AN ELECTROPHYSIOLOGICAL STUDY OF AFFECTIVE PROSODY PROCESSING

by

Valery Sramko
B.A., McGill University, 2005

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

In the
Department of Psychology

© Valery Sramko 2007

SIMON FRASER UNIVERSITY

Fall 2007

APPROVAL

Name: Valery Sramko

Degree: Master of Arts (Psychology)

Title of Thesis: An Electrophysiological Study of Affective
                       Prosody Processing

Chair: Dr. Michael Schmitt
                       Assistant Professor

                       Dr. Mario Liotti
                       Senior Supervisor
                       Associate Professor

                       Dr. Arlene Young
                       Supervisor
                       Associate Professor

External Examiner: Dr. Yue Wang
                       Assistant Professor
                       Department of Linguistics

Date Defended :   November 26, 2007

SFU SIMON FRASER UNIVERSITY
LIBRARY

Declaration of
Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the "Institutional Repository" link of the SFU Library website <www.lib.sfu.ca> at: <http://ir.lib.sfu.ca/handle/1892/112>) and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author's written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada

STATEMENT OF
ETHICS APPROVAL

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

(a) Human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

(b) Advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

(c) as a co-investigator, in a research project approved in advance,

or

(d) as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Bennett Library
Simon Fraser University
Burnaby, BC, Canada

ABSTRACT

The neural substrates of affective prosody are beginning to be understood. However, temporal processing and differences between emotions have yet to be established. High-density brain electrical activity was recorded while twenty healthy females heard and categorized semantically neutral sentences presented in five emotional prosodies (happy, sad, angry, fearful, and neutral). Accuracy was comparable between emotions, but reaction time was significantly different. Event-related potentials (ERPs) revealed an overall effect of prosody, with a main component peaking at 1 sec, localized in anterior temporal cortex (right greater than left). Main differences between emotions included an early effect peaking at 450 ms with increased activity in response to fear and localized in right dorsal premotor cortex. A later effect, peaking at 950 ms, showed sad-specific activity localized to left inferior premotor cortex. These results suggest that different emotions recruit distinct premotor regions during affective prosody recognition.

Keywords: emotion; affective prosody; event-related potentials
Subject Terms: emotion; nonverbal communication; prosodic analysis (linguistics); electroencephalography; neuropsychology

DEDICATION

To my parents, Jarmila and Rudolf Sramko,
for their support and encouragement.
Thank you.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: GENERAL INTRODUCTION AND LITERATURE REVIEW

1.1 Introduction: Affective Prosody and Communication

Human communication is complex and can take many forms. In addition to words, nonverbal cues, including facial expressions, gestures and body language, can provide important information about the speaker. One aspect of communication that until recently has received limited attention is affective (emotional) prosody. Prosody is a suprasegmental property of speech, often referred to as melody or intonation, which involves modulations in factors such as loudness, rate and pitch (Banse & Scherer, 1996). Prosody has several communicative functions, one of which is the expression of emotions or affective states. The expression of emotional prosody is typically a spontaneous, non-voluntary display of the inner state of the speaker (Wambacq & Jerger, 2004), and the processing of affective cues in speech acoustics is likewise considered an automatic process (Hird & Kirsner, 1998). The ability to process prosodic cues is a fundamental aspect of human communication with implications for social and psychological well-being, emotional competence, adjustment and maintenance of successful relationships (Carton, Kessler & Pape, 1999; McClure, 2000). Evidence for the importance of affective prosody in communication comes from studies with populations in whom this ability has not developed properly or is disturbed, such as individuals with autism (e.g., Paul, Augustyn, Klin & Volkmar, 2005; Shriberg et al., 2001), schizophrenia (e.g., Bozikas et al., 2006; Edwards, Pattison, Jackson & Wales, 2001) and Parkinson's disease (e.g., Davis, 2003). Difficulties in recognizing affective prosody influence social and emotional functioning as well as the ability to interpret the emotional states of others and to behave in a socially appropriate manner (Karow & Connors, 2003; Trauner, Ballantyne, Friedland & Chase, 1996). It is clear that the ability to correctly infer prosodic cues in speech has overall important implications for social-cognitive abilities in health and

disease. An understanding at the neural level can help shed light on this fundamental aspect of nonverbal communication. The neural substrates involved in affective prosody are beginning to be understood; however, much research is still needed in order to clarify the temporal processing of prosody, as well potential differences in the processing of different emotional intonations.

1.2   Language Representation in the Brain

Extensive research has been dedicated to the understanding of language. It is well-known that the dominant left cerebral hemisphere typically handles general language functions and linguistic processes, including phonetic, syntactic and semantic aspects of speech. Although language representation is now understood to be organized in large-scale architecture networks (see Vigneau, et al., 2006 for review), traditional neuropsychological models attribute comprehension of spoken language to Wernicke's area (Brodmann areas [BA] 39), in the posterior part of the left superior temporal gyrus, and production of speech to Broca's area (BA 44/45), in the left inferior frontal gyrus (Geschwind, 1970; Jay, 2003). Based on lesion data, Ross (1981) proposed a neuroanatomical model of prosody processing localized within distinct right-sided perisylvian regions organized in complete analogy to left-sided language areas. That is, the expression of affective prosody was said to rely on the Broca's homologue within the right inferior frontal cortex. On the other hand, comprehension of affective prosody was presumed to be bound to the Wernicke's homologue within the right superior temporal region.

1.3   Neural Substrates of Prosody

Considerable research effort has been directed towards the issue of hemispheric specialization in the processing of prosodic cues. Several streams of evidence suggest that the right cerebral hemisphere mediates the perception and recognition of affective prosody. Lesions within the right hemisphere have been linked with disproportionate impairments in the processing of emotionally-intoned speech (e.g., Blonder, Bowers & Heilman, 1991; Borod, Bloom, Brickman, Nakhutina & Curko, 2002; Bowers, Coslett,

Bauer, Speedie & Heilman, 1987). In line with these findings, functional neuroimaging studies have shown that evaluation of emotionally-intoned speech causes preferential activation within the right hemisphere, including the posterior middle temporal gyrus and superior temporal sulcus (Buchanan et al., 2000; Esslen, Pascual-Marqui, Hell, Kochi & Lehman, 2004; George et al., 1996; Imaizumi et al., 1997; Mitchell, Elliott, Barry, Cruttenden & Woodruff, 2003; Wildgruber, Pihan, Ackermann, Erb & Grodd, 2002), Despite the convergent findings pointing to the dominant role of the right hemisphere for speech prosody, recent evidence suggests a widespread network of predominantly bilateral regions associated with prosody processing. In addition to right temporal areas, comprehension of affective prosody appears to recruit both the right and left inferior frontal regions (Adolphs, 2002; Breitenstein, Daum & Ackermann, 1998; Imaizumi et al., 1997; Grandjean et al., 2005; Kotz et al., 2003; Mitchell et al., 2003; Wildgruber et al., 2002, 2004, 2005), as well as subcortical structures such as the basal ganglia (Breitenstein et al., 1998, Breitenstein, Van Lancker, Daum & Waters, 2001; Cancelliere and Kertesz, 1990; Kotz et al., 2003; Pell and Leonard, 2003). While the recognition of emotional prosody is a complex function that draws on multiple structures, the roles of such structures is unequal (Adolphs, Damasio & Tranel, 2002). Additional research is needed in order to fully understand the components of the system and their contribution to the processing of affective prosody

1.4 Emotion Recognition and Somatomotor Representations

An integral component of affective prosody processing involves the recognition of emotions being conveyed by speech acoustics. Recent theories of emotion recognition suggest that knowledge about emotions is strongly associated with knowledge of body states. In particular, the somatic marker theory postulates that somatomotor representations of one's own emotional states are activated not only during emotional experience and expression, but also in the perception or recognition of emotions in others (Damasio, 1994, 1996). The key idea in the hypothesis is that there is a link between intellectual, emotional events in the mind and body whereby somatic markers arise in bioregulatory processes and influence, both consciously and unconsciously, the processes of response to stimuli. The representation by the brain of its emotional state is strongly

3

coupled to bodily sensation. Thus, the markers are termed "somatic" because they relate to body-state structure and regulation even when they do not arise in the body proper but rather in the brain's representation of the body (Damasio, 1996). Consistent with the model, brain-injured patients with somatosensory cortex lesions have the greatest impairments in face emotion recognition (Adolphs, Damasio, Tranel, Cooper, & Damasio, 2000). As such, recognition of facial emotion requires the integrity of the right somatosensory cortices. It has been proposed that recognizing emotion in another person engages somatosensory representations that may simulate how one would feel if making the facial expression shown in the stimulus (Adolphs et al., 2000). Such a process is thought to help individuals discriminate emotional cues in the environment and guide their knowledge of what goes on in the minds of others. Adolphs and colleagues (2002) recently found that lesions within motor and somatomotor-related cortices were associated with compromised recognition of affective prosody, thus extending the somatomotor hypothesis to the auditory channel. This channel of emotion communication is of immediate survival value in situations, such as danger, in which visual cues (e.g., face or body postures) are not available, yet a quick action is required. In this evolutionary perspective, one might expect that different emotional channels would rely on a common central representation (somatomotor) to optimize adaptive behaviors. Whether such a system exists to aid in the recognition of affective prosody within a normal population has yet to be tested. Lesion correlation studies are limited by variability in etiology, lesion size, which typically does not respect physiological boundaries, and amount of reorganization that can take place, depending on time elapsed from the insult. In addition, they provide no information concerning the time and the order of activation of areas involved in affective prosody.

1.5   Temporal Processing of Prosody

The neural processing of affective prosody over time is poorly understood, yet such knowledge can provide important information on how prosody is processed by the brain. A number of studies using electroencephalography (EEG) have elicited the N400 component in paradigms with incongruent words and prosody (Schirmer & Kotz, 2003) and in cross-modal priming studies (Schirmer, Kotz & Friederici, 2002, 2005), reflecting

4

task-specific demands on semantic processing. Right-lateralized positivities have been reported using DC potentials in measuring emotional prosody (Pihan, Altenmuller & Ackermann, 1997; Pihan, Altenmuller, Hertrich & Ackermann, 2000). Such positivities are often related to attention-regulated motivation, not unlike the P300 elicited to rare, unexpected stimuli (Donchin and Coles, 1988; Schupp, Junghöfer, Weike & Hamm, 2004). Time-course analysis of emotional prosody in semantically congruent sentences revealed an early negative component (N1) at posterior electrode sites, a positive component (P2) at anterior sites, and a long-lasting negativity (Kotz & Paulmann, 2007). The positivity was elicited regardless of whether the prosody was task relevant or not, prompting the authors to suggest that it may instead be related to the process of extracting auditory/acoustic emotion-related parameters in order to assign significance during emotional speech comprehension. Using dynamic causal modeling, a recent study suggested that acoustic information associated with emotional prosody is first processed within the right temporal cortex before being processed in the bilateral frontal cortices in parallel (Ethofer at al., 2006). Such studies provide valuable information. However, the knowledge on time-course processing of affective prosody is limited and warrants more attention.

1.6 Differences Between Emotional Prosodies

An understanding of affective prosody processing is incomplete without the consideration of different emotions. Emotions are presumed to differ in terms of patterns of changes in motor expression, physiology, and subjective feeling (Scherer, 2000). Although defining acoustic features for different emotional intonations is somewhat difficult (Banse & Scherer, 1996), the idea of specificity of patterning remains viable since judges are able to reliably decode patterns of vocal expressions of emotion (Scherer, 1989). Research in the area of affective prosody is still in its infancy. However, initial evidence suggests differential processes and neural mechanisms for different emotional intonations. In one line of investigation, fMRI data revealed enhanced responses for angry prosody relative to neutral prosody in the auditory association cortex, specifically the superior temporal sulcus, which was unrelated to acoustic features of the stimuli (Grandjean et al., 2005; Sander et al., 2005). The authors hypothesized that this

area may be finely tuned to extract socially and affectively salient signals from co-specifics, just as facial expressions show an enhancement in sensory processing in the right mid-fusiform gyrus for fearful relative to neutral faces. In another fMRI study, Johnstone and colleagues (2006) found that happy intonations elicited more activation than angry voices in the right anterior and posterior middle temporal gyrus and inferior frontal gyrus. The enhanced response for happy intonations was interpreted as reflecting the salience of affiliative social signals. Unfortunately, neuroimaging studies that have included a number of emotional intonations either did not find any differences or failed to consider this issue (Ethofer et al., 2006; Wildgruber et al., 2002, 2005). A recent lesion study found differential processing of emotional intonation, showing that patients with frontal damage were mostly impaired in the comprehension of happy intonations, those with temporo-parietal damage in the assessment of sad intonations, while those with subcortical lesions had most difficulty with the comprehension of angry intonations (Rymarczyk & Grabowska, 2007). In an event-related study (Alter et al., 2003), valence judgements of emotionally intoned sentences (happiness, neutral, and cold anger) with congruent lexical content elicited distinctive brain potentials extending over most of the sentence rather than being highly localized. Specifically, a P200 component differentiated positive and neutral intonations, around 400 ms neutral intonations were differentiated from both positive and negative intonations, a difference which persisted over the course of the sentence, and finally positive and negative intonations were differentiated between 400 and 700 ms. It is clear from the review of different streams of evidence that emotion should be considered in attempts at understanding how they are processed by the brain.

