared the effect of PD across domains. Participants were also assessed for emotion recognition in facial expressions, taken as a control measure to exclude general deficits in emotion recognition. Furthermore, patients and controls underwent an extensive neuropsychological evaluation to ascertain that general cognitive or low-level perceptual problems would not account for putative impairments in emotion recognition.

PART 2

EMPIRICAL STUDIES

CHAPTER VI

Ageing and Musical Expertise Modulate Emotion Recognition in Music

6.1. INTRODUCTION

Emotion recognition in music relies on robust and fast-acting mechanisms, which are operational since early childhood. A poorly understood question concerns individual differences in these mechanisms. Why do we differ in emotional responses to music? Ageing and musical expertise may play significant roles. Advancing age produces many changes in cognitive and socio-emotional functioning (e.g., Hedden & Gabrieli, 2004; Salthouse, 2009; Samanez-Larkin & Carstensen, 2011), and one of these changes is in the recognition of emotional stimuli (e.g., Ruffman et al., 2008). Learning music can also induce long-lasting effects in brain structure and function (e.g., Habib & Besson, 2009; Hyde et al., 2009; Marques et al., 2007; Pantev & Herholz, in press). Whether these two experiential factors, ageing and musical expertise, modulate emotion recognition in music is only minimally investigated. We present two empirical studies that examine this question. In the first one, we determined how the recognition of positive and negative musical emotions changes across the full range of the adult life span[1]. In the second one, we examined whether these changes are independent of general age-related cognitive decline, i.e., whether they are primary in origin. We also compared expert musicians with musically untrained listeners[2].

AGEING AND EMOTION RECOGNITION | Research on age differences in emotion recognition focuses mostly on facial expressions (e.g., Calder et al., 2003; Isaacowitz et al., 2007; Orgeta, 2010; Williams et al., 2006; Williams et al., 2009). It has been shown that ageing is associated with decreased accuracy in the recognition of some emotions,

---

[1] Published in Lima, C. F., & Castro, S. L. (2011). Emotion recognition in music changes across the adult life span. *Cognition and Emotion, 25,* 585-598. doi:10.1080/02699931.2010.502449

[2] Submitted for publication in Castro, S. L., & Lima, C. F. (submitted). Emotion recognition in music is robust, and yet variable: The impact of age and musical expertise.

while the recognition of others remains stable. Which emotions undergo changes and which ones do not is still debated, but negative emotions seem to be particularly affected (Mill et al., 2009). Ruffman, Henry, Livingstone, and Philips (2008) conducted a meta-analytic review of 17 studies and concluded that older adults (around 71 years of age) are less accurate than younger ones (around 24 years of age) in recognizing negative facial expressions of fear, anger and sadness. They observed that changes for happiness and surprise are less consistent and smaller, and for disgust recognition accuracy tends to improve with age. Isaacowitz and colleagues (2007) also summarized results of previous studies and came to similar conclusions. Research on modalities other than faces is scarce, but available evidence indicates that age-related changes also occur for emotion recognition in speech prosody (e.g., Mitchell, 2007; Mitchell et al., 2011), nonverbal vocal expressions (Ryan et al., 2010), body postures (e.g., Ruffman et al., 2009), lexical stimuli (Isaacowitz et al., 2007) and pictures (St. Jacques, Dolcos, & Cabeza, 2008). This suggests that ageing may impact supramodal emotion processes to a certain extent. Notwithstanding, modality-specific effects might exist. For instance, Ruffman and colleagues (2008) found that faces and voices differ in the pattern of changes: decline for happiness is less consistent in faces than in voices, and decline for fear occurs in faces but not in voices. Research on ageing typically lies on comparisons between younger and older adults, on their sixties or seventies – "extreme age group" design (Isaacowitz & Stanley, 2011) –, and so the developmental trajectories during the middle years are much less well known. Studies that used finer gradations in age, though, indicate that age-related effects can be detected already in middle-age, around the forties (e.g., Calder et al., 2003; Isaacowitz et al., 2007; Paulmann et al., 2008b) or even earlier (Mill et al., 2009).

An important question concerns the origin of age differences in emotion recognition. Are they a by-product of domain-general cognitive decline, a consequence of motivational and experiential changes related to adult development, or the result of degradation in neural structures responsible for emotion processing? Ageing is associated with decline in general cognitive functions like attention, memory, processing speed and reasoning (e.g., Hedden & Gabrieli, 2004; Salthouse, 2009). This decline could account for the older adults' reduced accuracy in emotion recognition:

they might be less effective in dealing with the cognitive demands posed by emotion recognition tasks. According to this hypothesis, the decline should be greatest for emotions that are inherently more difficult to recognize – they would pose increased cognitive demands –, and emotion recognition should be associated with general cognitive abilities. For facial expressions and speech prosody, however, the emotions that undergo age-related changes do not appear to match those that are more difficult to recognize, and the changes persist after controlling for cognitive abilities (Mitchell, 2007; Orbelo et al., 2005; Ruffman et al., 2008; Ryan et al., 2010; Sullivan & Ruffman, 2004). Thus, in these modalities ageing seems to affect emotional functioning in a more direct manner. Other explanatory mechanisms have received more attention, namely changes in motivation and emotion regulation associated with adult development. This perspective postulates that older adults differences in emotion perception are primarily driven by changes in top-down regulatory mechanisms. According to the *socioemotional selectivity theory*, as we grow older, our ability to regulate emotions improves and emotional well-being becomes a priority over exploration and novelty, because time horizons start to be perceived as limited. These changes would lead to heightened ability and motivation to allocate cognitive resources towards positive information, and away from negative one. The term *positivity effect* is often used to refer to this developmental shift from a preference for negative input in youth to a preference for positive input at older ages (Carstensen & Mikels, 2005; Charles & Carstensen, 2010; Mather & Carstensen, 2005; Samanez-Larkin & Carstensen, 2011). This view is consistent with studies showing that ageing impacts relatively less the recognition of positive emotions than the recognition of negative ones (Laukka & Juslin, 2007; Mill et al., 2009; Mitchell et al., 2011). Evidence from neuroscience is convergent with the notion that changes in emotion regulation and motivation may influence emotion perception as we age. For instance, Williams and colleagues (2006) found that age-related behavioural changes in the recognition of facial expressions are accompanied by a shift in the medial prefrontal cortex activity (independent of grey matter loss). Specifically, the decline in the recognition of fearful faces was associated with increased activity in the medial prefrontal cortex, and the relative stability in the recognition of happy faces was associated with decreased activity in the same region. It was suggested that advancing age produces a shift towards greater prefrontal controlled inhibition over the processing of negative emotional input (due to regulatory strategies), and uninhibited processing of positive input. The greater recruitment of frontal cortex by older adults in response to emotional stimuli is a consistent finding in the literature (St. Jacques, Bessette-Symons, & Cabeza, 2009; St. Jacques et al., 2008). It has also been shown that amygdala activations decrease with age for negative but not for positive information (Gunning-Dixon et al., 2003; see also, Kisley, Wood, & Burrows, 2007). An alternative explanation for ageing effects on emotion recognition stresses the role of neuropsychological deterioration in brain structures implicated in emotion processing. The recognition of some emotions would undergo changes and others would remain stable because rates of deterioration differ across neural systems of emotion (e.g., Raz et al., 2005). For instance, decline in the cingulate cortex and in the amygdala might contribute to explain changes in the recognition of fear and sadness in faces, and the relative preservation of basal ganglia structures might explain the stability observed for disgust (Calder et al., 2003; Ruffman et al., 2008). According to Cacioppo, Berntson, Bechara, Tranel, and Hawkley (2011), brain decline could also account for the positivity effect: age-related deterioration of the amygdala, which is important to monitor negative information, would lead to the reduction of negative affect. Moreover, it is possible that the increased frontal activations in older adults reflect compensation for brain decline, instead of motivation-related regulatory strategies (St. Jacques et al., 2009). Overall, the relative contribution of general cognitive decline, motivational changes, and brain deterioration, for age-related differences in emotion recognition is still a contentious issue. Because most research was conducted with visual stimuli, studies on other modalities are crucial to examine the generality of the effects and to provide insights into their causes.

Does emotion recognition in music change with advancing age, undergoing similar developmental modifications as those of other emotional stimuli relevant for social functioning and communication? If so, are changes towards positivity, with responsiveness to negative emotions declining more than responsiveness to positive ones? Are the effects a consequence of general cognitive decline, or primary in origin? Here we are interested in these questions. It has been shown that musical emotions may exhibit some of the properties of other emotion domains (see Chapter III), but the modulations related to ageing remain largely unexplored. In a first systematic approach to the topic, Laukka and Juslin (2007) compared younger and older adults, over 65 years of age, in the recognition of emotions in music and in speech prosody. Musical emotions were communicated only through expressive cues in performance: the same excerpt (the theme from W. A. Mozart's Piano Sonata in A Major, K331) was played with differences in features like loudness, articulation, vibrato and phrasing, to express anger, fear, happiness, sadness and neutrality. Older adults had lower recognition rates than younger ones for the negative emotions of sadness and fear, both in music and in speech. No differences were observed for happiness, neutrality and anger. These results indicate that ageing may produce similar effects in emotion recognition in music and in speech prosody. The emotions which were hardest to recognize were not those that underwent age-related changes. Thus, general cognitive decline might not be the primary cause of the effects. However, several questions are to be examined. First, it is unknown how ageing impacts emotions expressed through music compositional structure, i.e., emotions encoded in the musical notation by the composer. Variations in structural features such as mode, tempo, pitch range and dissonance, determine the expressed emotion (e.g., Dalla Bella et al., 2001; Gabrielsson & Lindström, 2010; Hunter, Schellenberg, & Schimmack, 2008) and predict the listener's subjective judgments (e.g., Balkwill & Thompson, 1999). Second, it remains to be determined whether effects for music are observed already in middle-aged, as described for other emotion modalities (e.g., Calder et al., 2003; Isaacowitz et al., 2007; Paulmann et al., 2008b). Third, possible associations between emotion recognition in music and general cognitive abilities need to be inspected to determine directly whether age differences are mediated by general, emotion-unspecific, factors. Furthermore, the study by Laukka and Juslin (2007), as most studies, included an unbalanced number of positive and negative emotions (only one positive emotion, as compared with three negative ones). This undermines conclusions regarding effects of valence.

MUSICAL EXPERTISE AND EMOTION RECOGNITION IN MUSIC | Musical expertise might also be a source of individual differences in emotion recognition in music. Longitudinal and cross-sectional studies indicate that learning music changes brain morphology and cognitive processing (for a review, Habib & Besson, 2009). For instance, Hyde and colleagues (2009) showed that 15 months of lessons in early childhood (mean age at start of study = 6.32 years) leads to behavioural improvements in melodic and rhythmic discrimination tasks, and to structural changes in auditory and motor brain areas. Adult professional musicians, as compared to amateur musicians and non-musicians, have increased grey matter volume in the primary motor and somatosensory areas, premotor areas, anterior superior parietal areas, inferior temporal gyrus, cerebellum, Heschl's gyrus and inferior frontal gyrus (Gaser & Schlaug, 2003). Training impacts on non-musical cognitive abilities, such as discrimination of F_0 in speech (e.g., Moreno et al., 2009), nonverbal reasoning (Forgeard, Winner, Norton, & Schlaug, 2008) and executive control (Bialystok & DePape, 2009). It is plausible that it also fine-tunes mechanisms underlying the processing of musical emotions because musical practice focuses extensively on the ability to deal with fine-grained modulations of acoustic features for expressive purposes, as well as on the relations between musical structure and the instantiation of emotions. This hypothesis appears intuitive, but delineating which specific musical abilities are enhanced in musicians is not a trivial endeavor, particularly in light of findings showing that untrained listeners perform like experts in some music processing tasks (e.g., learning new musical idioms) – they acquire sophisticated musical knowledge just through exposure (for a review, see Bigand & Poulin-Charronnat, 2006). Indeed, the study by Bigand and colleagues (2005) suggests that emotional responses to music might be largely similar in musically trained and untrained listeners. They asked participants to group music excerpts according to the emotion they evoked, and the number of groups produced was independent of training. The only effect of musical expertise was in enhancing the consistency of responses across testing sessions in a test-retest situation. However, this study assessed musical emotions at an implicit level, and explicit task conditions might be more sensitive to expertise effects (Bigand & Poulin-Charronnat, 2006). However, results for explicit tasks are scarce. In an exploratory analysis (trained participants, n = 9), Livingstone, Muhlberger, Brown, and Thompson (2010) found a significant correlation between the number of years of training and recognition accuracy for music excerpts that conveyed emotions through both structural and expressive cues. These results suggest that training does play a role, though it is difficult to ascertain whether the effect is driven by a better ability of musicians to process the expressive and/or the structural emotion cues. Bhatara and colleagues (2011) found that musical training is associated with enhanced sensitivity to differences in expressive musical cues, namely timing and amplitude (n = 10 in one experiment + n = 10 in another experiment). Resnicow, Salovey, and Repp (2004), on the other hand, failed to find such an effect (number of trained participants unspecified). Hence, available results are mixed. To the best of our knowledge, no studies have addressed the impact of expertise for emotions expressed by structural features alone. Furthermore, prior research was not able to disentangle putative specific effects on musical emotions from domain-general cognitive benefits of training (e.g., Bialystok & DePape, 2009; Schellenberg, 2006)

THE CURRENT STUDIES | The goal of the two following studies is to determine how ageing and musical expertise modulate emotion recognition in instrumental music. We also aim to establish whether these effects are independent of general cognitive differences. Stimuli were music excerpts previously validated to express two positive emotions, happiness and peacefulness, and two negative ones, sadness and fear/threat (Vieillard et al., 2008). These excerpts were unfamiliar to the participants – they were composed for experimental purposes. Emotions were communicated through compositional structural cues only. Participants rated how much each excerpt expressed each of the four emotions on intensity scales. This task does not involve a selection of a single category, and so it is less prone to the response biases that can affect conventional forced-choice tasks (e.g., Isaacowitz et al., 2007). Additionally, it mitigates the high decisional demands posed by forced choices, which might augment the impact of older adults' general cognitive difficulties. Such a task was previously used in studies on emotion recognition in facial expressions (e.g., Adolphs, Schul, & Tranel, 1998), speech prosody (e.g., Péron et al., 2010), nonverbal vocal expressions (Sauter, Eisner, Calder, et al., 2010) and music (Gosselin et al., 2005, 2007). In Study 1, we focused primarily on age-related changes. We covered the full range of the adult life span, including younger, middle-aged and older participants. Based on the reviewed literature, we predicted that advancing age would produce decreased responsiveness to musical emotions, particularly to the negative ones (sadness and fear/threat). These changes would be significant already in middle-age, as previously observed for other modalities. Because a subset of participants had some degree of musical training, we also undertook a preliminary examination of whether that training is associated with enhanced sensitivity to musical emotions. In Study 2, musical expertise and age were manipulated orthogonally. We compared expert musicians and musically untrained listeners (hereafter, controls) from two age cohorts, young and middle adulthood. In addition to emotion recognition, they were assessed for domain-general cognitive abilities, and completed brief control measures of personality and socio-communicative traits, because these may influence emotion processes (e.g., Hamann & Canli, 2004; Mill et al., 2009). We expected an advantage of musicians over controls for emotion recognition. If ageing and expertise effects on musical emotions are primary in origin, they should be independent of general cognitive and personality measures. In both Study 1 and Study 2, multiple regression analyses were conducted to explore how structural cues of music excerpts predict participants' subjective emotion ratings, and whether age and expertise effects are linked with differences in weighting these cues.

6.2. STUDY 1

6.2.1. Method

PARTICIPANTS | A total of 114 healthy adults (67 female) volunteered to take part in this study. They were aged between 17 and 84 years, and were categorized into three groups with 38 participants each: younger, middle-aged and older adults. Table 3 presents their demographic and background characteristics. The younger adults were undergraduate and graduate students from University of Porto; middle-age and older ones came from several local communities, including senior universities. The three groups were matched for education level, as assessed by the number of years attending school ($p > .05$). All participants had normal or corrected-to-normal vision, and none reported head trauma, substance abuse, cognitive or hearing difficulties, nor major psychiatric or neurological illnesses. For older adults, possible cognitive decline was evaluated by using the cognitive screening test Mini-Mental State Examination, MMSE

(Folstein, Folstein, & McHugh, 1975; Portuguese version, Guerreiro, Silva, Botelho, Leitão, & Garcia, 1994); they performed near ceiling, 29.3 on average (maximum 30), and none scored below 27. Older participants also completed a pure-tone audiometric screening to guarantee that they had acceptable hearing thresholds (at least 30 dB HL for frequencies of 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz at left and right ears). Thirty-three participants reported having had some kind of formal musical training, including learning how to play an instrument; mean of 5.1 years of training (*SD* = 3.9; range = 1 – 14; age when training began, *M* = 11 years).

Table 3. Demographic and background characteristics of the participants in each age group (Study 1). *SD*s are presented in parentheses.

	Age group		
	Younger	Middle-aged	Older
Age (years)	21.8 (3.5)	44.5 (6.2)	67.2 (6.2)
Age range (years)	17 - 29	35 - 56	60 - 84
Gender	17F / 21M	21F / 17M	29F / 9M
Education (years)	15.5 (2.2)	17.0 (3.7)	15.6 (3.1)
Mini-Mental State Examination (/30)	-	-	29.3 (1.0)
Musical training (years)	5 (4.1)	4.8 (4.2)	5.8 (2.8)
Participants with musical training (*n*)	18	10	5

MUSIC STIMULI | The stimuli were 56 music excerpts validated for research on emotions by Vieillard and colleagues (2008). These excerpts were composed to express four emotions: happiness, peacefulness, sadness and fear/threat, 14 stimuli per category. They consist of a melody with accompaniment, follow the rules of the Western tonal system, and were produced in piano timbre by a digital synthesizer. The emotional information is expressed exclusively by music compositional structure, through variations in features such as mode, dissonance, pitch range, tone density, rhythmic regularity and tempo. Excerpts do not vary in performance-related expressive features like dynamics, vibrato, phrasing or articulation. The mean duration of the excerpts is 12.4 s. They were downloaded at the Internet site from the Isabelle Peretz Research Laboratory (www.brams.umontreal.ca/plab/publications/article/96). Perceptual validation procedures including categorization, gating and dissimilarity judgments, confirmed that the intended emotions are perceived rapidly and with high agreement rates between listeners, high accuracy (Vieillard et al., 2008). Emotions can be recognized in these excerpts universally, that is, in the absence of prior exposition to the Western music (Fritz et al., 2009). The have also been used successfully in neuropsychological studies with patients with focal brain damage (Gosselin et al., 2005, 2007, 2011), neurodegenerative disorders (Drapeau, Gosselin, Gagnon, Peretz, & Lorrain, 2009), as well as with neurotypical children in developmental studies (Hunter et al., 2011).

The stimuli were pseudo-randomized and divided into two blocks of 28 trials each; the presentation order of the blocks was counterbalanced across participants. Two additional stimuli were used as practice trials: 15 s from Debussy's Golliwog's Cakewalk, for happiness, and 18 s from Rachmaninoff's 18th Variation from the Rhapsody on a Theme by Paganini, for peacefulness.

PROCEDURE | Participants were tested individually in a single experimental session lasting about 30 minutes; exceptionally, some undergraduate students were tested in groups of two or three. They were told that they would listen to short excerpts of music expressing different emotional states, namely happy (*alegre*), peaceful (*sereno*), sad (*triste*), and/or fearful/threatening (*assustador*). All participants were well familiarized with the four emotion labels. They rated how much each stimulus expressed each of the four emotions on 10-point scales, from 0 (absent, *ausente*) to 9 (present, *presente*). It was stressed that participants should always respond to the four emotion scales, as they could perceive different emotions simultaneously, with similar or different degrees of intensity. For instance, even if they perceived a stimulus as happy, they should rate it not only with respect to happiness, but also with respect to the possible expression of sadness, fear and peacefulness. Each stimulus was presented only once and no feedback was given. The task started with practice trials, followed by the two blocks of stimuli. Stimulus presentation was controlled by SuperLab version 4.0 software (Abboud, Schultz, & Zeitlin, 2006), running on an Apple MacBook Pro computer. Excerpts were presented via high-quality loudspeakers, adjusted to a comfortable volume level for each participant. Responses were collected via a pencil-and-paper questionnaire. Participants were encouraged to respond fast and spontaneously (they did not need to wait until the end of the excerpt to respond). The inter-stimulus interval was fixed for younger and middle-aged participants, 6 s, and contingent upon response for older ones, because some of them needed more than 6 s to provide the four ratings.

In the same experimental session, participants completed a brief questionnaire exploring their demographic, background, and musical profile. Older adults also performed the audiometric screening and the MMSE before the experimental task.

6.2.2. Results

To compare emotion recognition across age groups we derived a measure of accuracy from the raw ratings, based on the emotion that received the highest rating. For each subject and excerpt, when the highest of the four ratings matched the intended emotion, the response was coded as an *accurate* categorization. When the highest rating corresponded to a non-intended emotion, the response was considered *inaccurate*. When the highest rating corresponded to more than one emotion (e.g., giving 8 for sadness and also for peacefulness, and lower ratings for the other two categories), the response was considered *ambivalent*. Ambivalent responses indicate that, in a given excerpt, participants perceived more than one emotion with the same strength. This derived measure of accuracy was used previously in studies on emotion perception in music (Gosselin et al., 2005, 2007; Vieillard et al., 2008) and in prosody (e.g., Adolphs et al., 2002). It controls for the possibility that participants of different ages used the scales differently, and takes into account judgments both for the intended and non-intended emotions: a response is accurate only if the rating for the intended emotion is higher than the ratings for all the non-intended emotions.

Table 4 presents the percentage of accurate categorizations for each emotion, and the distribution of inaccurate and ambivalent responses, separately for younger, middle-aged and older participants. Accuracy rates were always higher for the intended than for the non-intended categories, indicating that the highest ratings were assigned consistently to the intended emotions. Agreement rates ranged between 33% (fear in older participants) and 97% (happiness in younger ones), as can be seen in the diagonal lines of the Table 4 (bold cells). This confirms that the stimuli effectively communicated the four emotions.

Table 4. Distribution of responses (%) for each intended emotion as a function of age (Study 1). Diagonal cells in bold indicate accurate categorizations, i.e., match between the highest rating and the intended emotion. Ambivalent responses represent situations in which the highest rating was assigned to more than one emotion. Standard errors are presented in parentheses.

Age group / excerpt type	Distribution of responses (%)				
	Happy	Peaceful	Sad	Scary	Ambivalent
Younger					
Happy	**97 (1.1)**	1	0	0	2
Peaceful	12	**44 (4.9)**	33	0	11
Sad	1	9	**79 (3.6)**	3	8
Scary	6	1	8	**76 (2.9)**	8
Middle-aged					
Happy	**94 (1.9)**	2	1	0	4
Peaceful	9	**65 (4.6)**	11	0	14
Sad	2	24	**58 (4.8)**	2	14
Scary	13	5	22	**45 (5.5)**	14
Older					
Happy	**88 (3.1)**	5	1	0	5
Peaceful	13	**57 (4)**	15	0	15
Sad	2	33	**45 (4.3)**	2	18
Scary	18	10	23	**33 (4.1)**	15

EFFECTS OF AGE ON ACCURACY | To examine how age modulates accuracy, correct categorizations per emotion were arcsine transformed and submitted to an ANCOVA, with emotion as repeated-measures factor (happy, peaceful, sad and scary), age as between subjects-factor (younger, middle-aged and older), and years of musical training as covariate. Musical training was partialled out here and in all the analyses on age effects because the number of trained participants was unbalanced across age groups (see Table 3). Main effects and interactions were followed-up with *post hoc* Tukey HSD tests. Effect sizes are expressed as partial eta squared (η_p^2). The four emotions elicited significantly different accuracy rates: happiness reached the highest accuracy (93%, *p*s < .01), followed by sadness (61%) and peacefulness (55%), with similar rates ($p > .1$); fear had lower accuracy than happiness and sadness (52%, *p*s < .01), but not peacefulness [$p > .5$; main effect of emotion, $F(3,330) = 90.16$, $p < .0001$, $\eta_p^2 = .45$]. As predicted, age groups differed in emotion recognition, but not uniformly across categories [main effect of age, $F(2,110) = 11.78$, $p < .0001$, $\eta_p^2 = .17$; interaction Age x Emotion, $F(6,330) = 12.71$, $p < .0001$, $\eta_p^2 = .19$]. For the negative emotions, sadness and fear, middle-aged and older participants were less accurate than younger ones (*p*s < .001); middle-aged and older participants did not differ significantly (*p*s > .3). For the positive emotions, happiness and peacefulness, accuracy rates remained invariant across age groups (*p*s > .9), with the exception that for peacefulness middle-aged participants achieved better accuracy than younger ones ($p < .01$). We repeated this analysis with participants' raw ratings on the intended emotion scales (raw ratings are displayed in Appendix 1). The pattern of age effects was globally replicated: older participants provided lower ratings than younger ones for sad and scary excerpts (*p*s < .01), and ratings for happy and peaceful ones did not change (*p*s > 1); for scary excerpts, differences were significant already in middle-age ($p < .0001$), but for sad ones they were not [$p > .1$; main effect of age, $F(2,110) = 10.91$, $p < .0001$, $\eta_p^2 = .17$; interaction Age x Emotion, $F(6,330) = 15.28$, $p < .0001$, $\eta_p^2 = .22$]. The age differences were also replicated with accuracy rates corrected for possible response bias (unbiased accuracy rates are displayed in Appendix 2)[3]: older participants had lower accuracy than younger

[3] In forced-choice tasks, raw proportions of correct responses are not always the best measure of accuracy because they are vulnerable to response bias. Participants might be disproportionately likely to respond using some response categories rather than others. As a consequence, differences between groups might reflect this bias instead of true differences in the intended ability (for a review, see Isaacowitz et al., 2007). Our accuracy data are less prone to bias because they are derived from ratings on intensity scales and participants did not perform any choice between categories. However, to confirm that the pattern of age-related changes is not a by-product of possible response trends, we reanalyzed accuracy using unbiased hit rates, *Hu* (Wagner, 1993). Hu is a measure of accuracy which is insensitive to bias in responses. It represents the joint probability that a stimulus category is correctly recognized when it has been presented, and that a response is correct given that it has been used. It was calculated for each emotion and participant using the formula $Hu = A^2 / (B \times C)$, where A corresponds to the number of stimuli correctly identified, B to the number of presented stimuli (14 per category), and C to the total number of responses provided for that category (i.e., including misclassifications). Values were arcsine transformed before being analyzed.

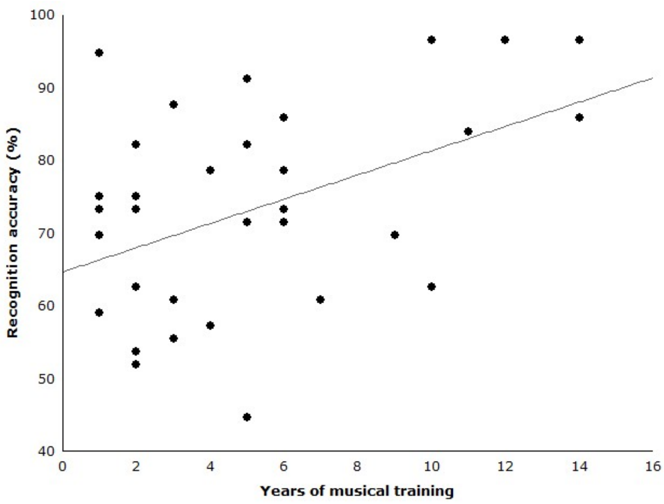
ones for sadness and fear (ps < .001), and were similar for happiness and peacefulness (ps > .05); the differences were significant already in middle-age for fear ($p < .001$), but not for sadness [$p > .3$; main effect of age, $F(2,110) = 11.97$, $p < .0001$, $\eta_p^2 = .18$; interaction Age x Emotion, $F(6,330) = 10.45$, $p < .0001$, $\eta_p^2 = .16$]. Furthermore, we looked for possible effects of gender, but they were non-significant (main effect of gender and interactions with age and emotion ns, ps > .3).

To further explore the developmental trajectories of the four emotions, we computed partial correlations between years of age and categorization accuracy for each emotion (controlling for years of musical training). Accuracy decreased linearly with advancing age for sadness ($r = -.44$, $p < .01$) and fear ($r = -.55$, $p < .01$), but not for happiness ($r = -.2$, $p > .05$) nor peacefulness ($r = .2$, $p < .05$). This pattern confirms that age was associated with valence-specific changes: decreased recognition for negative emotions, and relative invariance or slight improvement for positive ones.

DISTRIBUTION OF INACCURATE RESPONSES | We also analysed the pattern of inaccurate responses and whether it differed across age groups. For each emotion, an ANCOVA was carried out with the non-intended response categories as repeated-measures factor (three non-intended emotions and ambivalent responses), age as between-subjects factor, and musical training as covariate. For happy music, misclassifications were rare and included mainly ambivalent (4%) and peaceful responses [2%; 0.5% for sad and 0.1% for fear responses, ps < .01; main effect of category, $F(3,330) = 14.5$, $p < .0001$, $\eta_p^2 = .12$]. For peaceful music, inaccurate responses were frequent for sadness (20%, $p < .01$), and less for ambivalence (14%) and happiness [11%, $p > .3$; 0.3% for fear, $p < .01$; main effect of category, $F(3,330) = 35.9$, $p < .0001$, $\eta_p^2 = .25$]. For sad music, the most common non-intended responses were for peacefulness (22%, $p < .01$), though ambivalent responses were also frequent, 15% [$p < .01$; 2.1% for fear and 1.4% for happy responses, $p > .2$; main effect of category, $F(3,330) = 50.83$, $p < .0001$, $\eta_p^2 = .32$]. The confusion of peacefulness and sadness with each other has already been reported (Gosselin et al., 2005, 2007; Vieillard et al., 2008). A possible reason is that the two types of music share some features, like tempo. Finally, scary music was misclassified as sad (18%, $p < .01$), and less as ambivalent and happy, both 13% [$p > .9$; only 5% as peaceful, $p < .01$; main effect of category, $F(3,330) = 13.08$, $p < .0001$, $\eta_p^2 = .11$]. The confusion of scary music with sadness might be due to similarities in valence (both are negative). The distribution of non-intended responses was similar across age groups for happy and scary music [interactions Age x Category ns, *p*s > .1]. For peaceful and sad music, however, there were age-related differences: for peaceful music, younger participants gave more non-intended responses for sadness than middle-aged and older ones [*p*s < .01; middle-aged and older participants did not differ, $p > .9$; interaction Age x Category, $F(6,330) = 7.43$, $p < .0001$, $\eta_p^2 = .12$]; for sad music, younger participants gave less responses for peacefulness than the older age groups [*p*s < .01; middle-aged on older participants did not differ, $p > .1$; interaction Age x Category, $F(6,330) = 5.26$, $p < .0001$, $\eta_p^2 = .09$]. These differences suggest that younger participants were more likely to provide the highest rating to the scale *sadness,* while older participants were more likely to provide the highest rating to the scale *peacefulness.* In response to stimuli containing a certain degree of ambiguity and some shared structural cues, as it is the case of sad and peaceful music, older participants might have weighted positive information more strongly than negative one, and the opposite might have occurred for younger participants – age-related trend towards positivity (e.g., Carstensen & Mikels, 2005; Samanez-Larkin & Carstensen, 2011).

MUSICAL EXPERTISE AND EMOTION RECOGNITION | We examined whether musical expertise modulates emotion recognition in music by computing a partial correlation between years of training and global categorization accuracy (controlling for age in years). The longer participants were trained, the more accurate they were. This was observed when the analysis included only the subset of musically trained participants ($r = .48$, $p < .01$), and also when it was conducted on all participants, including untrained ones ($r = .33$, $p < .01$). A scatterplot of categorization accuracy rates as a function of years of training is presented in Figure 1.

Figure 1. Scatterplot of emotion recognition accuracy (%) of musically trained participants by years of training (Study 1).



MUSIC STRUCTURE CUES AS PREDICTORS OF RESPONSES | The excerpts communicated emotions only through structural compositional cues, and so listeners had to be sensitive to them to succeed in the task. Previous research has already shown that music cues predict participants' responses in emotion perception tasks (e.g., Balkwill & Thompson, 1999; Gabrielsson & Lindström, 2010; Juslin, 1997, 2000). We computed multiple simultaneous regression analyses to explore how participants of different age groups relied on different cues to respond. Each music stimulus was characterized for the following parameters: tempo (crochet beats per minute), mode (major, minor, or indefinite), tone density (number of melodic events / total duration), melodic pitch range (distance, in semitones, between the highest and the lowest tone of the melody) and pedal (with or without); scary excerpts were also evaluated by an expert musician for dissonance, unexpected events and rhythmic irregularities in five-point scales (1 = consonant, expected, regular; 5 = dissonant, unexpected, irregular; happy, sad and peaceful excerpts were all consonant, expected, and regular – they were given 1 on these parameters). The parameter values for each excerpt are presented in Appendix 3. All cues varied significantly across emotions (*p*s < .05). To keep the set of predictors small, only tempo, mode, tone density, pitch range, pedal and rhythmic irregularities entered the regression analyses; dissonance and unexpected events were excluded because they were very highly correlated with rhythmic irregularities (*r*s > .8, $p < .01$). The dependent variable was the mean raw ratings provided by participants on each emotion scale. Separate regressions were computed for each emotion and age group. The main findings are presented in Table 5, including beta weights for each cue and the proportion of variance explained by the whole model (adjusted R^2). Responses for the four emotion scales were significantly predicted by some combination of cues in younger, middle-aged and older participants (*p*s < .05). The proportion of explained variance ranged between .75 (fear/threat in older participants) and .91 (sadness in younger participants). The predictors reaching significant beta weights, or their direction (i.e., positive or negative), differed across emotions, as can be seen in Table 5. This is evidence that the participants relied on differentiated patterns of cues to provide their ratings on the four emotions (e.g., Balkwill & Thompson, 1999). Note that no less than two cues yielded significant beta weights for each emotion, showing that participants used a constellation of cues to respond, and not a single one. Mode and pedal were particularly informative to guide responses – they reached significant beta weights in all groups and emotions. The most important predictors of (higher) happy ratings were higher tone density, major mode, no pedal and regular rhythm; of peaceful ratings, major mode and pedal; of sad ratings, minor mode and pedal; and of scary ratings, lower tone density, minor mode, no pedal and rhythmic irregularities. These emotion-specific associations between musical cues and subjective ratings are in good agreement with previous descriptions (e.g., Gabrielsson, 2009; Gabrielsson & Lindström, 2010).

For happy and peaceful ratings, the three age groups were equally consistent in using music cues to respond, as indicated by highly similar predicted variances (see adjusted R^2 in Table 5). The cues reaching significant beta weights were also the same. These quantitative and qualitative similarities are consistent with the relative invariance observed across age groups in categorization accuracy for happiness and peacefulness. For sad and scary responses, the general predictive strength of the cues was somewhat lower in older than in younger participants (differences in adjusted R^2: -.08 for sadness; -.07 for fear). This suggests that age might be associated with slightly decreased consistency in how listeners rely on music structural parameters. There were also subtle differences in the weighting of the cues: for sad ratings, rhythmic irregularities were a marginally significant predictor in younger but not in middle-aged and older participants, and tone density was a marginally significant predictor in older but not in middle-aged and younger ones. Thus, is it possible that the age-related changes in categorization accuracy for sadness and fear are linked with a less consistent use of music cues, as well as with qualitative differences in the relative weighting of these cues.

Table 5. Results from multiple regression analyses on listeners' utilization of music compositional cues for subjective ratings, as a function of emotion and age (Study 1). Values correspond to beta weights and adjusted R^2.

	Emotion			
Age group / Cue	Happy	Peaceful	Sad	Scary
Younger				
Tone Density	$.27^*$	-.06	-.02	$-.38^*$
Pitch Range	-.01	-.06	-.07	.06
Mode	$-.53^*$	$-.41^*$	$.53^*$	$.44^*$
Pedal	$-.33^*$	$.79^*$	$.67^*$	$-.65^*$
Tempo	.11	-.08	-.07	-.13
Rhythmic Irregularities	$-.18^*$	-.09	$.12^!$	$.22^*$
Adjusted R^2	*.87**	*.87**	*.91**	*.82**
Middle-aged				
Tone Density	$.34^*$	-.02	-.02	$-.51^*$
Pitch Range	-.05	-.03	-.06	.1
Mode	$-.39^*$	$-.43^*$	$.63^*$	$.29^*$
Pedal	$-.35^*$	$.79^*$	$.57^*$	$-.77^*$
Tempo	.13	-.08	-.09	-.12
Rhythmic Irregularities	$-.26^*$	.05	-.09	$.26^*$
Adjusted R^2	*.86**	*.86**	*.89**	*.83**
Older				
Tone Density	$.43^*$	-.11	$-.19^!$	$-.4^*$
Pitch Range	-.03	-.01	-.07	-.07
Mode	$-.32^*$	$-.18^*$	$.55^*$	$.25^*$
Pedal	$-.28^*$	$.78^*$	$.34^*$	$-.65^*$
Tempo	.13	-.06	-.14	-.08
Rhythmic Irregularities	$-.19^*$	-.02	-.11	$.37^*$
Adjusted R^2	*.86**	*.83**	*.83**	*.75**

$^*p < .05$; $^! p \leq .1$

6.2.3. Discussion

We showed that emotion recognition in music changes across the adult life. Advancing age was associated with decreased accuracy in sadness and fear, as conveyed by compositional structural cues in instrumental music. These changes were significant already in middle-aged participants, not only in older ones. Sensitivity to happiness and peacefulness, on the other hand, remained relatively invariant. Multiple regression analyses revealed that music structure cues predict participants' subjective responses, and that there are age-related quantitative and qualitative differences in how these cues are used, notably for sadness and fear. We also found that the number of years of musical training is associated with enhanced categorization accuracy of musical emotions.

The pattern of changes for sadness and fear, and the concomitant stability for happiness and peacefulness, indicates that ageing modulates emotion recognition in music in a valence-specific manner: decreased responsiveness is observed for negative but not for positive emotions. This suggests that the positivity effect in ageing occurs for musical emotions, as previously observed in other domains (e.g., Kisley et al., 2007; Mitchell et al., 2011; Williams et al., 2006). Note that sadness in music might not always be a negative emotion, as it is in other contexts (e.g., Zentner et al., 2008). Listeners usually do not avoid listening to sad music and can even experience pleasure in response to it (e.g., Salimpoor et al., 2010). Nevertheless, sad excerpts in our study were at least relatively more negative than peaceful and happy ones. Indeed, an independent sample of 32 listeners (mean age = 33.3; $SD = 12.2$; range = 18 - 56) judged happy and peaceful music as similarly pleasant in a 10-point valence scale (z-transformed ratings, 0.5 and 0.3, respectively; $p > .3$), and sad and scary excerpts were considered unpleasant (-0.14 and -0.66, respectively; fear was considered more unpleasant than sadness, $p < .01$). Thus, the emotions that underwent age-related effects match those that are more unpleasant.

Because categorization accuracy for happiness was very high, 93% correct on average, it might be argued that the invariance observed across age groups reflects a ceiling effect. Notwithstanding, the observed rates still lie below the maximum score, and thus they would allow for variability in performance. Additionally, raw ratings for happiness were also highly stable across age groups (younger, 7.8; middle-aged, 7.7; older, 7.6; maximum 9), and this suggests true invariance, not lack of stimulus sensitivity to differences across listeners.

Can decline in general cognitive functioning underlie the age-related differences in emotion recognition? This would predict changes for the emotions which are hardest to recognize, and we did not find that: there were changes for a relatively easy emotion, sadness, and stability for a relatively difficult one, peacefulness. Furthermore, older participants had a near ceiling performance on the MMSE, which indicates that they did not have prominent global impairments. Changes for emotion recognition might thus be primary in origin. Still, we cannot rule out the possible role of subtle cognitive changes, namely in the case of middle-aged participants, who were not inspected for domain-general abilities. In Study 2 we examine directly whether there are associations between emotion recognition and general cognitive performance. We also compare expert musicians with untrained listeners to follow-up the effect of musical training unveiled in Study 1. In Study 2 we include two age groups, younger and middle-aged adults; we selected middle-aged instead of older participants because developmental changes in the middle-years are frequently neglected in the literature. Additionally, we observed relative stability in the pattern of changes from middle-age onward (for similar results in other modalities, see Isaacowitz et al., 2007).

6.3. STUDY 2

6.3.1. Method

PARTICIPANTS | Eighty participants took part in this study (none of them participated in Study 1). They were distributed into four groups according to musical expertise (musicians and controls) and age (younger and middle-aged), 20 per group (10 female). Table 6 presents their demographic and background characteristics. Younger participants were 23 years old on average (*SD* = 3.2), and middle-aged ones were 47.7 years old (*SD* = 4.4). Musicians were instrumentalists who played piano (*n* = 18), flute (*n* = 5), violin (*n* = 5), guitar (*n* = 3), double bass (*n* = 2), clarinet (*n* = 1), drums (*n* = 1), cello (*n* = 1), oboe (*n* = 1), viola (*n* = 1), trombone (*n* = 1) or accordion (*n* = 1); in addition to instrumental training, two of them had vocal training in classical singing. They had at least 8 years of formal musical training started in childhood, and practiced regularly their instruments at the moment of testing. Younger ones were advanced music students or professional musicians, and older ones were music teachers and/or orchestral performers. They were recruited from local music schools and orchestras, including Conservatório de Música do Porto, Escola Superior de Música e Artes do Espectáculo do Instituto Politécnico do Porto and Orquestra Sinfónica do Porto Casa da Música. Younger and older musicians were matched for the number of years of training, age of commencement and instrumental practice per week (*F*s < 1; see Table 6). Controls have never had formal music lessons nor played any instruments. They were recruited from several local communities. The four groups were matched for education level, as assessed by the number of years attending school (*F*s < 1). All participants had normal or corrected-to-normal vision and reported no head trauma, substance abuse nor major psychiatric or neurological illnesses. The groups rated themselves similarly for hearing acuity in a scale from 1, *very good,* to 6, *very bad* (M = 2.1; SD = .9; ps > .1). Participants were asked to self-rate their interest in music in a scale from 1, *very high*, to 6, *very low*; all reported that music is important to them, but musicians revealed stronger interest (1.03) than controls [2.18, $F(1,76) = 51.94$, $p < .001$, $\eta_p^2 = .4$]. Participants were financially compensated for their participation.

Table 6. Demographic and background characteristics of controls and musicians for each age group (Study 2). *SD*s are presented in parentheses.

	Controls		Musicians	
Characteristics	Younger	Middle-aged	Younger	Middle-aged
Age (years)	22.7 (2.8)	47 (4)	23.4 (3.6)	48.4 (4.8)
Education (years)	15.7 (1.5)	17.1 (4)	15.4 (1.8)	16.5 (3.6)
Musical training (years)	-	-	11.3 (3.1)	12.6 (3.2)
Age of commencement (years)	-	-	9.2 (2.5)	8.4 (4.5)
Instrumental practice (hours/week)	-	-	12.7 (11.6)	12.3 (11)
Montreal Cognitive Assessment (/30)	28.2 (1.3)	26.8 (1.9)	28.1 (1.6)	27.8 (1.2)
Raven's APM (problems solved, /36)	20.5 (4.7)	13.9 (5.1)	19.7 (4.9)	16.8 (4.5)
Stroop test[1]				
Baseline (words/s)	2.09 (0.3)	2.14 (0.3)	2.37 (0.3)	2.39 (0.3)
Conflict (colors/s)	1.04 (0.2)	0.87 (0.1)	1.08 (0.2)	1.04 (0.2)
Ten-Item Personality Inventory				
Extraversion (/7)	4.8 (1.5)	4.6 (1.7)	4.8 (1)	4.9 (1.9)
Agreeableness (/7)	5.3 (0.8)	5.5 (1)	5.2 (0.9)	5.5 (0.9)
Conscientiousness (/7)	5 (1.3)	5.5 (1.3)	4.2 (1.4)	5.1 (1.5)
Emotional stability (/7)	4.2 (1.4)	4.4 (2)	3.6 (1.2)	4.5 (1.5)
Openness to experience (/7)	5.6 (0.8)	5.6 (1)	5.6 (1)	6.1 (0.9)
Autism Spectrum Quotient[2] (/50)	17.5 (5.4)	18 (4.8)	17.5 (5.7)	17.5 (5.5)

[1] We computed the number of items named per s: number of correct responses/time taken to perform the task
[2] Higher values indicate higher autistic-like socio-communicative traits (cut-off for clinically significant levels of autistic traits, 32)

COGNITIVE AND PERSONALITY ASSESSMENT | The results of cognitive tests and personality scales are summarized in Table 6. Participants were screened for global cognitive dysfunction with the Montreal Cognitive Assessment (MOCA; www.mocatest.org; Portuguese version, Simões, Firmino, Vilar, & Martins, 2007), a brief instrument which inspects attention and concentration, executive functions, memory, language, visuoconstructional skills, conceptual thinking, calculations and orientation. Nonverbal general intelligence was assessed with a 20-minute timed version of the Raven's Advanced Progressive Matrices; this version correlates strongly with the untimed one (Hamel & Schmittmann, 2006). To examine executive control, we used a Stroop task (Trenerry, Crosson, Deboe, & Leber, 1995). In the version used here participants performed speeded naming in two conditions: baseline, consisting of reading words denoting colour names (blue, *azul;* pink, *rosa;* grey, *cinza*; green, *verde*);

and an incongruous condition, consisting of naming the ink of written words that denoted an incongruent color name (e.g., the word "blue" printed in green ink; Portuguese version, Castro, Martins, & Cunha, 2003). Musicians and controls did not differ in the MOCA and Raven's matrices (see Table 6; *p*s > .19). This indicates that they were similar for the general abilities assessed by these tests. In the Stroop task, musicians were faster than controls in both baseline [$F(1,76) = 16.2$, $p < .001$, $\eta_p^2 = .18$] and incongruous conditions [$F(1,76) = 6.7$, $p < .05$, $\eta_p^2 = .08$]. Their advantage in the incongruous condition only approached significance when baseline differences were controlled for [ANCOVA, $F(1,75) = 3.7$, $p = .06$, $\eta_p^2 = .05$]. This suggests that musicians' enhanced executive control was partially due to an advantage in processing speed; effects of musical expertise on executive control have been found by Bialystok and DePape (2009). Middle-aged participants scored lower than younger ones in the MOCA [$F(1,76) = 5.47$, $p < .05$, $\eta_p^2 = .07$], Raven's matrices [$F(1,76) = 19.5$, $p < .01$, $\eta_p^2 = .2$] and in the conflict condition of Stroop [$F(1,76) = 7.7$, $p < .01$, $\eta_p^2 = .09$]. These ageing effects were observed similarly for controls and for musicians (interactions Age x Expertise ns, *p*s > .05). The only measure that did not show age-related decline was the baseline condition of the Stroop test ($F < 1$; interaction Age x Expertise ns, $F < 1$).

