Emotional Prosody Processing in the Schizophrenia Spectrum.

Lucia Monserrat Alba-Ferrara

Ph.D.

Department of Psychology

Durham University

2011

CONTENT

Acknowledgements

My doctorate was supervised by Markus Hausmann and co-supervised by Susanne Weis. I consider myself lucky to have worked with such brilliant, dedicated scientists. They always give constructive advice as well as encouragement. And I would not have produced this work without their guidance and support.

Amanda Ellison has lead the inception of the TMS study and she has given me technical training in the technique. She has also been an excellent progress reviewer. Charles Fernyhough has advised me in the study of the phenomenological aspects of hallucination. I am deeply thankful to both of them.

My interest in neuroimaging has developed during my internship in Washington University, St Louis, Missouri, under the supervision of Prof. Marcus E. Raichle. I am extremely grateful to him and to the member of his team for the training I have received in his lab.

I am grateful for the financial support I have received from the Experimental Psychology Society, which has funded my clinical studies, my internship in the USA and my attendance to the OHBM 2010 meeting.

I would like to thank the postgraduate students at the Department of Psychology of Durham University for creating such a friendly and collaborative atmosphere. Specially, I appreciate all the help I have received from Marco Hirnstein, Rob Lee and John C. Thoresen, as they have always made themselves available to assist me in any requests. I also would like to thank all the patients and healthy controls who have given up many hours of their time to participate in the experiments in this thesis. Finally, I would like to dedicate this thesis to Seth MacFarlane, Trey Parker and Matt Stone as they have comforted me with truly memorable laughter in lonely times.

In loving memory of my auntie Teresa and my granny Martha.

L.M. Alba-Ferrara

Emotional Prosody Processing in the Schizophrenia Spectrum.

THESIS ABSTRACT

Emotional prosody processing impairment is proposed to be a main contributing factor for the formation of auditory verbal hallucinations in patients with schizophrenia. In order to evaluate such assumption, five experiments in healthy, highly schizotypal and schizophrenia populations are presented. The first part of the thesis seeks to reveal the neural underpinnings of emotional prosody comprehension (EPC) in a non-clinical population as well as the modulation of prosodic abilities by hallucination traits. By revealing the brain representation of EPC, an overlap at the neural level between EPC and auditory verbal hallucinations (AVH) was strongly suggested. By assessing the influence of hallucinatory traits on EPC abilities, a continuum in the schizophrenia spectrum in which high schizotypal population mirrors the neurocognitive profile of schizophrenia patients was established. Moreover, by studying the relation between AVH and EPC in non-clinical population, potential confounding effects of medication influencing the findings were minimized. The second part of the thesis assessed two EPC related abilities in schizophrenia patients with and without hallucinations. Firstly, voice identity recognition, a skill which relies on the analysis of some of the same acoustical features as EPC, has been evaluated in patients and controls. Finally, the last study presented in the current thesis, assessed the influence that implicit processing of emotional prosody has on selective attention in patients and controls. Both patients studies demonstrate that voice identity recognition deficits as well as abnormal modulation of selective attention by implicit emotion prosody are related to hallucinations exclusively and not to schizophrenia in general. In the final discussion, a model in which EPC deficits are a crucial factor in the formation of AVH is evaluated. Experimental findings presented in the previous chapters strongly suggests that the perception of prosodic features is impaired in patients with AVH,

resulting in aberrant perception of irrelevant auditory objects with emotional prosody salience which captures the attention of the hearer and which sources (speaker identity) cannot be recognized. Such impairments may be due to structural and functional abnormalities in a network which comprises the superior temporal gyrus as a central element.

INTRODUCTION

Chapter I: Contributions of emotional prosody comprehension deficits to the formation of auditory verbal hallucinations in schizophrenia.

1.1 *Abstract*

Deficits in emotional processing have been widely described in schizophrenia. Associations of positive symptoms with poor emotional prosody comprehension (EPC) have been reported at the phenomenological, behavioral, and neural levels. This review focuses on the relation between emotional processing deficits and auditory verbal hallucinations (AVH). We explore the possibility that the relation between AVH and EPC in schizophrenia might be mediated by the disruption of a common mechanism intrinsic to auditory processing, and that, moreover, prosodic feature processing deficits play a pivotal role in the formation of AVH. The review concludes with proposing a mechanism by which AVH are constituted and showing how different aspects of our neuropsychological model can explain the constellation of subjective experiences which occur in relation to AVH.

1.2. Introduction: Phenomenology of AVH and EPC

Emotional impairment in schizophrenia was first observed by Kraepelin (1919). Disturbances range from flat affect to intense bursts of inappropriate emotions such as anger and fear (Kraepelin, 1971 [1919]). The combination of these emotional disturbances points to one of the paradoxes of schizophrenia: whilst there can be a flattening of affect, as in negative symptoms, there can be also an increase in emotional arousal and reactivity, as in the positive symptoms of psychosis (Aleman & Kahn, 2005). Traditionally, deficits in emotion processing have been linked to negative symptoms such as inefficient social interaction, and apathy and avolition towards social stimuli (Hoekert, Kahn, Pijnenborg, & Aleman, 2007). However, a possible link between emotional processing difficulties, particularly for vocal stimuli, and the presence of hallucinations or delusions (Rossell & Boundy, 2005) has also been reported. In fact, dysfunctional emotion processing may also be associated with the positive symptoms of schizophrenia (Aleman & Kahn, 2005).

This review will focus on the auditory modality. Although deficits in the comprehension of facial emotion seem to relate to negative symptoms of schizophrenia, deficits in the comprehension of vocal emotion are thought to be specific for schizophrenic patients suffering from hallucinations. Additionally hallucinations in schizophrenia tend to be more often auditory than visual (Mueser, Bellack, & Brady, 1990). The congruency in modality between affective processing deficits and abnormal perception might suggest an underlying common mechanism. Auditory modality emotion recognition abilities are likely to be of central importance in the formation of AVH, and the study of this connection promises to reveal a mechanism linking both phenomena.

One of the central characteristics of AVH is that the voices patients hear are often spoken in emotional tones (Copolov, Mackinnon, & Trauer, 2004), most often with a negative emotional valence (Nayani & David, 1996), e.g., angry voices shouting abusive language. One interpretation of the link between AVH and EPC deficits lies in the coincidence in emotional valence between both phenomena. Patients with AVH seem to have difficulties correctly identifying happy prosody, possibly because of a tendency to attribute fear or sadness to any stimulus (Tsoi et al., 2008) and research in schizophrenia suggests that perception and labelling of emotions such as anger and fear may also be impaired (Allen et al., 2004). Regarding perceived intensity of prosodic emotions, it has also been discovered that patients with AVH generally tend to rate frightening stimuli as more intense than controls and patients without AVH (Rossell & Boundy, 2005). Given that patients who hear "malevolent" voices tend to rate them as "very powerful" more often than patients who hear benign or benevolent voices (Birchwood & Chadwick, 1997), these AVH patients might also rate fearful stimuli in EPC tasks as more intense. An alternative possibility is that comorbid persecutory delusions may contribute to this effect.

Contrariwise, an attentional bias towards negative affects in EPC tasks may increase the likelihood of experiencing AVH or even act as a trigger for them. Therefore, it is essential to determine the direction of causation in any such relation. On the one hand, emotionally negative AVH might lead to a negative mood state, which is reflected in a bias toward negative affects in EPC tasks. This has been shown in the visual modality, in that depressed patients show a bias to perceive neutral emotional facial expressions as negative (Hale, Jansen, Bouhuys, & van den

Hoofdakker, 1998). Thus, it is plausible that patients may live in an emotional negative state because of the intrusive hallucinations they experience, and this negative emotional state might bias judgements regarding the emotional expressions of others, such as in the evaluation of emotional prosody. On this interpretation, the bias in EPC may be based in the distress that results from the experience of emotionally negative AVH. In this causal model, AVH cause distress, and distress may lead to EPC deficits.

That interpretation notwithstanding, we intend to discuss how AVH might be associated with underlying EPC disturbances, with the direction of causation leading from emotion processing deficits to the psychotic symptom. We will consider how EPC deficits may contribute to the formation of AVH and to what extent aberrant auditory processing might be underlying both phenomena. In order to demonstrate this link we will consider behavioral and brain functional and structural findings connecting AVH and EPC, and we will integrate this evidence to suggest a new neuropsychological model of AVH with the aim of explaining the phenomenology of the abnormal experience.

1.3. The link between AVH and EPC at the behavioral level.

Until recently, the idea of an affective prosody impairment as a modular deficit in schizophrenia was controversial. It was assumed that impaired prosodic processing in schizophrenia patients merely reflects basic sensory deficits such as misperception of pitch and amplitude (Leitman et al., 2005). For example, Leitman et al. (Leitman et al., 2005) suggested that prosody processing deficits may be due, in part, to low-level pitch discrimination defects (Leitman et al., 2005). However, there

is evidence to suggest that prosody processing deficits in schizophrenia cannot be solely explained by pitch perception defects. In fact, pitch is a feature of emotional as well as linguistic prosody. If pitch perception impairment were the only cause of prosody processing deficits, linguistic as well as emotional prosody should be equally affected in schizophrenia, but contrary to this assumption, stress prosody comprehension seems to be preserved in patients with schizophrenia (Murphy & Cutting, 1990). In line with this observation, it is plausible that the dysfunction of low-level auditory processing, even if partially contributing to EPC impairment, is not sufficient to explain the complexity of the emotional prosody deficits in schizophrenia and that these may rather relate to a specific emotion processing deficit.

In addition to the specificity of EPC impairment as an emotion processing deficit, EPC difficulties are particularly prominent in the subgroup of patients with schizophrenia who have a tendency towards AVH. In a study comparing patients with and without AVH in an EPC discrimination task (Rossell & Boundy, 2005), it was observed that when EPC performance was tested with non-lexical speech sounds spoken in different tones of voice, only the AVH patient group were impaired relative to controls. When EPC stimuli contained prosodic as well as semantic elements, both patient groups showed significantly worse performance than controls. The authors interpret these findings as suggesting a dissociation between EPC for auditory stimuli with prosodic and semantic content, which further supports the idea of a specific connection between AVH and EPC. The additional processing of semantic content may mask the relation between EPC and AVH. Deficits in semantic processing are found in schizophrenia in general, without being linked to any

symptomatic manifestations of the disorder in particular (Rossell & David, 2006) while emotional prosodic deficits seem to be specific to the AVH subgroup (Rossell & Boundy, 2005).

A further factor mediating the formation of hallucinations is voice identity recognition. Prosodic features such as pitch intensity and pitch duration are important cues in differentiating voices (Belin, Fecteau, & Bedard, 2004). When prosodic features processing is impaired, recognition of voice identity becomes more difficult, and thus a voice may sound like someone else's (Cutting, 1990). In other words, voice identity recognition as well as vocal emotion recognition partly relies on processing of the same paralinguistic acoustical features (Belin et al., 2004), of which the most salient are fundamental frequency, intensity and interval (Perrot, Aversano, & Chollet, 2007). In fact, professional imitators use variations in pitch, tone and rhythm, to impersonate someone else (Orchard & Yarmey, 1995; Perrot et al., 2007). Variations in these same parameters can affect the emotion conveyed by the voice. Even so, it is important to note that some dissociations between voice identity recognition and emotional prosody comprehension have been found (Garrido et al., 2009), suggesting that both processes share a common paralinguistic ground. Following Cutting's approach (Cutting, 1990), Shea and colleagues suggested that vocal emotion may serve as one of the cues to determine the identity of the voice (Shea et al., 2007). Thus, prosody could be one of the contributing factors in identifying both others' as well as one's own voice, such as heard in inner speech. From a cognitive perspective, prosodic features impairment may lead to misattribution of inner speech via incapacity to use the prosodic cues to estimate the origin of the stimuli. If the patients cannot recognise the origin and identity of their

own inner voice, they might attribute the voice to someone else. It is important to note that many voices encountered in AVH have clear identities (Stephane, Thuras, Nasrallah, & Georgopoulos, 2003). Additionally, it has been claimed that misidentification of one's own inner speech relates to abnormal self-monitoring (Johns et al., 2001). Our interpretation explores the possibility that abnormal prosody processing may be a mechanism mediating misidentification of inner speech, which also contains prosodic cues helping the individual to recognize the identity of the speaker (Belin et al., 2004). Following this line of argument, inaccurate processing of the prosodic information encoded in inner speech might contribute to a misidentification of the identity of one's own inner voice.

Inaccurate decoding of prosodic features might also lay a foundation for the formation of certain AVH triggered by external sounds. In fact, a direct relationship has been reported between the timbre and pitch of real environmental sounds and how a degraded perception of them can be mistaken as an AVH (Hunter & Woodruff, 2004). A similar mechanism of degraded auditory perception might not only operate in the case of external environmental sounds but also for voices as features such as timbre are also involved in voice identification. It has been proposed that speaker identification is based on the extraction of paralinguistic features such as frequency and pitch (Formisano, De Martino, Bonte, & Goebel, 2008) (prosodic aspects of speech). EPC deficits, may account for the misidentification of self-generated auditory objects which lead to the formation of AVH. If this conclusion is correct, it might also explain why many AVH contain specific prosodic features and voice identities.

1.4. Neural underpinnings of AVH and their overlap with EPC

Studying gray matter abnormalities in patients with a history of AVH offers additional evidence about the neural mechanisms underlying the abnormal perception phenomenon. A recent study in a large sample of 99 schizophrenia patients found reduced gray matter density bilaterally in the superior temporal gyri which was correlated with hallucination severity (Nenadic, Smesny, Schlosser, Sauer, & Gaser, 2010). Another recent study investigating the link between AVH and gray matter volume reported significant structural differences between AVH patients and controls particularly in the left and right superior temporal gyrus (STG), left inferior frontal gyrus, left amygdala, and insula bilaterally (Garcia-Marti et al., 2008). The authors, basing their interpretation on a previous model of emotional dysfunction in schizophrenia (Aleman & Kahn, 2005), concluded that the amygdala and insula as key regions of the emotional brain may be involved in the emotional load of AVH (Garcia-Marti et al., 2008), particularly in the emotional prosody conveyed by AVH.

One of the pioneering neuroimaging studies applied positron emission tomography (PET) to assess the pattern of regional cerebral blood flow in schizophrenia patients when they were actively hallucinating. Activation was found in left temporal lobe and Broca's area, suggesting that certain language processes, such as speech production, inner speech, and AVH share similar neural mechanisms (McGuire, Shah, & Murray, 1993). A related functional magnetic resonance (fMRI) study assessed patients with schizophrenia while hallucinating and during auditory stimulation (Dierks et al., 1999) and found an activation of the superior and medial temporal gyrus and primary auditory cortex (predominantly in the left hemisphere for two patients and in the right hemisphere for the third patient) for both conditions.

Additionally, during hallucinations, the frontoparietal operculum, the hippocampus, and the amygdala showed an increase of the BOLD signal. The authors suggested that the activation of the primary auditory cortex during the hallucination may explain why hallucinations are perceived as real and often external sounds (Dierks et al., 1999), whereas the amygdala activation was considered to reflect an emotional response relating to the emotional content of the voices. Moreover, the authors interpreted the activation of the left temporal lobe as an indication of semantic processing in hallucinations, even though the relative contributions of semantic and prosodic aspects of language were not compared directly. Unfortunately, the lateralization of semantic (normally left-lateralized) and prosodic processes (normally right-lateralized) was not directly assessed in this study. Thus, these results cannot be related to evidence of a reduced lateralization of linguistic processing in schizophrenia particularly in those patients suffering from severe hallucinations (Sommer, Ramsey, & Kahn, 2001; Weiss et al., 2006).

Unlike the lateralization of linguistic processing, findings about the lateralization of prosody processing in schizophrenia are equivocal. While some evidence found a left lateralization of the normally right lateralized response to prosody (Mitchell, Elliott, Barry, Cruttenden, & Woodruff, 2004), a more recent study found increased right lateralization of emotional prosody in schizophrenic patients (Bach et al., 2009). The findings from the cited studies are limited by small sample sizes. Therefore, a recent meta-analysis merged (Jardri, Pouchet, Pins, & Thomas, 2011) data from ten neuroimaging studies of AVH and found that patients experiencing AVH showed increased activation in a bilateral temporo-frontal

network, including the inferior frontal gyrus and the middle and superior temporal gyrus bilaterally (Jardri et al., 2011).

Only a very few studies have directly assessed the neural link between EPC and AVH. Evidence comes from an fMRI study of patients with schizophrenia, in which neural activation in patients with and without history of AVH was compared while listening to external speech (Woodruff et al., 1997). This study found an increased activation particularly in the right middle temporal gyrus (MTG) for the patient group as a whole, relative to healthy controls during external speech stimulation. The authors suggested that this might indicate an inherent natural hyperresponsivity to emotional prosody of speech in the patient group. In fact, right hemisphere hyperresponsivity has been found in other clinical populations such as violence offenders, who also show deficits in the perception of emotional cues (Lee, Chan, & Raine, 2009). Woodruff at al. also scanned schizophrenic patients during the hallucinating state when they were simultaneously listening to external speech, finding that during AVH the response to external speech in the right MTG was decreased. This finding suggests that AVH compete with external speech for auditory processing resources within the temporal cortex, including those regions which in healthy population are known to be necessary for EPC (Woodruff et al., 1997). However, as it has been previously said, the lateralization of EPC in patients with schizophrenia remains controversial (Bach et al., 2009; Mitchell et al., 2004).

Several studies addressed the overlap of the neural correlates of AVH and perception of prosodic and semantic aspects of speech. For listening to emotional speech, one study (Sanjuan et al., 2007) found enhanced activation of the left middle and right superior temporal lobe, insula, thalamus, and middle and superior frontal

lobes in a group of patients with AVH in comparison to controls. The stimuli used in this study were based on patients' individual reports of the content of their AVH, where words were selected as stimuli based on their frequency in the reported AVH. The selected words contained an emotional semantic meaning and were recorded in an imperative prosodic tone. This study found an increase in brain activity in a fronto-temporal network involving the orbitofrontal cortex, the left medial temporal gyrus and right STG in the patient's group in comparison to controls, and this activation pattern was interpreted as related to AVH (Sanjuan et al., 2007). As this study did not test schizophrenia patients without AVH, however, the results should be interpreted with caution. Differences in the neural networks underlying semantic and prosodic aspects of speech processing between schizophrenia patients and healthy controls may be related to other aspects of the disorder and not specifically to AVH.

Along with the STG, the amygdala is one of the structures that seem to be involved in AVH as well as in EPC, specifically for anger and fear. Amygdala involvement in processing of angry voices may relate to the fact that angry voices "threaten" the listener, which is especially pertinent in schizophrenia given the negative focus of AVH outlined above. Further, it has been demonstrated that, at a resting state, the amygdala shows higher metabolism in patients with schizophrenia than in healthy controls. This might indicate a tonic hyperactivation of the amygdala in schizophrenia, as opposed to a phasic activation in the presence of salient emotional stimuli in healthy controls. Interestingly, hypometabolism of the amygdala during emotion judgment tasks has also been found in schizophrenia (for a review, see Aleman & Kahn, 2005). In fact, a study assessing patients' neural response to

nonverbal emotional auditory stimuli (crying and laughing) found that patients with a history of AVH showed reduced activity in the left amygdala and bilateral hippocampus during crying sounds in comparison with patients without such a history (Kang et al., 2008). A different study found increased amygdala activation in AVH patients during passive listening to emotional words in comparison to NAVH and controls (Escarti et al., 2010). However, this study did not control for occurrence of hallucinations during the scanning, which was a likely eventuality given that AVH patients obtained high scores in a scale measuring hallucinatory experience in the previous 24 hours to the testing session. Thus, the fMRI results might rather have been related to the experience of hallucinations itself than to the experimental stimuli. Possibly, the constant hypermetabolism of the amygdala in schizophrenic patients prevents the detection of differences in BOLD response between baseline and threatening stimuli conditions. Furthermore, tonic pattern of amygdala activation might explain the ubiquitous perception of threats some patients suffer from, resulting in constant (implicit) processing of emotional stimuli leading to lead to feelings of being emotionally overwhelmed that precedes development of AVH.

1.5. Structural and functional connectivity within the EPC and AVH networks

Studies of structural connectivity with diffusion tensor imaging (DTI) have revealed that AVH schizophrenia patients, compared with non-AVH schizophrenia patients, showed an increased connectivity in the temporo-parietal section of the arcuate fasciculus (Hubl et al., 2004; Shergill et al., 2007), which connects the STG and medial temporal gyrus (MTG) with the inferior frontal lobe (Glasser & Rilling,

2008). AVH patients also showed increased structural connectivity in the left cingulated bundle, which is part of the limbic system (Hubl et al., 2004; Shergill et al., 2007). The increased connectivity between language-related areas may lead to an abnormal hyperactivation of the circuit which processes external speech, increasing noise in the speech processing system and lowering the threshold for the intrusion of aberrant auditory perceptual input (Rotarska-Jagiela et al., 2009).

Additionally, a functional connectivity study showed that schizophrenic patients with AVH had poor functional integration between anterior cingulate cortex (ACC) and STG (Mechelli et al., 2007). The authors claim that this network underlies the evaluation of speech source (Mechelli et al., 2007). The aberrant activation of the neural correlates of speech perception without the regulation of the ACC might partially contribute to the formation of hallucinations via deficient source monitoring of inner speech. Moreover, increased structural connectivity within the temporo-frontal network might result in increased spreading of activation in the speech network, introducing noise and increasing excitability of the system (Rotarska-Jagiela et al., 2009), which, in addition to the poor functional integration in this network, might further contribute to abnormal perception. It is important to note that abnormal fluctuations in the BOLD signal can also be detected at resting state in patients with schizophrenia (Garrity et al., 2007), and it has been proposed that there is spontaneous activity in auditory sensory areas which might form the basis of auditory perception in the absence of external stimuli (Hunter et al., 2006).

In conclusion, AVH and EPC in schizophrenia seem to be based on overlapping neural networks, in which the STG seems to play a pivotal role. Interestingly, targeting the STG with TMS has been shown to reduce the severity of

AVH (Aleman, Sommer, & Kahn, 2007), suggesting a causal involvement of this structure in hallucinations. Additionally, a combined fMRI and TMS study found that even when the BOLD response during hallucinations appear to be distributed in a network comprising the left and right inferior frontal gyri and the left superior temporal gyrus, only TMS applied on the left superior temporal gyrus ameliorated the AVH (Hoffman et al., 2007). The authors concluded that this area plays a critical role in the genesis of AVH (Hoffman, et al., 2007). Additional evidence demonstrated that patients with a history of AVH have reduced gray matter density in the STG (Garcia-Marti et al., 2008; Nenadic et al., 2010) suggesting that this structure is constitutively involved in the generation of hallucinations. Additionally, there is evidence of abnormal structural and functional connectivity of STG with frontal areas in AVH patients which might come on top of the gray matter abnormalities of this area. It is important to note that numerous findings have associated this structure not just with AVH and prosodic processing but also with a variety of cognitive domains, including audiovisual integration, theory of mind and speech processing. Within the variety of neural networks the STG is temporally coupled to (Hein & Knight, 2008), for our model we will focus on the cases in which it interacts with the EPC and AVH networks.

1.6 Towards a new neuropsychological model of AVH

Some AVH seem to be triggered by external sounds and there is some supplementary evidence indicating that inaccurate perception of prosodic features of external sounds, such as timbre, amplitude, and timing, might contribute to the formation of AVH. Along these lines, a case study has shown that false perceptions may derive from misinterpretation of real sounds as exemplify in the case of a

patient who heard a real engine sound and parallelly perceived a voice with "mechanical timbre" (the voice sounded "like a machine") (Hunter & Woodruff, 2004). In this case, acoustical features of the AVH were already present in an original environmental sound (Hunter & Woodruff, 2004) and features corresponding to the timbre of sounds seem to be misprocessed in a way that triggers false perceptions. We suggest that this primary sensory processing impairment plays a necessary role in false perception, but it is not sufficient to form hallucinations *per se*.

While some AVH appear to be triggered by an external sound, other AVH appear to be generated intrinsically and in the absence of external stimuli. At the neurophysiological level, it has been shown at resting state there is spontaneous activity in the primary auditory cortex, superior temporal gyrus (STG), and anterior cingulate cortex (ACC) in healthy populations (Hunter et al., 2006). Such spontaneous activity establishes the possibility of this circuitry being involved in perception when external stimuli are absent. It has been proposed that disruption in the physiological modulation of this network at resting state, or some defect in the intrinsic functional connectivity of this network, could contribute to an enhanced sensitivity of the primary auditory cortex, which might lead to an increased sensitivity to irrelevant (auditory) signals, and thus to hallucinatory processes (Hunter et al., 2006). Considering that the primary auditory cortex has been shown to be prone to spontaneous activity, the existence of external stimulation might not be necessary for AVH and thus AVH might additionally be purely intrinsically generated. Thus, we hypothesize that such spontaneous fluctuation in the network might trigger the perception of AVH in absence of external stimuli in some cases,

while in other cases it might lowers the excitability threshold facilitating neural firing in the presence of external stimuli.

Regardless whether auditory stimuli are internally or externally generated, they are integrated into a common percept after sensory evaluation. Evidence suggests that the STG is the neural substrate of this integration process in the healthy population (Leitman et al., 2008) and it has been shown that STG, a pivotal structure underlying EPC (Hoekert, Bais, Kahn, & Aleman, 2008; Mitchell et al., 2004) , is also active during AVH (Stephane, Barton, & Boutros, 2001; Woodruff et al., 1997). The posterior section of STG in the left hemisphere is considered to be the location of Wernicke's area, the neural substrate of speech comprehension. Indeed, it may be the multifarious associations of the STG with numerous brain regions that explain why this structure is the focus of overlap between EPC and AVH. Functional connectivity analyses and DTI techniques have shown that STG is connected with auditory primary cortex, medial ventral frontal cortex, prefrontal cortex, ACC, and amygdala (Hunter et al., 2006; Mechelli et al., 2007).

Furthermore, the STG also forms part of an anatomically connected triadic network which also includes the amygdala and orbitofrontal cortices (Ghashghaei & Barbas, 2002). The amygdala may contribute to EPC specifically for the valences of fear and anger, and these two emotional valences are probably the most frequent affective tone of AVH. The amygdala receives inputs from the prefrontal cortex which regulates its activity via afferents suppressing amygdala output (Rosenkranz & Grace, 2001), but this amygdala regulation by prefrontal cortex is disrupted in schizophrenia (Leitman et al., 2008). Therefore, in schizophrenia, prefrontal circuitry may fail to down-regulate amygdala activity, which might therefore drive attention

towards certain features of prosody, particularly negative tones, coincidently with the emotional valence of most of the AVH that patients suffer from. The amygdala thus lacks prefrontal inhibitory control and becomes hyperresponsive, i.e., it fails to suppress this stimulus-driven (bottom-up) process. In combination, the inefficient top-down control of the amygdala, and dysfunction of bottom-up processes (sensorial misperception) may account for the mechanism underlying the relationship between AVH and EPC in schizophrenia. Amygdala disinhibition increases emotional reactivity (Siever, 2008) and consequently biases attention towards threatening stimuli, which ultimately causes abnormal emotion processing.

Finally, we suggest that a final element of the network mediating the contribution of EPC to AVH is the abnormal connectivity between ACC and STG. Evidence suggests that this abnormal connectivity causes difficulty judging whether a stimulus was internally or externally generated (Erkwoh et al., 2006), thus leading to misattribution of inner auditory objects such as inner speech.

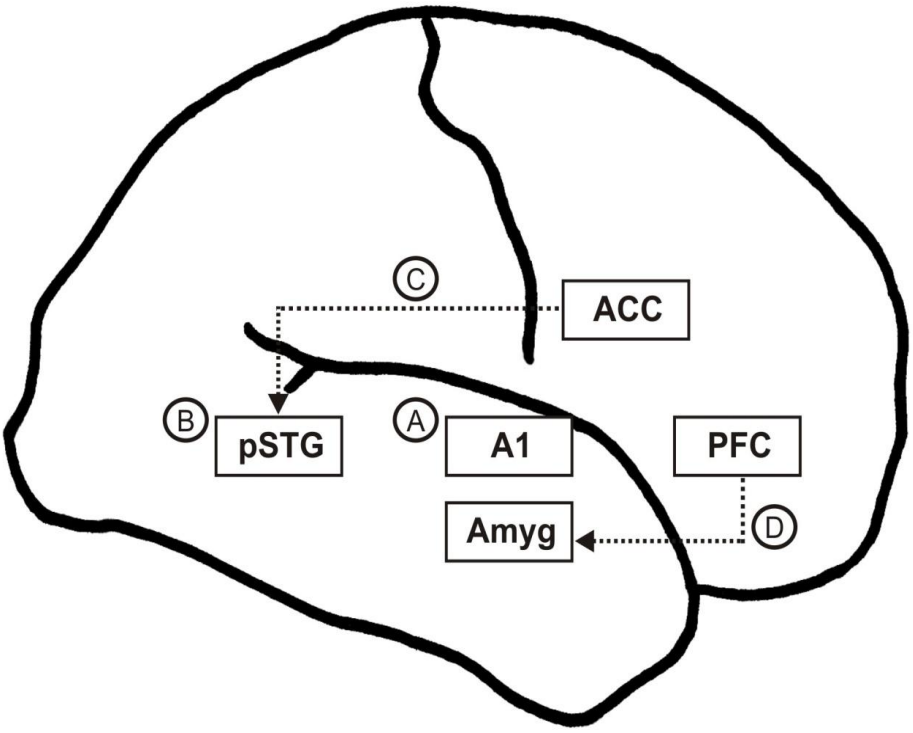

Figure 1. Hypothetical model of the neural formation of AVH in schizophrenia. Precursors of AVH originate at the primary auditory cortex (A1). Spontaneously generated intrinsic fluctuations in the activation of A1 can result in a misperception of external low level sound features and a perceptual experience of sounds even in the absence of external stimuli (A). A1 projects to the superior temporal gyrus. The posterior part of the superior temporal gyrus (pSTG) integrates prosodic features of external and intrinsically generated sounds. Structural and functional abnormalities in the pSTG in schizophrenia result in a degraded and aberrant integration of prosodic features (B). Additionally, a disruption of the functional connectivity between pSTG and anterior cingulate cortex (ACC) can lead to a misidentification of the source of both external and self-generated auditory objects (C). Finally, the amygdala is hyperactive due to aberrant top-down regulation by the prefrontal cortex (PFC), which results in increased emotionality and attention bias towards threatening auditory stimuli (D). These processes might independently or jointly contribute to the formation of AVH.