1.7 Asymmetries in Emotion Processing

Neuropsychological models of emotion processing suggest that right versus left hemisphere distinctions could be relevant to understanding the cerebral organization of emotions. A common asymmetry observed in emotion processing is the advantage of the right hemisphere in being faster and more accurate at encoding negative emotions such as fear and anger (Davidson, 1995). In contrast, the encoding of positive emotional stimuli, such as happiness, has failed to show a consistent pattern of hemispheric asymmetries, which has resulted in two separate theories of the laterality of emotional perception

(Smith & Fleming, 2006). The right-hemisphere hypothesis suggests that all emotions, regardless of valence are preferentially processed by the right hemisphere (e.g., Borod et al., 1998; Cicero et al., 1999). In contrast, the valence hypothesis posits that negative information is processed more efficiently by the right hemisphere but that positive information is processed more efficiently by the left hemisphere (e.g., Canli, Desmond, Zhao, Glover, & Gabrieli, 1998; Canli et al., 1999; Davidson & Irwin, 1999). Given the limited data on differences between emotional prosodies, it remains unclear whether emotional intonations follow a similar pattern of hemispheric asymmetries.

1.8 Advantage for Threat Detection

Several streams of evidence suggest that humans are particularly attuned to detecting emotional stimuli, and especially threatening stimuli, which tend to attract and hold attention (Ochsner & Schacter, 1995). Priming studies (e.g., Bargh, Chaiken, Govender, & Pratto, 1992) and emotional Stroop paradigms (e.g., Riemann & McNally, 1995; Williams, Mathews & MacLeod, 1996) reveal biases favouring the processing of affective information, particularly information that is relevant to topics of current concern. Visual search tasks reveal preferential attention to potentially dangerous stimuli (e.g., snakes, spiders, guns) over neutral stimuli (e.g., flowers, mushrooms, toasters; Fox, Griggs & Mouchlianitis, 2007). Similarly, threatening facial expressions are processed more efficiently in visual search tasks than positive or neutral expressions (e.g., Eastwood, Smilek, & Merikle, 2001; Fox et al., 2000; Hansen & Hansen, 1988; Öhman, Lundqvist, & Esteves, 2001). The detection of threat-relevant stimuli allows for the rapid perception of potential danger and engagement of coping strategies such as fighting, freezing or rapid escape (e.g., Armony & LeDoux, 2000; Calder, Lawrence, & Young, 2001; LeDoux, 1996). Thus, rapid and efficient judgements about the significance of social threat are important for species survival. One might expect a similar advantage for detecting threatening stimuli in speech prosody.

1.9 Gender Differences in Emotion Processing

Gender differences have been found for language processing as well as emotional processing. For instance, females tend to be more accurate than males at recognizing facial, gestural and vocal emotional expressions (Hall, 1978). Knapp and Hall (1997) proposed that gender differences in terms of measures of interpersonal sensitivity can help explain the female advantage in decoding nonverbal cues. Consistent findings suggest that females are better able to accurately sense and perceive one's personal, interpersonal, and social environment. As such, it may not be surprising that females are also better at detecting nonverbal cues which play an important role in communication. Additional support for a gender difference stems from neurophysiological research. Behavioural and ERP data reveal time-course differences in processing, showing that females integrate emotional prosody and word valence earlier then males (Schirmer et al., 2002). Additionally, brain lesion and neuroimaging studies have revealed gender differences in terms of language lateralization. Language representation appears to be more bilateral in females and more unilateral and left lateralized in males (e.g., Shaywitz, Shaywitz, Pugh, Constable & Skudlarski, 1995). Distinct brain areas also appear to be activated based on the gender of the voice that one is listening to. Specifically, in the male brain, the perception of female voices activates the right anterior superior temporal gyrus, near the superior temporal sulcus, whereas a male voice activates the mesio-parietal precuneus area (Sokhi, Hunter, Wilkinson & Woodruff, 2005). While no studies to date have investigated this cross-gender effect among females, a similar finding would not be unexpected. These findings highlight the importance of controlling the gender of the listener and the gender of the portrayer in experimental manipulations in order to better isolate effects.

1.10 Experimental Paradigms and Prosodic Stimuli

A variety of stimuli and experimental paradigms have been developed with the increasing interest in affective prosody as a research topic. However, due to the lack of standard or widely accepted measures, most investigators have generated their own stimuli in order to address their specific questions of interest. Dichotic listening tasks

have been particularly useful for understanding hemispheric processing of prosody and different emotional intonations (e.g., Bulman-Fleming & Bryden, 1994; Herrero & Hillix, 1990; Ley & Bryden, 1982; Reuter-Lorenz & Davidson, 1981). However, such tasks are limited in their ability to inform about the underlying pattern of regional activity or to delineate neural substrates recruited over time. Similarly, single words can reveal early effects (e.g., Schirmer & Kotz, 2003); however, in order to delineate later effects, the use of sentences appears to be preferable. Some investigators have employed unintelligible utterances (Kotz et al., 2003) or meaningless speech (Grandjean et al., 2005) spoken with emotional prosody, designed to remove semantic processing from the cognitive task. However, as Beaucousin et al. (2007) emphasize, such designs may paradoxically result in increased semantic demand. In support of this interpretation, the more the speech is unintelligible, the more the activity in the left perisylvian semantic areas appears to increase (Kotz et al., 2003; Meyer, Alter, Friederici, Lohmann & Von Cramon, 2002). A similar argument can be made about the use of stimuli for which semantic content is incongruent with prosody. Such paradigms are often well-suited for investigating brain responses to cognitive conflict (e.g., Mitchell et al., 2006) or answering questions of relative importance between different sources of information (e.g., Mehrabian, 1972). However, paradigms with conflicting semantic content and prosody can result in the taxing of the semantic system. Indeed, such stimuli have typically elicited a negative-going EEG component (N400) reflecting effortful meaning integration (Schirmer et al., 2002, 2005; Alter et al., 2003). In general, it is commonly accepted that when investigating emotional speech it is preferable to use stimuli with neutral lexical content in order to isolate acoustic features independent of lexical content (Wildgruber et al., 2002, 2004, 2005). In the developmental literature, the Diagnostic Analysis of Nonverbal Accuracy scale (DANVA; Nowicki & Duke, 1994) is a popular measure used. The DANVA includes a paralanguage subscale assessing nonverbal processing skills targeting 4 basic emotions: happiness, sadness, anger and fear. This measure has several desirable qualities such as the inclusion of a female and a male voice, intonations targeting a number of different emotions, and good psychometric properties. While the lexical content of this measure is neutral, it includes only one sentence, which questions the generalizability of findings beyond those of the particular stimulus. Thus, an ideal

measure would include a sampling of several semantically neutral sentences read in a number of emotional intonations in both a female and a male voice.

1.11 Electroencephalography (EEG)

To gain a better understanding of basic neural mechanisms involved in the processing of affective prosody, EEG is a well-suited technique. EEG measures the bioelectric activity of the brain noninvasively via electrodes placed on the surface of the scalp. EEG can attain temporal resolution better than 1 ms and spatial resolution of 2.5 cm at the cortical surface. Event-related potentials (ERPs) are a transient change of EEG voltages reflecting systematic brain activity which is triggered by a sensory stimulus or motor response (de Zubicaray, 2006). Given the exquisite time resolution, ERPs can be selectively averaged in response to different categories of rapidly occurring stimuli within the same experimental block. One of the advantages of ERPs is that they may elucidate the time course of prosody comprehension by allowing the differentiation of language subprocesses which are engaged sequentially in time, as reflected in language-related components (Alter et al., 2003). A limitation of EEG is that scalp-recorded electrical signals reflect activity generated in the cortical and subcortical grey matter and propagated to the head surface through tissue of inhomogeneous density and conductance. Several methods can be used to solve the inverse problem (i.e., the computation of images of electric neuronal activity based on extracranial measurements). Standardized low resolution brain electromagnetic tomography (sLORETA) projects EEG scalp measurements into brain space. This method produces 3D images of statistically standardized current density with exact point-wise localizations (Pascual-Marqui, Esslen, Kochi & Lehmann, 2002). Thus, the combination of EEG and source modeling techniques result in a method with excellent temporal and spatial resolution.

1.12 The Present Study

The purpose of the current study was to delineate common neural processes and substrates recruited during affective prosody processing, and to identify unique spatial-

temporal components that would differentiate emotional intonations using event-related potentials (ERPs). Several predictions were proposed.

1) First, in line with existing fMRI evidence (e.g., Ethofer et al., 2006; Wildgruber et al., 2005) we predicted an overall effect of affective prosody to result in activation of a right hemisphere network, particularly right temporal regions, independent of specific intonation.

2) Second, in line with Damasio's (1996) somatic marker hypothesis, it was predicted that recognition of emotions conveyed through affective prosody would also activate premotor areas in the listener, with inferior premotor cortex BA6 (mouth/larynx area) and adjacent Broca's area (BA 44) being the most likely candidates.

3) Third, based on previous evidence of hemispheric asymmetries of emotion in other channels, it was hypothesized that prosody would also be characterized by an asymmetric representation in the brain, depending on emotion.

4) Finally, it was predicted that emotions conveying threat (anger and fear) would be processed and recognized more quickly than other social emotions without threat content (sad, happy), resulting in shorter reaction times and earlier ERP activity.

In order to test the above hypotheses, high-density EEG activity was recorded while participants categorized semantically neutral sentences presented in five different emotional prosodies (happy, sad, angry, fearful, and neutral). Evoked electrical activity to the onset of the stimuli was compared across emotions. Effects of interest over the scalp were further explored with a dipole modelling method (sLORETA) allowing to localize the likely brain sources of the scalp effects.

Given the lack of existing adequate measures available to study affective prosody, the first part of the study involved developing and validating stimuli which served as the stimuli set in the main electrophysiology experiment.

CHAPTER 2: STIMULUS DEVELOPMENT AND VALIDATION

2.1 Methods

The Simon Fraser University Research Ethics Board approved this experiment. All participants gave their written informed consent before participating in this study and received course credit for their involvement.

2.1.1 Participants

Thirty one undergraduate students (28 females and 3 males) enrolled in introductory psychology courses participated in the rating of the stimuli. All participants had normal-to-corrected vision and hearing within normal limits. The sample included individuals between the ages of 18 and 27 (M=20.6, SD=2.4). English was the language that 87.1% of the sample was most comfortable in and all participants rated their overall competence in English as advanced to native-like.

2.1.2 Stimuli

Sixty semantically neutral English sentences approximately 4 sec in duration (e.g., The children spent all day picking strawberries in the field) were developed. Two trained actors, one male and one female, recorded all sentences in a happy, sad, angry, fearful and neutral intonation, yielding a total of 600 sentences. The stimuli were digitally recorded at a 16-bit/44.1-kHz sampling rate using a Sennheiser microphone. Sentences were normalized for average amplitude using audio-editing software (Cool Edit Pro) to avoid the possibility that subjects could infer the correct emotion solely based on the perceived loudness. Each participant rated a subset of 200 stimuli, which were

counterbalanced across the sample. Within each subset, sentences were presented in random order at a level that participants indicated as sufficiently loud.

2.1.3   Procedure

Participants completed a short questionnaire that gathered background information including age, gender, ethnicity and linguistic background. The experiment took place in a sound attenuated room with participants seated in front of a computer. Using headphones, participants listened to the recorded sentences and answered three questions for each sentence heard. The first question (Question 1) asked "What emotion (if any) is conveyed by the sentence?" (1 = happy, 2 = sad, 3 = scared, 4 = angry, 5 = none/neutral) with participants indicating their response via number keys on the keyboard. This question allowed for the assessment of interrater consensus regarding the emotion conveyed by the sentence. The next question asked either "How convincing was the sentence at conveying the emotion?" (Question 2a) or "How neutral of a neutral sentence is it?" (Question 2b) using a 5-point Likert scale (1 = not convincing/very bad to 5 = very convincing/very good) depending on whether raters initially indicated that the sentence conveyed an emotion or not. These two questions assessed the degree to which the sentence was effective at conveying an emotion, or lack of emotion, respectively. The last question (Question 3) asked "How artificial versus natural was the sentence?", also answered using a 5-point Likert scale (1 = very artificial, 5 = very natural) and assessed the degree to which that sentence was natural versus artificial-sounding. The experiment was self-paced and participants were able to replay sentences at any time by pressing the "r" key on the keyboard, which allowed them to become more familiar with the stimulus and facilitate ratings. A practice trial was first completed with the assistance of the experimenter prior to the experimental task in order to ensure that the participants understood the task and the keyboard responses. The experiment took approximately 50 minutes to complete.

2.2 Analysis

Fourteen sentences with the best inter-rater reliability were selected for the next phase of the study. Sentences judged as unconvincing and/or artificial-sounding were excluded in favor of those judged as convincing and natural-sounding. Each sentence selected showed greater than 70% agreement across all raters and sentences in each category had a mean accuracy rating of 90% or more across all participants. Table 2.1 shows the percent accuracy ratings for each condition.