Because personality characteristics can influence emotion processing (e.g., Hamann & Canli, 2004; Matsumoto et al., 2000; Mill et al., 2009), participants completed the Ten-Item Personality Inventory (TIPI), a brief questionnaire that inspects the Big-5 personality domains: extraversion, agreeableness, conscientiousness, emotional stability and openness to experience (Gosling, Rentfrow, & Swann Jr., 2003; Portuguese version, Lima & Castro, 2009). They also completed The Autism-Spectrum Quotient (AQ), a questionnaire that assesses socio-communicative traits associated with the autistic spectrum in neurotypical adults (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001; Portuguese version, Castro & Lima, 2009). Autistic traits correlate with structural and functional differences in brain regions involved in emotion processing, even in neurotypicals (Hagen et al., 2011; Martino et al., 2009). Both age and expertise groups had similar personality and socio-communicative characteristics, as assessed by these measures (*F*s < 1; see Table 6).

MUSIC STIMULI | As in Study 1, stimuli were taken from the set validated by Vieillard and colleagues (2008). Here were used 10 excerpts per emotion category, 40 in total. They were pseudo-randomized and divided into two blocks of 20 trials each.

PROCEDURE | Participants were tested in an individual session lasting about 90 minutes. The instructions and structure of the experimental task were the same as in Study 1.

Besides the experimental task and the cognitive and personality tests, participants also performed a forced-choice identification of seven emotional tones in speech prosody (these results are presented in Chapter VIII).

6.3.2. Results

Comparisons across groups were based on derived accuracy rates, as in Study 1. Table 7 presents the percentage of accurate categorizations for each emotion, and the distribution of inaccurate and ambivalent responses, separately for age and expertise groups. Agreement rates ranged between 42% (peacefulness in younger controls) and 96% (happiness in younger controls).

EFFECTS OF AGE AND MUSICAL EXPERTISE ON ACCURACY | Correct identifications per emotion were arcsine transformed and submitted to an ANOVA, with emotion as repeated-measures factor (happy, peaceful, sad and scary), and age (young and middle-age) and expertise (controls and musicians) as between-subject factors. Happiness reached the highest accuracy, 94%, and peacefulness the lowest, 50%; sadness and fear elicited intermediate and similar accuracies, 67% [*p*s < .001; main effect of emotion, $F(3,228) = 63.62$, $p < .001$, $\eta_p^2 = .46$].

Table 7. Distribution of responses (%) for each intended emotion, as function of age and expertise (Study 2). Diagonal cells in bold indicate accurate categorizations. Standard errors are presented in parentheses.

	Distribution of responses (%)									
	Controls					Musicians				
Age group/ excerpt type	Happy	Peaceful	Sad	Scary	Ambivalent	Happy	Peaceful	Sad	Scary	Ambivalent
Younger										
Happy	**96 (1.3)**	1	1	0	2	**95 (2.8)**	2	0	0	4
Peaceful	14	**42 (7.3)**	27	0	18	22	**46 (7.5)**	11	0	22
Sad	0	6	**83 (3)**	3	9	0	12	**70 (6.6)**	1	18
Scary	6	0	8	**77 (4.4)**	10	2	2	15	**70 (7)**	11
Middle-aged										
Happy	**91 (3.8)**	0	3	2	4	**95 (2.3)**	1	0	1	5
Peaceful	13	**51 (5.4)**	11	3	23	12	**64 (5.7)**	12	0	13
Sad	6	20	**52 (7.1)**	1	22	1	18	**64 (5.6)**	1	17
Scary	17	2	14	**49 (5.8)**	18	8	0	11	**72 (5.5)**	10

Concerning the impact of ageing, the pattern obtained in Study 1 was replicated: planned comparisons confirmed that middle-aged participants were less accurate than younger ones at categorizing the negative emotions, sadness (76% for younger vs. 58% for middle-aged, $p < .01$) and fear (73% for younger vs. 60% for middle-aged, $p < .05$). By contrast, the recognition of happiness remained stable ($p > .4$), and accuracy rates for peacefulness increased [44% for younger vs. 57% for middle-aged, $p < .05$; main effect of age ns, $p > .1$; interaction Age x Emotion, $F(3,228) = 7.3$, $p < .001$, $\eta_p^2 = .09$]. Note that the age-related changes for sadness and fear were observed in controls (*p*s < .01), but not in musicians [*p*s > .9; interaction Age x Expertise, $F(1,76) = 5.79$, $p < .05$, $\eta_p^2 = .07$]: younger and middle-aged musicians had similar accuracy rates. Our prediction that musical expertise would be associated with enhanced accuracy was confirmed, though for older musicians only: middle-aged musicians had higher accuracy (73%) than age-matched controls (61%, $p < .05$), but younger musicians (70%) and controls (73%) performed similarly [$p > .3$; main effect of expertise ns, $p > .1$; see above the interaction Age x Expertise]. This effect was independent of emotion category (interactions Expertise x Emotion, and Age x Expertise x Emotion ns, *p*s > .3)[4]. The pattern of age and expertise effects was globally replicated in ANOVAS on raw ratings (see raw ratings in Appendix 4) and on unbiased accuracy rates (see unbiased rate in Appendix 5), except that the age-related increase in responsiveness to peacefulness was not significant [raw ratings: middle-aged participants provided lower ratings than younger ones for sadness and fear, *p*s < .05, but for happiness and peacefulness ratings remained invariant, *p*s > .4, interaction Age x Emotion, $F(3,228) = 7.55$, $p < .001$, $\eta_p^2 = .09$; middle-aged musicians provided higher ratings on the intended emotions than controls, $p < .05$, interaction Age x Expertise, $F(1,76) = 7.01$, $p < .01$, $\eta_p^2 = .08$; unbiased rates: middle-aged participants were less accurate than younger ones for sadness and fear, *p*s < .05, but similar for happiness and peacefulness, *p*s > .05, interaction Age x Emotion, $F(3,228) = 6.65$, $p < .001$, $\eta_p^2 = .08$; middle-aged musicians performed better than age-matched controls, $p < .05$, interaction Age x Expertise,

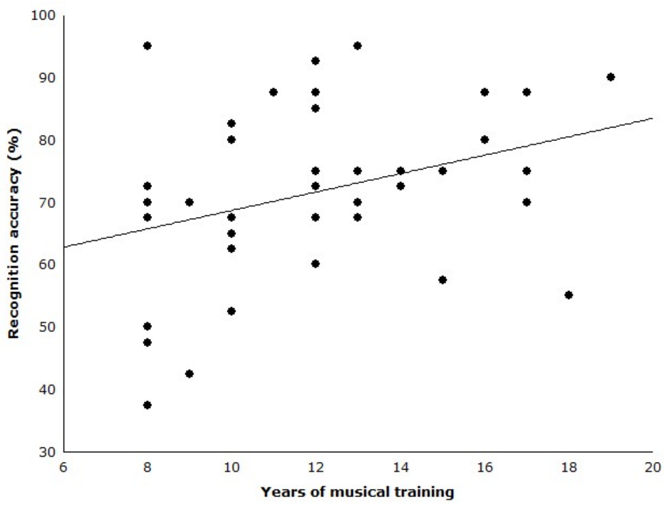
[4] We explored whether musicians' performance varied according to the played instrument. Because they were widely distributed across many different instruments (see Methods), we first categorized them by type of instruments [keyboard, $n = 18$; strings, $n = 12$; woodwinds, $n = 7$; others, $n = 3$ (drums, accordion, and trombone)] and then computed an ANOVA with accuracy rates as dependent measure. Performance was similar across instrument types ($p > .1$).

$F(1,76) = 4.35$, $p < .05$, $\eta_p^2 = .05$]. Gender did not impact results (main effect of gender ns, $F < 1$; interactions with age, expertise and emotion ns, ps > .05).

We also computed Pearson correlations between years of age and categorization accuracy for each emotion. These analyses were performed separately for controls and musicians because age affected differently both groups. For controls, recognition accuracy decreased with age for sadness ($r = -.56$, $p < .0001$) and fear ($r = -.52$, $p < .001$), but not for happiness ($r = -.2$, $p > .1$) and peacefulness ($r = .09$, $p > .5$). For musicians, recognition accuracy was not linked with age (ps > .05).

The correlation between the number of years of musical training and recognition rates obtained in Study 1 was replicated in this sample of expert musicians. The longer they were trained, the more accurate they were at categorizing the excerpts ($r = .33$, $p < .05$), as illustrated in Figure 2. The correlation was marginally significant when it was conducted on all participants, musicians and controls, $r = .21$, $p = .07$.

Figure 2. Scatterplot of emotion recognition accuracy (%) of musicians by years of training (Study 2).



DISTRIBUTION OF INACCURATE RESPONSES | An ANOVA was carried out for each emotion, with the non-intended response categories as repeated-measures factor (three non-intended emotions and ambivalent responses), and age and expertise as between-subject factors. The pattern of responses broadly mirrored the one reported in Study 1. For happy music, misclassifications were rare and included chiefly ambivalent responses, 4% [0.75% for each of the other categories, ps < .001; main effect of category, $F(3,228) = 8.74$, $p < .001$, $\eta_p^2 = .1$]. For peaceful music, non-intended responses were similarly frequent for ambivalence (19%), sadness (15%) and happiness [15%, $p > .3$; 1% for fear, ps < .001; main effect of category, $F(3,228) = 32.3$, $p < .001$, $\eta_p^2 = .3$]. For sad music, non-intended responses were common for ambivalence (16%) and peacefulness [14%, $p > .3$; 2% for happiness and 1% for fear, ps < .01; main effect of category, $F(3,228) = 54.98$, $p < .001$, $\eta_p^2 = .45$]. Scary music was misclassified mainly as sad and ambivalent (both 12%), and less as happy [8%, $p < .05$; 1% for peacefulness, ps < .05; main effect of category, $F(3,228) = 19.78$, $p < .001$, $\eta_p^2 = .2$]. The pattern of non-intended responses for happy and scary music was similar across age and expertise groups (interactions Age x Category, Expertise x Category, and Age x Expertise x Category ns, ps > .05). Regarding peaceful music, younger participants gave more sad responses than middle-aged ones, and controls (younger and older) gave more ambivalent than happy responses; these differences were not observed in musicians [interaction Age x Expertise x Category, $F(3,228) = 3.78$, $p < .05$, $\eta_p^2 = .05$; interactions Emotion x Age and Emotion x Expertise ns, ps > .2]. As for sad music, younger participants gave less responses for peacefulness than middle-aged ones [interaction Age x Category, $F(3,228) = 5.16$, $p < .001$, $\eta_p^2 = .06$; the other interactions were ns, ps > .05].

CORRELATIONS BETWEEN GENERAL COGNITIVE ABILITIES, PERSONALITY AND EMOTION RECOGNITION | The four groups were matched for background variables such as education and sex. However, they differed for cognitive abilities (see Methods): middle-aged participants were worse than younger ones for global cognitive functioning (MOCA), nonverbal intelligence (Raven's matrices) and executive control (Stroop test, conflict condition); and musicians had faster processing speed than controls and a trend for superior executive control (Stroop test, both conditions). Do these differences account for performance in emotion recognition? We computed correlation analyses between global categorization accuracy (score averaged across the four emotions) and scores on the cognitive measures. Correlation coefficients and *p* values are displayed in Table 8. No significant associations were found. This suggests that domain-general cognitive abilities did not influence emotion recognition in music[5]. Another potential confound was musicians' stronger interest in music, but this variable was also not related with recognition accuracy.

The four groups had similar personality characteristics, as assessed by the TIPI, and socio-communicative traits, as assessed by the AQ. We explored possible associations between these measures and emotion recognition, but all analyses yielded non-significant results (see Table 8).

Table 8. Correlations between cognitive and personality measures and global accuracy in emotion recognition in music.

	Emotion recognition accuracy	
Cognitive/personality measure	*r*	*p*
Montreal Cognitive Assessment	.1	.37
Raven's Advanced Progressive Matrices	.19	.1
Stroop test		
Baseline condition	.15	.18
Conflict condition	.07	.52
Ten-Item Personality Inventory		
Extraversion	.03	.8
Agreeableness	-.13	.27
Conscientiousness	-.05	.67
Emotional stability	-.17	.12
Openness to experience	.06	.62
Autism Spectrum Quotient	.1	.38
Interest in music	-.16	.17

[5] We further confirmed that age and expertise effects in emotion recognition were not explained by general cognitive abilities by computing ANCOVAs on categorization accuracy with MOCA, Raven's Matrices and Stroop test as covariates.

MUSIC STRUCTURE CUES AS PREDICTORS OF RESPONSES | We computed multiple simultaneous regression analyses to examine how age and expertise groups relied upon excerpts' compositional cues to respond. The procedure was the same as in Study 1. Table 9 depicts the main findings. Responses for the four emotions were significantly predicted by some combination of cues in younger and older controls and musicians (*p*s $< .05$). The proportion of explained variance ranged between .68 (fear/threat in middle-aged controls) and .92 (sadness in middle-aged musicians and in younger controls). The cues reaching significant beta weights for each emotion are in general agreement with the results of Study 1. The most important predictors of happy ratings were higher tone density, major mode, and no pedal (regular rhythm was also significant for middle-aged musicians); of peaceful ratings, major mode and pedal (slow tempo was also significant for middle-aged musicians); of sad ratings, minor mode, pedal and rhythmic irregularities; and of scary ratings, low tone density, minor mode and no pedal (differing from Study 1, rhythmic irregularities were significant only for middle-aged musicians).

Table 9. Results from multiple regression analyses on listener's utilization of music compositional cues for emotion ratings, as a function of emotion and age (Study 2). Values correspond to beta weights and adjusted R^2.

	Controls				Musicians			
Age group / Cue	Happy	Peaceful	Sad	Scary	Happy	Peaceful	Sad	Scary
Younger								
Tone Density	.25*	.06	-.15	-.32*	.16!	.15	.01	-.4*
Pitch Range	.0	.01	-.03	.06	.05	.01	-.06	-.04
Mode	-.5*	-.53*	.47*	.53*	-.71*	-.24*	.73*	.4*
Pedal	-.28!	.81*	.6*	-.55*	-.26*	.95*	.56*	-.83*
Tempo	.15	-.15	-.09	-.03	.11	-.13	-.06	-.15
Rhyth. Irreg.	-.15	.17!	-.02	.15	-.01	.06	-.24*	.04
Adjusted R^2	*.85**	*.83**	*.92**	*.79**	*.91**	*.90**	*.92**	*.74**
Middle-aged								
Tone Density	.23!	.1	-.04	-.43*	.21!	.14	.01	-.36*
Pitch Range	.04	.02	-.09	-.05	.02	-.02	-.02	-.04
Mode	-.51*	-.41*	.63*	.21	-.46*	-.41*	.63*	.35*
Pedal	-.2!	.91*	.37*	-.93*	-.33*	.82*	.65*	-.65*
Tempo	.2	-.08	-.2	-.21	.19	-.29*	-.09	.0
Rhyth. Irreg.	-.15	.11!	-.25*	.1	-.19*	.13	-.1	.25*
Adjusted R^2	*.84**	*.86**	*.84**	*.68**	*.89**	*.87**	*.92**	*.79**

Note. Rhyth. Irreg. = Rhythmic irregularities
* $p < .05$; ! $p \leq .1$

Two group comparisons are of particular interest to illuminate age and expertise effects on accuracy rates: younger and middle-aged controls, where valence-specific categorization accuracy differences were detected, and middle-aged musicians and controls, where the impact of expertise was significant. As observed in Study 1, the predicted variances (adjusted R^2) for younger and middle-aged controls were similar for happy and peaceful responses, and the cues reaching significant beta weights were also similar (see Table 9). Accordingly, no age-related decrements were observed in the categorization of these emotions. For sad and scary responses, the predictive strength of the cues was lower in middle-aged than in younger participants (differences in adjusted R^2: -.08 for sadness; -.11 for fear), suggesting that they used them less consistently. Additionally, some cues were weighted differently: rhythmic irregularities were a stronger predictor of sad ratings in middle-aged than in younger participants; and mode was a stronger predictor of scary ratings in younger than in middle-aged participants. These results lend further support to the hypothesis that the age-related changes in accuracy for sadness and fear are associated with quantitative and qualitative differences in the use of structural cues. Concerning the impact of expertise, the predictive strength of the cues was higher in middle-aged musicians than in controls (+.06, on average), and this might relate to the musicians' enhanced accuracy. Predictors yielding significant beta weights also differed: for happiness, rhythmic irregularities were a stronger predictor in musicians than in controls; for peacefulness, tempo was a stronger predictor in musicians than in controls; for sadness, rhythmic irregularities were a stronger predictor in controls than in musicians; and for fear, mode and rhythmic irregularities were stronger predictors in musicians than in controls (see Table 9). Thus, the effect of expertise might also be a consequence of differential weighting of music structure cues.

6.3.3. Discussion

In Study 2 we replicated in a different sample the pattern of valence-specific age-related changes found in Study 1: middle-aged participants were less accurate than younger ones at recognizing sadness and fear, but for happiness and peacefulness no decline was observed. Middle-aged participants performed worse than younger ones concerning global cognitive functioning, nonverbal intelligence and executive control. This is consistent with previous findings on neurocognitive ageing (e.g., Hedden & Gabrieli, 2004; Salthouse, 2009). However, these domain-general cognitive abilities were not associated with performance in emotion recognition, suggesting that ageing may modulate emotions directly. That is, the effects are not a by-product or general unspecific factors. Hearing loss is also an unlikely cause for the observed changes because younger and middle-aged participants did not differ for self-rated hearing status (see Methods). It would have been preferable to use an objective measure instead of self-assessment, but we think this is not a serious limitation because participants in our sample were relatively young and volume was adjusted individually to a comfortable hearing level.

With a more rigorous control of musical training, we replicated the correlation between the length of training and enhanced categorization accuracy. Additionally, expert musicians as a group were more accurate than controls, though this was observed for the older participants only. This effect was also independent of general cognitive factors. It might be linked with a more consistent reliance on music compositional cues, as revealed by the multiple regression analyses. That younger musicians, as a group, were not more accurate than controls is unanticipated: we can speculate that the null finding is a consequence of sampling error, or that the stimuli and task were insensitive to group differences in younger participants. Alternatively, it is possible that musically untrained younger adults can indeed reach a performance level as high as experts do in emotion recognition in music. The fact that expertise effects differed across age cohorts highlights the importance of examining different groups when general inferences are to be made concerning how musical training shapes cognition.

An additional finding of Study 2 was that age-related effects were not detected in musicians. Previous research on how environmental factors can influence cognitive ageing suggests that expertise in a certain domain might attenuate age-related changes for well-practiced tasks that are highly related with that domain (for a review, see

Kramer, Bherer, Colcombe, Dong, & Greenough, 2004). Because an important part of musicians' training and professional practice throughout their life concerns the expression and perception of musical emotions, this intensive and permanent practice might offset normative age-related trajectories. In fact, it is of paramount importance for musicians to maintain sensitivity to positive as well as to negative emotional expressions in music in order to be effective at communicating them. This incidental result deserves to be followed-up in future studies.

6.4. GENERAL DISCUSSION

The results of these two studies provide converging evidence that ageing and musical expertise modulate emotion recognition in music. Advancing age was associated with decreased responsiveness to the negative emotions of sadness and fear, while responsiveness to happiness and peacefulness remained relatively invariant. The differences were observed between younger and older adults, over sixty years of age (Study 1), but they were significant already in middle adulthood, around the forties (Studies 1 and 2). The number of years of musical training correlated with enhanced categorization accuracy (Studies 1 and 2), and musicians as a group were more accurate than untrained listeners, but this was observed for middle-aged participants only (Study 2). Ageing and expertise effects in emotion recognition cannot be accounted for by differences in domain-general cognitive abilities, personality traits and interest in music (Study 2). This suggests that they are primary in origin, that is, emotion-specific. The effects might be linked with differences in how listeners used musical structure cues to respond: multiple regression analyses revealed that older participants use music cues less consistently than younger ones to provide sad and fear responses, but they are similarly consistent for happy and peaceful responses (Studies 1 and 2); musicians were more consistent than controls in using the cues (Study 2).

AGEING AND EMOTION RECOGNITION IN MUSIC | Previous research on facial expressions (e.g., Calder et al., 2003; Ruffman et al., 2008), speech prosody (e.g., Laukka & Juslin, 2007; Mill et al., 2009; Paulmann et al., 2008b) and body postures (Ruffman et al., 2009; Ruffman et al., 2008) indicates that ageing is associated with decreased accuracy in emotion recognition, particularly for negative emotions, namely sadness and fear. To the best of our knowledge, the present studies are the first to examine whether these age-related changes also occur for emotions in music, as conveyed by music structure cues, such as mode and pitch range. Finding that responsiveness to sadness and fear on the basis of these cues decreases with age is evidence that music can undergo the same developmental processes as other socially relevant stimuli. Furthermore, because we covered the full range of the adult life span (Study 1), we were able to show that the changes are significant already in middle-age, as for other emotion modalities (Calder et al., 2003; Isaacowitz et al., 2007; Mill et al., 2009; Paulmann et al., 2008b). Our results add to those by Laukka and Juslin (2007), who reported that older adults (65+ years old) are less accurate than younger ones in categorizing sadness and fear in expressive cues of music performance (e.g., vibrato; phrasing). Thus, ageing affects how we recognize emotions both in music structure and in expressive performance. Finding that developmental shifts occur for music as for other modalities is also further evidence that ageing effects are supramodal to an important extent.

Contrasting with the age-related decrement in responsiveness to the negative musical emotions, the recognition of the two positive emotions, happiness and peacefulness, remained relatively invariant across adulthood. Stability for happiness was previously observed for musical expressiveness and speech prosody (Laukka & Juslin, 2007), and also for facial expressions (Calder et al., 2003; Williams et al., 2006). Notwithstanding, some studies report decline for this emotion in prosody and facial expressions (Isaacowitz et al., 2007; Paulmann et al., 2008b; Ruffman et al., 2009; Ruffman et al., 2008). Hence, the clear pattern of valence-specific effects that we uncovered here for music – decreased responsiveness for negative but not for positive emotions –, is not consensually established for other modalities. This might reflect modality-specific processes, but it might also result from the methodological characteristics of previous research. For example, most studies include an unbalanced number of negative and positive emotions, with happiness often being the only positive emotion, and this limits conclusions regarding valence differences. A strength of our design is the inclusion of a similar number of positive and negative emotional qualities. This would be feasible for other modalities, as people are able to recognize more than one positive emotion for instance in nonverbal vocal expressions (Sauter & Scott, 2007), but music might be a privileged method to explore in detail differentiated positive-valenced emotions. It is particularly effective at communicating them and positive responses predominate in music listening (e.g., Juslin et al., 2010; Juslin & Västfjäll, 2008; Sloboda & Juslin, 2010). In future studies it will be interesting to include more positive emotions, in addition to happiness and peacefulness, to further determine whether the age-related stability generalizes across positive musical emotions.

To understand the mechanisms that underlie age-related modulations in emotion recognition, it is crucial to determine whether they are primary in origin, or a consequence of general factors, notably cognitive decline and personality traits. In Study 2, we observed that advancing age is associated with decline in domain-general cognitive abilities, as has been reported in the literature (e.g., Salthouse, 2009), but no relation was found between these abilities and emotion recognition. We did not find associations with personality and socio-communicative traits as well. Furthermore, in Study 1 and 2, the emotions that underwent changes did not match those that were inherently more difficult to recognize, as would be expected if the origin of the changes were domain-general abilities. Altogether, these results strongly suggest that ageing can shape emotion recognition in music in a specific and direct manner. This is in line with findings for speech prosody (Mitchell, 2007; Orbelo et al., 2005) and facial expressions (Ruffman et al., 2008; Sullivan & Ruffman, 2004). For emotion recognition in these modalities, it has been proposed that top-down regulatory effects, and neuropsychological deterioration in structures important for emotion processing, might explain age-related changes. These mechanisms might also explain our results. Accumulating evidence suggests that age-related differences in emotion processing might be driven by shifts in regulatory and motivational mechanisms. According to the socioemotional selectivity theory, advancing age leads to enhanced ability and motivation to heighten attention towards positive information, and decrease attention towards negative information – positivity effect (Carstensen & Mikels, 2005; Charles & Carstensen, 2010; Mather & Carstensen, 2005; Samanez-Larkin & Carstensen, 2011). Consistently, we reported a pattern of decreased responsiveness to negative musical emotions, and stability for positive ones. This pattern fits coherently with studies showing that, with advancing age, brain responses decreases selectively for negative emotional stimuli (Gunning-Dixon et al., 2003; Kisley et al., 2007), and there is increased engagement of prefrontal systems which may implement regulatory effects over emotion processing (St. Jacques et al., 2009; St. Jacques et al., 2008; Williams et al., 2006). In a different vein, it has been hypothesized that age-related degradation in brain regions such as the amygdala and cingulate cortex contribute to explain the decline observed in the recognition of facial expressions of sadness and fear (Calder et al., 2003; Ruffman et al., 2008). Indeed, the volume of the amygdala was shown to decrease with advancing age (Curiati et al., 2009; Mu, Xie, Wen, Weng, & Shuyun, 1999; Walhovd et al., 2005). Neuropsychological research using the same stimuli and task as the present studies indicate that this structure is critical for the perception of fear and sadness in music (Gosselin et al., 2005, 2007). Thus, it is plausible that age-related changes for these emotions are linked with decline in the amygdala (see also, Cacioppo et al., 2011). On the other hand, the stability observed for the positive emotions might be related to the relative preservation of basal ganglia structures (e.g., Williams et al., 2006). Neuroimaging evidence showed that pleasant music recruits the striatum (e.g., Koelsch et al., 2006; Mitterschiffthaler et al., 2007). Note that some studies observed age-related shrinkage in the striatum, however (Raz et al., 2003). The present experiments are not ideally suited to specify the relative contribution of top-down regulatory mechanisms and neuropsychological deterioration for ageing effects in emotion recognition in music. Future studies combining behavioural measures with functional and structural brain imaging techniques will be critical to approach this question. These studies will also throw light on the differences that we observed in how younger and older participants relied on music structural cues to respond. For happy and peaceful responses, the utilization of the cues remained invariant with advancing age, but for sadness and fear older participants were less consistent than younger ones and weighted some cues differently. This might reflect deterioration of the neurocognitive mechanisms that support the processing of these emotions, but can also be accounted for by top-down regulatory and motivational processes: older participants might implement an active and controlled disengagement of music cues when they signal negative emotional meanings.

MUSICAL EXPERTISE AND EMOTION RECOGNITION IN MUSIC | Two findings indicate that musical training bolsters emotion recognition in music. First, in both studies, the length of musicians' training correlated significantly with enhanced recognition accuracy. This suggests that learning music operates gradual fine-tuning on mechanisms underlying musical emotions. Second, in Study 2, middle-aged musicians were more accurate than age-matched controls at categorizing the four emotions. As observed for age, we showed for the first time that the effects of expertise cannot be explained by differences in socio-educational background, general cognitive abilities and personality characteristics. The correlation between training and accuracy extends previous results by Livingstone and colleagues (2010), who reported a similar correlation for musical emotions portrayed by both structural and expressive cues, and by Bhatara and colleagues (2011), who focused on the expressive features timing and amplitude. Here we established that the effect also occurs for emotions embodied solely on music structure. Further support for the notion that musical training influences emotion processing in music comes from a study by Dellacherie, Roy, Hugueville, Peretz, and Samson (2011). These authors found that musical experience affects both behavioural and psychophysiological responses to dissonance in music. Musical training was associated with more unpleasant feelings and stronger physiological responses to dissonance, as assessed by skin conductance and electromyographic signals. Resnicow and colleagues (2004) failed to find effects of training, possibly because their design had no power enough to attain statistical significance – they included only three stimuli per emotion; furthermore, it was not reported how many trained participants were included.

It might be the case that the effect of expertise is small, a fact that would explain why it is not detected in some conditions (e.g., for younger musicians vs. controls in Study 2). Indeed, Bigand and colleagues (2005) observed that emotional responses to music in an implicit task are highly consistent and similar across trained and untrained participants. Training only increased the consistency of responses across testing sessions. Moreover, in a review of studies on the effects of musical expertise, Bigand and Poulin-Charronnat (2006) concluded that the effects of expertise on music processing are usually small, particularly in light of the huge differences in training that typically exist between groups. According to the authors, this is evidence that musical abilities, namely emotion processing, are acquired through exposure, not requiring explicit training. Our data is also evidence that training is not a prerequisite to be adept at perceiving musical emotions: all participants, including musically untrained ones, identified the four intended emotions with high agreement rates. This consistency was obtained with stimuli that were short, unfamiliar, all played in the same genre and timbre, and expressing emotions exclusively through music structure cues. Additionally, participants were able to effectively use these cues to make emotion differentiations, as indicated by the multiple regression analyses. By claiming that learning music impacts emotion recognition, we intend to highlight the plasticity of the system, and not to argue that it needs training to operate robustly.

How can musical training enhance the categorization of emotional qualities in music? Musicians, because of their extensive implicit and explicit practice with musical emotions, might develop more fine-grained, sharply defined and readily accessible categories to respond to emotion cues in music. Finding that they were more consistent than controls in using music structure cues to respond is in agreement with this hypothesis. Barrett and colleagues (Barrett, 2006b, 2009; Lindquist & Barrett, 2010) have suggested that individual differences in emotional responses and experience might reflect differences in the *granularity* of emotion concepts, which could be trained. According to this theoretical proposal, individuals higher in emotional granularity are more precise, specific and differentiated at categorizing affective states, while those with low granularity experience and describe emotions in a more global and undifferentiated fashion. Higher emotional granularity could be achieved through training and practice, just like wine and X-ray experts learn to perceive subtle differences that novices are unaware of. It is plausible that musical training contributes to increase the granularity of emotion concepts for music.

6.5. CONCLUSION

In two studies, we sought to determine the effects of ageing and musical expertise on emotion recognition in music. To our knowledge, this is the first examination of these effects that covers the full range of the adult life span, and that focus on emotions expressed through musical structure. We found that ageing modulates the recognition of musical emotions, with developmental shifts being observed early on in middle-age. Specifically, advancing age was associated with a gradual decrease in responsiveness to the negative emotions of sadness and fear in music, while the positive ones, happiness and peacefulness, remained invariant. This is evidence that emotion processing in music can follow the same developmental trends as stimuli important for biological survival and social functioning, such as facial expressions and speech prosody. We also showed that musical expertise is associated with more accurate categorization of musical emotions, thus contributing to the literature on the relation between musical experience and music perception. Both ageing and expertise effects were independent of domain-general cognitive or personality differences, and this suggests that they have a primary origin. Mechanisms supporting emotion recognition in music are robust, but also dynamic and variable: they respond plastically to the influence of experience.

CHAPTER VII

Development and Validation of a Set of Portuguese Stimuli for Research on Emotional Prosody

7.1. INTRODUCTION

Well-devised and validated stimuli are important for research on emotional speech prosody (e.g., Burkhardt et al., 2005; Pell, 2002; Ross et al., 1997). Because prosodic information is overlaid on the speech signal, sets of stimuli need to be developed for different languages. The availability of materials in different languages is also crucial to shed light on the role of language-specific and universal factors in the processing of prosody (e.g., Pell, Monetta, et al., 2009; Thompson & Balkwill, 2006). In this study we develop and validate a database of prosodic stimuli in Portuguese[6].

In the literature on emotional speech prosody, typically researchers ask trained actors or untrained speakers to read some verbal materials aloud while portraying different emotional states, such as anger, disgust, fear, happiness, sadness and surprise (e.g., Adolphs et al., 2002; Banse & Scherer, 1996; Dara et al., 2008; Juslin & Laukka, 2001; Scherer et al., 1991). Distinct types of verbal materials have been used, including complete sentences (e.g., Adolphs & Tranel, 1999; de Gelder & Vroomen, 2000; Kotz et al., 2003; Mitchell, 2007), single words (e.g., Wiethoff et al., 2008) or monosyllabic utterances (e.g., Mitchell & Ross, 2008). It has often been argued that these posed portrayals may result in intense, prototypical expressions, which reflect conventionalized stereotypical norms, more than the natural psychophysiological changes on voice that occur in authentic emotion episodes. As stated by Scherer (2003), though, the so-called natural public expressions are also acted to some degree, since we self-regulate vocal expressions during everyday life, and sociocultural norms impose production constrains to a significant extent (pull effects). It is also well documented

[6] Published in Castro, S. L., & Lima, C. F. (2010). Recognizing emotions in spoken language: A validated set of Portuguese sentences and pseudo-sentences for research on emotional prosody. *Behavior Research Methods, 42*, 74-81. doi:10.3758/BRM.42.1.74

that listeners can identify reliably emotions in acted stimuli, indicating that they are at least partially modelled on the basis of natural expressions. Furthermore, although scarce, available evidence suggests that the acoustic patterns associated with specific emotions are qualitatively similar in portrayals and in natural expressions (Juslin & Laukka, 2001; Laukka et al., 2008). Another approach to devise prosodic stimuli involves inducing emotional states in the speakers (e.g., through films or pictures), and record expressions under these states. While this approach would be better suited to obtain spontaneous samples, it is difficult to evoke strong and differentiated emotional reactions in laboratory settings, so that the resulting voice stimuli are usually low intense and with relatively undifferentiated affect (Juslin & Laukka, 2001). Here we use acted portrayals and endorse the assumption that they are suitable for research on emotion communication through prosody (Bänziger & Scherer, 2007).

A fundamental issue that has to be dealt with when creating prosodic materials is the interplay of prosody – the variable of interest – with the lexico-semantic content of the utterance. One strategy consists of removing semantic information through acoustic filtering procedures that suppress the segmental content. However, emotions are more difficult to recognize in these acoustically manipulated stimuli than in unaltered ones (Kotz et al., 2003). Alternatively, the influence of semantics can be controlled for by using standardized verbal materials, a procedure often called *standard content paradigm* (e.g., Juslin & Laukka, 2001, 2003). That is, the same verbal material is used to portray all emotions, and because content is similar across emotions, it is assumed that the effects observed in the listener's judgments reflect solely the impact of prosodic cues (Scherer, 2003). This approach is used in most studies. The characteristics of the standardized content itself deserve also to be considered because they may modulate the neurocognitive processes that are set into action. For instance, neuropsychological and fMRI studies using the Aprosodia Battery show that stimuli with different degrees of verbal complexity recruit different neurocognitive processes (Mitchell & Ross, 2008; Ross & Monnot, 2008). The Aprosodia Battery (Ross et al., 1997) employs six intonations (angry, disinterested, happy, sad, surprised and neutral) in three types of materials: short sentences with emotionally neutral semantic content ("I am going to the other movies"), a type of stimuli used frequently in the literature (e.g., Adolphs, Tranel,

& Damasio, 2001; Mitchell, 2007; Wildgruber et al., 2005); monosyllabic utterances (repeated *ba*); and asyllabic stimuli (prolonged *aaaaahhhhh*). It was observed that healthy listeners recognize emotions similarly well across conditions, about 70% correct on average, confirming that the processing of prosody can proceed independently of semantics. However, the relative contribution of the left and right hemispheres depended on the type of stimulus: lateral temporal lobes in both hemispheres were engaged by emotional prosody, but the contribution of the left hemisphere got smaller as linguistic complexity decreased (Mitchell & Ross, 2008; Ross & Monnot, 2008). The authors concluded that emotional prosody is right-lateralized to a significant extent, and that the contribution of the left hemisphere is related to the processing of verbal information. Accordingly, the dynamic dual-pathway model of language functions postulates a dissociation between semantic/syntactic processes versus prosodic ones: a left-lateralized temporo-frontal network supports syntactic and semantic processes, being the right hemisphere dominant for prosodic processes, depending on the linguistic demands of the stimuli or task – the higher these are, the greater is the co-involvement of the left hemisphere (Friederici & Alter, 2004). Hence, stimuli with different degrees of linguistic information can be useful to explore the processes that underlie emotional prosody. One can for example resort to meaningless speech and pseudo-sentences to create stimuli with reduced semantic content (e.g., Banse & Scherer, 1996; Pell, 2002; Péron et al., 2010). Pseudo-sentences are utterances that include pseudo-words. Their lexico-semantic content is substantially reduced, because pseudo-words have no meaning, yet they have the advantage of affording a good language-like quality because both phonetic-segmental and suprasegmental features of normal speech are present. Emotions can be recognized in pseudo-sentences with high accuracy (e.g., 78% correct in Pell, 2002), and this type of stimuli has been successfully used in neuropsychological (e.g., Dara et al., 2008; Péron et al., 2010), electrophysiological (e.g., Paulmann & Kotz, 2008b) and neuroimaging studies (Bach et al., 2008).

Emotional stimuli like facial expressions (e.g., Ekman & Friesen, 1978), pictures (e.g., Lang, Bradley, & Cuthbert, 2008), music (Vieillard et al., 2008) and nonverbal vocal expressions (Belin et al., 2008; Sauter, Eisner, Calder et al., 2010; Sauter, Eisner, Ekman et al., 2010; Sauter & Scott, 2007) can be used independently of the linguistic background of the participants. By contrast, emotional prosody has at least some degree of language specificity. This is why sets of prosodic stimuli have been devised for different languages, such as English (Pell, 2002; Ross et al., 1997), Polish (Staroniewicz & Majewski, 2009), Mandarin (Wu et al., 2006), Russian (Makarova & Petrushin, 2002) and German (Burkhardt et al., 2005). We devised a set of stimuli in European Portuguese, a research tool useful for research on emotional prosody with Portuguese-speaking listeners and also for cross-language studies. We resorted to the standard content paradigm, applied to two types of verbal materials that vary in the amount of lexico-semantic information afforded: sentences with emotionally neutral semantic content, and pseudo-sentences, composed of pseudo-words and a few function words. Both types of stimuli are frequently used in the literature and seem to be well suited to examine emotional prosody. We used whole sentences, instead of single words, because longer stretches of speech allow for a more fine-grained emotional differentiation than shorter ones. Two speakers recorded the stimuli portraying a neutral tone of voice and six emotions: anger, disgust, fear, happiness, sadness and surprise. The stimuli were analysed for primary acoustic measures, including duration, mean F_0 and F_0 variability. This analysis was conducted to confirm that the emotion-specific profiles are in general agreement with previous descriptions (e.g., Banse & Scherer, 1996; Juslin & Laukka, 2003). For perceptual validation, we presented the stimuli to 80 undergraduate students and collected three measures: emotion categorization accuracy, based on a conventional forced-choice task; reaction times (RTs); and perceived emotion intensity. As a secondary goal, we compared sentences and pseudo-sentences concerning these perceptual measures to examine whether differences in lexico-semantic information modulate emotion recognition. Based on the reviewed studies, we predicted that recognition accuracy and intensity would not differ across stimulus type because prosody appears to be processed equally well in stimuli with different degrees of semantic content (Mitchell & Ross, 2008; Ross & Monnot, 2008). However, we expected RTs to be shorter for pseudo-sentences than for sentences. Latencies are not usually measured for emotion recognition in prosody, but it seems reasonable to hypothesize that the reduced amount of lexico-semantic content in pseudo-sentences might speed up the recognition process: pseudo-sentences would pose less demands than sentences on the resources dedicated to the integration of semantics with prosody

(Schirmer & Kotz, 2006), and so the decision regarding the emotional expression would be reached sooner.

7.2. METHOD

7.2.1. Recording

SPEAKERS | After conducting pilot tests with several candidates, we selected 2 women (mean age = 18 years) to record the stimuli. They were native speakers of European Portuguese. Both had formal musical training, including singing lessons, and were chosen on the basis of their ability to modulate emotional prosody.

MATERIALS: SENTENCES AND PSEUDO-SENTENCES | A set of 16 short Portuguese sentences and 16 pseudo-sentences was devised. The complete list is presented in Table 10. The sentences were syntactically simple and composed of high-frequency words. Lexico-semantic content was emotionally neutral, as assessed by three independent judges. Mean length was 8 syllables (*SD* = 1.3; range = 6 - 11). The pseudo-sentences were derived from the sentences and respected the syntax, morphology and phonotactics of Portuguese. They were derived as follows: one to three phonemes of the content words of the sentences (nouns, adjectives and verbs carrying salient semantic information) were replaced, such that pseudo-words were formed; vowels were replaced by vowels and consonants by consonants, and the syllabic structure and stress of the original word were kept. For example, the sentence *O* ***q****uadr****o*** *está na pa****r****e****d****e* (the painting is on the wall) was transformed into *O* ***j****uadr****e*** *está na pa****fên****e*. Because the pseudo-sentences retained the phonological properties of Portuguese, they were highly language-like. That is, they were "nonsensical but an appropriate carrier of suprasegmental information" (Pell, 2002, p. 500).

Table 10. Sentences and pseudo-sentences included in the database.

Sentence (Translation)	Pseudo-sentence
Esta mesa é de madeira. (*This table is of wood.*)	Esta dêpa é de faneira.
O rádio está ligado. (*The radio is switched on.*)	O dárrio está guilado.
Aquele livro é de história. (*That book is of history.*)	Aquele jicro é de hisbólia.
A Terra é um planeta. (*The Earth is a planet.*)	A Pirra é um flaneto.
O cão trouxe a bola. (*The dog brought the ball.*)	O lão droube a nóma.
Ele chega amanhã. (*He arrives tomorrow.*)	Ele chena aguinhã.
Esta roupa é colorida. (*These clothes are colorful.*)	Esta souda é lacoripa.
Os jardins têm flores. (*The gardens have flowers.*)	Os bartins têm pléres.
As pessoas vão a concertos. (*People go to concerts.*)	As semoas vão a cambêrtos.
Há árvores na floresta. (*There are trees in the forest.*)	Há árjuques na plurisca.
Os tigres são selvagens. (*Tigers are wild animals.*)	Os lagres são siltávens.
O quadro está na parede. (*The painting is on the wall.*)	O juadre está na pafêne.
Alguém fechou as janelas. (*Someone closed the windows.*)	Alguém belhou as jalétas.
Os jovens ouvem música. (*Young people listen to music.*)	Os dófens mavem tézica.
O futebol é um desporto. (*Football is a sport.*)	O dutebel é um nesforpo.
Ela viajou de comboio. (*She traveled by train.*)	Ela jiavou de lantóio.

RECORDING PROCEDURE | We conducted two separate recording sessions, one for each speaker, lasting about four hours each. A list of the sentences and pseudo-sentences was given to them three days before the recording session so that they could familiarize themselves with the materials. They were told that each sentence and pseudo-sentence was to be produced varying intonation in order to portray neutrality and the six emotions (anger, disgust, fear, happiness, sadness and surprise). In the recording session proper, they were asked to say the materials with the clearest, but still natural and spontaneous target expression. The general instruction was: "Say this as if you were feeling [the target emotion]." To enhance expressiveness, we coached the speakers providing hypothetical scenarios (e.g., "imagine that you are very happy because you have just been given fantastically good news. Think of a situation where that really happened and put yourself in the mood, smile! [pause] Done? Then let's speak out these sentences in a happy manner!"). We did not give specific instructions on how to achieve the emotional tones (e.g., high- or low-pitched voice, slow or quick paced). When necessary, a stimulus was produced several times until being evaluated by the experimenters (César Lima and São Luís Castro) as a clearly recognizable instance of the intended expression. Scherer and colleagues (e.g., Banse & Scherer, 1996; Scherer, 2003; Scherer et al., 2003) acknowledged that some emotion categories might be produced in distinct manners, or subtypes, and that these might correspond to distinct acoustic profiles (e.g., anger can be produced in a hot explosive manner, or in a cold controlled way). Here the speakers tended to produce some subtypes of emotions more than others: anger was produced mainly in hot rather than cold form, sadness as quiet sadness rather than despair, and fear as a sustained or milder state rather than panic.