1.7 Conclusion:

The neuropsychological model proposed here accounts for the diverse phenomenology of AVH, considering that different aberrant processes may result in similar abnormal perceptions. We contribute to previous literature by integrating

phenomenological and behavioural aspects of AVH in an empirically verifiable neural model. Regardless of whether AVH are triggered by the misperception of external sounds, by spontaneously generated intrinsic fluctuations in the auditory system, or by a combination of both, hallucinations in schizophrenia may occur at the stage of transition from primary sensory processing regions to integrative regions such as STG. Deficits in perception and integration of stimulus features may lead to EPC deficits which contribute to the formation of AVH, for example, via voice identity misidentification. Given that the right STG is involved in voice identity recognition and EPC, which are in turn associated with AVH, future empirical research should focus on how the STG is functionally and structurally altered in AVH patients and how these alterations are linked to deficits that patients might present at the behavioral level (such as prosody decoding and voice identity recognition deficits). A second strand of further research should use neuroimaging data to investigate dynamical interactions between the auditory cortex and associative areas during hallucinations in order to demonstrate the validity of the model empirically. A comprehensive account of the neural process by which hallucinations are formed could lead to the development of clinical interventions that interrupt the formation of AVH, for example by applying TMS approaches (Aleman et al., 2007). Such interventions targeting key brain regions for AVH (i.e., rSTG and primary auditory cortex) with mechanism-based methods could diminish AVH precursor processes in patients resistant to the medication and thus reduce the human cost of the symptom.

PART I: Studies in non-clinical population

Chapter II: The neural correlates of emotional prosody comprehension: Disentangling simple from complex emotion.

The previous chapter described a model of the formation of auditory verbal hallucinations (AVH). Evidence from the literature highlighted deficits in emotional prosody decoding (EPC) as a main contributing factor for the formation of AVH. Moreover, the model described the neural overlapping between AVH and EPC. However, findings regarding the neural representation of EPC are often contradictory. While evidence in the literature consistently indicates an involvement of the superior temporal gyrus in EPC tasks, there is still controversy regarding the lateralization of EPC in the posterior temporal lobes. Moreover, the role of the prefrontal cortex in EPC is still a matter of debate. While the prefrontal cortex appears to be associated with EPC in some fMRI studies, others do not find this region involved in EPC tasks.

In order to test the AVH model, it is necessary to clarify the neural representation of the main component (EPC). For that reason an fMRI study of EPC is presented in the current chapter. Factors modulating the prefrontal cortex involvement in EPC are addressed.

2.1. Abstract

Emotional prosody comprehension (EPC), the ability to interpret another person's feelings by listening to their tone of voice, is crucial for effective social communication. Previous studies assessing the neural correlates of EPC have found inconsistent results, particularly regarding to the involvement of the medial prefrontal cortex (mPFC). It remained unclear whether the involvement of the mPFC is linked to an increased demand related to the socio-cognitive components of EPC such as mental state attribution and if basic perceptual processing of EPC can be performed without the contribution of this region. fMRI was used to delineate neural activity during the perception of prosodic stimuli conveying simple and complex emotion. Emotional trials in general, as compared to neutral ones, activated a network comprising temporal and lateral frontal brain regions, while complex emotion trials specifically showed an additional involvement of the mPFC, premotor cortex, frontal operculum and left insula. These results indicate that the mPFC and premotor areas are not crucial to EPC per se; however, the mPFC supports socio-cognitive skills necessary to interpret complex emotion such as inferring mental states. Additionally, the premotor cortex involvement may reflect the participation of the mirror neuron system for prosody processing of complex emotion.

2.2 Introduction:

Human everyday verbal communication involves not only semantic but also non-linguistic, information being carried by the voice (Belin et al., 2004). This phenomenon, known as prosody, comprises acoustic features such as pitch, amplitude, segment and pause duration and allows for the encoding and decoding of emotions in speech (Belin et al., 2004), a skill which is necessary to ensure effective social communication (Beatty, Orbelo, Sorocco, & Ross, 2003).

We will denote the act of decoding emotion cues conveyed by prosody as emotional prosody comprehension (EPC). EPC does not represent a single construct. There are qualitative differences between simple emotions and more complex emotional states. Therefore, EPC is a multi-level mechanism, from the decoding of simple emotions such as fear, happiness or anger to the assessment of complex mental states. Furthermore, EPC is regarded as one of the precursor of emotional theory of mind (Chakrabarti & Baron-Cohen, 2006). Evolutionary, simple emotions evolved for "their adaptive value in dealing with fundamental life tasks" (Ekman, 1992). They are shared with other primates, include a distinctive, universal physiological response (Ekman, 1992) and are characterized by automated and complex changes involving facial and vocal expressions (Zinck, 2008). They only last for a limited period of time, are highly stereotypical and involve very limited cognitive processing (Zinck, 2008). In contrast, complex emotions, and especially social emotions such as pride, guilt and embarrassment, require the interpretation of social intentions (van Hooren et al., 2008), consideration of other people, comprehension of social norms and recognition of personal responsibility for the

consequences of a situation (Bauminger, Edelsztein, & Morash, 2005). They require the monitoring of attitudes and opinions of others regarding our own behaviour, are culturally dependent, and rely upon the evaluation of others (Capps, Yirmiya, & Sigman, 1992). Non-social complex emotions, such as thoughtfulness, boredom or interest are belief-based rather than situation-based and reflect the inner thoughts of an individual (Shaw et al., 2005). An important difference between complex emotion and simple emotion is based on the fact that complex emotions involve adjudicating a cognitive state as well as an emotion and are context and culture dependent (Golan, Baron-Cohen, Hill, & Golan, 2006; Tamietto, Latini Corazzini, de Gelder, & Geminiani, 2006). The cognitive content is an essential constituent of the emotion and it is a relevant part of what causes the emotion (Zinck & Newen, 2008). Thus, complex emotions are a cognitively enriched extension of simple emotion (Zinck & Newen, 2008) and additional cognitive elaboration is necessary to process complex mental states (Johnson-Laird, 1989).

At the behavioural level, studies have sought to determine whether emotion comprehension for simple emotion and complex emotion is differentially affected by neurological impairments and childhood development, which might imply separate neural processes. However, findings have been equivocal. One study in children with learning disabilities revealed difficulties in understanding complex social emotions such as pride or guilt together with a preserved ability to recognize simple emotions such as happiness or sadness (Bauminger et al., 2005), suggesting that both emotional processes might be neuronally dissociated. In agreement with these finding, another study using facial stimuli found a similar dissociation between the comprehension of simple and complex emotion in patients with schizophrenia

(Kington, Jones, Watt, Hopkin, & Williams, 2000). On the other hand, a recent study looking at the detection of sarcasm (a complex emotion) and simple emotion from vocal cues found that performance in both tasks was highly correlated in both a control and a schizophrenic patients group (Leitman, Ziwich, Pasternak, & Javitt, 2006).

Looking at functional brain imaging data, some clinical studies have indicated overlapping brain areas involved with simple and complex emotion comprehension deficits (Adolphs, Baron-Cohen, & Tranel, 2002; Adolphs, Damasio, & Tranel, 2002). To the best of our knowledge, however, there have not yet been any studies in healthy participants which have directly compared the brain networks involved with simple and complex emotion comprehension from speech cues. Such studies might be especially interesting, as simple and complex emotion comprehension might be mediated by different brain areas, even if they appear correlated on the behavioural level. Previous neuroimaging studies have shown that, EPC in general is supported by a temporo-frontal network (Ethofer et al., 2006; Leitman et al., 2010; Schirmer, Zysset, Kotz, & Yves von Cramon, 2004; Wildgruber et al., 2005). However, the role of each of the neural components in the network, particularly the involvement of prefrontal lobes, is still under debate.

Some authors have claimed that the involvement of prefrontal regions in EPC depends on linguistic features of the stimuli. One study (Mitchell, Elliott, Barry, Cruttenden, & Woodruff, 2003) found that concurrent semantic content of prosodic cues resulted in increased activation of the inferior frontal gyrus, while activation of the posterior lateral temporal lobe during prosody decoding remained constant independently of the semantic load of the stimuli (Mitchell et al., 2003).

Interestingly, it has been suggested that increased (semantic) processing demands may therefore have little effect on the auditory cortex response, but may modulate the frontal lobe response (Mitchell, 2006). Conversely, it has been proposed that taking away the labelling element of typical EPC tasks (such as classifying the stimuli into a category represented with a word like "happy") and asking participants to discriminate EPC instead (make same/different judgements about the emotion conveyed in pairs of sentences) reduces frontal lobe activity (Mitchell, 2006), indicating that the demand on frontal lobe resources is reduced when EPC are reduced to purely perceptual judgements.

An alternative model of emotion perception proposes that in order to decode other person's emotions, postural, facial or vocal cues are observed, which activate engrams to simulate a similar emotion (Bastiaansen, Thioux, & Keysers, 2009). Such an internal simulation facilitates the sensation of the emotional state in an embodied way, which then is interpreted and attributed to other individuals. If this model is correct, the recruitment of a mirror neuron system for the perception of emotions would be necessary. In fact, the role of a mirror system for emotion decoding from facial emotion (Jabbi, Swart, & Keysers, 2007; Seitz et al., 2008) as well as prosody (Aziz-Zadeh, Sheng, & Gheytanchi, 2010; Ramachandra 2009; Warren et al., 2006) have been proposed. Some of these studies suggested that the engagement of the mirror system depends on the empathic characteristic of the participants (Aziz-Zadeh et al., 2010). For example, in a prosody decoding task, activations in the bilateral superior, middle and inferior frontal gyri, as well as the anterior insula and bilateral perisylvian activation inversely correlated with empathic ability (Aziz-Zadeh et al., 2010). The extensive activation including the bilateral superior, middle and inferior

frontal gyri may relate to inner simulation of the emotional state of others (Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003) which might be particularly crucial for more demanding emotions in which the inference of intentions is required (Mason & Just, 2011). Thus, it is plausible to predict that premotor activation would be more prominent for complex social emotions in comparison to simple.

The mirror neuron system is proposed to help people to understand the mental states of other on the basis of our own mental state, which is the first step for theory of mind (ToM) (Frith & Frith, 2006). It has been proposed that making inferences about social interactions (a task which requires ToM) relies upon the integrity of the orbitofrontal as well as the medial prefrontal cortex (Frith & Frith, 2006; Mah, Arnold, & Grafman, 2004). ToM skills may be particularly needed in EPC for complex and social emotion because they imply to adjudicate inner thoughts to the individual experiencing the mental state in the case of the non-social complex emotion, whereas social emotion require the interpretation of social cues, taking the dyadic relation in which the emotion emerges. If this is true, the neural network underlying ToM should also underlie EPC for complex and social emotions.

The present study examines the neural correlates of EPC of simple and complex emotion from vocal cues. This investigation proposes that EPC for both simple and complex emotion share common neural components, but additional socio-cognitive modules are recruited for complex emotion. It is hypothesised that the neural correlates of the complex emotion comprehension differ from those of simple emotion due to the requirement of taking the emotional perspective of others [33] which might partly rely on mental state decoding skills (Lee, Harkness, Sabbagh, & Jacobson, 2005). Specifically, we predict that EPC for complex emotion

involves activation of the orbitofrontal and medial PFC as part of the social brain (Blakemore, 2008) as well as the premotor cortex as a part of the mirror neuron system, indicating that the involvement of the PFC in EPC depend on the complexity of social judgments involved in the task.

2.3. Methods

Participants

A group of twenty male students and academics staff were recruited from the Department of Psychology at Durham University. Only male participants were recruited because women present larger variability in functional brain organization, partly due to hormonal fluctuations across the menstrual cycle (Weis & Hausmann, 2010; Weis, Hausmann, Stoffers, & Sturm, 2010). Also, emotional state is affected hormonally, for example, across the menstrual cycle (Rubinow & Schmidt, 2006), causing unwanted additional variability. Students received course credits for taking part in the study. One participant had to be excluded from the analysis due to artefacts caused by an orthodontic brace. The mean age of the remaining 19 male participants was 24.8 years (*SD* = 8.79 years, age range: 18 to 51 years). All participants were native English speakers and reported not to have any history of psychiatric disorders, hearing impairment, history of drug or alcohol abuse, long periods of unconsciousness or head injuries. Ethical approval for the study was obtained from the Ethics Sub-Committee of the Psychology Department at Durham University.

Stimuli and task

The EPC stimuli used in the current study were selected from Banse and Scherer's study of vocal emotion expression (Banse & Scherer, 1996) and comprised numbers spoken in different tones of voice. Stimuli were created by the Linguistic Data Consortium (LDC), an open consortium of universities, companies and government research laboratories, hosted by the University of Pennsylvania. The utterances were recorded by six professional actors (three male, three female) recorded on two channels, with a sampling rate of 22.05K, and two microphones, a stand-mounted boom Shure SN94 and a headset Sennheiser HMD 410. Sound files were encoded in interleaved 16-bit PCM, high-byte-first format. Further details about the EPC stimuli can be found at http://www.ldc.upenn.edu.

From the full set of numbers, stimuli spoken in three simple emotional tones (happy, sad, angry) and three complex/social tones (proud, guilty, bored) were selected. Simple and complex emotion stimuli were presented in two separate functional MRI runs, administered in a counterbalanced order across participants. For both conditions, also numbers spoken in a neutral tone of voice and silent trials (baseline condition) were employed, resulting in a total of 160 utterances (32 per emotional category, 32 neutral stimuli and 32 silent trials per run). The mean duration of each sound stimulus was 2010 ms (range: 1800 to 2080 ms). Stimuli were presented using E-Prime (Psychology Software Tools, Philadelphia) via Phillips digital stereo headphones. The order of trials was selected based on statistical efficiency computations by an fMRI simulator software taking into account the shape and timing of the canonical hemodynamic response function employed for modeling the time course of the BOLD response in SPM. More details

about this software can be found at http://www.cabiatl.com/CABI/resources/fmrisim/.

In both runs, participants were asked to identify the emotion conveyed by the tone of voice for each trial and to indicate their forced-choice response by pressing one of four keys (one for each emotional valence) of a five key response box. A picture of the response box indicating which key corresponded to which response was continuously presented on the screen. Participants were asked to respond as fast and as accurate as possible.

Image Acquisition

Functional MRI images were acquired with a Phillips Achieva 3T scanner with a SENSE standard 8-channel birdcage head coil. The functional gradient-echo echoplanar T2*-weighted images (EPI) were acquired with an echo time (TE) of 30ms, a flip angle of 90°, a field of view (FOV) of 192 mm and an in plane resolution of 64 x 64 voxels. Each functional image consisted of 28 axial slices (4 mm thickness with 0.5 gap), which covered the whole cerebral cortex. In order for participants to be able to hear the auditory stimuli during the functional runs, we used a sparse imaging procedure with a repetition time (TR) of 8 s, including an effective acquisition time (TA) of 2 s interleaved with a silent gap of 6 s. Auditory stimuli were delivered binaurally via MRI-compatible headphones, and were presented with a varying jitter of 2.5 to 3.5 seconds relative to scan onset. For each participant, a high resolution T1-weighted anatomical scan was acquired using a TR of 9.6 seconds, TE of 4.6, FA of 8°, FOV 256 mm x 256 mm x 150 mm with 150 slices of 1.0 mm thickness.

Image Processing

Functional images were preprocessed and analyzed with Statistical Parametric Mapping (SPM8; Wellcome Department of Imaging Neuroscience, London, UK, www.fil.ion.ucl.ac.uk) software implemented in MATLAB 7.8.0 (Mathworks Inc., Sherborn, MA). The first four images of each run were discarded to ensure signal stabilization. Images were realigned applying a rigid body spatial transformation of each of the BOLD volumes onto the fifth volume of the first run in order to remove movement artifacts. Functional images were co-registered with the anatomical scan and were stereotactically normalized into Montreal Neurological Institute (MNI) space on the basis of the structural T1-weighted 3D volume. Then, functional images were re-sliced at resolution of 3 x 3 x 3 mm and smoothed by a Gaussian filter of 8 x 8 x 8 FWHM.

Analysis

A statistical analysis on the basis of the general lineal model was performed using SPM8. In an event-related design, for each of the different emotional tones of voices as well as for the neutral stimuli, the expected hemodynamic response was modelled by the canonical hemodynamic response function (HRF; Friston et al., 1998) and its temporal derivative, as implemented in SPM8, with the silent trials serving as a baseline. Subsequently, parameter estimates of the HRF regressor for each of the different conditions were calculated from the least mean squares fit of the model to the time series. Parameters estimates for the temporal derivative were not further considered in any contrast.

The resulting contrast images were subjected to one sample t-tests subsequently explored at a threshold of $p<0.005$. Correction for multiple comparisons to $p < 0.05$ was achieved using a cluster extent threshold procedure first described by Slotnick et al. (Slotnick et al., 2003; Slotnick and Schacter, 2004). As reported in a previous study (Slotnick & Schacter, 2004), the cluster extent threshold procedure relies on the fact that given spurious activity or noise (voxel-wise type-I error), the probability of observing increasingly large (spatially contiguous) clusters of activity systematically decreases (Slotnick & Schacter, 2004). Therefore, the cluster extent threshold can be enforced to ensure an acceptable level of corrected cluster-wise Type I error. For an individual voxel Type I error of $p < 0.005$, this procedure identified a cluster extent of 18 contiguous resampled voxels as necessary to correct for multiple voxel comparisons across the whole brain at $P < 0.05$.

In a further analysis, a new model was set up adding pitch as a parametric regressor to each of the conditions in the general linear model to exclude confounding effects of this variable.

2.4. Results

Behavioural data

During both tasks all four response categories were discriminated significantly above chance level of 25% (simple emotion task [Mean accuracy ±simple emotion]: Happy: 73.92 ± 0.01; Angry 77.25 ± 0.01%; Sad: 71.21 ± 0.01%; Neutral 70.90 ± 0.01%. complex emotion task: Proud 53.45 ± 0.01%; Guilty 55.23 ± 0.01%; Bored 59.37 ± 0.01% and Neutral 57.03 ± 0.01%) (all $t > 21.3$, $p < 0.0001$). Then, data were collapsed across all simple emotions for the SE task and all complex

emotions for the CE task. The reason for collapsing the data lies in the fact that this study was designed to investigate the average neural correlates of simple and complex emotion, independently of the specific emotional valences. A paired sample t-test showed that participants performed significantly more accurate ($t(18)=14.88$ $p< 0.001$) in the simple emotion (M = 73.93, SE = 0.01) run in comparison to the complex emotion run (M =56.22, SE = 0.01)

To investigate whether stimuli of the simple and complex emotion conditions differ not only in the emotional complexity but also in low level acoustical features, simple and complex emotion conditions were also compared according to sound amplitude, duration and pitch mean. These acoustical features were extracted from the stimuli sound files using Praat software for the analysis of speech in phonetics available at http://www.fon.hum.uva.nl/praat/. Paired t-tests revealed no differences between conditions in amplitude, $t(70)=1.06$ $p= 0.295$, and duration, $t(70)=0.83$ $p= 0.412$. However, there was a significant difference in pitch between simple and complex emotion. The analysis revealed that simple emotion stimuli (M = 246.22, SD = 96.36) have an average higher pitch than complex emotion stimuli (M = 166.93, SD = 59.10, $t(70)=6.16$ $p< 0.001$).

Functional imaging data

Emotional versus neutral trials

When neutral tone trials were compared to emotional trials across both runs (see table 1 and figure 2), stronger activations for emotional as opposed to neutral trials were observed within the temporal lobe, in the middle and superior temporal gyri bilaterally, extending into the left temporal pole and the right insula. Within the

frontal lobe, increased BOLD response was found in the inferior frontal operculum bilaterally and in the left pars triangularis. Additionally, the left precentral gyrus was activated. Further significantly activated clusters were observed in the right inferior parietal gyrus and right precuneus, the left putamen and the right cerebellum.

Table 1: Subtraction of neutral from emotional trials. Local maxima of the brain regions activated more for simple and complex emotion as opposed to neutral trials at p < 0.05 corrected for multiple comparisons. Coordinates refer to the MNI system.

	structure	cluster size	Z score	*x*	*y*	*z*	BA peak
temporal lobe	R middle temporal	577	4.48	51	-37	1	21
	L superior temporal	222	4.14	-60	-22	4	22
	L superior temporal pole	40	3.49	-54	11	-14	38
	R insula	66	4.02	33	11	-8	48/38
frontal lobe	L precentral	633	4.82	-36	-28	61	4
	L inferior tri frontal	27	3.57	-36	32	1	47
	L inferior frontal operculum	119	3.46	-36	17	22	48
	R inferior frontal operculum	28	3.13	51	17	19	44
parietal lobe	R inferior parietal	33	3.38	33	-52	49	40
	R precuneus	30	3.27	12	-73	43	7
subcortical structures	L putamen	101	4.09	-24	17	-5	
	R cerebellum	30	3.29	33	-73	-22	

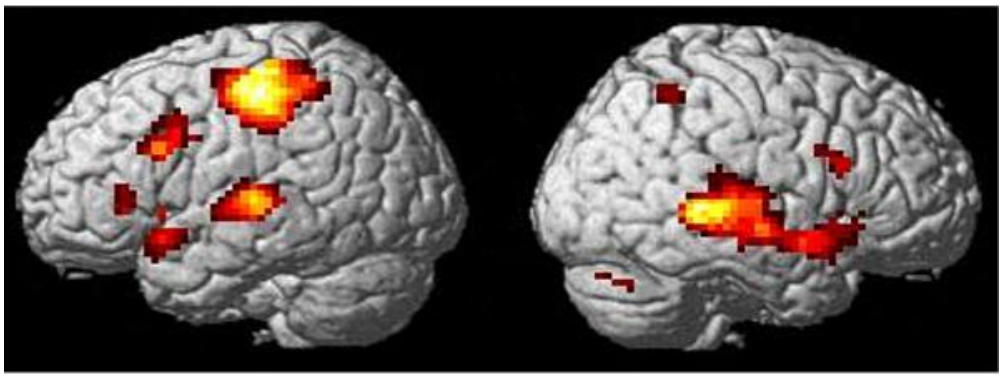

Figure 2. Regions involved in prosody for simple and complex emotion. Brain regions showing significantly stronger activations for simple and complex emotion as opposed to neutral trails. Activations are shown for p < 0.05, corrected for multiple comparisons.

Complex versus simple emotions

When brain activation during perception of simple emotion trials was subtracted from activation during perception of complex emotion trails, there was an increased BOLD response within the frontal lobe, where significantly activated clusters were located in the middle orbito-frontal cortex, right frontal operculum, left supplementary motor area and in the superior medial frontal gyrus (BA 9/32) (see table 2 and Fig. 3). Within the temporal lobes, an increase of activations was found in the right inferior temporal gyrus, the left superior temporal and left fusiform gyrus, the left insula and the right hippocampus. Further significant activations were observed bilaterally in somatosensory association cortex of the parietal lobes, the left thalamus and the right cerebellum.

The reverse contrast did not show any significantly increased activations for simple as compared to complex emotion

Table 2: Subtraction of simple from complex emotion trials. Local maxima of the brain regions activated more for complex as opposed to simple emotion at p < 0.05 corrected for multiple comparisons. Coordinates refer to the MNI system.

	Structure	**cluster size**	**Z score**	***x***	***y***	***z***	**BA peak**
temporal lobe	L insula	129	3.78	-30	20	-11	38
	R parahippocampal cortex	42	3.73	27	2	-26	36
	R inferior temporal	26	3.43	51	-34	-17	20
	L middle temporal	77	3.3	-63	-31	-8	21
	L fusiform	163	3.29	-30	-43	-23	37
frontal lobe	L middle frontal	140	4.02	-42	20	40	44
	R frontal operculum	38	3.34	48	17	40	44
	L superior medial frontal	299	3.32	-6	41	46	9/32
	R supplementary motor	73	3.24	15	5	58	6
	L middle orbito-frontal	33	3.23	-36	50	-5	47/10
	L precentral	19	3.17	-36	-1	61	6
parietal lobe	L superior parietal	44	3.46	-18	-61	67	7
	R superior parietal	25	3.3	21	-55	64	5
subcortical structures	L thalamus	450	4.11	-6	-28	-17	
	cerebellum	320	3.62	12	-79	-32	

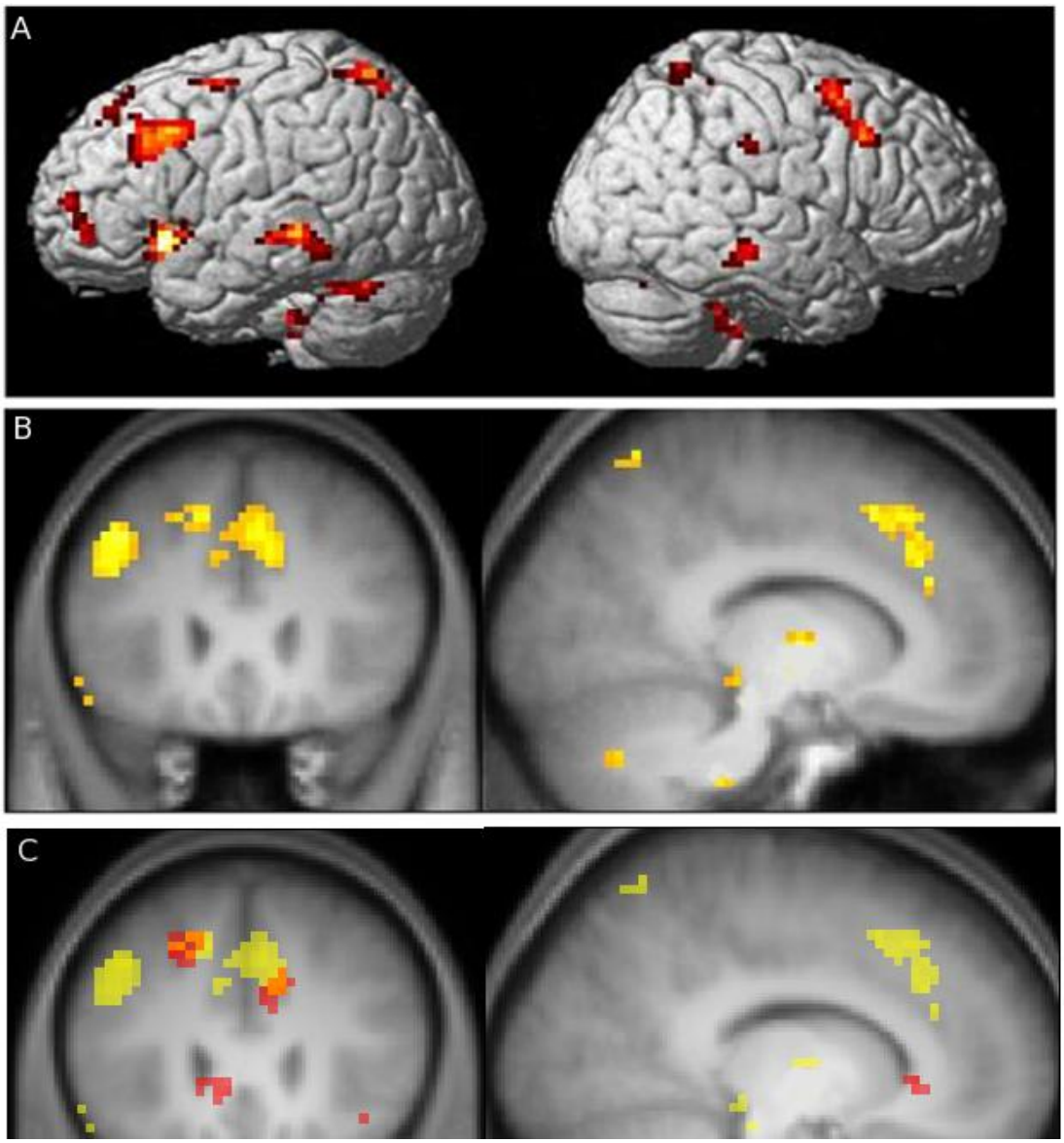

Figure 3. Regions involved in prosody for complex emotion. Brain regions showing significantly stronger activations for complex as opposed to simple emotion. The activation maps (at $p < 0.05$, corrected for multiple comparisons) are shown overlaid onto a canonical brain rendered in three dimensions (A). The anatomical location of the medial frontal activation (at $p < 0.05$, corrected for multiple comparisons) is shown overlaid onto the mean high-resolution T1 scan of the group (B). In (C) activations for the standard analysis are shown in yellow and activations corrected for confounding effects of pitch are shown in red with the overlap of the standard analysis and the analysis corrected for pitch shown in orange.

Complex versus simple emotions controlled for pitch

Stimuli of the simple and complex emotion condition did not only differ in the emotional complexity but also in pitch, a basic acoustical feature. Therefore, we added pitch to the model as a parametric regressor of no interest. Even when controlling for possible confounding effects of pitch, the key network previously described underlying complex emotion is still active. In comparison to the previous analysis, this analysis revealed activations in the right and left superior frontal gyrus (BA 9/32) extending towards medial regions, the right parahippocampal gyrus, the left thalamus and the cerebellum.

Table 3: Subtraction of simple from complex emotion trials controlling for pitch. Local maxima of the brain regions activated more for complex as compared to simple emotion controlled for confounding effects of pitch at $p < 0.05$ corrected for multiple comparisons. Coordinates refer to the MNI system.

Structure	cluster size	T score	*x*	*y*	*z*	BA peak
L Thalamus	23	5.81	-9	-7	1	
L superior frontal gyrus	32	4.22	-15	26	43	9\32
L cerebellum	55	3.94	-30	-49	-29	44
Mid. Cingulum/Supp medial Frontal	20	3.79	6	26	31	32
R parahippocampal	19	3.49	24	-22	-23	30/36

2.5 Discussion

This study was conducted to reveal differences between the neural correlates of EPC for complex as opposed to simple emotion. Disentangling the brain representation for these different types of emotion should contribute to the prevailing controversies regarding the involvement of frontal brain regions in EPC. Moreover, controlling for possible confounding effects of pitch, demonstrated for the first time

that the frontal network underlying EPC for complex emotion is relatively independent of this confound.

Perception of emotional versus neutral trials

In agreement with previous fMRI findings (Kotz et al., 2003; Leitman et al., 2010; Wildgruber, Ackermann, Kreifelts, & Ethofer, 2006), our data showed that EPC in general (pooling across simple and complex emotion relative to neutral trials), is supported by a temporo-frontal network, comprising the middle and superior temporal gyri, left temporal pole, right insula, Broca's area and its right hemisphere homologue, as well as the left motor cortex. Within this network, it is especially the right lateral temporal lobe and particularly the right superior temporal gyrus (rSTG) that has been shown to be crucial for prosody decoding (Friederici & Alter, 2004; Ross & Monnot, 2008a). The additional involvement of left lateral temporal regions in the EPC task might not be related to EPC per se, but rather to explicit verbal labelling of emotional valences (Mitchell & Ross, 2008).

After sensory evaluation of prosodic features in the STG, the output presumably is transferred towards more anterior regions for further processing, as it has been proposed by an analysis of effective connectivity (Ethofer et al., 2006). Along the pathway towards anterior regions, an involvement of the left temporal pole in EPC was found. This paralimbic structure has been considered responsible for coupling visceral emotional responses with representation of complex auditory stimuli (Olson, Plotzker, & Ezzyat, 2007). Another border structure between the temporal and frontal lobes in which EPC is processed is the insula. The results show

right insula activation during EPC, interpreted as related to amalgamating interoception of body states and emotion (Sander & Scheich, 2005).