CHAPTER 3: EEG EXPERIMENT

3.1 Methods

The Simon Fraser University Research Ethics Board approved this experiment. All participants gave their written informed consent before participating in this study and received course credit or a monetary incentive for their involvement.

3.1.1 Participants

Twenty-one university females participated in the EEG experiment; however, one participant was excluded from analyses due to excessive artifact in her EEG data. Participants were between the ages of 18 and 38 (M=24.0, SD=5.0), were metal-free, and had normal-to-corrected vision and hearing within normal limits. In order to control for gender effects, the current study was limited to females. The sample was predominantly right-handed, but included one left-handed participant whose data made no significant difference in the results. Three participants indicated a history of psychiatric disorders with active treatment at the time of the experiment. Comparisons with and without their data revealed no significant differences, thus it was incorporated into the main analyses. The remainder of the sample showed no signs of neurological, psychiatric or hearing disorders. Nineteen of the participants had English as their native language, which was rated as their most fluent language. The remaining participant had Spanish as her native language, but rated her overall competence in English as advanced to native-like.

3.1.2 Procedure

Participants were asked to come to the laboratory with clean, dry hair free of hair styling products. Each participant was informed about the nature of the study and gave their written informed consent prior to beginning the experiment. Each participant was

asked to complete a short inventory (Background and Medical Questionnaire) asking about background information such as age, education, birthplace, languages, handedness and a brief medical history emphasizing neurological and mental health problems. The experiment was carried out in a sound-attenuated booth. Participants sat approximately 45 cm from the computer screen, with headphones on their ears and a game pad in their hands. The setting for ambient light was standardized across participants. The volume was adjusted at a comfortable level for each participant. Participants heard sentences played over the headphones and were asked to indicate as quickly and accurately as they could which emotion, if any, was being conveyed by the stimuli. The task was a forced-choice measure with each of the five prosodies corresponding to one of five response keys on the game pad. In order to control for the use of different fingers corresponding to different emotions, the emotions were counterbalanced across the hands. Half of the participants used the left index to indicate "sad"; the left middle finger for "happy"; the right index for "fearful", the right middle finger for "angry"; and the right thumb for "neutral". The other half used the left index for "fearful"; the left middle finger for "angry"; the right index for "sad"; the right middle finger for "happy"; and the left thumb for "neutral". Thus, across all participants each condition was equally represented by both hands, thus eliminating a hand effect.

In order to prompt the participant to the upcoming stimulus, participants first heard a 500 ms sound (500 Hz tone) followed by a 500 ms pause/buffer prior to hearing each sentence. The duration of each sentence was approximately 4000 ms (range 3400-4600), during which participants were required to select their response. This was followed by a random time interval of 400-800 ms preceding the onset of the next trial. Participants were instructed to remain still and avoid overt facial expressions. They were also instructed not to blink from the time they heard the pre-stimulus cue (tone) until they made their response. In order to limit eye movements, participants were asked to keep their eyes on the central fixation (i.e., "+") on the computer screen throughout the experiment.

The experimental task was preceded by three short practice trials to allow participants to learn the task, the key responses, and the timing of eye blinks. In the first practice trial, participants viewed emotion words (e.g., "HAPPY") on the computer

screen and were required to press the corresponding button on the game pad, until 100% accuracy was reached. Visual feedback followed each trial. In the second practice trial, participants listened to sample sentences played in five different intonations (happy, sad, angry, fearful, and neutral) and were required to identify the emotion conveyed by each sentence and indicate their response using the corresponding buttons on the game pad, until 80% accuracy was reached. Visual feedback again followed each trial. In the final practice trial, participants listened to sample sentences played in five different intonations (happy, sad, angry, fearful, and neutral) and were again required to indicate the emotion conveyed by each sentence by using the response buttons, but no feedback was provided in this trial in order to mimic the experimental task. This trial was repeated until participants were comfortable with the pace of the experiment and the timing of their eye blinks.

The experiment consisted of 6 blocks of approximately 6 minutes in duration. Each block had 70 unique stimuli composed of the 14 selected sentences recorded by the female actor in 5 different intonations (happy, sad, angry, fearful and neutral). There was a total of 420 stimuli in the entire experiment. Sentences within each block were presented in pseudo-random order with the constraints that there were no more than 2 consecutive stimuli of the same intonation played at a time, and semantically identical sentences were never played consecutively within a block nor between two consecutive blocks. The duration of the experimental task was approximately 40 minutes. Short breaks were given between each block.

3.1.3   EEG Recording and Analysis

EEG activity was recorded continuously from the scalp through 128 sintered Ag-AgCl electrodes embedded in an elastic cap (Electro-Cap International), which provides very low noise, low offset voltages and very stable DC performance. Electrodes were positioned in an equiradial layout relative to the vertex (i.e., each electrode was radially equidistant from Cz, the vertex location on the scalp). Water-soluble conductive electrode gel (SignaGel) was used with no additional skin preparation, given that active electrodes would make this redundant. Two reference electrodes were placed on the left and right mastoids. In order to monitor eye blinks, an extra electrode was placed below the left eye.

Electrode impedances were kept below 40 KOhm. EEG signals were amplified between 0.16 and 128 Hz by BioSemi Active-Two amplifiers and sampled at 256 Hz (bandwidth 52 Hz). Brain activity was recorded and analyzed offline using BESA software version 5.1.8. The EEG amplifier bandpass was 0.5 to 30Hz. Trials contaminated by eye movements or muscle activity were rejected from analyses. Event-related activity was selectively averaged for the different emotions (correct hits), time-locked to sentence onset. These average epochs were created by averaging from the onset of the stimulus, which included a 200 ms pre-stimulus baseline and a 3000 ms ERP time window. Grand-averages for each condition were computed by averaging single-subject ERP averages. In order to examine commonalities in brain activity across all conditions, these averages were pooled together to form a single grand-average.

In order to explore differences between emotional intonations, ERP waveforms and topographical maps of grand-averages for each emotion were inspected and compared for latency and amplitude of peak voltage activity in the main observed components. Regions of interest (ROIs) were selected based on peak voltages and neighbouring electrodes showing similar voltage amplitudes and the windows of interest were centered around this activity. Mean voltage amplitudes in the selected time windows were extracted and rectified and then employed as a parameter in the ERP analysis. In order to correct for sphericity violations (p < .05), the Greenhouse-Geisser correction was used in relevant cases.

In order to localize the source of the EEG activity, the scalp data was projected into brain space using a transformation matrix implemented in sLORETA. For the localization of the overall effect of prosody, the ERP grand-average of all conditions combined was projected into brain space. Also, in order to localize the source of observed scalp differences between emotions, individual subject ERP averages for each emotion were projected into brain space. SLORETA includes the computation of sample-by-sample t-statistics (every 2 sec) using iterative permutation tests (5000) to reject type 1 errors. Significance value was set at t = 2.58, p < 0.01, 2-tailed, corrected for multiple comparisons. The maximum of the current density obtained during the selected time window was taken as the source of the particular component.

3.2 Results

3.2.1 Behavioural Results

Mean and median reaction times (RT), standard deviations and mean percent accuracy for each condition is listed in Table 3.1. A repeated measures ANOVA revealed no significant differences between mean percent accuracy for angry, happy, sad, fearful, and neutral sentences (F (2.5, 49.4) = 1.95, p = .14), indicating comparable recognition rates across conditions. A main effect of condition (F (4, 80) = 20.06, p <.0001) on mean reaction time (RT) was observed using a repeated measures ANOVA. Bonferroni-corrected pairwise comparisons revealed that mean reaction time was significantly faster for angry than for happy (p = .001), sad (p = .001), fearful (p = .008), and neutral (p =.002) intonations. Fearful intonations had significantly faster mean reaction time than did happy intonations (p = .008). Likewise, neutral intonations had significantly faster reaction times than happy intonations (p = .007). Figure 3.1 shows mean RT for each condition.

3.2.2 Overall Effect of Affective Prosody

Inspection of ERP waveforms and scalp topographies of the grand-average of all conditions combined revealed a main slow wave with a dipolar distribution consisting of a negative peak over lateral-inferior frontal scalp locations and a less pronounced positive peak over dorsal-posterior scalp sites (see Figure 3.2 for scalp topography). The global field power was dominated by a slow wave peaking around 1 sec, which was extracted as the first component in a principal component analysis (PCA as implemented in BESA 5.1.8) accounting for 88.1% of the variance in the 0-2000 msec time window (see Figure 3.2 for global field power and PCA). With a more restricted time window around the peak (800-1000 ms), virtually all the variance (99.9%) was accounted for by the first PCA component.

Source localization (sLORETA) revealed that the anterior part of the inferior temporal poles, particularly in the right hemisphere (MNI coordinates 40, 15, -40)

accounted for most of the activity during the 800-1000 ms time window (see Figure 3.3 for distribution of activity in brain space).

3.2.3 Differences Between Emotions in ERP Scalp Topography

Scalp topography distribution for all emotional intonations can be found in Figure 3.4.

400-500 ms Time Window

Inspection of ERP waveforms and difference waves among the various emotions revealed that fearful intonations evoked a slow wave with greater amplitude over right dorsal frontal scalp, peaking at around 450 ms. Mean voltage amplitude in a 400-500 ms time window was computed for two regions of interest (ROIs) each created by collapsing 2 neighbouring sensors (left frontal: C28, C31; right frontal: C15, C09), and compared for the different conditions (happy, sad, fearful, angry and neutral). A repeated measures ANOVA in this time window revealed a main effect of emotion (F (2.8, 56.4) = 3.83, p = .017) (see Figure 3.5 for mean amplitude for each ROI by emotion). Planned comparisons (correction for multiple comparisons, p<.0125) confirmed that fearful intonations elicited greater amplitude than neutral (p=.009), angry (p = .012) and happy (p = .012) intonations, while fearful and sad intonations did not differ from each other.

500-600 ms Time Window

A second effect of interest peaked around 550 ms in left temporal/prefrontal regions with greater amplitude for fearful intonations. Mean voltage amplitude in a 500-600 ms time window was compared for different conditions (happy, sad, fearful, angry and neutral), and scalp regions of interest (left temporal/prefrontal: D09, D22; right temporal/prefrontal: B25, B28) (see Figure 3.6 for mean amplitude for each ROI by emotion). A significant interaction was found (F (4, 76) = 2.71, p = .036). Main effect of emotion was not found to be significant when each region was analyzed separately (right: F(1, 19) = .90, p > .05; left: F(1, 19) = 1.19, p > .05).

20

800-1100 ms Time Window

A difference wave peaking around 950 ms over frontal electrode sites appeared to show differences in amplitude between emotions. A repeated measures ANOVA with conditions (happy, sad, fearful, angry and neutral), and scalp regions of interest (left frontal: C29, C30; right frontal: C08, C16) on mean voltage amplitude in an 800-1100 ms time window revealed a main effect of emotion ($F(2.3, 46.3) = 4.41$, $p = .014$) and an interaction between ROI and emotion ($F(2.4, 46.3) = 3.38$, $p = .034$), which qualified the main effect of emotion. In order to interpret the effects, analyses were conducted for each region of interest separately. In the left frontal region, Bonferroni-corrected pairwise comparisons revealed significant differences between happy intonations and angry ($p = .019$), fearful ($p = .020$), and sad ($p = .031$) intonations. In contrast, in the right frontal regions, significant differences between happy intonations and angry ($p = .025$), fearful ($p = .001$), and neutral ($p = .008$) intonations were revealed. See Figure 3.7 for mean amplitude for each ROI by emotion during this time window.

3.2.4 Differences Between Emotions in Brain Space (sLORETA)

400-500 ms Time Window

In the first epoch explored, current density differences were found in right superior premotor cortex (Brodmann area [BA] 6, MNI coordinates 35, -5, 55). Significantly greater current density was observed for fearful intonations relative to both neutral intonations (maximum t-value: 4.29 at 468 ms) and angry intonations (maximum t-value: 2.88 at 457 ms). No significant current density differences were observed between fearful and happy or sad intonations, or between angry and neutral intonations (for all: $t < 1.96$, $p > .05$). See Figure 3.8 for source localization during this time window.

500-600 ms Time Window

The effects observed in the early time window in right dorsal frontal cortex appeared to persist in this later epoch, although the current density peak difference was shifted 2 cm more anteriorly, in right middle frontal gyrus (BA6/8, MNI coordinates 35, 15, 50) for fearful relative to angry intonations (maximum t-value: 3.08 at 574 ms) and fearful versus neutral intonations (maximum t-value: 3.42 at 574 ms). As for the earlier time window,

no significant differences were observed between happy and fearful intonations, fearful and sad intonations, or angry and neutral intonations (for all: t < 1.96, p >.05). See Figure 3.9 for source localization during this time window.

800-1100 ms Time Window

Greater current density in left inferior premotor cortex (BA 6/44, MNI coordinates -55, -5, 5) significantly differentiated sad from happy intonations (maximum t-value = 4.24 at 934 ms), sad from neutral intonations (maximum t-value: 4.69 at 953 ms), and neutral from happy intonations (maximum t-value = 4.24 at 934 ms). No significant differences were found in the contrasts between happy and fearful intonations, or happy and angry intonations (for both: t < 1.96, p >.05). See Figure 3.10 for source localization during this time window.