The stimuli were recorded in the sound-insulated booth of the Speech Laboratory at University of Porto, Faculty of Psychology and Education, using Pro Tools LE version 5.1.1 (Digidesign, Avid Technology) software and a high-quality microphone attached to an Apple Macintosh G4 computer. Digitization was done at a 48-kHz sampling rate and 16-bit resolution. Individual files were prepared with the more recognizable productions of each sentence and pseudo-sentence from each speaker according to the judgment of the experimenters and informed, in case of doubt or disagreement, by judgments from a small number of lab colleagues. The sound files were normalized for maximum peak intensity using Sound Studio version 3.5.5 (Kwok, 2008); 50 ms of silence were added at the beginning and at the end of the utterance. A total of 448 stimuli were submitted to the validation procedure (32 sentences/pseudo-sentences x 7 intonations x 2 speakers).

7.2.2. Validation

PARTICIPANTS | Eighty undergraduate students (72 females; mean age = 21.8 years; *SD* = 6.1) participated in the perceptual validation experiment for course credit. They were divided into four groups of 20 participants; each group judged either the sentences or the pseudo-sentences produced by one of the speakers (between-subjects design). All participants were native speakers of European Portuguese and reported no hearing impairments or speech disorders, no major psychiatric or neurological illnesses, and no head trauma or substance abuse.

PROCEDURE | Participants were tested individually in one experimental session lasting 25 minutes. Each of the four between-subjects conditions included 112 stimuli: all the sentences, or pseudo-sentences, produced by one speaker (16 x 7). In each condition, the stimuli were pseudo-randomized and divided into two blocks of 56 trials; the presentation order of the blocks was counterbalanced across participants. Participants performed two consecutive judgments for each stimulus. First, they categorized as fast and accurately as possible the emotion expressed by the tone of voice (forced-choice), and then they made an intensity judgment on a 7-point scale, evaluating how well the expression was displayed – that is, how representative the stimulus was of the chosen category. The stimuli were presented via high-quality headphones and the responses were given by pressing pre-specified buttons on a seven-button response pad, model RB-730, from Cedrus Corporation. Seven labels for the expression categories were affixed below the buttons of the response pad in the following order: neutral, *neutro;* happy, *alegre;* sad, *triste;* angry, *zangado;* fearful, *assustado;* disgusted, *repulsivo;* and surprised, *surpreendido*; the numbers 1 (low intensity) to 7 (high intensity) were affixed above the buttons. Participants were instructed to respond with the index finger of their dominant hand and to keep the hand in the same position, at the centre of the response pad, throughout the session.

The presentation of the stimuli and the recording of RTs and responses were controlled with the SuperLab version 4.0 (Abboud et al., 2006) running on a Macintosh PowerBook G4 computer. Each trial was as follows: a fixation point (+) appeared on the screen for 500 ms; the stimulus was then presented while the fixation point remained visible; after the participant categorized the emotion, the intensity scale appeared on screen, prompting the participant to make the intensity judgment; after the response, a blank screen was presented for 3,000 ms (inter-stimulus interval). The session started with four practice trials to familiarize the participants with the stimuli and task, and to adjust headphone volume. RTs were measured from stimulus onset until the first button press (corresponding to the emotion categorization).

SELECTION | For each stimulus, the mean percentage of correct identifications of the target emotion was computed. The database included only the stimuli that reached an accuracy rate of at least 43%, about three times the level of chance (14%), and of less than 43% in any of the other categories. Our criterion was more liberal than the one used by other researchers (e.g., 60% correct in Pell, 2002), in order to have a varied set of stimuli in the database, namely some that fell within a range of relative ambiguity; these might be useful for studies that manipulate emotion ambiguity (e.g., Bach et al., 2009). From the initial set of stimuli, 368 were selected, and 80 were excluded. The mean number of stimuli per emotion was 26 (SD = 6.4). The stimuli included in the database were submitted to more detailed analyses, which are presented below.

7.3. RESULTS AND DISCUSSION

This database for research on emotional prosody consists of 190 sentences and 178 pseudo-sentences spoken in European Portuguese, which portray a neutral tone and six emotions: anger, disgust, fear, happiness, sadness and surprise. Acoustic and perceptual characteristics of each stimulus are detailed in the Appendix 6. The database itself can be obtained from the authors upon request.

ACOUSTIC ANALYSIS | The stimuli were measured for mean F_0, F_0 variability (SD) and duration (ms), using Praat (Boersma & Weenink, 2009). These acoustic attributes are presented for each emotion category in Table 11. To examine possible differences across emotions, and between sentences and pseudo-sentences, raw acoustic values were z-transformed for each speaker; this minimizes the variance related with speakers' voice idiosyncrasies, but preserves within-speaker variability across categories and stimulus type. Sentences and pseudo-sentences were acoustically similar (Fs < 1). Emotions differed for the three measures, as revealed by one-way ANOVAs [main effect of category for mean F_0, $F(6, 354) = 293.3$, $p < .0001$, $\eta_p^2 = .83$; F_0 variability, $F(6, 354) = 64.7$, $p < .0001$, $\eta_p^2 = .52$; and duration, $F(6, 354) = 18.36$, $p < .0001$, $\eta_p^2 = .24$; interactions Category x Stimulus Type ns, $p > .14$]. Mean F_0 was higher for happiness, followed by surprise, and lower for disgust, neutrality and sadness; anger and fear were in the middle range. F_0 variability was largest for happiness and surprise, intermediate for anger and disgust, and smallest for fear, neutrality and sadness. Duration was longer for disgust and sadness, and shorter for anger; the remaining emotions, neutrality, happiness, fear and surprise, had intermediate duration. These differences are consistent with previous descriptions for stimuli in other languages (Banse & Scherer, 1996; Juslin & Laukka, 2001, 2003; Pell & Skorup, 2008). To examine whether acoustic measures provide sufficient information to discriminate between expression categories in our set of stimuli, we computed a discriminant analysis taking mean F_0, F_0 variability and duration as independent variables; the dependent variable was the expression category. The model was able to predict significantly the category membership of the stimuli [Wilks's $\lambda = .1$; $F(18, 1015) = 72.72$, $p < .0001$], with an overall accuracy of 60% (chance level, 14%). Concerning specific emotions, accuracy was 68% for anger, 86% for disgust, 73% for fear, 57% for happiness, 62% for neutrality, 56% for sadness and 31% for surprise.

Table 11. Mean F_0, F_0 variability (*SD*) and duration of the sentences and pseudo-sentences included in the database.

Emotion category	F0 (Hz)		F0 variability (*SD*)		Duration (ms)	
	Sentences	PSentences	Sentences	PSentences	Sentences	PSentences
Neutrality	211	214	38	42	1,477	1,484
Happiness	357	350	81	83	1,479	1,513
Sadness	206	205	38	39	1,532	1,530
Anger	313	307	67	62	1,266	1,277
Fear	302	317	37	43	1,445	1,457
Disgust	279	276	67	66	1,634	1,580
Surprise	332	323	71	80	1,435	1,446

Note. PSentences = Pseudo-sentences; Values are averaged across speakers.

RECOGNITION ACCURACY | Inter-participant reliability in responses was very high for both speakers and stimulus types: Cronbach's α for speaker A was .98 in sentences and .97 in pseudo-sentences; for speaker B, it was .95 in sentences and .96 in pseudo-sentences. The percentage of correct categorizations (true positives or *sensitivity*) was computed for each expression category, separately for sentences and pseudo-sentences. These rates are presented in the diagonal cells of Table 12, collapsed over the two speakers; recognition accuracy was similar across speakers ($p > .2$). Rates were high, ranging between 50% (disgust in sentences) and 88% (neutrality in sentences). We also computed the percentage of correct rejections for each response category (true negatives or *specificity*), i.e., the number of times that participants correctly rejected a given category (e.g., happy) when stimuli corresponding to the other, non-intended, categories (e.g., anger) were presented (see Table 12, cells in italics). These values were also high, ranging between 91% (surprise in sentences and pseudo-sentences) and 99% (fear in sentences). High sensitivity indicates that the stimuli expressing different emotions tended to be correctly categorized when they were presented. High specificity, on the other hand, indicates that listeners tended to recognize a specific emotion only when the stimulus was indeed an instance of that emotion.

Table 12. Distribution of responses (%, rows) for each intended expression in sentences and pseudo-sentences. Recognition accuracy (sensitivity) is indicated in bold, and correct rejection rate (specificity) in italics. Standard errors are presented in parentheses.

Intended expression		Response						
	n	Neutral	Happy	Sad	Angry	Fearful	Disgusted	Surprised
Sentences	190							
Neutrality	30	**88 (2.1)**	0	9	2	0	1	1
Happiness	29	1	**75 (2.8)**	1	2	1	2	19
Sadness	31	11	1	**84 (2.1)**	3	1	0	0
Anger	32	8	3	0	**77 (2.2)**	2	6	3
Fear	24	3	2	5	1	**65 (2.4)**	2	23
Disgust	12	8	22	1	4	2	**50 (1.9)**	15
Surprise	32	5	3	1	1	3	1	**87 (1.9)**
Correct Rejections		*94*	*97*	*97*	*98*	*99*	*98*	*91*
Pseudo-sentences	178							
Neutrality	30	**83 (2)**	1	13	3	0	0	1
Happiness	17	2	**59 (3.2)**	2	3	4	3	27
Sadness	28	13	2	**82 (2.2)**	2	1	1	0
Anger	31	4	2	1	**74 (2.2)**	4	10	4
Fear	17	4	4	7	1	**56 (2.3)**	5	24
Disgust	23	2	17	0	5	2	**60 (2.2)**	13
Surprise	32	4	4	0	1	3	2	**85 (2)**
Correct Rejections		*95*	*95*	*96*	*98*	*98*	*96*	*91*

Note. Values averaged across speakers

Stimuli expressing sadness, surprise and neutrality yielded the highest categorization accuracy (above 80%, *p*s < .0001), followed by anger (76%) and happiness (67%); stimuli expressing fear (60%) and disgust (55%) yielded the lowest accuracy (*p*s < .05). This was supported by an ANOVA on arcsine transformed accuracy rates, with category and stimulus type as between-subjects factors [main effect of category, $F(6, 354) = 42.42$, $p < .0001$, $\eta_p^2 = .42$]; Tukey HSD tests were used for *post hoc* comparisons. Disgust was particularly difficult to recognize. This had already been indicated by the fact that fewer of these stimuli reached the criteria for inclusion in the database (12/32 sentences and 23/32 pseudo-sentences from the original set). The difficulty in recognizing disgust in speech prosody was previously observed by Scherer and colleagues (Banse & Scherer, 1996; Scherer et al., 1991), who suggested that this emotion might be better expressed in nonverbal vocal expressions or vocal emblems, such as "yuck!", than in the tone of voice imparting over longer stretches of speech, as in the whole sentences used here. In fact, Belin, Fillion-Bilodeau, and Gosselin (2008) and Sauter and colleagues (2010) found that disgust is recognized with high accuracy in vocal expressions (81% and 94%, respectively). Another reason why this emotion may not be easily recognized in prosody is that in natural communicative settings it is likely to be expressed in facial expressions rather than in speech.

Concerning the pattern of errors, disgust was confused mainly with happiness and surprise, as can be seen in Table 12. Such a confusion appears unexpected because these emotions contrast in valence. However, as disgust was inherently difficult to recognize, this may have prompted listeners to guide their judgments on these stimuli by their prominent acoustic features, like large F_0 variability, which is also a characteristic of happiness and surprise (see Table 11); anger also has large F_0 variability, but it is highly distinctive because of its short duration. This may explain why disgust stimuli were rarely misclassified with anger. Happiness was confused primarily with surprise, possibly because they both have high mean F_0 and F_0 variability, and are not contrasting in terms of valence. Fear was also confused with surprise, possibly because both emotions are identical for duration and share a dimension of unexpectedness. The miscategorization of happiness and fear with surprise has already been reported (e.g., Adolphs et al., 2002). Sadness and neutrality were confused with each other, as observed in previous studies (e.g., Juslin & Laukka, 2001; Pell, 2002). They are similar for intensity (both were relatively low intense, see below), and also for mean F_0 and F_0 variability, which is low. The percentages of misclassifications for stimuli expressing anger and surprise were uniformly distributed across categories (see Table 12).

As predicted, all emotions were recognized equally well in sentences and in pseudo-sentences, with the exception of happiness, which was better recognized in sentences [main effect of stimulus type, $F(1, 354) = 11.24$, $p < .001$, $\eta_p^2 = .03$; interaction Category x Stimulus Type, $F(6, 354) = 2.83$, $p < .05$, $\eta_p^2 = .05$; *post hoc* comparisons yielded ns results for all emotions, *p*s > .36, except for happiness, $p < .01$]. The effect found for happiness is unanticipated, given that sentences and pseudo-sentences were identical for primary acoustic features. It might be related to subtle uncontrolled differences in fine acoustic details, such as in the distribution of energy in the spectrum.

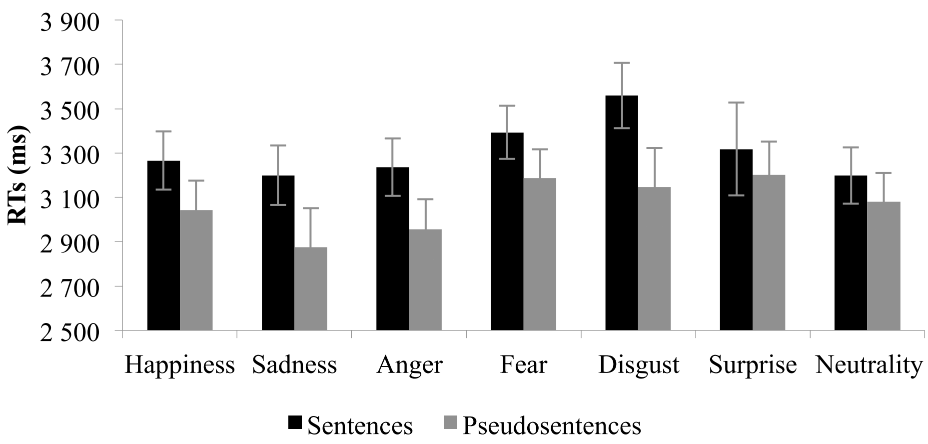
Summing up, categorization accuracy for our stimuli was high, both for sentences, 75% correct, and for pseudo-sentences, 71% correct. This indicates that both types of vocal materials produced highly recognizable emotions, and so they are suitable for research on emotional prosody. Accuracy rates in this study are close or higher than the ones reported in previous studies. For example, Adolphs and colleagues (2002) obtained 81% correct for 5 emotions in stimuli consisting of semantically neutral English sentences, Banse and Scherer (1996) 48% for 14 emotions in pseudo-sentences composed of phonemes from several Indo-European languages, Juslin and Laukka (2001) 56% for semantically neutral English and Swedish sentences, and Pell (2002) around 78% for 6 emotions in pseudo-sentences that resemble English.

LATENCIES | The time taken to correctly categorize the target expressions was computed for each stimulus; RTs deviating more than 3 *SD*s from the grand mean were excluded from the analysis. Figure 3 presents RTs for each expression category, for sentences and for pseudo-sentences. In line with with our hypothesis, emotions were recognized faster in pseudo-sentences (3,070 ms) than in sentences (3,310 ms), as indicated by an ANOVA with expression category and stimulus type as between-subjects factors [main effect of stimulus type, $F(1, 354) = 9.3$, $p < .005$, $\eta_p^2 = .03$;

interaction Category x Stimulus Type ns, $F < 1$]. The extra time needed to respond to sentences (+240 ms) might index the integration of semantics with prosody. The lexico-semantic information was emotionally neutral and irrelevant to the task, but it may have been automatically processed by left temporo-frontal networks dedicated to linguistic processing (Friederici & Alter, 2004; Mitchell & Ross, 2008; Ross & Monnot, 2008), and then integrated with the suprasegmental prosodic cues, as occurs in everyday communicative contexts (Schirmer & Kotz, 2006). Because for pseudo-sentences semantic information available was residual, resources dedicated to the integration were substantively decreased and responses could be produced faster.

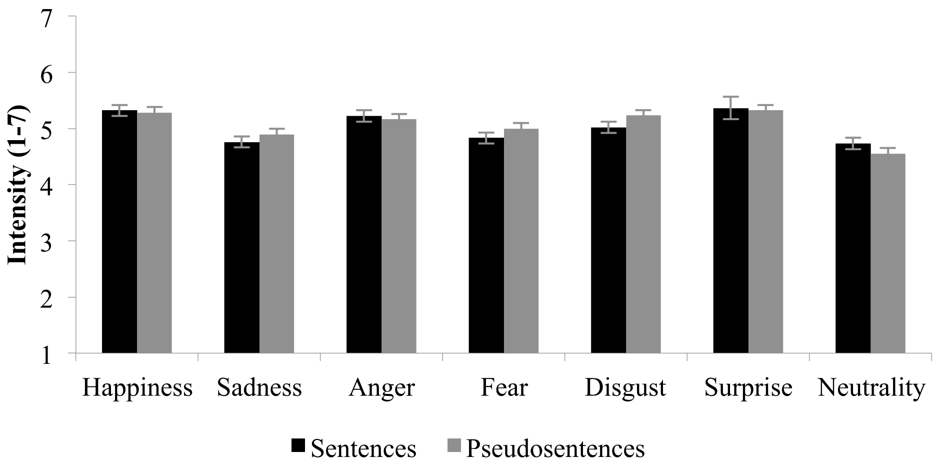
RTs did not differ significantly across emotions (main effect of category ns, $p > .3$).

Figure 3. Mean RTs to correctly categorize each expression in sentences and pseudo-sentences. Error bars present standard errors.



INTENSITY JUDGMENTS | Figure 4 depicts intensity rates for each expression category, as a function of stimulus type. These rates are based on data from correct categorizations only. Intensity was generally high, 5 on average (7-point scale), and sentences and pseudo-sentences elicited similar judgments, i.e., they were similarly considered as good exemplars of the different emotion categories (main effect of stimulus type ns, $F < 1$). This confirms that both types of stimuli are effective at communicating emotions through prosody. Stimuli expressing happiness (5.3), anger (5.2), disgust (5.1) and surprise (5.3) were considered more intense than those expressing neutrality (4.6), sadness (4.8) and fear [4.9, *p*s < .05; main effect of category, $F(6, 354) = 11.68$, $p < .0001$, $\eta_p^2 = .17$; interaction Category x Stimulus Type ns, $F < 1$].

Figure 4. Mean intensity rates by emotional expression for sentences and pseudo-sentences. Error bars present standard errors.



7.4. CONCLUSION

We herein devised and validated a stimulus set in European Portuguese for research on emotional prosody. We used standardized verbal materials, which consisted of 16 sentences with emotionally neutral semantic content, and 16 pseudo-sentences composed of pseudo-words with a few function words. Two speakers recorded these materials to portray seven intonations: anger, disgust, fear, happiness, sadness, surprise and neutrality. The most ambiguous stimuli were discarded to ensure the quality of the portrayals that were included in the database. The final set contains 190 sentences and 178 pseudo-sentences, which underwent acoustic analysis and perceptual validation.

The acoustic profiles found for the different emotions are in good agreement with previous descriptions for other languages (Banse & Scherer, 1996; Juslin & Laukka, 2001, 2003), and are similar for sentences and pseudo-sentences. Recognition accuracy was high, 75% correct for sentences and 71% for pseudo-sentences (chance level, 14%). Stimuli were judged as highly intense, indicating that they communicate the intended emotions clearly.

Emotions were recognized equally well in sentences and in pseudo-sentences, and intensity judgments were also similar across stimulus type. This is further evidence that the neurocognitive mechanisms that support prosody processing are effective in dealing with utterances with lexico-semantic content and without it. That is, the processing of prosody can occur independently of semantics (Mitchell & Ross, 2008; Ross & Monnot, 2008). The lexico-semantic information contained in sentences, even being neutral, produced a processing cost detected in RTs: emotions took longer to be recognized in sentences than in pseudo-sentences, probably because sentences required an integration of semantics with prosody (Schirmer & Kotz, 2006). Altogether, these findings are consistent with evidence that the processing of verbal and prosodic information rely on dissociated but interacting neurocognitive mechanisms (Friederici & Alter, 2004; Ross & Monnot, 2008).

Future developments of this database will benefit from including more speakers, namely male and older ones. The speaker's identity does not seem to play an important role on prosody processing (Juslin & Laukka, 2001; Paulmann & Kotz, 2008b; Pell, 2002), but increasing the diversity of voices will enhance the ecological validity of the stimulus set. It will also be important to validate the stimuli for listeners with different demographic characteristics. For example, it is known that ageing impacts emotion recognition in prosody (e.g., Mill et al., 2009; Mitchell, 2007; Paulmann et al., 2008b), and here perceptual validation relied exclusively on young listeners (undergraduate students).

The set of stimuli presented in this study is suitable for behavioural, neuropsychological, neuroimaging and electrophysiological research on prosody, language and communication. It will be useful, for instance, for cross-language experiments aiming to examine universal versus language-specific factors in emotion recognition in prosody (e.g., Pell, Monetta, et al., 2009; Pell, Paulmann, et al., 2009), and for inclusion in neuropsychological tools for the clinical assessment of pragmatic skills in brain-damaged and neuropsychiatric patients. We used a subset of sentences to investigate whether the processing of emotional prosody is modulated by musical expertise (Chapter VIII) and by basal ganglia dysfunction in PD (Chapter IX).

CHAPTER VIII

Speaking to the Trained Ear: Musical Expertise is Associated with Enhanced Recognition of Emotions in Speech Prosody

8.1. INTRODUCTION

Language and music share important features. Both are finely structured systems of expression and communication based on perceptually discrete units organized into flowing acoustic streams. Both are universal in the human species and have ancient origins (e.g., Mithen, 2007). The comparative analysis of language and music is an active area of research in cognitive neuroscience: which components of these faculties are subserved by domain-general shared mechanisms, and which ones are domain-specific? Answers to this question will illuminate debates on modularity and on the evolutionary origins of music (Hauser & McDermott, 2003; McDermott, 2008). They will also have far-reaching implications regarding how music can be used to treat language processing problems (Patel, 2011). Musical abilities such as tonal encoding of pitch and beat perception and synchronization might depend on domain-specific mechanisms (Patel, 2008a; Peretz, 2006, 2008; Peretz & Coltheart, 2003; Zentner & Eerola, 2010), though this remains controversial (e.g., McDermott, 2009; Patel, 2010; Trehub & Hannon, 2006). Neuroimaging and behavioural studies indicate that several other components of linguistic and musical processing are domain-general (for a review, Patel, 2008a). This is the case of syntactic operations (Koelsch, Gunter, Wittfoth, & Sammler, 2005; Patel, 2003), perception of phrase boundaries (Knösche et al., 2005), pitch processing (Liu et al., 2010; Marques et al., 2007; Schön et al., 2004) and conceptual processing (Koelsch et al., 2004; Schön, Ystad, Kronland-Martinet, & Besson, 2010). It has been suggested that music is able to express emotions because it engages neural resources dedicated to emotional speech (Juslin et al., 2010; Juslin & Västfjäll, 2008; Nieminen et al., 2011; Patel, 2008b; Peretz, 2010). Does emotion recognition in music and in speech prosody rely on shared mechanisms? To examine this hypothesis, we determined whether musical expertise, which was shown to impact on emotion recognition in music, also impacts on emotion recognition in speech[7].

Musicians spend a massive amount of time in music learning activities and deliberate practice, usually 10 or more years until they achieve expert performance (e.g., Ericsson & Lehmann, 1996; Ericsson & Towne, 2010). A pivotal part of this experience concerns the production and perception of fine-grained modulations of complex acoustic patterns for expressive and emotional purposes. In fact, learning music sharpens up the accuracy with which emotions are perceived in music, both when they are communicated by structural compositional cues and by expressiveness in performance (Bhatara et al., 2011; Lima & Castro, 2011; Livingstone et al., 2010). If emotion recognition in music and in speech recruits shared mechanisms, then musical expertise should be associated with enhanced processing of emotional prosody. To the best of our knowledge, only two studies tested directly this hypothesis, and they yielded conflicting results. Thompson, Schellenberg, and Husain (2004) found suggestive evidence of a positive effect in a series of three experiments. In the first one, nine musically trained undergraduates performed better than 12 untrained ones in a forced-choice identification of four emotions (happiness, sadness, fear and anger) in tone sequences which were melodic analogues of spoken sentences, i.e., that mimicked speech prosody. However, the untrained participants performed at chance level for all emotions – they might not have understood the task, or the stimuli might not have been effective in conveying the emotions. In the second experiment, the task was the same but the set of stimuli included also spoken sentences in English and in a language unknown to the participants, Tagalog. Musicians (n = 28) were better than controls (n = 28) for two emotions, sadness and fear, but not for happiness and anger; the main effect of expertise was not significant. Finally, in the third experiment, six-years-old children assigned to keyboard lessons (n = 10) for a year had better accuracy than children that received no lessons (n = 13) at discriminating pairs of emotions, but only for a subset of pairs: the effect was observed for discrimination between fear and anger, not between happiness and sadness. In the three experiments, results were based on prosodic stimuli produced by a single female speaker, and this limits their generalizability. More recently,

[7] Published in Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion, 11,* 1021-1031. doi:10.1037/a002452

Trimmer and Cuddy (2008) failed to find any effect of musical expertise on prosody. One hundred undergraduate students, who varied in the number of years of musical training (range = 0 – 17; M = 6.5 years), listened to sentences spoken in English and to gliding tone analogues derived from the sentences expressing four emotions through prosody: happiness, sadness, fear and anger (plus neutrality). They rated the prominence of each emotion on 11-point scales. Musical training did not correlate with the recognition of emotional prosody, and it was not a significant predictor in regression analyses. Interestingly, results on emotional prosody were predicted by a measure of emotional intelligence, specifically by the experiential dimension of the Mayer-Salovey-Caruso Emotional Intelligence Test (that assesses perceiving emotions and facilitating thought about emotions; the test also includes a strategic dimension, dealing with understanding and managing one's own emotions). The authors concluded that learning music does not modulate emotion recognition in speech prosody. They speculated that the correlation with emotional intelligence might indicate that recognizing emotional prosody depends on a higher order cross-modal system, unrelated to music training, which would be sensitive to emotions in a modality-independent manner. However, it was not specified how many participants were indeed musically trained. Furthermore, the length of their training, 6.5 years on average, is relatively low in comparison with studies on the effects of musical expertise on language processing with adult participants, where musicians typically have an average of 12 or more years of training (e.g., Kolinsky, Cuvelier, Goetry, Peretz, & Morais, 2009; Marques et al., 2007; Parbery-Clark, Skoe, Lam, & Kraus, 2009; Schön et al., 2004; Thompson et al., 2004). The vocal expression and discrimination of emotions is a basic biological function, arguably shaped by natural selection, thus being in all likelihood efficient even in the absence of specific training (Hauser & McDermott, 2003; Patel, 2008a; Scherer et al., 2003). Hence, a fairly extensive period of musical experience might be required to induce adaptive plastic changes in the underlying neural networks which are large enough to be detected in adults with an offline behavioural task.

Taking into account the inconsistencies of the available evidence, the hypothesis that musical expertise impacts on emotional speech prosody remains open. We set out to examine it in the present study. Forty highly trained musicians, with 12 years of training on average, were compared with forty musically naïve adults in the recognition of emotional speech prosody. With respect to previous studies, important differences were introduced in the procedure. First, we covered a wider range of emotions. The six categories usually analysed in emotion perception research were included: anger, disgust, fear, happiness, sadness and surprise (plus neutrality). Second, the stimuli underwent previous perceptual and acoustic validation (Chapter VII; Castro & Lima, 2010), so that it was known beforehand that they were effective in conveying the intended expressions. Third, besides recognition accuracy, two additional measures were taken: RTs and intensity ratings. RTs were collected to control for possible speed-accuracy trade-offs; because musicians are trained to examine carefully acoustic signals, their putative enhanced accuracy might be a consequence of longer response latencies (Chartrand & Belin, 2006). Intensity ratings were collected to control for the possibility that musicians have increased responsiveness to emotional salience in general. Fourth, participants were from two age cohorts, young and middle adulthood. This allows us to examine whether the putative effect of expertise is general and long-lasting, or whether it is contingent upon being undergoing intensive formal training at the moment of testing, as was the case of most young musicians. Previous research indicates that low-level acoustic properties of speech, such as F_0 and duration, predict listeners' emotion categorizations (e.g., Banse & Scherer, 1996; Juslin & Laukka, 2001). We conducted multiple regression analyses to explore how acoustic cues were used to provide subjective responses by musicians and by controls. Finally, participants were assessed for domain-general cognitive abilities and personality traits to tease apart effects on prosody from general, emotion-unspecific, consequences of training (e.g., Bialystok & DePape, 2009; Schellenberg, 2006). If the recognition of emotions in speech prosody is plastically responsive to musical expertise, then an advantage of highly trained musicians over untrained listeners should be found. This would be evidence in favour of partly shared mechanisms in the domains of speech and music.

8.2. METHOD

PARTICIPANTS | We tested 80 participants distributed into four groups according to musical expertise (musicians and controls) and age (younger and middle-aged), 20 per group (10 women). They also took part in the Study 2 presented in Chapter VI. Their demographic and background characteristics are displayed in Table 6 (Chapter VI, p. 124). Younger participants ranged between 18 and 30 years of age, and middle-aged ones between 40 and 60 years. Musicians were instrumentalists with at least 8 years of continuous musical training (*M* = 12 years; *SD* = 3.2; range = 8 – 19); in both age groups, musicians were similar for years of training, age of training onset, and average hours of weekly instrumental practice. Controls have never had music lessons nor played any instruments.

To assess domain-general cognitive abilities, we used the MOCA, Raven's Advanced Progressive Matrices, and a Stroop task (as an index of cognitive control). As measures of personality and socio-communicative traits, we used the TIPI and AQ questionnaires. Details concerning these measures and how scores differed across groups are described on pages 124-125 (for raw scores, see Table 6, p. 124). In summary, musicians and controls performed similarly on MOCA and Raven's matrices – they did not differ for general intellectual functioning. Musicians had enhanced executive control, as assessed by the conflict condition of the Stroop test, though this effect was partly due to their advantage in processing speed (baseline condition of Stroop). Middle-aged participants performed worse than younger ones for all the cognitive measures, except for the baseline condition of Stroop. No differences between age and expertise groups were found regarding personality and socio-communicative traits.

EXPERIMENTAL STIMULI AND TASK | The stimuli were selected from the database on emotional prosody in Portuguese that we devised and submitted to perceptual and acoustic validation (for details, see Chapter VII). For the current study we used 70 sentences with emotionally neutral semantic content (e.g., "Ela viajou de comboio", *She traveled by train*; "O quadro está na parede", *The painting is on the wall*), 10 tokens per emotional tone (anger, disgust, fear, happiness, sadness, surprise and neutrality). The database contains stimuli produced by two speakers, and care was taken that both were similarly represented in each emotion category to promote variability (5 stimuli from one speaker and 5 from the other speaker).

Participants were told that they would listen to short sentences that were emotionally neutral regarding semantic content, and were asked to focus their attention solely on the tone of voice. The labels of the seven emotional tones were introduced and briefly explained to ensure that they were adequately understood. The task had the same structure as the one used in the validation of the database (Chapter VII). Participants performed two consecutive judgments for each sentence: a forced-choice categorization of the emotional tone, wherein they were asked to respond as accurately and quickly as possible; and an intensity judgment, rating how much the expression was present in the stimulus, on a scale from 1 (low intensity) to 7 (high intensity). Responses were given on a seve-button response pad from Cedrus Corporation, attached to an Apple MacBook Pro computer running SuperLab version 4.0 (Abboud et al., 2006). RTs were measured only for the categorization response (from stimulus onset until the button press). Responses were given with the index finger of the dominant hand, that had to be kept in the same position, the center of the response pad, throughout the session. Each trial contained the following events: a fixation point (+) for 500 ms; the presentation of the sentence with a blank screen until the emotion was categorized; the 7-point scale then appeared on the screen until intensity was rated; the inter-stimulus interval was 3,000 ms. The stimuli were pseudo-randomized and divided into two blocks of 35 trials each; the presentation order was counterbalanced across participants. They were presented via high-quality headphones individually adjusted to a comfortable hearing level. The task started with four practice trials, during which feedback was given, to familiarize the participants with the stimuli; the practice stimuli were not included in the experimental blocks.

8.3. RESULTS

RECOGNITION ACCURACY | A response was considered correct when it matched the intended expression of the utterance. The percentages of correct identifications for each emotion and group are presented in the diagonal cells of Table 13; the distribution of inaccurate responses is presented in rows. Overall accuracy was around 72% correct. Recognition rates ranged between 26% (disgust in middle-aged controls) and 92% (surprise in younger musicians). Low rates for the recognition of disgust are frequently reported in speech prosody research (Banse & Scherer, 1996; Scherer et al., 1991) and were also found in the validation study of these stimuli (Chapter VII).

First we examined how musical expertise and age influenced categorization accuracy. Raw scores were arcsine transformed and submitted to an ANOVA with emotion as repeated-measures factor (anger, disgust, fear, happiness, sadness, surprise and neutrality), expertise (musicians and controls) and age (younger and middle-aged) as between-subjects factors. Accuracy was higher in musicians (77%) than in controls (68%), as indicated by a significant main effect of expertise [$F(1,76) = 9.87$, $p < .01$, $\eta_p^2 = .11$]. The advantage of musicians was similar across age groups and emotions (interactions Expertise x Emotion, Expertise x Age, and Expertise x Age x Emotion ns, ps > .05)[8]. It was also independent of the speaker who produced the stimuli, as examined in an additional ANOVA with speaker as repeated-measures factor (main effect of speaker and interaction Speaker x Expertise ns, ps > .05).

[8] As shown in Table 13, the recognition of sadness was not enhanced in middle-aged musicians (79% correct) vs. controls (85% correct). Given that the main effect was robust and that the interactions were not significant, this probably reflects measure variability. We also sought to examine whether the type of instrument played exerted a role on the effect, but this was limited by the fact that musicians were widely distributed by a varied set of instruments (see page 123). In an exploratory fashion, we categorized musicians according to instrumental group [keyboard, n = 18; strings, n = 12; woodwinds, n = 7; others, n = 3 (drums, accordion, and trombone)] and computed an ANOVA on accuracy rates; the performance was similar across instrumental groups (p > .3).

Table 13. Distribution of responses for each emotion (%, rows), as a function of expertise and age. Diagonal cells in bold indicate correct identifications. Standard errors are presented in parentheses.

	Distribution of responses													
	Controls							Musicians						
Age group /Emotion	Anger	Disgust	Fear	Happiness	Sadness	Surprise	Neutrality	Anger	Disgust	Fear	Happiness	Sadness	Surprise	Neutrality
Younger														
Anger	**69 (5.6)**	9	2	5	1	7	9	**84 (3.6)**	6	2	3	0	2	5
Disgust	3	**55 (7.3)**	0	17	0	17	9	3	**69 (5.5)**	0	15	1	7	8
Fear	3	1	**80 (5.5)**	1	1	12	4	1	0	**82 (3.4)**	2	1	13	2
Happiness	1	1	1	**71 (5.8)**	1	22	4	2	2	0	**76 (5.4)**	1	22	1
Sadness	1	2	5	0	**83 (4.2)**	0	10	2	0	4	0	**87 (3.4)**	0	8
Surprise	1	0	2	5	0	**85 (3)**	9	0	1	0	3	0	**92 (2.9)**	5
Neutrality	1	1	1	0	14	1	**84 (2.9)**	2	1	0	0	8	0	**90 (2.9)**
Middle-aged														
Anger	**50 (5.3)**	15	2	17	1	6	11	**72 (5)**	9	2	7	1	3	7
Disgust	8	**26 (5.6)**	1	36	1	21	9	9	**42 (5.6)**	1	23	0	22	4
Fear	5	6	**48 (6.3)**	3	5	28	7	2	4	**65 (6.7)**	4	3	19	5
Happiness	1	2	1	**62 (4.2)**	1	32	4	3	1	1	**67 (3.3)**	1	28	1
Sadness	2	2	3	0	**85 (3.9)**	1	8	2	2	4	3	**79 (5.1)**	0	11
Surprise	1	3	1	6	1	**81 (4.1)**	9	1	0	1	5	1	**88 (2.7)**	6
Neutrality	1	3	3	1	20	2	**72 (6)**	2	2	0	1	15	0	**81 (2.4)**

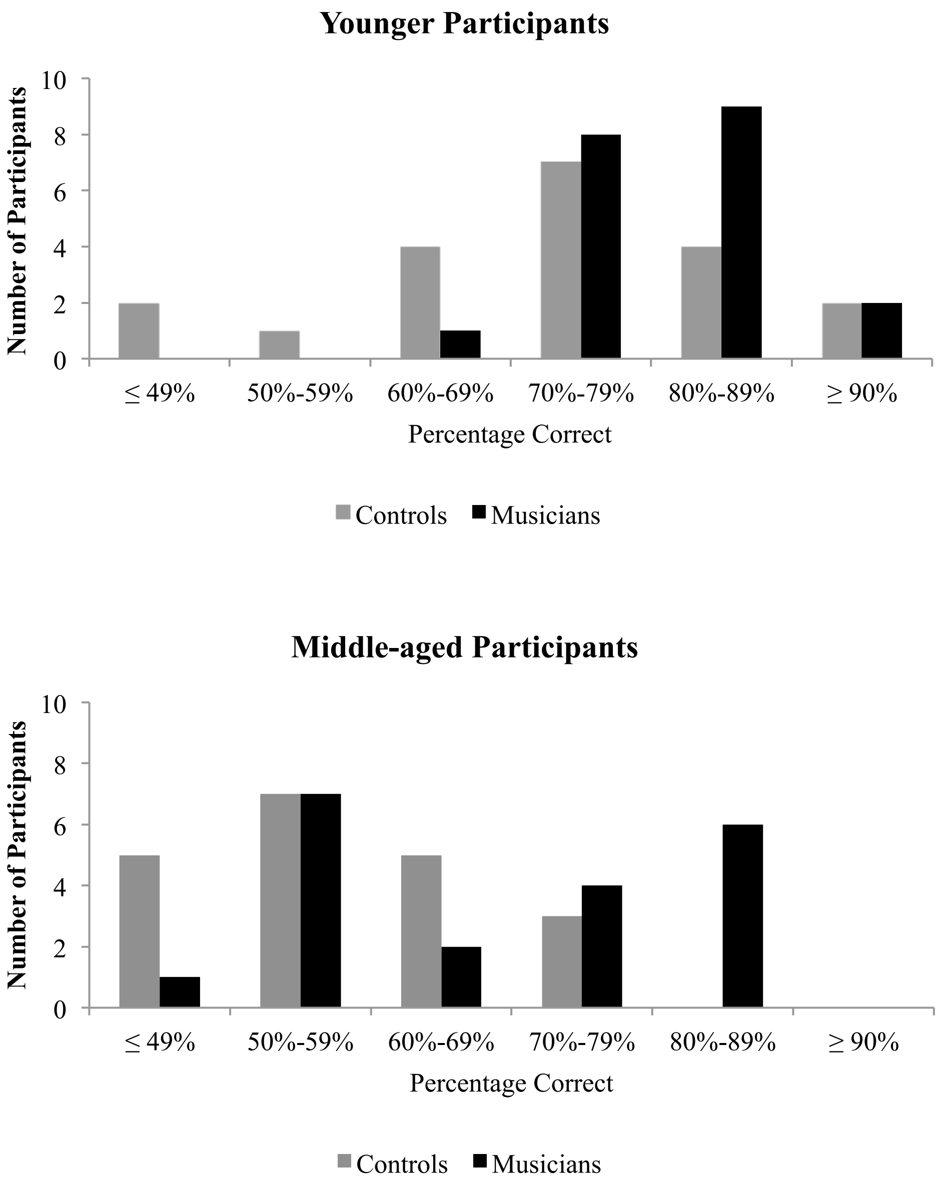
Concerning age differences, middle-aged participants were less accurate than younger ones, but not for all emotions [main effect of age, $F(1,76) = 23.98$, $p < .001$, $\eta_p^2 = .24$; interaction Age x Emotion, $F(6,456) = 4.16$, $p < .001$, $\eta_p^2 = .05$]. Age-related changes were significant for disgust (62% for younger vs. 34% for middle-aged, $p < .01$) and fear (81% for younger vs. 57% for middle-aged, $p < .01$), and approached significance for anger (76% for younger vs. 61% for middle-aged, $p = .06$), as revealed by *post hoc* Tukey HSD tests. No differences were detected for happiness, sadness, surprise and neutrality (*p*s > .4). This effect adds to recent evidence that emotion recognition in prosody changes with age (e.g., Mill et al., 2009; Paulmann et al., 2008b).

The effects of expertise and age were not an artefact of differences in domain-general cognitive abilities, because they were replicated when the scores on the MOCA, Raven's Matrices and Stroop test (both baseline and incongruous conditions) entered as covariates in an ANCOVA [main effect of expertise, $F(1,72) = 10.57$, $p < .01$, $\eta_p^2 = .13$; main effect of age, $F(1,72) = 7.16$, $p < .01$, $\eta_p^2 = .09$; interaction Age x Emotion, $F(6,432) = 3.21$, $p < .01$, $\eta_p^2 = .04$; all other interactions ns, *p*s > .2; age-related decline was significant for disgust and fear, *p*s < .01, marginal for anger, $p = .07$, and ns for the remaining tones, *p*s > .4]. The effects were also replicated in an additional analysis on accuracy scores corrected for possible response biases, *Hu* (Wagner, 1993); unbiased accuracy rates are displayed in Appendix 7. This ANOVA revealed significant main effects of expertise [$F(1,76) = 9.86$, $p < .01$, $\eta_p^2 = .11$] and age [$F(1,76) = 24.64$, $p < .001$, $\eta_p^2 = .24$], as well as the interaction between age and emotion [$F(6,456) = 4.14$, $p < .001$, $\eta_p^2 = .05$]. The age-related decrement was significant for anger, disgust, fear and happiness (*p*s < .01), not for sadness, surprise and neutrality (*p*s > .6). Note that, differently from the main analysis on uncorrected scores, with *Hu* age differences were observed for happiness. This indicates that the stability observed with the uncorrected scores may reflect a bias of middle-aged participants to use the response category *happiness* (see below the pattern of inaccurate responses).

In general, some emotions were better recognized than others [main effect of emotion, $F(6,456) = 37.18$, $p < .001$, $\eta_p^2 = .33$]: sadness was better recognized than all the other emotions except fear (*p*s < .05), and disgust was the most difficult one (*p*s < .05; see

Table 13). To confirm that the categorization rates were above chance level, we compared the obtained accuracy for each emotion and group with the accuracy expected by chance alone (14.3%). Pairwise *t*-tests ($df = 19$) revealed that, in the four groups, all emotions were identified above chance (all *p*s < .0001, except disgust in middle-aged participants, where $p = .05$).

Figure 5. Distribution of younger and middle-aged participants (controls and musicians) as a function of performance level.



To get an in-depth view on the effect of musical expertise, we analysed whether it was general across musicians, or rather produced by a small subset of highly skilled participants. Figure 5 plots the distribution of musicians and controls across performance levels, separately for younger and middle-aged groups. The number of musicians and controls across levels was significantly different ($\chi^2 = 14.1$, $df = 5$, $p < .05$). Most of the musicians clustered in the upper levels of performance (72.5% musicians vs. 40% controls above 70% correct), and most of the controls clustered below (60% controls vs. 27% musicians with less than 70% correct; $\chi^2 = 8.56$, $df = 1$, $p < .01$).

We also examined whether the number of years of musical training was associated with global categorization accuracy. The longer participants were trained, the better they performed; this was observed when the correlation analysis included all participants ($r = .27$, $p < .05$), but not when it included only musicians ($r = -.07$, $p > .6$).

The pattern of inaccurate responses for each emotion was briefly analysed (see Table 13, rows). Anger was confused mainly with disgust, happiness, or neutrality, and disgust with happiness or surprise. Fear and happiness were both misidentified as surprise, and surprise as neutrality and happiness. Neutrality and sadness were confused with each other. This pattern of confusions was observed in previous research with other stimuli (e.g., Adolphs et al., 2002; Juslin & Laukka, 2001; Pell, 2002), as well as in the larger stimulus set from where the stimuli used herein were selected (Chapter VII). To examine whether the distribution of inaccurate responses varied across groups, we computed separate ANOVAS for each emotion, with inaccurate emotion categories as repeated-measures factor (six categories), and expertise and age as between-subjects factors. Musicians and controls did not differ for the pattern of misclassifications (main effects of expertise and interactions Expertise x Age ns, *F*s < 1). This confirms that both groups perceived the general emotional properties of the stimuli in a qualitatively similar fashion. The misidentifications of younger and middle-aged participants were similar for fear, sadness and surprise (interactions Age x Emotion ns, *F*s < 1), but not for the four remaining emotions. Middle-aged participants categorized anger as happiness more often than younger ones [$p < .01$; interaction Age x Emotion, $F(5,380)$

= 2.43, $p < .05$, $\eta_p^2 = .03$], and disgust more often as happiness and surprise [*p*s < .01; interaction Age x Emotion, $F(5,380) = 5.05$, $p < .01$, $\eta_p^2 = .06$]. They also confused happiness more often with surprise [$p < .01$; interaction Age x Emotion, $F(5,380) = 2.58$, $p < .05$, $\eta_p^2 = .03$], and neutrality with sadness [$p < .01$; interaction Age x Emotion, $F(5,380) = 3.06$, $p < .05$, $\eta_p^2 = .04$]. No other effects were significant (*p*s > .05).