Activations in Broca's area and its homotopic region during EPC are in line with several previous neuroimaging studies (Aziz-Zadeh et al., 2010; Bach et al., 2008; Wildgruber et al., 2005). An involvement of the right IFG has been suggested for explicit evaluative judgements of emotional prosody, whereas the activation of the left IFG may reflect integration of vocal and verbal information (Schirmer & Kotz, 2006). Further activation of the frontal lobe was found in the left precentral gyrus. This activity should not be related to the motor response (button pressing) because such activation should have been identical for emotional and neutral trials. An alternative interpretation of the motor activation triggered by emotional stimuli relates to a preparation of motor responses to perceived emotion, such as the mimic of a communicative gesture to respond to the perceived emotion (Warren et al., 2006).

Finally, the activation of the right inferior parietal gyrus and precuneus is consistent with previous findings showing a role of this region in explicit emotional stimuli, as compared to phonetic/semantic stimuli (Bach, Grandjean, et al., 2008; Mitchell et al., 2003). This activation has been interpreted as higher order analyses of auditory signals (Bach, Grandjean, et al., 2008) in polymodal areas of the parietal cortex.

Perception of complex versus simple emotions

Although contradictions exist (Ethofer et al., 2006; Ischebeck, Friederici, & Alter, 2008), some previous studies found the mPFC to be involved in emotional

prosody (Beaucousin et al., 2007; Sander et al., 2005). Our key finding revealed that the same regions of mPFC reported in previous studies (Beaucousin et al., 2007; D. Sander, et al., 2005) were activated specifically for the complex emotion task, suggesting that the mPFC is one of the key structures for decoding complex emotion. Moreover, an additional analysis, in which pitch was added as a potential confounder, still reveals mPFC activation, indicating that the mPFC, even if connected to temporal regions processing pitch, is not modulated by this acoustical feature. Instead, the involvement of the mPFC presumably reflects higher cognitive processing intrinsic to complex emotion such as inferring other's mental states. Given that the mPFC is involved in social cognitive processes, such as recognition of conspecifics and understanding of other's emotions intentions and beliefs (Blakemore, 2008), it is likely that the recruitment of mPFC in EPC tasks depends on the extent to which participants engage with another's perspective when attempting to decode their emotions. Decoding non-social complex emotions is based on the interpretation of cognitive beliefs which caused the current mental state, and decoding complex social emotion require the interpretation of social intentions. The mPFC activation underlies the interpretation of cognitive beliefs and social intentions, which are necessary for complex emotion, but only contingent to comprehend simple emotion.

The mPFC activation in the EPC task may be related to ToM strategies applied to the comprehension of complex emotion. In order to interpret social complex emotions, individuals may need to simulate the feelings of other people in their own mind in order to understand them. In accordance with this finding, the mPFC has been particularly associated to understanding intentions of others

(Willems et al., 2010), affective evaluation of imagined objects (Cunningham, Johnsen, & Waggoner, 2010), and it is also a component of the network supporting modality independent emotion perception (Peelen, Atkinson, & Vuilleumier, 2010). Moreover, recruitment of this region was found in emotional speech comprehension, and it has been interpreted as related to inferring and sharing other's emotion (Beaucousin et al., 2007).

Alternatively it could be argued that the mPFC activation is due to increased task demands during complex emotion as compared to simple emotion perception. As has been said, the acoustical features of prosody for complex emotion are less differentiated; thus processing EPC for complex emotion strongly relies upon cognitive interpretation, in detriment of perceptual processing resulting in a greater degree of uncertainty. This would be in line with the behavioural data showing a higher accuracy for simple emotion in comparison to complex emotion trials. However, tasks with high effortful cognitive demands have been shown to rely on the recruitment of more dorso-lateral regions of the PFC (Duncan & Owen, 2000), while the medial PFC seems to be rather involved in low demand situations, such as the absence of a task requiring deliberative processing as it has been shown in the brain default network (Raichle et al., 2001). Thus, this alternative explanation seems unlikely.

Noteworthy, the mPFC might be necessary but not sufficient for ToM. In addition to the mPFC, the present study revealed activation in the supplementary motor area as well as in the somatosensory association cortex, which is in line with the mirror neuron system role in ToM. In fact, ToM tasks focussed on indentifying

beliefs and emotional states have shown to recruit the mPFC, the IFG and somatosensory association cortex (Hooker, Verosky, Germine, Knight, & D'Esposito, 2008), which was interpreted as the use of internal affective representations to understand other's emotions (Hooker et al., 2008).

Besides the somatosensory and premotor cortices, the right frontal operculum and insula were additionally recruited for complex and social in comparison to simple emotion. However, these structures might be involved in differential aspects of EPC for complex emotion. The somatosensory and premotor cortex form a modality independent representation of emotions (Adolphs, 2010) and it has been shown that these regions are needed for the processing of facial (Adolphs, Damasio, Tranel, & Damasio, 1996) as well as vocal affect (Banissy et al., 2010; Warren et al., 2006), meaning that they form part of a general mirror neuron system for emotion processing. In contrast to the multimodal representation of emotion in somatosensory and premotor cortex, the right inferior frontal operculum seems to be related with the processing of vocal emotion in particular. The right frontal operculum, a part of the audio motor loop, comprises the engrams of orofacial movements necessary for an automatic motor mapping of prosody (Hoekert et al., 2008). Interestingly, more empathic individuals recruit motor regions to a greater extent during EPC tasks (Aziz-Zadeh et al., 2010) thus, the recruitment of the Broca's area homologue during complex emotion perception might be driven by the increased need of empathizing during complex emotion comprehension. Interestingly, Broca's homologue was also present in the subtraction of neutral from emotion trials (simple and complex together) probably because the complex trials drove this region activation. Finally, the anterior insula, known for linking the

perception of emotional stimuli with visceral responses, would be involved in sensing one's own bodily state (Keysers & Gazzola, 2006). It is likely that the somatosensory and premotor cortex simulate the perceived emotion, the anterior insula adds visceral reaction, the inferior frontal gyrus activates orofacial movement engrams to respond to the perceived emotion, and the mPFC disentangles one's owns mental states from those of others.

Complex emotion comprehension also revealed activations within the temporal lobe, such as the left middle temporal gyrus. The right middle temporal gyrus has a role in prosody decoding (Ethofer et al., 2006; Ischebeck et al., 2008). The left lateralized response of this region might be related to the more linguistic aspects of the stimuli. Other temporal activations, such as those of the right parahippocampal cortex and fusiform area are in agreement with a study of multimodal emotion perception in which medial temporal regions are triggered by the amygdala in the presence of emotionally salience stimuli (Baumgartner, Lutz, Schmidt, & Jancke, 2006). Medial temporal regions process memory for emotional arousing material automatically, being a gate between emotion and cognition (Baumgartner et al., 2006).

As a final consideration, it is noteworthy that the subtraction of complex from simple emotions did not reveal any activation. This null finding indicates that simple emotion does not involve any cognitive perceptual process not conveyed by complex emotion.

In sum, the present study revealed that EPC for complex emotion and for simple emotion share the same emotional-perceptual network. However, additional

social and cognitive neural components are recruited when processing complex emotion. By controlling for pitch differences between complex and simple emotion, the present study demonstrated that prefrontal involvement in EPC for complex emotion is relative independent of low level acoustical features. Key structures as mPFC and rSTG, and somatosensory association cortex are crucial for EPC of complex emotions. This neural network is very similar to the network that has been found in studies focussing on ToM. It is possible that inconsistent involvement of the mPFC as well as the somatosensory cortex in EPC is due to the extent in which participants try to infer belief and intention of external agents. Since making inferences about social intentions and mental states for the comprehension for simple emotions is plausible but not necessary, this skill is essential for the comprehension of complex emotions.

Chapter III. Decoding emotional prosody: Resolving differences in functional neuroanatomy from fMRI and lesion studies using TMS[1]

In the previous chapter, the neural network for EPC was revealed. Moreover, particular prefrontal components were shown to have a role in EPC for complex emotion but not for simple emotion. Thus, a temporo-frontal network for EPC was found, which is in line with the previous fMRI studies. However, there are still controversies in the literature regarding the neural correlates of EPC, particularly regarding the lateralization of prosody decoding in the posterior lateral temporal region. Findings from fMRI studies antagonise those revealed by lesions studies. While the neuroimaging literature proposes a more bilateral temporal involvement in prosody decoding, lesion studies emphasise the causal role of the right temporal region in prosody decoding.

It is important to note that functional imaging is correlative in nature and when comparing imaging and lesion data, imaging data indicate brain regions which may plausibly be involved (either directly or coincidently) in a task, but not those that are 'necessary' (Price, Mummery, Moore, Frakowiak, & Friston, 1999). In short, current neuroimaging methods do not allow us to discern causality, only correlation (Paus, 2005). On the other hand, Inferences about normal neural function based on lesion studies are not robust to problems either as lesions often lead to compensatory reorganisation, and their foci may spread across more than one region (Pascual-

[1] This chapter has been accepted for publication as an original article by Brain Stimulation. Ref. No.: BRS-D-11-00047

Leone, Bartres-Faz, & Keenan, 1999). An ability to directly inquire about the causal contribution of different brain regions is greatly needed, and recent years have seen this need partly assuaged by TMS (Pascual-Leone, Walsh, & Rothwell, 2000). TMS can be used to transiently disrupt the function of a given cortical target. It therefore allows us to empirically test specific neuropsychological models and constructs (Pascual-Leone et al., 2000) free from several confounds inherent to neuroimaging and lesion studies, including the lack of mono-directional causal brain-behaviour relationship (Sack & Linden, 2003). The following study has been design to address the lateralization of the posterior superior temporal gyri in emotional prosody decoding using TMS.

3.1 Abstract

Prosody conveys information about the emotional state and intention of others. Lesion studies have shown that damage to the right posterior temporal region is associated with prosody decoding deficits. Dissimilarly to findings from lesion studies, neuroimaging data show substantial bilateral perisylvian activation. This study aimed to investigate the involvement of the left and right superior temporal gyrus (STG) in prosodic and semantic processing using transcranial magnetic stimulation. These two regions of interest were chosen for their correspondence to Wernicke's area in the left hemisphere and its analogue in the right. Offline TMS with a stimulation frequency of 1 Hz and intensity of 60% of stimulator output (approx 1.1 Tesla) with one pulse applied per second for 10 minutes (600 pulses) was performed. Directly after TMS on the right STG, the left STG or sham-stimulation, participants completed a prosody decoding or a semantic judgment task (whether the tone/meaning was happy or sad). Reaction times (RT) for the prosodic task were significantly slower when TMS was applied in the right STG in comparison to left STG and sham conditions. TMS over both right and left STG delayed RT in the semantic task, significantly when the tone of voice was incongruent with the meaning. Our data strongly suggests that left temporal regions are not crucial to the basic task of prosody decoding *per se* however, the analogous region on the right is. Hence, involvement of the left STG in prosodic decoding revealed in previous imaging data is incidental.

3.2 Introduction

Emotional prosody is a crucial higher-order language function that encompasses nonverbal aspects of speech necessary for recognizing and conveying emotions in speech. In addition to prosody, speech also conveys linguistic or semantic content. The neural underpinnings of semantic aspects of language have been extensively assessed (Heim, 2005). It is accepted that lexico-semantic decoding predominantly relies upon the processing that occurs in the left superior temporal gyrus (STG) (Demonet et al., 1992; Kuperberg et al., 2000; Scott, Blank, Rosen, & Wise, 2000). Dissociations between emotional prosody decoding and semantic processing were found in patients with left hemisphere lesions (Heilman, Scholes, & Watson, 1975). Taking into account the impairment in the expression of emotional prosody (Ross & Mesulam, 1979) and in the comprehension of emotional prosody (Heilman et al., 1975) that patients with right hemisphere brain damage suffer from, Ross proposed that emotional prosody relies on the integrity of the right hemisphere (Ross, 1981). Further neuropsychological investigations have determined that patients with specifically right posterior lateral temporal lobe lesions have deficits in the comprehension (rather than the production or repetition) of emotional prosody (Behrens, 1988; Walker, Pelletier, & Reif, 2004) with patients with analagous lesions on the left producing near normal performance (Blonder, Bowers, & Heilman, 1991).

Even though lesion studies have indicated the importance of the right posterior hemisphere for decoding emotional prosody, the neural underpinning of this function is not without controversy. In fact, most functional imaging experiments of speech prosody, show substantial *bilateral* perisylvian activations

(Kotz et al., 2003; Mitchell & Ross, 2008). For instance, bilateral activation in the posterior superior temporal gyri (Buchanan et al., 2000; George et al., 1996; Kotz et al., 2003; Zatorre, Evans, Meyer, & Gjedde, 1992) in affective prosodic aspects of language has led to the interpretation of the left hemisphere as a contributor towards phonetic segmental processing of the vocal stimuli (Wildgruber et al., 2006). Alternatively, it has been claimed that bilateral perisylvian activation indicates an increase in task demand (Kotz et al., 2003). In other words, when the task demand exceeds the processing capacity of the right perisylvian area its homotopic region needs to be recruited. Controversy exists as to the exact causal substrate of prosody decoding, particularly in relation to the lateralization of the superior temporal cortex in this task. According to some studies, bilateral activation during prosody decoding is circumscribed to the middle superior temporal cortex adjacent to the Sylvian fissure (Belin, Zatorre, Lafaille, Ahad, & Pike, 2000; Wiethoff et al., 2008) as well as bilateral anterior and posterior regions (Buchanan et al., 2000; Kotz et al., 2003; Mitchell & Ross, 2008; Wittfoth et al., 2010). However, with respect to more posterior aspects of the superior temporal cortex, some imaging studies in this field report highly lateralized effects for posterior STG and identify a contribution to prosody decoding particularly for the right posterior superior temporal cortex (Ethofer et al., 2006; Ethofer et al., 2009). Thus, we consider that the fMRI literature on prosody decoding cannot entirely determine the causal involvement of the posterior right STG disentangling it from posterior left STG. The controversy between neuroimaging and lesion studies in this domain is thus still on-going.

In order to examine this apparent disagreement between neuropsychological and neuroimaging findings, we propose to use a neurodisruptive technique

(transcranial magnetic stimulation (TMS) to determine the absolute involvement of the posterior right and left STG to investigate any dissociations in involvement in semantic and prosodic judgements. Functional imaging is correlative in nature and when comparing imaging and neuropsychological data, imaging data indicate brain regions which may plausibly be involved (either directly or coincidently) in a task, but not those that are 'necessary' (Price et al., 1999). Also, Inferences about normal neural function based on lesion studies are not robust to problems either as lesions often lead to compensatory reorganisation, and their foci may spread across more than one region (Pascual-Leone et al., 1999).

In some previous neuroimaging studies, prosodic processing has not been totally disentangled from semantic processing due to the fact that speech conveys both kinds of information (Kotz et al., 2003; Mitchell et al., 2003). Therefore, the nature of the speech stimuli, in addition to the correlational nature of fMRI findings, make difficult to interpret whether the obtained brain responses reflect linguistic or non-linguistic processes. Our study uses the same stimuli for both conditions, but different instructions (i.e. judging the semantic content ignoring the tone of voice and vice versa) creating a prosodic and a semantic judgment task, allowing us to test for a double dissociation between the neural correlates of semantic and prosodic processing. Moreover, our study includes incongruent sentences which are necessary to identify the source from which participants derived the emotion cues. In other words, if semantics and prosody were consistently congruent we would not be able to disentangle whether the participant's responses are due to the meaning of the sentence or the tone of voice.

Our study uses the same experimental stimuli applied by Mitchell and colleagues (Mitchell, 2006; Mitchell et al., 2003). By doing so, we aimed to make our results directly comparable with this prior fMRI research. One problem in the literature consists in the diversity of tasks and stimuli applied to measure prosody decoding. For example, some studies used in single words (Buchanan et al., 2000; Ethofer et al., 2006) whether others used whole sentences (Kotz et al., 2003; Wildgruber et al., 2005). It is believed that auditory stimuli length modulates activity in the primary and secondary auditory cortex (Hu et al., 2010), thus the difference in stimuli length might have contribute towards contradictory findings. We used TMS to target the posterior part of the superior temporal gyri in both hemispheres. The right superior temporal gyrus is the contralateral equivalent to Wernicke's area (which is located in the left superior temporal gyrus and is known for its role in semantic processing) (Demonet et al., 1992) and as previously mentioned has been implicated in prosodic processing in both neuropsychological and neuroimaging studies. Moreover, it is currently believed that prosodic processing in the right hemisphere mirrors semantic processing in the left (Ross, 1981). For that reason we decided to target the posterior STG. The selection of this region of interest (ROI) for TMS as well as the design of a paradigm to evaluate the interhemispheric neuroanatomy of emotional prosody decoding by targeting homotopic regions in both hemispheres are the novel aspects of our study.

3.3 Methods

MATERIALS AND METHODS

Participants

Eleven healthy native English speakers (age *MEAN* 34 *SD* 10, 4 females) participated in the study. All participants reported to be right handed and having normal hearing. Participants gave their signed informed consent in accordance with the Declaration of Helsinki and with the approval of Durham University Ethics Advisory Committee, and could leave the experiment at any point. Subject selection complied with current guidelines for rTMS research (Machii, Cohen, Ramos-Estebanez, & Pascual-Leone, 2006; Rossi, Hallett, Rossini, & Pascual-Leone, 2009).

TMS

TMS was applied over two regions of interest, left and right posterior superior temporal gyri, chosen for their correspondence to Wernicke's area in the left hemisphere and its homologue in the right (see fig 4). Positions were located prior to and maintained during each experimental session, using frameless stereotaxy (Brainsight™, Rogue Research, Montreal, Canada) on each subject's anatomical MRI scan previously collected at the Newcastle Magnetic Resonance Centre.

Offline rTMS with a stimulation frequency of 1 Hz was performed using a 70 mm diameter figure-of-eight coil connected to a Magstim™ Super Rapid magnetic stimulator (Magstim, Whitland, Dyffed, Wales, UK). The coil was held tangentially on the skull over the ROI in a constant position with the handle pointing medially parallel to the horizontal and mid-sagittal plane by the experimenter at all times. The stimulation intensity was set at 60% of stimulator output (approx 1.1 Tesla) with one pulse applied per second for 10 minutes (600 pulses). According to the literature, this procedure should affect the neural activity of the ROI for approximately six minutes (Kosslyn et al., 1999).

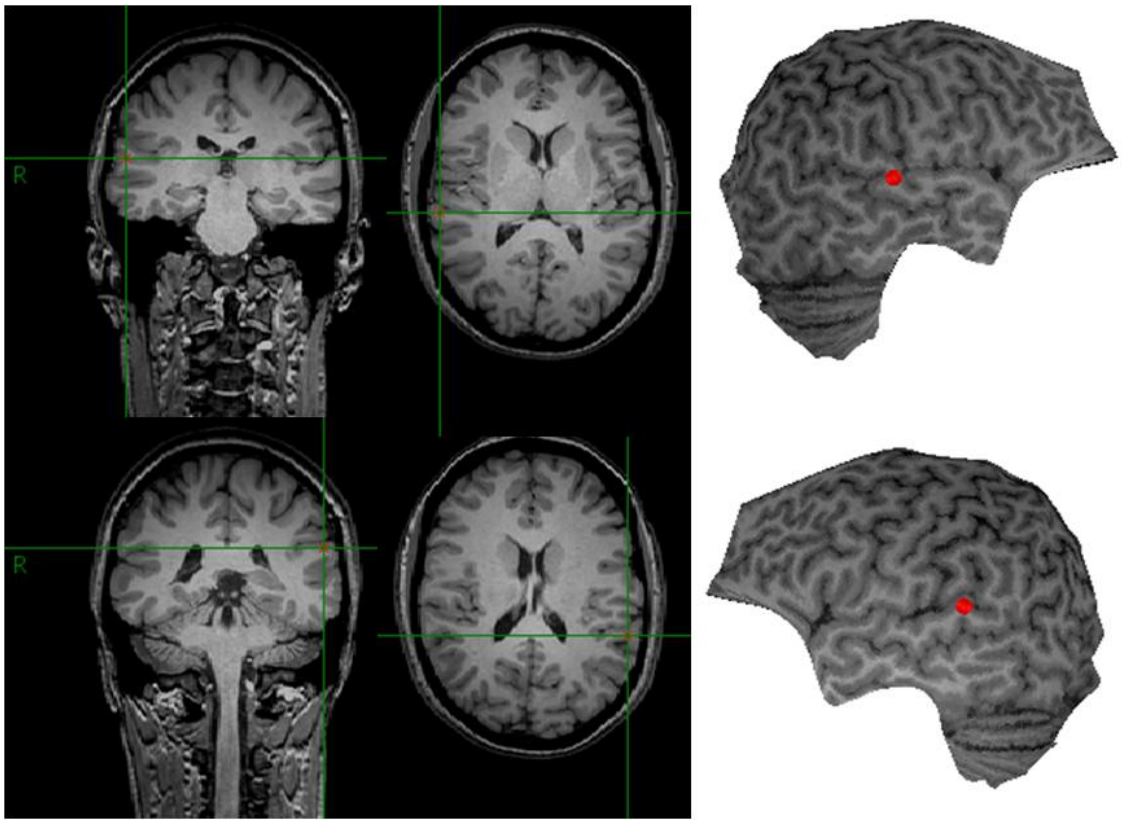

Figure 4. Stimulated areas were localized using each subject's MPrages co-registered to their skull co-ordinates using Brainsight software.

TMS procedure

Participants were seated in a comfortable chair in front of a computer screen and fitted with a swimming cap to allow marking of the stimulation sites. A chin rest was used to minimize head movements during the experimental blocks. Directly after rTMS on the right STG, the left STG or sham-stimulation (in which a non-discharging coil was held to one or other ROI whilst a discharging coil was in close proximity resulting in similar conditions to TMS application but without the magnetic pulse), participants completed one block of the task for approximately 4 min. Subsequently, there was a 30 min break before the next block so that neural activity returned to baseline because it has been demonstrated that the effects of rTMS in neural excitability that outlast the period of stimulation may last for several

minutes (Pascual-Leone et al., 1998; Pascual-Leone, Valls-Sole, Wassermann, & Hallett, 1994). The order of the experimental tasks and the stimulation blocks were counterbalanced over subjects and sessions. The entire study was completed in two sessions separated by seven day's interval. In each session, TMS on the right STG, left STG and sham-stimulation was applied and the 3 blocks which comprise the prosodic or the semantic tasks were performed. Sessions lasted approximately 1.5 hours per subject.

Emotional prosody tasks

Sentences of happy or sad semantic content (i.e. 'she was delighted to be pregnant', 'the dog had to be put down'), pronounced in a happy or sad tone of voice by a male native phonetician of British English were used as stimuli. Half of the sentences had congruency between prosodic and semantic valences (i.e. sentences with happy meaning spoken in happy tone of voice) whether the other half were incongruent. The sentences were approximately the same length (+ 100 ms) to avoid variability in decision time, and were of a consistent style and format (duration *MEAN* = 2.1511 *SD* = 0.2558). Three stimuli lists were produced. Each of the stimuli list contained equal numbers of happy and sad content sentences spoken in happy and sad tones of voice. Stimuli list and stimulation side (left, right and sham) were counterbalanced. The stimuli were the same as those used previously in the fMRI study of Mitchell et al (Mitchell et al., 2003).

The tasks were developed and presented using E-prime software (Psychology Software Tools; PA, USA). Each trial lasted 4 seconds including the sentence and the inter-trial interval. Each experimental block consisted of 60 trials.

The stimuli were presented through 2 loudspeakers, located on each side of the PC screen. Participants had to respond via a key press on a PST serial response box. The index finger (left button) was used to respond for "sad", and the middle finger (right button) was used to indicate "happy". For both tasks the same combined semantic-prosodic stimuli were used. In the semantic tasks, participants were asked to focus on the meaning of the sentence, ignoring the tone of voice and answer whether the content of the sentence was happy or sad. In the prosodic task, participants were asked to focus on tone of voice and ignore the meaning, and indicate whether the intonation was happy or sad. Participants were instructed to respond as fast as they could, but without sacrificing accuracy.

Statistical Analysis

Our main analysis comprised a 2 x 2 (task [semantic, prosodic] x TMS site [right STG, left STG]) repeated measures ANOVA using normalised reaction times (normalized RT = (RT [TMS] – RT [sham]) / RT [sham]) for comparison across task and stimulation site in order to investigate relative differences in TMS effect across conditions. This is a standard analysis which takes into account each participant's performance with respect to their relative control and so the effect of TMS across site and task can be easily compared (see (Bjoertomt, Cowey, & Walsh, 2002; Ellison, Battelli, Cowey, & Walsh, 2003)). This analysis was also completed to investigate trends in error rates.

Performance within each task was further analysed by four one factor repeated measures (TMS [sham rt v left TMS rt v right TMS rt) ANOVAs in which

the effect of congruency was also taken into account (i.e. congruent prosody, incongruent prosody, congruent semantic, incongruent semantic).

3.4 RESULTS

Error data analysis:

During both tasks and all stimulation sides (including sham) the accuracy was almost perfect (Prosody task [Mean accuracy, SE]: left STG TMS: 93 %, 0.02%; right STG TMS 92 %, 0.01%; Sham TMS: 92%. 0.02%. Semantic task: left STG TMS 98%, 0.01%; right STG TMS 97%; 0.01%; Sham TMS 98%; 0.01%. Normalized accuracy data was analyzed though a 2 x 2 ANOVA. There was no main effect of task ($F_{(1,10)}$=.083, p = 0.779, η^2 = .008) or TMS ($F_{(1,10)}$=1.558, p = 0.240, η^2 = .135). There was no interaction between TMS side and task ($F_{(1,10)}$=0.094, p = 0.765, η^2 = .009). Therefore, no further analysis on error rates was carried out. All errors were removed from the dataset for analysis of reaction times.

Reaction time analysis: This analysis reveal that there was no main effect for task ($F_{(1,10)}$=.366, p = 0.559, η^2 = .04). However, there was a main effect for TMS indicating that TMS had a greater effect ($F_{(1,10)}$=11.81, p = 0.006, η^2 = .54) on reaction times when applied over the right STG in comparison to the left. There was a significant interaction between TMS side and task ($F_{(1,10)}$=5.35, p = 0.043, η^2 = .35) (see Fig 5.). A post hoc bonferroni test comparing the effect of TMS on right and left STG for the prosodic task showed that TMS had a smaller effect over the left STG in comparison to the right ($t_{(10)}$=-4.18, p<0.001). However, there was no significant difference between TMS effect over the right or left STG in the semantic task. ($t_{(10)}$=-124, p<0.904) (see Figure 5).

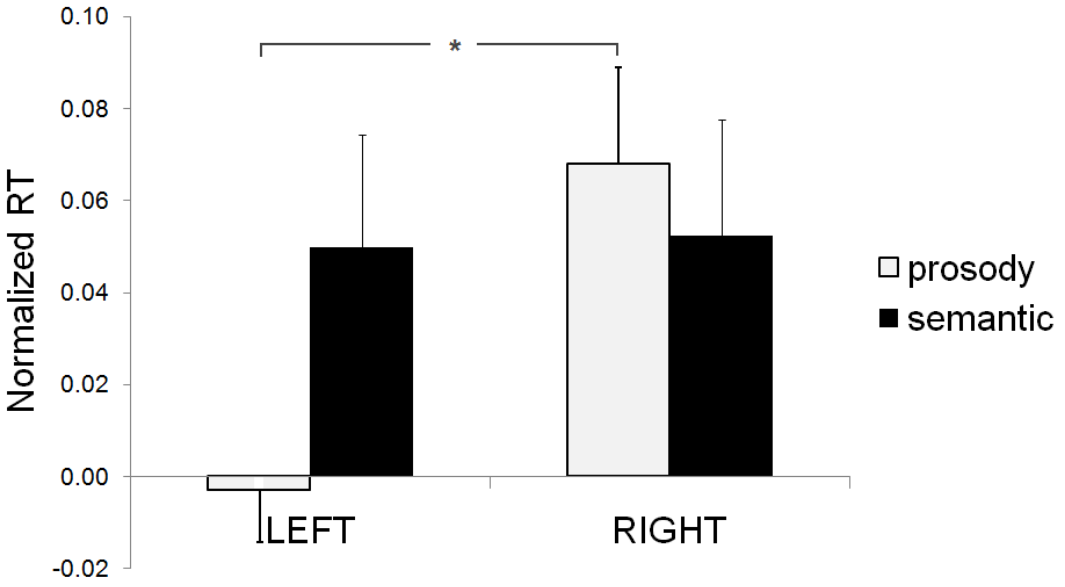

Figure 5. Normalized RT for the prosodic and semantic tasks after left and right stimulation. Bars indicate standard error (SE). A significant interaction was observed between left and right TMS for the prosodic task

The TMS effect on right STG in the prosodic task was found to be significant using a one-tailed comparison to baseline (0) ($t(10) = 3.232$, $p = 0.0045$) as was the effect of TMS over left and right STG in the semantic task ($t(10) = 2.017$, $p = 0.035$ & $t(10) = 2.068$, $p = 0.033$ respectively)

Our further within task analysis revealed there was no significant effect of TMS in congruent trials ($F_{(2,20)} = 2.18$, $p = 0.139$ $\eta^2 = .17$) in the semantic task. However, TMS did significantly affect reaction times for incongruent trials (($F_{(2,22)} = 5.91$, $p = 0.009$ $\eta^2 = .34$) in this task. As can be seen in Figure 6, post-hoc Fisher's least significant difference (LSD) pair-wise comparisons revealed that TMS significantly increased reaction times over right STG ($p = .015$) and also over left TMS ($p = .018$).

In the prosodic task, there was no significant effect of TMS in congruent trials ($F_{(2,20)}$ =2.196, p < 0.137 η^2 = .19). TMS did affect incongruent trials, however, with a main effect of TMS ($F_{(2,20)}$ = 4.095, p < 0.032 η^2 = .22) with a post-hoc significant difference (LSD) only between reaction times when TMS is applied over the right STG and sham TMS.

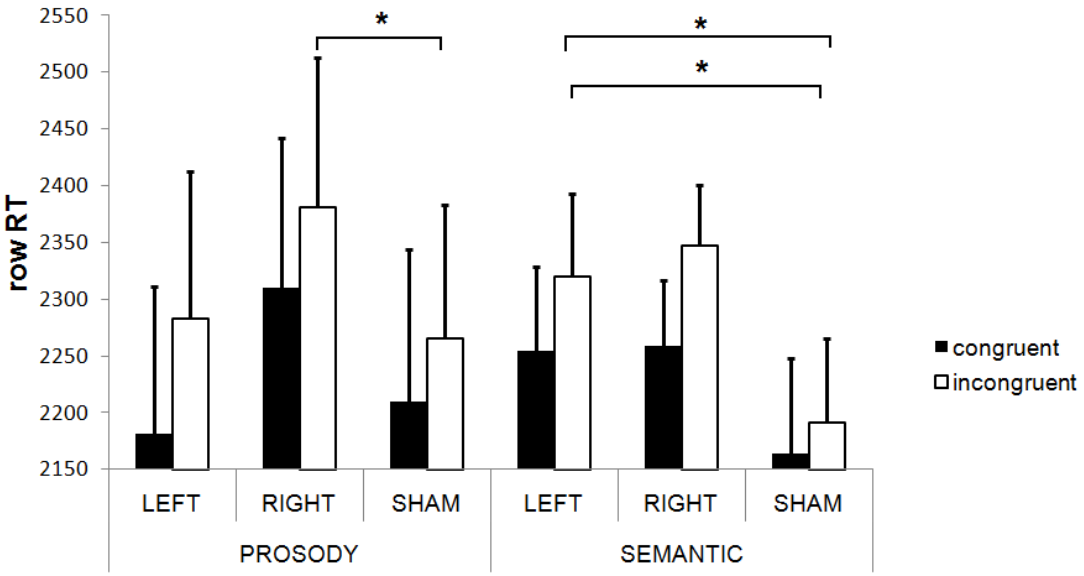

Figure 6. RTs (with *SE*) with congruency and side for each task.