The results of statistical analyses in brain space are reported in Tables 3.2 and 3.3.

3.3 Discussion

Affective prosody is an important nonverbal means of communication with implications for social competence. An understanding of the spatial-temporal correlates of neural processing involved in the recognition of affective prosody as well as emotion-specific features allows for a more complete characterization of how affective prosody is processed. The current study employed a high-density EEG system to investigate the neural processes associated with an auditory emotion categorization task involving five different emotional intonations (happy, sad, angry, fearful, and neutral) in semantically neutral sentences.

3.3.1 Summary of Results

Behaviourally, sentences with an angry intonation were faster to be recognized than all other intonations. Moreover, fearful and neutral intonations were faster to be identified than happy intonations, which cannot be attributed to differences in ease of recognition, since accuracy was comparable across all conditions. The overall effect of

emotional prosody revealed a main peak in event-related neural activity around 1 sec. This main component was localized in the inferior part of the anterior temporal poles, with greater activity on the right compared to the left. Event-related potentials revealed three main peaks (450 ms, 550 ms, and 950 ms) with differential effects between emotions. The earliest time window (400-500 ms) showed scalp differences in mean amplitude for right and left frontal regions. Brain source analysis indicated that *fearful* intonations were accompanied by faster and greater electrical activity in a right superior frontal region, with maximum activity in the superior part of the precentral gyrus (premotor cortex BA6) and extending anteriorly towards the middle frontal gyrus (BA6/8) in the second time window (500-600 ms). Importantly, the largest scalp difference in mean amplitudes between emotions was observed over bilateral inferior frontal regions during a later time window (800-1100 ms). Brain source analysis revealed that *sad* intonations evoked greater current density than neutral or happy intonations in the inferior portion of *left* premotor cortex centered in BA6 (mouth/larynx) and extending anteriorly to include Broca's area (BA44).

3.3.2 Accuracy and Speed of Emotion Recognition

One important way in which emotions can differentiate themselves is in terms of accuracy and speed of recognition. Behavioural findings, particularly the consistency of ranking of reaction time and accuracy in emotion categorization, have been used as evidence for the universality of facial expressions (Ekman, 1979), as well as cultural differences (Huang, Tang, Helmeste, Shioiri & Someya, 2001; Mandal, Harizuka, Bhushan & Mishra, 2001). No known studies have reported and analyzed differences in reaction time and/or accuracy between emotional intonations, despite the inclusion of different emotional prosodies in a number of studies. In the present study both angry and fearful intonations were recognized faster than happy intonations. Within the face literature, there is evidence for threatening faces to demonstrate superior fear conditioning (Ohman & Mineka, 2001) and an advantage in detection (Ohman, Lundqvist, & Esteves, 2001). However, the bulk of the evidence shows a response latency advantage for happy facial expressions (Ekman, 1979; Leppänen & Hietanen, 2003). This reaction time advantage for recognizing happy faces cannot be explained by

uniqueness of features since schematic faces with equal number of critical features for each emotion show a similar effect (Kirita & Endo, 1995). While this happy-specific advantage for facial expressions has yet to be explained, differences in emotion ranking between different emotion channels (e.g., face and prosody) may reflect complementary processing of information. The auditory channel may be more attuned to detecting and decoding emotional signals that are beyond the boundaries of visual perception, such as in the dark, at a distance, or when vision is occluded by intervening objects. This may particularly hold true for emotional intonations with survival value conveying threat, such as angry and fearful prosodies, which command a quick fight-or-flight response. In a study on the acoustical attributes of emotional prosody, Sobin and Albert (1999) found that fear was reliably recognized by listeners, but was the only emotion not fully captured by acoustical features, despite a lack of heterogeneity of acoustic patterns across samples. Overdetection and a low threshold for fear, understood from an evolutionary framework where fear is considered a defensive emotion, were offered as potential explanations for their results (Sobin & Albert, 1999). In addition to shedding light on how different emotions are processed by the brain, the ranking order in which emotions are recognized can potentially inform other areas of research, such as questions of universality of affective prosodies. This area deserves consideration in future studies and replication of the current findings is needed in order to confirm their significance.

3.3.3   The Right Cerebral Hemisphere in Emotion Processing

The right cerebral hemisphere has repeatedly been implicated in the processing of emotion, although such a role is somewhat controversial due to inconsistent findings (Mitchell et al., 2003). Support for the dominant role of the right hemisphere in emotion processing has been found across various modalities and tasks, including facial emotion recognition (Borod et al., 1998), emotional scene recognition (Gardner, Ling, Flamm & Silverman, 1975), emotional memory (Burton, Gilliam, Flynn & Labar, 1999), facial emotion expression (Nakamura et al., 1999), and the ability to express appropriate emotional mood (Andersson, Krogstad & Finset, 1999). Bowers, Bauer and Heilman (1993) suggested that the right hemisphere houses a "non-verbal affect lexicon", which encompasses not just prosodic comprehension and expression, but also subserves the

production and interpretation of facial expressions and gestures. The general superiority of the right hemisphere in emotional and non-verbal processing may be based on the greater corticolimbic connectivity of this hemisphere (Liotti & Tucker, 1995). By confirming evidence that emotional prosody is processed more quickly and efficiently by the right hemisphere for fear intonations, the current study provides further support for the hypothesis that emotional processing is relatively right-lateralised.

3.3.4   Temporal Regions and Emotion Recognition

Typically, the processing of affective prosody is found to be mediated by the right superior and middle temporal gyri. The localization of the activity within the anterior temporal pole is slightly more ventral than areas identified in recent neuroimaging studies (e.g., Kotz et al., 2003; Mitchell et al., 2003), which may reflect the lower resolution for source localization based on scalp topography. However, if the activity is indeed localized in the inferior temporal poles, it may reflect higher-order processing of complex stimuli and attempts at analyzing the gestalt of the sentences. The left anterior temporal pole has been linked to the processing of complex auditory stimuli in both humans and non-human primates, showing a preference for intelligible speech (Scott, Blank, Rosen & Wise, 2000) and a specialization for species-specific vocalizations (Rauschecker, 1998; Kass & Hackett, 1999; Romanski et al., 1999). Moreover, stories activate this region when compared to unlinked sentences (Mazoyer et al., 1993; Fletcher et al., 1995), whereas sentences show activations when compared to random strings of words (Bottini et al., 1994), suggesting a role in linguistic processing beyond lexical and semantic analysis of individual words. It appears that this area is involved in extracting the meaning of the whole (Maguire, Frith & Morris, 1999). On the other hand, the right inferior temporal gyrus has been implicated in visual perception (Ishai, Ungerleider, Martin & Haxby, 1999; Herath, Kinomura & Roland, 2001) and multimodal sensory integration (Mesulam, 1998). Taken together, the inferior temporal gyri appear to mediate the processing of complex stimuli. Further research is needed in order to clarify the role of inferior temporal regions in the processing of affective prosody.

3.3.5 Somatomotor Representations and Emotion Recognition

According to the somatic marker hypothesis put forward by Damasio and colleagues (1996) emotion recognition and subjective experience (i.e., feeling) may both require re-activation of somatomotor and somato-visceral representations of their own body state during the experience of that emotion. Support for this comes from lesion and imaging data for the recognition of facial expressions (Adolphs et al., 2000), and more recently from lesion data in the recognition of affective prosody (Adolphs et al., 2002). Support for the somatic marker hypothesis also comes from work on "mirror" neurons in monkey frontal premotor cortex, showing that the same neurons are activated by performing an action toward a target, or viewing another monkey performing the same action (Rizzolatti, Fadiga, Gallese & Fogassi, 1996). The fear-specific effect in the right superior frontal region (BA 6/8) observed in the current study may reflect the need to re-activate or have access to fear-specific somatomotor representations needed in order to recognize this emotional intonation. Based on the known somatotopy of the human precentral gyrus, where the trunk, lower limbs and upper limbs are represented dorsally and the face/mouth and larynx are represented ventrally, we venture the speculation here that "embodied" fear may involve a faster and stronger reactivation of a premotor cortex subdivision representing body parts more directly involved in fight or flight responses (e.g., trunk, as related to posture, and limbs for running and fighting). This is consistent with recent neuroimaging work showing somatotopically organized motor-premotor activations during action observation, for instance, in dancers or swimmers (Buccino et al., 2001; Cross, Hamilton & Grafton, 2006). The right hemisphere dominance may be explained by the general superiority of the right hemisphere in emotional and non-verbal processing.

3.3.6 Left Inferior Frontal Gyrus and Emotion Recognition

Activity within the inferior frontal gyrus has previously been found during the recognition of affective prosody. Specifically, activation in this region has been observed in paradigms with conflicting emotional prosody and lexico-semantic cues (Mitchell, 2006; Schirmer et al., 2004; Schirmer & Kotz, 2006), which may reflect increased

semantic processing demands. Schirmer and colleagues (2004) have also suggested that this brain region is the site of interaction between emotional prosody and word valence. In the current study, increased event-related activity in the left frontal regions (BA6/44) was observed for sad intonations, despite the use of semantically neutral sentences. Given that the current study did not use stimuli with incongruent lexical content and speech prosody, a conflict resolution explanation would not readily explain the observed findings. Moreover, Kotz et al. (2003) found frontal activation in both normal speech and prosodic speech (filtered speech), suggesting that frontal activations cannot solely reflect the processing of specific linguistic information (syntax or lexico-semantic) as this information was eliminated in the prosodic speech condition.

Grimshaw (1998) argued that prosody response selection may always require some left hemisphere input because of the need to apply an emotion label. Schirmer & Kotz (2006) further elaborated that the response to emotional prosody alone may be localized to the right lateral temporal lobe. However, explicit emotional judgements are needed to activate frontal lobe regions such as the inferior frontal gyrus, which are important for assigning a verbal label and integration into language processing. In the current study, participants engaged in a forced-choice emotion categorization task, thus the involvement of the left inferior gyrus is not unexpected given these past findings. However, such an explanation would not help account for the differential recruitment of left frontal regions for sad intonations.

One possible explanation is that recognition of sad prosody may entail reactivation of premotor cortex representing face/mouth and the vocalization apparatus and relevant for facial expressions of sadness and crying vocalizations. Interestingly, the most effective visual stimuli to promote sadness are visual scenes containing sad or crying faces rather than objects or scenes. In contrast, fear or disgust can easily be elicited by scenes, animals (e.g., a snake, a spider, a bear) or objects (e.g., a gun). An alternative and possibly complementary interpretation of this sad-specific effect comes from neuroimaging studies of induced sadness and depression. The left ventral prefrontal cortex (BA45) has been found to have elevated resting state activity in acutely depressed patients (Drevets & Raichle, 1992), and it is also activated during transient sadness induced in healthy subjects using autobiographical scripts such as "think sad thoughts"

(Pardo, Pardo & Raichle, 1993) or other verbal instructions to facilitate rather than suppress mood provocation (Bunge, Ochsner, Desmond, Glover & Gabrieli, 2001). The same region is also activated in tasks requiring effortful retrieval of semantic information, such as in the verb generation task (Buckner, Raichle & Petersen, 1995) and it may therefore reflect ruminations in case of dysphoria or sad mood. Drevets (2001) proposes that the left ventrolateral PFC has a special status as convergence zone/interface between cognitive systems (e.g., language/semantic and working memory) and the limbic system, as demonstrated by direct connections to the anterior cingulate cortex, insula, orbitofrontal cortex and amygdala (Drevets, 2001), and it may be a main area of conscious access to emotional processing. This interpretation is supported by the efficacy of rapid rate Transcranial Magnetic Stimulation (rTMS) over left dorsolateral Prefrontal cortex in transient relief of severe depression. rTMS may prevent ruminations but also have a transynaptic effect on limbic cortex (Paus, Castro-Alamancos & Petrides, 2001). A follow-up study using fMRI would help clarify the current findings.

3.3.7 Theories on Emotion and Prosodic Processing

The hemispheric asymmetries between different emotional intonations observed in the current study fail to follow a predictable pattern lateralization based on traditional neuropsychological models of emotion processing. Overall, a right hemisphere advantage for processing emotions was observed, providing some support for the right hemisphere theory. Moreover, fearful intonations elicited quicker and greater activation in the right hemisphere, when compared to the left hemisphere and to other emotions, which is in line with both the right-hemisphere theory and the valence theory of emotion processing. On the other hand, sad intonations showed preferential processing within the left hemisphere, which is somewhat contradictory to both the right-hemisphere theory and the valence theory of emotion processing, as neither posits an advantage for negative emotions in the left hemisphere. This suggest that the recognition of affective prosody follows a unique pattern of lateralization.