INTENSITY JUDGMENTS AND LATENCIES | The mean intensity ratings provided to correct identifications were submitted to an ANOVA with emotion as repeated-measures factor, and expertise and age as between-subjects factors. Ratings were similar across age and expertise groups (main effects of expertise and age ns, *p*s > .05), with the exception that musicians tended to give lower ratings than controls for disgust stimuli [4.5 vs. 5.4, respectively, $p = .09$; interaction Expertise x Emotion, $F(6,456) = 2.93$, $p < .01$, $\eta_p^2 = .04$; intensity ratings are presented in Appendix 8]. Stimuli were judged to express emotions with high intensity, 5.3 on average (maximum = 7): younger and middle-aged controls, means 5.6 and 5.3, respectively; younger and middle-aged musicians, means 5 and 5.1[9]. Intensity ratings differed across emotions [main effect of emotion, $F(6,456) = 30.31$, $p < .001$, $\eta_p^2 = .29$]: disgust (4.9) was rated as less intense than happiness (5.4) and surprise (5.5, *p*s < .05). The four remaining emotions were rated similarly (anger and sadness, 5.3; fear and neutrality, 5.2; *p*s > .1). The other interactions were not significant (*F*s < 1). The similarities between groups for intensity judgments indicate that the advantage of musicians in categorization accuracy does not reflect general differences in the perceived emotional salience of the stimuli.

With respect to RTs, we calculated an ANOVA to verify whether the advantage of musicians was associated with longer response latencies. Inaccurate responses and outliers (RTs exceeding the mean of each participant by 2 *SD*s) were not included in

[9] Because the stimuli were judged as highly intense, a question that arises is whether the effect of expertise on accuracy would also be observed at lower levels of intensity. To approach this concern, we split the 10 stimuli in each emotion into a more intense and a less intense group ($n = 5$ in each), and repeated the ANOVA including intensity as repeated-measures factor. The main effect of expertise was again significant, $F(1,76) = 10.23$, $p < .01$, $\eta_p^2 = .12$, and it did not interact with intensity, $p > .05$. This result indicates that the advantage of musicians over controls was similar across intensity levels. Nevertheless, future studies will benefit from an experimental manipulation of this factor.

this analysis. On average, participants took 3,666 ms to respond (latencies are displayed in Appendix 9). Middle-aged participants were slower (4,079 ms) than younger ones [3,253 ms; main effect of age, $F(6,456) = 30.31$, $p < .001$, $\eta_p^2 = .29$], but musicians did not differ significantly from controls (3,812 vs. 3,512 ms, respectively, $p > .1$). There was a main effect of emotion [$F(6,456) = 5.4$, $p < .001$, $\eta_p^2 = .07$]. Anger (3,701 ms), fear (3,747 ms), happiness (3,660 ms), sadness (3,601 ms), surprise (3,503 ms), and neutrality (3,543 ms) took the same time to be categorized (ps > .1), but disgust (3,909 ms) took longer than some of them (happiness, sadness, surprise and neutrality; ps < .05). None of the interactions reached significance (ps > .3). The correlation between accuracy and latencies was not significant, $r = -.14$, $p > .05$, confirming the absence of a speed-accuracy trade-off.

ACOUSTIC CUES AS PREDICTORS OF EMOTION CATEGORIZATION | Constellations of low-level acoustic attributes of voice predict the subjective emotion judgments performed by the listeners (e.g., Banse & Scherer, 1996; Juslin & Laukka, 2001; Sauter, Eisner, Calder, et al., 2010). We conducted multiple regression analyses to explore how the responses of musicians and controls were predicted by patterns of acoustic cues. First, each emotion portrayal was measured regarding major voice cues related to F_0, intensity, voice quality and temporal aspects. Specifically, we covered 10 acoustic parameters: F_0 mean, variability (*SD*), minimum, maximum, and small-scale perturbations (jitter); intensity mean and variability (*SD*); relative proportion of high-frequency energy (cut-off 500 Hz); total duration and proportion of pauses (ratio between the duration of silent parts and the total duration of the stimulus). The proportion of high-frequency energy is an index of voice quality, so that an increased amount of energy above a certain cut-off frequency is associated with a more sharp and less soft voice (e.g., Juslin & Laukka, 2003). Praat software (Boersma & Weenink, 2009) was used to compute these measures. The results of the acoustic analyses are displayed in Appendix 10. Acoustic cues were taken as predictors in the regression analyses, with the exception of intensity *SD* and F_0 maximum. Intensity *SD* was excluded because it was invariant across emotions ($p > .1$; all the other cues assumed different values across emotion categories, ps < .05), and F_0 maximum because it correlated very highly with F_0 mean and *SD*, rs > .8 (Juslin & Laukka, 2001). The dependent variable was the number of participants who choose each emotion category (e.g., for a certain stimulus, if 20 participants responded happy, the happiness score for that stimulus was 20; if none of the participants responded happy, the score was 0; Banse & Scherer, 1996). Two simultaneous multiple regression analyses were calculated for each emotion, one for controls and another for musicians. Table 14 presents the main results of these analyses: beta weights and the proportion of variance explained by the acoustic measures (adjusted R^2).

Table 14. Results from multiple regression analyses on the listeners' utilization of acoustic cues (lines) for the categorization of emotions (columns), separately for controls and musicians. Values correspond to beta weights. Adjusted R^2 are also shown.

Voice cue	Anger	Disgust	Fear	Happiness	Sadness	Surprise	Neutrality
Controls							
F_0 mean	.26	-.24	.83*	.42*	-.83*	.21	-.47*
F_0 *SD*	-.19	.3	-.76*	.21	.15	.27	-.01
F_0 minimum	-.3*	.03	-.09	-.25$^!$	.17	.32*	.01
Jitter	.02	-.05	.28*	-.2$^!$	-.1	-.02	.09
Intensity (mean)	-.17	-.02	.11	-.01	-.28*	-.01	.4*
HF 500 Hz	.54*	.18	-.15	-.04	-.21*	-.07	-.09
Total duration	-.38*	.3*	.09	.16	-.12	-.03	.1
Pause proportion	-.03	.27$^!$	.05	.16	-.09	-.26*	.06
Adjusted R^2	.42*	.18*	.41*	.39*	.49*	.32*	.30*
Musicians							
F_0 mean	.23	-.27	.83*	.42*	-.84*	.17	-.47*
F_0 *SD*	-.14	.33$^!$	-.72*	.18	.19	.28	-.08
F_0 minimum	-.28*	.01	-.1	-.22$^!$	.19	.34*	-.01
Jitter	-.01	-.05	.3*	-.17	-.12	-.06	.1
Intensity (mean)	-.16	-.04	.08	.01	-.27*	-.04	.4*
HF 500 Hz	.5*	.09	-.15	-.01	-.22*	-.12	-.04
Total duration	-.4*	.32*	.04	.19$^!$	-.13	-.04	.13
Pause proportion	.01	.23$^!$	.06	.09	-.04	-.25$^!$	.01
Adjusted R^2	.41*	.17*	.38*	.32*	.45*	.31*	.36*

Note. HF = high-frequency energy
* $p < .05$; $^!$ $p \leq .1$

All emotion categories were significantly predicted by some combination of acoustic cues (*p*s < .05). The explained variance ranged between .17 (disgust in musicians) and .49 (sadness in controls). An inspection of Table 14 shows that the constellation of cues yielding significant beta weights was unique for each emotion, thus confirming that listeners relied on differentiated acoustic profiles to categorize emotions (e.g., Juslin & Laukka, 2001, 2003). Concerning the impact of musical expertise, these analyses revealed two important findings. First, the predictive strength of acoustic measures was similar in controls and musicians [the explained variance, adjusted R^2, was the same across groups, .36 and .34, respectively; $t(12) = .31$, $p > .7$], indicating that both groups were similarly efficient and consistent in the utilization of the low-level cues analysed here. Thus, the musicians' enhanced categorization accuracy appears to be driven by mechanisms other than general advantages in low-level auditory processing. Second, the emotion-specific patterns of cues that predicted responses were largely common in both groups – cues reaching significant beta weights for anger, fear, sadness, surprise and neutrality were the same in musicians and controls; for happiness, jitter was slightly more predictive in controls and duration in musicians, and for disgust F_0 *SD* was slightly more predictive in musicians than in controls. The similarities between groups in the relative weighting of acoustic cues suggest that the inference rules that musicians used to recognize emotional prosody are not qualitatively different from the ones used by controls.

8.4. DISCUSSION

Does emotion recognition in speech and in music engage common mechanisms? We examined this question by determining whether musical expertise modulates the processing of emotional speech prosody. A robust effect of expertise was found: musicians were more accurate than musically naïve listeners at identifying emotions in speech. The effect was general across the seven emotional tones tested, was observed in two age cohorts (young and middle adulthood), and was widespread across the participants. It is not a by-product of differences in socio-educational background, domain-general cognitive abilities or personality traits. It cannot also be accounted for by trade-offs with latencies, or by increased responsiveness to emotional salience – RTs and intensity judgments were similar in musicians and controls. Additionally, both groups perceived the general emotional properties of the stimuli in a qualitatively similar way – they did not differ in the pattern of misclassifications, nor in the profile of acoustic cues that were predictive of their categorization responses. Thus, musical expertise is associated with specific cross-domain benefits in the ability to categorize emotions in speech prosody. This suggests that extensive musical training drives adaptive plasticity in the neurocognitive mechanisms underlying emotional prosody. Another finding was an age-related decrement in accuracy for the recognition of anger, disgust and fear (not for happiness, sadness, surprise or neutrality). This is in line with previous evidence that the recognition of emotions in speech prosody, particularly negative ones, changes with age with an early onset in middle-age (Mill et al., 2009; Paulmann et al., 2008b).

The effect of musical expertise on emotional speech reported herein clarifies previous conflicting results. The initial findings by Thompson and colleagues (2004) were replicated and extended in important ways: (a) with well-controlled stimuli that were accurately perceived by musicians and controls and that were produced by more than on speaker, the advantage of musicians was general, not restricted to a subset of emotions; (b) the effect was obtained in a wider range of emotions, specifically in the six categories usually tested in emotion perception research – anger, disgust, fear, happiness, sadness and surprise; (c) it was observed in younger as well as middle-aged musicians, thus indicating that it is general and long-lasting; and (d) potential confounds not analysed previously were excluded, namely possible differences in latencies and in perceived intensity. In contrast, the present results stand at odds with the null findings reported by Trimmer and Cuddy (2008). A likely explanation for this discrepancy is the level of musical training of the participants in both studies. Here, we investigated expert musicians with a minimum of eight years of continuous formal training, 12 years on average, whereas in the Trimmer and Cuddy's study the trained subjects had a lower level of expertise, 6.5 years on average. Indeed, in other studies where positive effects of musical expertise were found, musicians also had quite an extended practice: for instance, 13 years in Thompson et al.'s (2004), at least 14 years in Marques et al.'s (2007), 15 years in Schön et al.'s (2004), and 16 years in Kolinsky et al.'s (2009) and in Parbery-Clark et al.'s (2009). Such an extensive training might be required to detect experience-dependent behavioural differences in adults, especially in domains tapping important adaptive abilities such as emotional vocalizations (Hauser & McDermott, 2003; Patel, 2008a; Scherer et al., 2003). There are studies reporting effects of musical expertise after short training periods, including one year or even less (e.g., Fujioka, Ross, Kakigi, Pantev, & Trainor, 2006; Moreno et al., 2009; Thompson et al., 2004), but these are mostly based on neural measures with children participants, who are more plastic and presumably have a less varied exposure to life experiences. Another factor that may have played a role is that here the chances of observing the effect were increased because more emotions, and more stimuli per emotion, were included than in Trimmer and Cuddy's study (2008).

From a neurocognitive standpoint, the effect of expertise suggests that recognizing emotions in speech prosody and in music recruits common mechanisms. This converges with other evidence that music and language depend on domain-general shared mechanisms to a significant extent (e.g., Besson et al., 2011; Koelsch et al., 2005; Moreno et al., 2009; Patel, 2008a; Schön et al., 2010; Wong, Skoe, Russo, Dees, & Kraus, 2007). It also fits nicely with the theoretical framework by Juslin and colleagues on the mechanisms underlying emotional responses to music (Juslin et al., 2010; Juslin & Västfjäll, 2008). One of them, emotional *contagion*, contends precisely that music conveys emotional expressions because it mimics the sound patterns of vocal emotions, thereby engaging neural resources dedicated to voice; by contagion, the perceived expression would trigger the corresponding emotional experience. Therefore, musical training, by operating changes on musical emotions, would concomitantly change prosodic ones. Arguing that emotion processing in speech and music engage common mechanisms does not entail that the overlapping is total, i.e., that there are no differences across domains. As noted by Juslin and colleagues (ibd.), the richness of emotional reactions to music does not depend on the affinity with speech alone, involving other mechanisms as well, namely brainstem reflexes, episodic memory, musical expectancies, visual imagery, evaluative conditioning and rhythmic entrainment. Furthermore, Ilie and Thompson (2006) showed that there might be functional differences in affective responses for speech and music, though these remain largely unspecified.

Future studies will be valuable to determine the stage(s) of processing where mechanisms for emotional processing are shared, or the locus of the cross-domain transfer. Considering Schirmer and Kotz's (2006) model on vocal emotional processing, it might be as early as in low-level auditory processes, later in the integration of emotionally significant acoustic cues, or in higher order cognitive processes. Recently, Strait, Kraus, Skoe and Ashley (2009) found that musical expertise influences auditory subcortical responses to vocal signs of emotion: brainstem potentials to an infant's unhappy cry were enhanced and faster in musicians. There is also evidence that musicians have earlier and larger brainstem responses than untrained participants to linguistic stimuli (Musacchia, Sams, Skoe, & Kraus, 2007; see also Wong et al., 2007). Thus, one locus for the speech-music interaction might be at an early sensory level of processing. However, the early modulations uncovered in these studies can be driven by top-down mechanisms (Kraus & Chandrasekaran, 2010; Strait et al., 2009). Additionally, it is unclear whether they relate to the behavioural effect reported here, where integrative, explicit categorization is involved. Globerson, Lavidor, Golan, Kishon-Rabin, and Amir (2010) found that basic pitch discrimination abilities are indeed associated with emotion categorization in prosody, but Leitman and colleagues (2010) and we (Carvalho, Lima, & Castro, 2011) failed to find such an association in healthy adults. Importantly, in multiple regression analyses, we observed that musicians were not more consistent than controls in how they relied upon low-level stimuli's acoustic cues to perform emotion categorizations. This suggests that their enhanced accuracy is not reducible to general advantages in basic auditory mechanisms. The effect might occur at processing stages involving the conceptual system of emotion. Barrett and colleagues (Barrett, 2006b, 2009; Lindquist & Barrett, 2010) have proposed that individuals differ in the granularity of their emotion concepts: persons with lower emotional granularity perceive affective states as broad and undifferentiated categories with low specificity, and those with higher granularity perceive more precise and differentiated emotional states. They further suggest that emotional granularity is malleable and can be trained, in the same way that wine experts are trained to perceive subtle differences that novices are unaware of. Analogously, musicians might develop sharper (granular) emotion concepts for music because an important part of their training and professional activity as performers or teachers concerns the expression and perception of musical emotions. Given that the acoustic codes of emotions are similar in speech and music (Bowling et al., 2010; Curtis & Bharucha, 2010; Juslin & Laukka, 2003), perceiving emotional speech would benefit from finely grained concepts for musical emotions. Specifying which are the shared stages of emotional processing in speech and music will also explain why musicians were superior to controls in the identification of disgust and surprise, emotions that are not frequently reported in response to music (Juslin, 2009; Juslin & Laukka, 2003; Zentner et al., 2008). If the effect results from an advantage on basic auditory mechanisms, then musicians would be expected to perform generally better, irrespective of emotion category. However, enhanced recognition of these arguably less trained emotions is also compatible with a higher level locus of the effect: if musicians develop sharply defined concepts for a specific set of emotions, this might indirectly benefit their performance on other categories by restricting the space of the categorization problem (for example, when presented with disgust stimuli, musicians might reject happiness or fear as response candidates more easily than controls, thus producing more finely tuned, and accurate, identification).

Two shortcomings of the current study should be discussed. The first concerns *causality*. Like most research on the effects of musical expertise, we implemented a quasi-experimental design (Bialystok & DePape, 2009; Bigand et al., 2005; Brochard, Dufour, & Després, 2004; Gaser & Schlaug, 2003; Kolinsky et al., 2009; Marques et al., 2007; Pantev et al., 1998; Ro, Friggel, & Lavie, 2009; Schellenberg & Moreno, 2009). Hence, we cannot definitely ascertain that musicians' enhanced performance stems from training – they could have been particularly skilled already before being trained. However, this possibility is unlikely because musicians and controls in our sample were carefully matched for background socio-educational variables, domain-general cognitive abilities, as well as for personality traits, and this makes the groups highly comparable. Furthermore, studies that assessed children pre- and post-training provided strong evidence for the causal relationship between music lessons and changes in brain structure and function (Fujioka et al., 2006; Hyde et al., 2009; Moreno et al., 2009). Thus, differences reported in cross-sectional designs are likely to be a true contingence of training, in other words, a consequence of music-driven plasticity. This fits in well with research on expert performance suggesting that high levels of proficiency in a certain domain are a result of extended deliberate practice rather than of static initial predispositions (for reviews, Ericsson & Lehmann, 1996; Ericsson & Towne, 2010). Another possibility worth considering regarding causality is that musicians' enhanced categorization of prosody might result from non-musical dimensions of training (Schellenberg, 2005, 2006). That is, musical training, like chess or math lessons, is an extra out-of-school activity that is school-like in that it involves demanding "cognitive, adult-supervised activities that require serious and concerted effort on the part of the participant in order to acquire knowledge" (Schellenberg, 2006, p. 465). Such broad, music-unspecific, learning experiences are likely to foster general cognitive abilities by exaggerating the schooling effect. Note, though, that in the present study the positive effect of musical expertise was not general, as this hypothesis would predict: musicians had a selective advantage in recognizing emotional speech, but performed like controls on domain-general cognitive measures; musicians were better on the Stroop test, but this did not account for the effect on prosody (see Results). Therefore, we are confident that the participants' extensive experience in the specific domain of music is the most likely cause of the observed effect.

The second shortcoming concerns *specificity.* Participants were tested for emotion recognition in prosody and we cannot exclude that musical expertise affects also the recognition of emotional stimuli in other domains, such as facial expressions. This would be expected if the impact of expertise were on putative general emotion processes. Indeed, some studies suggest that there are supramodal mechanisms in emotion recognition. Borod and colleagues (2000) reported a positive correlation between emotion recognition in prosody and in faces, even though it turned out to be non-significant when demographic and cognitive characteristics of the participants were entered as covariates. Peelen, Atkinson and Vuilleumier (2010) observed that perceiving emotions in different modalities recruits some common, modality-independent, brain regions, namely the medial prefrontal cortex and the left superior temporal sulcus. Only future studies will clarify whether musical expertise modulates specifically emotions in the auditory domain, or whether it affects general supramodal processes. Available evidence points tentatively to the former hypothesis: previous studies showed that musical training does not affect performance on measures of emotional intelligence, which are meant to cover general emotion skills (Resnicow et al., 2004; Trimmer & Cuddy, 2008). Furthermore, different lines of evidence indicate that auditory emotions involve mechanisms distinct from those involved in other modalities. Bänziger, Grandjean and Scherer (2009) compared the perception of emotions in prosody, faces and bodies, and a principal component analysis revealed two separate factors in emotion perception ability, one for the visual modality and another for the auditory one. Emotion recognition in facial expressions and in voice were shown to be neurally dissociable (Adolphs et al., 2001; Drapeau et al., 2009; Hornak, Rolls, & Wade, 1996; Rohrer et al., in press). Moreover, the encoding of emotional information in vocalizations may have a specific phylogenetic status (e.g., Hauser & McDermott, 2003; Juslin & Laukka, 2003; Scherer et al., 2003).

8.5. CONCLUSION

This study substantiates a processing interplay between speech and music: we established a robust effect of musical expertise in the identification of emotional speech prosody, which was general across emotion categories and age cohorts. We believe that this paves the way for further systematic research on the neurocognitive relations between vocal and musical emotion processing, with implications for both applied and fundamental psychological science. If speech and music set into action common mechanisms, music therapy might be a useful device in language rehabilitation programs. Music would be useful to strengthen, via adaptive plasticity, the neurocognitive processes that are defective in individuals with prosodic problems (e.g., Golan, Baron-Cohen, & Hill, 2006; Pell & Leonard, 2003). From a fundamental research standpoint, future studies will specify which cognitive and neural mechanisms are common across domains, as well as the magnitude of the overlapping, i.e., to what extent it is total or partial. This will contribute to the understanding of the complex relations between language, music, and emotions in the mind/brain system.

CHAPTER IX

A Dissociation Between Emotion Recognition in Music and in Speech Prosody in Parkinson's Disease

9.1. INTRODUCTION

The recognition of emotions in speech prosody can be impaired in patients with PD (Gray & Tickle-Degnen, 2010), and this impairment has been attributed to the dysfunction in the basal ganglia[10](e.g., Dara et al., 2008; Pell & Leonard, 2003; Schröder, Nikolova, & Dengler, 2010). Does PD compromise similarly the recognition of emotions in music? Answering this question will shed light on the degree of the neural overlap across domains. It will also have implications for our understanding of the neurocognitive substrates of musical emotions and, more generally, of the disease clinical phenomenology. In this study, we compared for the first time the impact of PD in emotion recognition in music and in speech prosody[11].

The neuropathological hallmark of PD is the progressive loss of dopamine neurons in the substantia nigra pars compacta and the subsequent striatal dopamine deficiency, which leads to a cascade of dysfunctions in the basal ganglia pathways (e.g., Obeso et al., 2008). These pathological changes are ultimately responsible for the development of the cardinal motor syndrome that characterizes PD clinically, that may include rigidity, resting tremor, akinesia (lack of movement), bradykinesia (slowing of movement) and postural instability (e.g., Assogna, Pontieri, Caltagirone, & Spalletta, 2008; Dickson et al., 2009; Kehagia, Barker, & Robbins, 2010; Utter & Basso, 2008). Non-motor symptoms are also common, namely cognitive and emotional difficulties, because the basal ganglia are richly interconnected with the rest of the brain and modulate the

[10] The basal ganglia are a group of subcortical nuclei that include the globus pallidus (external and internal segments), subthalamic nucleus, substantia nigra (compact and reticular parts) and the striatum, which is composed of the caudate, putamen and nucleus accumbens. These nuclei are strongly interconnected with the cerebral cortex, thamalus and other brain regions (e.g., Nolte, 2002)
[11] Submitted for publication in Lima, C. F., Garrett, C., & Castro, S. L. (submitted). Perceiving emotions in music and in speech prosody is dissociated in Parkinson's disease.

activity of diverse cortical regions: they receive broad cortical inputs and, after processing that information, project back to the cortex via thalamus. They are involved in several functionally segregated parallel circuits, concretely motor, oculomotor, prefrontal, association and limbic loops (e.g., Gale, Amirnovin, Williams, Flaherty, & Eskandar, 2008; Nambu, 2008; Obeso et al., 2008; Utter & Basso, 2008). Frequent non-motor symptoms are depression, apathy and anxiety, as well as impairments in executive control, attention, visuospatial abilities, working memory and learning. More severe features, such as psychotic symptoms (e.g., visual hallucinations) and full dementia might also occur, mainly in the later stages of the disease (e.g., Chaudhuri, Healy, & Schapira, 2006; Chaudhuri & Schapira, 2009; Kehagia et al., 2010; Watson & Leverenz, 2010).

Neuropsychological research indicates that PD patients recognize emotions in speech prosody less accurately than healthy matched controls (e.g., Ariatti, Benuzzi, & Nichelli, 2008; Breitenstein, Lancker, Daum, & Waters, 2001; Dara et al., 2008; Pell & Leonard, 2003; Péron et al., 2010; Sheila Scott, Caird, & Williams, 1984; Yip, Lee, Ho, Tsang, & Li, 2003). In a meta-analytic review of 28 comparisons, Gray and Tickle-Dengnen (2010) concluded that the size of this effect is highly heterogeneous across samples, but it can be observed with different tasks, namely forced-choice, emotion discrimination and rating tasks. The patients' difficulties tend to be more severe for negative emotions (anger, disgust, fear and sadness) than for relatively positive ones (happiness and surprise). Additionally, they are not secondary to comorbid depression, and may be detected when the patients are optimally medicated during testing as well as when they are in a hypodopaminergic state. Difficulties in processing emotional prosody were also observed in patients with focal lesions in the basal ganglia (Paulmann et al., 2008a) and with Huntington's disease (Speedie et al., 1990), a neurodegenerative genetic disorder that affects the striatum. Moreover, some fMRI studies found activations in the putamen and caudate when participants categorize emotions in speech prosody (Frühholz et al., in press; Kotz et al., 2003; Leitman, Wolf, et al., 2010), though others did not (Johnstone, Reekum, Oakes, & Davidson, 2006; Mitchell & Ross, 2008). Therefore, it has been suggested that the basal ganglia are part of the neurocognitive network that subserves vocal emotional processing (Frühholz et al., in press; Paulmann et al., 2011). Their functional role remains unclear, however. Electrophysiological studies indicate that they are not mandatorily involved in early implicit prosodic processes: patients with focal lesions in the basal ganglia show normal ERP responses to emotional prosody in implicit tasks, even though they are impaired in explicit emotion categorization (Paulmann et al., 2011; Paulmann et al., 2008a; but see, Schröder et al., 2009). One hypothesis is that they contribute, at an intermediary stage of processing, to the meaningful perceptual sequencing of dynamic, time-dependent information; the striatum would act as a binding mechanism for the multiple emotionally significant acoustic parameters that unfold in the speech stream, paving the way for further cortical processes of emotion recognition (Kotz & Schwartze, 2010; Meyer, Steinhauer, Alter, Friederici, & Cramon, 2004; Paulmann & Pell, 2010b; Pell & Leonard, 2003). Alternatively, the basal ganglia may contribute primarily for the later stages of prosodic processing. Through the prefrontal-striatal loop, they would be implicated in the executive and working memory processes that are necessary for higher order prosody recognition, namely for the decision-making and labelling operations required by off-line emotion recognition tasks (Benke, Bösch, & Andree, 1998; Mitchell & Bouças, 2009). In fact, according to the meta-analysis by Gray and Tickle-Degnen (2010), the prosodic difficulties found in PD might be partly associated with the patients' executive and working memory impairments. This possibility awaits further empirical specification.

Whether PD has a detrimental impact on emotion processing in music remains poorly investigated. To the best of our knowledge, only one study examined this question up to now. Using a forced-choice task, Tricht, Smeding, Speelman, and Schmand (2010) compared 20 patients and 20 controls at recognizing happiness, sadness, fear and anger in 32 music excerpts (eight per emotion) taken from different periods of the musical history (baroque, classical and romantic periods) and with different instrumentations (solo, chamber and orchestral music). They found that patients were less accurate than controls at categorizing anger and fear, whereas they were normal for happiness and sadness. The authors further observed that this pattern was independent of the magnitude of participants' depression symptoms, and could not be explained by executive dysfunction or by low-level music perception defects, as inspected by two subtests (Scale and Meter) of the Montreal Battery of Evaluation of Amusia (MBEA; Peretz et al., 2003). However, some caveats limit the interpretation of these results: performance reached ceiling levels for two of the four emotions (happiness and sadness); the fact that the stimuli were taken from the classic repertoire makes it difficult to tease apart emotional and familiarity effects; and the number of negative and positive emotions was unbalanced (3 vs. 1, respectively), precluding conclusions regarding effects of valence. Additionally, whilst this study suggests the involvement of the basal ganglia for negative musical emotions, findings from neuroimaging suggest that those structures are involved in positively-valenced responses to music. The striatum is activated while we listen to pleasant music (Blood & Zatorre, 2001; Koelsch et al., 2006), and these activations are associated with the release of endogenous dopamine, both during reward anticipation and during the proper experience of pleasure evoked by music (Koelsch, 2010; Salimpoor et al., 2011). Altogether, these studies indicate that the basal ganglia might play a role for emotion recognition in music, but its specific contribution is unclear. Omar and colleagues (2011) also found an association between emotion recognition in music and the volume of the striatum in frontotemporal lobar degeneration, but they did not analyse individual emotion effects.

The main goal of this study is to compare directly how idiopathic PD affects emotion recognition in speech prosody and in music. An identical profile of impairments across domains will suggest that they recruit shared neural substrates – probably the same basal ganglia pathways – in a functionally similar manner. By contrast, differential effects for prosody and music will indicate that emotion recognition in these domains is dissociated, that it, they are partly independent of each other. This would be evidence that domain-specific mechanisms are engaged. Non-demented and non-depressed PD patients in the early stages of the disease were compared with healthy controls of similar age, education and musical training. We examined four emotions in each domain, two negative and two relatively positive ones. The emotion recognition task was identical across domains: the participants listened to each music excerpt, or spoken sentence, four times, and on each time they rated how much it expressed one of the four emotions. Such a rating task is highly sensitive (e.g., Adolphs et al., 1998; Adolphs & Tranel, 1999; Adolphs et al., 2001) and puts fewer demands on working memory and decisional/executive processes than conventional forced-choice tasks. It has rarely been used with PD patients (Gray & Tickle-Degnen, 2010; Pell & Leonard, 2003). On the basis of the reviewed literature, for speech prosody we hypothesized that PD patients could display a deficit, particularly for the recognition of negative emotions, and that this deficit might be related to executive abilities. For music, no predictions could be made with certainty in light of the limited available data. Participants were also tested for the recognition of emotions in facial expressions, taken as a control measure to disentangle auditory-specific from general emotion processes (Patel, 2008b). Typically, PD patients are not impaired for facial expressions (e.g., Adolphs et al., 1998; Cohen, Gagné, Hess, & Pourcher, 2010; Pell & Leonard, 2005), or show highly selective impairments for specific emotions (e.g., anger in the study by Lawrence, Goerendt, & Brooks, 2007; disgust in the study by Suzuki, Hoshino, Shigemasu, & Kawamura, 2006). Participants underwent a comprehensive neuropsychological assessment in order to determine whether domain-general cognitive processes, namely executive and working memory abilities, or low-level processing defects, would account for the putative impairments in emotion recognition. Additionally, we explored a possible dissociation between music perception and emotion recognition in music. Evidence concerning this topic is not clear (Stewart, Kriegstein, Dalla Bella, Warren, & Griffiths, 2008). Peretz and colleagues (1998) examined the brain-damaged patient I.R., who had normal emotion recognition in music, but impaired music perception and recognition, suggesting separate pathways. On the other hand, McDonald and Stewart (2008) found that amusic individuals (impaired music perception) have abnormal emotional responses to music, suggesting association: they reported fewer psychological changes linked with listening to music, and stronger negative feelings about imposed music. Here, we examined music perceptual-cognitive abilities with the MBEA (Peretz et al., 2003).

9.2. METHOD

PARTICIPANTS | Twenty-four patients with idiopathic PD, non-demented and non-depressed, were recruited through the Movement Disorders Unit at the Department of

Neurology of Hospital de S. João, Porto, Portugal. The diagnosis was confirmed by an experienced neurologist based on the criteria of the United Kingdom Parkinson's Disease Brain Bank (Hughes, Daniel, Kilford, & Lees, 1992). All patients were physically independent. According to the 5-stage scale by Hoehn and Yahr (1967), the severity of their motor symptoms ranged from Stage 1 (unilateral symptoms only) to Stage 3 (balance impairment, mild to moderate disease). They scored 17.9 (maximum 56) on the motor scale of the Unified Parkinson's Disease Rating Scale (UPDRS; Fahn, Elton, & UPDRS program members, 1987). Motor symptoms were characterized as left dominant (n = 9), right dominant (n = 10) or bilateral (n = 5). Disease subtype (Schiess, Zheng, Soukup, Bonnen, & Nauta, 2000) was classified as tremor-dominant (n = 6), akinetic-rigid (n = 12) or mixed (n = 6). The patients were optimally medicated during the testing sessions (*on* state) as follows: carbidopa/L-dopa (n = 23), dopamine agonist (n = 19), MAO-B inhibitor (n = 13), COMT inhibitor (n = 5), amantadine (n = 2) and anticholinergics (n = 1). Four of them were also taking antidepressants (sertraline, n = 2; fluoxetine, n = 1; trazodone, n = 1). The patients were compared with 25 healthy controls of similar age, education and musical training. Patients and controls had no other known major neurological or psychiatric conditions. The clinical and demographic characteristics of the participants are detailed in Table 15.

This study was approved by the ethics committee of the Hospital de S. João, and written informed consent was obtained according to the declaration of Helsinki. Participants were financially compensated for their time. They were assessed in two individual sessions lasting about two hours each; the sessions with patients were scheduled for the time of the day at which their motor symptoms were typically least severe.

NEUROPSYCHOLOGICAL ASSESSMENT AND MUSIC PERCEPTION | The results of the neuropsychological and music perception tests are also summarized in Table 15. Dementia was excluded using the MMSE (inclusion criterion ≥ 24; Folstein et al., 1975; portuguese version, Guerreiro et al., 1994), and depression using the Geriatric Depression Scale (inclusion criterion ≤ 9; one patient and one control scored borderline, 10; Yesavage et al., 1983; portuguese version, Barreto, Leuschner, Santos, & Sobral, 2003). All participants had acceptable hearing thresholds, as determined by a pure-tone audiometric screening (minimum 30 dB HL for frequencies of 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz in the better ear).

We assessed verbal intelligence with the Vocabulary test of the Wechsler Adult Intelligence Scale (WAIS-III; Wechsler, 2008), in which participants are asked to provide verbal definitions of words (e.g., *What is a boat?*); scores in this test correlate strongly with the full scale IQ. Auditory working memory was assessed with the Digit Span test of the same scale: in the forward condition, participants repeat increasingly longer sequences of numbers in the same order as the experimenter; in the backward condition, they repeat sequences of numbers in the reserve order. To examine executive functions, one-minute phonemic (letters F, A, S) and semantic (animals) verbal fluency tests were used (Lezak, Howieson, & Loring, 2004), as well as a Stroop task. In this Stroop task, participants were asked to name colours as fast as possible in two conditions: without incongruity (baseline), where stimuli consisting of XXXX were printed in blue, pink, green or grey; and with incongruity, where stimuli consisting of written words denoting these colour names (*azul*, blue; *rosa,* pink; *verde*, green; *cinza,* grey) were printed in an incongruent ink colour (e.g., the word *azul*, blue, printed in green ink; 112 items in each condition). In the later condition the written word was to be ignored and the ink to be attended to. The time needed to name colours without incongruity was taken as a proxy for processing speed, and the time difference between the incongruous and baseline conditions – interference cost –, was taken as a proxy for cognitive control. To ensure that participants had no low-level defects at processing speech and faces, we used, respectively, the nonword minimal pair discrimination test of the Portuguese Psycholinguistic Assessments of Language Processing in Aphasia (short-form; Castro et al., 2007) and the Benton Facial Recognition Test (short-form; Benton, Sivan, Hamsher, Varney, & Spreen, 1994). In the nonword discrimination test, participants perform same/different judgments on pairs of nonwords; in half of the pairs, the two nonwords are identical (e.g., *sápa/sápa*), and in the other half they differ in a single phoneme (e.g., *dul/gul*). In the Benton's test, participants are asked to match a target face with the identical face(s) out of six options.

Table 15. Background and neuropsychological information of the patients and controls. *SD*s are presented in parentheses.

Characteristics	Patients (*n* = 24)	Controls (*n* = 25)	*p*
Age (years)	61.8 (11.8)	59.2 (11.9)	.45
Education (years)	7.8 (4.3)	9.6 (5.0)	.17
Gender	17M/7F	12M/13F	**< .05**
Musical training (years)	0.1 (0.4)	0.5 (1.3)	.16
Hoehn and Yahr stage	1.9 (0.5)	-	-
Motor UPDRS[1]	17.9 (9.1)	-	-
Disease duration (years)	8.3 (4.9)	-	-
Age of disease onset	53.5 (10.3)	-	-
MMSE (/30; cut-off 24)	28.3 (1.5)	28.9 (0.9)	.11
Vocabulary WAIS-III (/66; raw score)	38.4 (13.3)	40.3 (13.5)	.62
Digit Span WAIS-III			
Forward (/9)	5.1 (1.3)	5.4 (1.2)	.5
Backward (/9)	4.1 (1.4)	4.1 (0.9)	.99
Verbal Fluency (number of items)			
Letters F, A, S	25.5 (10)	27.7 (9.5)	.42
Animals	14.5 (5.2)	17.1 (5.5)	.1
Stroop test[2]			
Congruous colour naming (s/item)	1.1 (0.4)	0.9 (0.2)	**< .05**
Interference (incongruous - congruous)	1.1 (0.8)	0.5 (0.3)	**< .01**
Geriatric Depression Scale (/30; cut-off 9)	3.6 (2.7)	2.5 (2.7)	.17
Benton Facial Recognition Test (/27)	21 (3.4)	21.8 (3)	.41
PALPA-P[3] minimal pairs discrimination (/32)	31.5 (.8)	31.6 (0.6)	.38
Montreal Battery of Evaluation of Amusia			
Scale (/30)	21.9 (6.4)	21.3 (5.8)	.72
Rhythm (/30)	22.7 (4.4)	23 (4.5)	.79
Meter (/30)	22.1 (4.3)	21.3 (5.5)	.55

[1] Unified Parkinson's Disease Rating Scale
[2] Values correspond to the time taken, in s, per item
[3] Psycholinguistic Assessments of Language Processing in Aphasia, Portuguese version

As can be seen in Table 15, the two groups scored similarly in all the neuropsychological tests, with the exception of the Stroop task. This indicates that the patients sampled here had well-preserved general cognitive abilities, probably because they were in relatively early stages of the disease. Concerning the Stroop task, patients took longer than controls to name colours without incongruity. This is likely to be a consequence of the nonspecific slowing down that is associated with the dopamine depletion in the basal ganglia (Cohen et al., 2010; Gauntlett-Gilbert & Brown, 1998).

They also had greater interference cost, which indicates executive impairment. Executive difficulties occur frequently in the early stages of PD and are due to the dysfunction in the prefrontal-striatal loop (e.g., Kehagia et al., 2010).

We assessed perceptual-cognitive musical abilities with three tests of the MBEA (Peretz et al., 2003), covering melody- and duration-related musical dimensions: Scale (melody-based), Rhythm and Meter tests (both duration-based). The Scale test assesses the tonal encoding of pitch by examining sensitivity to out-of-key tones. The participant performs same/different judgments on 30 pairs of short melodies composed according to the rules of the Western tonal system; in half of them the two melodies are identical, and in the other half the second melody differs from the first one on a single scale-violated tone. The Rhythm test inspects sensitivity to changes in duration; a similar task (same/different judgments) is used where the difference of the melodies lies in the duration of two adjacent tones, so that the rhythmic grouping by temporal proximity is changed. The Meter test inspects the perception of the hierarchical organization of musical beats by assessing the ability to distinguish duple from triple meters; participants listen to 30 stimuli, half of them in duple meter and the other half in triple meter. The task consists of categorizing each stimulus as a march or as a waltz. The performance on these tests was good, above 70% correct. Importantly, no differences were observed between patients and controls (see Table 15). Hence, PD does not seem to compromise music perception in non-emotional contexts.

EMOTIONAL STIMULI AND TASK | The music stimuli consisted of 40 short excerpts of instrumental music, which were taken from the database by Vieillard and colleagues (2008). They express sadness, fear, happiness or peacefulness (10 stimuli per emotion). The speech stimuli consisted of 40 spoken sentences with emotionally neutral semantic content, taken from the database on emotional prosody previously validated (Chapter VII). They express sadness, fear, happiness or surprise (10 stimuli per emotion). The selection of the music and speech stimuli was based on a pilot study in order to guarantee that both modalities expressed emotions with similar accuracy and intensity. This was done to increase cross-domain comparability. Fifteen individuals who did not participate in the neuropsychological study (mean age = 26.4; *SD* = 10.8; range = 19 –

51) performed a forced-choice task and provided intensity judgments (6-point scale) on a larger set of stimuli; the final set used herein was selected so that both domains had identical recognition accuracy (about 90% correct for happiness, and 80% correct for the remaining emotions) and perceived intensity (about 4 for all emotions, maximum 6). Facial expressions, which were used as a control measure of emotion recognition in the visual modality, also consisted of a set of 40 stimuli expressing sadness, fear, happiness or surprise. They were selected from the Karolinska Directed Emotional Faces database (KDEF; Lundqvist, Flykt, & Öhman, 1998) and were matched with music and speech for recognition accuracy and intensity (based on the same pilot study). Faces were colour frontal views of male and female amateur actors, who had no beards, moustaches, earrings, eyeglasses or visible make-up. Goeleven, Raedt, Leyman, and Verschuere (2008) confirmed that the faces convey the intended emotions effectively.

The task was identical across modalities. The 40 stimuli in each modality were presented four times in randomized order. On each presentation, participants rated on a 7-point scale, from 0 (*not at all*) to 6 (*very much*), how much the stimulus expressed only one emotion: sadness, fear, happiness, or peacefulness/surprise. Thus, each stimulus was rated with respect to the four possible emotions, not only the intended one. This procedure was used before in neuropsychological studies (e.g., Adolphs et al., 1998; Adolphs & Tranel, 1999; Adolphs et al., 2001; Pell & Leonard, 2003). The four emotion labels were explained and exemplified to ensure that they were adequately understood. There were six practice trials to familiarize the participants with the rating scale and the stimuli. Responses were collected using a seven-button response pad from Cedrus Corporation, model RB-730, attached to an Apple MacBook Pro computer running SuperLab version 4.0 (Abboud et al., 2006). On each trial, the participants were presented with the stimulus while the emotion and the scale to be rated appeared on the screen. Even though there was no time limit, they were encouraged to respond fast and intuitively (not after long deliberation); in the case of speech and music, they were told that they did not need to wait until the end of the stimulus to respond. After the response had been given, the experimenter advanced to the next trial by pressing the space bar. Care was taken that participants knew which emotion they were rating and that they used the scale correctly. The presentation order of the emotion recognition tasks was counterbalanced across participants. RTs were also collected to make sure that the observed results are not due to speed-accuracy trade-offs; they were measured from stimulus onset until the button press corresponding to the rating.

MEASURES AND STATISTICAL ANALYSES | Emotion recognition data were analysed according to two measures: *correlations* between participant ratings and the ratings given by a reference sample of controls; and *derived accuracy*, based on the emotion that received the highest rating. The first measure has been used by Adolphs and colleagues (e.g., Adolphs et al., 1998; Adolphs & Tranel, 1999; Adolphs et al., 2001; Adolphs et al., 1995) and Gosselin and colleagues (2005), whose procedures on the related statistical analyses we also adopted. The ratings provided by each participant (patients and controls) on the four categories for a given stimulus were first correlated with the mean ratings given for that stimulus by an independent sample of 35 healthy adults (norms; mean age = 19.6; *SD* = 3.3; range = 18 – 38)[12]. This measure is an index of the entire range of emotions perceived in each stimulus. It provides lower variance and avoids possible floor and ceiling effects (e.g., Adolphs et al., 1998; Adolphs et al., 2001; Gosselin et al., 2005). The higher the correlation, the more typical is the performance – in the sense that the rating profile is closer to the profile of the independent, normative sample. Pearson correlations were computed for each stimulus and participant. These correlations were Z-transformed, averaged across the 10 stimuli expressing a specific emotion, and inverse Z-transformed to obtain the mean correlation for that category. Then the mean correlations for each category (Z-transformed means) were submitted to standard statistical analyses, that is, the patients were compared with the healthy matched controls in ANOVAs and ANCOVAs (see below). For the second measure, accuracy was derived according to the emotion that received the highest rating (an identical procedure was used in the Chapter VI, p. 113): when the highest rating corresponded to the intended emotion of the stimulus, the response was considered correct; when it did not, the response was considered an error; when it was assigned to

[12] This sample was composed of undergraduate students who performed the same emotion recognition tasks as the participants in the neuropsychological study. As part of other experiment, they also performed pitch discrimination tasks in order to explore whether low-level auditory abilities correlate with emotion processing in speech prosody and in music; no significant correlations were found (Carvalho et al., 2011).

more than one emotion with identical magnitude, the response was considered ambivalent. This measure is an index of sensitivity to the intended emotions.

We analysed separately the results for each emotion modality by computing ANOVAs with emotion as repeated-measures factor (sadness, fear, happiness and peaceful/surprise), and group (PD and controls) as between-subjects factor. We also calculated ANCOVAs with the Stroop task as covariate (baseline and interference) in order to account for variability in processing speed and executive control. This was necessary because the patients performed worse than controls on the Stroop task (see Table 15), and performance on this test correlated with performance on emotion recognition (music: baseline, $r = -.42$, interference, $r = -.45$; speech: baseline, $r = -.54$, interference, $r = -.6$; faces: baseline, $r = -.39$, interference, $r = -.47$; all ps $< .01$). Patients also differed from controls regarding the number of women and men, but gender was not associated with emotion recognition (ps $> .12$), and so it was not included in the analyses. Main effects and interactions were followed-up in *post hoc* tests using Fisher's least-significant difference (LSD) comparisons. Pearson correlations were computed to examine associations between emotion recognition, music perception and clinical variables. Complementary analyses that we conducted to further clarify some aspects of results are detailed directly in the Results' section (see below).