3.5 Discussion

The purpose of the present study was to clarify the inter-hemispheric neural correlates of emotional prosody decoding. We have found a critical involvement of right STG for the emotional prosody task, unlike its contralateral side. We also found both right and left STG were involved in the semantic task. On further investigation however, it would seem that there is bilateral involvement of STG for trials in which

the tone and meaning are incongruent when participants are asked to make a purely semantic judgement. Such investigation shed light on how incongruency may increase semantic demand resulting in bilateral involvement.

Our main finding was a prominent distinction in the involvement of the right and left STG in emotional prosody decoding. The right STG seems to have a causal contribution in the prosodic task, as TMS over this region showed a disruptive effect, unlike TMS over its contralateral homologue. Partial eta squared showed that 35 % of the variance that was found in the analysis was associated with the effect that TMS had on the task, and this large effect size strongly suggests a dissociation between the right and left STG in prosody decoding. Our finding sheds light onto the inter-hemispheric localization of emotional prosody decoding, and is in line with lesion studies demonstrating a causal involvement of the right STG in emotional prosody decoding (Ross, 1981; Ross & Monnot, 2008a)

According to the fMRI literature, not only the right but also the left STG appears to be associated with emotional prosody decoding tasks (Bach, Grandjean, et al., 2008; Mitchell, 2006; Mitchell & Ross, 2008; Wildgruber et al., 2006). The current study used the same stimuli as were used in a previous fMRI study that uncovered bilateral activations (Mitchell et al., 2003) however, TMS has now shown that left STG does not have a causal role. The fMRI indicated left STG involvement in emotional prosody decoding has been interpreted as related to explicit labelling of emotional valences during prosody tasks or to automatic linguistic processing depending on the semantic load of the stimuli (Mitchell & Ross, 2008). Lesion studies have shown that the more complex the linguistic information embedded in the emotional prosody stimuli, the more frequent emotional prosody decoding

deficits were amongst patients with *left* hemisphere lesions (Behrens, 1985; Mitchell & Ross, 2008; Pell, 2006; Ross, Thompson, & Yenkosky, 1997). In contrast, patients with right hemisphere lesions experienced difficulty independently of the stimuli linguistic load. One interpretation for this result lies in the confounds associated with lesions studies. Functional and structural changes in homotopic regions in the cortex contralateral to a lesion have been reported in the literature, mainly linked to neural connections between the areas (Rema & Ebner, 2003). Thus effects due to brain re-organization of a cognitive function cannot be dismissed. We consider that lesion studies claiming an involvement of the left STG in emotional prosody decoding (Behrens, 1985; Pell, 2006; Ross, et al., 1997) should be interpreted taking into account brain plasticity phenomena specifically differentiation of regions homotopic to the lesion.

Due to its role in facial processing, Van Rijn et al. (van Rijn et al., 2005) used TMS to investigate the role of the right frontal operculum in prosodic processing. TMS has previously been used to investigate the role of the right frontal operculum in prosodic and the time course of involvement of this region and the right STG was further investigated by Hoekert et al. (Hoekert et al., 2008), who found that stimulation of both ROIs resulted in longer RT in comparison to a control condition. Moreover, a recent study (Hoekert, Vingerhoets, & Aleman, 2010) found increased RT in the emotional prosody decoding task during TMS over both left and right inferior frontal gyrus as compared to sham condition. They interpreted this finding as a demonstration of the critical involvement of both right and left inferior frontal gyri in emotional prosody decoding. The current study builds on these findings by investigating language processing in the temporal lobe and also including a control measure to rule out

non-specific effects of TMS. In contrast to activity in the frontal cortex, prosodic processing would seem to be lateralised to the right STG.

Our second finding shows that stimulation of either hemisphere (left STG and right STG) delayed processing of the incongruent trials of the semantic task in comparison to the baseline, and this difference was not found for congruent trials alone. Additionally, there was no difference between RTs between right and left STG. Thus, our finding should be interpreted taking into account that our stimuli contained high emotional load (emotional meaning) from which participants judged the emotional content. We interpret the bilateral involvement of the STG as related to the semantic emotional salience of the stimuli, in which the left STG would be in charge of lexico-semantic processing in general and the right STG would contribute to the processing of the emotional load convey by the meaning of the sentences. In line with our results, a considerable number of lesion studies reported that patients with right brain damage at various loci present difficulties in the perception of lexically based emotional stimuli (Borod, Andelman, Obler, Tweedy, & Welkowitz, 1992; Rymarczyk & Grabowska, 2007; Trauner, Ballantyne, Friedland, & Chase, 1996). In agreement with the lesion studies, evidence obtained with dichotic listening paradigms highlights the contribution of the right hemisphere in the semantic processing of emotions (Ely, Graves, & Potter, 1989). It may be possible that higher task demands can increase the recruitment of brain regions (Kotz et al., 2003) therefore, the judgement of incongruent trials may engage a different a neural network than the congruent trials (Mitchell, 2006; Mitchell et al., 2003). In other words, greater task demands drive the involvement of STG in a semantic task (Jobard, Vigneau, Mazoyer, & Tzourio-Mazoyer, 2007; Roskies, Fiez, Balota,

Raichle, & Petersen, 2001) therefore resulting in a greater effect of neurodisruption in incongruent trials. In our study, the stimuli employed concurrently conveyed semantic and prosodic information. By manipulating the instructions and asking participants to judge prosody whilst ignoring semantics (and vice versa), we intended to bias their attention towards the processing of prosodic features. However, we acknowledge the possibility of implicit background effects of the prosodic information. Further research including sentences with emotionally neutral meaning is urgently needed. We propose future studies investigating the neural underpinnings of semantic decoding of utterances with and without emotional content.

3.6. Conclusion

Our study strongly suggested that left temporal regions are not crucial to the basic task of prosody decoding in the absence of semantic processing, however, the analogous region on the right is. Hence, previous imaging data indicate incidental involvement of the left STG in prosodic decoding but our TMS data shows it is not necessary for pure prosody. Furthermore, it has also shown that left as well as right STG is involved in the semantic judgement of sentences with emotional meaning. This may be because right STG was involved on emotional grounds (as all sentences had emotional meaning) and left STG on semantic grounds. We anticipate further research addressing the neural correlates of semantic judgement with and without emotional meaning.

Chapter IV: Investigating the links between schizotypal personality trait and emotional prosody comprehension, a pilot study

In the previous chapters, the neural correlates of emotional prosody comprehension (EPC) in healthy population were evaluated. EPC deficits are widely observed in schizophrenia and it has been proposed that such deficits are linked with the emergence of auditory verbal hallucinations (AVH) in this mental disorder. Although impairment in cognitive and affective domains, such as EPC deficits, seem to be intrinsically related to schizophrenia, it is difficult to extricate some confounds which might have an impact in the observed impairment. Research in schizophrenia needs to control for the confounding effects of antipsychotic medication that is used to treat the disorder, for the qualitative and quantitative variability in symptoms during the course of the illness, and for the potential existence of a generalized performance deficit intrinsic to the disorder. There have been some methodological approaches aiming to overcome such confounding effects.

One way to disentangle confounds of medication, years of hospitalization and general executive (performance) deficits consist in examining whether a particular deficit is associated with a specific symptom. Another possible solution consists in designing cross-sectional paradigms evaluating the schizophrenia spectrum, for example by measuring levels of schizotypy in non-clinical population. By implementing such strategies, pathophysiological mechanisms associated with the core cognitive and social impairments of the schizophrenia spectrum could be dissociated confounds related to the recurrent or chronic psychosis, such as cognitive

and social deficits seeing in chronic schizophrenia (Siever & Davis, 2004). For all those reason, the following study presents a novel paradigm in which the relation between EPC deficits and hallucination traits is assessed in non-clinical schizotypal population. The relation of EPC with traits linked to other symptoms of schizophrenia is also evaluated.

4.1 Abstract

Emotional prosody comprehension (EPC) impairment has been reported in certain psychiatric conditions such as schizophrenia. Such impairment has traditionally been linked to the negative symptoms of psychoses. However, new findings suggested that EPC deficits are also associated with the positive symptoms of psychoses. This study aimed to test whether this association extended to positive schizotypal traits in healthy populations, and to examine how this normal personality trait modulates EPC. Fifty healthy participants completed the O-LIFE schizotypal personality questionnaire and were allocated by a median split to either of two groups, loading high and low in the unusual experiences subscale, which measures personality traits associated with positive symptoms of the schizophrenia spectrum. Participants' EPC speed was tested for both 'simple' and 'complex' emotions, because both types of emotions have shown to be separable domains and can be differentially affected in certain clinical populations. In general, reaction times (RT) for complex emotions were longer than RT for simple emotions. Moreover, there was a main effect of group; that is the high unusual experience group obtained slower RT than did the low unusual experience group. Such difference was slightly more pronounced for complex emotion. The results suggest a relation between positive schizotypal traits and EPC, which might be mediated by the complexity of the emotional valences. However, due to the small sample size, the results should be interpreted with caution. The present study sheds light on how personality traits associated with psychoses can impair the exchange of social signals diminishing interpersonal relations.

4.2 Introduction

Social cognition concerns the ability to construct representations of the relations between ourselves and others (Adolphs, 2001). It comprises the ability to perceive and understand social and emotional cues, the capacity to recognize social and emotional information, knowledge of different social behaviours and their consequences in particular social environments and the ability to empathise, the latter being a key element of theory of mind (ToM) (Bauminger et al., 2005). In each of these socio-cognitive scenarios, integrative processing of emotion cues is required to successfully interact with peers. Decoding and encoding these emotion cues has an important informative function regarding the intentions and feelings of the sender.

Facial and vocal expressions are important channels of non-verbal communication. Fleeting changes in the countenance of a face and in the tone of a voice are essential aspects of human social interaction (Russell, Bachorowski, & Fernandez-Dols, 2003). In everyday interactions, humans are persistently exposed to verbal communication, meaning that massive amounts of social information are carried by the voice (Belin et al., 2004). Speech encodes semantic information, but also carries non-linguistic information collectively known as prosody. These prosodic elements comprise acoustic features such as pitch, amplitude and segment and pause duration, and one of their key uses is to encode affective information (Belin et al., 2004). Thus, prosody allows for the encoding and decoding of feelings in speech, a skill which is necessary to ensure effective communication (Beatty, et al., 2003). This act of decoding emotion cues conveyed by prosody may be called emotional prosody comprehension (EPC); however, this label should not misleadingly be interpreted as representing a single construct. EPC is in fact a multi-

level construct, that begins with the decoding and encoding of simple emotions, followed by the assessment of complex mental states, culminating in the highest level of complexity in emotional ToM (also known as empathy) (Singer, 2006). There are qualitative differences between the decoding of simple emotions and complex emotions.

Ontogenetically, simple emotions are thought to be innate. They are predominantly negative in valence (except for happiness) because they are involved with self-preservation and fight-flight behaviours (Ross & Monnot, 2011). Moreover, they include a distinctive, universal physiological response (Ekman, 1992), and their facial and vocal displays are recognized across cultures (Pell, Monetta, Paulmann, & Kotz, 2009; Ross, Prodan, & Monnot, 2007) (for more details about this topic see Chapter II). In contrast, complex emotions are learned emotions acquired through development. Some examples of complex emotion are thoughtfulness and interest, which reflect the inner thoughts of an individual (Shaw et al., 2005), and are in fact mental states involving the adjudication of a cognitive state as well as a feeling (for more details see chapter II). Within complex emotion there is a special category called social emotions. This special type of complex emotion requires consideration of other people, comprehension of social norms, and recognition of personal responsibility for the consequences of a situation (Bauminger, et al., 2005). Even if complex emotion in general might derive from simple emotion, additional cognitive elaboration is necessary to process such complex mental states (Johnson-Laird & Oatley, 1989).

Notable inter-individual differences in ability to comprehend simple and complex emotions occur, even in healthy young adults free from neurological or

psychiatric dysfunction. Individual differences in the expression of emotions impact on feelings of well-being, guide social relationships, and establish hierarchies (Gross & John, 2003; Keltner & Haidt, 1999). In the extreme, many authors proposed that deficits in the ability to express and interpret emotion cues are known risk factors for mental illnesses such as schizophrenia (Davidson et al., 1999; Done, Crow, Johnstone, & Sacker, 1994; Keshavan et al., 2010; Litter & Walker, 1993; Niendam, Jalbrzikowski, & Bearden, 2009; Olin & Mednick, 1996; van 't Wout et al., 2007). One particularly strong candidate for the modulatory factors that drive these normal variations are continua of personality traits.

One example of such a personality trait is schizotypy. From the standpoint of a "fully dimensional" approach (i.e., an individual-differences approach), schizotypy represents a continuously distributed trait throughout the population. It is associated with normal functioning for most of its extent, but relates to schizophrenia at its upper extreme (McCreery & Claridge, 2002). In this extreme group, patients with schizophrenia show significantly lower accuracy in EPC tasks (Edwards, Pattison, Jackson, & Wales, 2001; Haskins, Shutty, & Kellogg, 1995; Hoekert et al., 2007; Matsumoto et al., 2006; Schneider, Gur, Gur, & Shtasel, 1995). The consistency of these findings and their detection at early stages (in first episode patients) may indicate that they are a trait marker of schizophrenia (Edwards et al., 2001). In addition to enhancing our understanding of normal EPC, studies of schizotypal personality traits also inform our understanding of abnormal EPC in schizophrenia, since this approach excludes potentially confounding epiphenomena surrounding the disease such as medication, long-term psychiatric care, etc. Some studies have already demonstrated the relationship between schizotypal personality traits and

different aspects of socio-cognitive functioning in healthy young adults. These studies support the concept of a continuum of socio-cognitive ability between schizophrenia and schizotypal personality (Fernyhough, Jones, Whittle, Waterhouse, & Bentall, 2008; Henry, Bailey, & Rendell, 2008; Jahshan & Sergi, 2007; Kerns, 2005; Langdon & Coltheart, 2004; Mohanty et al., 2008; Pickup, 2006; Shean, Bell, & Cameron, 2007; van 't Wout, Aleman, Kessels, Laroi, & Kahn, 2004). Although these findings have included comprehension of prosodic emotion cues (Henry et al., 2008; Kerns, 2005; Shean, et al., 2007; van 't Wout, et al., 2004), they are limited to perception of simple emotions. Given the differences between simple and complex emotions, particularly the additional cognitive operations needed to decode complex emotions, there are reasons to suspect the performance in EPC will be lower for complex than for simple emotion. Moreover, as high schizotypy is associated with difficulties in social cognition, this group may have even more severe difficulties in EPC for complex emotion in comparison to simple emotion. It is thought that the more complex the emotional process, the more difficult it is to decode and/or the greater the need for additional cognitive operations, and also the more prone EPC is to the effects of psychological ill-health (e.g. schizophrenia) or extreme personality characteristics (e.g. schizotypal). In other words, inter-subject variability in EPC tasks, in particular EPC for complex emotions, may be negatively affected in people high in positive schizotypal traits.

The constellation of behaviours labelled as ‘schizotypal’ comprises several dimensions. Mason et al. (1995) proposed four dimensions of schizotypy. Their first dimension refers to anomalous perceptual experiences, which they consider the most essential marker of the schizophrenia spectrum. This unusual perceptual experiences

dimension indicates a trait toward hallucinatory and magical thinking, and is consistent with the positive symptoms of psychosis (Mason, Claridge, & Jackson, 1995). The other dimensions relate more closely to negative symptoms of psychoses. There is a cognitive disorganization dimension, which relates to attentional deficits; an introvertive anhedonia subscale, which measures lack of enjoyment of social sources of pleasure; and an impulsive non-conformity subscale, which relates to antisocial behaviour. Some studies in non-clinical populations have shown that social cognition, including ToM and EPC, is most closely related to the positive schizotypy (Fyfe, Williams, Mason, & Pickup, 2008; Kerns, 2005; Langdon & Coltheart, 2004; Pickup, 2006). In fact, proneness toward unusual experiences has also been linked to subtle ToM impairment in healthy populations (Pickup, 2006). Explanations for the association between unusual experiences in non-clinical populations and emotional processing deficits include the suggestion that individuals with such difficulties tend to be less successful in coping with stressful situations and thus withdraw into a world of fantasy (Shean et al., 2007).

In accordance with the literature reviewed above, it was expected that people with a high tendency to unusual perception will have poorer EPC abilities than those people low in unusual perception. In contrast, it was expected that the other schizotypal personality traits should not have a significant impact on EPC, since they are more typically associated with negative schizotypy. Specifically, a significant interaction between block emotion (simple emotion and complex emotion) and Group is predicted by which the group high in unusual experiences shows particularly low performance for complex emotion compared to simple emotion. The effects are expected especially in response time because response times have been

shown to constitute a more sensitive measure to detect differences in EPC than has accuracy (Chevallier, Noveck, Happe, & Wilson, 2009). As social cognition has direct influences on the quality of interpersonal relationships and psychological well-being, our research has implications for the psychology of individual differences and for psychological medicine.

4.3 Methods

Participants:

A group of undergraduate students (N = 43) and non-academic staff (N = 7) from the Department of Psychology at Durham University (N = 50) participated in the study. Students were recruited from the participant pool of the Department of Psychology, which consists of a large body of undergraduate students registered on psychology modules. These students earn course credit by participating in psychological research. Non-academic staff were recruited via an advertisement on the university's intranet. The final sample consisted of 29 females and 21 males (mean age 27 years, SD = 11 years). All participants were native speakers of British English. The exclusion criteria preventing participation in the study included current or past psychiatric disease, known hearing impairment, history of drug or alcohol abuse, long periods of unconsciousness and head injuries. Self report screening was used to assess the exclusion criteria. Participants signed a consent form certifying that they did not match any of the exclusion criteria listed. Ethical approval for the conduct of the study was obtained from the Ethics Sub-Committee of the Psychology Department at Durham University.

Measures

Emotional prosody decoding.

The stimuli used in the experimental task were created by the Linguistic Data Consortium (LDC), an open consortium of universities, companies and government research laboratories, hosted by the University of Pennsylvania. Utterances were recorded on two channels, with a sampling rate of 22.05K, and two microphones, a stand-mounted boom Shure SN94 and a headset Sennheiser HMD 410. Resultant files were encoded in interleaved 16-bit PCM, high-byte-first format. The EPC stimuli used in the current study comprised numbers spoken in different tones of voice, selected from Banse and Scherer's study of vocal emotion expression in which fourteen separate valences were identified (Banse & Scherer, 1996). The utterances were recorded by six professional actors (three females). More details about the development work by the LDC can be found at http://www.ldc.upenn.edu.

From the full stimulus set, number stimuli spoken in four simple emotion categories (happy, sad, fearful and angry) and four complex tones (proud, guilty, interested and bored) were selected. Neutral intonation stimuli were also used, giving a total of 240 utterances (24 per emotional category and 48 for the neutral categories, which were included as a baseline condition in both the simple and complex emotion blocks). The mean duration of each spoken number stimulus was 2 seconds. Stimuli were presented using E-Prime (Psychology Software Tools, Philadelphia) via digital stereo headphones Sennheiser HD201. Different emotional valence trials appeared randomly within each block, separated by an inter-trial interval of 4 seconds.

Simple and complex emotion stimuli were presented in two separate EPC blocks, administered in counterbalanced order. In both EPC blocks, participants were

asked to identify the emotion conveyed by tone of voice for each trial, and indicate their forced-choice response by means of a response box. This response box had five keys, each labelled with the simple or complex emotion tones, plus neutral. After the first block, the labels were changed according to which block was next. Participants were asked to respond as fast as they could, but without sacrificing accuracy. RTs in EPC were measured.

The Oxford Liverpool Inventory of Feelings and Experiences (O-Life).

As described above, this scale was used to measure extent of schizotypal personality traits. O-Life is a self-report measure of schizotypy developed by Mason and colleagues (Mason et al., 1995). This scale has been validated in 508 healthy participants and each of the subscales has been shown to have high internal consistency (full psychometric data and norms for the O-LIFE are given in Mason et al, (1995)). The scale contains a total of 105 questions. The normative data of the cited validation study indicated that the mean for the unusual perception subscale is 9.7 with a standard deviation of 6.7.

The Positive and Negative Affect Schedule.

The PANAS was administered to measure emotional status over the two weeks immediately before the testing session. The scale consists of 10 positive affects (interested, exited, strong, enthusiastic, proud, alert, inspires, determined, attentive and active) and 10 negative affects (distressed, upset, guilty, scares, hostile, irritable, ashamed, nervous, jittery and afraid). For each affect, participants are asked to indicate to what extent they feel that affect on a scale from 1 to 5, where 1 is very slightly or not at all and 5 is extremely (Watson, 1988).

4.4 Results

Median split in the unusual experience subscale

All participant were allocated to one of two groups by using a median split based on their score in the unusual experiences subscale of the O-LIFE questionnaire, resulting in a group high (N = 25, 10 males) and low in unusual experiences (N = 25, 10 males). A paired sample t-test showed that the groups high (13.8 ± 1.1; range: 0-8) and low in unusual experiences (4.2 ± 2.4; range: 9-34) differed significantly from one another ($t(24) = -6.88$, $p < .001$). Group (high and low in unusual experiences) was included in all subsequent analyses as a between-subjects factor.

Moreover, since mood status is known to have an impact on EPC, thus becoming a potential confound (Brand et al., 2009; Velez Feijo, Rieder, & Chaves, 2008), the PANAS score was included as a covariate in the statistical design. The groups did not significantly differ in age, sex or verbal IQ (see Table 4). Greenhouse-Geisser procedure was used if the assumption of sphericity was violated.

Table 4. Summary of participant's demographic and cognitive profile. Group comparisons, using independent samples t-test, of high Vs low Unusual Perception (UE). Group Means and Standard Deviations are reported. No significant differences were found.

	Groups		
	High UE (N=25)(10 male)	**Low UE (N=25)(10 male)**	**Independent t test**
Age	26.48 (12.23)	27.04 (9.69)	(t = 0.18, p=.861)
Verbal IQ	117.34 (4.35)	115.95 (4.22)	(t = -1.09, p=.282)

Accuracy

Accuracy data were subjected to a 2 x 2 mixed ANCOVA with Emotion Block (simple emotion, complex emotion) as within-participants factor, Group (high unusual experiences, low unusual experiences) as between-participants factor and PANAS score as covariate. The results revealed a main effect of Emotion Block ($F_{(1,20)}$ = 49.05, p < .001) with higher scores for simple (80.91%, ± 7.98) than complex emotions (63.39%, ± 10.81). Neither the main effect of Group not the interaction between Group and Emotion Block approached significance (all F < 1.85, n.s.). The covariate show neither significant main effects neither nor interaction (all F < 0.90, n.s.).

Response times

The ANCOVA analysis performed on accuracy data was repeated with response times. There was a main effect of Emotion Block ($F_{(1,48)}$ = 5.67, p < .05) indicating faster RTs for simple emotion (2000 ms ± 415) in comparison to complex emotion (2436 ms ± 616). Moreover, there was a main effect of Group indicating that the high unusual experience group (2394 ms ± 544) obtained slower RT in comparison to the low unusual experience group (2041 ms ± 392) (F = 7.46, p < .05). The interaction between Emotion Block and Group only approached significance ($F_{(1,48)}$ = 3.04, p = .088; η^2 = .06). It should be noted, however, that this trend was in the predicted direction, that is group differences were slightly stronger for complex emotions than simple emotion. The covariate did not show significant main effect neither interaction (all F < 0.01, n.s.).

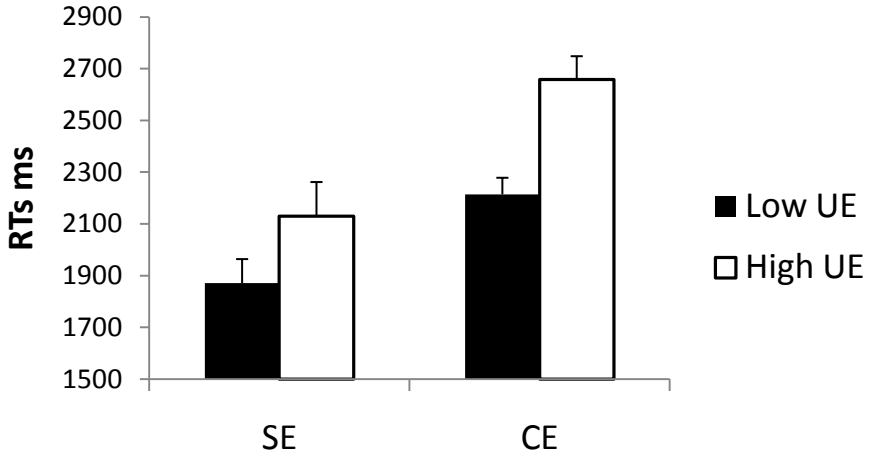

Figure 7. Mean response times in milliseconds for the EPC in simple emotion (SE) and complex emotion (CE) for both low (black bars) and high (white bars) unusual experience groups divided by median split. Bars represent standard errors.

Previous research in normal populations has shown differences in cognitive profiles between 'high' and 'low' schizotypal groups only in extreme groups, although in much larger samples (Cohen, Iglesias, & Minor, 2009; Fernyhough et al., 2008; Jahshan & Sergi, 2007; Kerns, 2005; Pickup, 2006; van 't Wout et al., 2004). It could be that the grouping procedure based on the median split of the present study was not sufficient to obtain significant group differences. Therefore, we designed more extreme groups by including only those participants who were 1 SD below or above the mean score of 9.16 in the unusual experiences subscale, resulting in a group high (17.75 ± 6.09; range: 13-34, N = 12,) and low in unusual experiences (2.08 ± 1.24, range: 0-4, N = 12,). The remaining 26 subjects within 1 *SD* from the group mean were excluded from the analyses.

Accuracy rates and RTs were subjected to a 2 x 2 mixed ANCOVA with Emotion Block (simple emotion, complex emotion) as within-participants factor,

Extreme Group (extreme high unusual experiences, extreme low unusual experiences) as between-participants factor and PANAS score as covariate. The results of the accuracy data revealed a main effect of Emotion block ($F_{(1,20)} = 49.05$, $p < .001$) with higher accuracy rates for simple than complex emotion, but there was no other main effect or interaction (all $F < 1.63$, ns). The results of RTs revealed main effect of Emotion Block ($F_{(1,20)} = 6.52$, $p < .05$), with slower response times for complex emotion than simple emotion. However, the main effect of Extreme Group only approached significance ($F_{(1,20)} = 3.19$, p= .09), with numerically faster RTs for the extreme group low in unusual experience (1999 ms ±391) than the extreme group high in unusual experience (2303 ms ± 427). In addition, there was a significant interaction between Extreme Group and Emotion Block ($F_{(1,20)} = 6.23$, $p < .05$). Alpha-adjusted post-hoc paired sample t-tests showed that response times in the low unusual experience extreme group did not significantly differ between simple emotion (1933 ms ± 290) and complex emotion (2042 ms ± 493), ($t(11) = 1.24$, ns). However, the high unusual experience extreme group showed faster RTs in simple emotion (2042 ms, ± 322) that in complex emotion (2500 ms ± 533) ($t(11) = 4.88$ $p < .001$). However, given the small sample sizes the results of the extreme group comparison should be interpreted with caution and rather have an exploratory character.

Other O-LIFE subscales

Corresponding 2 x 2 mixed ANCOVAs for the other subscales (i.e., median-splits for Impulsive Non-conformity, Cognitive Disorganization and Introvertive Anhedonia) did not reveal any significant Group effects interaction, neither for accuracies (all $F < 2.40$, n.s.) nor response times (all $F < 1.74$, n.s.).

4.5 Discussion

The present study demonstrated that RTs are longer for complex emotions than for simple emotions. Moreover, the grouping approach based on the median split showed increased RTs in EPC for the high unusual experience group than for the low unusual experience group. Additionally, the exploratory extreme-group approach suggested an interaction between unusual experience and RT in the EPC blocks in that the high unusual experience group showed longer RTs for complex emotion than for simple emotion. However, this interaction did not reach significance in the median split analysis.

The differences in RTs and accuracies between complex and simple emotion suggest that complex emotions such as boredom and interest are intrinsically more difficult to understand than simple emotions because they reflect mental states which arise later in development in comparison to basic simple emotion. Moreover, expression of complex emotions may be more subtle than those of simple emotions (Kington et al., 2000). Complex emotion processes acoustical features such as pitch, amplitude and timing that are less specific than those for simple emotion (Bryant & Fox Tree, 2005), making the judgment of vocal complex emotional expression more difficult. Processing of simple and complex emotions are also mediated by different neural correlates (Adolphs, Baron-Cohen, et al., 2002) (see also chapter II). The modularity of these domains, or in other words how these two groups of emotion can be dissociated, has been already demonstrated in clinical populations (Adolphs, Baron-Cohen, et al., 2002; Kington, et al., 2000). The present findings are in line with evidence showing that processing simple and complex emotions is mediated by different neural pathways (Adolphs, Baron-Cohen, et al., 2002; Kington et al., 2000).

The results also revealed that differences in RTs between groups on EPC were found with the unusual experiences subscale. This effect was not found in accuracy because, at it has already been mentioned, such measure do not seem to be sensitive enough to detect differences between groups in EPC performance (Chevallier et al., 2009). The association between unusual experiences and EPC difficulties has been previously demonstrated non-clinical schizotypal population (Shean et al., 2007). It is possible that healthy people are advanced in decoding emotions, as this ability is necessary to interact effectively with other individuals. However, participants scoring high on unusual experience, a trait related to the positive symptoms of schizophrenia, find EPC generally more difficult, regardless whether simple or complex emotions are processed. In fact, it has been proposed that individuals who have difficulty perceiving emotion "are more likely to cope with stressful interactions by withdrawing into a world of fantasy" (Shean et al., 2007). It might be the case that aberrant auditory processing, particularly false auditory perception, underlies both EPC difficulties and unusual experience. However, regardless of whether the mechanism underlying unusual experience and EPC are associated, the influence that personality traits have on social skills may be of central importance in understanding why some individuals are more successful in prosocial behaviour than others. The present study also shed light on why even in healthy populations EPC can fail, impoverishing interpersonal relationships (Pickett, Gardner, & Knowles, 2004). Importantly, differences in EPC performance between groups are unlikely to be due to a psychometric or demographic confound on which the participants were measured, as the groups did not statistically differ from each other in IQ and age. Moreover, the groups were matched in gender.