Alternatively, it has been proposed that the lateralization of prosodic processing may vary as a function of the acoustic parameters of prosody, such as fundamental frequency (F_0), intensity or duration. For example, it has been suggested that F_0 is

processed in the right hemisphere, while intensity and duration are processed in the left hemisphere (e.g., Van Lancker & Sidtis, 1992; Zatorre, Belin, & Penhune, 2002). It may be that the recognition of sad intonations relies largely on acoustic cues processed in the left hemisphere, for instance duration of the speech sample. On the other hand, fearful intonations, which show a right hemisphere bias, may rely largely on the processing of pitch information. A recent study suggests that pitch, among other acoustic features, is particularly important for distinguishing fearful intonations from other intonations, whereas rate of utterance appears important for distinguishing sad intonations (Sobin & Alpert, 1999).

3.3.8 Strengths, Weaknesses and Limitations

There are several important issues to consider in relation to the stimuli used. Affective prosody is a complex function where acoustic cues are inherently related to expressing and perceiving affective prosody. The stimuli in this study were normalized for mean amplitude over the duration of the entire stimulus in order to remove the possible cue of perceived loudness in recognizing the emotional prosodies. However, the stimuli were not processed or manipulated in any other way. Thus, the paradigm used does not allow for a differentiation between neural processes associated with the acoustic cues of the stimuli and the cognitive task of recognizing and categorizing the emotions conveyed by the stimuli. The goal of the current study was not to address this question, but rather to address the question of differences between emotional prosodies, an issue that has been relatively ignored. As the current results highlight, it is important to consider emotion-specific differences in the processing of affective prosody, which can provide valuable information about the processing of emotions in the audiory domain.

Affective prosody also draws on linguistic information. Disentangling the different levels of linguistic and paralinguistic information is challenging, yet this is an important concern when studying affective prosody. Attempts were made to reduce the semantic processing of the stimuli; however, this was done while trying to balance other goals of the study. These involved employing a set of natural-sounding vocal stimuli that would generalize to everyday situations while maintaining a low semantic demand. As such, the use of spoken utterances was preferred to nonsensical stimuli. Moreover, the

same lexical sentences were used for each emotional intonation and pooled together, minimizing the differences between sentences and maximizing their shared attributes. Neutral lexical content, rather than emotionally-valenced content, was aimed at minimizing the semantic demand of the stimuli and helping to isolate the prosodic effects. Repeating the same sentences over a number of trials may also have reduced the semantic or linguistic processing of the stimuli, thus allowing participants to focus on prosodic aspects of the utterances. Despite these implementations, the results cannot be unequivocally attributed to the processing of emotional prosody alone, free of linguistic processing. In order to better clarify the results and isolate the processing of the emotional content of the prosodic stimuli from the acoustic and linguistic features, follow-up studies that control each of these variables would be needed.

There are many ways to conceptualize emotions. Basic emotion models represent the most popular conceptualization of emotions. Such models suggest fundamental emotions such as anger, fear, joy and sadness that have evolved as adaptive emotional strategies and each have their own eliciting conditions and specific physiological, expressive, and behavioral reaction patterns (Scherer, 2000). In contrast, dimensional models suggest that emotional states can be characterized by one or more dimensions, such as activation/arousal, valence or pleasantness. Although debates continue with regards to the conceptualization of emotions, the use of dimensional models in physiological and neuropsychological emotion research is growing, with links between phylogenetically continuous approach-avoidance mechanisms and positive-negative valence being made (Davidson, 1992). The use of a dimensional characterization of affective prosody may be well-suited for understanding the neural mechanisms underlying the processing of affective cues in speech and may provide further clarification in how different emotional prosodies are processed by the brain.

The current study only included females, in order to control for effects of gender as a confounding variable. While this allows for the isolation of effects without the loss of statistical power, the findings need to be extended to males. Moreover, given past findings of differences in cross-gender effects between the gender of the listener and that of the portrayer (Sokhi et al., 2005), employing a cross-gender paradigm may provide additional insight in how prosody is processed. Also, as brain laterality can be affected by

handedness (Brunswick & Rippon, 1994), the sample was predominantly right-handed, save for one participant who was left-handed. Thus, the current conclusions are limited to right-handed individuals.

3.3.9 Future Directions

Functional neuroimaging could be used to clarify the present findings. The increased spatial resolution of fMRI would provide better localization of activity, which could confirm the activation of premotor regions for mouth/laryx or limbs, depending on the emotional intonation. Similarly, it could help confirm or rule-out the alternative interpretation of rumination/convergence zone between cognitive systems proposed for sad intonations. Future studies could also investigate the recruitment of these areas in two parallel streams of information. Emotional information can be conveyed by various means of communication, such as propositional content, speech intonation, facial expression, and gestures. Information on the processing of emotional information can be gleaned from studies on multimodal integration, a topic of growing research interest. For instance, it has been shown that when conflict exists between lexical content and affective prosody, prosody tends to be the most important source of information upon which decisions are made (Mehrabian, 1972). In studies of audio-visual integration, information in one stream can bias the processing of information in another stream, even under explicit instructions to ignore one of the modalities (de Gelder & Vroomen, 2000; Massaro & Egan, 1996). However, prosody in general is congruent with the semantic meaning of an utterance (Wambacq & Jerger, 2004). Moreover, as Ethofer and colleagues note (2006), categorization of emotions based on two conflicting streams of information may not reflect common demands on emotional information processing in everyday life. Thus, it would be helpful to study two streams of emotional information and identify common neural substrates recruited when these streams are not in conflict.

Past studies suggest that personality traits are related to differences in the neural processing of emotions. For instance, introverted individuals show right hemisphere dominance for affective prosody processing, while extraverted individuals show left hemisphere dominance (McNeely & Netley, 1998). This is in line with findings from the developmental literature where withdrawal is associated with higher relative right frontal

activity, whereas approach behaviour is associated with higher relative left activity (Davidson & Fox, 1989). In the recognition of emotional faces, high-trait socially anxious individuals show enhanced sensitivity for fearful faces, whereas low-trait individuals show sensitivity for happy faces (Richards et al., 2002). A similar effect has yet to be explored in the recognition of affective prosody. The consideration of personality traits such as trait and state anxiety, neuroticism, depression, shyness and so forth may mediate the relationship between emotions and brain responses. The inclusion of individual traits in future studies on the processing of affective prosody may yield important information in how emotions are processed by the brain.

An understanding of the mechanisms involved in the processing of affective prosody in health can yield important clues about deficits observed among certain clinical populations. For instance, one may expect a delayed or diminished response within the somatomotor cortex during the recognition of affective prosody among individuals with autism. Such a hypothesis would need to be tested directly. While the causes remain unknown, several streams of research support the hypothesis of a dysfunctional mirror neuron system in accounting for some of the observed social and emotional deficits characteristic of autism (e.g., Hadjikhani et al., 2006; Oberman et al., 2005). Thus, the understanding of affective prosody in a normal population can shed light on subsequent investigations among individuals with deficits in nonverbal emotion processing. In turn, investigating the abnormalities among a clinical population could add knowledge to the understanding of important mechanisms and substrates involved in emotion processing in general.

3.3.10 Conclusion

In conclusion, the processing of affective prosody in the current study showed differentiations between emotions. Behaviorally, angry and fearful intonations were recognized more quickly, suggesting an evolutionary advantage for threat detection in auditory stimuli. Event-related potentials revealed an overall effect of prosody, with activity peaking around 1 sec and primarily localized in the inferior part of the right anterior temporal pole, consistent with past studies indicating a relatively right-lateralized dominance for affective prosody, and emotional processing in general. An extended early

fear-specific ERP effect (400-600 ms) was localized in the superior part of the precentral gyrus and middle frontal gyrus, which may be related to the need to re-activate or have access to fear-specific somatomotor representations in order to recognize this intonation. A later sad-specific ERP effect (800-1100 ms) was localized in the inferior part of the left premotor cortex and extended to Broca's area, which may also be related to the need to re-activate premotor cortex in order to recognize this intonation. Alternatively, the sad-specific effect may be related to rumination and working memory associated with sad thoughts. The current results suggest that different emotions recruit distinct premotor regions during affective prosody recognition. These results also highlight the importance of considering spatial-temporal differences between emotional intonations when studying affective prosody processing.

3.4 Figures

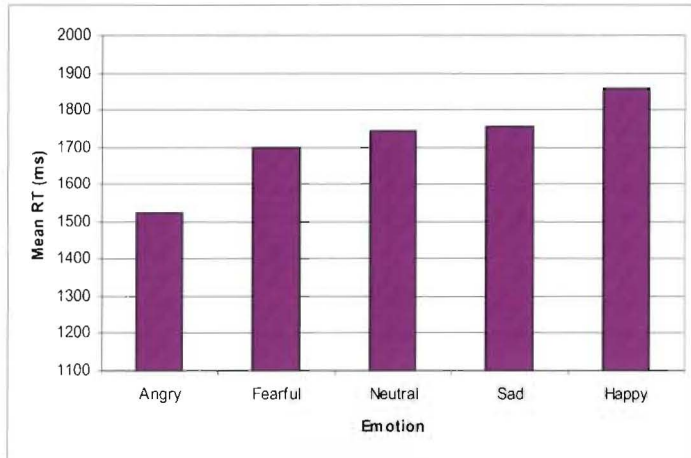**Figure 3. 1   Mean reaction time for each emotion**

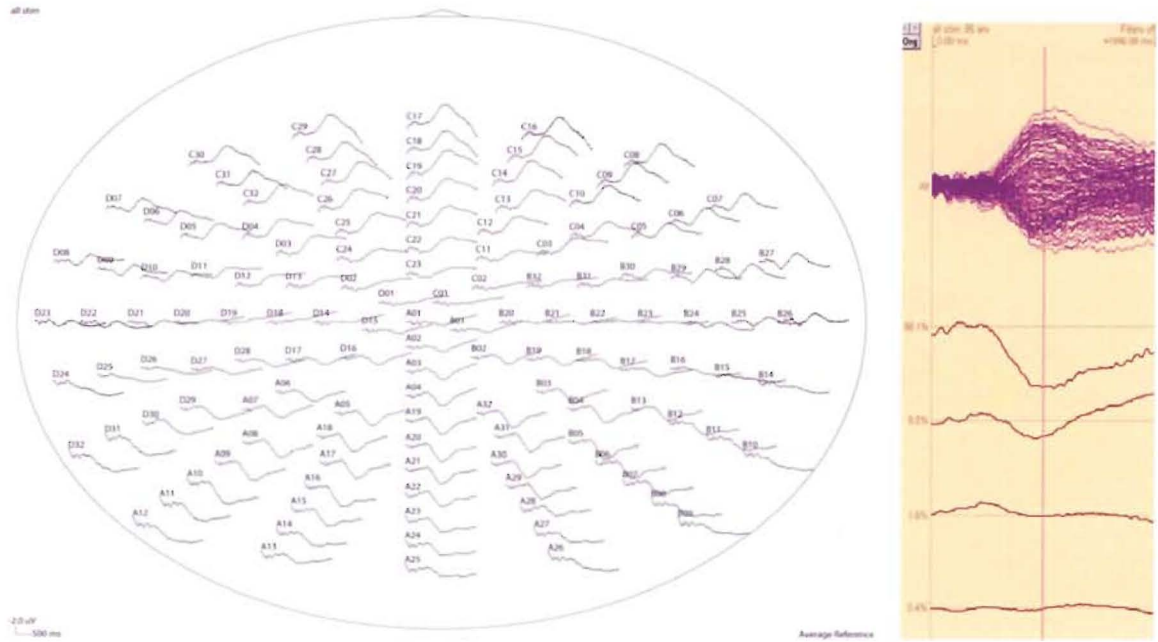**Figure 3. 2   Scalp topography and Global Field Power/PCA for grand-average of all emotions**