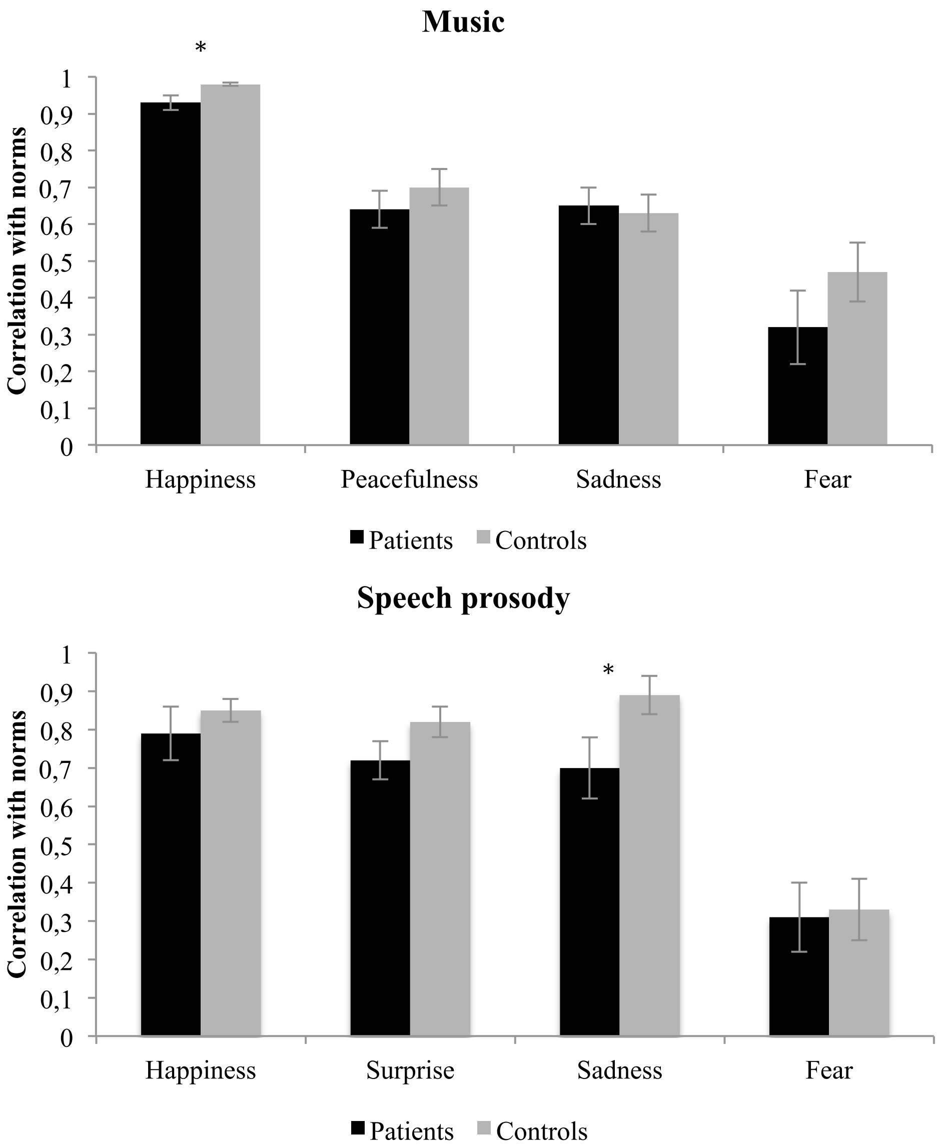
9.3. RESULTS

Emotion recognition in music and in speech prosody

MUSIC | We first focused on the correlations between participant ratings and the independent norms. These correlations are depicted for patients and matched controls in the upper panel of Figure 6. Generally, happy music elicited the highest correlation with the norms ($r = .95$, ps $< .001$), followed by peaceful (.67) and sad music (.64, $p = .4$); fear elicited the lowest correlation [.40, ps $< .01$; main effect of emotion, $F(3,141) =$

131.8, $p < .01$, $\eta_p^2 = .74$]. The patients were selectively impaired in the recognition of happiness: their correlations were significantly lower than those of matched controls in happy music only ($p < .001$), and they were similar in the other emotions [ps > .16; interaction Group x Emotion, $F(3,141) = 2.91$, $p < .05$, $\eta_p^2 = .06$]. The impairment is not reducible to the patients' executive dysfunction and general slowing, because it remained significant ($p < .01$) after controlling for these variables in an ANCOVA [interaction Group x Emotion, $F(3,135) = 3.35$, $p < .05$, $\eta_p^2 = .07$].

Figure 6. Correlations between participant and normative ratings for each emotion, in music (upper panel) and in speech prosody (lower panel), for patients and matched controls. Error bars present standard errors.



We also examined possible trade-offs between judgment performance and latencies (latencies for music are displayed in Appendix 11). Patients and controls took the same time to respond [5,653 ms and 4,891 ms, respectively; $F(1,47) = 2.78$, $p > .1$, $\eta_p^2 = .06$] and no associations were found between RTs and the magnitude of the correlations with norms (averaged across emotions, $r = -.11$, $p > .4$); this was also the case for happiness, in which patients had impaired performance ($r = -.27$, $p > .05$). Thus, the quality of the ratings was independent of the time taken to respond (consequently, it was also independent of how long participants listened to the stimuli).

We then analysed derived accuracy data. The upper part of Table 16 displays the percentage of correct categorizations in the diagonal cells (bold), and the distribution of misidentifications and ambivalent responses in rows. The percentage of categorizations was higher for the intended (correct) than for the non-intended response categories. The exception was sadness in controls, which was confused with peacefulness. The trend to confuse sadness with peacefulness has been observed in previous studies with the same stimuli (Gosselin et al., 2005), and is particularly clear in older listeners (see Chapter VI), as is the case of the participants sampled herein. Overall, happiness had the highest accuracy (88% correct, *p*s < .001), followed by peacefulness (47%, *p*s < .05); sadness (33%) and fear (31%) had similar scores [$p > .9$; main effect of emotion, $F(3,141) = 66.03$, $p < .001$, $\eta_p^2 = .58$]. Patients and controls showed similar accuracy rates for sadness and fear (*p*s > .3), but the patients were less accurate for the positive emotions: this was significant for peacefulness ($p < .05$) and marginal for happiness [$p = .06$; interaction Group x Emotion, $F(3,141) = 2.9$, $p < .05$, $\eta_p^2 = .06$][13]. After controlling for performance on the Stroop task, the impaired recognition was significant both for happiness and peacefulness [*p*s < .05; interaction Group x Emotion, $F(3,135) = 3.77$, $p < .05$, $\eta_p^2 = .08$]. Latencies did not correlate with categorization rates (averaged across emotions, $r = -.01$, $p > .9$; for happiness, $r = -.1$, $p > .4$; for peacefulness, $r = .19$, $p > .1$).

[13] Because accuracy rates for sadness and fear were relatively low (see Table 16), one cannot dismiss the possibility that the absence of differences between patients and controls for these emotions reflects a floor effect. To approach this concern, we reanalysed the data considering only the subset of excerpts that reached the highest categorization accuracies (5 for sadness, 41% correct on average, and 5 for fear, 38% correct). This analysis confirmed that the patients did not differ from controls (*p*s > .6). The fact that we also did not find differences in the correlations with norms, a measure immune to floor effects, lends further support to the notion that PD did not affect the recognition of sadness and fear in music.

We also examined possible group differences in the distribution of the non-intended categorizations (the three non-intended categories and ambivalent responses entered as repeated-measures factor in additional ANOVAs). For peacefulness, sadness and fear, the distribution of non-intended responses was similar across groups (interaction Group x Category, *p*s > .05). The groups were also similar for happy music, except that patients provided more ambivalent categorizations than controls [$p < .05$; interaction Group x Category, $F(3,141) = 6.79$, $p < .05$, $\eta_p^2 = .13$; see Table 16].

Altogether, these results reveal that PD affected the recognition of happiness and peacefulness in music, but left intact the recognition of sadness and fear. This selective impairment for the positive emotions was not secondary to the patients' executive dysfunction and general slowing. It was also not explained by trade-offs with latencies.

SPEECH PROSODY | Correlations between participant ratings and the norms for speech prosody are presented in the lower panel of Figure 6. They were higher for happiness ($r = .82$), surprise (.77) and sadness (.79, *p*s > .4), and lower for fear [.32, *p*s < .001; main effect of emotion, $F(3,141) = 41.47$, $p < .001$, $\eta_p^2 = .5$]. The patients were nominally worse than controls, but the main effect and the interaction only approached significance [main effect of group, $F(1,47) = 2.92$, $p = .09$, $\eta_p^2 = .06$; interaction Group x Emotion, $F(3,141) = 2.05$, $p = .1$, $\eta_p^2 = .04$]. On the basis of prior empirical evidence that PD might affect some emotions more than others (e.g., Dara et al., 2008; Lawrence et al., 2007; Suzuki et al., 2006), we followed-up the marginal interaction. *Post hoc* comparisons across specific emotions revealed that the patients were significantly impaired in the recognition of sadness: their correlations with norms for sad sentences were lower than controls' ($p < .01$); for the other emotions there were no significant group differences (*p*s > .4). This pattern is probably related to the executive dysfunction and general slowing of the patients, because when performance on the Stroop task entered the analysis as a covariate (both baseline and interference indexes), the main effect of group ($F < 1$, $p > .8$) and the interaction Group x Emotion ($F > 1$, $p > .7$) were far from reaching significance; the selective effect for sadness also disappeared after partialling out performance on Stroop ($p > .7$). When the ANCOVA included only the interference index of Stroop as covariate (executive control), not the baseline (processing speed), the effects also turned out non-significant (main effect of group, $F < 1$, $p > .8$; interaction Group x Emotion, $F < 1$, $p > .7$; selective difference for sadness, $p > .3$). This is evidence that cognitive control exerted a specific role for the patients' difficulties at processing prosody.

Table 16. Distribution of responses for each emotion (%, rows) in music and in speech prosody, for patients and controls. Diagonal cells in bold indicate correct identifications. Standard errors are presented in parentheses.

	Music				
	Distribution of responses				
Emotion	Happiness	Peacefulness	Sadness	Fear	Ambivalent
Patients					
Happiness	**82 (3.6) ***	5	0	0	13 *
Peacefulness	15	**39 (5) ***	14	4	28
Sadness	3	25	**37 (6.6)**	5	29
Fear	13	10	19	**30 (5.9)**	28
Controls					
Happiness	**95 (2)**	2	0	0	2
Peacefulness	18	**55 (5.9)**	9	1	18
Sadness	4	32	**29 (4.9)**	6	28
Fear	20	8	18	**31 (6.5)**	22
	Speech prosody				
Emotion	Happiness	Surprise	Sadness	Fear	Ambivalent
Patients					
Happiness	**60 (7)**	11 *	1	1	27
Surprise	23 *	**48 (6.6)**	3	1	24
Sadness	5	8	**50 (6.9)**	7	30
Fear	13	27	4	**27 (5.1)**	29
Controls					
Happiness	**48 (6.3)**	26	0	0	26
Surprise	10	**56 (5)**	4	2	28
Sadness	4	4	**63 (6.2)**	6	22
Fear	10	28	10	**.27 (4.6)**	25

* Patients were significantly different from controls

Patients had marginally longer latencies than controls [3,972 ms and 3,212 ms, respectively; $F(1,47) = 3.89$, $p = .05$, $\eta_p^2 = .08$; latencies for speech prosody are displayed in Appendix 11], but no associations were found between the time taken to respond and the magnitude of the correlations with norms (averaged across emotions, $r = -.22$, $p > .1$; for sadness, $r = -.13$, $p > 3$).

Derived accuracy rates are depicted in the lower part of Table 16. Categorizations were higher for the intended than for the non-intended emotions of the sentences, with the exception of fear, which was confused with surprise (for similar trends in misclassifications, see Chapter VII and VIII). Accuracy was higher for happiness (54% correct), surprise (52%) and sadness (56%, ps > .5) than for fear [27%, ps < .01; main effect of emotion, $F(3,141) = 13.86$, $p < .01$, $\eta_p^2 = .23$]. The patients' categorizations for sadness were nominally lower than controls' (-13% correct), but the difference was not significant ($p > .1$). There were also no effects of PD for happiness, surprise and fear [main effect of group, $F < 1$, $p > .7$; interaction Group x Emotion, $F(3,141) = 2.18$, $p = .09$, $\eta_p^2 = .04$][14]. Additionally, latencies and accuracy were not associated ($r = -.15$, $p > .2$).

The analyses on the distribution of non-intended responses revealed that patients and controls were similar for sadness and fear (interactions Group x Category ns, ps > .4). For happiness, though, patients responded with surprise less often than controls [$p < .01$; interaction Group x Category, $F(3,141) = 3.16$, $p < .05$, $\eta_p^2 = .06$; see Table 16], and for surprise they responded more often with happiness [$p < .01$; interaction Group x Category, $F(3,141) = 3.02$, $p < .05$, $\eta_p^2 = .06$]. This pattern remained unaltered after partialling out performance on the Stroop test [happiness, $p < .01$; interaction Group x Category, $F(3,135) = 5.83$, $p < .01$, $\eta_p^2 = .11$; surprise, $p < .01$; interaction Group x Category, $F(3,135) = 5.22$, $p < .05$, $\eta_p^2 = .1$]. These differences in misclassifications suggest that patients were more likely than controls to provide higher ratings for happiness (vs. surprise) whenever the sentences were not negative. To examine whether this bias influenced the effect of PD on correct categorizations, we reanalysed the data using *Hu* (Wagner, 1993); unbiased rates are displayed in Appendix 12. The ANOVA

[14] To confirm that the absence of differences between patients and controls for fear was not due to the low accuracy rates obtained in this emotion (see Table 16), we redone the ANOVA including only the subset of stimuli that reached the highest categorization accuracies (34% correct on average); no effects of PD were detected ($p > .6$).

also failed to find significant effects of group on accuracy [main effect of group, $F(1,47) = .37$, $p > .5$, $\eta_p^2 = .01$; interaction Group x Emotion, $F(3,141) = .85$, $p > .4$, $\eta_p^2 = .02$].

Overall, these results show that PD patients had only a subtle impairment on speech prosody, particularly for sadness, which appears to be a consequence of their executive dysfunction. Patients also had qualitative differences in the distribution of non-intended responses.

The pattern of results for speech prosody and music was replicated when the analyses were repeated without the four patients who were taking antidepressants. For music, patients had impaired recognition of happiness in the correlations' data [$p < .01$; interaction Group x Emotion, $F(3,129) = 2.8$, $p < .05$, $\eta_p^2 = .06$], and of happiness ($p = .07$) and peacefulness ($p < .05$) in accuracy data [marginally significant interaction Group x Emotion, $F(3,129) = 2.44$, $p = .07$, $\eta_p^2 = .05$]; this pattern remained unaltered after including the Stroop task as a covariate [correlations: happiness, $p < .05$, interaction Group x Emotion, $F(3,123) = 2.64$, $p = .05$, $\eta_p^2 = .06$; accuracy: happiness and peacefulness, *p*s < .05, interaction Group x Emotion, $F(3,123) = 2.38$, $p = .07$, $\eta_p^2 = .05$]. For prosody, patients had only impaired recognition of sadness in the correlations' data [$p < .05$; interaction Group x Emotion ns, $F(3,129) = 1.24$, $p > .3$, $\eta_p^2 = .03$], which turned out to be non-significant after controlling for executive functioning [$p > .1$; interaction Group x Emotion non-significant, $F(3,123) = .11$, $p > .9$, $\eta_p^2 = .0$]; in accuracy data, there no differences between patients and controls (*p*s > .1).

Summing up, the recognition of positive emotions was impaired in music but not in speech, and the recognition of sadness was impaired in speech but not in music. To provide direct statistical support for this cross-domain dissociation, we carried out an additional ANOVA including domain as repeated-measures factor; this analysis was computed on the correlations' data for sadness, fear and happiness (the remaining emotion category was not included because it was not identical across modalities, peacefulness-music/surprise-speech). The interaction Modality x Group x Emotion was significant [$F(2,94) = 5.05$, $p < .01$, $\eta_p^2 = .1$]: *post hoc* comparisons confirmed that, for happiness, PD patients were impaired in music ($p < .05$) but were normal in speech ($p > .6$), whereas for sadness, they were impaired in speech ($p < .01$) but were normal in music ($p > .8$); for fear, they were normal both in music and speech (*p*s > .3).

Emotion recognition in facial expressions

Figure 7 presents the correlations between participant and normative ratings for facial expressions. They were highest for happiness ($r = .82$, *p*s < .001), intermediate for sadness (.79) and surprise (.77, $p > .07$), and lowest for fear [.32, *p*s < .001; main effect of emotion, $F(3,141) = 331.68$, $p < .001$, $\eta_p^2 = .88$]. Differences between patients and controls did not to reach statistical significance in either analyses: in the ANOVA, for the main effect of group, $F(1,47) = 3.9$, $p = .06$, $\eta_p^2 = .08$; interaction Group x Emotion, $F < 1$, $p > .6$; in the ANCOVA including performance on the Stroop task, for the main effect of group, $F < 1$, $p > .54$; interaction Group x Emotion, $F < 1$, $p > .59$]. Patients and controls had similar latencies [3,381 ms and 2,849 ms, respectively, $F(1,47) = 1.95$, $p > .1$, $\eta_p^2 = .04$; latencies for faces are displayed in Appendix 11]; they were not associated with the magnitude of the correlations with norms ($r = -.17$, $p > .2$).

Figure 7. Correlations between participant and normative ratings for each emotion in facial expressions, for patients and controls. Error bars present standard errors.

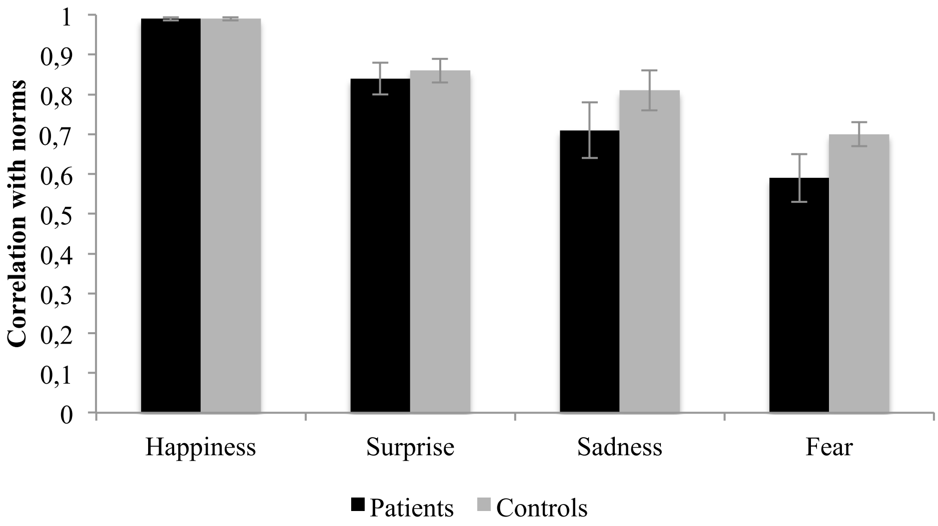


Table 17 presents the percentage of correct categorizations in diagonal cells (bold), and the distribution of misidentifications and ambivalent responses in rows. As in the correlations' data, accuracy was highest for happiness, (92% correct, *p*s $< .001$), intermediate for sadness (61%) and surprise (65%, $p > .33$), and lowest for fear [35%, *p*s $< .001$; main effect of emotion, $F(3,141) = 50.91$, $p < .001$, $\eta_p^2 = .52$]. Patients did not differ from controls [main effect of group, $F(1,47) = 2.53$, $p > .1$, $\eta_p^2 = .05$; interaction Group x Emotion ns, $F < 1$, $p > .4$]. RTs did not correlated with accuracy ($r = -.09$, $p > .5$). The distribution of the non-intended responses was similar across groups for all emotions (interactions Group x Category ns, *p*s > .2).

These results show that our sample of PD patients had normal recognition of emotions in facial expressions.

Table 17. Distribution of responses for each emotion (%, rows) in facial expressions (rows), for patients and controls. Diagonal cells in bold indicate correct identifications. Standard errors are given in parentheses.

	Distribution of responses				
Emotion	Happiness	Surprise	Sadness	Fear	Ambivalent
Patients					
Happiness	**92 (3)**	4	0	0	4
Surprise	3	**62 (5.7)**	2	10	23
Sadness	3	11	**58 (6.2)**	8	21
Fear	3	25	13	**27 (4.6)**	32
Controls					
Happiness	**92 (3.2)**	2	0	0	6
Surprise	2	**68 (5.8)**	4	5	20
Sadness	1	8	**64 (6.6)**	8	21
Fear	0	18	13	**43 (4.5)**	24

CORRELATIONS BETWEEN EMOTION RECOGNITION, MUSIC PERCEPTION AND CLINICAL VARIABLES | We explored possible associations between the three emotion recognition modalities: music, speech and facial expressions. We also examined whether emotion recognition in the two experimental modalities – music and speech – correlate with music perceptual-cognitive abilities, with the magnitude of participants' depression symptoms, and with clinical variables, notably disease duration and severity. These analyses were based on the z-transformed correlations with norms, averaged across the four emotions for each modality. Table 18 presents the correlation coefficients, as well as the partial coefficients that were calculated taking into account demographic and cognitive variables. Partial correlations were computed to guarantee that the detected associations are not artefacts of variability in domain-general cognitive abilities and demographic characteristics.

A correlation was found between emotion recognition in music and speech, even when demographic and cognitive variables were partialled out[15]. This indicates that the two modalities engage common processes to a significant extent. Emotion recognition in facial expressions was also correlated with both auditory modalities, a result that may reflect the operation of modality-independent levels of emotion processing (Peelen et al., 2010).

After controlling for demographic and cognitive variables, performance on the music perceptual-cognitive tests (MBEA) did not correlate with emotion recognition in music, with the exception of the Scale test. Further analyses conducted for specific emotions revealed that the Scale test correlated only with the recognition of fear ($r = .37$, $p < .05$), not with the other emotions (ps > .1). This suggests that sensitivity to scale-violated tones is selectively important to perceive fear in music, probably because out-of-key notes were a salient feature of scary music (Vieillard et al., 2008). Emotion recognition in music and speech was not significantly associated with depression symptoms, severity of the disease (as assessed by the Hoehan and Yahr stage and by the motor UPDRS), age of onset, and disease duration (see Table 18).

[15] We also examined associations between music and speech separately for the three specific emotions which were common across domains. Correlations were significant for sadness ($r = .31$, $p < .05$) and fear ($r = .28$, $p = .05$), and marginally significant for happiness ($r = .23$, $p = .1$).

Table 18. Correlations between emotion recognition, music perception and clinical variables of PD patients.

	Emotion recognition			
	Music		Speech Prosody	
Variable	*r*	Partial *r*[1]	*r*	Partial *r*
Emotion recognition in speech prosody	.61*	.37*	-	-
Emotion recognition in faces	.67*	.51*	.63*	.4*
Montreal Battery of Evaluation of Amusia				
Scale	.61*	.36*	.42*	.08
Rhythm	.54*	.28	.33*	.07
Meter	.27	.08	.19	.01
Geriatric Depression Scale	-.09	-.11	-.11	-.11
Hoehn and Yahr stage	.04	.17	-.22	-.13
Motor UPDRS[2]	-.12	-.1	-.41*	-.45
Disease duration (years)	-.34	-.34	-.49*	-.49
Age of disease onset (years)	-.51*	-.34	-.29	-.49

* $p < .05$
[1] Controlled for demographic (age, education and gender) and cognitive (MMSE, Vocabulary, Digit Span, Verbal Fluency and Stroop test) characteristics
[2] Unified Parkinson's Disease Rating Scale

9.4. DISCUSSION

We showed that PD affects differently emotion recognition in music and in speech prosody. Patients were impaired in the recognition of positive emotions in music, namely happiness and peacefulness, and normal in the recognition of the negative emotions of sadness and fear. For speech prosody, they were impaired in the recognition of sadness, and normal in the recognition of fear, happiness and surprise. Impairments for music were not explained by low-level music processing problems, depression symptoms and domain-general cognitive difficulties, and so they appear to be a primary feature of PD. For speech, the impairments were related to the patients' executive dysfunction. To our knowledge, this is the first report of a neuropsychological dissociation between emotion recognition in music and speech prosody. It substantiates that the neurocognitive mechanisms involved in the two auditory emotion domains are not fully identical. We also found a dissociation between (normal) music perceptual-cognitive abilities and (impaired) emotion recognition in music, an observation suggesting that these two components of music rely on distinct neural systems.

It has been held that emotion processing in speech and music engages common neurocognitive mechanisms (Juslin et al., 2010; Juslin & Västfjäll, 2008; Patel, 2008b; Peretz, 2010). Previous studies gave some support in favour of this hypothesis: the acoustic encoding of specific emotions has similarities across domains (Bowling et al., 2010; Curtis & Bharucha, 2010; Juslin & Laukka, 2003), musical expertise modulates the recognition of emotional prosody (Chapter VIII; Thompson et al., 2004), and the processing of pitch, a low-level emotion cue in both domains, depends on shared mechanisms (e.g., Liu et al., 2010; Nan et al., 2010). Our findings add to these by showing that emotion recognition in music and speech correlate, even after controlling for cognitive and demographic variables. However, perceiving emotions is a higher order process that comprises multiple cognitive operations and neural systems (e.g., Juslin et al., 2010; Koelsch, 2010; Leitman, Wolf, et al., 2010; Schirmer & Kotz, 2006). So far, the extent to which all or only part of them is similarly engaged by both domains was unspecified. Finding that impairments in music and speech are dissociated in PD is clear evidence that the overlap is partial – there are differences between domains. Importantly, these differences do not seem to be merely related to general emotion processes, as patients had normal processing of emotions in the visual modality (facial expressions). The claim that mechanisms are shared only partly is consistent with Ilie and Thompson's (2006) finding that analogous manipulations of acoustic parameters elicits both similar and opposing effects in affective ratings for music and speech. The hallmark neuropathological changes in early PD are in basal ganglia circuitries (e.g., Obeso et al., 2008), and so they are a likely locus for the dissociation between domains. The basal ganglia receive various cortical inputs and have a broad influence on several neuronal pathways (e.g., Utter & Basso, 2008). They might be engaged differently by emotions in music and in speech, as we discuss below.

We observed for the first time that PD has a detrimental impact on the recognition of happiness and peacefulness in instrumental music. This impairment does not seem to be a consequence of low-level processing abnormalities because the patients performed as good as controls in well validated tests of music perceptual-cognitive abilities. It is not reducible to the patients' general executive difficulties, and cannot be accounted for by familiarity effects, as the excerpts were unfamiliar to all participants (they were composed for experimental purposes, not taken from the classic repertoire). Thus, the integrity of the basal ganglia pathways may play a critical and direct role for the normal recognition of positive emotion categories in music. PD can affect other brain systems during the course of the disease, such as the brainstem, thalamus, entorhinal cortex, cingulate cortex, cortical association areas and the amygdala (e.g., Braak, Ghebremedhin, Rüb, Bratzke, & Tredici, 2004; Braak et al., 2003), and these neuropathological changes correlate with cognitive impairment (e.g., Kövari et al., 2003). Notwithstanding, they are likely to have played a minimal role here because the sampled patients were in relatively early stages of the disease, performed normally in most of the domain-general cognitive measures, and were non-depressed. Additionally, extra-basal ganglia pathology, namely in the amygdala, would produce a pattern of results different from ours. Neuropsychological studies on patients with amygdala damage, using the same music stimuli as we used here, found impaired recognition of fear and sadness, with relatively spared recognition of happiness and peacefulness (Gosselin et al., 2005, 2007, 2011). Hence, whilst our findings suggest the involvement of the basal ganglia in the recognition of positive musical emotions, Gosselin et al.'s suggest that the involvement of the amygdala in negative musical emotions. This pattern is coherent with fMRI evidence. Koelsch and colleagues (2006) showed that pleasant consonant music, consisting of joyful instrumental dance-tunes, activates the ventral striatum (among other structures), and that unpleasant music (permanently dissonant) activates the amygdala. Mitterschiffthaler, Fu, Dalton, Andrew, and Williams (2007) reported activations in the ventral and dorsal striatum for happy music, and in the amygdala for sad music. In their study on how PD affects the recognition of musical emotions, Tricht and colleagues (2010) found that responses were normal for happy music. However, this was probably due to a ceiling effect, as all but two participants obtained the maximum score for this emotion. On the other hand, they found impaired recognition of fear, while we did not. A straightforward explanation for the discrepancy is lacking, but it is possible that the sample of patients in their study had more extra-basal ganglia neuropathology than ours, including in the amygdala, what would explain the difficulties for fear. Indeed, their patients had longer disease duration than ours (12 vs. 8 years), and they also had more severe motor symptoms, as indicated by the higher score in the motor scale of the UPDRS (26 vs. 18).

Reward and motor functions of the basal ganglia might underlie the selective impairment for positive musical emotions. Human and non-human studies link dopaminergic activity in the basal ganglia with reward-related responses (e.g., Graybiel, 2005; Nambu, 2008), and these are altered in PD (e.g., Schott et al., 2007), both in patients undertaking dopaminergic replacement therapy and in those not medicated (Rowe et al., 2008). Patients might have experienced abnormal reward responses to the positive music stimuli due to their dopaminergic dysfunction. Recent evidence that the activation of the striatum by pleasurable music reflects the modulation of dopaminergic activity lends support to this hypothesis (Koelsch, 2010; Salimpoor et al., 2011). Testing patients *on* and *off* dopaminergic therapy will shed light on the role of this neurotransmitter for the impairment at recognizing musical emotions in PD. The basal ganglia, particularly the striatum and substantia nigra, are also involved in the interface between motivational responses and motor functions (Mitterschiffthaler et al., 2007; Utter & Basso, 2008). It has been suggested that these structures support the activation of (pre)motor representations during the perception of pleasant auditory stimuli (Koelsch et al., 2006). Therefore, the dysfunction of these reward and motor circuitries might culminate in altered outputs for the cortical regions where higher order evaluative and integrative processes take place, thus leading to the observed behavioural impairments. The reliance of emotion recognition on reward and motor processes is consistent with simulation and embodied models. These models emphasize that perceiving emotions is inextricably linked with the somato-motor processes that are involved in the experience of the corresponding emotion (e.g., Adolphs, 2006, 2010; Barrett, 2009).

Regarding speech prosody, the patients tested here were only slightly impaired, particularly for sadness. Several prior studies have found broader and more pronounced defects (e.g., Dara et al., 2008; Gray & Tickle-Degnen, 2010; Paulmann & Pell, 2010b; Pell & Leonard, 2003; Yip et al., 2003), but others have found no defects at all

(Caekebeke, Jennekens-Schinkel, Linden, Buruma, & Roos, 1991; Kan, Kawamura, Hasegawa, Mochizuki, & Nakamura, 2002; Mitchell & Bouças, 2009). Importantly, when our participants' executive abilities were controlled for (Stroop task), the prosodic impairment was no longer significant. Thus, the patients' general executive problems, which probably reflect the dysfunction in the prefrontal-striatal loop (e.g., Kehagia et al., 2010), played an important role in the difficulties in processing prosody. This finding adds to different lines of evidence suggesting that the basal ganglia are chiefly involved in the later stages of emotion recognition in prosody, where executive processes are preponderant. First, other researchers have also observed that in PD executive function correlates with the recognition of emotional prosody (Breitenstein, Lancker, Daum, et al., 2001; Pell & Leonard, 2003). Second, it was suggested that some subgroups of PD patients show normal processing of prosody because the impairment is detected only when the executive dysfunction and working memory impairments are prominent (Benke et al., 1998; Caekebeke et al., 1991; Mitchell & Bouças, 2009). Third, in ERP experiments, Paulmann and colleagues (2008a, 2011) found that patients with focal lesions in the basal ganglia, despite being impaired for explicit emotion recognition in prosody, have normal brain responses in the P200 component, which reflects emotional salience detection, in a following negative-going brain wave, associated with meaning-related processes, as well as in a right lateralized positive component that indexes emotional prosody expectancy violations. This pattern suggests that the basal ganglia are not critical during early online prosodic processes, but only during explicit offline judgments, which rely greatly on executive and working memory operations. In a related vein, a fMRI study by Bach and colleagues (2008) revealed that the basal ganglia activations in response to prosody are stronger under explicit (forced-choice emotion categorization) vs. implicit (gender judgment) task instructions. Note that, in the measure of derived accuracy, the patients had also some differences in the patterns of misclassifications: they categorized happy sentences as surprised less often than controls, and surprised sentences as happy more often than controls. These differences remained unaltered when the patients' executive dysfunction was controlled for, and so they could point to a more specific origin for the prosodic difficulties in PD. However, we think that such an inference cannot be made with certainty from the current results because we were unable to detect any effects for these emotions, neither in the intended accuracy rates after correcting for response bias, nor in the correlations with norms, which are a sensitive index of the entire range of emotions perceived in a given stimulus. Future studies manipulating the executive demands of the task will better specify whether the role of basal ganglia for emotional prosody is exclusively attentional/executive, or whether it is also a low-level and early one, more directly related to prosody processing.

Why were positive emotions impaired in music but not in speech prosody? A possible answer lies in the functional specificities of both domains. We listen to music typically to modulate emotional states and because it is a rewarding experience. Music is most adept in communicating and inducing positive emotions (e.g., Zentner et al., 2008), and it can produce intensely pleasurable responses (e.g., Blood & Zatorre, 2001). Thus, music might engage reward and emotion systems more readily than speech prosody does. This is in accordance with Ilie and Thompson's (2006) observation that listeners provide higher ratings of valence and energetic arousal for music than for speech. Furthermore, music may be more effective than speech in recruiting motor systems. Brain motor regions are activated even during passive music listening (Chen, Penhune, & Zatorre, 2008). Children as young as 5-24 months engage in rhythmic movement in response to music more than in response to speech (Zentner & Eerola, 2010). Moreover, music is frequently used for movement coordination and dancing in social contexts (e.g., Koelsch, 2010). On the basis of these findings, we can hypothesize that the integrity of emotion, reward and motor systems is relatively more important for emotion recognition in music than in speech. Speech prosody, on the other hand, is used primarily for communicative purposes. In the natural communicative situations encountered during daily life, we integrate prosodic information with co-occurring processes related to lexico-semantic information in order to make sense out of the interlocutor's spoken utterances (Schirmer & Kotz, 2006). Hence, processing prosody usually requires integrative and selection processes to a greater extent than music does, that is, executive functions could be more preponderant for prosody than for music. The debate on the parallels between both domains will benefit from accounting for these functional specificities to better explain the engagement of both common and specific neurocognitive mechanisms.

A drawback of this study is that, although three of the emotion categories were identical for music and speech prosody (sadness, fear and happiness), the other one was not: peacefulness was examined in music, and surprise in speech (peacefulness is frequently reported in response to music but not to speech prosody, and surprise is frequently reported in response to speech prosody but not to music; Zentner et al., 2008). Apart from happiness, it was difficult to find a non-negative emotion category that can be similarly well expressed in both domains, and our priority was to keep a balanced number of negative and positive emotions in order to better ascertain effects of valence. Further work will need to include a greater number of equivalent emotions across domains. Notwithstanding, this limitation does not compromise our conclusion that PD affects differently music and prosody: we observed cross-domain dissociations in two of the emotion categories which were identical, happiness and sadness. Two other interrelated issues regarding comparability across domains should be discussed. One is that stimuli were longer in music than in speech, and this could limit direct comparisons (it would be difficult to tease apart domain-related from duration-related effects). However, care was taken to guarantee that music and speech stimuli were comparable in perceptual/functional dimensions, which is perhaps more important than equivalence in low-level physical features: the stimuli were pretested and selected to express emotions with similar recognition accuracy and intensity in music and speech, as well as in facial expressions. Furthermore, we did not find any association between how long participants listened to the stimuli and their performance on emotion recognition (see correlations between latencies and emotion recognition measures in the Results' section). That is, the stimulus' duration did not appear to impact on performance. The other issue concerns the possibility that cross-domain dissociations are an artefact of differential task difficulty (e.g., participants might have found a certain emotion more difficult to recognize in one domain than in the other). This alternative hypothesis is not compatible with our results, though: happiness was better recognized in music than in speech, and it was precisely in music that patients were impaired (not in speech); and sadness was better recognized in speech than in music, and it was in speech that patients were impaired. Therefore, we consider that the results reported here reflect true differences in how emotions are perceived in music and speech.

The present study also revealed that music perception is dissociated from emotion recognition in music. Patients were as good as controls in detecting out-of-key tones, rhythmic changes, and differences in meter, as assessed by three tests of the MBEA. This observation indicates that PD does not compromise these abilities. Yet, patients had impaired recognition of musical emotions, as discussed above. These results extend those by Gosselin and colleagues (2005), who reported that patients with temporal medial resection have no problem in detecting timing errors in music, but are impaired in the perception of emotions in music. A dissociation in the opposite direction was found by Peretz and colleagues (1998): the brain-damaged patient I.R. had severe impairments in music perception and relatively spared emotional judgments (see also Omar, Hailstone, Warren, Crutch, & Warren, 2010). Such a double dissociation is evidence that the neurocognitive resources allocated for music perception are distinct from those allocated for musical emotions. Additionally, we observed that performance in music perception tests did not correlate with emotion recognition in music. The exception was the Scale test (task: detect out-of-key tones), which correlated with the recognition of fear. This suggests that achieving better tonal encoding of pitch facilitates emotion recognition when out-of-key tones are an important emotion cue, as was the case of scary excerpts (Vieillard et al., 2008). Hence, albeit music perception and emotion recognition appear to be neurally autonomous systems, they might interact in some situations. This interaction might contribute to explain why individuals with amusia have abnormal emotional responses to music (McDonald & Stewart, 2008).

9.5. CONCLUSION

In this study we uncovered a neuropsychological dissociation between emotion recognition in music and in speech prosody in PD. Patients had impaired recognition of positive emotions in music, but not in speech. Basal ganglia circuitries, namely those related to reward/dopamine and motor functions, might be critical for the recognition of these musical emotions. In contrast, the recognition of sadness was impaired in speech, but normal in music. This impairment could be attributed to general executive difficulties, which are probably due to a dysfunction in the prefrontal-striatal loop. From a clinical perspective, the present findings corroborate and add to previous studies showing that even cognitively well-preserved and non-depressed PD patients might experience difficulties in perceiving emotions, particularly in the auditory domain. While these difficulties may go unnoticed by clinicians and familiars in such independent patients, it is possible that they contribute to interpersonal distress in communicative settings, reducing the quality of social interactions. From a basic science perspective, the dissociation between speech and music provides evidence that the neurocognitive mechanisms subserving emotion recognition in both domains do not fully overlap – specific mechanisms can be engaged.

CHAPTER X

General Conclusions

We began this thesis with the idea that making sense out of sound is a defining feature of human nature. The sounds of voice and music are pervasive in our auditory environment during all our life, and we respond to them since very early stages of development, possibly even before birth (e.g., Belin et al., 2004; Zentner & Eerola, 2010). They are complex, multidimensional and powerful communication tools. The overarching goal of the empirical work presented herein was to investigate how emotions are perceived in speech prosody and in music. We set out to examine the impact of ageing and musical training in emotion recognition in music, how emotions are recognized in speech prosody in Portuguese, and the extent to which common neurocognitive mechanisms are put into action by emotional speech and musical emotions. To that end, we conducted a series of five studies using behavioural and neuropsychological methodologies. These studies relate to broader issues regarding emotional communication through acoustic signals, individual differences in emotion processing, neurocognitive ageing, neural plasticity, modularity of music and language as neurocognitive systems, and clinical phenomenology of a neurodegenerative disorder, PD. The main findings are outlined below. Future research directions are also delineated.

10.1. SUMMARIZING

Do ageing and musical expertise modulate emotion recognition in music?

A consistent observation in the literature is that ageing produces changes in the recognition of emotion stimuli, with older adults being less accurate than younger ones for some emotion categories (e.g., Ruffman et al., 2008). That musical training drives changes in neurocognitive functioning via neural plasticity is also well established (e.g., Habib & Besson, 2009). In two studies, we examined whether these two experiential factors, ageing and musical expertise, influence emotion recognition in music. The emotion categories investigated were happiness, peacefulness, sadness and fear. They were communicated through musical structure features (e.g., mode, dissonance, pitch range) in short instrumental excerpts (Vieillard et al., 2008). An emotion rating task was employed, in which participants judged how much each excerpt expressed the four emotion categories. In the first study, we tested participants across the full range of the adult life span: younger, middle-aged and older adults. Advancing age was associated with decreased sensitivity to sadness and fear, and sensitivity to happiness and peacefulness remained relatively intact. This pattern of age-related effects – changes for negative but not for positive emotions – was significant already in middle-age, around the forties, not only later in life. Furthermore, in the subset of participants with musical training, we found that the number of years of training was associated with better accuracy in recognizing musical emotions. In the second study, we followed-up this effect of training by directly comparing musicians and untrained listeners. Additionally, we aimed at determining whether ageing effects on musical emotions are primary in origin or a by-product of general cognitive decline. Participants varying orthogonally in training (musicians and controls) and age (young and middle-age) were tested, and they were also examined for general cognitive abilities. The pattern of ageing effects was replicated, as was the correlation between years of training and enhanced sensitivity to musical emotions. As a group, musicians outperformed untrained controls, but this was significant only for the middle-aged participants. Both ageing and expertise effects were independent of differences in general cognitive abilities. Thus, these experiential factors seem to exert a direct and specific impact on musical emotions. In both studies, we explored associations between the excerpts' structural cues and the participants' subjective ratings. Responses for the four emotions could be significantly predicted by some combination of musical cues. Ageing and expertise effects may be linked with quantitative and qualitative differences in how listeners rely on the musical cues to respond.

The effect of ageing on musical emotions uncovered here might be accounted for by neuropsychological deterioration in neural structures of emotion (e.g., Cacioppo et al., 2011; Ruffman et al., 2008), as well as by motivational changes related to adult development – older participants could be more able than younger ones to implement top-down regulatory strategies to attenuate the processing of negative information and promote positive information (e.g., Samanez-Larkin & Carstensen, 2011; Williams et al., 2006). Finding that emotion recognition in music undergoes developmental changes adds to evidence that musical emotions share several of the properties of stimuli important for biological survival and social functioning. The fact that the effect was significant in middle-age extends results from recent studies showing that modulations in emotion processing start early in adult life (e.g., Mill et al., 2009; Paulmann et al., 2008b). The observed effect of expertise indicates that the neurocognitive mechanisms underlying musical emotions change plastically in response to musical training. This contributes to the literature on the impact of musical experience on music processing. More generally, in these two studies we showed that ageing and musical training are two sources of individual variation in how listeners perceive emotions communicated by music.

How are emotions recognized in emotional prosody in Portuguese?

Although the recognition of emotions in speech prosody is governed by universal principles to a significant extent, it has been shown that language-specific factors also play a role (e.g., Pell, Monetta, et al., 2009; Scherer et al., 2001; Thompson & Balkwill, 2006). It is thus important to devise prosodic stimuli and to conduct perceptual studies for different languages. In this thesis we developed and validated a set of Portuguese stimuli for research on emotional prosody, which is now available to the research community. The standard content paradigm was used, in that the same verbal materials were recorded to express different emotions through prosodic cues alone. Verbal materials consisted of a set of 16 short sentences with emotionally neutral semantic content, and of 16 pseudo-sentences derived from the sentences. They were recorded by two female native speakers of European Portuguese so as to express neutrality, happiness, sadness, anger, fear, disgust and surprise. Each stimulus was evaluated in a perceptual experiment by 20 listeners, who performed a forced-choice emotion recognition task and provided intensity judgments. The stimuli that reached accuracy rates above three times the level of chance (circa 14%) were included in the database. The final database contains 190 sentences and 178 pseudo-sentences. Acoustic and perceptual characteristics are available for each stimulus. We observed that the acoustic profiles of each emotion, namely differences in F_0, F_0 variability and duration, are in general agreement with previous research in other languages (e.g., Banse & Scherer, 1996; Juslin & Laukka, 2001, 2003). Furthermore, in a discriminant analysis, these acoustic features provided sufficient information to discriminate between emotion categories with an accuracy of 60%. Perceptual data showed that emotions are recognized with similarly high accuracy in sentences, 75% correct, and in pseudo-sentences, 71%. The time needed to identify emotions, though, was longer in sentences than in pseudo-sentences, a processing cost presumably reflecting lexico-semantic and integrative operations, which were residual in pseudo-sentences. Intensity judgements were also similarly high in both types of materials, indicating that they both were clear exemplars of the intended emotions.

Therefore, the database provided herein contains a set of prosodic stimuli which are effective at communicating neutrality and six emotion categories. It adds to the bulk of studies that have been conducted to devise materials for different languages (e.g., Burkhardt et al., 2005; Makarova & Petrushin, 2002; Pell, 2002; Ross et al., 1997; Staroniewicz & Majewski, 2009; Wu et al., 2006). In addition to its potential value for basic research on emotional prosody, these materials may also be useful for the assessment of pragmatic skills in clinical settings.

Does emotion recognition in speech prosody and in music recruit shared neurocognitive mechanisms?