The extreme-group analysis suggests an interaction between Emotion block (simple vs. complex) and Group (low vs. high unusual experience extreme groups), even though this interaction did not reach significance in the median-split analysis. The high unusual experience extreme group showed particularly slow RTs when judging the complex emotions compared to simple emotion. Although this effect should be interpreted with caution due to small sample size, this observation is in line with previous studies with much larger sample sizes showing that patients suffering from schizophrenia are particularly impaired in areas related to social functioning, which may reflect a continuum in the schizophrenia spectrum (Bora, Gokcen, & Veznedaroglu, 2008; Bozikas et al., 2006; Couture, Lecomte, & Leclerc, 2007; Hooker & Park, 2002a, 2002b; Nelson, Combs, Penn, & Basso, 2007). EPC deficits in general have previously been reported in schizophrenia (Bozikas et al., 2006; Edwards et al., 2001; Kington et al., 2000; Leentjens, Wielaert, van Harskamp, & Wilmink, 1998; Ross et al., 2001). Moreover, even when schizophrenia patients seem to perform relatively similarly to healthy controls in decoding simple emotion, they may present difficulties at decoding complex emotion (Kington et al., 2000). The difference in RTs between high and low unusual experience groups in EPC in general, in addition to the trend suggested by the extreme-groups approach in which EPC for complex emotion was particularly compromised in the high unusual experience group, seems to be partially in line with the idea of social cognition impairment in high schizotypy (Cohen et al., 2009; Henry et al., 2008; Kerns, 2005; Shean et al., 2007; van 't Wout et al., 2004). Mental state decoding skills might be critical to solve the EPC task, particularly for complex emotion which may rely less on perceptual processing and more on higher cognitive functions (for an extended

discussion on this topic see chapter II). Moreover, participants with high scores in unusual experience might have had extra difficulties decoding complex emotion because they might not have the same flexibility to switch from one point of view to another, which is a necessary step to understand social emotions (Takahashi et al., 2004). This lack of flexibility has been highlighted as characteristic of people who experience delusions and persecutory ideation (Colbert, Peters, & Garety, 2009; Guillaume et al., 2007; Ross, Freeman, Dunn, & Garety, 2009).

Notably, because of the exploratory nature of the analysis, the finding of the interaction between Group and Emotion block should not be taken as conclusive. For this analysis the sample was divided using median split; and for an additional exploratory analysis using extreme-groups approach. This procedure was done following the recommendations for practice of extreme-groups approach by Preacher and colleagues (2005). Still, this analysis allows for the presentation of preliminary evidence which suggests a link between social cognition and unusual experiences. However, as the interaction was not consistent across analysis, alternative explanations to this finding should be sought.

The small size of the present sample may affect the reliability of the result. With the exception of the study by Pickup (2006) who obtained a sample of 62 participants, previous studies using the extreme-group approach in schizotypy collected data from more than 200 participants (Cohen et al., 2009; Fernyhough et al., 2008; Jahshan & Sergi, 2007; Kerns, 2005; van 't Wout et al., 2004), resulting in extreme groups of at least 20 individuals. It is difficult to establish how the low and high groups differed in schizotypal traits between studies due to fact that different scales were used to measure schizotypy. Due to the small size of the present sample,

our participants might not have presented schizotypal traits quantitative and qualitative related to schizophrenia symptoms. However, Cochrane and colleagues (2010) assessed schizophrenia patients with the Unusual experience subscale, showing that the mean score of their patients in Unusual experience was similar to the scores of the high group (median split analysis) of the present study (in schizophrenia patients M = 14.9; the present study high group M = 13.8). Taken together, it seems that the present high unusual experience group show similar unusual experiences traits than medicated schizophrenia patients. A second interpretation for the inconsistent interaction consists of other factors, besides positive symptoms, which may contribute to the impairment in decoding complex emotion in schizophrenia. Perhaps additional neurocognitive impairment is the basis of impaired specifically complex emotion decoding. Differences in neurocognitive impairment between the present sample and previous studies in schizophrenia patients cannot be ruled out. Alternatively, it is important to note that even if the unusual experience dimension strongly relates to the positive symptoms of schizophrenia, there may be not just quantitative but also qualitative differences between unusual experiences trait and positive symptoms. For example, even when healthy population may experience auditory hallucinations, the phenomena differs between patients and non-clinical population terms of controllability and emotional valence (Jenner, Rutten, Beuckens, Boonstra, & Sytema, 2008), and such differences may be critical to EPC for complex emotion.

This study shows that the only schizotypal dimension associated with difficulties in EPC is Unusual experiences, as the other schizotypal subscales did not reveal any association with the experimental task. This is in agreement with some

schizophrenia research, in which deficits in EPC seem to be associated with hallucinations in particular and not to any other symptom of psychosis (Rossell & Boundy, 2005; Shea et al., 2007). In contradiction to the cited studies, some research in schizophrenia linked EPC difficulties to neurocognitive deficits such as poor executive function and attention impairment (Bozikas, Kosmidis, Anezoulaki, Giannakou, & Karavatos, 2004; Roux, Christophe, & Passerieux, 2010). However, executive impairment in schizophrenia seems to underlie poor outcome in many different cognitive domains (Fossati, Amar, Raoux, Ergis, & Allilaire, 1999). The present sample was integrated by a majority of high functioning university students, thus it is highly unlikely that any of the participants had substantial executive function impairment. However, differences between groups in EPC skills were still present, which are difficult to explain by differences in executive function between groups. Differences in EPC between groups seem to be associated to unusual experiences. A recent study assessing schizophrenia patients and schizotypal population indicated that the unusual experience subscale is correlated with the positive symptoms of schizophrenia as measured in the Scale for the Assessment of Positive Symptoms (Andreasen, 1984b) in both schizophrenia and non-clinical population (Cochrane, Petch, & Pickering, 2010). It should be noted, however, that a relationship between negative traits as measured in the O-LIFE and negative symptoms of psychosis was not found by Cochrane et al. Taken together, this evidence indicates that Unusual experience is the O-LIFE subscale that mirrors clinical symptoms more closely. If this is the case, one can speculate that the present study found significant differences in the EPC task in relation to this subscale because it is the most relevant to schizophrenia. In fact, unusual experience subscale

does not have a clear corresponding trait in the standard models of normal psychology, unlike the others subscales (Cochrane et al., 2010). Unusual experience seems intrinsically and exclusively related to schizophrenia.

Taken together, the findings suggest that complex emotion and simple emotion are distinct, and their comprehension is affected by a personality trait associated with the positive symptoms of psychoses. The suggested modulation of EPC by the positive schizotypy dimension provides additional knowledge about how a personality trait can be on the basis of inter-individual variability in social cognition skills even in a healthy population. To the best of our knowledge, this is the first attempt to demonstrate differences in the processing of complex emotion in the absence of differences in processing simple emotion in relation to positive schizotypal traits. It will require further advances in socio-cognitive neuroscience to evaluate how personality traits such as schizotypy might mediate the relationship between emotional processing, and to reveal their underlying neural correlates mechanisms of such modulation.

PART II: Studies in schizophrenia patients

Chapter V: Voice identity recognition failure in schizophrenia patients with auditory verbal hallucinations.

The first part of the thesis assessed emotional prosody comprehension in a non-clinical population. The neural underpinning of EPC in healthy population was revealed. Moreover, a link between EPC difficulties and certain schizotypy traits has been suggested, which is in line with literature indicating that prosody processing in schizophrenia is affected, particularly for patients who suffer from AVH (Rossell & Boundy, 2005; Shea et al., 2007).

Interestingly, voice identity recognition and EPC rely on partly similar mechanisms. Specifically, both rely in part on the same acoustical features (pitch, amplitude, tempo). Voice identity recognition in schizophrenia has been rarely studied; and the few studies assessing it have varied in methodology to such an extent that it is still unknown whether there are deficits in this capacity in schizophrenia (Drakeford et al., 2006; Waters & Badcock, 2009; Zhang et al., 2008). The main reason for studying voice identity recognition in schizophrenia relies on the fact that the voices heard in hallucinations usually have the vocal characteristics of someone known to the patient, instead of sounding like the patient's own voice. Misidentification of the source of the voice is a crucial phenomenological feature of AVH. It has been hypothesise that this misidentification might have its roots in difficulties in recognizing the identity of a speaker. Difficulties to recognize one's own inner speech due to incapacity to differentiate individual voices, might lead to source misattribution of inner speech. Using a validated paradigm for the identification of famous voices (Damjanovic & Hanley, 2007), which has been shown to be sensitive to the assessment of phonagnosia patients (Garrido et al.,

2009; Hailstone, Crutch, Vestergaard, Patterson, & Warren, 2010), this chapter will test the existence of voice identity recognition impairment in AVH patients.

5.1 Abstract

Cognitive models propose that auditory verbal hallucinations arise through inner speech misidentification. However, such models cannot explain why the voices in hallucinations often have identities different from the hearer. This study investigated whether a general voice identity recognition difficulty might be related to auditory verbal hallucinations. Thirteen schizophrenia patients with hallucinations (AVH), twelve schizophrenia patients without hallucinations (NAVH) and thirteen healthy controls were tested on recognition of famous voices. Signal detection theory was used to calculate perceptual sensitivity and response criterion measures. AVH patients obtained fewer hits and had lower perceptual sensitivity to detect famous voices than healthy controls. The NAVH group did not differ from Controls and AVH in any measure. There were no differences between groups in false alarm rate or response criterion. The results indicate that AVH patients are impaired at voice identity recognition due to decreased sensitivity, which may result in inner speech misidentification.

5.2 Introduction

Auditory verbal hallucinations (AVH) are one of the most striking symptoms of schizophrenia, affecting up to 70% of schizophrenia patients during the course of the illness (Bentall, 1990). Despite the vast amount of research carried out in the area, the mechanisms of formation of AVH are still poorly understood. A prominent phenomenological feature of AVH is the perception of voices which have a specific identity (Stephane et al., 2003). Moreover, the voices appear to be generated by a person other than the self and often have the acoustical features (such as pitch, tempo, amplitude and even accent) of a particular individual, different to the hearer's own (Jones & Fernyhough, 2007). The voice identity specificity of AVH challenged theories claiming a misattribution of inner speech as the foundations of AVH, as these theories still need to account for the mechanism by which inner speech conveys acoustic properties different from the hearer's own voice. The missing link could be an additional voice identity recognition deficit in patients with AVH, as difficulties in recognizing one's own voice could contribute to the attribution of self-generated material to an external source.

In line with inner speech approaches, some research has intended to address whether AVH result from impaired self monitoring of verbal material (Johns et al., 2001). In the study by Johns and colleagues (2001), AVH patients and controls were asked to read words aloud while wearing headphones which transmitted the vocal input back to the participant. In some of the trials the transmission of the speech was distorted (acoustic features such as pitch were modified). In other trials, participants heard someone else's voice instead of their own as they spoke. Control trials were also included in which participants could hear their voice without any

modification. Immediately after saying the words, participants identified the source of the voice they heard as own, alien or unsure via a key press. The results showed that patients with positive symptoms, compared to patients without positive symptoms and healthy controls, misidentify their own voice as that of someone else when presented during the distorted feedback trials, even when they had the option to answer "unsure" about the source of speech (Johns et al., 2001). This finding was interpreted by the authors in terms of abnormal self monitoring. However, there might be alternative explanations of this misidentification of self-generated speech.

For example, Allen and colleagues tested AVH patients with distorted recordings of adjectives spoken in their own or other person's voice (Allen et al., 2004). As this paradigm did not involve generating verbal material by the subjects at the moment of testing, the task could be performed without the use of verbal self-monitoring. The authors propose that the misattribution was due to an externalising bias when processing unusual perceptual information. Finally, the authors dismissed the idea that the misattribution was due to a general deficit of voice discrimination, arguing that AVH patients tend to attribute the utterances to external sources instead of choosing the options "their own voice" or "unsure". However, it is important to consider that one's voice sounds different when it is heard from a recording (Békésy, 1949); thus the existence of a voice discrimination deficits cannot be ruled out by Allen's paradigm (Allen et al., 2004). During self-generated speech, sound reaches the inner ear by way of two separate paths. Air-conducted sound is transmitted from the surrounding environment through the external auditory canal, eardrum and middle ear to the cochlea. Bone-conducted sound reaches the cochlea directly through the tissues of the head. The voice heard during self-generated speech is

perceived by the combination of sound carried along both paths, resulting in a deeper and more resonant sound. Instead, listening to an external sound is performed through the air conducted path solely. Thus, in Allen's study the misattribution bias might be explained by taking into account that the participant's own voice reproduced by an external device sounds different in comparison to self speech production. In fact, the loss of spectral information due to the applied pitch distortion results in stimuli which are harder to recognize. Moreover, the recorded voice is not transmitted by bone-conduction, resulting in a mismatch between the internal representation of one's own voice and perception of the recorded voice. The mentioned loss of acoustical information of the recorded stimuli, in addition to the voice identity recognition deficit in AVH would result in the misattribution of the stimuli as alien. In other words, as the recorded voice sounds strange, it may result in a bias towards attributing the voice to an external source. While controls might be able to compensate for the mismatch of their voices with the recorded stimuli, the task might become critical for those with a voice discrimination deficit. Thus, applying a voice identity recognition paradigm in AVH patients free from the bone-conducted pathway confound, would show whether there is a genuine voice misidentification deficit in this population.

A recent study has assessed voice recognition using a novel paradigm (Zhang et al., 2008). This study presented AVH patients with personally familiar voices and with voices of strangers. Participants had to decide whether the voices were familiar or unfamiliar. This study found impairment in voice recognition in AVH patients in comparison to non-AVH patients and healthy controls. Unfortunately, the authors did not test whether the patients could recognize the

identity of the speaker, instead they only assessed familiarity. Thus, it is not entirely clear whether a voice identity recognition difficulty (phonagnosia) is present in this group of patients or if alternatively the results indicate a voice familiarity problem. In fact, a study showed that patients with damage in the right inferior lateral parietal lobe have difficulties in recognizing familiar voices but are unimpaired in voice discrimination abilities, while the opposite pattern was shown in patients with right temporal lobe damage, suggesting that these two domains are neuronally dissociated (Vanlancker, Cummings, Kreiman, & Dobkin, 1988). Additionally, because signal detection measures were not calculated, it is difficult to infer from Zhang et al. (2008) whether the differences in performance between AVH patients and controls are due to a decreased sensitivity to detect familiar over unfamiliar voices or to a bias towards a conservative approach at solving the task , such as a tendency to classify any voice as unfamiliar (Snodgrass & Corwin, 1988). If AVH patients cannot distinguish the identity of the speaker by the tone of voice, it may be the case that even their own voice may not be recognized, resulting in the source of the voice being perceived as alien.

It is the aim of the present study to establish whether schizophrenia patients with AVH have an impairment in voice identity recognition by using an established paradigm in the assessment of voice identity recall as well as of phonagnosia (Damjanovic & Hanley, 2007; Garrido et al., 2009; Hanley & Damjanovic, 2009). It is hypothesized that AVH patients will identify fewer famous voices in comparison to non-AVH patients and healthy controls. Moreover, in AVH patients, a tendency to find the voices rather unfamiliar independently whether they belonged to known or unknown speakers is predicted.

5.3 Methods

Participants

Twenty-five individuals (20 males) who met the DSM-IV-TR criteria (American Psychiatric Association, 2000) for schizophrenia were recruited from several outpatient clinics from Northumberland, Tyne and Wear NHS Foundation Trust and Tees, Esk and Wear Valleys NHS Foundation Trust. The psychiatric diagnosis was confirmed by an independent psychiatrist. All patients were taking antipsychotic drugs (haloperidol, risperidone, olanzapine, aripiprazole or clozapine). Exclusion criteria for patients were multiple diagnoses such as the presence of co-morbidities with axis I disorders of the DSM, or existence of a neurological condition.

Additionally, thirteen healthy participants (8 males) were recruited via advertisement in the local post office. They were screened for history of psychiatric illness, head injury, epilepsy, and drug use. All patients and controls were native English speakers and were permanent residents in the UK. After receiving a detailed description of the study, written informed consent was obtained from each participant. The study was approved by the regional NHS ethics committee and Durham University Ethics Advisory Committee. Subjects received £30 for participating in the study.

At the beginning of the testing session, all participants completed a hearing screening and were assessed with the National Adult Reading Test (NART; (Nelson & Willson, 1991), a test devised to estimate premorbid intellectual performance (IQ) with a high test–retest reliability (Morrison, Sharkey, Allardyce, Kelly, &

McCreadie, 2000). The screening revealed normal hearing in all participants. The results of the NART and the Edinburgh handedness questionnaire are shown in table 5. There were no differences in handedness and verbal IQ between groups (H=.528, ns and H=.900, ns respectively).

Table 5: Demographic features and hearing test. Characteristics of patient groups and controls. Schizophrenia patients with hallucinations (AVH), schizophrenia patients without hallucinations (non-AVH) and healthy controls. Group Means and Standard Deviations are reported. The three groups did not differ significantly in education, age or verbal IQ. Absolute hearing threshold expressed in dB SPL.

demographic data	**Groups**			
	AVH (N=13; 10 male)	Non-AVH (N=12; 10 male)	Controls (N=13; 9 males)	Analysis (Kruskal-Wallis test)
Age (years)	41.73 (2.62)	37.83 (2.87)	42.69 (3.09)	(H =1.47) (p=0.489)
NART/Verbal IQ	112.14 (1.79)	110.42 (1.74)	112.69 (1.24)	(H=0.900) (p=0.648)
Hearing test	50.00 (1.24)	55.00 (1.82)	50.00 (1.16)	(H=0.639)(p=.735)

Assessment of psychopathology

Interviews were conducted by a qualified clinical psychologist, using a semi-structured diagnostic interview Comprehensive Assessment of Symptoms and History (CASH) (Andreasen, Flaum, & Arndt, 1992). This interview includes the Scale for the Assessment of Positive Symptoms (SAPS, with 34 items measured on an ordinal scale ranging from 0 [absent] to 5 [severe]) (Andreasen, 1984b) and the Scale for the Assessment of Negative Symptoms (SANS, with 21 items) (Andreasen, 1984a). Details about the scales can be found in Table 1. Twelve patients who were not currently experiencing hallucinations (as defined by a score of 1 or below in SAPS hallucination global score) were allocated to the non-hallucinators group (NAVH). Patients who reported hallucinations (scoring at least 3 on the SAPS hallucinations global score) were allocated to the hallucinators group (AVH). None

of the patients scored between 1 and 3 in this scale. The AVH group subsequently completed the auditory hallucination subscale corresponding to the PSYRATS (Haddock, McCarron, Tarrier, & Faragher, 1999). This subscale consists of 11 items measuring frequency, duration, severity and intensity of distress caused by hallucinations, as well as the controllability, loudness, location, negative content, degree of negative content, beliefs about origin of voices and disruption they cause in daily life. A five-point ordinal scale is used to rate symptom scores (0-4).

The three groups did not differ significantly in education, age or verbal IQ. However, there were differences between the NAVH and AVH group in their mean SAPS global scores for delusion (U = 26.5, p <.005) (see table 6).

Table 6: Symptomatology. AVH (*n*=13) who were not experiencing hallucinations during the testing session, as defined by a score of one or below in the SAPS hallucinations global score. Non-AVH (NAVH)(n=13) scored at least 3 in the SAPS hallucinations global score. Group Means and Standard Deviations are reported. Between-groups comparisons analyzed with non-parametric Mann–Whitney *U* test, *df* = 37

symptoms rating	**Groups**		
	AVH (N=13)(10 male)	Non-AVH (N=12)(10 male)	Analysis (Mann-Whitney U test)
Duration of illness	13.73 (2.22)	15.17 (2.01)	(U = 86)(P=0.844)
SANS_ total	10.88 (1.51)	10.04 (0.84)	(U=77)(p=0.525)
Affective flattening	1.53 (0.25)	1.71 (0.4)	(U=86)(P=0.862)
Alogia	1.43 (0.27)	1.79 (0.47)	(U=81)(p=0.655)
Avolition	2.27 (0.27)	2.29 (0.44)	(U=89)(p=0.98)
Anhedonia	2.57 (0.25)	2.17 (0.42)	(U=78)(p=0.552)
Attention	2.6 (0.3)	2.92 (0.43)	(U=70)(p=0.330)
SAPS total	9.9 (0.69)	4.92 (0.87)	(U=21)(P=0.001)*****
Hallucinations	3.83 (0.24)	0.96 (0.23)	(U=1.5)(P<0.001)******
Delusions	2.97 (0.33)	1.17 (0.31)	(U=26.5)(P=0.002)*****
Bizarre behaviour	1.47 (0.25)	1.42 (0.31)	(U=86)(P=0.840)
Positive formal thought	1.38 (0.33)	1.63 (0.3)	(U=79.5)(P=0.598)
PSYRATS (hallucination subscale)	25.87 (1.73)	1.75 (1.18)	(U=2)(p=<0.001)******

Materials

The stimuli used in the experimental task were taken from the Damjanovic Famous Voices dataset (Damjanovic & Hanley, 2007). They consisted of 96 voice samples (half female) taken from television interviews, each of which lasted approximately seven seconds. Half of the samples were from famous people, such as David Beckham and Margaret Thatcher, and the other half from non-famous people. The tone of voice of each sample was neutral and its content did not offer any clues regarding the identity or occupation of the speaker. The audio tracks from both famous and non-famous voices were presented via headphones. The whole task lasted approximately twenty-five minutes. Participants could opt to take breaks during the task. Stimuli were presented using E-prime.

Procedure

The participants entered a quiet testing room and were seated in front of a laptop. At the beginning of the session, a short hearing test was conducted using monaural white-noise bursts (duration 1s), presented via headphones with various sound-pressure levels (steps of 10 dB). More details about the hearing test can be found in (Hirnstein, Hausmann, & Lewald, 2007). This test revealed normal performance in all participants. Then, participants performed the experimental task. They were informed that they would be presented with a sequence of voices, some of which belonged to people who were well known in the British media, whilst others belonged to non-famous individuals. For each voice, participants had to decide by key press whether the voice belonged to somebody famous or not. If they considered the voice to be famous, they were asked to orally classify their responses as one of

the following categories: remember (R), know (K) or guess (G). Along the lines of previous research (Maylor, 1995), the experimenter asked the participants, if they could recall the person's name (R), or if they could associate some facts about the person (such as profession or recall of an event the celebrity has taken part in) even though they could not recall the person's name (K). If participants thought they have heard the voice before but they could not recall anything about the person, this was registered by the experimenter as G. Including G as a possible response is necessary to separate confident from unconfident K responses by giving the subject the possibility to answer G in case of very low confidence. However, G responses should be interpreted with caution as they do not strictly reflect recognition (Gardiner, Ramponi, & Richardson-Klavehn, 2002).

Analysis

For the statistical analyses only those trials where participants expressed confidence in their memory (correct R and K responses) were included. All responses were converted to proportions by dividing the number of responses per category (R and K) by 48 (total of famous voices). The total number of hits was defined as the sum of R and K responses for each participant and the total number of false alarms comprised all non-famous voices categorized as famous. G responses were excluded from the analyses. Because the measure of hits only provides information about accuracy without assessing the capacity to discriminate between target and distracters, an additional analysis was performed. We calculated signal detection theory (STD) measures in order to quantify the participant's ability to discriminate famous from non-famous voices and their response criterion in answering.

The sensitivity index A' was calculated as:

$$A' = \frac{1}{2} + \frac{(HIT - FA)(1 + HIT - FA)}{4HIT(1 - FA)}$$

(Donaldson, 1996). The A' index, an indicator of discrimination of famous over non-famous voices can vary between 0 and 1, with values of 1 indicating perfect discrimination of famous from non-famous voices and values around 0.5 indicating chance performance. A', a measure of sensitivity, represents the strength of the internal representation of a signal. In a noisy environment the detectability of a signal can increase by enhancing the signal itself or by filtering the distracters. It has been suggested that the measure A' is more appropriate than d' because A' is less affected by the differences in response criteria essential to the model (see also (Donaldson, 1993; Gardiner & Gregg, 1997; Gardiner et al., 2002).

The response criterion B''D was calculated as:

$$B''_D = \frac{(1 - HIT)(1 - FA) - (HIT)(FA)}{(1 - HIT)(1 - FA) + (HIT)(FA)}$$

This index provides a measure of the participant's tendency to categorize each voice as famous. The values for B"D can vary between -1, a response bias of classifying a voice as famous (lax criterion), and +1, a response bias of classifying a voice as non-famous (strict criterion), with a value of zero representing a neutral bias. Dependent variables (hits, false alarms, A', and B''D) were subjected to separate one-way ANOVA's with Group (AVH patients, non-AVH patients, and healthy controls) as between subjects variable.

5.4 Results

Hit rates

The overall means for recognition performance are displayed in Table 3. There was a significant effect of Group on hit rates, $F(2,35) = 3.96$, $p < .05$. Post-hoc Sidak corrected comparisons tests showed that the AVH group (M =18.9, SD = .90) obtained significantly lower hit rates than healthy controls (M =34.94, SD =1.45, $p < .05$). The non-AVH group (M =22.4, SD =2.02) did not differ from the AVH group or from the controls (see Fig. 8).

False alarms

The corresponding one-way ANOVA did not reveal a significant group effect in false alarms rates, $F(2,35) = .77$, p = ns.

Sensitivity index A'

The one-way ANOVA showed a significant group effect for A' ($F(2,35) = 3.44$, $p < .05$). Post-hoc Sidak corrected comparisons showed a significant difference between the AVH group (M = 0.69, SD = 0.16,) and healthy controls (M =.81, SD = .05, $p < .05$). The non-AVH group (M = .73, SD = .10) did not differ from the AVH group and healthy controls (see Fig. 8).

Response criterion B''D

For the response criterion B"D, the one-way ANOVA revealed that the effect of Group was not significant ($F(2,35) = 1.09$, ns).

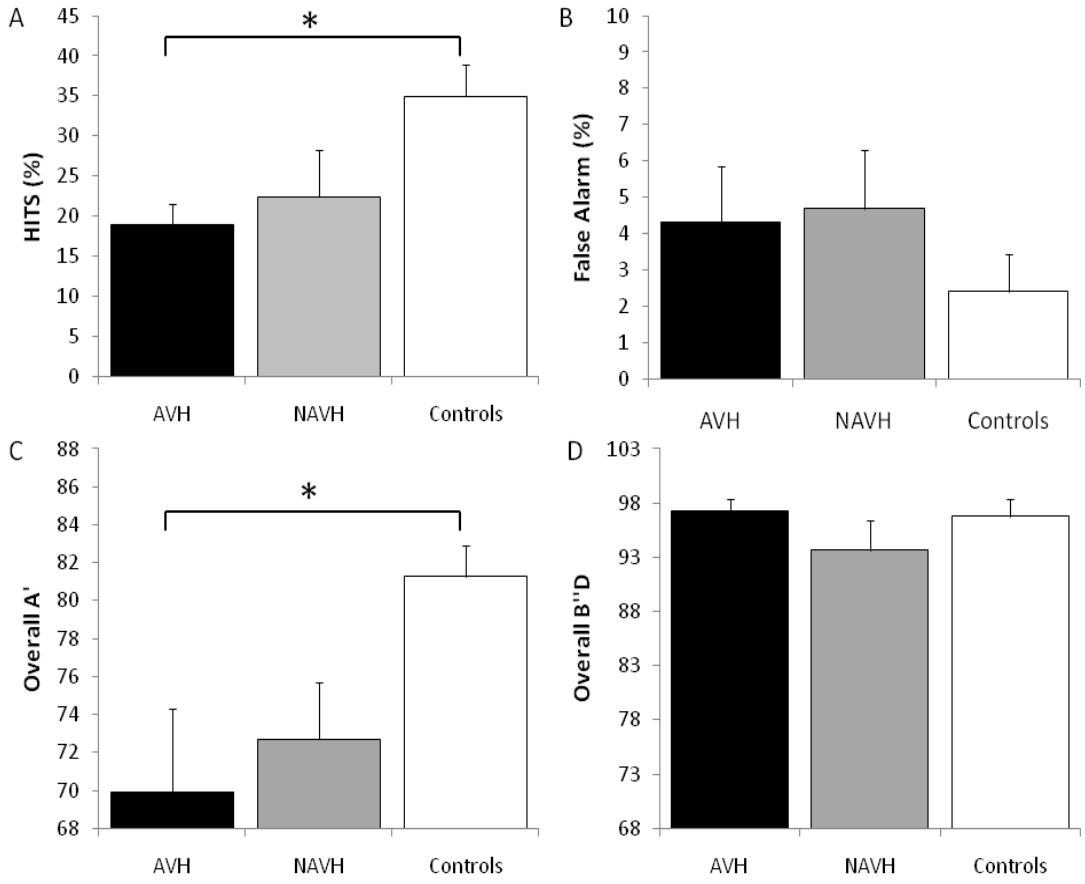

Figure 8. Performance and SDT measures for the schizophrenia patients with hallucinations (AVH), healthy controls (controls) and schizophrenia patients without hallucinations (NAVH) groups. Error bars depict standard errors, *p <.005. A) Mean hits (famous voices correctly identified). B) Mean false alarms (non-famous voices incorrectly identified). C) Mean response criterion B''D. D) Mean discrimination index A'.

5.5 Discussion

The results of the healthy control group replicate previous studies using the same stimuli in terms of hit rates and sensitivity index A' for famous voices (Damjanovic & Hanley, 2007; Garrido et al., 2009). The healthy control group did not significantly differ from the NAVH group in hit rates and sensitivity index A'. However, AVH patients revealed significantly lower hit rates and sensitivity index A' than healthy controls. AVH patients also revealed numerically, but not significantly lower hit rates and A' than NAVH patients. Given that NAVH patients

also did not differ significantly from healthy controls in both measurements, these finding suggests that the performance of NAVH patients is between those of AVH patients and healthy controls. The SDT measure B''D and FA rates did not reveal any differences between groups suggesting that all groups applied a similar response criterion while performing the task. Overall, group differences in hit rates and A' indicated that especially AVH patients have difficulties in recognizing famous voices.

A potential link between hallucinations and defective voice identity recognition in patients with schizophrenia is supported by Zhang et al., (2008) who assessed voice familiarity in schizophrenia patients with and without AVH (for more details on this study see introduction). This study found that in a task in which participants were presented the recorded voices of their own relatives and friends, and they have to classify them as familiar or non-familiar, only the AVH group was impaired, not NAVH patients and healthy controls. Taken together, evidence for a voice familiarity deficit in AVH patients solely, in addition to a voice identity recognition deficit in AVH patients compared to healthy controls, point towards a specific link between voice recognition impairment and hallucinations.

The cognitive mechanism by which voice recognition is impaired in AVH patients remains unknown. However, a possible interpretation is that AVH patients are less accurate than healthy controls at extracting a signal (a particular famous voice) from a noisy environment (small individual differences of acoustical features in voices). Thus, this interpretation implies that the AVH group is characterised by low signal to noise internal representations of individual voices. It has been demonstrated that in schizophrenia there are high levels of internal noise, which

results in reduced signal to noise ratio leading to poor prosodic stimuli recognition (Bach, Buxtorf, Grandjean, & Strik, 2008). Both, emotional prosody and voice identity recognition relies in part on pitch analysis (Perrot et al., 2007). In fact, difficulties in low level auditory perception such as pitch discrimination have also been found in schizophrenia (Leitman et al., 2005; Leitman et al., 2007; Matsumoto, et al., 2006). Altogether, it seems that there are intrinsic perceptual limitations of voice processing in schizophrenia that may relate to hallucinations in particular.