Figure 3. 3   Distribution of activity in brain space
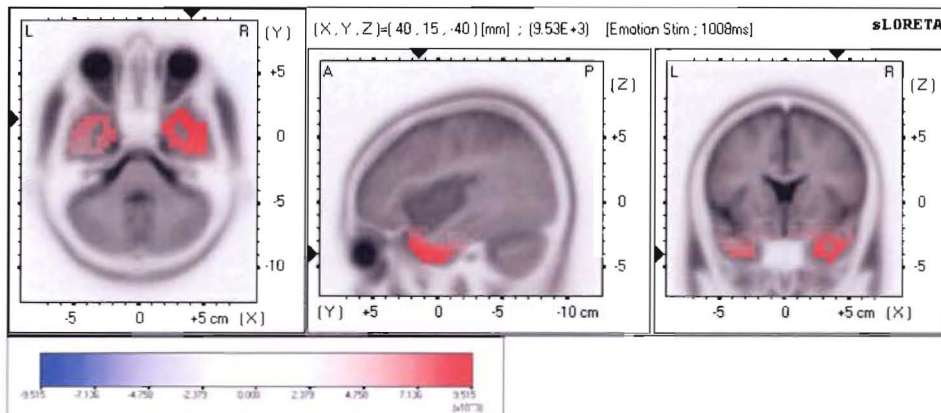
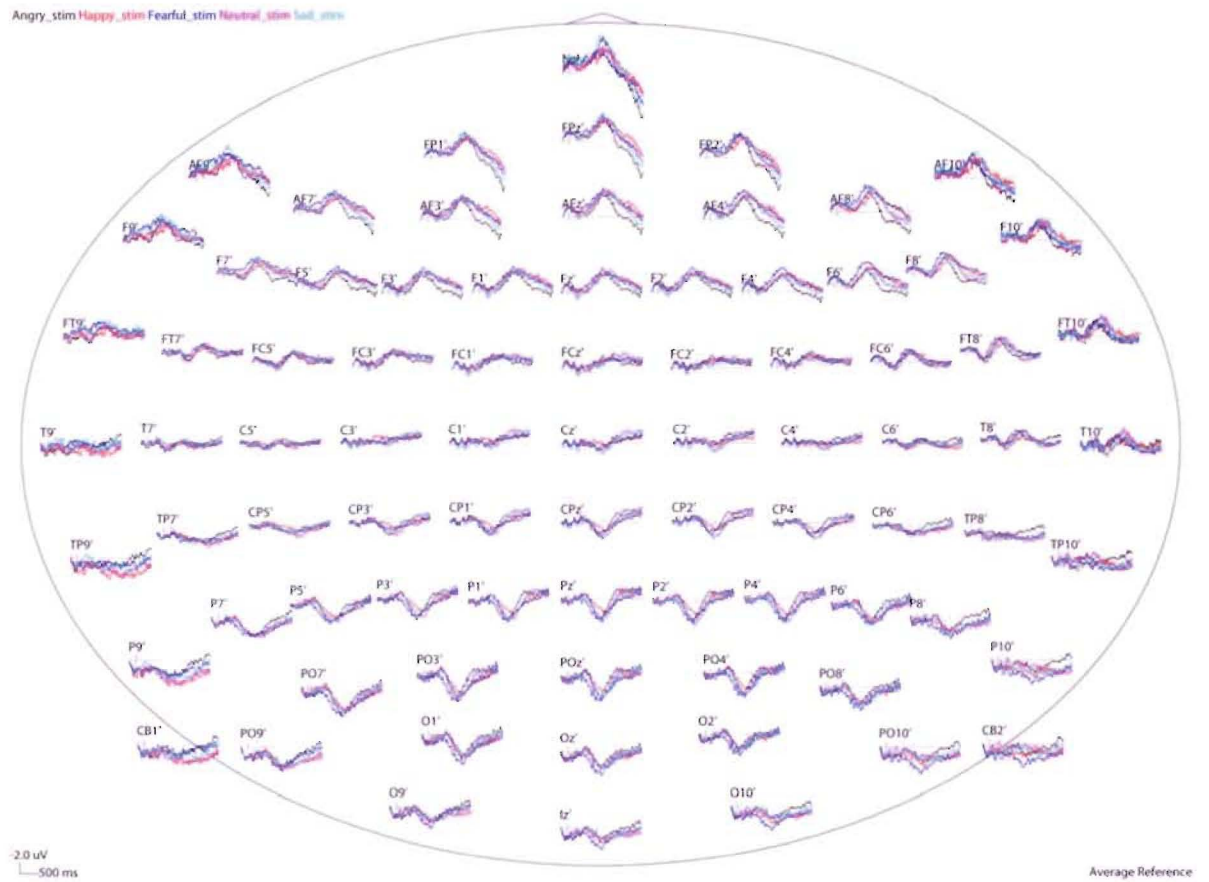**Figure 3. 4   Scalp topography for each emotion**

Figure 3. 5   Mean amplitude for left and right frontal regions (400-500 ms)
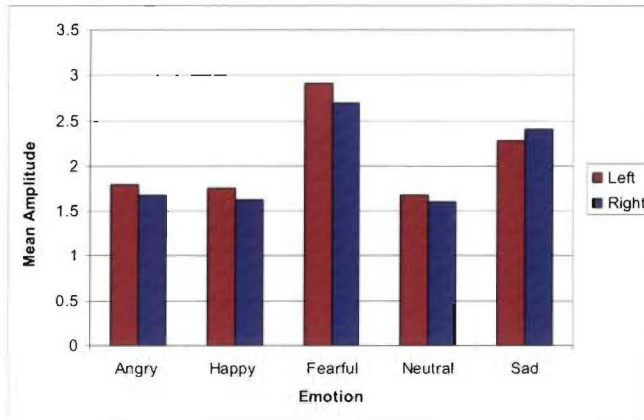

Figure 3. 6   Mean amplitude for left and right temporal/prefrontal regions (500-600 ms)
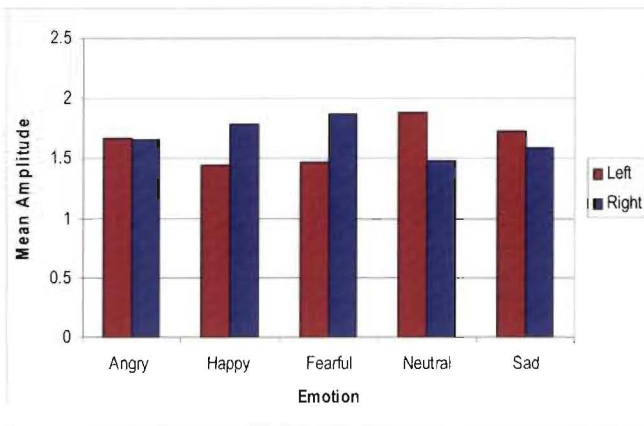

Figure 3. 7   Mean amplitude for left and right frontal regions (800-1100 ms)

Figure 3. 8   Source localization of peak differences in 400-500 ms time window

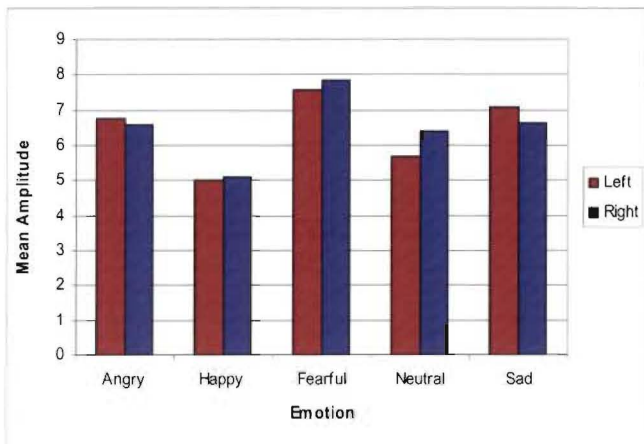*Fearful versus Neutral*

Fearful versus Angry
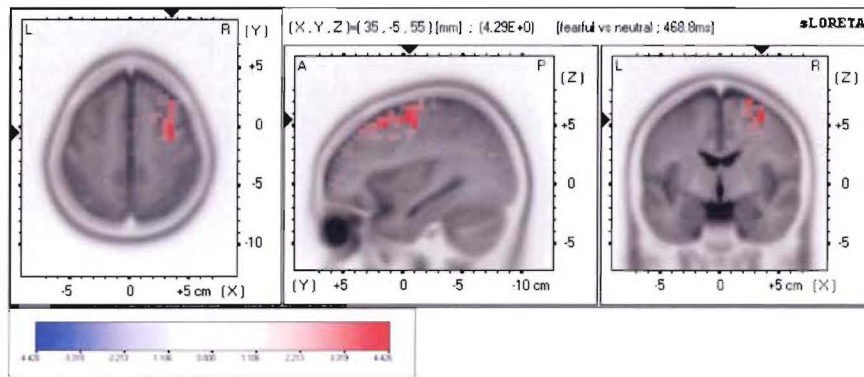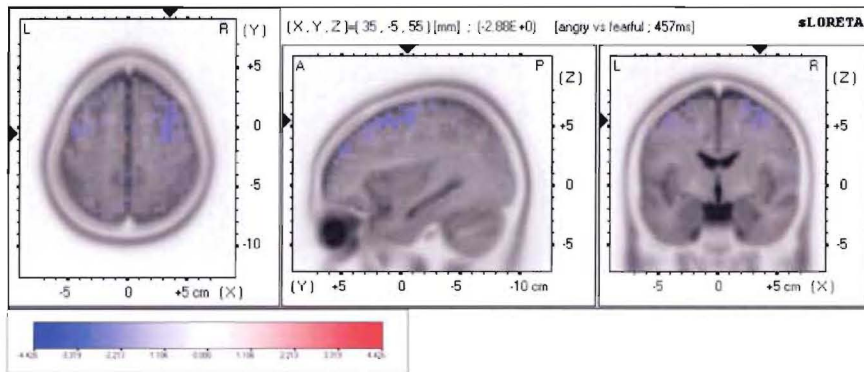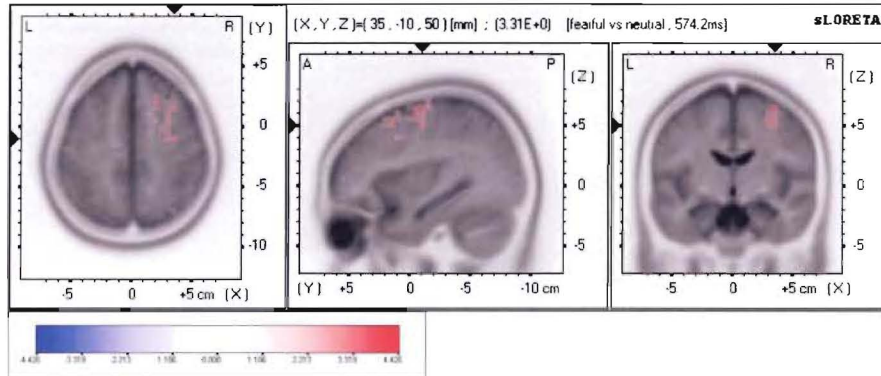
Figure 3. 9   Source localization of peak differences in 500-600 ms time window

Fearful versus Neutral

Fearful versus Angry
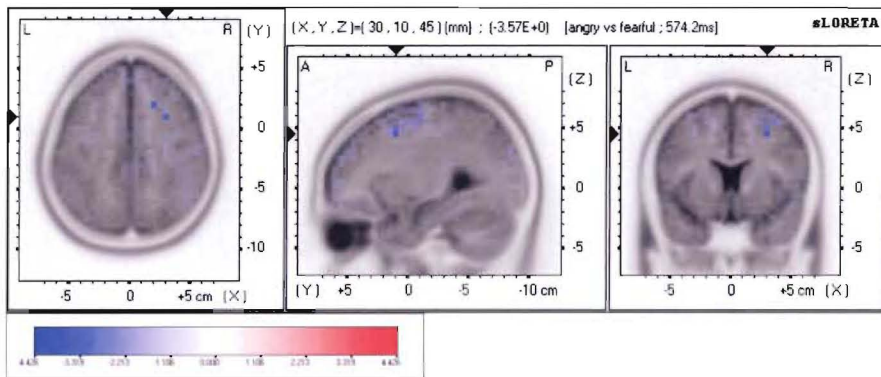
Figure 3. 10   Source localization of peak differences in 800-1100 ms time window

Sad versus Happy
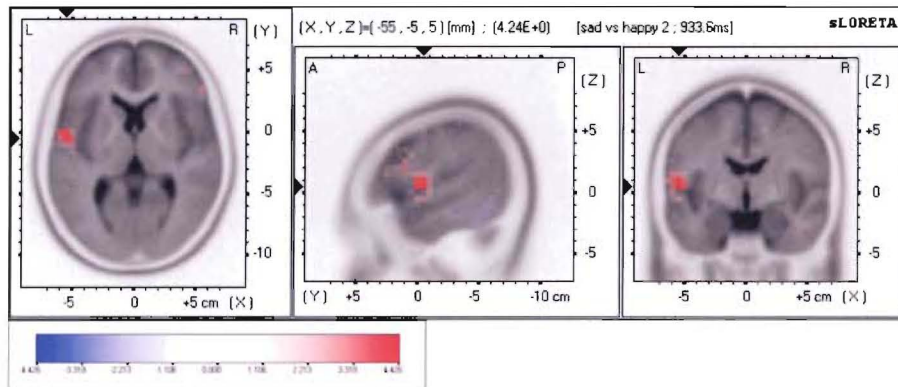
Sad versus Neutral
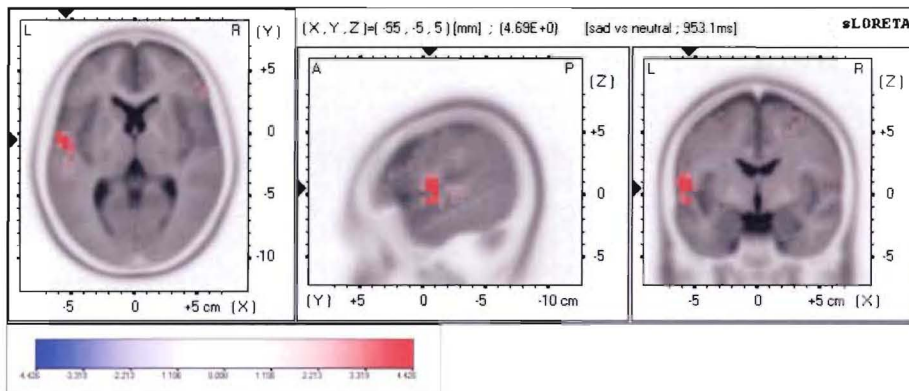
Neutral versus Happy
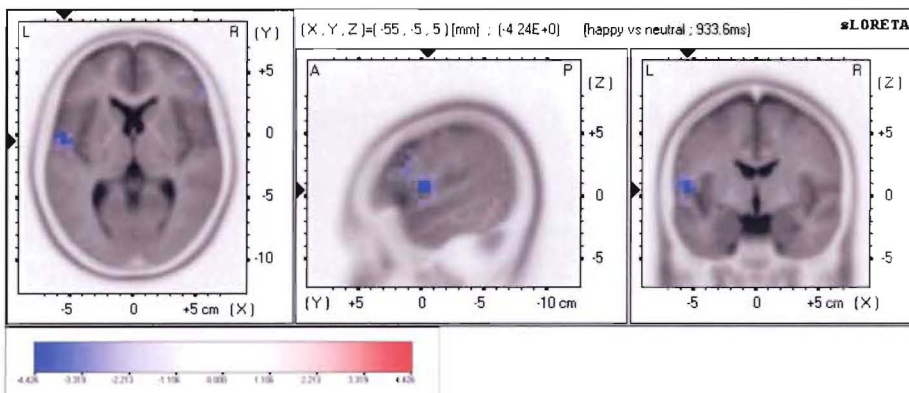
3.5   Tables