The relations between language and music as neurocognitive and communicative systems has attracted increased attention in cognitive neuroscience (e.g., Besson et al., 2011; Patel, 2008a, 2011; Schön & Besson, 2001). A highly debated hypothesis is whether processing emotions in both domains engages common neurocognitive mechanisms (e.g., Juslin & Laukka, 2003; Juslin et al., 2010; Juslin & Västfjäll, 2008; Peretz, 2010). One of the main goals of this thesis was to investigate this hypothesis. To that end, we conducted two studies. In one of them, we examined cross-domain transfer effects from musical expertise to emotional speech prosody. In the other one, a comparative design was adopted, in which we examined whether PD impacts emotion recognition in speech and in music in a similar manner.

If emotion recognition in speech and music engages common mechanisms, training in one domain (music) can be expected to produce benefits in the other domain (speech) via neural plasticity. We compared highly trained musicians and untrained listeners from two age cohorts (young and middle-age) in the recognition of neutrality and six emotion categories in speech prosody. A forced-choice task was employed, with intensity judgments and latencies also being collected. Musicians in both age groups were more accurate than controls across all emotions. Their enhanced performance cannot be accounted for by potential confounds, namely general cognitive differences, trade-offs with latencies, as well as by differences in intensity ratings, personality traits, or in socio-educational background. Furthermore, musicians and controls were similar in how they relied upon the stimuli's acoustic features to respond, a fact that suggests that differences in accuracy are not due to the musicians' advantage in low-level speech processing. This robust finding contributes to clarify previous results on the effects of musical training on emotional prosody, which were inconclusive (Thompson et al., 2004; Trimmer & Cuddy, 2008). It supports the hypothesis that mechanisms are at least partly shared across domains. Our study also highlights that musical training is a source of individual differences in a pragmatic skill important for interpersonal functioning – the processing of emotions in speech prosody. From a clinical standpoint, this raises the promising possibility that music-therapy may be a valuable tool to bolster speech and communicative functions in patients with prosodic impairments.

Another strategy to examine the extent to which mechanisms are shared for emotion processing in speech and music consists of investigating a neuropathological condition that affects one domain and determine whether it affects the other domain similarly (Patel, 2008b). We used this approach by examining patients with PD. Several studies have associated the basal ganglia dysfunction in PD with difficulties in the recognition of emotional speech prosody (e.g., Gray & Tickle-Degnen, 2010; Pell & Leonard, 2003), but only one study has explored how this disease impacts the recognition of musical emotions (Tricht et al., 2010). Critically, no direct cross-domain comparisons were conducted so far. We determined the impact of PD in the recognition of four emotions in speech and music using emotion rating tasks. The pattern of impairments was distinct in both domains. This is evidence that they are partly independent in some way. The recognition of positive emotions was impaired in music, but normal in speech. The impairment for music presumably reflects the patient's dysfunction in the basal ganglia pathways, namely in those related to reward and motor functions. The recognition of sadness was impaired in speech, but normal in music. The patients' impairments for speech prosody were associated with their executive dysfunction, suggesting that the basal ganglia may contribute chiefly for later stages of prosodic processing, when controlled and executive processes are preponderant. This is the first report of a neuropsychological dissociation between emotion recognition in speech prosody and music. It is in line with previous findings showing that similar acoustical manipulations in speech and music produce common but also domain-specific effects in affective ratings (Ilie & Thompson, 2006). It adds to a previous report of a dissociation between emotion recognition in music and in nonverbal vocal expressions in frontotemporal lobar degeneration (Omar et al., 2011). Thus, both types of vocal emotion cues, speech prosody and nonverbal expressions, may be partly independent of musical emotions. We also uncovered a dissociation between music perception abilities, which were normal, and emotion recognition in music, which was impaired. In close convergence with previous studies, this indicates that perceptual and emotional musical functions are subserved by independent neural systems (Gosselin et al., 2005; Peretz et al., 1998). Additionally, our study contributes to the knowledge on the clinical phenomenology of PD. It shows that, even non-depressed and cognitively well-preserved patients in relatively early stages of the disease, may show specific difficulties in executive functions and in the recognition of emotional expressions in auditory signals.

The pattern that emerges from these two studies is that emotion recognition in speech prosody and in music engages a combination of domain-general and domain-specific mechanisms. This highlights the notion that both domains comprise complex constellations of subcomponents, with some of them being shared, and others being dependent on the domain specificities. Future conceptualizations and research on the music-language relations will thus need to focus on the commonalities across domains, but also to account for their differences.

10.2. FUTURE DIRECTIONS

Science is fascinating because it provides answers, but also because each provisory answer raises a multitude of new questions. As such, the studies presented in this thesis can be seen as stimulating starting points for new avenues of research. For instance, we observed that ageing is associated with decreased sensitivity to negative musical emotions, but positive ones remain relatively stable. Is this *positivity effect* a modality-independent general property of emotional functioning as we get older, or is it specific to certain modalities? A trend towards positivity has been found in ageing studies on memory and attention processes (Mather & Carstensen, 2005; Mather & Knight, 2005), emotional experience (Carstensen, Pasupathi, Mayr, & Nesselroade, 2000), and on the recognition of facial expressions (Williams et al., 2006), speech prosody (Mitchell et al., 2011) and emotion pictures (Kisley et al., 2007). However, it is far from being established. Many studies have found age-related decline in the recognition of positive emotions, and stability for negative ones; this was found for facial expressions (Calder et al., 2003; Ruffman et al., 2008), prosody (Paulmann et al., 2008b) and lexical stimuli (Isaacowitz et al., 2007). Future studies on different emotion modalities will yield interesting insights on this issue. The inclusion of a balanced number of positive and negative emotion categories, and a match between them for arousal properties, are methodological requirements that should be met. A related question concerns the stage of processing at which ageing operates changes. Is it in early and automatic stages of emotion processing, or rather in later controlled ones? This cannot be answered with offline emotion recognition tasks, as the ones that predominate in the literature and that we used here. Behavioural studies that examine emotion recognition under different levels of cognitive load (e.g., dual task) will be valuable to determine the extent to which the effects of ageing are modulated by the degree of availability of controlled cognitive resources. Neuroimaging studies exploring emotion processing at different appraisal levels (e.g., implicit vs. explicit emotion processing tasks) may be illuminating too. It also remains to be examined at which stage of processing musical expertise modulates emotion recognition.

With regard to the music-language relations for emotion processing, our results suggest that it might be time to start looking beyond the general, though important, question of whether mechanisms are shared or distinct, and focus on more specific and nuanced aspects of the neurocognitive architecture and dynamics that underlie emotions across domains. For instance, does music engage reward and emotion systems more readily than speech does, as we discussed about in the study on PD patients? Do we allocate different attentional strategies when appreciating emotions in speech and in music, as Ilie and Thompson (2006) speculated to explain why similar acoustic manipulations may lead to different consequences in affective ratings? Is the conceptual system for emotions shared across domains? Another question concerns the fact that most comparative research so far has focused on the relation between music and speech prosody (e.g., Bowling et al., 2010; Curtis & Bharucha, 2010; Ilie & Thompson, 2006; Juslin & Laukka, 2003; Thompson et al., 2004). This leaves other vocal cues of emotion unexplored, namely nonverbal vocal expressions, such as laughter, sighs, crying and screams. Are there acoustic similarities for emotions in music and in nonverbal expressions? Does musical training hone emotion recognition in nonverbal expressions, as it does for speech prosody? To which extent the neurocognitive mechanisms that support these expressions are the same that respond to music (Omar et al., 2011)? The proper acoustic and neurocognitive relations between speech prosody and nonverbal expressions remain to be explored (Sauter, Eisner, Calder, et al., 2010).

10.3. CONCLUDING

In a series of empirical studies, we (1) found that ageing and musical training modulate emotion recognition in music, (2) developed and validated a database of speech stimuli in Portuguese for research on emotional prosody, and (3) showed that emotion processing in speech prosody and music depends on a combination of shared and domain-specific neurocognitive mechanisms. We hope that this contribution paves the way for further basic and applied research on how our minds and brains make emotional meaning out of sound.

In the end of this thesis – and in the new beginning that it opens –, these Walt Whitman’s verses from poem Song of Myself (1855/2005) serve as inspiration: *This day before dawn I ascended a hill and look'd at the crowded heaven, / And I said to my spirit When we become the enfolders of those orbs, / and the pleasure and knowledge of every thing in them, shall we / be fill'd and satisfied then? / And my spirit said No, we but level that lift to pass and continue beyond.*

REFERENCES

Abboud, H., Schultz, W. H., & Zeitlin, V. (2006). SuperLab, Stimulus presentation software (Version 4.0). San Pedro, California: Cedrus Corporation.

Adolphs, R. (2006). How do we know the minds of others? Domain-specificity, simulation, and enactive social cognition. *Brain Research, 1079*, 25-35. doi:10.1016/j.brainres.2005.12.127

Adolphs, R. (2010). Social cognition: Feeling voices to recognize emotions. *Current Biology, 20*, R1071-R1072. doi:10.1016/j.cub.2010.11.019

Adolphs, R., Damasio, H., & Tranel, D. (2002). Neural systems for recognition of emotional prosody: A 3-D lesion study. *Emotion, 2*, 23-51. doi:10.1037/1528-3542.2.1.23

Adolphs, R., Schul, R., & Tranel, D. (1998). Intact recognition of facial emotion in Parkinson's disease. *Neuropsychology, 12*, 253-258. doi:10.1037/0894-4105.12.2.253

Adolphs, R., & Tranel, D. (1999). Intact recognition of emotional prosody following amygdala damage. *Neuropsychologia, 37*, 1285-1292. doi:10.1016/S0028-3932(99)00023-8

Adolphs, R., Tranel, D., & Damasio, H. (2001). Emotion recognition from faces and prosody following temporal lobectomy. *Neuropsychology, 15*, 396-404. doi:10.1037/0894-4105.15.3.396

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. R. (1995). Fear and the human amygdala. *The Journal of Neuroscience, 15*, 5879-5891.

Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., … Sobel, N. (2003). Dissociated neural represenations of intensity and valence in human olfaction. *Nature Neuroscience, 6*, 196-202. doi:10.1038/nn1001

Ariatti, A., Benuzzi, F., & Nichelli, P. (2008). Recognition of emotions from visual and prosodic cues in Parkinson's disease. *Neurological Sciences, 29*, 219-227. doi:10.1007/s10072-008-0971-9

Assogna, F., Pontieri, F. E., Caltagirone, C., & Spalletta, G. (2008). The recognition of facial emotion expressions in Parkinson's disease. *European Neuropsychopharmacology, 18*, 835-848. doi:10.1016/j.euroneuro.2008.07.004

Bach, D. R., Buxtorf, K., Grandjean, D., & Strik, W. K. (2009). The influence of emotion clarity on emotional prosody identification in paranoid schizophrenia. *Psychological Medicine, 39*, 927-938. doi:10.1017/S0033291708004704

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*, 919-927. doi:10.1016/j.neuroimage.2008.05.034

Balkwill, L.-L., & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception, 17*, 43-64.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*, 614-636. doi:10.1037/0022-3514.70.3.614

Bänziger, T., Grandjean, D., & Scherer, K. R. (2009). Emotion recognition from expressions in face, voice, and body: The Multimodal Emotion Recognition Test (MERT). *Emotion, 9*, 691-704. doi:10.1037/a0017088

Bänziger, T., & Scherer, K. R. (2007). Using actor portrayals to systematically study multimodal emotion expression: The GEMEP Corpus. *Lectures Notes in Computer Science, 4738*, 476-487. doi:10.1007/978-3-540-74889-2_42

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*, 5-17. doi:10.1023/A:1005653411471

Barreto, J., Leuschner, A., Santos, F., & Sobral, M. (2003). Escala de depressão geriátrica: Tradução portuguesa da Geriatric Depression Scale, de Yesavage et al. (1983) [Portuguese translation of the Geriatric Depression Scale, by Yesavage et al. (1983)]. In A. Mendonça, C. Garcia & M. Guerreiro (Eds.), *Escalas e Testes na Demência: Grupo de Estudos de Envelhecimento Cerebral e Demência [Scales and tests in dementia: Research group on brain aging and dementia]*. Lisbon: Grupo de Estudos de Envelhecimento Cerebral e Demência.

Barrett, F. S., Grimm, K. J., Robins, R. W., Wildschut, T., Sedikides, C., & Janata, P. (2010). Music-evoked nostalgia: Affect, memory, and personality. *Emotion, 10*, 390-403. doi:10.1037/a0019006

Barrett, L. F. (2006a). Are emotions natural kinds? *Perspectives on Psychological Science, 1*, 28-58. doi: 10.1111/j.1745-6916.2006.00003.x

Barrett, L. F. (2006b). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review, 10*, 20-46. doi:10.1207/s15327957pspr1001_2

Barrett, L. F. (2009). Variety if the spice of life: A psychological construction approach to understanding variability in emotion. *Cognition and Emotion, 23*, 1284-1306. doi:10.1080/02699930902985894

Barrett, L. F., & Kensinger, E. (2010). Context is routinely encoded during emotion perception. *Psychological Science, 21*, 595-599. doi:10.1177/0956797610363547

Beaucousin, V., Lacheret, A., Turbelin, M. R., Morel, M., Mazoyer, B., & Tzourio-Mazoyer, N. (2007). FMRI study of emotional speech comprehension. *Cerebral Cortex, 17*, 339-352. doi:10.1093/cercor/bhj151

Belin, P., Bestelmeyer, P., Latinus, M., & Watson, R. (in press). Understanding voice perception. *British Journal of Psychology*. Advance online publication. doi: 10.1111/j.2044-8295.2011.02041.x

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

Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods, 40*, 531-539. doi:10.3758/BRM.40.2.531

Benke, T., Bösch, S., & Andree, B. (1998). A study of emotional processing in Parkinson's disease. *Brain and Cognition, 38*, 36-52. doi:10.1006/brcg.1998.1013

Benton, A. L., Sivan, A. B., Hamsher, K., Varney, N., & Spreen, O. (1994). *Contributions to neuropsychological assessment: A clinical manual* (2nd ed.). New York: Oxford University Press.

Berckmoes, C., & Vingerhoets, G. (2004). Neural foundations of emotional speech processing. *Current Directions in Psychological Science, 13*, 182-185. doi:10.1111/j.0963-7214.2004.00303.x

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology, 2*, 1-12. doi:10.3389/fpsyg.2011.00094

Bhatara, A., Tirovolas, A. K., Duan, L. M., Levy, B., & Levitin, D. J. (2011). Perception of emotional expression in musical performance. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 921-934. doi:10.1037/a0021922

Bialystok, E., & DePape, A. M. (2009). Musical expertise, bilingualism, and executive functioning. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 565-574. doi:10.1037/a0012735

Bigand, E., & Poulin-Charronnat, B. (2006). Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. *Cognition, 100*, 100-130. doi:10.1016/j.cognition.2005.11.007

Bigand, E., Vieillard, S., Madurell, F., Marozeau, J., & Dacquet, A. (2005). Multidimensional scaling of emotional responses to music: The effect of musical expertise and of the duration of the excerpts. *Cognition and Emotion, 19*, 1113-1139. doi:10.1080/02699930500204250

Blair, R. J., Mitchell, D., Richell, R., Kelly, S., Leonard, A., Newman, C., … Scott, S. K. (2002). Turning a deaf ear to fear: Impaired recognition of vocal affect in psychopathic individuals. *Journal of Abnormal Psychology, 111*, 682-686. doi:10.1037/0021-843X.111.4.682

Blonder, L. X., Bowers, D., & Heilman, K. M. (1991). The role of the right hemisphere in emotional communication. *Brain, 114*, 1115-1127. doi:10.1093/brain/114.3.1115

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences, 98*, 11818-11823. doi:10.1073/pnas.191355898

Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience, 2*, 382-387. doi:10.1038/7299

Boersma, P., & Weenink, D. (2009). Praat: Doing phonetics by computer (Version 5.1.05). Retrieved from http://www.praat.org/

Borod, J. C., Pick, L. H., Hall, S., Sliwinski, M., Madigan, N., Obler, L., … Tabert, M. (2000). Relationships among facial, prosodic, and lexical channels of emotional perceptual processing. *Cognition and Emotion, 14*, 193-211. doi:10.1080/026999300378932

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

Bowling, D. L., Gill, K., Choi, J. D., Prinz, J., & Purves, D. (2010). Major and minor music compared to excited and subdued speech. *The Journal of the Acoustical Society of America, 1*, 491-503. doi:10.1121/1.3268504

Braak, H., Ghebremedhin, E., Rüb, U., Bratzke, H., & Tredici, K. D. (2004). Stages in the development of Parkinson's disease-related pathology. *Cell and Tissue Research, 318*, 121-134. doi:10.1007/s00441-004-0956-9

Braak, H., Tredici, K. D., Rüb, U., Vos, R. d., Steur, E., & Braak, E. (2003). Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiology of Aging, 24*, 197-211. doi:10.1016/S0197-4580(02)00065-9

Breitenstein, C., Lancker, D. V., & Daum, I. (2001). The contribution of speech rate and pitch variation to the perception of vocal emotions in a German and an American sample. *Cognition and Emotion, 15*, 57-79. doi:10.1080/0269993004200114

Breitenstein, C., Lancker, D. V., Daum, I., & Waters, C. H. (2001). Impaired perception of vocal emotions in Parkinson's disease: Influence of speech time processing and executive functioning. *Brain and Cognition, 45*, 277-314. doi:10.1006/brcg.2000.1246

Brochard, R., Dufour, A., & Després, O. (2004). Effect of musical expertise on visuospatial abilities: Evidence from reaction times and mental imagery. *Brain and Cognition, 54*, 103-109. doi:10.1016/S0278-2626(03)00264-1

Brown, S. (2000). The "musilanguage" model of music evolution. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 271-300). Cambridge: MIT Press.

Brück, C., Kreifelts, B., Kaza, E., Lotze, M., & Wildgruber, D. (2011). Impact of personality on the cerebral processing of emotional prosody. *NeuroImage, 58*, 259-268. doi:10.1016/j.neuroimage.2011.06.005

Bruckert, L., Bestelmeyer, P., Latinus, M., Rouger, J., Charest, I., Rousselet, G. A., … Belin, P. (2010). Vocal attractiveness increases by averaging. *Current Biology, 20*, 116-120. doi:10.1016/j.cub.2009.11.034

Brunswik, E. (1956). *Perception and the representative design of psychological experiments* (2nd ed.). Berkeley, California: University of California Press.

Budd, M. (1985). *Music and the emotions: The philosophical theories*. London: Routledge.

Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W., & Weiss, B. (2005, September). *A database of German emotional speech.* Paper session presented at the 9th European Conference on Speech Communication and Technology, Lisbon, Portugal.

Cacioppo, J. T., Berntson, G. G., Bechara, A., Tranel, D., & Hawkley, L. C. (2011). Could an aging brain contribute to subjective well being?: The value added by a social neuroscience perspective. In A. Tadorov, S. T. Fiske & D. Prentice (Eds.), *Social Neuroscience: Toward understanding the underpinnings of the social mind* (pp. 249-262). New York: Oxford University Press.

Caekebeke, J. F., Jennekens-Schinkel, A., Linden, M. E. v. d., Buruma, O. J. S., & Roos, R. A. C. (1991). The interpretation of dysprosody in patients with Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry, 54*, 145-148. doi:10.1136/jnnp.54.2.145

Calder, A., Keane, J., Manly, T., Sprengelmeyer, R., Scott, S. K., Nimmo-Smith, I., & Young, A. W. (2003). Facial expression recognition across the adult life span. *Neuropsychologia, 41*, 195-202. doi:10.1016/S0028-3932(02)00149-5

Carstensen, L., & Mikels, J. (2005). At the intersection of emotion and cognition. *Current Directions in Psychological Science, 14*, 117-121. doi:10.1111/j.0963-7214.2005.00348.x

Carstensen, L., Pasupathi, M., Mayr, U., & Nesselroade, J. (2000). Emotional experience in everyday life across the adult life span. *Journal of Personality and Social Psychology, 79*, 644-655. doi:10.1037/0022-3514.79.4.644

Carton, J. S., Kessler, E. A., & Pape, C. L. (1999). Nonverbal decoding skills and relationship well-being in adults. *Journal of Nonverbal Behavior, 23*, 91-100. doi:10.1023/A:1021339410262

Carvalho, S., Lima, C. F., & Castro, S. L. (2011, March). O processamento da altura tonal prediz o reconhecimento de emoções em prosódia? [Do pitch processing abilities predict emotion recognition in speech prosody?]. Poster session presented at the 6th annual meeting of the Associação Portuguesa de Psicologia Experimental, Coimbra, Portugal.

Castro, S. L., Caló, S., Gomes, I., Kay, J., Lesser, R., & Coltheart, M. (2007). *PALPA-P, Provas de Avaliação da Linguagem e da Afasia em Português [Tasks for the assessment of language processing and aphasia in Portuguese, PALPA-P]*. Lisbon: CEGOC-TEA.

Castro, S. L., & Lima, C. F. (2009). QA- Quociente de Espectro Autista Versão Adultos (16+ anos), versão portuguesa [Autism Spectrum Quotient Questionnaire for Adults, Portuguese version]. Retrieved from http://www.autismresearchcentre.com/arc_tests

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*, 74-81. doi:10.3758/BRM.42.1.74

Castro, S. L., & Lima, C. F. (submitted). Emotion recognition in music is robust, and yet variable: The impact of age and musical expertise.

Castro, S. L., Martins, L., & Cunha, L. (2003). Neuropsychological screening with a Portuguese Stroop test [Abstract]. *The Clinical Neuropsychologist, 17*, 104.

Charles, S. T., & Carstensen, L. (2010). Social and emotional aging. *Annual Review of Psychology, 61*, 383-409. doi:10.1146/annurev.psych.093008.100448

Chartrand, J. P., & Belin, P. (2006). Superior voice timbre processing in musicians. *Neuroscience Letters, 405*, 164-167. doi:10.1016/j.neulet.2006.06.053

Chaudhuri, K. R., Healy, D. G., & Schapira, A. H. V. (2006). Non-motor symptoms of Parkinson's disease: Diagnosis and management. *Lancet Neurology, 5*, 235-245. doi:10.1016/S1474-4422(06)70373-8

Chaudhuri, K. R., & Schapira, A. H. V. (2009). Non-motor symptoms of Parkinson's disease: Dopaminergic pathophysiology and treatment. *Lancet Neurology, 8*, 464-474. doi:10.1016/S1474-4422(09)70068-7

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex, 18*, 2844-2854. doi:10.1093/cercor/bhn042

Cohen, A. J. (2010). Music as a source of emotion in film. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications*. New York: Oxford University Press.

Cohen, H., Gagné, M. H., Hess, U., & Pourcher, E. (2010). Emotion and object processing in Parkinson's disease. *Brain and Cognition, 72*, 457-463. doi:10.1016/j.bandc.2010.01.001

Colibazzi, T., Posner, J., Wang, Z., Gorman, D., Gerber, A., Yu, S., … Russell, J. A. (2010). Neural systems subserving valence and arousal during the experience of induced emotions. *Emotion, 10*, 377-389. doi:10.1037/a0018484

Cunningham, J. G., & Sterling, R. S. (1988). Developmental change in the understanding of affective meaning in music. *Motivation and Emotion, 12*, 399-413. doi:10.1007/BF00992362

Curiati, P. K., Tamashiro, J. H., Squarzoni, P., Duran, F. L. S., Santos, L. C., Wajngarten, M., … Alves, T. C. T. F. (2009). Brain structural variability due to aging and gender in cognitively healthy elders: Results from the São Paulo Ageing and Health Study. *American Journal of Neuroradiology, 30*, 1850-1856. doi:10.3174/ajnr.A1727

Curtis, M. E., & Bharucha, J. J. (2010). The minor third communicates sadness in speech, mirroring its use in music. *Emotion, 10*, 335-348. doi:10.1037/a0017928

d' Errico, F., Henshilwood, C., Lawson, G., Vanhaeren, M., Tillier, A. M., Soressi, M., … Julien, M. (2003). Archaeological evidence for the emergence of language, symbolism, and music - An alternative multidisciplinary perspective. *Journal of World Prehistory, 17*, 1-70. doi:10.1023/A:1023980201043

Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition, 80*, B1-B10. doi:10.1016/S0010-0277(00)00136-0

Dara, C., Monetta, L., & Pell, M. D. (2008). Vocal emotion processing in Parkinson's disease: Reduced sensitivity to negative emotions. *Brain Research, 1188*, 100-111. doi:10.1016/j.brainres.2007.10.034

Darwin, C. (1872/2009). *The expression of the emotions in man and animals*. London: Penguin Books.

Davidson, R. J., Scherer, K. R., & Goldsmith, H. H. (2002). *Handbook of Affective Sciences* (1st ed.). Oxford: Oxford University Press.

de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion, 14*, 289-311. doi:10.1080/026999300378824

Dellacherie, D., Roy, M., Hugueville, L., Peretz, I., & Samson, S. (2011). The effect of musical experience on emotional self-reports and psychophysiological responses to dissonance. *Psychophysiology, 48*, 337-349. doi: 10.1111/j.1469-8986.2010.01075.x

Delplanque, S., Grandjean, D., Chrea, C., Coppin, G., Aymard, L., Cayeux, I., … Scherer, K. R. (2009). Sequential unfolding of novelty and pleasantness appraisals of odors: Evidence from facial electromyography and autonomic reactions. *Emotion, 9*, 316-328. doi:10.1037/a0015369

Dickson, D. W., Fujishiro, H., Orr, C., DelleDonne, A., Josephs, K. A., Frigerio, R., … Ahlskog, J. E. (2009). Neuropathology of non-motor features of Parkinson disease. *Parkinsonism and Related Disorders, 15S3*, S1-S5. doi:10.1016/S1353-8020(09)70769-2

Doherty, C. P., Fitzsimons, M., Asenbauer, B., & Staunton, H. (1999). Discrimination of prosody and music by normal children. *European Journal of Neurology, 6*, 221-226. doi:10.1111/j.1468-1331.1999.tb00016.x

Drapeau, J., Gosselin, N., Gagnon, L., Peretz, I., & Lorrain, D. (2009). Emotional recognition from face, voice, and music in dementia of the Alzheimer type. *Annals of the New York Academy of Sciences, 1169*, 342-345. doi: 10.1111/j.1749-6632.2009.04768.x

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

Ekman, P. (1992b). An argument for basic emotions. *Cognition and Emotion, 6*(3-4), 169-200. doi:10.1080/02699939208411068

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russel's mistaken critique. *Psychological Bulletin, 115*, 268-287. doi:10.1037/0033-2909.115.2.268

Ekman, P. (1999). Basic emotions. In T. Dalgleish & M. Power (Eds.), *Handbook of Cognition and Emotion*. Sussex: Wiley.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: Investigator's guide*. Palo Alto, California: Consulting Psychologists Press.

Elfenbein, H. A. (2006). Learning in emotion judgments: Training and the cross-cultural understanding of facial expressions. *Journal of Nonverbal Behavior, 30*(1), 21-36. doi:10.1007/s10919-005-0002-y

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin, 128*, 203-235. doi:10.1037/0033-2909.128.2.203

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology, 47*, 273-305. doi:10.1146/annurev.psych.47.1.273

Ericsson, K. A., & Towne, T. J. (2010). Expertise. *WIREs Cognitive Science, 1*, 404-416. doi:10.1002/WCS.47

Ethofer, T., Anders, S., Erb, M., Herbert, C., Wiethoff, S., Kissler, J., … Wildgruber, D. (2006). Cerebral pathways in processing of affective prosody: A dynamic causal modeling study. *NeuroImage, 30*, 580-587. doi:10.1016/j.neuroimage.2005.09.059

Ethofer, T., Keifelts, B., Wiethoff, S., Wolf, J., Grodd, W., Vuilleumier, P., et al. (2008). Differential influences of emotion, task, and novelty on brain regions underlying the processing of speech melody. *Journal of Cognitive Neuroscience, 21*, 1255-1268. doi:10.1162/jocn.2009.21099

Ethofer, T., Ville, D. V. D., Scherer, K. R., & Vuilleumier, P. (2009). Decoding of emotional information in voice-sensitive cortices. *Current Biology, 19*, 1028-1033. doi:10.1016/j.cub.2009.04.054

Eugène, F., Lévesque, J., Mensour, B., Leroux, J. M., Beaudoin, G., Bourgouin, P., & Beauregard, M. (2003). The impact of individual differences on the neural circuitry underlying sadness. *NeuroImage, 19*, 354-364. doi:10.1016/S1053-8119(03)00121-6

Evans, P., & Schubert, E. (2008). Relationship between expressed and felt emotions in music. *Musicae Scientiae, XII*, 75-99. doi:10.1177/102986490801200105

Fahn, S., Elton, R. L., & UPDRS program members (1987). Unified Parkinson's Disease Rating Scale. In S. Fahn, C. D. Marsden, M. Goldstein & D. B. Calne (Eds.), *Recent development in Parkinson's disease* (Vol. 2, pp. 153-163). New Jersey: Macmillan Healthcare Information.

Fichtel, C., Hammerschmidt, K., & Jürgens, U. (2001). On the vocal expression of emotion: A multi-parametric analysis of different states of aversion in the squirrel monkey. *Behaviour, 138*, 97-116. doi:10.1163/15685390151067094

Filipic, S., Tillmann, B., & Bigand, E. (2010). Judging familiarity and emotion from very brief musical excerpts. *Psychonomic Bulletin & Review, 17*, 335-341. doi:10.3758/PBR.17.3.335

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189-198. doi:10.1016/0022-3956(75)90026-6

Forgeard, M., Winner, E., Norton, A., & Schlaug, G. (2008). Practicing a musical instrument in childhood is associated with enhanced verbal ability and nonverbal reasoning. *PLoS One, 3*, e3566. doi:10.1371/journal.pone.0003566

Freeman, T. W., Hart, J., Kimbrell, T., & Ross, E. D. (2009). Comprehension of affective prosody in veterans with chronic posttraumatic stress disorder. *The Journal of Neuropsychiatry and Clinical Neurosciences, 21*, 52-58. doi:10.1176/appi.neuropsych.21.1.52

Friederici, A. D., & Alter, K. (2004). Lateralization of auditory language functions: A dynamic dual pathway model. *Brain and Language, 89*, 267-276. doi: 10.1016/S0093-934X(03)00351-1

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., … Koelsch, S. (2009). Universal recognition of three basic emotions in music. *Current Biology, 19*, 573-576. doi:10.1016/j.cub.2009.02.058

Frühholz, S., Ceravolo, L., & Grandjean, D. (in press). Specific brain networks during explicit and implicit decoding of emotional prosody. *Cerebral Cortex*. Advance online publication. doi:10.1093/cercor/bhr184

Fujioka, T., Ross, B., Kakigi, R., Pantev, C., & Trainor, L. J. (2006). One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain, 129*, 2593-2608. doi:10.1093/brain/awl247

Gabrielsson, A. (2002). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae, Special Issue 2001-2002*, 123-147.

Gabrielsson, A. (2009). The relationship between musical structure and perceived expression. In S. Hallam, I. Cross & M. H. Thaut (Eds.), *The Oxford Handbook of Music Psychology* (pp. 141-150). New York: Oxford University Press.

Gabrielsson, A., & Lindström, E. (2010). The role of structure in the musical expression of emotions. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 367-400). New York: Oxford University Press.

Gagnon, L., & Peretz, I. (2003). Mode and tempo relative contributions to "happy-sad" judgments in equitone melodies. *Cognition and Emotion, 17*, 25-40. doi:10.1080/02699930302279

Gale, J. T., Amirnovin, R., Williams, Z. M., Flaherty, A. W., & Eskandar, E. N. (2008). From symphony to cacophony: Pathophysiological of the human basal ganglia in Parkinson disease. *Neuroscience and Biobehavioral Reviews, 32*, 378-387. doi:10.1016/j.neubiorev.2006.11.005

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *The Journal of Neuroscience, 23*, 9240-9245. doi:0270-6474/03/239240-06$15.00/0

Gauntlett-Gilbert, J., & Brown, V. J. (1998). Reaction time deficits and Parkinson's disease. *Neuroscience and Biobehavioral Reviews, 22*, 865-881. doi:10.1016/S0149-7634(98)00014-1

de Gelder, B. (2010). The grand challenge for Frontiers in Emotion Science. *Frontiers in Psychology, 1*, 187. doi:10.3389/fpsyg.2010.00187

de Gelder, B., & den Stock, J. V. (2011). The bodily expressive action stimulus test (BEAST). Construction and validation of a stimulus basis for measuring perception of whole body expression of emotions. *Frontiers in Psychology, 2,* 181. doi:10.3389/fpsyg.2011.00181

Globerson, E., Lavidor, M., Golan, O., Kishon-Rabin, L., & Amir, N. (2010, May). *Psychoacoustic abilities as predictors of vocal emotion recognition*. Paper session presented at the Speech Prosody 5th International Conference, Chicago, USA.

Goeleven, E., Raedt, R. D., Leyman, L., & Verschuere, B. (2008). The Karolinska Directed Emotional Faces: A validation study. *Cognition and Emotion, 22*, 1094-1118. doi:10.1080/02699930701626582

Golan, O., Baron-Cohen, S., & Hill, J. (2006). The Cambridge Mindreading (CAM) Face-Voice Battery: Testing complex emotion recognition in adults with and without Asperger Syndrome. *Journal of Autism and Developmental Disorders, 36*, 169-183. doi: 10.1007/s10803-005-0057-y

Gorelick, P., & Ross, E. D. (1987). The aprosodias: Further functional-anatomical evidence for the organisation of affective language in the right hemisphere. *Journal of Neurology, Neurosurgery, and Psychiatry, 50*, 553-560. doi:10.1136/jnnp.50.5.553

Gosling, S. D., Rentfrow, P. J., & Swann Jr., W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality 37*, 504-528. doi:10.1016/S0092-6566(03)00046-1

Gosselin, N., Peretz, I., Hasboun, D., Baulac, M., & Samson, S. (2011). Impaired recognition of musical emotions and facial expressions following anteromedial temporal lobe excision. *Cortex, 47*, 1116-1125. doi:10.1016/j.cortex.2011.05.012

Gosselin, N., Peretz, I., Johnsen, E., & Adolphs, R. (2007). Amygdala damage impairs emotion recognition from music. *Neuropsychologia, 45*, 236-244. doi:10.1016/j.neuropsychologia.2006.07.012

Gosselin, N., Peretz, I., Noulhiane, M., Hasboun, D., Beckett, C., Baulac, M., & Samson, S. (2005). Impaired recognition of scary music following unilateral temporal lobe excision. *Brain, 128*, 628-640. doi:10.1093/brain/awh420

Gosselin, N., Samson, S., Adolphs, R., Noulhiane, M., Roy, M., Hasboun, D., … Peretz, I. (2006). Emotional responses to unpleasant music correlates with damage to the parahippocampal cortex. *Brain, 129*, 2585-2592. doi:10.1093/brain/awl240

Grandjean, D., Bänziger, T., & Scherer, K. R. (2006). Intonation as an interface between language and affect. *Progress in Brain Research, 156*, 235-247. doi:10.1016/S0079-6123(06)56012-1

Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). The voices of wrath: Brain responses to angry prosody in meaningless speech. *Nature Neuroscience, 8*, 145-146. doi:10.1038/nn1392

Gray, H. M., & Tickle-Degnen, L. (2010). A meta-analysis of performance on emotion recognition tasks in Parkinson's disease. *Neuropsychology, 24*, 176-191. doi:10.1037/a0018104

Graybiel, A. M. (2005). The basal ganglia: Learning new tricks and loving it. *Current Opinion in Neurobiology, 15*, 638-644. doi:10.1016/j.conb.2005.10.006

Guerreiro, M., Silva, A. P., Botelho, M. A., Leitão, O., & Garcia, C. (1994). Adaptação à população portuguesa da tradução do "Mini Mental State Examination" (MMSE) [Adaptation to the Portuguese population of the translation of Mini Mental State Examination (MMSE)]. *Revista Portuguesa de Neurologia, 1*, 9-10.

Gunning-Dixon, F., Gur, R., Perkins, A., Schroeder, L., Turner, T., Turetsky, B., … Gur, R. E. (2003). Age-related differences in brain activation during emotional face processing. *Neurobiology of Aging, 24*, 285-295. doi:10.1016/S0197-4580(02)00099-4

Habib, M., & Besson, M. (2009). What do music training and musical experience teach us about brain plasticity? *Music Perception, 26*, 279-285. doi:10.1525/MP.2009.26.3.279

Hagen, E., Nummenmaa, L., Yu, R., Engell, A. D., Ewbank, M. P., & Calder, A. (2011). Autism spectrum traits in the typical population predict structure and function in the posterior superior temporal sulcus. *Cerebral Cortex, 21*, 493-500. doi:10.1093/cercor/bhq062

Hall, J. A., Andrzejewski, S. A., & Yopchick, J. E. (2009). Psychosocial correlates of interpersonal sensitivity: A meta-analysis. *Journal of Nonverbal Behavior, 33*, 149-180. doi:10.1007/s10919-009-0070-5

Hamann, S., & Canli, T. (2004). Individual differences in emotion processing. *Current Opinian in Neurobiology, 14*, 233-238. doi:10.1016/j.conb.2004.03.010

Hamel, R., & Schmittmann, V. D. (2006). The 20-minute version as a predictor of the Raven Advanced Progressive Matrices test. *Educational and Psychological Measurement, 66*, 1039-1046. doi:10.1177/0013164406288169

Hammerschmidt, K., & Jürgens, U. (2007). Acoustical correlated of affective prosody. *Journal of Voice, 21*, 531-540. doi:10.1016/j.jvoice.2006.03.002

Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: Effects of enculturation and formal training on development. *Trends in Cognitive Sciences, 11*, 466-472. doi:10.1016/j.tics.2007.08.008

Hauser, M. D., & McDermott, J. (2003). The evolution of the music faculty: A comparative perspective. *Nature Neuroscience, 6*, 663-668. doi:10.1038/nn1080

Hedden, T., & Gabrieli, J. (2004). Insights into the ageing mind: A view from cognitive neuroscience. *Nature Reviews Neuroscience, 5*, 87-97. doi:10.1038/nrn1323

Helmholtz, H. L. F. v. (1863/1954). *On the sensations of tone*. New York: Dover Publications.

Hillecke, Nickel, A., & Bolay, H. V. (2005). Scientific perspectives on music therapy. *Annals of the New York Academy of Sciences, 1060*, 271-282. doi:10.1196/annals.1360.020

Hoehn, M. M., & Yahr, M. D. (1967). Parkinsonism: Onset, progression, and mortality. *Neurology, 17*, 427-442.

Hooker, C., & Park, S. (2002). Emotion processing and its relationship to social functioning in schizophrenia patients. *Psychiatry Research, 112*, 41-50. doi:10.1016/S0165-1781(02)00177-4

Hornak, J., Rolls, E. T., & Wade, D. (1996). Face and voice expression identification in patients with emotional and behavioral changes following ventral frontal lobe damage. *Neuropsychologia, 34*, 247-261. doi:10.1016/0028-3932(95)00106-9

Hughes, A. J., Daniel, S. E., Kilford, L., & Lees, A. J. (1992). Accuracy of clinical diagnosis of idiopathic Parkinson's disease: A clinico-pathological study of 100 cases. *Journal of Neurology, Neurosurgery, and Psychiatry, 55*, 181-184. doi:10.1136/jnnp.55.3.181

Hunter, P., Schellenberg, E. G., & Schimmack, U. (2008). Mixed affective responses to music with conflicting cues. *Cognition and Emotion, 22*, 327-352. doi:10.1080/02699930701438145

Hunter, P., Schellenberg, E. G., & Schimmack, U. (2010). Feelings and perceptions of happiness and sadness induced by music: Similarities, differences, and mixed emotions. *Psychology of Aesthetics, Creativity, and the Arts, 4*, 47-56. doi:10.1037/a0016873

Hunter, P. G., Schellenberg, E. G., & Stalinski, S. M. (2011). Liking and identifying emotionally expressive music: Age and gender differences. *Journal of Experimental and Child Psychology, 110*, 80-93. doi:10.1016/j.jecp.2011.04.001

Hyde, K. L., Lerch, J. P., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). Musical training shapes structural brain development. *The Journal of Neuroscience, 29*, 3019-3025. doi:10.1523/JNEUROSCI.5118-08.2009

Ilie, G., & Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception, 23*, 319-329.

Isaacowitz, D., Lockenhoff, C., Lane, R., Wright, R., Sechrest, L., Riedel, R., & Costa, P. T. (2007). Age differences in recognition of emotion in lexical stimuli and facial expressions. *Psychology and Aging, 22*, 147-159. doi:10.1037/0882-7974.22.1.147

Isaacowitz, D., & Stanley, J. T. (2011). Bringing an ecological perspective to the study of aging and recognition of emotional facial expressions: Past, current, and future methods. *Journal of Nonverbal Behavior*. Advance online publication. doi:10.1007/s10919-011-0113-6

Izard, C. E. (2009). Emotion theory and research: Highlights, unanswered questions, and emerging issues. *Annual Review of Psychology, 60*, 1-25. doi:10.1146/annurev.psych.60.110707.163539

James, W. (1884). What is an emotion? *Mind, 9*, 188-205.

Jaywant, A., & Pell, M. D. (2010). Listeners impressions of speakers with Parkinson's disease. *Journal of the International Neuropsychological Society, 16*, 49-57. doi:10.1017/S1355617709990919

Johnsen, E. L., Tranel, D., Lutgendorf, S., & Adolphs, R. (2009). A neuroanatomical dissociation for emotion induced by music. *International Journal of Psychophysiology, 72,* 24-33. doi:10.1016/j.ijpsycho.2008.03.011

Johnson-Laird, P. N., & Oatley, K. (2010). Emotions, music, and literature. In M. Lewis, J. M. Haviland-Jones & L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 102-113). New York: The Guilford Press.

Johnstone, T., Reekum, C. M. v., Oakes, T. R., & Davidson, R. J. (2006). The voice of emotion: An fMRI study of neural responses to angry and happy vocal expressions. *Social Cognitive and Affective Neuroscience, 1*, 242-249. doi:10.1093/scan/nsl027

Juslin, P. N. (1997). Perceived emotional expression in synthesized performances of a short melody: Capturing the listeners' judgment policy. *Musicae Scientiae, 1*, 225-256.

Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1797-1813. doi:10.1037/0096-1523.26.6.1797

Juslin, P. N. (2009). Emotional responses to music. In S. Hallam, I. Cross & M. H. Thaut (Eds.), *The Oxford Handbook of Music Psychology* (pp. 131-140). New York: Oxford University Press.

Juslin, P. N., & Laukka, P. (2001). Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion, 1*, 381-412. doi:10.1037/1528-3542.1.4.381

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin, 129*, 770-814. doi:10.1037/0033-2909.129.5.770

Juslin, P. N., Liljeström, S., Västfjäll, D., Barradas, G., & Silva, A. (2008). An experience sampling study of emotional reactions to music: Listener, music, and situation. *Emotion, 8*, 668-683. doi:10.1037/a0013505

Juslin, P. N., Liljeström, S., Västfjäll, D., & Lundqvist, L. (2010). How does music evoke emotions? Exploring the underlying mechanisms. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of Music and Emotion: Theory, Research, Applications* (pp. 605-642). New York: Oxford University Press.

Juslin, P. N., & Sloboda, J. A. (2010). Introduction: Aims, organization, and terminology. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 3-12). New York: Oxford University Press.

Juslin, P. N., & Timmers, R. (2010). Expression and communication of emotion in music performance. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of Music and Emotion: Theory, Research, Applications* (pp. 453-489). New York: Oxford University Press.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences, 31*, 559-621. doi:10.1017/S0140525X08005529

Kan, Y., Kawamura, M., Hasegawa, Y., Mochizuki, S., & Nakamura, K. (2002). Recognition of emotion from facial, prosodic and written verbal stimuli in Parkinson's disease. *Cortex, 38*, 623-630. doi:10.1016/S0010-9452(08)70026-1

Kehagia, A. A., Barker, R. A., & Robbins, T. W. (2010). Neuropsychological and clinical heterogeneity of cognitive impairment and dementia in patients with Parkinson's disease. *Lancet Neurology, 9*, 1200-1213. doi:10.1016/S1474-4422(10)70212-X

Keltner, D., & Ekman, P. (2002). Introduction: Expression of emotion. In R. J. Davidson, K. R. Scherer & H. H. Goldsmith (Eds.), *Handbook of Affective Sciences*. Oxford: Oxford University Press.

Khalfa, S., Della Bella, S., Roy, M., Peretz, I., & Lupien, S. J. (2003). Effects of relaxing music on salivary cortisol level after psychological stress. *Annals of the New York Academy of Sciences, 999*, 374-376. doi:10.1196/annals.1284.045

Khalfa, S., Guye, M., Peretz, I., Chapon, F., Girard, N., Chauvel, P., & Liégeois-Chauvel, C. (2008). Evidence of lateralized anteromedial temporal strucures involvement in musical emotion processing. *Neuropsychologia, 46*, 2485-2493. doi:10.1016/j.neuropsychologia.2008.04.009

Kisley, M. A., Wood, S., & Burrows, C. L. (2007). Looking at the sunny side of life: Age-related change in an event-related potential measure of the negativity bias. *Psychological Science, 18*, 838-843. doi:10.1111/j.1467-9280.2007.01988.x

Kivy, P. (1991). *Music alone: Philisophical reflections on the purely musical experience*. Ithaca New York: Cornell University Press.

Kivy, P. (2002). *Introduction to a philosophy of music*. New York: Oxford University Press.