Contrary to the present study, previous research found differences in response criterion between groups. For example, Allen et al. (2004) found a bias towards externalizing self-generated speech schizophrenia patients, especially for those with positive symptoms. The lack of group differences in the response bias in the present study might be partly due to the different methodological approach. The present study used famous voices instead of recorded self-generated speech. As mentioned in the introduction, one's own voice whilst spoken sounds deeper and more resonant than when displayed from an external apparatus due to the loss of bone-conducted pathway. The externalization bias found in the previous study may have been due to the loss of the bone-conducted pathway when listening to recordings of self-generated speech, and AVH patients might have been particularly sensitive to such pathway. Given that there is a qualitative difference between the perception of one's own voice whilst spoken than when reproduced from an external apparatus, it is possible that the recorded voices in Allen et al. study (2004) were not perceived as one's own, resulting in a bias towards externalizing. However, the bias would not be expected when listening to externally presented famous voices, as the bone-conducted pathway confound would be absent.

Regardless of whether group differences in the response criterion exist or not, which in Allen et al.'s study probably occurred as a result of the relatively unfamiliar own voice, this study and the present study found group differences with particularly low performance of voice recognition in patients with schizophrenia with positive symptoms. A general deficit of voice identity recognition linked to hallucinations is proposed.

Extending the theory of misattribution of inner speech

Misattribution of inner speech has been considered as a main contributing factor towards the formation of AVH (Frith, 1987; Johns et al., 2001; Shea, et al., 2007). This theory assumes that inner speech is not recognized as self-generated and thus is perceived as hallucination. The present study provides perhaps the first empirical evidence that inner speech is misattributed because there is a general voice identity recognition deficit.

At first glance it might be difficult to understand why AVH and NAVH groups did not differ significantly in voice identity recognition. If a voice identity recognition deficit is critical for the development of hallucinations both groups should actually differ in discrimination rates. However, it should be noted that both groups did significantly differ with respect to the frequency and severity in delusions (and the total scale of positive symptoms) (see table 1). In fact, it has been reported that that the majority of AVH patients suffer from delusions (Liddle, 1992). Thus, the results of the present study suggest that the voice identity recognition deficit needs to be accompanied by delusions in order to develop hallucinations. Since hallucinations can be linked to poor voice identity recognition, such difficulty may

result in the perception of a voice whose identity is unknown. Concurrent delusional thinking might then result in attributing a specific identity to the unknown voice, which results in the hallucinations being identity specific.

The present study cannot fully rule out that the reported group differences are affected by different media experiences and exposure to famous voices as used in the present study. However, this explanation is rather unlikely because voices were taken from popular celebrities of the British media and all participants reported watching TV and listening to the radio regularly (see supplementary material for a list of famous voices used in the study). Group-specific general auditory deficits can also not account for group differences in voice recognition performance, as normal hearing was confirmed in all participants.

To summarize, the present study showed that AVH patients performed particularly low in voice identity recognition, suggesting that a deficit in voice identity recognition is a general feature in schizophrenia patient suffering from hallucinations. However, AVH patients did not differ from NAVH patients in this capacity, strongly suggesting that voice identity recognition is necessary but not sufficient to develop auditory verbal hallucinations. The results rather suggest that additional delusional think is a critical event which has to co-occur in order to produce hallucinations. The present study provides the first empirical evidence in order to support the theory of misattribution of inner speech, extending and specifying the cognitive mechanisms of hallucinations.

Chapter VI: Emotional prosody modulates attention in schizophrenia patients with hallucinations

In the previous chapter the link between a voice identity recognition deficit and AVH in schizophrenia was discussed. Such a deficit explains one of the phenomenological features of AVH; that is the fact that the voices heard in hallucinations have specific identities that are different from the hearer. However, the AVH model presented in chapter I mentioned other phenomenological feature of hallucinations; that is its emotional (often negative) salience. Chapter I also proposes that the reason for such phenomenological feature may be based on an attentional bias towards emotional features of auditory objects. More precisely, emotional salience of auditory objects may catch attention favouring the processing of these inputs in detriment of non-salient objects. Although attentional capture by emotionally salience stimuli also occurs in healthy individuals, they may be able to modulate this involuntary attentional capture by using top-down control. Instead, it might be the case that AVH patients lack of top-down control and are recurrently captured by salient auditory objects (Hugdahl, 2009). However, the link between AVH and implicit attention towards salient emotional prosody has not been tested directly.

The present chapter will assess the effects of emotional prosody distractors during an irrelevant task in schizophrenia patients, using a validated dichotic listening paradigm. The dichotic listening paradigm would also provide evidence regarding the lateralization of prosody in schizophrenia. It may be speculated that EPC deficits in schizophrenia patients are due to cerebral abnormalities underlying such processes.

6.1 Abstract

Recent findings demonstrate that emotional prosody attracts attention involuntarily. Healthy participants overcome irrelevant salient stimuli by using attentional control. Attentional control is impaired in schizophrenia and there is evidence suggesting that such impairment may be a mechanism contributing to the formation of hallucinations. The present study aims to investigate the effects that implicit emotional prosody exerts on involuntary attention in schizophrenia. Fifteen schizophrenia patients with hallucinations, 12 schizophrenia patients without hallucinations and 16 healthy controls completed a dichotic listening paradigm, in which an emotional vocal outburst was paired with a neutral vocalization spoken in male and female voices. Participants were asked to attend to either the left or right ear and report only the sex of the speaker. Schizophrenia patients without hallucination and healthy controls revealed longer response times when emotional prosody was presented in the attended right ear than in the attended left ear. Responses were faster when the attended ear and the ear which received the emotional prosody differed. Schizophrenia patients with hallucinations, however, did not benefit from attending to one ear when the emotion was presented to the other ear. The findings suggest that in healthy controls and schizophrenia patients without hallucinations, emotional prosody processing is lateralized to the right hemisphere as the right ear/left hemisphere present difficulties handling both, sex discrimination and emotional prosody processing. Schizophrenia patients with hallucinations on the other hand seem to have difficulties in top-down control, as reflected by selective attention, when interfered by automatic emotional processing (bottom-up input), as a result of aberrant lateralization in emotional prosody.

6.2 Introduction

The tone of voice, known as emotional prosody, conveys information about the emotional state and intention of others. Emotional prosody is a non-verbal component of language which allows us to encode and decode the feelings expressed in speech by means of variation in pitch amplitude and tempo. The ability to process emotional prosody is necessary to establish effective communications and also to maintain successful social interactions (Beatty et al., 2003). In fact, it has been demonstrated that in some neuropsychiatric disorders, deficits in comprehending emotional prosodic cues correlate with difficulties maintaining effective social interactions (Beatty et al., 2003). For example, emotional prosody decoding is known to be impaired in schizophrenia (Borod et al., 1989; Borod et al., 1990), leading to poor outcome in social functioning (Kee, Green, Mintz, & Brekke, 2003). However, emotional prosody decoding has usually been assessed *explicitly*, that is, participants were explicitly paying attention to stimuli they were asked to classify (Buchanan, et al., 2000; Kotz et al., 2003). Implicit processing of emotional prosody on the other hand has received little attention. Since in most everyday situations, humans are not specifically asked to focus on emotional prosody during conversations, investigating implicit prosody processing is also interesting in terms of ecological validity. Only a few studies investigated implicit processing of emotional prosody in healthy participants (Aue, Cuny, Sander, & Grandjean, 2010; Sander et al., 2005).

Implicit emotional processing occurs because emotionally enhanced stimuli capture attention with and without the involvement of conscious or voluntary processes (e.g., (Bradley & Lang, 2000; Lipp & Waters, 2007); see (Vuilleumier,

2005), for a review). This involuntary and automatic attraction of attention towards certain objects is a typical bottom-up process, controlled by external stimulus presentation, and not under participants' control (Pessoa, Kastner, & Ungerleider, 2002). Its counterpart, voluntary attention, refers to the direction of attention under control of the individual, for example when attention is being focussed on the basis of instructions, such as when participants are asked to make an explicit judgment of the emotional tone conveyed in prosodic stimuli. This is a typical top-down controlled mechanism, requiring the participants' attentional effort (Pattyn, Neyt, Henderickx, & Soetens, 2008).

A few studies on prosody have manipulated involuntary and voluntary attention, to investigate the influence that emotional salience exerts over top-down control (Aue et al., 2010; Grandjean et al., 2005; Sander, et al., 2005). These studies assessed implicit processing of emotional prosody using similar tasks, in which participants listened to male and female voices in angry or neutral tone dichotically presented via headphones to the participants' left and right ears in a balanced order. The participants were asked to attend to either the left or right ear. In order to assess only implicit emotional prosody, participants were asked to make judgements about the speaker's sex of the attended ear regardless of the emotional tone conveyed by the voice. The results revealed that angry prosody attracts attention and provokes behavioural and physiological changes such as variations in skin conductance and body temperature, even when it is not voluntarily attended (Aue et al., 2010). Functional magnetic resonance imaging (fMRI) evidence showed that angry, relative to neutral stimuli, generally led to increased activation in the right superior temporal sulcus regardless of the ear the angry stimulus was presented to and which ear was

attended, suggesting again that even when emotional prosody is unattended, it is still processed (Grandjean et al., 2005). Moreover, reaction times in the sex discrimination task were faster when an angry voice was presented to the left ear, but only when the left ear had to be attended to, demonstrating that when the emotional stimuli converges on the side of the attended ear, such a distractor is more difficult to ignore (Sander et al., 2005). This finding can be interpreted as a superiority of the left ear/right hemisphere in processing auditory emotional stimuli (Grimshaw, Kwasny, Covell, & Johnson, 2003), meaning that even if prosody is task-irrelevant, the dominant hemisphere might automatically process the emotional input. Right-hemisphere dominance has been repeatedly reported for emotional prosody processing in healthy controls (Ross et al., 1997; Wildgruber et al., 2005).

Behavioural studies suggest that schizophrenia patients show deficits in explicit judgements of emotional prosody (Bach et al., 2009; Bozikas et al., 2006; Hoekert et al., 2007; Rossell & Boundy, 2005). However, less is known about the implicit perception of emotional prosody in schizophrenia (when voluntary attention towards the stimuli is not required). A recent study compared schizophrenia patients and controls in two conditions. In the explicit prosody processing condition, participants had to attend to semantically neutral words and judge the emotion conveyed by the tone of voice. In the implicit prosody condition, participants listened to words with emotional meanings spoken in different prosodic tones and they had to judge the semantic content, which was either congruent or incongruent with respect to the prosodic tone (Roux et al., 2010). The authors found that schizophrenia patients revealed higher error rates during implicit prosody processing in incongruent trials than healthy controls. In response times, however, both

schizophrenia patients and healthy controls obtained slower responses for incongruent trials than congruent trials. According to the authors, the results demonstrate that schizophrenic patients are still sensitive to implicit processing of prosody, however, they refrain from interpreting the increased error rate in patients because controls had a floor effect on this measure (Roux et al., 2010). The increase in error rates may indicate a lack of top-down control (voluntary attention) in the presence of bottom-up salience due to the prosodic features of the stimuli which capture the attention interfering with the semantic task.

It is possible that schizophrenia patients with particularly pronounced positive symptoms such as hallucinations have problems with processing implicit emotional prosody. For example, it has been proposed that the capture of involuntary attention by negative emotional stimuli provides a mechanism for the formation, maintenance and exacerbation of positive symptoms (Mohanty et al., 2008). Indeed, schizophrenia patients with positive symptoms as well as highly schizotypal individuals with positive trait symptoms showed deficits in implicit emotional processing (Mohanty et al., 2008; Phillips, Senior, & David, 2000). For instance, individuals with positive schizotypy showed poor attentional performance in an emotional Stroop task in which they had to judge the ink colour of emotional words (Mohanty et al., 2008). Importantly, none of these studies investigated how emotional stimuli capture attention using prosodic cues. Moreover, since hallucinations in schizophrenia tend to be more often auditory than visual (Mueser et al., 1990), one may expect that implicit processing of prosody is particularly disturbed in schizophrenia patients with auditory hallucinations. A study investigating implicit emotional prosody in schizophrenic patients (Roux et al.,

2010), did, however, not differentiate between patients with prominent positive and negative symptoms. The present study therefore addresses this issue by comparing hallucinators (AVH) with non-hallucinators (NAVH).

The difficulties in implicit emotional prosody may result from aberrant brain organization. Specifically, AVH patients may display an atypical lateralization of emotional prosody, which then leads to impaired processing of emotional prosody. In fact, atypical lateralization has been considered one of the main causes subjacent to cognitive deficits in schizophrenia (Crow, 1990). Lateralization of prosody processing in schizophrenia has been extensively debated in the literature (Bach et al., 2009; Mitchell & Crow, 2005; Mitchell et al., 2004). However, results in this topic are controversial. Mitchell and colleagues (2004) claim that lateralization might be reversed in schizophrenia patients with hallucinations who displayed greater involvement of the left temporal lobe in prosody processing, whereas Bach and colleagues (2009) found enhanced right lateralization to prosody in these patients. However, the findings from these studies are limited by small samples sizes. Atypical lateralization of linguistic components of speech (i.e., right hemisphere superiority) has also been described in schizophrenia (Sommer et al., 2001; Weiss et al., 2006; Woodruff et al., 1997; Youn, Park, Kim, Kim, & Kwon, 2003), and this abnormal lateralization of linguistic (Sommer et al., 2001; Weiss et al., 2006; Woodruff et al., 1997; Youn et al., 2003) as well as prosodic elements of speech (Mitchell & Crow, 2005; Mitchell et al., 2004) seems to be related to the presence of hallucination. Perhaps, a reduction of the typical right hemisphere lateralization for emotional prosody is particularly present in patients who suffer from hallucinations.

The present study aims to investigate the effects that emotional prosody exerts on involuntary attention using a dichotic listening paradigm which allows us to investigate lateralization of emotional prosody. The task of the present study is adopted from Grandjean et al. (2005), Sander et al. (2005) and Aue et al. (2010). One main goal was to determine whether emotional salience can modulate the allocation of attention. It is hypothesized that emotionally neutral trials will be more efficiently processed than emotional trials. Moreover, the present study aimed to test whether healthy controls and NAVH patients successfully use voluntary (top-down) attention control to ameliorate the impact of bottom-up salience when the emotional prosody distracters are presented on the non-attended side. Specifically, it is predicted that when emotional stimuli are presented to the unattended ear, implicit processing of emotional prosody will not interfere with the sex-discrimination task in both (healthy and NAVH) control groups. Interference is likely, however, if both processes converge, particularly when stimuli are presented to the right ear corresponding to the (non-dominant) left hemisphere. Given that lateralization of emotional prosody is assumed to be abnormal in schizophrenia, and particularly related to hallucinations, it is hypothesized that contrary to the NAVH and healthy control groups, AVH patients would not benefit from attending towards the ear opposite to the emotional stimuli.

6.3 Methods

Participants

Twenty seven (21 males) individuals who met the DSM-IV criteria (American Psychiatric Association, 2000) for schizophrenia were recruited from several

outpatient clinics from Northumberland, Tyne and Wear NHS foundation trust and Tees, Esk and Wear Valleys *NHS* Foundation *Trust*. The psychiatric diagnosis was confirmed by an independent psychiatrist. All patients were native speakers of British English. All patients were taking atypical antipsychotic drugs (haloperidol, risperidone, olanzapine, aripiprazole or clozapine). Exclusion criteria for patients were multiple diagnoses such as the presence of co-morbidities with axis I disorders of the DSM, or existence of neurological condition. Additionally, sixteen healthy participants (11 males) were recruited via advertisement in the local post office. They were screened for history of psychiatric illness, head injury, epilepsy, and drug use.

Assessment of psychopathology

Interviews were conducted by a qualified clinical psychologist, using a semi-structured diagnostic interview Comprehensive Assessment of Symptoms and History (CASH) (Andreasen, et al., 1992). This interview includes the Scale for the Assessment of Positive Symptoms (SAPS, with 34 items measured on an ordinal scale ranging from 0 [absent] to 5 [severe]) (Andreasen, 1984b) and the Scale for the Assessment of Negative Symptoms (SANS, with 21 items) (Andreasen, 1984a). Twelve patients who were not currently experiencing hallucinations (as defined by a score of 1 or below in SAPS hallucination global score) were allocated to the non-hallucinators group (non-auditory verbal hallucinators or NAVH). Patients who reported hallucinations (scoring at least 3 on the SAPS hallucinations global score) were allocated to the hallucinators group (auditory verbal hallucinators or AVH). None of the patients scored between 1 and 3 in this scale. The AVH group subsequently completed the auditory hallucination subscale corresponding to the

PSYRATS (Haddock et al., 1999). This subscale consists of 11 items measuring frequency, duration, severity and intensity of distress caused by hallucinations, as well as the controllability, loudness, location, negative content, degree of negative content, beliefs about origin of voices and disruption they cause in daily life. A five-point ordinal scale is used to rate symptom scores (0-4). Finally, all subjects were assessed with the National Adult Reading Test (NART) (Nelson & Willson, 1991), a test devised to estimate premorbid intellectual performance (IQ) with a high test–retest reliability in schizophrenia (Morrison et al., 2000). The means and standard deviation of all measures in all three groups and all (sub-) scales can be found in Table 7.

There were no differences in age ($F(2,42) = 0.72$, ns), education as measured by highest qualification achieved (1 = primary school, 2 = secondary school, 3 = 0 levels, 4 = A levels, 5 = university degree; $F_{(2, 42)} = 0.35$, ns), estimated verbal IQ ($F_{(2,42)} = 0.51$, ns) and handedness measured by the Edinburgh Handedness Questionnaire ($F_{(2,42)} = 1.10$, ns) between controls, AVH and NAVH patients. Between the patient groups, Mann-Whitney U-test revealed significant differences in the total score ($U = 21.00$, $Z = -3.37$, $p < .05$), the hallucinations global subscale ($U = 1.50$, $Z = -4.37$, $p < .001$), and delusions global subscale ($U = 26.50$, $Z = -3.18$, $p < .05$) of the SAPS, confirming the initial division into AVH and NAVH patients. None of the other psychopathology measures revealed significant differences between patient groups (all $U < 70.50$, ns).

Table 7: Clinical, neuropsychological and demographic characteristics of the three groups. SANS and SAPS: Scale for the Assessment of Negative and positive Symptoms respectively. PSYRATS: Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales. *****p* **< .001; ****p*** **< .05.**

Measures	Non-AVH Mean(SD)	AVH mean (SD)	Controls Mean (SD)
N	**12** (2 women)	**15** (4 women)	**16** (5 women)
Premorbid verbal IQ	**110.42** (1.74)	**112.14** (1.79)	**112.69** (1.24)
handedness scale	**77.50** (16.43)	**67.33** (15.96)	**93.13** (5.06)
Age	**37.83** (2.87)	**41.73** (2.62)	**42.69** (3.09)
highest qualification achieved	**2.25** (0.37)	**2.07** (0.28)	**2.44** (0.33)
duration of illness	**15.17** (2.01)	**15.33** (2.22)	
SANS total	**10.88** (1.51)	**10.40** (0.84)	
affective flattening	**1.71** (0.40)	**1.53** (0.25)	
Alogia	**1.79** (0.47)	**1.43** (0.27)	
Avolition	**2.29** (0.44)	**2.27** (0.27)	
Anhedonia	**2.17** (0.42)	**2.57** (0.25)	
Attention	**2.92** (0.43)	**2.60** (0.30)	
SAPS total	**4.92** (0.87)	**9.90** (0.69)**	
Hallucinations	**0.96** (0.23)	**3.83** (0.24)**	
Delusions	**1.17** (0.31)	**2.97** (0.33)*	
bizarre behaviour	**1.42** (0.31)	**1.47** (0.25)	
positive formal thought	**1.38** (0.33)	**1.63** (0.30)	
PSYRATS (hallucination subscale)	-	**25.87** (1.73)	

Materials and procedure

Stimuli

The stimulus used in the experimental task were taken from the "Montreal Affective Voices" data set (Belin, Fillion-Bilodeau, & Gosselin, 2008), which constitutes an

auditory equivalent of affective faces set by Ekman & Friesen (1986). In the present study thirty nonverbal affect bursts consisting in non-linguistic vocal sounds corresponding to emotions of anger, happiness and neutral expression recorded by ten different actors (five male and five females) were used. Since interaural intensity differences and length differences are known confounds in dichotic listening tasks (Hugdahl, Westerhausen, Alho, Medvedev, & Hamalainen, 2008), voice stimuli were edited to make them equal in duration (900 ms) and amplitude (80 DB) using Audacity audio editing software (http://audacity.sourceforge.net). Male and female speakers were equally distributed across conditions (see Belin et al., 2008, for more details about the stimuli validation).

Procedure

After receiving a complete description of the study, written informed consent was obtained from each participant. The study was approved by the regional ethics committee from the NHS and Durham University Ethics Advisory Committee. All participants received £30 for participating in the study. Hearing screenings at the beginning of the testing session revealed that all participants had normal hearing.

Two voice stimuli were simultaneously presented, one to each ear, resulting in five possible combinations: Left angry/Right neutral, Left neutral/Right angry, Left happy/Right neutral, Left neutral/Right happy, and Left neutral/Right neutral. Every trial consisted of one female and one male voice. Participants were instructed to selectively attend to the voice presented to either the left or right ear and to decide on the sex of the speaker pronouncing the vocal sound in the attended ear. Participants indicated their forced-choice response by means of a response box. Half

of the participants used the left response key to indicate a female voice and the right to indicate a male voice. The response keys were counterbalanced across participants. A total of 144 trials were presented to each participant excluding five practice trials at the beginning of the experiment. In one block (72 trials), participants focused on voices presented to their right ear; in another block (72 trials), they attended to voices presented to their left ear. Block order was counterbalanced across participants. Participants listened to the voice stimuli and gave their response during the inter-trial interval of 3000 ms. The task took eight minutes excluding a short break between experimental blocks. Accuracy and reaction times (RT) were analyzed. For the RT analysis, incorrect trials and outliers were excluded (every individual trial from a single subject that was above or below three SD from the median). The aim of the present study was to investigate whether involuntary attention modulates emotional prosody in general. Given that previous research has demonstrated that involuntary attention is similarly modulated by negative as well as positive emotional prosody (Grandjean, Sander, Lucas, Scherer, & Vuilleumier, 2008; Roux et al., 2010), trials with happy and angry stimuli were collapsed for further analyses.

6.4 Results

In the statistical analysis degrees of freedom were epsilon-corrected (Greenhouse-Geisser) when sphericity was violated (Mauchly). Post hoc t-tests were alpha-adjusted (Bonferroni) for multiple comparisons.

Accuracy (%)

Participants' overall accuracy (across all groups and conditions) in the sex discrimination task was significantly above chance (M = 73 % ± SD = .05, $t(40)$ = 20.1, $p < .001$, one-sample t-test, chance level was 50%). Accuracy (%) were subjected to a 2 x 3 x 3 mixed ANOVA with Ear attended (left, right) and Trial-type (neutral binaurally (baseline), emotion on the left ear LEE), emotion on the right ear (REE)) as within-participants factors and Group (AVH, NAVH, healthy controls) as between-participants factor. The ANOVA revealed a main effect of Trial type ($F_{(2,76)}$ = 3.34, $p < .05$). Alpha adjusted pairwise comparisons indicate that the accuracy during baseline (75.00 % ± .01) is higher than during the REE condition (72.00 % ± .01, $p < .05$). The LEE condition (73.40 % ± .01) did not differ from REE and baseline. The main effect of Ear attended did not approach significance ($F_{(1,38)} = 0.42$, ns). Moreover, there was a significant interaction between Ear attended and Trial type ($F_{(2,76)}$ = 7.94, $p < .05$). Alpha-adjusted post hoc t-tests revealed that in the REE condition, participants obtained lower accuracies when the right ear was attended (69.40 % ± 1.27) than when the left ear was attended (75.00 % ±1.07, $t(41) = -3.09$, $p < .05$), suggesting that if the emotion is presented to the attended ear, there is an increased difficulty to ignore the affective distractor while performing the sex discrimination task (see Figure 9a). No other post hoc test approached significance (all $t \leq 2.45$, ns). No other main effect or interaction approached significance, (all $F \leq 1.49$, ns).

Reaction times (RTs)

Identical to the accuracy data analysis, a 2 x 3 x 3 ANOVA was calculated for reaction times (RTs). The ANOVA revealed a significant main effect of Ear attended ($F_{(1,38)} = 9.33$, $p < .05$). Attending to the right ear (1269 ms ± 39) resulted in longer RTs than attending to the left ear (1205 ms ± 43). Moreover, there was significant main effect of Trial type ($F_{(2,76)} = 9.23$, $p < .001$). Alpha-adjusted pairwise comparisons showed that participants responded more quickly to the baseline condition (1187 ms ± 42) than to the LEE condition (1260 ms ± 38; $p < .05$) and REE condition (1256 ms ± 40; $p < .05$), again suggesting that emotional prosody interferes with the sex discrimination task. The LEE condition did not differ from REE. Additionally, the main effect of Group was significant ($F_{(1,38)} = 5.45$, $p < .05$). Alpha-adjusted pairwise comparisons revealed the only group difference between AVH (1354 ms ± 65) and healthy controls, the latter obtaining faster RTs (1062 ms ± 65, $p < .05$). Other group differences did not approach significance.

Again, there was a significant interaction between Ear attended and Trial type ($F_{(2,76)} = 8.44$, $p < .001$). Alpha-adjusted post hoc t-tests revealed a significant effect of Ear attended for the REE condition ($t(40) = 4.35$, $p < .001$), indicating slower responses when the right ear was attended. The Ear attended did neither differ for LEE nor the baseline condition (see Figure 9b).

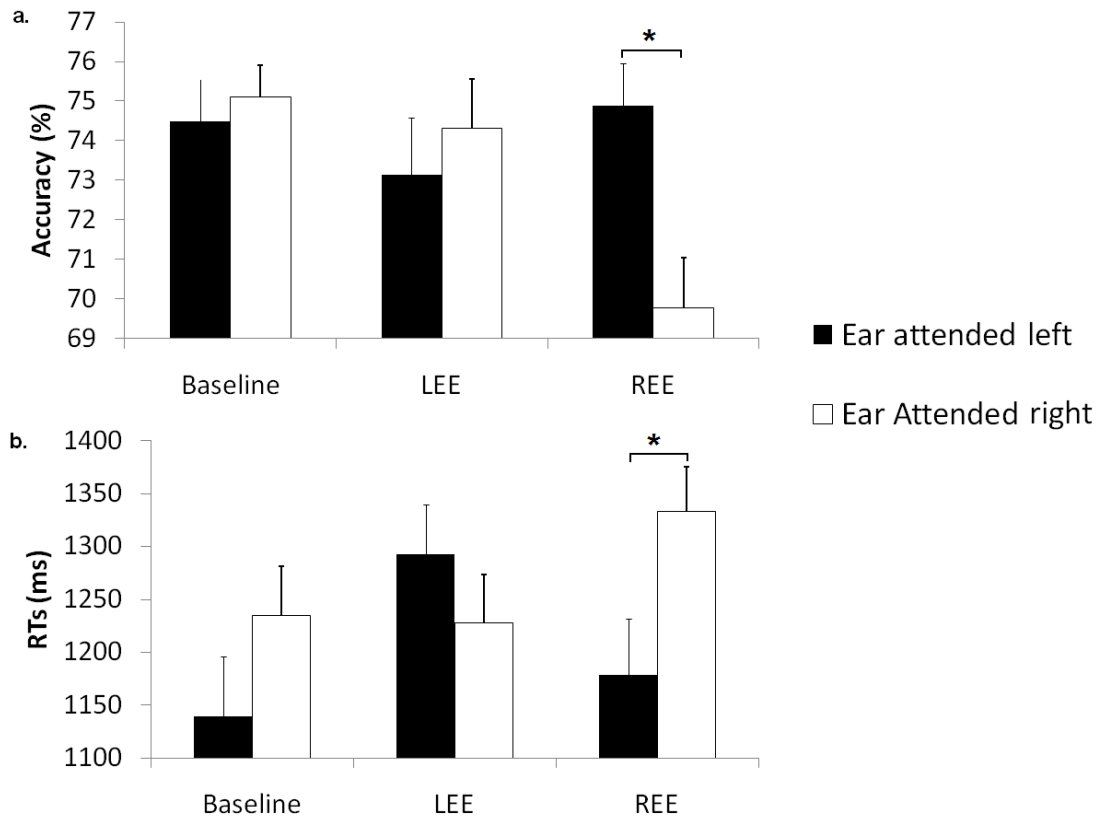

Figure 9. Mean accuracies (%) (Figure 9a) and mean reaction times (ms) (Figure 9b) and standard errors in the sex discrimination task across groups (schizophrenia patients with hallucinations (AVH), without hallucinations (NAVH) and healthy controls) in three different conditions (baseline, trials with emotion presented to the left ear (LEE) and trials with emotions presented to the right ear (REE). Black columns represent trials in which the left ear was attended. White columns represent right ear attended trials.

Finally, there was a significant three-way interaction ($F_{(4,76)} = 6.74$, $p < .001$). To investigate the nature of this effect, three separate ANOVAs (one for each group) were performed. There was no main effect of Ear attended in any of the groups (all $F \leq 3.82$, ns). The main effect of Trial type was only significant in healthy controls ($F_{(2,28)} = 5.50$, $p < .05$), and in the AVH group ($F_{(2,28)} = 8.74$, $p < .001$). In the control group, alpha-adjusted post hoc test showed faster responses in the baseline condition (1026 ms ± 68) than in the REE condition (1085 ms ± 68, $t(14) = 3.09$, $p < .05$) but the baseline condition did not differ from LEE (1076 ms ± 68, $t(14) = 1.36$, ns). The REE and LEE conditions did also not differ significantly (see Figure 10). In the

AVH group, the baseline condition (1309 ms ± 73) revealed faster RTs than REE (1395 ms ± 71, $t(14) = 3.01$, $p < .05$) and LEE (1384 ms ± 77, $t(14) = 3.52$, $p < .05$). REE did not differ from LEE ($t(14) = 1.62$, ns). The main effect of Trial type in the NAVH group was not significant ($F_{(2,20)} = 0.37$, ns). Most importantly, the interaction between Ear attended and Trial type was significant in healthy controls ($F_{(2,28)} = 28.07$, $p < .001$) and NAVH ($F_{(2,20)} = 15.45$, $p < .001$) but not in the AVH group ($F_{(2,28)} = 2.98$, ns).

To explain the significant interaction between Ear attended and Trial type in both control groups (healthy controls and NAVH), alpha-adjusted post hoc t-tests showed that RTs were slower when the Ear attended and the emotion coincided on the right (healthy controls: $t(14) = 3.70$ $p < .05$; NAVH: $t(10) = 3.34$, $p < .05$), not when the left ear was attended in the REE condition (all $t \leq 2.55$, ns). The comparison between attended ears was not significant in LEE and the baseline condition (all $t \leq 0.85$, ns). To compare convergent (Ear attended and side of emotion co-occurring at the same hemispace) and divergent condition (Ear attended and the side of emotion occurring at opposite hemispace), additional alpha-adjusted post hoc t-tests were performed. The analyses revealed that both control groups showed a significant difference between LEE and REE, when the left ear was attended (healthy controls $t(14) = 4.88$, $p < .001$; NAVH ($t(11) = -2.72$, $p < .05$), not when the right ear was attended.