Table 2. 1   Accuracy ratings for selected stimuli set

Condition	Mean %	Min %	Median %
Angry	100	100	100
Neutral	100	100	100
Sad	100	100	100
Happy	96	73	100
Fearful	91	75	95

Table 3. 1   Reaction time (RT) and accuracy for each emotion

Condition	Mean RT (ms)	STD RT (ms)	Median RT (ms)	Mean %
Angry	1525.17	445.06	1438.24	98.64
Fearful	1698.51	335.55	1621.65	96.37
Neutral	1744.20	396.76	1642.49	96.49
Sad	1754.25	402.89	1639.97	96.20
Happy	1857.04	420.76	1762.74	97.68

Table 3. 2   Statistical results for source localization in the 400-600 ms time window

Region	MNI x y z	Fearful vs Neutral			Fearful vs Angry			Fearful vs Happy			Fearful vs Sad			Angry vs Neutral		
		Max t	Max Latency	Time Extent	Max t	Max Latency	Time Extent	Max t	Max Latency	Time Extent	Max t	Max Latency	Time Extent	Max t	Max Latency	Time Extent
R superior premotor PrC 6	35 -5 55	4.29	468ms	410-500	2.88	457ms	441-500	-	-	-	-	-	-	-	-	-
PrC/MFG 6/8	35 15 50	3.08	574 ms	500-589	3.42	574 ms	500-600	-	-	-	-	-	-	-	-	-

Table 3. 3   Statistical results for source localization in the 800-1100 ms time window

Region	MNI x y z	Sad vs Happy			Sad vs Neutral			Neutral vs Happy			Angry vs Happy			Fearful vs Happy		
		Max t	Max Latency	Time Extent	Max t	Peak Latency	Time Extent	Max t	Max Latency	Time Extent	Max t	Max Latency	Time Extent	Max t	Max Latency	Time Extent
L inferior premotor PrC/ IFG BA 6/44	-55 -5 5	4.24	934ms	867-1100	4.69	953ms	869-957 1000-1100	4.24	934	875-1100	-	-	-	-	-	-

3.6 Reference List

Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion in Neurobiology, 12*, 169-177.

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

Adolphs, R., Damasio, H., Tranel, D., Cooper, G., & Damasio, A. R. (2000). Role for somatosensory cortices in the visual recognition of emotion as revealed by 3-D lesion mapping. *Journal of Neuroscience, 20*, 2683-2690.

Alter, K., Rank, E., Kotz, S. A., Toepel, U., Besson, M., Schirmer, A., & Friederici, A. D. (2003). Affective encoding in the speech signal and in event-related brain potentials. *Speech Communication, 40*, 61-70.

Andersson, S., Krogstad, J. M., & Finset, A. (1999). Apathy and depressed mood in acquired brain damage: Relationship to lesion localization and psychophysiological reactivity. *Psychological Medicine, 29*, 447–456.

Armony, J. L., & LeDoux, J. (2000). How danger is encoded: Toward a systems, cellular, and computational understanding of cognitive-emotional interactions in fear. In M. S. Gazzaniga (Ed.). *The new cognitive neurosciences* (2nd ed., pp. 1067–1079). Cambridge, MIT Press.

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

Bargh, J. A., Chaiken, S., Govender, R. & Pratto, F. (1992). The generality of the attitude activation effect. *Journal of Personality and Social Psychology, 62*, 893-912.

Beaucousin, V., Lacheret, A., Turbellin, M.-R., Morel, M., Mazoyer, B., & Tzourio-Mazoyerm N. (2007). FMRI study of emotional speech comprehension. *Cerebral Cortex, 17*, 339-352.

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

Borod, J. C., Bloom, R. L., Brickman, A. M., Nakhutina, L., & Curko, E. A. (2002). Emotional processing deficits in individuals with unilateral brain damage. *Applied Neuropsychology, 9*, 23–36.

Borod, J. C., Cicero, B. A., Obler, L. K., Welkowitz, J., Erhan, H. M., Santschi, C.,
Grunwald, I. S., Agosti, R. M., & Whalen, J. R. (1998). Right hemisphere
emotional perception: Evidence across multiple channels. *Neuropsychology, 12,*
446–458.

Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., et al. (1994).
The role of the right hemisphere in the interpretation of figurative aspects of
language. *Brain, 117,* 1241–1253.

Bowers, D., Bauer, R. M., & Heilman, K. M. (1993). The nonverbal affect

Bowers, D., Coslett, H. B., Bauer, R. M., Speedie, L. J., & Heilman, K. M. (1987).
Comprehension of emotional prosody following unilateral hemispheric lesions:
Processing defect versus distraction defect. *Neuropsychologia, 25,* 317-328.

Bozikas, V. P., Kosmidis, M. H., Anezoulaki, D., Giannakou, M., Andreou, C., &
Karavatos, A. (2006) Impaired perception of affective prosody in schizophrenia.
Journal of Neuropsychiatry Clinical Neuroscience, 18, 81-85.

Breitenstein, C., Daum, I., & Ackermann, H. (1998). Emotional processing following
cortical and subcortical brain damage: Contribution of the fronto-striatal circuitry.
Behavioural Neurology, 11, 29–42.

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

Brothers, L., & Ring, B. (1992). A neuroethological framework for the representation of
minds. *Journal of Cognitive Neuroscience, 4,* 107-118.

Brunswick, N., & Rippon, G. (1994). Auditory event-related potentials, dichotic listening
performance and handedness as indices of lateralisation in dyslexic and normal
readers. *International Journal of Psychophysiology, 18,* 265–275.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Gallese, V., Seitz, R. J., Zilles, K.,
Rizzolatti, G., & Freund, H. J. (2001). Action observation activates premotor and
parietal areas in a somatotopic manner: An fMRI study. *The European Journal of
Neuroscience, 13,* 400-404.

Buchanan, T. W., Lutz, K., Mirzazade, S., Specht, K., Jon Shah, N., Zilles, K., & Jancke, L. (2000). Recognition of emotional prosody and verbal components of spoken language : An fMRI study. *Cognitive Brain Research, 9,* 227-238.

Buckner, R. L., Raichle, M. E., & Petersen, S. E. (1995). Dissociation of human prefrontal cortical areas across different speech production tasks and gender groups. *Journal of Neurophysiology, 74,* 2163-2173.

Bulman-Fleming, M. B., & Bryden, M. P. (1994). Simultaneous verbal and affective laterality effects. *Neuropsychologia, 32,* 787-797.

Bunge, S. A., Ochsner, K. N., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (2001). Prefrontal regions involved in keeping information in and out of mind. *Brain, 124,* 2074-2086.

Burton, L. A., Gilliam, D., Flynn, S., & Labar, D. (1999). Affective verbal memory in patients with temporal lobe epilepsy. *Applied Neuropsychology, 6,* 115–120.

Calder, A., Lawrence, A., & Young, A. W. (2001). Neuropsychology of fear and loathing. *Nature Reviews Neuroscience, 2,* 352–363.

Cancelliere, A., & Kertesz, A. (1990). Lesion localization in acquired deficits of emotional expression and comprehension. *Brain and Cognition, 13,* 133–147.

Canli, T., Desmond, J. E., Zhao, Z., Glover, G., & Gabrieli, J. D. E. (1998). Hemispheric asymmetry for emotional stimuli detected with fMRI. *NeuroReport, 9,* 3233-3239.

Canli, T., Zhao, Z., Desmond, J. E., Glover, G., & Gabrieli, J. D. E. (1999). fMRI identifies a network of structures correlated with retention of positive and negative emotional memory. *Psychobiology, 27,* 441-452.

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.

Cicero, B.A., Borod, J.C., Santschi, C., Erhan, H.M., Obler, L.K., Agosti,

Cross, E. S., Hamilton, A. F. D., & Grafton, S. T. (2006). Building a motor simulation de novo: Observation of dance by dancers. *Neuroimage, 31,* 1257-1267.

Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain.* New York: Grosset/Putman.

45

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 351,* 1413-1420.

Davidson, R. J. & Fox, N. (1989). Frontal brain asymmetry predicts infants' response to maternal separation. *Journal of Abnormal Psychology, 98,* 127-131.

Davidson, R. J. (1992). Prolegomenon to the structure of emotion: Gleanings from neuropsychology. *Cognition and Emotion, 6,* 245-268.

Davidson, R. J. (1995). Cerebral asymmetries, emotion, and affective style. In R. J. Davidson and K. Hugdahl (Eds.), *Brain asymmetry* (pp. 361-387). Cambridge, MA: MIT Press.

Davidson, R. J., & Irwin, W. (1999). The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Sciences, 3,* 11-20.

Davis, K. L. (2003). Comprehension of emotion in neurodegenerative disorders. *Sciences and Engineering, 63,* 3489.

de Zubicaray, G. I. (2006). Neuroimaging and clinical neuropsychological practice. In P. J. Snyder, P. D. Nussbaum and D. L. Robins (Eds.) *Clinical neuropsychology: A pocket handbook for assessment (2nd ed.)* (pp. 56-74). Washington, D. C.: American Psychological Association.

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of cognitive updating? *Behavioral and Brain Sciences, 11,* 357-427.

Drevets, W. C. (2001). Neuroimaging and neuropathological studies of depression: Implications for the cognitive-emotional features of mood disorders. *Current Opinion in Neurobiology, 11,* 240-249.

Drevets, W. C., & Raichle, M. E. (1992). Neuroanatomical circuits in depression: Implications for treatment mechanisms. *Psychopharmacology Bulletin, 28,* 261-274.

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Perception and Psychophysics, 63,* 1004–1013.

Edwards, J., Pattison, P. E., Jackson, H. J., & Wales, R. J. (2001). Facial affect and
  affective prosody recognition in first-episode schizophrenia. *Schizophrenia*
  Research, 48, 235-253.

Ekman, P. (1979). About brows: Emotional and conversational signals. In M. von
  Cranach, K. Foppa, W. Lepenies, and D. Ploog (Eds.), *Human Ethology* (pp. 169-
  202). Cambridge: Cambridge University Press.

Engels, A. S., Heller, W., Mohanty, A., Herrington, J. D., Banich, M. T., Webb, A. G., &
  Miller, G. A. (2007). Specificity of regional brain activity in anxiety during
  emotion processing. *Psychophysiology, 44,* 352-363.

Esslen, M., Pascual-Marqui, R.D., Hell, D., Kochi, K., & Lehman, D., (2004). Brain
  areas and time course of emotional processing. *Neuroimage, 21,* 1189-1203.

Ethofer, T., Anders, S., Erb, M., Herbert, C., Wiethoff, S., Kissler, J., Grodd, W., &
  Wildgruber, D. (2006). Cerebral pathways in processing of affective prosody: A
  dynamic causal modeling study. *NeuroImage, 30,* 580-587.

Fletcher, P. C., Happe, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S. J., et al.
  (1995). Other minds in the brain: A functional imaging study of 'theory of mind'
  in story comprehension. *Cognition, 57,* 109–128.

Fox, E., Griggs, L., & Mouchlianitis, E. (2007). The detection of fear-relevant stimuli:
  Are guns noticed as quickly as snakes? *Emotion, 7,* 691-696.

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A., & Dutton, K. (2000). Facial
  expressions of emotion: Are angry faces detected more efficiently? *Cognition and*
  Emotion, 14, 61–92.

Gardner, H., Ling, P. K., Flamm, L., & Silverman, J. (1975). Comprehension and
  appreciation of humourous material following brain damage. *Brain, 98,* 399-412.

George, M. S., Parekh, P. I., Rosinsky, N., Ketter, T. A., Kimbrell, T. A., Heilman, K.
  M., Herscovitch, P., & Post., R. M. (1996). Understanding emotional prosody
  activates right hemisphere regions. *Archives of Neurology, 53,* 665-670.

Geschwind, N. (1970). The organization of language and the brain. *Science, 170,* 940-
  944.