Knösche, T. R., Neuhaus, C., Haueisen, J., Alter, K., Maess, B., Witte, O. W., & Friederici, A. D. (2005). Perception of phrase structure in music. *Human Brain Mapping, 24*, 259-273. doi:10.1002/hbm.20088

Koelsch, S. (2010). Towards a neural basis of music-evoked emotions. *Trends in Cognitive Sciences, 14*, 131-137. doi:10.1016/j.tics.2010.01.002

Koelsch, S. (2011). Toward a neural basis of music perception: A review and updated model. *Frontiers in Psychology, 2*, 1-20. doi:10.3389/fpsyg.2011.00110

Koelsch, S., Fritz, T., Cramon, D. Y. v., Muller, K., & Friederici, A. D. (2006). Investigating emotion with music: An fMRI study. *Human Brain Mapping, 27*, 239-250. doi:10.1002/hbm.20180

Koelsch, S., Gunter, T., Wittfoth, M., & Sammler, D. (2005). Interaction between syntax processing in language and in music: An ERP study. *Journal of Cognitive Neuroscience, 17*, 1565-1577. doi:10.1162/089892905774597290

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience, 7*, 302-307. Doi:10.1038/nn1197

Koelsch, S., Offermanns, K., & Franzke, P. (2010). Music in the treatment of affective disorders: An exploratory investigation of a new method for music-therapeutic research. *Music Perception, 27*, 307-316. doi:10.1525/MP.2010.27.4.307

Koelsch, S., & Siebel, W. A. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences, 9*, 578-584. doi:10.1016/j.tics.2005.10.001

Kolinsky, R., Cuvelier, H., Goetry, V., Peretz, I., & Morais, J. (2009). Music training facilitates lexical stress processing. *Music Perception, 26*, 235-246. doi:10.1525/MP.2009.26.3.235

Konecni, V. J. (2008). Does music induce emotion? A theoretical and methodological analysis. *Psychology of Aesthetics, Creativity, and the Arts, 2*, 115-129. doi:10.1037/1931-3896.2.2.115

Korpilahti, P., Jansson-Verkasalo, E., Mattila, M.-L., Kuusikko, S., Suominen, K., Rytky, S., … Moilanen, I. (2007). Processing of affective speech prosody is impaired in Asperger Syndrome. *Journal of Autism and Developmental Disorders, 37*, 1539-1549. doi:10.1007/s10803-006-0271-2

Kotz, S. A., Meyer, M., Alter, K., Besson, M., Cramon, D. Y. v., & Friederici, A. D. (2003). On the lateralization of emotional prosody: An event-related functional MR investigation. *Brain and Language, 86*, 366-376. doi:10.1016/S0093-934X(02)00532-1

Kotz, S. A., & Schwartze, M. (2010). Cortical speech processing unplugged: A timely subcortico-cortical framework. *Trends in Cognitive Sciences, 14*, 392-399. doi:10.1016/j.tics.2010.06.005

Kövari, E., Gold, G., Herrmann, F., Canuto, A., Holf, P., Bouras, C., & Giannakopoulos, P. (2003). Lewy body densities in the entorhinal and anterior cingulate cortex predict cognitive deficits in Parkinson's disease. *Acta Neuropathologica, 106*, 83-88. doi:10.1007/s00401-003-0705-2

Kramer, A. F., Bherer, L., Colcombe, S. J., Dong, W., & Greenough, W. T. (2004). Environmental influences on cognitive and brain plasticity during aging. *Journal of Gerontology: Medical Sciences, 59A*, 940-957. doi:10.1093/gerona/59.9.M940

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

Kwok, L. (2008). Sound Studio (Version 3.5.5). New York City: Felt Tip inc.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8*. Gainesville, Florida: University of Florida.

Latinus, M., & Belin, P. (2011). Human voice perception. *Current Biology, 21*, R143-R145. doi:10.1016/j.cub.2010.12.033

Laukka, P. (2005). Categorical perception of vocal emotion expressions. *Emotion, 5*, 277-295. doi:10.1037/1528-3542.5.3.277

Laukka, P., & Juslin, P. N. (2007). Similar patterns of age-related differences in emotion recognition from speech and music. *Motivation and Emotion, 31*, 182-191. doi:10.1007/s11031-007-9063-z

Laukka, P., Linnman, C., Ahs, F., Pissiota, A., Frans, Ö., Faria, V., … Furmark, T. (2008). In a nervous voice: Acoustic analysis and perception of anxiety in social phobics' speech. *Journal of Nonverbal Behavior, 32*, 195-214. doi:10.1007/s10919-008-0055-9

Lawrence, A. D., Goerendt, I. K., & Brooks, D. J. (2007). Impaired recognition of facial expressions of anger in Parkinson's disease patients acutely withdrawn from dopamine replacement therapy. *Neuropsychologia, 45*, 65-74. doi:10.1016/j.neuropsychologia.2006.04.016

Leitman, D. I., Laukka, P., Juslin, P. N., Saccente, E., Butler, P., & Javitt, D. C. (2010). Getting the cue: Sensory contributions to auditory emotion recognition impairments in schizophrenia. *Schizophrenia Bulletin, 36*, 545-556. doi:10.1093/schbul/sbn115

Leitman, D. I., Wolf, D. H., Ragland, J. D., Laukka, P., Loughead, J., Valdez, J. N., … Gur, R. C. (2010). "It's not what you say, but how you say it": A reciprocal temporo-frontal network for affective prosody. *Frontiers in Human Neuroscience, 4*, 1-13. doi:10.3389/fnhum.2010.00019

Levenson, R. W. (1999). The intrapersonal functions of emotion. *Cognition and Emotion, 13*, 481-504. doi:10.1080/026999399379159

Levenson, R. W. (2011). Basic emotion questions. *Emotion Review, 3*, 1-8. doi:10.1177/1754073911410743

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological Assessment* (4th ed.). New York: Oxford University Press.

Lima, C. F., & Castro, S. L. (2009). TIPI - Inventário de Personalidade de 10 Itens, versão portuguesa [Ten-Item Personality Inventory, portuguese version]. Retrieved from http://homepage.psy.utexas.edu/homepage/faculty/gosling/scales_we.htm#Ten%20Item%20Personality%20Measure%20(TIPI)

Lima, C. F., & Castro, S. L. (2011). Emotion recognition in music changes across the adult life span. *Cognition and Emotion, 25*, 585-598. doi:10.1080/02699931.2010.502449

Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion, 11,* 1021-1031. doi:10.1037/a002452

Lima, C. F., Garrett, C., & Castro, S. L. (submitted). Perceiving emotions in music and in speech prosody is dissociated in Parkinson's disease.

Lindner, J., & Rosén, L. (2006). Decoding of emotion through facial expression, prosody and verbal content in children and adolescents with Asperger's syndrome. *Journal of Autism and Developmental Disorders, 36*, 769-777. doi:10.1007/s10803-006-0105-2

Lindquist, K. A., & Barrett, L. F. (2010). Emotional complexity. In M. Lewis, J. M. Haviland-Jones & L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 513-530). New York: The Guilford Press.

Liu, F., Patel, A. D., Fourcin, A., & Stewart, L. (2010). Intonation processing in congenital amusia: Discrimination, identification and imitation. *Brain, 133*, 1682-1693. doi:10.1093/brain/awq089

Livingstone, S. R., Muhlberger, R., Brown, A. R., & Thompson, W. F. (2010). Changing musical emotion: A computational rule system for modifying score and performance. *Computer Music Journal, 34*, 41-64.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces – KDEF [CD]. Department of Clinical Neuroscience, Psychology section, Karolinska Institutet. ISBN 91-630-7164-9.

Makarova, V., & Petrushin, V. A. (2002, September). *RUSLANA: A database of Russian Emotional Utterances.* Paper session presented at the 7th International Conference on Spoken Language Processing, Denver, Colorado, USA.

Marques, C., Moreno, S., Castro, S. L., & Besson, M. (2007). Musicians detect pitch violations in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience, 19*, 1453-1463. doi:10.1162/jocn.2007.19.9.1453

Martino, A. D., Shehzad, Z., Kelly, C., Roy, A. K., Gee, D. G., Uddin, L. Q., … Milham, M. P. (2009). Relationship between cingulo-insular functional connectivity and autistic traits in neurotypical adults. *American Journal of Psychiatry, 166*(8), 891-899. doi:10.1176/appi.ajp.2009.08121894

Mather, M., & Carstensen, L. (2005). Aging and motivated cognition: The positivity effect in attention and memory. *Trends in Cognitive Sciences, 9*, 496-502. doi:10.1016/j.tics.2005.08.005

Mather, M., & Knight, M. (2005). Goal-directed memory: The role of cognitive control in older adults' emotional memory. *Psychology and Aging, 20*, 554-570. doi:10.1037/0882-7974.20.4.554

Matsumoto, D., & Hwang, H. S. (2011). Evidence for training the ability to read microexpressions of emotion. *Motivation and Emotion, 35*, 181-191. doi:10.1007/s11031-011-9212-2

Matsumoto, D., Keltner, D., Shiota, M. N., O'Sullivan, M., & Frank, M. (2010). Facial expressions of emotion. In M. Lewis, J. M. Haviland-Jones & L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 211-234). New York: The Guilford Press.

Matsumoto, D., LeRoux, J., Wilson-Cohn, C., Raroque, J., Kooken, K., Ekman, P., … Goh, A. (2000). A new test to measure emotion recognition ability: Matsumoto and Ekman's japanese and caucasian brief affect recognition test (JACBART). *Journal of Nonverbal Behavior, 24*, 179-209. doi:10.1023/A:1006668120583

McDermott, J. (2008). The evolution of music. *Nature, 453*, 287-288. doi:10.1038/453287a

McDermott, J. (2009). What can experiments reveal about the origins of music? *Current Directions in Psychological Science, 18*, 164-168. doi:10.1111/j.1467-8721.2009.01629.x

McDonald, C., & Stewart, L. (2008). Uses and functions of music in congenital amusia. *Music Perception, 25*, 345-355. doi:10.1525/MP.2008.25.4.345

Menon, V., & Levitin, D. J. (2005). The rewards of music listening: Response and physiological connectivity of the mesolimbic system. *NeuroImage, 25*, 175-184. doi:10.1016/j.neuroimage.2005.05.053

Meyer, L. (1956). *Emotion and meaning in music*. Chicago: Chicago University Press.

Meyer, M., Steinhauer, K., Alter, K., Friederici, A. D., & Cramon, D. Y. v. (2004). Brain activity varies with modulation of dynamic pitch variance in sentence melody. *Brain and Language, 89*, 277-289. doi:10.1016/S0093-934X(03)00350-X

Milders, M., Fuchs, S., & Crawford, J. (2003). Neuropsychological impairments and changes in emotional and social behaviour following severe traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology, 25*, 157-172. doi:10.1076/jcen.25.2.157.13642

Mill, A., Allik, J., Realo, A., & Valk, R. (2009). Age-related differences in emotion recognition ability: A cross-sectional study. *Emotion, 9*, 619-630. doi:10.1037/a0016562

Mitchell, R. L. C. (2007). Age-related decline in the ability to decode emotional prosody: Primary or secondary phenomenon? *Cognition and Emotion, 21*, 1435-1454. doi:10.1080/02699930601133994

Mitchell, R. L. C., & Bouças, S. B. (2009). Decoding emotional prosody in Parkinson's disease and its potential neuropsychological basis. *Journal of Clinical and Experimental Neuropsychology, 31*, 553-564. doi:10.1080/13803390802360534

Mitchell, R. L. C., Kingston, R. A., & Bouças, S. B. (2011). The specificity of age-related decline in interpretation of emotion cues from prosody. *Psychology and Aging, 26*, 406-414. doi:10.1037/a0021861

Mitchell, R. L. C., & Ross, E. D. (2008). fMRI evidence for the effect of verbal complexity on lateralization of the neural response associated with decoding prosodic emotion. *Neuropsychologia, 46*, 2880-2887. doi:10.1016/j.neuropsychologia.2008.05.024

Mithen, S. (2007). *The singing Neanderthals: The origins of music, language, mind, and body*. Cambridge: Harvard University Press.

Mitterschiffthaler, M. T., Fu, C. H. Y., Dalton, J. A., Andrew, C. M., & Williams, S. C. R. (2007). A functional MRI study of happy and sad affective states induced by classical music. *Human Brain Mapping, 28*, 1150-1162. doi:10.1002/hbm.20337

Mohn, C., Argstatter, H., & Wilker, F. W. (in press). Perception of six basic emotions in music. *Psychology of Music*. Advance online publication. doi:10.1177/0305735610378183

Monnot, M., Nixon, S., Lovallo, W., & Ross, E. (2001). Altered emotional perception in alcoholics: Deficits in affective prosody comprehension. *Alcoholism: Clinical and Experimental Research, 25*, 362-369. doi:10.1111/j.1530-0277.2001.tb02222.x

Montag, C., Reuter, M., & Axmacher, N. (in press). How one's favorite song activates the reward circuitry of the brain: Personality matters! *Behavioural Brain Research*. Advance online publication. doi:10.1016/j.bbr.2011.08.012

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex, 19*, 712-723. doi:10.1093/cercor/bhn120

Mu, Q., Xie, J., Wen, Z., Weng, Y., & Shuyun, Z. (1999). A quantitative MR study of the hippocampal formation, the amygdala, and the temporal horn of the lateral ventricle in healthy subjects 40 to 90 years of age. *American Journal of Neuroradiology, 20*, 207-2011.

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

Nambu, A. (2008). Seven problems on the basal ganglia. *Current Opinion in Neurobiology, 18*, 595-604. doi:10.1016/j.conb.2008.11.001

Nan, Y., Sun, Y., & Peretz, I. (2010). Congenital amusia in speakers of a tone language: Association with lexical tone agnosia. *Brain, 133*, 2635-2642. doi:10.1093/brain/awq178

Nieminen, S., Istók, E., Brattico, E., Tervaniemi, M., & Huotilainen, M. (2011). The development of aesthetic responses to music and their underlying neural and psychological mechanisms. *Cortex, 47*, 1138-1146. doi:10.1016/j.cortex.2011.05.008

Nolte, J. (2002). *The human brain: An introduction to its functional anatomy* (5th ed.). St. Louis: Mosby, Inc.

North, A. C., & Hargreaves, D. J. (2010). Music and marketing. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications*. New York: Oxford University Press.

Obeso, J. A., Marin, C., Rodriguez-Oroz, C., Blesa, J., Benitez-Temiño, B., Mena-Segovia, J., … Olanow, C. W. (2008). The basal ganglia in Parkinson's disease: Current concepts and unexplained observations. *Annals of Neurology, 64*, S30-S46. doi:10.1002/ana.21481

Omar, R., Hailstone, J. C., Warren, J. E., Crutch, S. J., & Warren, J. D. (2010). The cognitive organization of music knowledge: A clinical analysis. *Brain, 133*, 1200-1213. doi:10.1093/brain/awp345

Omar, R., Henley, S., Bartlett, J., Hailstone, J., Gordon, E., Sauter, D., … Warren, J. D. (2011). The structural neuroanatomy of music emotion recognition: Evidence from frontotemporal lobar degeneration. *NeuroImage, 56*, 1814-1821. doi:10.1016/j.neuroimage.2011.03.002

Orbelo, D., Grim, M., Talbott, R., & Ross, E. (2005). Impaired comprehension of affective prosody in elderly subjects is not predicted by age-related hearing loss or age-related cognitive decline. *Journal of Geriatric Psychiatry and Neurology, 18*, 25-32. doi:10.1177/0891988704272214

Orgeta, V. (2010). Effects of age and task difficulty on recognition of facial affect. *Journal of Gerontology: Psychological Sciences, 65B*, 323-327. doi:10.1093/geronb/gbq007

Pantev, C., & Herholz, S. C. (in press). Plasticity of the human auditory cortex related to musical training. *Neuroscience and Biobehavioral Reviews*. Advance online publication. doi:10.1016/j.neubiorev.2011.06.010

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature, 392*, 811-814. doi:10.1038/33918

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear & Hearing, 30*, 653-661. doi:10.1097/AUD.0b013e3181b412e9

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6*, 674-681. doi:10.1038/nn1082

Patel, A. D. (2008a). *Music, language, and the brain*. New York: Oxford University Press.

Patel, A. D. (2008b). A neurobiologial strategy for exploring links between emotion recognition in music and speech. *Behavioral and Brain Sciences, 31*, 589-590. doi:10.1017/S0140525X0800544X

Patel, A. D. (2010). Music, biological evolution, and the brain. In M. Bailar (Ed.), *Emerging disciplines* (pp. 91-144). Texas: Rice University Press.

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology, 2*, 1-14. doi:10.3389/fpsyg.2011.00142

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

Paulmann, S., & Kotz, S. A. (2008b). An ERP investigation on the temporal dynamics of emotional prosody and emotional semantics in pseudo- and lexical-sentence context. *Brain and Language, 105*, 59-69. doi:10.1016/j.bandl.2007.11.005

Paulmann, S., Ott, D. V. M., & Kotz, S. A. (2011). Emotional speech perception unfolding in time: The role of the basal ganglia. *PloS ONE, 6*, e17694. doi:10.1371/journal.pone.0017694

Paulmann, S., & Pell, M. D. (2010a). Contextual influences of emotional speech prosody on face processing: How much is enough? *Cognitive, Affective, & Behavioral Neuroscience, 10*, 230-242. doi:10.3758/CABN.10.2.230

Paulmann, S., & Pell, M. D. (2010b). Dynamic emotion processing in Parkinson's disease as a function of channel availability. *Journal of Clinical and Experimental Neuropsychology, 32*, 822-835. doi:10.1080/13803391003596371

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

Paulmann, S., Pell, M. D., & Kotz, S. A. (2008b). How aging affects the recognition of emotional speech. *Brain and Language, 104*, 262-269. doi:10.1016/j.bandl.2007.03.002

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

Peelen, M. V., Atkinson, A. P., & Vuilleumier, P. (2010). Supramodal representations of perceived emotions in the human brain. *The Journal of Neuroscience, 28*, 10127-10134. doi:10.1523/ JNEUROSCI.2161-10.2010

Pell, M. D. (1998). Recognition of prosody following unilateral brain lesion: Influence of functional and structural attributes of prosodic contours. *Neuropsychologia, 36*, 701-715. doi:10.1016/S0028-3932(98)00008-6

Pell, M. D. (2002). Evaluation of nonverbal emotion in face and voice: Some preliminary findings on a new battery of tests. *Brain and Cognition, 48*, 499-504.

Pell, M. D. (2005). Prosody-face interactions in emotional processing as revealed by the facial affect decision task. *Journal of Nonverbal Behavior, 39*, 193-215. doi:10.1007/s10919-005-7720-z

Pell, M. D. (2006). Cerebral mechanisms for understanding emotional prosody in speech. *Brain and Language, 96*, 221-234. doi:10.1016/j.bandl.2005.04.007

Pell, M. D., Cheang, H. S., & Leonard, C. L. (2006). The impact of Parkinson's disease on vocal-prosodic communication from the perspective of listeners. *Brain and Language, 97*, 123-134. doi:10.1016/j.bandl.2005.08.010

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*, 275-288. doi:10.3758/CABN.3.4.275

Pell, M. D., & Leonard, C. L. (2005). Facial expression decoding in early Parkinson's disease. *Cognitive Brain Research 23*, 327-340. doi:10.1016/j.cogbrainres.2004.11.004

Pell, M. D., Monetta, L., Paulmann, S., & Kotz, S. A. (2009). Recognizing emotions in a foreign language. *Journal of Nonverbal Behavior, 33*, 107-120. doi:10.1007/s10919-008-0065-7

Pell, M. D., Paulmann, S., Dara, C., Alasseri, A., & Kotz, S. A. (2009). Factors in the recognition of vocally expressed emotions: A comparison of four languages. *Journal of Phonetics, 37*, 417-435. doi:10.1016/j.wocn.2009.07.005

Pell, M. D., & Skorup, V. (2008). Implicit processing of emotional prosody in a foreign versus native language. *Speech Communication, 50*, 519-530. doi:10.1016/j.specom.2008.03.006

Peppé, S., Cleland, J., Gibbon, F., O'Hare, A., & Castilla, P. M. (2011). Expressive prosody in children with autism spectrum conditions. *Journal of Neurolinguistics, 24*, 41-53. doi:10.1016/j.jneuroling.2010.07.005

Peretz, I. (2006). The nature of music from a biological perspective. *Cognition, 100*, 1-32. doi:10.1016/j.cognition.2005.11.004

Peretz, I. (2008). Musical disorders: From behavior to genes. *Current Directions in Psychological Science, 17*, 329-333. doi:10.1111/j.1467-8721.2008.00600.x

Peretz, I. (2010). Towards a neurobiology of musical emotions. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of Music and Emotion: Theory, research, applications* (pp. 99-126). Oxford: Oxford University Press.

Peretz, I., Blood, A. J., Penhune, V. B., & Zatorre, R. J. (2001). Cortical deafness to dissonance. *Brain, 124*, 928-940. doi:10.1093/brain/124.5.928

Peretz, I., Champod, A. S., & Hyde, K. L. (2003). Varieties of musical disorders: The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences, 999*, 58-75. doi:10.1196/annals.1284.006

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience, 6*, 688-691. doi:10.1038/nn1083

Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. *Cognition, 68*, 111-141. doi:10.1016/S0010-0277(98)00043-2

Péron, J., Grandjean, D., Jeune, F. L., Sauleau, P., Haegelen, C., Drapier, D., … Vérin, M. (2010). Recognition of emotional prosody is altered after subthalamic nucleus deep brain stimulation in Parkinson's disease. *Neuropsychologia, 48*, 1053-1062. doi:10.1016/j.neuropsychologia.2009.12.003

Petkov, C., Kang, X., Alho, K., Bertrand, O., Yund, E. W., & Woods, D. (2004). Attentional modulation of human auditory cortex. *Nature Neuroscience, 7*, 658-663. doi:10.1038/nn1256

Phan, K. L., Wager, T., Taylor, S., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage, 16*, 331-348. doi:10.1006/nimg.2002.1087

Pinheiro, A. P., Galdo-Álvarez, S., Rauber, A., Sampaio, A., Niznikiewicz, M., & Gonçalves, Ó. F. (2011). Abnormal processing of emotional prosody in Williams syndrome: An event-related potentials study. *Research in Developmental Disabilities, 32*, 133-147. doi:10.1016/j.ridd.2010.09.011

Pinker, S. (1997). *How the mind works*. London: Penguin Books.

Pittam, J., & Scherer, K. R. (1993). Vocal expression and communication of emotion. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (1st ed., pp. 185-198). New York: Guilford Press.

Pons, F., Harris, P. L., & Rosnay, M. d. (2004). Emotion comprehension between 3 and 11 years: Developmental periods and hierarchical organization. *European Journal of Developmental Psychology, 1*, 127-152. doi:10.1080/17405620344000022

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology, 17*, 715-734. doi:10.1017/S0954579405050340

Power, M., & Dalgleish, T. (2008). *Cognition and emotion: From order to disorder* (2nd ed.). Sussex: Psychology Press.

Raz, N., Lindenberger, U., Rodrigue, K., Kennedy, K., Head, D., Williamson, A., … Acker, J. D. (2005). Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cerebral Cortex, 15*, 1676-1689. doi:10.1093/cercor/bhi044

Raz, N., Rodrigue, K. M., Kennedy, K. M., Head, D., Gunning-Dixon, F., & Acker, J. D. (2003). Differential aging of the human striatum: Longitudinal evidence. *American Journal of Neuroradiology, 24*, 1849-1856.

Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology, 84*, 1236-1256. doi:10.1037/0022-3514.84.6.1236

Resnicow, J. E., Salovey, P., & Repp, B. H. (2004). Is recognition of emotion in music performance an aspect of emotional intelligence? *Music Perception, 22*, 145-158.

Ro, T., Friggel, A., & Lavie, N. (2009). Musical expertise modulates the effects of visual perceptual load. *Attention, Perception, & Psychophysics, 71*, 671-674. doi:10.3758/APP.71.4.671

Rohrer, J. D., Sauter, D., Scott, S. K., Rossor, M. N., & Warren, J. D. (in press). Receptive prosody in nonfluent primary progressive aphasias. *Cortex*. Advance online publication. doi:10.1016/j.cortex.2010.09.004

Ross, E. D. (1981). The aprosodias: Functional-anatomic organization of the affective components of language in the right hemisphere. *Archives of Neurology, 38*, 561-569.

Ross, E. D., & Monnot, M. (2008). Neurology of affective prosody and its functional-anatomic organization in right hemisphere. *Brain and Language, 104*, 51-74. doi:10.1016/j.bandl.2007.04.007

Ross, E. D., & Monnot, M. (2011). Affective prosody: What do comprehension errors tell us about hemispheric lateralization of emotions, sex and aging effects, and the role of cognitive appraisal. *Neuropsychologia, 49*, 866-877. doi:10.1016/j.neuropsychologia.2010.12.024

Ross, E. D., Thompson, R. D., & Yenkosky, J. (1997). Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions. *Brain and Language, 56*, 27-54. doi:10.1006/brln.1997.1731

Rowe, J. B., Hughes, L., Ghosh, B. C. P., Eckstein, D., Williams-Gray, C. H., Fallon, S., … Owen, A. M. (2008). Parkinson's disease and dopaminergic therapy - differential effects on movement, reward and cognition. *Brain, 131*, 2094-2105. doi: 10.1093/brain/awn112

Ruffman, T., Halberstadt, J., & Murray, J. (2009). Recognition of facial, auditory, and bodily emotions in older adults. *Journal of Gerontology: Psychological Sciences, 64B*, 696-703. doi:10.1093/geronb/gbp072

Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience and Biobehavioral Reviews, 32*, 863-881. doi:10.1016/j.neubiorev.2008.01.001

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161-1178. doi:10.1037/h0077714

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145-172. doi:10.1037/0033-295X.110.1.145

Russell, J. A., Bachorowski, J.-A., & Fernández-Dols, J. M. (2003). Facial and vocal expressions of emotion. *Annual Review of Psychology, 54*, 329-349. doi:10.1146/annurev.psych.54.101601.145102

Ryan, M., Murray, J., & Ruffman, T. (2010). Aging and the perception of emotion: Processing vocal expressions alone and with faces. *Experimental Aging Research, 36*, 1-22. doi:10.1080/03610730903418372

Rymarczyk, K., & Grabowska, A. (2007). Sex differences in brain control of prosody. *Neuropsychologia, 45*, 921-930. doi:10.1016/j.neuropsychologia.2006.08.021

Salimpoor, V., Benovoy, M., Larcher, K., Dagher, A., & Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience, 14*, 257-262. doi:10.1038/nn.2726

Salimpoor, V. N., Benovoy, M., Longo, G., Cooperstock, J. R., & Zatorre, R. J. (2010). The rewarding aspects of music listening are related to degree of emotional arousal. *PLoS One, 4*, e7487. doi:10.1371/journal.pone.0007487

Salthouse, T. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging, 30*, 507-514. doi:10.1016/j.neurobiolaging.2008.09.023

Samanez-Larkin, G. R., & Carstensen, L. (2011). Socioemotional functioning and the aging brain. In J. Decety & J. T. Cacioppo (Eds.), *The handbook of social neuroscience*. New York: Oxford University Press.

Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). Emotion and attention interactions in social cognition: Brain regions involved in processing anger prosody. *NeuroImage, 28*, 848-858. doi:10.1016/j.neuroimage.2005.06.023

Särkämö, T., Tervaniemi, M., Laitinen, S., Forsblom, A., Soinila, S., Mikkonen, M., … Hietanen, M. (2008). Music listening enhances cognitive recovery and modd after middle cerebral artery stroke. *Brain, 131*, 866-876. doi: 10.1093/brain/awn013

Sauter, D. (2006). *An investigation into vocal expressions of emotions: The roles of valence, culture, and acoustic factors* (Unpublished doctoral dissertation). University College London, London.

Sauter, D., & Eimer, M. (2009). Rapid detection of emotion from human vocalizations. *Journal of Cognitive Neuroscience, 22*, 474-481. doi:10.1162/jocn.2009.21215

Sauter, D., Guen, O. L., & Haun, D. B. M. (in press). Categorical perception of emotional expressions does not require lexical categories. *Emotion*.

Sauter, D. A., Eisner, F., Calder, A. J., & Scott, S. K. (2010). Perceptual cues in nonverbal vocal expressions of emotion. *The Quarterly Journal of Experimental Psychology, 63*, 2251-2272. doi:10.1080/17470211003721642

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

Sauter, D. A., & Scott, S. K. (2007). More than one kind of happiness: Can we recognize vocal expressions of different positive states? *Motivation and Emotion, 31*, 192-199. doi: 10.1007/s11031-007-9065-x

Schellenberg, E. G. (2005). Music and cognitive abilities. *Current Directions in Psychological Science, 14*, 317-320. doi: 10.1111/j.0963-7214.2005.00389.x

Schellenberg, E. G. (2006). Long-term positive associations between music lessons and IQ. *Journal of Educational Psychology, 98*, 457-468. doi:10.1037/0022-0663.98.2.457

Schellenberg, E. G., & Moreno, S. (2009). Music lessons, pitch processing, and g. *Psychology of Music, 38,* 209-221. doi: 10.1177/0305735609339473

Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin, 99*, 143-165. doi: 10.1037/0033-2909.99.2.143

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice, 9*, 235-248. doi:10.1016/S0892-1997(05)80231-0

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

Scherer, K. R. (2004). Which emotions can be induced by music? What are the underlying mechanisms? And how can we measure them? *Journal of New Music Research, 33*, 239-251. doi:10.1080/0929821042000317822

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information, 44*, 695-729. doi: 10.1177/0539018405058216

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology, 32*, 76-92. doi: 10.1177/0022022101032001009

Scherer, K. R., Banse, R., Wallbott, H. G., & Goldbeck, T. (1991). Vocal cues in emotion ecoding and decoding. *Motivation and Emotion, 15*, 123-148. doi:10.1007/BF00995674

Scherer, K. R., Johnstone, T., & Klasmeyer, G. (2003). Vocal expression of emotion. In R. Davidson, K. R. Scherer & H. Goldsmith (Eds.), *Handbook of the affective sciences* (pp. 433-456). New York: Oxford University Press.

Scherer, K. R., & Scherer, U. (in press). Assessing the ability to recognize facial and vocal expressions of emotion: Construction and validation of the emotion recognition index. *Journal of Nonverbal Behavior*. Advance online publication. doi:10.1007/s10919-011-0115-4

Schiess, M. C., Zheng, H., Soukup, V. M., Bonnen, J. G., & Nauta, H. J. W. (2000). Parkinson's disease subtypes: Clinical classification and ventricular cerebrospinal fluid analysis. *Parkinsonism and Related Disorders, 6*, 69-76. doi:10.1016/S1353-8020(99)00051-6

Schirmer, A., Escoffier, N., Zysset, S., Koester, D., Striano, T., & Friederici, A. D. (2008). When vocal processing gets emotional: On the role of social orientation in relevance detection by the human amygdala. *NeuroImage, 40*, 1402-1410. doi:10.1016/j.neuroimage.2008.01.018

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

Schirmer, A., Kotz, S. A., & Friederici, A. D. (2002). Sex differentiaties the role of emotional prosody during word processing. *Cognitive Brain Research, 14*, 228-233. doi:10.1016/S0926-6410(02)00108-8

Schirmer, A., Kotz, S. A., & Friederici, A. D. (2005). On the role of attention for the processing of emotions in speech: Sex differences revisited. *Cognitive Brain Research, 24*, 442-452. doi:10.1016/S0926-6410(02)00108-8

Schirmer, A., Striano, T., & Friederici, A. D. (2005). Sex differences in the preattentive processing of vocal emotional expressions. *NeuroReport, 16*, 635-639. doi:10.1016/j.cogbrainres.2005.02.022

Schirmer, A., Zysset, S., Kotz, S. A., & Cramon, D. Y. v. (2004). Gender differences in the activation of inferior frontal cortex during emotional speech perception. *NeuroImage, 21*, 1114-1123. doi:10.1016/j.neuroimage.2003.10.048

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*, 2869-2877. doi:10.1016/j.neuropsychologia.2010.05.029

Schön, D., & Besson, M. (2001). Comparison between language and music. *Annals of the New York Academy of Sciences, 930*, 232-258. doi: 10.1111/j.1749-6632.2001.tb05736.x

Schön, D., Gordon, R., Campagne, A., Magne, C., Astésamo, C., Anton, J. L., & Besson, M. (2010). Similar cerebral networks in language, music and song perception. *NeuroImage, 51*, 450-461. doi:10.1016/j.neuroimage.2010.02.023

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology, 41*, 341-349. doi:10.1111/1469-8986.00172.x

Schön, D., Ystad, S., Kronland-Martinet, R., & Besson, M. (2010). The evocative power of sounds: Conceptual priming between words and nonverbal sounds. *Journal of Cognitive Neuroscience, 22*, 1026-1035. doi:10.1162/jocn.2009.21302

Schott, B. H., Niehaus, L., Wittmann, B. C., Schutze, H., Seidenbecher, C. I., Heinze, H.-J., & Düzel, E. (2007). Ageing and early-stage Parkinson's disease affect separable neural mechanisms of mesolimbic reward processing. *Brain, 130*, 2412-2424. doi:10.1093/brain/awm147

Schröder, C., Nikolova, Z. T., & Dengler, R. (2010). Changes of emotional prosody in Parkinson's disease. *Journal of the Neurological Sciences*, *289,* 32-35. doi:10.1016/j.jns.2009.08.038

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. S. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain, 123*, 2400-2406. doi: 10.1093/brain/123.12.2400

Scott, S., Caird, F., & Williams, B. (1984). Evidence for an apparent sensory speech disorder in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry, 47*, 840-843. doi:10.1136/jnnp.47.8.840

Scott, S. K., Sauter, D., & McGettigan, C. (2010). Brain mechanisms for processing perceived emotional vocalizations in humans. In M. B. Stefan (Ed.), *Handbook of Behavioral Neuroscience* (Vol. 19, pp. 187-197). London: Academic Press.

Silver, H., Goodman, C., Knoll, G., & Isakov, V. (2004). Brief emotion training improves recognition of facial emotions in chronic schizophrenia: A pilot study. *Psychiatry Research, 128*, 147-154. doi:10.1016/j.psychres.2004.06.002

Simões, M., Firmino, H., Vilar, M., & Martins, M. (2007). Montreal Cognitive Assessment (MOCA) - Versão experimental portuguesa [portuguese experimental version]. Retrieved from http://www.mocatest.org

Sloboda, J. A., & Juslin, P. N. (2010). At the interface between the inner and outer world: Psychological perspectives. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of Music and Emotion: Theory, Research, Applications* (pp. 73-97). New York: Oxford University Press.

Smith, C. A., & Lazarus, R. S. (1990). Emotion and adaptation. In L. Pervin (Ed.), *Handbook of personality: Theory and research* (1st ed.). New York: Guilford Press.

Snowdon, C. T., & Teie, D. (2010). Affective responses in tamarins elicited by species-specific music. *Biology Letters, 6*, 30-32. doi:10.1098/rsbl.2009.0593

Speedie, L., Brake, N., Folstein, S., Bowers, D., & Heilman, K. (1990). Comprehension of prosody in Huntington's disease. *Journal of Neurology, Neurosurgery, and Psychiatry, 53*, 607-610. doi:10.1136/jnnp.53.7.607

Spencer, H. (1857). The origin and function of music. *Fraser's Magazine, 56*, 396-408.

St. Jacques, P., Bessette-Symons, B., & Cabeza, R. (2009). Functional neuroimaging studies of aging and emotion: Fronto-amygdalar differences during emotional perception and episodic memory. *Journal of the International Neuropsychological Society, 15*, 819-825. doi:10.1017/S1355617709990439

St. Jacques, P., Dolcos, F., & Cabeza, R. (2008). Effects of aging on functional connectivity of the amygdala during negative evaluation: A network analysis of fMRI data. *Neurobiology of Aging, 31*, 315-327. doi:10.1016/j.neurobiolaging.2008.03.012

Staroniewicz, P., & Majewski, W. (2009). Polish emotional speech database - recording and preliminary validation. *Lectures Notes in Computer Science, 5641*, 42-49. doi:10.1007/978-3-642-03320-9_5

Stewart, L., Kriegstein, K. v., Dalla-Bella, S., Warren, J. D., & Griffiths, T. D. (2008). Disorders of musical cognition. In S. Hallam, I. Cross & M. H. Thaut (Eds.), *Oxford Handbook of Music Psychology* (pp. 184-196). New York: Oxford University Press.

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

Sullivan, S., & Ruffman, T. (2004). Social understanding: How does it fare with advancing years? *British Journal of Psychology, 95*, 1-18. doi:10.1348/000712604322779424

Suzuki, A., Hoshino, T., Shigemasu, K., & Kawamura, M. (2006). Disgust-specific impairment of facial expression recognition in Parkinson's disease. *Brain, 129*, 707-717. doi: 10.1093/brain/awl011

Taler, V., Baum, S., Chertkow, H., & Saumier, D. (2008). Comprehension of grammatical and emotional prosody is impaired in Alzheimer's disease. *Neuropsychology, 22*, 188-195. doi:10.1037/0894-4105.22.2.188

Tavares, G. M. (2004). *Poesia 1* [Poetry 1]. Lisboa: Relógio D'Água Editores.

Thaut, M. H. (2005). The future of music in therapy and medicine. *Annals of the New York Academy of Sciences, 1060*, 303-308. doi:10.1196/annals.1360.023

Thaut, M. H., Gardiner, J. C., Holmberg, D., Horwitz, J., Kent, L., Andrews, G., … McIntosh, G. R. (2009). Neurologic music therapy improves executive function and emotional adjustment in traumatic brain injury rehabilitation. *Annals of the New York Academy of Sciences 1169*, 406-416. doi:10.1111/j.1749-6632.2009.04585.x

Thaut, M. H., & Wheeler, B. (2010). Music therapy. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications*. New York: Oxford University Press.

Thompson, W. F. (2009). *Music, though, and feeling: Understanding the psychology of music*. New York: Oxford University Press.

Thompson, W. F., & Balkwill, L.-L. (2006). Decoding speech prosody in five languages. *Semiotica, 158*, 407-424. doi:10.1515/SEM.2006.017

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

Tracy, J. L., & Robins, R. W. (2008). The automaticity of emotion recognition. *Emotion, 8*, 81-95. doi:10.1037/1528-3542.8.1.81

Trauner, D., Ballantyne, A., Friedland, S., & Chase, C. (1996). Disorders of affective and linguistic prosody in children after early unilateral brain damage. *Annals of Neurology, 39*, 361-367. doi:10.1002/ana.410390313

Trehub, S. E., & Hannon, E. E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition, 100*, 73-99. doi:10.1016/j.cognition.2005.11.006

Trenerry, M. R., Crosson, B., Deboe, J., & Leber, W. R. (1995). *The Stroop Neuropsychological Screening Test Manual*. Tampa, Florida: Psychological Assessment Resources.

Tricht, M. J. v., Smeding, H. M. M., Speelman, J. D., & Schmand, B. A. (2010). Impaired emotion recognition in music in Parkinson's disease. *Brain and Cognition, 74*, 58-65. doi:10.1016/j.bandc.2010.06.005

Trimmer, C. G., & Cuddy, L. L. (2008). Emotional intelligence, not music training, predicts recognition of emotional speech prosody. *Emotion, 8*, 838-849. doi:10.1037/a0014080

Uekermann, J., Abdel-Hamid, M., Lehmkämper, C., Vollmoeller, W., & Daum, I. (2008). Perception of affective prosody in major depression: A link to executive functions? *Journal of the International Neuropsychological Society, 14*, 552-561. doi:10.10170S1355617708080740

Utter, A. A., & Basso, M. A. (2008). The basal ganglia: An overview of circuits and function. *Neuroscience and Biobehavioral Reviews, 32*, 333-342. doi:10.1016/j.neubiorev.2006.11.003

Vieillard, S., Peretz, I., Gosselin, N., Khalfa, S., Gagnon, L., & Bouchard, B. (2008). Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cognition and Emotion, 22*, 720-752. doi:10.1080/02699930701503567

Vuoskoski, J. K., & Eerola, T. (2011). The role of mood and personality in the perception of emotions represented by music. *Cortex, 47*, 1099-1106. doi:10.1016/j.cortex.2011.04.011

Vytal, K., & Hamann, S. (2010). Neuroimaging support for discrete neural correlates of basic emotions: A voxel-based meta-analysis. *Journal of Cognitive Neuroscience, 22*, 2864-2885. doi:10.1162/jocn.2009.21366

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

Walhovd, K. B., Fjell, A. M., Reinvang, I., Lundervold, A., Dale, A. M., Eilertsen, D. E., … Fischl, B. (2005). Effects of age on volumes of cortex, white matter and subcortical structures. *Neurobiology of Aging, 26*, 1261-1270. doi:10.1016/j.neurobiolaging.2005.05.020

Watson, G. S., & Leverenz, J. B. (2010). Profile of cognitive impairment in Parkinson's disease. *Brain Pathology, 20*, 640-645. doi:10.1111/j.1750-3639.2010.00373.x

Wechsler, D. (2008). *WAIS -III - Escala de Inteligência de Wechsler para Adultos - 3.ª edição [portuguese version of the Wechsler Adult Intelligence Scale - Third Edition, WAIS-III]*. Lisbon: CEGOC-TEA.

Whitman, W. (1855/2005). *Leaves of grass*. New York: Oxford University Press.

Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., & Ethofer, T. (2008). Cerebral processing of emotional prosody: Influence of acoustic parameters and arousal. *NeuroImage, 39*, 885-893. doi:10.1016/j.neuroimage.2007.09.028

Wildgruber, D., Ackermann, H., Kreifelts, B., & Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies. *Progress in Brain Research, 156*, 249-268. doi:10.1016/S0079-6123(06)56013-3

Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., & Ackermann, H. (2005). Identification of emotional intonation evaluated by fMRI. *NeuroImage, 24*, 1233-1241. doi:10.1016/j.neuroimage.2004.10.034

Williams, L. M., Brown, K. J., Palmer, D., Liddell, B. J., Kemp, A. H., Olivieri, G., … Gordon, E. (2006). The mellow years?: Neural basis of improving emotional stability over age. *The Journal of Neuroscience, 26*, 6422-6430. doi:10.1523/JNEUROSCI.0022-06.2006

Williams, L. M., Mathersul, D., Palmer, D. M., Gur, R. C., Gur, R. E., & Gordon, E. (2009). Explicit identification and implicit recognition of facial emotions: I. Age effects in males and females across 10 decades. *Journal of Clinical and Experimental Neuropsychology, 31*, 257-277. doi:10.1080/13803390802255635

Wittfoth, M., Schröder, C., Schardt, D. M., Dengler, R., Heinze, H.-J., & Kotz, S. A. (2010). On emotional conflict: Interference resolution of happy and angry prosody reveals valence-specific effects. *Cerebral Cortex, 20*, 383-392. doi:10.1093/cercor/bhp106

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

Wu, T., Yang, Y., Wu, Z., & Li, D. (2006, June). *MASC: A speech corpus in Mandarin for emotion analysis and affective speaker recognition.* Paper session presented at the IEEE Odyssey 2006 Workshop on speaker and language recognition, San Juan, Puerto Rico.

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research, 17*, 37-49. doi:10.1016/0022-3956(82)90033-4

Yip, J., Lee, T., Ho, S.-L., Tsang, K.-L., & Li, L. (2003). Emotion recognition in patients with idiopathic Parkinson's disease. *Movement Disorders, 18*, 1115-1122. doi:10.1002/mds.10497

Zatorre, R. J. (2001). Neural specializations for tonal processing. *Annals of the New York Academy of Sciences, 930*, 193-210. doi:10.1111/j.1749-6632.2001.tb05734.x

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences, 6*, 37-46. doi:10.1016/S1364-6613(00)01816-7

Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences, 107*, 5768-5773. doi:10.1073/pnas.1000121107

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion, 8*, 494-521. doi:10.1037/1528-3542.8.4.494

Zentner, M., & Kagan, J. (1996). Perception of music by infants. *Nature, 383*, 29. doi:10.1038/383029a0

APPENDICES

1. Raw ratings for each musical emotion in Study 1, Chapter VI

Raw ratings (minimum 0, maximum 9) for the intended and non-intended emotions, as a function of age group. Diagonal cells in bold indicate ratings on the excerpts' intended emotion. Standard errors are presented in parentheses.

Age group / excerpt type	Ratings across response categories			
	Happy	Peaceful	Sad	Scary
Younger				
Happy	**7.8 (0.1)**	1.5	0.2	0.2
Peaceful	2.8	**5.3 (0.3)**	4.4	0.3
Sad	0.4	3.6	**7.3 (0.2)**	1.8
Scary	1.2	0.8	2.9	**6.3 (0.2)**
Middle-aged				
Happy	**7.7 (0.2)**	2.2	0.4	0.2
Peaceful	2.4	**6.4 (0.2)**	3.2	0.2
Sad	0.9	4.3	**6.2 (0.2)**	0.7
Scary	1.6	1.1	3	**4.2 (0.4)**
Older				
Happy	**7.6 (0.2)**	1.9	0.4	0.2
Peaceful	2.2	**5.9 (0.3)**	3.1	0.2
Sad	0.8	5	**5.5 (0.3)**	0.5
Scary	2	1.9	3	**2.8 (0.3)**

2. Unbiased accuracy rates for each musical emotion in Study 1, Chapter VI

Hu for each musical emotion, as a function of age group. Values vary between 0 and 1. *Zero* indicates that none of the stimuli of a given category was identified correctly, and that such response category was never used correctly; *one* indicates that all the stimuli of a given category were identified correctly, and that such response category was always used correctly. Standard errors are presented in parentheses.