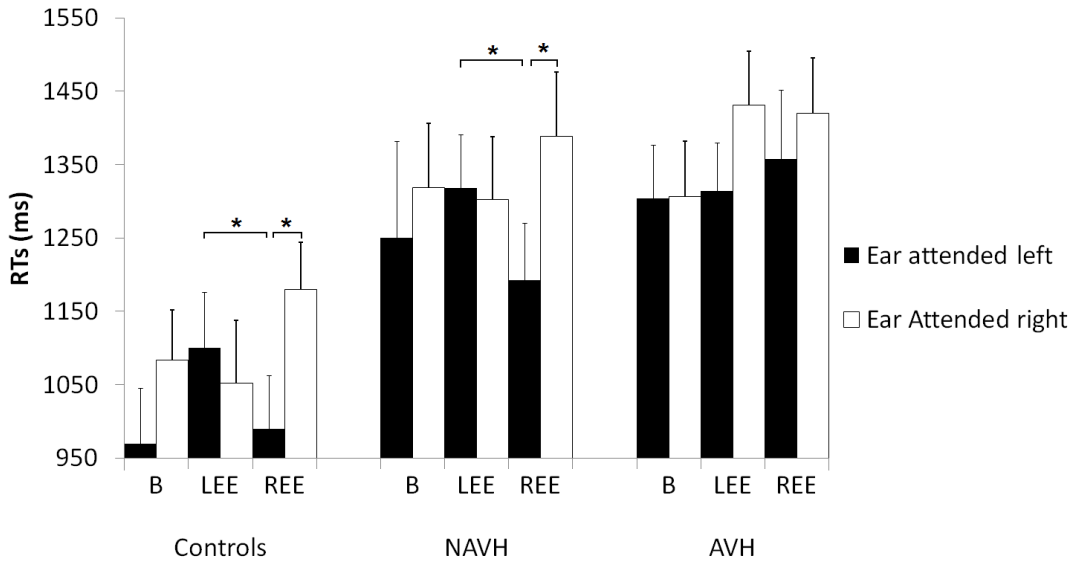

Figure 10 Mean reaction times (ms) and standard errors *(SE)* in milliseconds in the sex discrimination tasks for the different groups for the different groups (schizophrenia patients with hallucinations (AVH), without hallucinations (NAVH) and healthy controls) in three different conditions (baseline, trials with emotion presented to the left ear (LEE) and trials with emotions presented to the right ear (REE). Black columns represent trials in which the left ear was attended. White columns represent right ear attended trials. Post hoc tests are alpha-adjusted for multiple comparisons, $p < .05$.

6.5 Discussion

This study was conducted to reveal how involuntary attention is modulated by emotional prosody in AVH schizophrenia patients. By including two control groups (healthy controls and NAVH schizophrenia patients), this study controls for possible confounding effects of schizophrenia in general, demonstrating for the first time that abnormal modulation of involuntary attention by emotional prosody is associated with AVH in particular.

In general, the present study revealed faster responses in the sex discrimination task when participants' attended the left ear, in comparison to the right ear. The left ear advantage following dichotic presentation is in agreement with

previous studies reporting that the right hemisphere is superior in sex discrimination by vocal cues (Lattner, Meyer, & Friederici, 2005; Sokhi, Hunter, Wilkinson, & Woodruff, 2005). Moreover, the emotionally neutral baseline condition resulted in faster responses in comparison to emotional trials. This is in agreement to previous findings by Aue et al. (2010) suggesting that attention is involuntarily captured by auditory emotional stimuli, and thus, resulting in interference with the explicit sex discrimination task.

The present study also revealed lower accuracies and prolonged response times when the emotion was presented to the attended right ear. At first glance, this effect might be interpreted as interference caused by the convergence of the attended ear and the emotion presented in the same hemispace, which may occur because one cerebral hemisphere does not have the capacity to process both tasks in parallel. This interference seems to be less pronounced when the attended ear and ear of emotion differ. Participants were also less distracted by emotionally neutral voices. These findings seem to be in line with two previous studies (Grandjean et al., 2005; Sander, et al., 2005), suggesting that it is more difficult to allocate voluntary attention when the emotional distractor drives involuntary attention toward irrelevant but salient features of the target stimuli. Top-down control seems to be less demanding when the emotional distractor is presented to the unattended ear, and thus can be more easily ignored. Sander and colleagues (2005) suggest that the consequence of orienting attention to the side of the emotional stimulus imposes the evaluation of a salient social signal. Orienting attention away from the task-relevant hemispace probably makes the task-irrelevant stimulus and its automatic processing less salient, and consequently causes less interference.

It is noteworthy that this effect was only found when the emotion was presented in the right ear, suggesting that the (non-dominant) left hemisphere is particularly sensitive to this specific effect. This finding contradicts to Sanders and colleagues (2005) who observed the same convergence effect in the right hemisphere. Although it is difficult to understand where the difference between studies comes from, it should be noted that Sanders et al (2005) did not find a significant behavioural asymmetry in the sex-discrimination task. Moreover, they cited a study which found bilateral activation in the superior temporal sulcus for the task, whereas the results of the present study strongly suggest right hemisphere superiority in sex discrimination by vocal cues. The right hemisphere superiority for sex discrimination (Lattner, et al., 2005; Sokhi, et al., 2005) and emotional prosody (Friederici & Alter, 2004; Ross, 2010; Ross & Mesulam, 1979; Ross, et al., 1997) may explain the longer response times when the right ear was attended corresponding to non-dominant left hemisphere, as an interhemispheric transfer from the non-dominant left to the dominant right hemisphere is required. In other words, interference effects due to a convergent processing of information within one hemisphere together with the additional costs caused by the interhemispheric transfer of information to the task-related superior hemisphere might explain the slower response times in the present study.

Interference differences in schizophrenia patients with hallucinations

The main finding of the present study is that the above described interference effect differed between groups. Both control groups (healthy controls and NAVH patients) revealed a decrease in performance when the emotional voice was presented to the attended right ear. This effect was not present in schizophrenia

patients with hallucinations. The AVH group did not show any convergent or divergent effects between the attentional focus and the side where emotional prosody was presented. This finding suggests that (a) AVH patients had difficulties controlling their selective attention during the presence of an emotional distractor and (b) AVH patients are not typically right lateralized for emotional prosody.

The first suggestion implies that, although all groups are driven by the emotional salience of distractor, both control groups seem to overcome such difficulty when the salient stimulus is divergent to the ear attended. This phenomenon is particularly present, if the emotional distractor does not converge with the dominant attended ear. This is not the case for the AVH group, in which top-down control cannot benefit from the divergent effect. In a previous study, phonological processing was assessed with a dichotic listening paradigm that included conditions in which participants focused attention to either the left or right ear as well as a non-forced condition in which no specific attentional instruction was given (Loberg, Jorgensen, & Hugdahl, 2004). Two groups of schizophrenia patients, one with ongoing AVH and the other with a history of AVH were compared to healthy controls. The results revealed that difficulties in top-down control, as measured in the forced attention condition, occurred in both schizophrenia groups, suggesting that diminished top-down control is particularly present in AVH (Loberg et al., 2004). The present study, however, did not show a general top-down deficit in AVH patients in comparison to NAVH patients, as both groups did not differ from another in overall performance. Instead, the present finding is partially in line with the literature supporting that even if voluntary attention is preserved in schizophrenia, it may fail in the presence of salient bottom-up competitors (Hahn et

al., 2010). Indeed, the AVH group obtained shorter RTs in the baseline condition than in the emotion conditions, demonstrating that while AVH patients in general responded more slowly to the sex discrimination task (including neutral baseline trials), they particularly struggled when emotional prosody stimuli were present. Thus, difficulty in top-down control arises in the AVH group particularly when emotional bottom-up divergent competitors interfere with the explicit sex discrimination task. It seems that emotional salience of prosodic stimuli captures attention of AVH patients in detriment of executive control, regardless whether both interfering stimuli are in the focus of attention.

The second suggestion that AVH patients are not typically right lateralized for emotional prosody is based on the observation that AVH patients do not benefit from trials in which the attended ear does not correspond with the ear in which emotional prosody is presented. AVH patients might have a more bilateral brain representation of emotional prosody, thus they do not benefit from having the emotional distractor in the putative non-dominant side while focusing attention on the dominant side (which may help to dismiss the distractor in normal population). Although, the present study included two potentially right lateralized processes (emotional prosody and sex discrimination), we are inclined to believe that the atypical lateralization rather refers to emotional prosody because of the significant main effect of Ear attended. The left ear advantage was present for all groups. General abnormal lateralization in schizophrenia has been previously reported (Hirano et al., 2008; Levitan, Ward, & Catts, 1999; Ngan et al., 2003; Ribolsi et al., 2009; Sommer et al., 2001; van Hoof, 2002; van Veelen et al., 2011; Woodruff et al., 1997; Youn et al., 2003), and some of these studies suggest that such abnormal

lateralization is particularly present in schizophrenia patients with hallucinations (Levitan et al., 1999; Sommer et al., 2001; Woodruff et al., 1997; Youn et al., 2003).

It should be noted that all patients of the present study took antipsychotic medication. Thus, it is difficult to fully ascertain whether the differences found between patient groups are not influenced by potential confounding effects of antipsychotic drugs. In fact, it has been proposed that the ability to ignore irrelevant stimuli, which is compromised in schizophrenia (Leumann, Feldon, Vollenweider, & Ludewig, 2002), might improve with the intake of medication, particularly atypical antipsychotic drugs such as clozapine and olanzapine (Simosky, Freedman, & Stevens, 2008). However, the groups did not differ in premorbid IQ, education, years of illness, age and handedness; thus, alternative explanations for the findings, such as cognitive and demographic differences between groups, are less likely.

Finally, the present findings have implications for neuropsychological models of hallucinations. There is evidence of an attentional bias toward emotionally stimuli in schizophrenia patients (Green, Williams, & Davidson, 2001; Waters, Badcock, & Maybery, 2006). However, the present results extend the finding of abnormal attentional bias toward prosodic emotional stimuli presented in the non-dominant side and divergent to the focus of attention to the subgroups of patients with AVH. It is possible that emotional salience of prosodic stimuli captures attention even when attention should be oriented away from the distractor, resulting in emotional inputs which access processing in detriment of non-emotional competitors. This attentional bias towards irrelevant prosodic emotional stimuli might therefore impair the ability to focus on appropriate aspects of the surrounding acoustical environment (Javitt, 2009), leading to abnormal perceptual processing.

Hallucinations are an example of aberrant perceptual process, which usually convey emotional salience. Such a bottom-up saliency may diminish top-down control. It has been proposed that AVH may involve a shift of attention towards the perceived voices (Hugdahl et al., 2007). Clinical interventions to ameliorate this symptom should focus on directing attention away from the voices.

6.6 Conclusion

The present study provides evidence that healthy controls as well as schizophrenia patient can implement top-down control. Moreover, healthy controls and NAVH patients showed that bottom-up salience (as manipulated by emotional distractors) modulates top-down control, and disadvantages caused by distractors can be overcome depending on whether the distractors are presented to the non-dominant hemisphere and are divergent from the focus of attention. On the contrary, in AVH patients, selective attention cannot overcome divergent emotional distractors presented to the putative non-dominant side, suggesting that bottom-up salience may capture attention revealing inefficient top-down control. Such finding points towards a failure of the typical lateralization process as the core mechanism underlying abnormal modulation of attention by emotional prosody in AVH patients.

GENERAL DISCUSSION

CHAPTER VII

7.1 Foreword

The aim of this thesis was to create a new neuropsychological model of auditory verbal hallucinations (AVH) based on testable hypotheses and fitting the phenomenological characteristic of these abnormal perceptions, such as the subjective nature of feelings and experiences associated with hallucinations (see chapter I). Before introducing and evaluating the validity of the new model, it is necessary to summarize previous theoretical frameworks with their valuable contributions, and also with their shortcomings.

One of the most prominent cognitive models of AVH proposed that source monitoring defects are a rich theoretical framework for the study of hallucinations (Laroi & Woodward, 2007). According to this perspective, there is some resemblance between inner speech and AVH. It is proposed that even considering that inner speech is ubiquitous to all populations, and it may not be abnormal in itself in patients with hallucinations (AVH patients), these patients may not experience inner speech as self-generated. Instead, AVH patients may attribute inner speech to an external source and location. In other words, the cognitive capacity responsible for monitoring inner speech, verbal self-monitoring, might be impaired in AVH patients resulting in this "internal dialogue" being apprehended as "alien"(Jones & Fernyhough, 2007).

The misattribution of inner speech model for AVH has some weaknesses. For example, this theoretical framework is unable to explain why patients do not consider their AVH as vivid inner thoughts; instead they describe them as

perception. In fact many patients claim to “hear the voices with their ears” (patients, personal communication, August, 2010). A second weakness of the misattribution of inner speech model consists in its failure to account for non-verbal auditory hallucinations (such as noises) or even for AVH which have prosodic non-human characteristics. An example of the last one was found by Hunter who described a patient hearing a voice with a mechanistic timbre and an “engine sound” (Hunter & Woodruff, 2004). Moreover, inner speech theories have had difficulties explaining why, if AVH are misattributed inner thoughts, they have person specific characteristics different from the individual experiencing them (see chapter V). It could be argued that inner thoughts may not only involve inner speech, in the sense of imagined speech without observable output (talking to oneself with one's voice), but also auditory mental imagery such as the recall of a previously heard familiar voice. Even if hallucinations were produced by auditory mental imagery, this still does not explain why AVH patients produce such a variety of verbal imagery embodied in non-self speakers instead of their own-inner voice (Hoffman, Varanko, Gilmore, & Mishara, 2008) . In fact, patients are very unlikely to report that the voice they hear during hallucinations is their own.

In order to prevail over some of the mentioned shortcomings of the cognitive model, a different approach was postulated. This approach, called the neurological and perceptual model, intended to provide a bridge between the phenomenology and the neural correlates of AVH. The neurological and perceptual model starts from the idea that auditory hallucinations might arise from a functional disturbance of the auditory system. This idea is supported by a study of intrinsic functional connectivity in healthy participants showing spontaneous fluctuation of

activity in a speech sensitive auditory network which was interpreted as a connate vulnerability of the brain towards the production of aberrant auditory activity (Hunter et al., 2006). In line with the cited study, corollary discharges in the auditory system as the neural cause of perceptions in the absence of external stimulation (Penfield & Perot, 1963) has also been taken into account by this theoretical approach.

The pioneer work of Penfield showed that epilepsy patients can experienced visual and auditory hallucinations as a result of ictal brain activity in sensory areas (Penfield & Perot, 1963). When the patient's temporal lobe was electrically stimulated, half of the patients reported to experience hallucinations (Penfield & Perot, 1963) (i.e., a patient clamming to *hear* a melody, as distinct from thinking of it or remembering it). The authors claimed that if stimulation in the associative auditory cortex is applied, a sound might be heard; but if the electrode is moved a few millimeters into neighboring cortex, a voice might be perceived instead of a sound. Finally, it was proposed that a specific external stimulus might precipitate a seizure attack in a vulnerable system. For example, the aura of a seizure might start with a tingling on the skin which could be precipitated by touching that area (Penfield & Perot, 1963). This phenomenon is known as psychical precipitation.

The neurological perceptual model also emphasizes the role of aberrant perceptual processing as a main contributing factor of hallucination. Several low level auditory processing defects have been found in schizophrenia patients, such as impairment in pitch perception (Leitman et al., 2007) sound duration (Todd, Michie, & Jablensky, 2003) and sound localization (Erkwoh et al., 2006; Guterman & Klein, 1991; Heilbrun, Blum, & Haas, 1983). Such low level sensorial deficits determine

the ability to perform more complex perceptual and cognitive operations (Javitt, 2009). Bottom-up processes have a central role in the formation of hallucinations (Javitt, 2009).

Summarizing the previous theoretical frameworks of AVH, the neurological and perceptual model explains why hallucinations can be spontaneously generated and perceived as objective and real, and still takes into account external stimuli that can expedite the phenomenon in a prone system. However, it does not specify why AVH are mostly negative in content neither why AVH usually have the voice characteristics of a particular individual. In other words, very often the voices perceived in AVH have the vocal qualities of somebody the patient knows, such a relative or a friend. The need to create a new model of AVH comes from the unattainability of the current theoretical frameworks to explain the full constellation of phenomenological aspects conveyed in hallucinations. The new model should address why some AVH can be heard in the absence of external stimuli and still being experienced as generated by an external force; why AVH often have "person specific properties", such as the particular pitch, accent and timbre of recognizable identities; why AVH have mostly negative emotional content; and why patients described them as perceptions instead of thoughts. Finally, new models should also be empirically testable as well as physiologically plausible.

7.2 The neural underpinnings of EPC

This thesis presents an empirically testable new model of AVH, in which a link between AVH and emotional prosody comprehension (EPC) deficits (chapter I) is of central importance. AVH and EPC are associated at the phenomenological,

behavioural and neural level. At the neural level, AVH and EPC appear to be supported by partly overlapped networks of which the superior temporal gyrus (particularly on the right hemisphere) is a nodal structure for both processes (chapter I). While there is a vast amount of literature demonstrating the neural correlates of AVH, and a meta-analysis has shown consistencies between studies (Jardri et al., 2011) the literature about the brain representation of EPC is more controversial. In fact, there are some contradictory findings regarding both the involvement of the prefrontal cortex (PFC) and the lateralization of the temporal lobe in EPC tasks. Before introducing the clinical data, it is necessary to clarify how EPC is represented in the healthy brain, as findings from the literature are not conclusive. A neuroimaging and a neurodisruptive study (chapters II and III) were designed to reveal the neural underpinning of EPC.

The fMRI study presented in chapter II aimed to clarify the role of the PFC (particularly on its medial regions) as well as the somatosensory cortex in EPC. The need for this study comes from the fact that even when (bilateral) STG activation during EPC is a consistent finding in fMRI studies, the role of PFC and somatosensory cortex (as part of the mirror neuron system) has varied greatly between studies. Chapter II explores the possibility that perhaps different types of emotion (simple vs. complex) may explain inconsistencies in the involvement of the prefrontal regions, as the current EPC literature overlooked the effects that different types of emotion may have in the recruitment of differential neural networks. The results in chapter II showed that while bilateral STG involvement was consistent across the task regardless of the type of emotion processed, the recruitment of mPFC and somatosensory cortex was linked to the decoding of complex emotions in

particular, suggesting that the different demands in empathizing and mentalizing skills are crucial for the recruitment of both regions in EPC (chapter II). In other words, the mPFC and somatosensory cortex have been shown to be particularly important for EPC for social and complex emotion, because they rely on inferring other people's mental states to a greater extent in comparison to simple emotion (for more details see chapters II and IV). This finding is in line with studies suggesting that the somatosensory cortex as well as the mPFC form part of a specific social cognition network and are recruited when mentalizing about emotion in particular (Adolphs, 2001; Hooker et al., 2008). The present finding (chapter II) also differs from a study suggesting that the mPFC is involved in EPC for simple emotion (Beaucousin et al., 2007). It is possible that the emotional (semantic) meaning of the stimuli used by Beaucousin and colleagues may have caused the recruitment of the mPFC in the basic EPC task. However, when the stimuli contained minimal semantic content, the simple task would not rely on the mPFC (chapter II). As has been explained, findings regarding the role of mPFC in EPC are not as robust the consistent latero-temporal involvement in such domain. Although the fMRI study revealed the importance of the lateral temporal lobe in the basic EPC task, it is difficult to disentangling whether the highlighted structure was only necessary or also sufficient to perform EPC (for a discussion on this topic see chapter III).

Even though the involvement of the lateral posterior temporal lobe in EPC is a robust finding in neuroscience literature, and the STG in particular has a role in this domain regardless of the complexity of emotion (chapter II), the lateralization of EPC within the temporal lobe is still a matter of debate (see chapter III). While neuroimaging studies proposed bilateral temporal lobe involvement in EPC (chapter

II), lesion studies propose a right temporal lobe representation of this function (Ross & Mesulam, 1979; Ross & Monnot, 2008b, 2011; Ross et al., 1997) particularly when the stimuli do not contain semantic information (Friederici & Alter, 2004). It was demonstrated that, when measured with fMRI, bilateral activation is associated to EPC (chapter II), probably due to the fact that fMRI does not establish a relation of causality between a task and the brain network supporting it. Instead, the TMS study (chapter III) showed that only the right temporal lobe (specifically the right STG) is causally involved in EPC, which is in agreement with the cited lesion studies. The current finding (chapter III) also suggests that previous fMRI studies claiming a bilateral temporal representation of EPC (Grandjean et al., 2005; Kotz et al., 2003; Wildgruber et al., 2005) may have revealed areas associated with, but not essential to the EPC task. Taken the results of both, fMRI and TMS studies together, the STG (particularly on the right hemisphere) is pivotal to the basic EPC task. Such disclosure has implications for the proposed AVH model (chapter I), as one of the most consistent structural abnormalities in schizophrenia is reduction of gray matter in the left and right lateral temporal lobe (Honea, Crow, Passingham, & Mackay, 2005; Nenadic et al., 2010). Moreover, by examining difficulties in EPC at the behavioral level in the schizophrenia spectrum, it may be possible to infer which neural circuits are affected in this population.

7.3 EPC in the schizophrenia spectrum

As hypothesized in chapter I, it is proposed that emotional prosody might be a critical factor in the formation of hallucinations. By assessing the schizophrenia spectrum, it is possible to investigate the link between AVH and EPC deficits with non-clinical high schizotypy population. This approach has the advantage of

overcoming potential confounds associated to the illness such as the effect of medication, hospitalization and general cognitive decline. Evidence presented in chapter IV, showed that non-clinical schizotypal participants with high scores in AVH traits appear to have increased RTs in EPC for simple emotion in comparison to non-clinical schizotypal participants with low scores in AVH traits. Such findings may reflect a continuum of schizotypal participants and schizophrenia patients, in which the STG is compromised resulting in poor EPC. Moreover, the exploratory additional analysis of extreme groups, in which the group with high scores in AVH traits showed a difficulty in EPC for complex emotion in absence of difficulties in EPC for simple emotion in comparison to the low AVH traits group (chapter IV), might denote additional mPFC and somatosensory cortex abnormalities in the high AVH traits group. However, because of the exploratory nature of the study and the small sample size, the results should be interpreted with caution.

The deficit in EPC, which usually relies on the (right) STG (chapters II and III), was observed in the high AVH traits group (chapter IV). This is in agreement with the schizophrenia literature. According to the literature, along with the left medial temporal cortex, the left and right lateral temporal cortex, including the arcuate fasciculus (bilaterally), left Heschl's gyrus, left uncinate fasciculus, and right primary auditory cortex, show consistent volume reduction in schizophrenia patients (Honea et al., 2005). Moreover, two more recent investigations linked bilateral STG gray matter reduction specifically with AVH (Garcia-Marti et al., 2008; Nenadic et al., 2010) (for more information about these studies see chapter I). Thus, it is possible that volume reduction in the STG can be linked to the EPC deficits that patients with AVH suffer from. The role of the right STG in AVH and EPC has been

extensively discussed in the current model, and the recruitment of the right STG during hallucinatory activity has been interpreted as related to the prosodic features that hallucinations conveyed (chapter I). However, some structural MRI studies report a bilateral or even prominently left abnormality of this temporal region in patients with hallucinations (Seok et al., 2007). Importantly, since the right STG is associated with the prosodic features of AVH, the left STG, which comprises Wernicke's area on its posterior part, might be underlying the semantic components of AVH. In order to formulate a complete explanation of the role of the right and left STG in schizophrenia, it is necessary to assess the laterality of language functions in schizophrenia patients, and particularly on those who suffer from hallucinations.

7.4 Prosody related processes in schizophrenia patients with and without AVH

As has been demonstrated in chapter III, the right but not the left posterior STG is crucial for EPC in the healthy population; however, the lateralization of this domain in schizophrenia patients with AVH is less clear (see chapter I). While there is certain degree of consensus that there is a reduction of the typical left lateralization of semantic component of language in these patients (Sommer et al., 2001; van Veelen et al., 2011; Weiss et al., 2006) contradictory results were found in the assessment of EPC (Bach et al., 2009; Mitchell et al., 2004). Since chapters II and III only present evidence for EPC lateralization in healthy controls, some cues can be extracted from a paradigm evaluating modulation of attention by implicit prosody using dichotic listening (DL) in patients of schizophrenia with and without hallucinations (chapter VI). The paradigm consisted in an explicit voice gender

discrimination task, which relies on the processing of prosodic features such as pitch, using forced choice DL in which emotional prosody was conveyed by the voice presented in the right or left ear (for more details about the task see chapter VI). This study strongly suggested that while implicit emotional prosody is lateralized to the right hemisphere in healthy controls as well as in schizophrenia patient without hallucinations (NAVH), such a function is not typically lateralized in AVH patients (chapter VI). It has been proposed that hallucinations derive from a failure in the lateralization of language processes (Crow, 1990). If this is the case, it will be difficult to extricate the phenomenological correlates of lateralized neural responses to hallucination (i.e. whether a right STG activation during AVH corresponds to the prosodic or semantic aspects of the perceived voices). Both right and left STG present structural abnormalities in schizophrenia patients and both sides may be implicated in AVH. Thus, it can be expected that the cognitive modules underpinned by the STG, such as EPC, were affected.

Findings form behavioural studies of the current thesis demonstrated that, in addition to EPC abnormalities in AVH patients (chapter VI) and in a high schizotypal population with positive symptoms traits (chapter IV), schizophrenia patients with hallucinations also present difficulties in voice identity recognition (chapter V). A previous fMRI study showed that difficulties of AVH patients to recognize familiar voices are related to abnormalities in the neural network underlying voice identity recognition, particularly in the STG (Zhang et al., 2008) which is also involved in EPC (chapters II and III). Voice identity recognition is partially mediated by the same perceptual mechanism as EPC. However, common acoustical parameters might not be enough to produce the correlation in both

cognitive modules. Basic sensory processing deficits do not suffice to explain EPC deficits in schizophrenia (for a discussion on this topic see chapter I), and they are also very unlikely to explain voice identity recognition. The association between voice identity recognition and EPC may be due to the fact that voice identity is in part recognized by the emotional prosodic tone and style that characterizes the speaker (Adami, 2007). Furthermore, a nodal brain structure supporting both capacities may again be the STG.

So far, it has been explained that the STG, a brain structure active during hallucinations, is also implicated in EPC and voice identity recognition. Both cognitive capacities are conveyed by variations in pitch, amplitude and tempo (Spreckelmeyer, Kutas, Urbach, Altenmuller, & Munte, 2009). Thus, it is not surprising that voice identity recognition and EPC share some common neural mechanisms. In a model formulated by Belin and colleagues, it has been proposed that the right superior temporal sulcus (STS), which separates the STG from the MTG, is a nodal structure for the analysis of voice identity (Belin et al., 2004; Warren, Scott, Price, & Griffiths, 2006). Taken these results together, it is agreed that the posterior right STS is central for voice identity perception (Belin et al., 2004; Warren et al., 2006) as well as for EPC (chapters II and III), and this structure presents bilateral functional and structural abnormalities in schizophrenia, particularly in schizophrenia patients with hallucinations (Garcia-Marti et al., 2008; Honea et al., 2005; Nenadic et al., 2010). Again, there is evidence that voice identity recognition is impaired in AVH patients (chapter V); and a central structure in this domain (the right STG) is affected in these patients. Thus, the brain representation of voice identity recognition in AVH patients might differ from controls.

The study presented in chapter V does not assess the neural representation of voice identity recognition in AVH patients. However, some indirect evidence could be extracted from the study presented in chapter VI. That is, in the dichotic listening paradigm, AVH patients showed abnormalities when emotional prosody presented in an unattended non-dominant side, interfered with a sex discrimination task revealing an atypical pattern of lateralization (chapter VI). However, AVH patients showed the typical left ear advantage for the sex discrimination task in general, suggesting that, at least some features of identifying voices may not be atypically lateralized in schizophrenia patients with hallucinations (chapter VI). However, a previous fMRI study assessing voice familiarity in AVH patients may not be in line with our findings (Zhang et al., 2008). The cited study, which also revealed poor performance in AVH patients at recognizing familiar voices in comparison to controls and NAVH patients, found abnormal lateralization of this function in AVH patients. Indeed, since the controls activated the right middle frontal gyrus and the right STG after the subtraction of unfamiliar voices from the familiar ones, the AVH group demonstrated less cerebral activation in the right STG and increased BOLD response in the left STG in the same contrast (Zhang et al., 2008). It may be interpreted that the recruitment of the left STG for this right hemisphere task compensates for the deficiencies of the right STG in AVH patients. Thus, AVH patients are not typically lateralized for voice familiarity, which is supposed to be a right hemisphere function (Hailstone et al., 2010). The model of AVH presented in this thesis (chapter I) makes the assumption that AVH patients present difficulties at recognizing the identity of voices, which was demonstrated in chapter V, and such difficulties may

rely on structural and functional abnormalities of the STG, a structure supporting this cognitive function.

Voice identity recognition deficits can explain one of the main phenomenological features of hallucinations, namely the identity specificity of the perceived voices (see chapters I and V). It has been assumed that, in the process of the formation of AVH, auditory mental imagery (such as inner-speech) occurs in a first stage, following which such mental imageries are (mis)attributed to a source (Evans, McGuire, & David, 2000). Importantly, a few studies in AVH patients which used a paradigm whereby participants listened to recordings of their own or another person's voice and had to judge the source of the utterances, found a bias towards externalizing self-generated speech in this population in comparison to controls (Allen et al., 2004; Johns et al., 2001) (for a discussion about methodological issues related to measuring own voice recognition from recorded tapes see chapter V). Moreover, the externalizing bias occurred only when their self-generated material was distorted through pitch manipulation (Allen et al., 2004; Johns et al., 2001) suggesting that perhaps the acoustical features of the auditory mental imagery necessary to identify the source of the voice may be ambiguous and/ or degraded in AVH patients. The data presented in chapter V is in line with the idea that the misidentification of the voices in AVH exists because there is defective processing of voice identity, as AVH patients showed difficulties to recognize famous voices. If this is the case, there may be a perceptual deficit linked to hallucination, and an abnormal cognitive process (misattribution of the voices to a source) may be a secondary process (chapter V).