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,* 145-145.

Grimshaw, G.M. (1998). Integration and interference in the cerebral hemispheres: Relations with hemispheric specialization. *Brain and Cognition, 36,* 108–127.

Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin, 4,* 665-670.

Hansen, C. H., & Hansen, R. D. (1988). Finding the face in the crowd: An anger superiority effect. *Journal of Personality and Social Psychology, 54,* 917–924.

Herath, P., Kinomura, S., & Roland, P. E. (2001). Visual recognition: Evidence for

Herrero, J. V., & Hillix, W. A. (1990). Hemispheric performance in detecting prosody: A competitive dichotic listening task. *Perceptual and Motor Skills, 71,* 479-486.

Hird, K., & Kirsner, K. (1998). Control processes in prosody. In K. Kirsner, M. Anderson, C. McLeod, M. Mayberry, A. O'Brien & C. Speelman (Eds.), *Implicit and explicit mental processes* (pp. 201-218). Mahwah, NJ: Lawrence Erlbaum.

Huang, Y., Tang, S., Helmeste, D., Shioiri, T., & Someya, T. (2001). Differential judgement of static facial expressions of emotions in three cultures. *Psychiatry and Clinical Neurosciences, 55,* 479-483.

Imaizumi, S., Mori, K., Kiritani, S., Kawashima, R., Sugiura, M., Fukuda, H., Itoh, K., Kato, T., Nakamura, A., Hatano, K., Kojima, S., & Nakamura, K. (1997). Vocal identification of speaker and emotion activates different brain regions. *Neuroreport, 8,* 2809-2812.

Ishai, A., Ungerleider, L. G., Martin, A., & Haxby, J. V. (1999). Distributed representation of objects in the human ventral visual pathway. *Proceedings of the National Academy of Sciences, 96,* 9379–9384.

Jay, T. B. (2003). *The psychology of language.* Toronto: Prentice Hall.

Johnstone, T., van Reekum, C. M., 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.

Karow, C. M., & Connors, E. C. (2003). Affective communication in normal and brain-damaged adults: An overview. *Seminars in Speech and Language, 24,* 69-91.

Kass, J. H., & Hackett, T. A. (1999). 'What' and 'where' processing in auditory cortex. *Nature Neuroscience, 2,* 1045-1047.

Kirita, T., & Endo, M. (1995). Happy-face advantage in recognizing facial expressions. *Acta Psychologica, 89,* 149–163.

Knapp, M. L. & Hall, J. A. (1997). *Nonverbal communication in human interaction (4th ed.),* Belmont, CA: Wadworth, Thompson Learning.

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

Kotz, S. A., Meyer, M., & Paulmann, S. (2006). Lateralization of emotional prosody in the brain: An overview and synopsis on the impact of study design. *Progress in Brain Research, 156,* 285-294.

Kotz, S. A., Meyer, M., Alter, K., Besson, M., von Cramon, D. Y., & Friederici, A. D. (2003). On the lateralization of emotional prosody: An event-related functional MR investigation. *Brain and Language, 86,* 366-376.

LeDoux, J. (1996). *The emotional brain.* New York: Simon & Schuster.

Leppänen, J. M., & Hietanen, J. K. (2003). Affect and face perception: Odors modulate the recognition advantage of happy faces. *Emotion, 3,* 315–326.
lexicon. *Neuropsychology, 7,* 433–444.
lexicon: Theoretical perspectives from neuropsychological studies of two distinctive mechanisms from a PET study. *Human Brain Mapping, 12,* 110–119.

Ley, R., & Bryden, M. P. (1982). A dissociation of right and left hemisphere effects for recognizing emotional tone and verbal content. *Brain and Cognition, 1,* 3-9.

Liotti, M., & Tucker, D. M. (1995). Emotion in asymmetric corticolimbic networks. In R. J. Davidson and K. Hugdahl (Eds.), *Brain asymmetry* (pp.389-423). London: MIT Press.

Maguire, E. A., Frith, C. D., & Morris, R. G. M. (1999). The functional neuroanatomy of comprehension and memory: The importance of prior knowledge. *Brain, 122,* 1839-1850.

Mandal, M. K., Harizuka, S., Bhushan, B., & Mishra, R. C. (2001). Cultural variation in hemifacial asymmetry of emotion expressions. *The British Journal of Social Psychology, 40,* 385-398.

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., et al.
(1993). The cortical representation of speech. *Journal of Cognitive Neuroscience,
5,* 467–479.

McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression
processing and their development in infants, children, and adolescents.
Psychological Bulletin, 126, 38-59.

McNeely, H., & Netley, C. (1998). Right hemisphere lateralization of emotional prosody
recognition predicts introverted personality in left handers. *Brain and Cognition,
37,* 51-54.

Mehrabian, A. (1972). *Non-verbal communication.* New York: Aldine-Atherton.

Mesulam, M. M. (1998). From sensation to cognition. *Brain, 121,* 1013–1052.

Meyer, M., Alter, K., Friederici, A. D., Lohmann, G., & Von Cramon, D. Y. (2002).
fMRI reveals brain regions mediating slow prosodic modulations in spoken
sentences. *Human Brain Mapping, 17,* 73-88.

Mitchell, R. L. (2006). How does the brain mediate interpretation of incongruent auditory
emotions? The neural response to prosody in the presence of conflicting lexico-
semantic cues. *European Journal of Neuroscience, 24,* 3611-3618.

Mitchell, R. L. C., Elliott, R., Barry, M., Cruttenden, A., & Woodruff, P. W. R. (2003).
The neural response to emotional prosody, as revealed by functional magnetic
resonance imaging. *Neuropsychologia, 41,* 1410-1421.

Nakamura, K., Kawashima, R., Ito, K., Sugiura, M., Kato, K., Nakamura, A., Hatano, K.,
Nagumo, S., Kubota, K., Fukuda, H., & Kojima, S. (1999). Activation of the right
inferior frontal cortex during assessment of facial emotion. *Journal of
Neurophysiology, 82,* 1610–1614.

Nowicki, S., Jr., & Duke, M., (1994). Individual differences in nonverbal communication
of affect: The Diagnostic Analysis of Nonverbal Accuracy Scale (DANVA).
Journal of Nonverbal Behavior, 18, 9-35.

Ochsner, K. N., & Schacter, D. L. (1995).). Emotion in asymmetric corticolimbic
networks. In R. J. Davidson and K. Hugdahl (Eds.), *Brain asymmetry* (pp.163-
193). London: MIT Press.

Ohman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review, 108,* 483-522.

Ohman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology, 80,* 381-396.

Ohman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology, 80,* 381–396.

Pardo, J. V., Pardo, P. J., & Raichle, M. E. (1993). Neural correlates of self-induced dysphoria. *American Journal of Psychiatry, 150,* 713-719.

Pascual-Marqui, R. D., Esslen, M., Kochi, K., & Lehmann, D. (2002). Functional mapping of electric neuronal activity with zero localization error: Standardized low resolution brain electromagnetic tomography (sLORETA). *8th International Conference on Functional Mapping of the Human Brain,* June 2-6, 2002, Sendai, Japan.

Paul, R., Augustyn, A., Klin, A., & Volkmar, F. R. (2005). Perception and production of prosody by speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorder, 35,* 205-220.

Paus, T., Castro-Alamancos, M. A., & Petrides, M. (2001). Cortico-cortical connectivity of the human mid-dorsolateral frontal cortex and its modulation by repetitive transcranial magnetic stimulation. *European Journal of Neuroscience, 14,* 1405-1411.

Pell, M.D., & Leonard, C.L. (2003). Processing emotional tone from speech in Parkinson's disease: A role for the basal ganglia. *Cognitive Affective and Behavioral Neuroscience, 3,* 275–288.

Pihan, H., Altenmuller, E., & Ackermann, H. (1997). The cortical processing of perceived emotion: A DC potential study on affective speech prosody. *Neuroreport, 8,* 623-627.

Pihan, H., Altenmuller, E., Hertrich, I., & Ackermann, H. (2000). Cortical activation patterns of affective speech processing depend on concurrent demands on the subvocal rehearsal system. A DC-potential study. *Brain, 123,* 2338-2349.

R.M., Welkowitz, J., & Grunwald, I.S. (1999). Emotional versus nonemotional
lexical perception in patients with right and left brain damage. *Neuropsychiatry,
Neuropsychology, and Behavioral Neurology, 12,* 255–264.

Rauschecker, J. P. (1998). Cortical processing of complex sounds. *Current Opinions in
Neurobiology, 8,* 516-521.

Reuter-Lorenz, P., & Davidson, R. J. (1981). Differential contributions of the 2 cerebral
hemispheres to the perception of happy and sad faces. *Neuropsychologia, 19,* 609-
613.

Richards, A., French, C. C., Calder, A. J., Webb, B., Fox, R., & Young, A. W. (2002).
Anxiety-related bias in the classification of emotionally ambiguous facial
expressions. *Emotion, 2,* 273-287.

Riemann, B. & McNally, R. J. (1995). Cognitive processing of personally relevant
information. *Cognition and Emotion, 9,* 325-340.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the
recognition of motor actions. *Cognitive Brain Research, 3,* 131-141.

Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S., & Rauschecker,
J. P. (1999). Dual streams of auditory afferents target multiple domains in the
primate prefrontal cortex. *Nature Neuroscience, 2,* 1131-1136.

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.

Rymarczyk, K., & Grabowska, A. (2007). Sex differences in brain control of prosody.
Neuropsychologia, 45, 921-930.

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.

Scherer, K. R. (1989). Vocal correlates of emotion arousal and affective distrubance. In
H. Wagner and A. Manstead (Eds.), *Handbook of psychophysiology: Emotion and
social behavior.* (pp. 165-197). London: Wiley.

Scherer, K. R. (2000). Psychological models of emotion. In J. C. Borod (Ed.), *The
neuropsychology of emotion* (pp. 137-162). New York: Oxford University Press.

Schirmer, A. & Kotz, S. A. (2003). ERP evidence for a sex-specific stroop effect in

emotional speech. *Journal of Cognitive Neuroscience, 15*, 1135-1148.

Schirmer, A., Kotz, S. A., & Friederici, A. D. (2002). Sex differentiates the role of

emotional prosody during word processing. *Cognitive Brain Research, 14*, 228-

233.

Schirmer, A., Kotz, S. A., & Friederici, A. D. (2005). On the role of attentions for the

processing of emotions in speech: Sex differences revisited. *Cognitive Brain

Research, 24*, 442-452.

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.

Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2004). The selective

processing of briefly presented affective pictures: An ERP analysis.

Psychophysiology, 41, 441–449.

Scott, S. K., Blank, C. C., Rosen, S., & Wise, J. S. (2000). Identification of a pathway for

intelligible speech in the left temporal lobe. *Brain, 123*, 2400-2406.

Shaywitz, B. A., Shaywitz, S. E., Pugh, R. T., Constable, R. T., & Skudlarski, R. K.

(1995). Sex differences in the functional organization of the brain for language.

Nature, 373, 607-609.

Shriberg, L. D., Paul, R., McSweeny, J. L., Klin, A., Cohen, D. J., & Volkmar, F. R.

(2001). Speech and prosody characteristics of adolescents and adults with high

functioning autism and Asperger syndrome. *Journal of Speech, Language and

Hearing Research, 44*, 1097-1115.

Smith, S. D., & Bulman-Fleming, M. B. (2006). Hemispheric asymmetries for the

conscious and unconscious perception of emotional words. *Laterality, 11*, 304-

330.

Sobin,C., & Alpert, M. (1999). Emotion in speech: The acoustic attributes of fear, anger,

sadness, and joy. *Journal of Psychologuistic Research, 28*, 347-365.

Sokhi, D.S., Hunter, M. D., Wilkinson, I. D., & Woodruff, P. W. (2005). Male and

female voices activate distinct regions in the male brain. *Neuroimage, 27*, 572-

578.

Trauner, D. A., 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.

Van Lancker, D., & Sidtis, J. J. (1992). The identification of affectiveprosodic stimuli by left- and right-hemisphere-damaged subjects: All errors are not equal. *Journal of Speech and Hearing, 35,* 963–970.

Vigneau, M., Beaucousin, V., Herve, P. Y., Duffau, H., Crivello, F., Houde, O., Mazoyer, B., & Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *Neuroimage, 30,* 1414-1432.

Wambacq, I. J. A., & Jerger, J. F. (2004). Processing of affective prosody and lexical-semantics in spoken utterances as differentiated by event-related potentials. *Cognitive Brain Research, 20,* 427-437.

Wildgruber, D., Hertrich, I., Riecker, A., Erb, M., Anders, S., Grodd, W., & Ackermann, H. (2004). Distinct frontal regions subserve evaluation of linguistic and emotional aspects of speech intonation. *Cerebral Cortex, 14,* 1384-1389.

Wildgruber, D., Pihan, H., Ackermann, H., Erb, M., & Grodd, W. (2002). Dynamic brain activation during processing of emotional intonation: Influence of acoustic parameters, emotional valence and sex. *Neuroimage, 15,* 856–869.

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.

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin, 120,* 3-24.

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Science, 1,* 37–46.