Age group / excerpt type	*Hu* scores
Younger	
Happy	.8 (0.02)
Peaceful	.4 (0.04
Sad	.5 (0.03)
Scary	.7 (0.04)
Middle-aged	
Happy	.8 (0.02)
Peaceful	.5 (0.04)
Sad	.4 (0.05)
Scary	.4 (0.05)
Older	
Happy	.7 (0.03)
Peaceful	.3 (0.03)
Sad	.3 (0.03)
Scary	.3 (0.04)

3. Structural characteristics of the musical excerpts used in Study 1 and Study 2, Chapter VI

Structural characteristics of the excerpts used in the Study 1 (all, n = 56) and Study 2 (coloured in grey, n = 40). Note: Tempo corresponds to crochet beats per minute; Mode, 1 = major, 2 = minor, 3 = indefinite; Pedal, 0 = without, 1 = with; Tone Dens. = Tone density (number of melodic events / total duration); M.P. Range = Melodic pitch range (number of semitones between the highest and the lowest tone of the melody); Diss. = Dissonance (minimum 1, maximum 5); Un. = Unexpected events (minimum 1, maximum 5); Rhy. Irr. = Rhythmic Irregularities (minimum 1, maximum 5).

#	Emotion	Stimulus	Tempo	Mode	Pedal	Tone Dens.	M.P. Range	Diss.	Un.	Rhy. Irr.
1	Happy	G01	112	1	0	3.4	17	1	1	1
2	Happy	G02	143	1	0	4.4	21	1	1	1
3	Happy	G03	91	1	0	3.7	18	1	1	1
4	Happy	G04	180	1	0	5.0	20	1	1	1
5	Happy	G05	100	1	0	5.5	22	1	1	1
6	Happy	G06	195	1	0	5.2	17	1	1	1
7	Happy	G07	180	1	0	3.2	14	1	1	1
8	Happy	G08	143	1	0	3.7	21	1	1	1
9	Happy	G09	143	1	0	2.5	10	1	1	1
10	Happy	G10	160	1	0	3.8	12	1	1	1
11	Happy	G11	140	1	0	4.5	19	1	1	1
12	Happy	G12	107	1	0	2.6	12	1	1	1
13	Happy	G13	120	1	0	3.3	19	1	1	1
14	Happy	G14	112	1	0	5.7	19	1	1	1
16	Sad	T01	40	2	1	0.8	5	1	1	1
17	Sad	T02	40	2	1	0.7	14	1	1	1
18	Sad	T03	40	2	1	0.4	12	1	1	1
19	Sad	T04	54	2	1	0.5	9	1	1	1
20	Sad	T05	60	2	1	0.9	5	1	1	1
21	Sad	T06	54	2	1	0.6	8	1	1	1
22	Sad	T07	42	2	1	0.8	10	1	1	1

15	Sad	T08	40	2	1	0.6	7	1	1	1
23	Sad	T09	54	2	1	1.3	8	1	1	1
24	Sad	T10	40	2	1	0.58	3	1	1	1
25	Sad	T11	48	2	1	0.8	10	1	1	1
26	Sad	T12	60	2	1	0.5	9	1	1	1
27	Sad	T13	48	2	1	0.7	7	1	1	1
28	Sad	T14	48	2	1	1.0	10	1	1	1
29	Scary	P01	100	2	0	3.6	4	2	3	3
30	Scary	P02	140	2	0	2.2	7	2	2	1
31	Scary	P03	100	2	0	1.4	18	2	4	3
32	Scary	P04	100	2	0	0.9	2	4	5	4
33	Scary	P05	100	2	0	3.4	33	2	1	2
34	Scary	P06	96	2	0	0.9	43	5	2	2
35	Scary	P07	100	2	0	1.6	13	2	3	4
36	Scary	P08	96	2	0	4.3	19	2	3	4
37	Scary	P09	44	2	0	0.7	7	1	1	1
38	Scary	P10	120	2	0	1.1	12	2	4	3
39	Scary	P11	44	3	0	1.4	4	2	5	4
40	Scary	P12	75	2	0	1.9	28	2	5	2
41	Scary	P13	72	2	0	2.4	8	2	4	1
42	Scary	P14	172	2	0	3.1	36	1	3	4
43	Peaceful	A01	60	1	1	0.4	12	1	1	1
44	Peaceful	A02	72	1	1	0.9	20	1	1	1
45	Peaceful	A03	72	1	1	0.7	9	1	1	1
46	Peaceful	A04	54	1	1	0.8	13	1	1	1
47	Peaceful	A05	69	1	1	0.7	10	1	1	1
48	Peaceful	A06	33	1	1	1.4	13	1	1	1
49	Peaceful	A07	69	1	1	0.9	5	1	1	1
50	Peaceful	A08	80	1	1	1.2	15	1	1	1
51	Peaceful	A09	72	1	1	1.1	12	1	1	1
52	Peaceful	A10	80	1	1	2.0	24	1	1	1
53	Peaceful	A11	88	1	1	1.6	14	1	1	1
54	Peaceful	A12	75	1	1	0.9	12	1	1	1
55	Peaceful	A13	75	1	1	1.1	6	1	1	1
56	Peaceful	A14	72	1	1	1.6	17	1	1	1

4. Raw ratings for each musical emotion in Study 2, Chapter VI

Raw ratings (minimum 0, maximum 9) for the intended and non-intended emotions, as a function of age and expertise groups. Diagonal cells in bold indicate ratings on the excerpts' intended emotion. Standard errors are presented in parentheses.

Age group/ excerpt type	Ratings across response categories							
	Controls				Musicians			
	Happy	Peaceful	Sad	Scary	Happy	Peaceful	Sad	Scary
Younger								
Happy	**7.6 (0.2)**	1.6	0.2	0.2	**7.2 (0.4)**	1.7	0.2	0.2
Peaceful	2.8	**5.5 (0.4)**	4.6	0.2	4.3	**4.9 (0.5)**	3	0
Sad	0.4	3.3	**7.4 (0.2)**	2.1	0.4	4.6	**7 (0.3)**	0.9
Scary	1	0.4	3.7	**7.1 (0.3)**	0.8	1.2	3.5	**5.2 (0.5)**
Middle-aged								
Happy	**6.7 (0.5)**	2.1	0.6	0.4	**7.4 (0.3)**	1.6	0.6	0.6
Peaceful	3.9	**5 (0.5)**	2.7	0.4	2.9	**5.2 (0.4)**	3.2	0.4
Sad	1.5	3.9	**5.5 (0.5)**	0.5	0.7	4.4	**6 (0.3)**	1.3
Scary	1.8	0.6	2.9	**3.7 (0.4)**	1.5	0.8	2.8	**5.3 (0.4)**

5. Unbiased accuracy rates for each musical emotion in Study 2, Chapter VI

Hu for each musical emotion, as a function of age and expertise groups. Values vary between 0 and 1. Standard errors are presented in parentheses.

Age group / excerpt type	*Hu* scores	
	Controls	Musicians
Younger		
Happy	.8 (0.03)	.8 (0.04)
Peaceful	.4 (0.07)	.4 (0.07)
Sad	.6 (0.05)	.5 (0.06)
Scary	.7 (0.04)	.7 (0.07)
Middle-aged		
Happy	.7 (0.05)	.8 (0.03)
Peaceful	.4 (0.05)	.5 (0.05)
Sad	.4 (0.06)	.5 (0.05)
Scary	.5 (0.06)	.7 (0.05)

6. Acoustic and perceptual characteristics of the prosodic stimuli presented in Chapter VII

Acoustic (duration, F_0 mean, F_0 *SD*) and perceptual (recognition accuracy, intensity) characteristics of each sentence and pseudo-sentence included in the database of prosodic stimuli.

Sentences

In Stimulus, *s* refers to sentence, *p* to pseudo-sentence, *A* to one speaker, and *B* to the other speaker.

#	Stimulus	Content	Duration (ms)	F_0 (Hz)	F_0 variability (*SD*, Hz)	Accuracy (%)	Intensity (1-7)
1	1sA_angry1	estaMesa	1278	339	99	75	4.5
2	2sA_angry2	oRadio	1245	323	81	80	5
3	3sA_angry3	aqueleLivro	1367	347	85	95	5.8
4	4sA_angry4	aTerra	1306	355	68	70	4.9
5	5sA_angry5	oCao	1166	299	99	80	3.6
6	6sA_angry6	eleChega	1066	370	95	75	5.3
7	7sA_angry7	estaRoupa	1224	349	85	80	4.4
8	8sA_angry8	osJardins	1415	333	75	70	3.5
9	9sA_angry9	asPessoas	1712	358	86	70	4.9
10	10sA_angry10	haArvores	1426	368	65	75	5.3
11	11sA_angry11	osTigres	1572	279	76	75	4.8
12	12sA_angry12	oQuadro	1301	371	78	80	5.9
13	13sA_angry13	alguemFechou	1455	336	93	95	5.7
14	14sA_angry14	osJovens	1477	338	80	95	5.7
15	15sA_angry15	oFutebol	1487	377	82	70	4.9
16	16sA_angry16	elaViajou	1323	329	93	85	5.2
17	17sB_angry17	estaMesa	1116	266	53	75	5.3
18	18sB_angry18	oRadio	1079	289	56	95	5.9
19	19sB_angry19	aqueleLivro	1186	284	51	55	4.6
20	20sB_angry20	aTerra	1068	288	53	80	5.6
21	21sB_angry21	oCao	1059	278	65	75	4.8
22	22sB_angry22	eleChega	897	252	41	70	5.2

23	23sB_angry23	estaRoupa	1127	305	54	55	5.5
24	24sB_angry24	osJardins	1160	283	47	90	6.1
25	25sB_angry25	asPessoas	1378	324	47	70	5.5
26	26sB_angry26	haArvores	1202	293	61	85	6.1
27	27sB_angry27	osTigres	1387	281	56	55	5.2
28	28sB_angry28	oQuadro	1200	262	60	95	4.7
29	29sB_angry29	alguemFechou	1296	305	52	85	6.1
30	30sB_angry30	osJovens	1198	302	48	75	6.3
31	31sB_angry31	oFutebol	1215	264	24	65	4.6
32	32sB_angry32	elaViajou	1137	264	42	80	5.9
33	33sA_disgust1	oRadio	1661	270	76	60	5.3
34	34sA_disgust2	oCao	1538	263	69	45	4.2
35	35sA_disgust3	haArvores	1596	289	76	55	4.5
36	36sA_disgust4	osTigres	1908	279	77	45	5.2
37	37sA_disgust5	oQuadro	1800	286	78	60	4.3
38	38sA_disgust6	alguemFechou	1852	281	67	45	5.4
39	39sB_disgust7	oCao	1376	272	62	50	5.6
40	40sB_disgust8	estaRoupa	1317	266	64	45	4.9
41	41sB_disgust9	asPessoas	1875	289	75	60	5
42	42sB_disgust10	haArvores	1545	294	54	45	5.8
43	43sB_disgust11	oQuadro	1530	287	57	50	5.1
44	44sB_disgust12	oFutebol	1608	278	51	45	5
45	45sA_fear1	estaMesa	1687	276	28	55	5.4
46	46sA_fear2	oRadio	1501	336	45	75	4.5
47	47sA_fear3	aqueleLivro	1488	310	43	55	3.9
48	48sA_fear4	eleChega	1803	290	15	55	4.5
49	49sA_fear5	osJardins	2104	313	33	50	5.1
50	50sA_fear6	asPessoas	1772	310	45	65	4.4
51	51sA_fear7	haArvores	1954	291	38	80	5.2
52	52sA_fear8	osTigres	1659	290	35	55	4
53	53sA_fear9	oQuadro	2009	309	34	65	4.7
54	54sA_fear10	alguemFechou	1759	290	58	75	5.6
55	55sA_fear11	oFutebol	1645	255	60	50	3.4
56	56sA_fear12	elaViajou	1173	287	21	55	4.7
57	57sB_fear13	estaMesa	1086	300	28	85	5.6
58	58sB_fear14	oRadio	1129	281	34	80	5.2
59	59sB_fear15	aqueleLivro	1125	308	33	75	5.2
60	60sB_fear16	oCao	1063	323	54	55	5.7
61	61sB_fear17	eleChega	913	326	27	60	5.1
62	62sB_fear18	estaRoupa	1173	311	40	45	4.1
63	63sB_fear19	osJardins	1212	325	45	60	4.4
64	64sB_fear20	haArvores	1245	290	30	75	5.2
65	65sB_fear21	osTigres	1400	305	41	80	5.3
66	66sB_fear22	oQuadro	1200	302	27	60	4.6
67	67sB_fear23	alguemFechou	1320	311	37	65	5.2
68	68sB_fear24	oFutebol	1300	314	41	70	5.1
69	69sA_happy1	estaMesa	1451	360	83	85	4.8
70	70sA_happy2	oRadio	1436	386	94	75	4.3
71	71sA_happy3	aqueleLivro	1588	379	102	90	4.7
72	72sA_happy4	oCao	1368	377	98	75	5.5
73	73sA_happy5	eleChega	1207	385	68	100	6.1

74	74sA_happy6	estaRoupa	1321	373	93	90	4.7
75	75sA_happy7	osJardins	1628	376	84	90	5.9
76	76sA_happy8	asPessoas	1772	391	85	90	5.7
77	77sA_happy9	haArvores	1574	360	76	75	4.9
78	78sA_happy10	osTigres	1700	385	102	55	4.3
79	79sA_happy11	oQuadro	1461	373	85	95	4.7
80	80sA_happy12	alguemFechou	1702	351	75	90	5.9
81	81sA_happy13	osJovens	1623	355	77	75	4.7
82	82sA_happy14	oFutebol	1647	371	83	90	5.7
83	83sA_happy15	elaViajou	1592	375	83	90	4.6
84	84sB_happy16	oRadio	1348	338	59	50	5.3
85	85sB_happy17	aqueleLivro	1370	311	69	60	5.3
86	86sB_happy18	aTerra	1271	356	81	65	5.6
87	87sB_happy19	oCao	1286	341	79	60	6.1
88	88sB_happy20	eleChega	1178	376	94	65	6.3
89	89sB_happy21	estaRoupa	1402	348	83	60	5.5
90	90sB_happy22	osJardins	1611	346	87	75	5.9
91	91sB_happy23	asPessoas	1660	320	79	75	5.2
92	92sB_happy24	haArvores	1470	345	93	70	5.8
93	93sB_happy25	oQuadro	1451	353	87	45	4.8
94	94sB_happy26	alguemFechou	1551	348	82	90	5.9
95	95sB_happy27	osJovens	1547	335	76	70	5.4
96	96sB_happy28	oFutebol	1368	332	80	60	5.2
97	97sB_happy29	elaViajou	1296	321	70	60	5.4
98	98sA_neutral1	estaMesa	1459	210	33	100	4.9
99	99sA_neutral2	oRadio	1280	213	54	95	4.7
100	100sA_neutral3	aqueleLivro	1536	208	51	100	4.4
101	101sA_neutral4	aTerra	1555	204	29	90	5.5
102	102sA_neutral5	oCao	1433	196	32	95	5.2
103	103sA_neutral6	eleChega	1088	203	25	90	4.9
104	104sA_neutral7	estaRoupa	1483	205	33	95	4.8
105	105sA_neutral8	osJardins	1575	239	79	80	4.6
106	106sA_neutral9	asPessoas	1792	217	56	95	5.3
107	107sA_neutral10	haArvores	1481	205	33	90	4.3
108	108sA_neutral11	osTigres	1842	196	28	100	4.6
109	109sA_neutral12	oQuadro	1438	194	26	95	4.7
110	110sA_neutral13	alguemFechou	1576	217	71	75	4.4
111	111sA_neutral14	osJovens	1795	208	70	100	5.5
112	112sA_neutral15	oFutebol	1838	223	84	90	5
113	113sA_neutral16	elaViajou	1358	191	21	95	5.2
114	114sB_neutral17	estaMesa	1609	215	24	100	5
115	115sB_neutral18	aqueleLivro	1282	221	40	85	4.5
116	116sB_neutral19	aTerra	1379	205	19	85	5.1
117	117sB_neutral20	oCao	1190	200	20	80	4.5
118	118sB_neutral21	eleChega	952	209	20	100	4.6
119	119sB_neutral22	estaRoupa	1294	218	23	90	4.7
120	120sB_neutral23	asPessoas	1671	215	20	70	5
121	121sB_neutral24	haArvores	1400	208	16	80	4.6
122	122sB_neutral25	osTigres	1682	209	54	80	4.1
123	123sB_neutral26	oQuadro	1361	221	53	70	3.3
124	124sB_neutral27	alguemFechou	1631	205	17	80	4.9

125	125sB_neutral28	osJovens	1438	212	22	65	4.7
126	126sB_neutral29	oFutebol	1603	236	74	60	4.7
127	127sB_neutral30	elaViajou	1303	225	26	95	4.5
128	128sA_sad1	estaMesa	1756	197	22	90	5.4
129	129sA_sad2	aqueleLivro	1756	203	31	85	4.4
130	130sA_sad3	aTerra	1720	191	19	70	5.1
131	131sA_sad4	oCao	1329	185	31	70	4.4
132	132sA_sad5	eleChega	1427	190	24	80	5.4
133	133sA_sad6	estaRoupa	1574	194	37	65	4.2
134	134sA_sad7	osJardins	1701	231	81	75	4.1
135	135sA_sad8	asPessoas	1846	233	97	75	4.7
136	136sA_sad9	haArvores	1594	198	50	70	4.4
137	137sA_sad10	osTigres	1996	178	26	80	4.4
138	138sA_sad11	oQuadro	1599	184	25	80	4.8
139	139sA_sad12	alguemFechou	1750	206	75	75	5.2
140	140sA_sad13	osJovens	1760	181	20	85	5.2
141	141sA_sad14	oFutebol	1932	210	47	55	4.3
142	142sA_sad15	elaViajou	1555	171	17	85	5.4
143	143sB_sad16	estaMesa	1360	213	19	85	4.9
144	144sB_sad17	oRadio	1294	218	55	95	4.5
145	145sB_sad18	aqueleLivro	1367	216	30	90	4.7
146	146sB_sad19	aTerra	1274	204	14	80	4.6
147	147sB_sad20	oCao	1278	205	24	95	4.9
148	148sB_sad21	eleChega	1101	208	34	95	4.6
149	149sB_sad22	estaRoupa	1397	207	21	100	4.6
150	150sB_sad23	osJardins	1491	244	74	95	5.2
151	151sB_sad24	asPessoas	1550	213	41	95	4.8
152	152sB_sad25	haArvores	1423	208	24	95	5.2
153	153sB_sad26	osTigres	1603	210	42	65	4.6
154	154sB_sad27	oQuadro	1290	208	23	100	4.8
155	155sB_sad28	alguemFechou	1494	207	21	90	5.3
156	156sB_sad29	osJovens	1475	215	25	100	5
157	157sB_sad30	oFutebol	1435	237	80	90	4.1
158	158sB_sad31	elaViajou	1374	218	45	85	4.4
159	159sA_surprise1	estaMesa	1615	272	60	95	6
160	160sA_surprise2	oRadio	1340	314	92	95	6
161	161sA_surprise3	aqueleLivro	1530	251	39	100	5.5
162	162sA_surprise4	aTerra	1436	307	97	100	6.1
163	163sA_surprise5	oCao	1360	291	66	95	5.8
164	164sA_surprise6	eleChega	1222	321	99	85	5.3
165	165sA_surprise7	estaRoupa	1547	341	102	95	6.1
166	166sA_surprise8	osJardins	1655	305	43	100	6.2
167	167sA_surprise9	asPessoas	1859	327	58	95	5.8
168	168sA_surprise10	haArvores	1586	295	59	100	6.5
169	169sA_surprise11	osTigres	1739	321	98	100	5
170	170sA_surprise12	oQuadro	1544	330	87	100	5.5
171	171sA_surprise13	alguemFechou	1802	325	100	80	5
172	172sA_surprise14	osJovens	1652	329	110	90	5.7
173	173sA_surprise15	oFutebol	1519	295	114	95	4.8
174	174sA_surprise16	elaViajou	1525	306	92	85	5.2
175	175sB_surprise17	estaMesa	1243	319	44	85	4.2

176	176sB_surprise18	oRadio	1238	330	55	75	5.5
177	177sB_surprise19	aqueleLivro	1349	337	59	85	4.9
178	178sB_surprise20	aTerra	1187	343	61	80	5.6
179	179sB_surprise21	oCao	1154	351	64	65	5.5
180	180sB_surprise22	eleChega	1061	339	62	60	4.9
181	181sB_surprise23	estaRoupa	1239	350	57	70	5.2
182	182sB_surprise24	osJardins	1458	360	64	85	5.1
183	183sB_surprise25	asPessoas	1555	358	77	80	5.6
184	184sB_surprise26	haArvores	1345	344	65	85	4.6
185	185sB_surprise27	osTigres	1541	385	51	90	5.2
186	186sB_surprise28	oQuadro	1311	383	62	80	5.3
187	187sB_surprise29	alguemFechou	1432	393	65	85	6
188	188sB_surprise30	osJovens	1316	360	55	90	5
189	189sB_surprise31	oFutebol	1337	346	63	70	3.6
190	190sB_surprise32	elaViajou	1211	383	51	80	4.8

Pseudo-sentences

#	Stimulus	Content	Duration (ms)	F_0 (Hz)	F_0 variability (SD) (Hz)	Accuracy (%)	Intensity (1-7)
1	1pA_angry1	estaDepa	1299	323	87	65	4.3
2	2pA_angry2	oDarrio	1358	306	67	85	4.9
3	3pA_angry3	aqueleJicro	1392	313	89	70	5.4
4	4pA_angry4	aPirra	1285	405	58	85	5.4
5	5pA_angry5	oLao	1107	297	89	70	3.8
6	6pA_angry6	eleChena	1093	386	70	70	5.8
7	7pA_angry7	estaSouda	1261	352	73	85	4.7
8	8pA_angry8	osBartins	1468	363	97	75	5
9	9pA_angry9	haArjuques	1452	308	66	85	5.1
10	10pA_angry10	osLagres	1551	308	81	70	4.8
11	11pA_angry11	oJuadre	1332	356	62	75	4.4
12	12pA_angry12	alguemBelhou	1491	356	81	90	5.5
13	13pA_angry13	osDofens	1577	354	81	85	4.7
14	14pA_angry14	oDutebel	1426	369	73	55	4.4
15	15pA_angry15	elaJiavou	1308	330	92	90	5.3
16	16pB_angry16	estaDepa	1171	266	44	80	6.1
17	17pB_angry17	oDarrio	1146	290	52	60	5.3
18	18pB_angry18	aqueleJicro	1162	264	46	50	5.1
19	19pB_angry19	aPirra	1081	286	36	85	5.8
20	20pB_angry20	oLao	1050	262	45	75	5.1
21	21pB_angry21	eleChena	1023	294	59	50	5.3
22	22pB_angry22	estaSouda	1201	263	54	65	5.2
23	23pB_angry23	osBartins	1236	278	60	55	5.5
24	24pB_angry24	asSemoas	1295	292	44	75	5.1
25	25pB_angry25	haArjuques	1200	278	58	85	5.1
26	26pB_angry26	osLagres	1338	259	49	70	6.1
27	27pB_angry27	oJuadre	1165	250	39	70	4.2
28	28pB_angry28	alguemBelhou	1298	290	54	95	5.3

29	29pB_angry29	osDofens	1398	274	40	70	5.6
30	30pB_angry30	oDutebel	1253	277	37	65	5.2
31	31pB_angry31	elaJiavou	1179	268	39	95	6.3
32	32pA_disgust1	estaDepa	1682	252	68	60	5.1
33	33pA_disgust2	oDarrio	1604	262	87	50	4.8
34	34pA_disgust3	aqueleJicro	1853	267	89	65	4.5
35	35pA_disgust4	aPirra	1709	302	78	65	5.2
36	36pA_disgust5	oLao	1444	263	85	50	4.3
37	37pA_disgust6	eleChena	1330	282	93	55	5.4
38	38pA_disgust7	estaSouda	1636	304	73	65	5
39	39pA_disgust8	asSemoas	2066	321	79	55	5.3
40	40pA_disgust9	haArjuques	1684	283	85	45	4.4
41	41pA_disgust10	oJuadre	1846	291	72	60	5.1
42	42pA_disgust11	oDutebel	1911	309	75	75	5.9
43	43pB_disgust12	oDarrio	1548	250	67	50	5.1
44	44pB_disgust13	aqueleJicro	1357	236	48	80	6
45	45pB_disgust14	aPirra	1319	277	55	70	4.7
46	46pB_disgust15	eleChena	1169	288	59	75	5.9
47	47pB_disgust16	estaSouda	1458	255	53	55	5.7
48	48pB_disgust17	osBartins	1437	257	44	65	5.2
49	49pB_disgust18	asSemoas	1446	287	29	75	5.5
50	50pB_disgust19	oJuadre	1513	270	61	50	5
51	51pB_disgust20	alguemBelhou	1506	280	41	50	5.1
52	52pB_disgust21	osDofens	1859	278	48	50	6.2
53	53pB_disgust22	oDutebel	1468	258	53	70	5.5
54	54pB_disgust23	elaJiavou	1501	281	69	45	5.4
55	55pA_fear1	estaDepa	1703	273	29	55	5
56	56pA_fear2	oDarrio	1541	320	56	45	5.2
57	57pA_fear3	aqueleJicro	1487	298	42	60	5
58	58pA_fear4	osBartins	1763	358	64	50	4.8
59	59pA_fear5	haArjuques	1928	302	40	45	4.4
60	60pA_fear6	osLagres	1928	297	53	45	4.8
61	61pA_fear7	alguemBelhou	1961	292	38	65	5.4
62	62pA_fear8	oDutebel	1821	360	76	50	4
63	63pB_fear9	estaDepa	1137	317	33	70	5.6
64	64pB_fear10	oDarrio	1124	340	39	55	5.7
65	65pB_fear11	oLao	1033	331	34	60	5.8
66	66pB_fear12	eleChena	1049	354	39	55	5.1
67	67pB_fear13	estaSouda	1165	309	42	45	3.7
68	68pB_fear14	haArjuques	1303	327	30	75	5.2
69	69pB_fear15	oJuadre	1276	319	49	50	5.2
70	70pB_fear16	alguemBelhou	1365	325	40	50	5.3
71	71pB_fear17	elaJiavou	1186	271	32	70	5
72	72pA_happy1	estaDepa	1456	355	89	50	4.4
73	73pA_happy2	oDarrio	1402	367	81	50	4.4
74	74pA_happy3	aqueleJicro	1602	344	91	50	4.6
75	75pA_happy4	aPirra	1330	367	78	45	4

76	76pA_happy5	estaSouda	1350	355	93	45	4.4
77	77pA_happy6	osBartins	1566	359	77	75	5.3
78	78pA_happy7	asSemoas	1758	372	80	75	5.3
79	79pA_happy8	haArjuques	1527	325	91	75	5
80	80pA_happy9	osLagres	1681	374	94	50	4.8
81	81pA_happy10	oJuadre	1627	366	72	80	5.7
82	82pA_happy11	alguemBelhou	1735	358	81	45	5.6
83	83pA_happy12	elaJiavou	1662	350	81	70	4.9
84	84pB_happy13	oDarrio	1436	325	71	45	6.3
85	85pB_happy14	oLao	1213	338	81	60	5.9
86	86pB_happy15	eleChena	1285	346	86	50	6.4
87	87pB_happy16	osBartins	1506	310	84	75	6.4
88	88pB_happy17	asSemoas	1598	337	75	65	6.3
89	89pA_neutral1	estaDepa	1541	205	29	90	4.9
90	90pA_neutral2	oDarrio	1468	196	66	70	4.5
91	91pA_neutral3	aqueleJicro	1734	204	51	90	4.7
92	92pA_neutral4	aPirra	1499	196	37	80	4.4
93	93pA_neutral5	oLao	1300	191	26	75	4.3
94	94pA_neutral6	eleChena	1184	191	26	75	4.4
95	95pA_neutral7	estaSouda	1617	219	45	90	4.1
96	96pA_neutral8	osBartins	1647	260	102	75	4.8
97	97pA_neutral9	asSemoas	1705	237	79	90	5.2
98	98pA_neutral10	haArjuques	1503	200	23	60	4
99	99pA_neutral11	osLagres	1841	232	99	95	4.9
100	100pA_neutral12	oJuadre	1535	208	51	85	4.7
101	101pA_neutral13	alguemBelhou	1698	220	62	95	4.5
102	102pA_neutral14	osDofens	1833	206	43	90	5.3
103	103pA_neutral15	oDutebel	1714	220	53	90	4.8
104	104pA_neutral16	elaJiavou	1445	195	43	90	4.4
105	105pB_neutral17	estaDepa	1369	220	25	60	4.6
106	106pB_neutral18	aqueleJicro	1396	208	30	60	3.1
107	107pB_neutral19	aPirra	1210	219	19	85	5.4
108	108pB_neutral20	oLao	1226	211	26	85	4.5
109	109pB_neutral21	eleChena	926	214	19	90	5.3
110	110pB_neutral22	estaSouda	1348	214	18	90	3.8
111	111pB_neutral23	osBartins	1439	215	19	90	4.5
112	112pB_neutral24	asSemoas	1602	219	22	80	5.1
113	113pB_neutral25	haArjuques	1428	229	22	70	4.4
114	114pB_neutral26	osLagres	1503	233	58	90	4.6
115	115pB_neutral27	oJuadre	1423	208	35	70	4.2
116	116pB_neutral28	alguemBelhou	1383	233	77	80	4.1
117	117pB_neutral29	oDutebel	1319	226	23	90	4.3
118	118pB_neutral30	elaJiavou	1679	207	20	95	4.8
119	119pA_sad1	estaDepa	1615	186	18	80	5.4
120	120pA_sad2	aqueleJicro	1883	191	42	75	4.2
121	121pA_sad3	aPirra	1711	189	17	70	3.9
122	122pA_sad4	oLao	1298	179	28	60	4.2

123	123pA_sad5	eleChena	1401	189	21	70	4.1
124	124pA_sad6	estaSouda	1668	182	28	70	4.8
125	125pA_sad7	asSemoas	1838	205	77	85	4.6
126	126pA_sad8	haArjuques	1868	206	61	90	4.7
127	127pA_sad9	oJuadre	1645	181	28	65	4.1
128	128pA_sad10	alguemBelhou	1888	205	87	70	4.5
129	129pA_sad11	oDutebel	1929	212	76	60	3.6
130	130pA_sad12	elaJiavou	1555	166	14	65	5.1
131	131pB_sad13	estaDepa	1369	212	25	90	5.4
132	132pB_sad14	oDarrio	1293	218	25	95	5.6
133	133pB_sad15	aqueleJicro	1335	213	23	90	4.7
134	134pB_sad16	aPirra	1357	229	72	85	5.5
135	135pB_sad17	oLao	1245	206	24	70	5.1
136	136pB_sad18	eleChena	1052	211	19	85	5.2
137	137pB_sad19	estaSouda	1400	227	49	95	4.8
138	138pB_sad20	osBartins	1460	235	76	95	4.8
139	139pB_sad21	asSemoas	1511	219	55	95	5.3
140	140pB_sad22	haArjuques	1488	208	19	90	5.3
141	141pB_sad23	osLagres	1646	228	65	85	5.6
142	142pB_sad24	oJuadre	1350	211	19	90	4.8
143	143pB_sad25	alguemBelhou	1481	233	50	90	5.1
144	144pB_sad26	osDofens	1570	198	40	95	5.7
145	145pB_sad27	oDutebel	1512	208	30	90	5.5
146	146pB_sad28	elaJiavou	1465	203	16	85	5.3
147	147pA_surprise1	estaDepa	1601	324	111	95	6.2
148	148pA_surprise2	oDarrio	1361	311	93	95	5.8
149	149pA_surprise3	aqueleJicro	1533	330	105	95	5.7
150	150pA_surprise4	aPirra	1437	295	86	85	5.9
151	151pA_surprise5	oLao	1280	274	64	90	5.6
152	152pA_surprise6	eleChena	1198	323	95	80	5.9
153	153pA_surprise7	estaSouda	1537	314	89	95	6.1
154	154pA_surprise8	osBartins	1676	324	90	85	5.5
155	155pA_surprise9	asSemoas	1781	298	67	85	5.8
156	156pA_surprise10	haArjuques	1601	287	43	90	5.9
157	157pA_surprise11	osLagres	1723	322	100	95	5.2
158	158pA_surprise12	oJuadre	1620	297	98	85	5
159	159pA_surprise13	alguemBelhou	1833	275	84	95	5.4
160	160pA_surprise14	osDofens	1829	293	84	100	4.8
161	161pA_surprise15	oDutebel	1606	295	111	70	4.3
162	162pA_surprise16	elaJiavou	1500	299	93	80	4.3
163	163pB_surprise17	estaDepa	1252	345	79	100	6.5
164	164pB_surprise18	oDarrio	1189	344	78	85	5.1
165	165pB_surprise19	aqueleJicro	1266	332	54	55	3.5
166	166pB_surprise20	aPirra	1164	344	73	80	4.6
167	167pB_surprise21	oLao	1123	333	69	70	4.6
168	168pB_surprise22	eleChena	950	333	61	55	4.9
169	169pB_surprise23	estaSouda	1332	337	65	95	5.8

170	170pB_surprise24	osBartins	1439	363	67	95	6.1
171	171pB_surprise25	asSemoas	1566	353	85	85	5.5
172	172pB_surprise26	haArjuques	1322	356	61	85	5.3
173	173pB_surprise27	osLagres	1509	342	60	90	5.2
174	174pB_surprise28	oJuadre	1385	338	71	90	4.8
175	175pB_surprise29	alguemBelhou	1408	331	74	80	4.9
176	176pB_surprise30	osDofens	1556	335	72	85	5.9
177	177pB_surprise31	oDutebel	1395	360	89	70	4.4
178	178pB_surprise32	elaJiavou	1285	343	97	90	5.8

7. Unbiased accuracy rates for each prosodic emotion in Chapter VIII

Hu for each prosodic emotion, as a function of age and expertise groups. Values vary between 0 and 1. Standard errors are presented in parentheses.

Age group / emotion	*Hu* scores	
	Controls	Musicians
Younger		
Anger	.6 (0.06)	.8 (0.04)
Disgust	.5 (0.07)	.6 (0.06)
Fear	.7 (0.06)	.8 (0.04)
Happy	.5 (0.06)	.6 (0.05)
Sadness	.7 (0.04)	.8 (0.03)
Surprise	.6 (0.05)	.6 (0.03)
Neutrality	.6 (0.05)	.7 (0.04)
Middle-aged		
Anger	.4 (0.05)	.6 (0.06)
Disgust	.2 (0.04)	.3 (0.05)
Fear	.4 (0.06)	.6 (0.07)
Happy	.3 (0.03)	.4 (0.04)
Sadness	.7 (0.04)	.6 (0.05)
Surprise	.4 (0.04)	.5 (0.03)
Neutrality	.5 (0.05)	.6 (0.04)

8. Intensity ratings for each prosodic emotion in Chapter VIII

Intensity ratings (minimum 1, maximum 7) for each prosodic emotion, as a function of age and expertise groups. Standard errors are presented in parentheses.

Age group / emotion	Intensity ratings	
	Controls	Musicians
Younger		
Anger	5.6 (0.2)	5.1 (0.2)
Disgust	5.6 (0.2)	4.4 (0.3)
Fear	5.8 (0.2)	4.9 (0.3)
Happy	5.4 (0.3)	5.4 (0.2)
Sadness	5.7 (0.3)	4.8 (0.4)
Surprise	5.8 (0.3)	5.3 (0.2)
Neutrality	5.6 (0.4)	5.2 (0.3)
Middle-aged		
Anger	5.2 (0.3)	5.1 (0.3)
Disgust	5.1 (0.3)	4.6 (0.3)
Fear	5.2 (0.3)	5.1 (0.4)
Happy	5.3 (0.2)	5.5 (0.2)
Sadness	5.4 (0.3)	5.2 (0.3)
Surprise	5.6 (0.2)	5.4 (0.2)
Neutrality	5.2 (0.4)	4.9 (0.3)

9. RTs for each prosodic emotion in Chapter VIII

RTs for each prosodic emotion, as a function of age and expertise groups. Standard errors are presented in parentheses.

Age group / emotion	RTs	
	Controls	Musicians
Younger		
Anger	3,074 (106)	3,419 (185)
Disgust	3,131 (117)	3,764 (189)
Fear	3,111 (122)	4,509 (173)
Happy	3,078 (74)	3,411 (187)
Sadness	2,936 (100)	3,512 (229)
Surprise	3,006 (107)	3,247 (171)
Neutrality	2,960 (111)	3,390 (195)
Middle-aged		
Anger	4,141 (218)	4,171 (251)
Disgust	4,243 (273)	4,498 (328)
Fear	4,171 (185)	4,197 (211)
Happy	3,956 (277)	4,194 (262)
Sadness	3,758 (280)	4,201 (221)
Surprise	3,768 (182)	3,991 (147)
Neutrality	3,963 (182)	3,861 (239)

10. Acoustic characteristics of the prosodic stimuli used in Chapter VIII

Detailed acoustic characteristics of the prosodic stimuli. Note: Dur = Duration (ms); F_0 mean = fundamental frequency mean (Hz); F_0 var. = fundamental frequency variability (*SD*, Hz); Int. = mean intensity (dB); Int. var. = intensity variability (*SD*, dB); F_0 min. = fundamental frequency minimum (Hz); F_0 max. = fundamental frequency maximum (Hz); Jit. = Jitter (%); P.p. = Pause proportion (ratio between portions of silence and total duration); HFE = high-frequency energy (relative energy above vs. below 500 Hz).

#	Stimulus	Dur	F_0 mean	F_0 var.	Int.	Int. var.	F_0 min.	F_0 max.	Jit.	P.p.	HFE
1	4sA_angry4	1306	355	68	78	16	210	447	1.55	.31	4.50
2	5sA_angry5	1166	299	99	76	11	176	469	1.55	.28	2.17
3	7sA_angry7	1224	349	85	80	9	169	473	1.28	.11	0.47
4	14sA_angry14	1477	338	80	76	12	174	524	1.30	.35	1.33
5	16sA_angry16	1323	329	93	77	11	163	525	0.87	.21	2.43
6	18sB_angry18	1079	289	56	79	13	188	381	1.62	.21	4.33
7	24sB_angry24	1160	283	47	80	8	204	382	2.40	.10	1.24
8	26sB_angry26	1202	293	61	79	9	178	448	2.24	.30	3.38
9	28sB_angry28	1200	262	60	77	14	115	332	2.03	.37	4.00
10	29sB_angry29	1296	305	52	77	8	190	379	2.06	.05	4.00
11	33sA_disgust1	1661	270	76	76	8	171	469	2.10	.13	1.18
12	35sA_disgust3	1596	289	76	79	10	182	523	1.86	.10	3.20
13	36sA_disgust4	1908	279	77	76	8	191	507	1.83	.22	1.63
14	37sA_disgust5	1800	286	78	78	13	203	494	1.42	.26	4.13
15	38sA_disgust6	1852	281	67	78	9	195	495	0.92	.17	2.00
16	39sB_disgust7	1376	272	62	79	8	150	373	2.44	.14	0.92
17	40sB_disgust8	1317	266	64	78	11	90	333	2.12	.33	0.08
18	41sB_disgust9	1875	289	75	78	12	80	410	1.99	.20	0.26
19	43sB_disgust11	1530	287	57	76	11	151	399	3.46	.34	0.47
20	44sB_disgust12	1608	278	51	74	12	197	395	2.25	.61	0.25
21	46sA_fear2	1501	336	45	80	9	224	403	2.84	.13	0.22
22	50sA_fear6	1772	310	45	78	9	218	433	3.24	.12	0.06
23	51sA_fear7	1954	291	38	78	8	212	371	2.21	.10	0.17
24	53sA_fear9	2009	309	34	78	11	213	452	1.84	.20	0.12

25	54sA_fear10	1759	290	58	79	7	77	491	2.34	.07	0.09
26	57sB_fear13	1086	300	28	78	10	224	328	1.38	.22	0.08
27	59sB_fear15	1125	308	33	77	12	201	325	3.21	.35	0.13
28	63sB_fear19	1212	325	45	79	10	234	387	1.79	.16	0.09
29	65sB_fear21	1400	305	41	79	9	213	359	1.94	.10	0.70
30	68sB_fear24	1300	314	41	77	11	269	482	4.28	.33	0.13
31	69sA_happy1	1451	360	83	80	7	189	435	1.08	.06	0.53
32	71sA_happy3	1588	379	102	77	9	188	491	0.92	.15	0.50
33	72sA_happy4	1368	377	98	80	8	187	499	0.96	.18	1.53
34	75sA_happy7	1628	376	84	79	8	192	495	1.55	.18	0.19
35	82sA_happy14	1647	371	83	76	13	199	508	1.41	.38	1.50
36	91sB_happy23	1660	320	79	78	13	181	481	1.74	.20	0.72
37	92sB_happy24	1470	345	93	79	10	81	525	1.92	.07	2.00
38	94sB_happy26	1551	348	82	77	9	176	464	1.16	.13	5.00
39	95sB_happy27	1547	335	76	77	10	178	489	1.58	.17	0.82
40	97sB_happy29	1296	321	70	78	9	170	479	1.69	.16	1.13
41	101sA_neutral4	1555	204	29	79	15	135	300	2.80	.29	0.36
42	103sA_neutral6	1088	203	25	80	5	153	284	1.93	.00	0.16
43	105sA_neutral8	1575	239	79	78	7	152	477	2.83	.09	0.23
44	108sA_neutral11	1842	196	28	78	7	145	262	3.87	.04	0.27
45	112sA_neutral15	1838	223	84	79	11	149	485	3.15	.22	0.58
46	119sB_neutral22	1294	218	23	79	10	175	262	2.98	.13	0.25
47	114sB_neutral17	1609	215	24	79	8	166	258	1.11	.05	0.37
48	115sB_neutral18	1282	221	40	80	9	175	494	1.30	.14	0.79
49	121sB_neutral24	1400	208	16	79	8	172	243	2.55	.22	1.18
50	124sB_neutral27	1631	205	17	78	7	175	233	1.54	.05	1.00
51	128sA_sad1	1756	197	22	76	8	166	243	2.09	.16	0.44
52	137sA_sad10	1996	178	26	75	6	149	251	3.36	.12	0.05
53	138sA_sad11	1599	184	25	76	12	148	253	1.93	.35	0.92
54	140sA_sad13	1760	181	20	77	8	149	233	1.81	.11	0.32
55	142sA_sad15	1555	171	17	77	7	147	238	2.03	.11	0.33
56	144sB_sad17	1294	218	55	77	10	170	509	2.39	.22	1.14
57	143sB_sad16	1367	216	30	79	9	178	272	1.37	.08	0.14
58	146sB_sad19	1278	205	24	78	12	169	247	2.68	.37	0.19
59	149sB_sad22	1101	208	34	78	10	179	265	2.98	.23	0.53
60	150sB_sad23	1491	244	74	77	10	182	499	2.34	.30	0.24
61	160sA_surprise2	1340	314	92	79	9	168	528	1.52	.09	0.86
62	163sA_surprise5	1360	291	66	78	11	204	464	1.24	.21	1.80
63	164sA_surprise6	1222	321	99	80	6	212	522	1.21	.10	0.72
64	168sA_surprise10	1586	295	59	77	11	218	516	2.43	.19	4.50
65	172sA_surprise14	1652	329	110	79	8	152	493	1.91	.11	0.32
66	175sB_surprise17	1243	319	44	78	9	239	406	1.85	.10	0.59
67	178sB_surprise20	1187	343	61	79	12	226	440	1.78	.21	0.50
68	182sB_surprise24	1458	360	64	78	8	236	464	1.53	.08	0.36
69	185sB_surprise27	1541	385	51	79	9	281	499	1.24	.14	1.41
70	187sB_surprise29	1432	393	65	79	7	268	486	1.07	.05	0.90

11. RTs for prosody, music and facial expressions in Chapter IX

RTs for emotion recognition in speech prosody, music and facial expressions, for PD patients and healthy controls. Standard errors are presented in parentheses.

Modality / emotion	RTs	
	Patients	Controls
Music		
Happiness	4,482 (250)	3,706 (253)
Peacefulness	5,972 (367)	5,374 (349)
Sadness	6,169 (423)	5,505 (334)
Fear	6,053 (407)	5,040 (359)
Speech prosody		
Happiness	3,649 (357)	2,878 (184)
Surprise	3,992 (331)	3,193 (202)
Sadness	3,938 (350)	3,047 (225)
Fear	4,332 (316)	3,574 (269)
Facial expressions		
Happiness	2,694 (206)	2,125 (165)
Surprise	3,437 (346)	3,094 (277)
Sadness	3,654 (382)	2,822 (219)
Fear	3,792 (368)	3,401 (313)

12. Unbiased accuracy rates for speech prosody in Chapter IX

Hu for speech prosody, for patients and controls. Values vary between 0 and 1. Standard errors are presented in parentheses.

Prosodic emotion	*Hu* scores	
	Patients	Controls
Happiness	.4 (0.06)	.4 (0.06)
Surprise	.3 (0.05)	.3 (0.04)
Sadness	.4 (0.07)	.5 (0.06)
Fear	.2 (0.05)	.2 (0.04)