7.5 The formation of AVH

This thesis proposes a novel explanation regarding the identity of the voice in AVH (chapter V). The new model suggests that because AVH patients have difficulties in voice identity recognition (chapter V), the identity of the voices perceived during hallucinations might be ambiguous and thus difficult to recognize as self-generated. In addition to voice identity recognition difficulties, delusional thinking may attribute a particular identity to the ambiguous voice which then results in the AVH with the phenomenological characteristic that patients usually report (i.e. the voice of someone they know). In other words, it is hypothesized that delusions compensates for the faulty input of a stimuli (in this case a perceived voice whose identity is ambiguous), "filling in" the gaps by projecting an identity to the AVH. Interestingly, AVH and delusions generally co-occur in psychotic population (see table VI), and it has been proposed that AVH precede the emergence of delusions, the latter being a cognitive response to the former (Smeets et al., 2010). The experiment in chapter V showed that AVH patients had fewer hits and reduced sensitivity at recognizing the identity of the speakers in comparison to healthy controls, although the groups did not differ in response criteria. Such findings point towards the existence of a defect in voice identity recognition in AVH schizophrenia patients. The voice identity recognition deficit is proposed to be the link by which AVH patients might misidentify their own inner speech (chapter I). Moreover, it has also been proposed that additional contributions from delusional thinking might be necessary to attribute inner speech to an external source resulting in hallucinations being perceived as “alien”. Since previous theoretical models neglected the role of

perceptual deficits in the misattribution of the voices heard in hallucination, the current thesis demonstrated that such deficits are on the basis of AVH formation.

Another aspect of the model presented in chapter VI, which was not investigated by previous models of AVH, is the mechanism by which hallucinations obtain their emotional tone. It has been demonstrated that salient emotional prosodic stimuli capture involuntary attention in patients as well as in controls, while healthy controls as well as NAVH patients have the ability to overcome the attentional capture when the emotional stimuli are presented away from the focus of attention. AVH patients fail to overcome such emotional distractor (chapter VI). In other words, the modulation of involuntary attention by emotional prosody is abnormal in AVH patients, in comparison to healthy controls and NAVH. That is, the AVH group cannot ignore an emotional stimulus even when such a distractor is outside the focus of selective attention, showing difficulties in top-down control, when interfered by automatic emotional processing (bottom-up input). Such difficulty in the AVH group seems to rely on atypical lateralization of emotional prosody perception (chapter VI). The present thesis demonstrated abnormal capture of attention by implicit, irrelevant emotional prosodic stimuli in AVH patients at the behavioral level (chapter VI). Although the dichotic listening paradigm gives information about the inter-hemispheric representation of the explicit sex discrimination task as well as the implicit emotional prosody processing, evidence regarding the intra-hemispheric organization of such functions can only be found in the literature.

In line with the notion of psychical precipitation proposed by Penfield (1963), intrinsic physiological activity of the primary auditory cortex (Hunter et al.,

2006) and amygdala (Aleman & Kahn, 2005) may conform the neural basis resulting in abnormal perception of irrelevant emotional auditory objects, and the behavioral consequences of this abnormality were shown in chapter VI. It has been shown that the amygdala receives inputs from the prefrontal cortex, which regulates its activity via afferents suppressing amygdala output (Rosenkranz & Grace, 2001). This amygdala regulation by prefrontal cortex is disrupted in schizophrenia (Anticevic, Repovs, & Barch, 2011; Leitman et al., 2008). Therefore, it is hypothesised that in schizophrenia, prefrontal circuitry may fail to down-regulate amygdala activity, which might therefore drive attention towards incoming affective cues, even when such emotional cues are irrelevant. At the behavioural level, amygdala deregulation was demonstrated by the impossibility of AVH patients to stop processing non-attended emotional prosody distractors (chapter VI, for an extended discussion in this topic see also Chapter I). As the perception of prosodic features is abnormal in AVH patients, due in part to abnormalities in the STG, supra-segmental information might be degraded. Then, the amygdala, deregulated due to poor coupling to the PFC (Anticevic et al., 2011) might impose a negative emotional tone to the already deteriorated auditory perception. In fact, difficulties in EPC shown by a high AVH traits group (chapter IV) could be linked to amygdala deregulation in addition to STG abnormalities, as both structures are involved in EPC (chapter II). The aberrant auditory perception with negative emotional tone might be misidentified with respect to its source, perhaps due to the difficulties in voice identity recognition explained in chapter V, becoming a hallucination. Unfortunately, intrinsic functional connectivity between the auditory cortex, the amygdala and the PFC was not evaluated, thus validity of the statements proposing abnormal oscillation of physiological activity in

the auditory cortex, amygdala hyperactivation at baseline level, and lack of amygdala inhibition by the prefrontal cortex in AVH patients still remains unknown.

Chapter I presented a model of AVH in which lack of top-down control in the presence of bottom-up irrelevant emotional inputs contributed to the formation of hallucinations. Additionally, it was also suggested that such detriment of top-down control of bottom-up emotional distractors may be specific for negative valences. This would be in line with a principal phenomenological feature of AVH, in which the salient emotional tone is often negative. Unfortunately, the specificity of negative emotional valence in the emotional bias could not be addressed by the present thesis. Chapter VI suggests a general abnormal top-down control in the presence of emotional bottom-up distractors, regardless of their emotional valence. However, because of statistical power issues, valence specificity was not investigated.

7.6 Summary of the model.

In the proposed model, auditory verbal hallucinations can emerge in the absence of external auditory objects (as a result of spontaneous physiological fluctuations in the auditory cortex) or they can be triggered by external auditory stimuli. In the latter case, it is proposed that salient emotional features of external sounds capture the hallucinator's involuntary attention, resulting in these emotional features gaining access to processing in detriment of more relevant external stimuli.

In a second stage, perceived auditory objects (independent of whether they were internally generated or triggered by external stimuli) are processed in the posterior right superior temporal gyrus (STG). This is the key region for EPC (see chapters II and III). It is proposed that this area is responsible for attaching an

emotional prosodic tone to the AVH. Moreover, a more anterior part of the right STG is meant to host the capacity of voice identity recognition (Belin et al., 2004; Warren et al., 2006). It has been demonstrated in chapter V that voice identity recognition is impaired in AVH patients, which might contribute towards a misidentification of inner speech resulting in an inner voice perceived as alien.

Finally, it may be the case that the STG is aberrantly connected to the anterior cingulate cortex (ACC) in AVH patients (Hunter et al., 2006). The ACC has been implicated in self non-self differentiation and sense of ownership, such as experiencing one's own body and thought as personal (Northoff & Bermpohl, 2004). Aberrant functional connectivity between ACC and STG in AVH patients may underlie the impossibility of some of these patients to identify their auditory perceptions as self-generated. Such mechanism may contribute to the misidentification of one's own voice in AVH patients reported in the literature (Allen et al., 2004; Belin et al., 2004; Johns et al., 2001; Warren, et al., 2006). Although this capacity was not directly tested in the present thesis, chapter V found defective voice identity recognition in AVH patients, which in addition to self non-self differentiation deficits may result in a spontaneous auditory perception whose source is misattributed.

7.7 General conclusion

This thesis introduced a new model of AVH. This novel neuropsychological model of the formation of AVH integrates aspects from a previous cognitive model (the misattribution of inner speech) and the perceptual (neurological) model and adds assumptions which were empirically tested. Previous models' weaknesses in

explaining the phenomenology of AVH have been analyzed, and moreover new formulations to overcome such flaws have been postulated. Firstly, by revealing the neural correlates of EPC and proposing an overlap between emotional prosody decoding and AVH (particularly in the STG), this thesis provides some cues with respect to the origin of the prosodic features conveyed by hallucinated voices; that is the demonstration of implicit emotionally salient auditory objects driving attention in detriment of more relevant stimuli which is abnormal in AVH patients (chapter VI). Additionally, by evaluating the presence of a voice identity recognition deficit in AVH, the present model offers some hints regarding why AVH have personal identities. Moreover, the present model takes into account intrinsic physiological fluctuation between the auditory cortex, the STG and the ACC, which can explain why AVH may emerge in the absence of external sounds (as a form of spontaneous ictal activity in primary sensory areas) or even being triggered by external sounds (as psychical precipitation followed by aberrant processing of prosodic features), and in both cases been experienced as perceptions instead of thoughts (Hunter et al., 2006). Finally, the present model is supported by evidence from the literature at the neural level, and it is empirically testable.

The present integrative model provides a bridge between the neuropathology and the phenomenology of hallucinations. Such model can be used as a foundation to generate therapeutic strategies to ameliorate this symptom at the neural level (i.e. by targeting regions of interest with neurodisruptive techniques) as well as at the behavioral level (i.e. by training patients to withdraw attention from the voices). This model provides a testable framework from which to generate hypotheses for further research.

REFERENCES

Adami, A. G. (2007). Modeling prosodic differences for speaker recognition. *Speech Communication, 49*(4), 277-291.

Adolphs, R. (2001). The neurobiology of social cognition. *Current Opinion in Neurobiologyl, 11*(2), 231-239.

Adolphs, R. (2010). Social Cognition: Feeling Voices to Recognize Emotions. *Current Biology, 20*(24), R1071-R1072.

Adolphs, R., Baron-Cohen, S., & Tranel, D. (2002). Impaired recognition of social emotions following amygdala damage. *Journal of Cognitive Neuroscience, 14*(8), 1264-1274.

Adolphs, R., Damasio, H., & Tranel, D. (2002). Neural systems for recognition of emotional prosody: a 3-D lesion study. *Emotion, 2*(1), 23-51.

Adolphs, R., Damasio, H., Tranel, D., & Damasio, A. R. (1996). Cortical systems for the recognition of emotion in facial expressions. *Journal of Neuroscience, 16*(23), 7678-7687.

Aleman, A., & Kahn, R. S. (2005). Strange feelings: do amygdala abnormalities dysregulate the emotional brain in schizophrenia? *Progress in Neurobiology, 77*(5), 283-298.

Aleman, A., Sommer, I. E., & Kahn, R. S. (2007). Efficacy of slow repetitive transcranial magnetic stimulation in the treatment of resistant auditory hallucinations in schizophrenia: a meta-analysis. *Journal of Clinical Psychiatry, 68*(3), 416-421.

Allen, P. P., Johns, L. C., Fu, C. H., Broome, M. R., Vythelingum, G. N., & McGuire, P. K. (2004). Misattribution of external speech in patients with hallucinations and delusions. *Schizophrenia Research, 69*(2-3), 277-287.

Andreasen. (1984a). *Modified Scale for the Assessment of Negative Symptoms (SANS)*: University of Iowa, Iowa City

Andreasen. (1984b). *Scale for the Assessment of Positive Symptoms (SAPS)*: University of Iowa, Iowa City.

Andreasen, Flaum, M., & Arndt, S. (1992). The Comprehensive Assessment of Symptoms and History (CASH). An instrument for assessing diagnosis and psychopathology. *Archives of General Psychiatry, 49*(8), 615-623.

Anticevic, A., Repovs, G., & Barch, D. M. (2011). Emotion Effects on Attention, Amygdala Activation, and Functional Connectivity in Schizophrenia. *Schizophrenia Bulletin*. In press

Aue, T., Cuny, C., Sander, D., & Grandjean, D. (2010). Peripheral responses to attended and unattended angry prosody: A dichotic listening paradigm. *Psychophysiology*. 48(3):385-92

Aziz-Zadeh, L., Sheng, T., & Gheytanchi, A. (2010). Common premotor regions for the perception and production of prosody and correlations with empathy and prosodic ability. *PLoS ONE, 5*(1), e8759.

Bach, D. R., Buxtorf, K., Grandjean, D., & Strik, W. K. (2008). The influence of emotion clarity on emotional prosody identification in paranoid schizophrenia. *Psychological medicine*, 1-12.

Bach, D. R., Grandjean, D., Sander, D., Herdener, M., Strik, W. K., & Seifritz, E. (2008). The effect of appraisal level on processing of emotional prosody in meaningless speech. *Neuroimage, 42*(2), 919-927.

Bach, D. R., Herdener, M., Grandjean, D., Sander, D., Seifritz, E., & Strik, W. K. (2009). Altered lateralisation of emotional prosody processing in schizophrenia. *Schizophrenia Research, 110*(1-3), 180-187.

Banissy, M. J., Sauter, D. A., Ward, J., Warren, J. E., Walsh, V., & Scott, S. K. (2010). Suppressing sensorimotor activity modulates the discrimination of auditory emotions but not speaker identity. *Journal of Neuroscience, 30*(41), 13552-13557.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*(3), 614-636.

Bastiaansen, J. A., Thioux, M., & Keysers, C. (2009). Evidence for mirror systems in emotions. *Archives of Philosopical Transactions of the Royal Society London Biological Science, 364*(1528), 2391-2404.

Baumgartner, T., Lutz, K., Schmidt, C. F., & Jancke, L. (2006). The emotional power of music: how music enhances the feeling of affective pictures. *Brain Research, 1075*(1), 151-164.

Bauminger, N., Edelsztein, H. S., & Morash, J. (2005). Social information processing and emotional understanding in children with LD. *Journal of Learning Disabilities, 38*(1), 45-61.

Beatty, W. W., Orbelo, D. M., Sorocco, K. H., & Ross, E. D. (2003). Comprehension of affective prosody in multiple sclerosis. *Multiple Sclerosis, 9*(2), 148-153.

Beaucousin, V., Lacheret, A., Turbelin, M. R., Morel, M., Mazoyer, B., & Tzourio-Mazoyer, N. (2007). FMRI study of emotional speech comprehension. *Cerebral Cortex, 17*(2), 339-352.

Behrens, S. J. (1985). The perception of stess and lateralization of prosody. [Article]. *Brain and Language, 26*(2), 332-348.

Behrens, S. J. (1988). The role of the right hemisphere in the production of linguistic stress. *Brain Language, 33*(1), 104-127.

Békésy, G. (1949). The structure of the middle ear and the hearing of one's own voice by bone conduction. *Journal of the Acoustical Society of America, 21*, 217-232

Belin, P., Fecteau, S., & Bedard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends in Cognitive Sciences, 8*(3), 129-135.

Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal Affective Voices: a validated set of nonverbal affect bursts for research on auditory affective processing. *Behavioral Research Methods, 40*(2), 531-539.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature, 403*(6767), 309-312.

Bentall, R. P. (1990). The illusion of reality: a review and integration of psychological research on hallucinations. *Psychological Bulletin, 107*(1), 82-95.

Birchwood, M., & Chadwick, P. (1997). The omnipotence of voices: testing the validity of a cognitive model. *Psychological Medicine, 27*(6), 1345-1353.

Bjoertomt, O., Cowey, A., & Walsh, V. (2002). Spatial neglect in near and far space investigated by repetitive transcranial magnetic stimulation. *Brain, 125*(Pt 9), 2012-2022.

Blakemore, S. J. (2008). The social brain in adolescence. *Nat Rev Neurosci, 9*(4), 267-277.

Blonder, L. X., Bowers, D., & Heilman, K. M. (1991). The role of the right hemisphere in emotional communication. *Brain, 114 (Pt 3)*, 1115-1127.

Bora, E., Gokcen, S., & Veznedaroglu, B. (2008). Empathic abilities in people with schizophrenia. [Article]. *Psychiatry Research, 160*(1), 23-29.

Borod, J. C., Alpert, M., Brozgold, A., Martin, C., Welkowitz, J., Diller, L., Peselow, E., Angrist, B. & Lieberman, A. (1989). A preliminary comparison of flat affect schizophrenics and brain-damaged patients on measures of affective processing. *Journal of Communication Disorders, 22*(2), 93-104.

Borod, J. C., Andelman, F., Obler, L. K., Tweedy, J. R., & Welkowitz, J. (1992). Right hemisphere specialization for the identification of emotional words and sentences: evidence from stroke patients. *Neuropsychologia, 30*(9), 827-844.

Borod, J. C., Welkowitz, J., Alpert, M., Brozgold, A. Z., Martin, C., Peselow, E., & Diller, L. (1990). Parameters of emotional processing in neuropsychiatric disorders: conceptual issues and a battery of tests. *Journal of Communication Disorders, 23*(4-5), 247-271.

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

Bozikas, V. P., Kosmidis, M. H., Anezoulaki, D., Giannakou, M., & Karavatos, A. (2004). Relationship of affect recognition with psychopathology and cognitive performance in schizophrenia. *Journal of the International Neuropsychological Society, 10*(4), 549-558.

Bradley, M. M., & Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology, 37*(2), 204-215.

Brand, J. G., Burton, L. A., Schaffer, S. G., Alper, K. R., Devinsky, O., & Barr, W. B. (2009). Emotional recognition in depressed epilepsy patients. *Epilepsy & Behavior. 15*(3), 333-338.

Bryant, G. A., & Fox Tree, J. E. (2005). Is there an ironic tone of voice? *Language and speech, 48*(Pt 3), 257-277.

Buchanan, T. W., Lutz, K., Mirzazade, S., Specht, K., Shah, N. J., Zilles, K., & Jancke, L. (2000). Recognition of emotional prosody and verbal components of spoken language: an fMRI study. *Brain Reseach: Cognitive Brain Research, 9*(3), 227-238.

Capps, L., Yirmiya, N., & Sigman, M. (1992). Understanding of Simple and Complex Emotions in Nonretarded-Children with Autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 33*(7), 1169-1182.

Carr, L., Iacoboni, M., Dubeau, M. C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Science U S A, 100*(9), 5497-5502.

Chakrabarti, B., & Baron-Cohen, S. (2006). Empathizing: neurocognitive developmental mechanisms and individual differences. *Progress in Brain Research, 156*, 403-417.

Chevallier, C., Noveck, I., Happe, F., & Wilson, D. (2009). From acoustics to grammar: Perceiving and interpreting grammatical prosody in adolescents with Asperger Syndrome. [Article]. *Research in Autism Spectrum Disorders, 3*(2), 502-516.

Cochrane, M., Petch, I., & Pickering, A. D. (2010). Do measures of schizotypal personality provide non-clinical analogues of schizophrenic symptomatology? *Psychiatry Research, 176*(2-3), 150-154.

Cohen, A. S., Iglesias, B., & Minor, K. S. (2009). The neurocognitive underpinnings of diminished expressivity in schizotypy: What the voice reveals. [Article]. *Schizophrenia Research, 109*(1-3), 38-45.

Colbert, S. M., Peters, E. R., & Garety, P. A. (2009). Delusions and belief flexibility in psychosis. *Psychology and psychotherapy*. 83(Pt 1):45-57

Copolov, D. L., Mackinnon, A., & Trauer, T. (2004). Correlates of the affective impact of auditory hallucinations in psychotic disorders. *Schizophrenia Bulletin, 30*(1), 163-171.

Couture, S., Lecomte, T., & Leclerc, C. (2007). Personality characteristics and attachment in first episode psychosis - Impact on social functioning. [Article]. *Journal of Nervous and Mental Disease, 195*(8), 631-639.

Crow, T. J. (1990). Temporal lobe asymmetries as the key to the etiology of schizophrenia. *Schizophrenia Bulletin, 16*(3), 433-443.

Cunningham, W. A., Johnsen, I. R., & Waggoner, A. S. (2010). Orbitofrontal cortex provides cross-modal valuation of self-generated stimuli. *Social Cognitive and Affective Neurosci*.

Cutting, J. C. (1990). The Right Cerebral Hemisphere and Psychiatric Disorders. *Oxford University Press*.

Damjanovic, L., & Hanley, J. R. (2007). Recalling episodic and semantic information about famous faces and voices. *Memory and Cognition, 35*(6), 1205-1210.

Davidson, M., Reichenberg, A., Rabinowitz, J., Weiser, M., Kaplan, Z., & Mark, M. (1999). Behavioral and intellectual markers for schizophrenia in apparently healthy male adolescents. *American Journal of Psychiatry, 156*(9), 1328-1335.

Demonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., Rascol, A., & Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain, 115 (Pt 6)*, 1753-1768.

Dierks, T., Linden, D. E., Jandl, M., Formisano, E., Goebel, R., Lanfermann, H., & Singer, W. (1999). Activation of Heschl's gyrus during auditory hallucinations. *Neuron, 22*(3), 615-621.

Donaldson, W. (1993). Accuracy of D' and a' as Estimates of Sensitivity. *Bulletin of the Psychonomic Society, 31*(4), 271-274.

Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory and Cognition, 24*(4), 523-533.

Done, D. J., Crow, T. J., Johnstone, E. C., & Sacker, A. (1994). Childhood antecedents of schizophrenia and affective illness: social adjustment at ages 7 and 11. *British Medical Journal, 309*(6956), 699-703.

Drakeford, J. L., Edelstyn, N. M., Oyebode, F., Srivastava, S., Calthorpe, W. R., & Mukherjee, T. (2006). Auditory recognition memory, conscious recollection, and executive function in patients with schizophrenia. *Psychopathology, 39*(4), 199-208.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neuroscience, 23*(10), 475-483.

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*(2-3), 235-253.

Ekman, P. (1992). Are there basic emotions? *Psychological review*, *99*(3), 550-553.

Ellison, A., Battelli, L., Cowey, A., & Walsh, V. (2003). The effect of expectation on facilitation of colour/form conjunction tasks by TMS over area V5. *Neuropsychologia, 41*(13), 1794-1801

Ely, P. W., Graves, R. E., & Potter, S. M. (1989). Dichotic listening indices of right hemisphere semantic processing. *Neuropsychologia*, *27*(7), 1007-1015.

Erkwoh, R., Rodon, A., Nacken, A., Lampe, K., Doring, W. H., Vorlander, M., Volmer, A., Harke, K., Kunert, H & Hoff, P. (2006). Directional hearing and functional imaging in schizophrenia. *Neurology Psychiatry and Brain Research, 13*(1), 1-8.

Escarti, M. J., de la Iglesia-Vaya, M., Marti-Bonmati, L., Robles, M., Carbonell, J., Lull, J. J., Garcia-Marti, G., Manjon, J., Aguilar, E., Aleman, A., & Sanjuan, J. (2010). Increased amygdala and parahippocampal gyrus activation in schizophrenic patients with auditory hallucinations: An fMRI study using independent component analysis. *Schizophrenia Research, 117*(1), 31-41.

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*(2), 580-587.

Ethofer, T., Kreifelts, B., Wiethoff, S., Wolf, J., Grodd, W., Vuilleumier, P., & Wildgruber, D. (2009). Differential influences of emotion, task, and novelty on brain regions underlying the processing of speech melody. *Journal of Cognitive Neuroscience, 21*(7), 1255-1268.

Evans, C. L., McGuire, P. K., & David, A. S. (2000). Is auditory imagery defective in patients with auditory hallucinations? *Psychological Medicine, 30*(1), 137-148.

Fernyhough, C., Jones, S. R., Whittle, C., Waterhouse, J., & Bentall, R. P. (2008). Theory of mind, schizotypy, and persecutory ideation in young adults. *Cognitive neuropsychiatry*, *13*(3), 233-249.

Formisano, E., De Martino, F., Bonte, M., & Goebel, R. (2008). "Who" is saying "what"? Brain-based decoding of human voice and speech. *Science, 322*(5903), 970-973.

Fossati, P., Amar, G., Raoux, N., Ergis, A. M., & Allilaire, J. F. (1999). Executive functioning and verbal memory in young patients with unipolar depression and schizophrenia. *Psychiatry Research, 89*(3), 171-187.

Friederici, A. D., & Alter, K. (2004). Lateralization of auditory language functions: a dynamic dual pathway model. *Brain and Language, 89*(2), 267-276

Frith, C. D. (1987). The Positive and Negative Symptoms of Schizophrenia Reflect Impairments in the Perception and Initiation of Action. *Psychological Medicine, 17*(3), 631-648.

Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron, 50*(4), 531-534.

Fyfe, S., Williams, C., Mason, O. J., & Pickup, G. J. (2008). Apophenia, theory of mind and schizotypy: perceiving meaning and intentionality in randomness. *Cortex, 44*(10), 1316-1325.

Garcia-Marti, G., Aguilar, E. J., Lull, J. J., Marti-Bonmati, L., Escarti, M. J., Manjon, J. V., Moratal, D., Roble, M., & Sanjuan, J. (2008). Schizophrenia with auditory hallucinations: a voxel-based morphometry study. *Progress in Neuropsychopharmacology and Biologcal Psychiatry, 32*(1), 72-80.

Gardiner, J. M., & Gregg, V. H. (1997). Recognition memory with little or no remembering: Implications for a detection model. *Psychonomic Bulletin & Review, 4*(4), 474-479.

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (2002). Recognition memory and decision processes: A meta-analysis of remember, know, and guess responses. *Memory, 10*(2), 83-98.

Garrido, L., Eisner, F., McGettigan, C., Stewart, L., Sauter, D., Hanley, J. R., Schweinberger, S., Warren, J., & Duchaine, B. (2009). Developmental phonagnosia: a selective deficit of vocal identity recognition. *Neuropsychologia, 47*(1), 123-131.

Garrity, A. G., Pearlson, G. D., McKiernan, K., Lloyd, D., Kiehl, K. A., & Calhoun, V. D. (2007). Aberrant "default mode" functional connectivity in schizophrenia. *American Journal of Psychiatry, 164*(3), 450-457.

George, M. S., Parekh, P. I., Rosinsky, N., Ketter, T. A., Kimbrell, T. A., Heilman, K. M., . . . Post, R. M. (1996). Understanding emotional prosody activates right hemisphere regions. *Archives in Neurology, 53*(7), 665-670.

Ghashghaei, H. T., & Barbas, H. (2002). Pathways for emotion: interactions of prefrontal and anterior temporal pathways in the amygdala of the rhesus monkey. *Neuroscience, 115*(4), 1261-1279.

Glasser, M. F., & Rilling, J. K. (2008). DTI tractography of the human brain's language pathways. *Cerebral Cortex, 18*(11), 2471-2482.

Golan, O., Baron-Cohen, S., Hill, J. J., & Golan, Y. (2006). The "reading the mind in films" task: complex emotion recognition in adults with and without autism spectrum conditions. *Social neuroscience, 1*(2), 111-123.

Grandjean, D., Sander, D., Lucas, N., Scherer, K. R., & Vuilleumier, P. (2008). Effects of emotional prosody on auditory extinction for voices in patients with spatial neglect. *Neuropsychologia, 46*(2), 487-496.

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*(2), 145-146.

Green, M. J., Williams, L. M., & Davidson, D. J. (2001). Processing of threat-related affect is delayed in delusion-prone individuals. *British Journal of Clinical Psychology, 40*(Pt 2), 157-165.

Grimshaw, G. M., Kwasny, K. M., Covell, E., & Johnson, R. A. (2003). The dynamic nature of language lateralization: effects of lexical and prosodic factors. *Neuropsychologia, 41*(8), 1008-1019.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of personality and social psychology, 85*(2), 348-362.

Guillaume, F., Guillem, F., Tiberghien, G., Martin, F., Ganeva, E., Germain, M., Pampoulova, T., & Lalonde, P. (2007). Use of the process dissociation procedure to study the contextual effects on face recognition in schizophrenia: familiarity, associative recollection and discriminative recollection. *Psychiatry Research, 149*(1-3), 105-119.

Guterman, Y., & Klein, E. (1991). The Role of Head Movement and Pinnae in Auditory Localization in Schizophrenia and Psychosis. *Schizophrenia Research, 6*(1), 67-73.

Haddock, G., McCarron, J., Tarrier, N., & Faragher, E. B. (1999). Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales (PSYRATS). *Psycological Medicine, 29*(4), 879-889.

Hahn, B., Robinson, B. M., Kaiser, S. T., Harvey, A. N., Beck, V. M., Leonard, C. J., Kappenman, E., Luck S.J., & Gold, J. M. (2010). Failure of schizophrenia patients to overcome salient distractors during working memory encoding. *Biological Psychiatry, 68*(7), 603-609.

Hailstone, J. C., Crutch, S. J., Vestergaard, M. D., Patterson, R. D., & Warren, J. D. (2010). Progressive associative phonagnosia: A neuropsychological analysis. *Neuropsychologia, 48*(4), 1104-1114.

Hale, W. W., 3rd, Jansen, J. H., Bouhuys, A. L., & van den Hoofdakker, R. H. (1998). The judgement of facial expressions by depressed patients, their partners and controls. *Journal of Affective Disorders, 47*(1-3), 63-70.

Hanley, J. R., & Damjanovic, L. (2009). It is more difficult to retrieve a familiar person's name and occupation from their voice than from their blurred face. *Memory, 17*(8), 830-839.

Haskins, B., Shutty, M. S., & Kellogg, E. (1995). Affect processing in chronically psychotic patients: development of a reliable assessment tool. *Schizophrenia research, 15*(3), 291-297.

Heilbrun, A. B., Jr., Blum, N., & Haas, M. (1983). Cognitive vulnerability to auditory hallucination. Preferred imagery mode and spatial location of sounds. *British Journal of Psychiatry, 143*, 294-299.

Heilman, K. M., Scholes, R., & Watson, R. T. (1975). Auditory affective agnosia. Disturbed comprehension of affective speech. *Journal of Neurology, Neurosurgery and Psychiatry, 38*(1), 69-72.

Heim, S. (2005). The structure and dynamics of normal language processing: insights from neuroimaging. *Acta Neurobiologiae Experimentalis (Warsaw), 65*(1), 95-116.

Hein, G., & Knight, R. T. (2008). Superior temporal sulcus--It's my area: or is it? *Journal of Cognitive Neuroscience, 20*(12), 2125-2136.

Henry, J. D., Bailey, P. E., & Rendell, P. G. (2008). Empathy, social functioning and schizotypy. [Article]. *Psychiatry Research, 160*(1), 15-22.

Hirano, S., Hirano, Y., Maekawa, T., Obayashi, C., Oribe, N., Kuroki, T., Kanba, S., & Onitsuka, T. (2008). Abnormal neural oscillatory activity to speech sounds in schizophrenia: A magnetoencephalography study. *Journal of Neuroscience, 28*(19), 4897-4903.

Hirnstein, M., Hausmann, M., & Lewald, J. (2007). Functional cerebral asymmetry in auditory motion perception. *Laterality, 12*(1), 87-99.

Hoekert, M., Bais, L., Kahn, R. S., & Aleman, A. (2008). Time course of the involvement of the right anterior superior temporal gyrus and the right fronto-parietal operculum in emotional prosody perception. *PLoS ONE, 3*(5), e2244.

Hoekert, M., Kahn, R. S., Pijnenborg, M., & Aleman, A. (2007). Impaired recognition and expression of emotional prosody in schizophrenia: Review and meta-analysis. *Schizophrenia Research, 96*(1-3), 135-145

Hoekert, M., Vingerhoets, G., & Aleman, A. (2010). Results of a pilot study on the involvement of bilateral inferior frontal gyri in emotional prosody perception: an rTMS study. [Article]. *BioMed Central Neuroscience, 11*.

Hoffman, R. E., Hampson, M., Wu, K., Anderson, A. W., Gore, J. C., Buchanan, R. J., Constable R.T., Haukins, K.A., Sahay, N. & Krystal, J. H. (2007). Probing the pathophysiology of auditory/verbal hallucinations by combining functional magnetic resonance imaging and transcranial magnetic stimulation. *Cerebral Cortex, 17*(11), 2733-2743.

Hoffman, R. E., Varanko, M., Gilmore, J., & Mishara, A. L. (2008). Experiential features used by patients with schizophrenia to differentiate 'voices' from ordinary verbal thought. *Psychological Medicine, 38*(8), 1167-1176.

Honea, R., Crow, T. J., Passingham, D., & Mackay, C. E. (2005). Regional deficits in brain volume in schizophrenia: a meta-analysis of voxel-based morphometry studies. *American Journal of Psychiatry, 162*(12), 2233-2245.

Hooker, C., & Park, S. (2002a). Emotion processing and its relationship to social functioning in schizophrenia patients. [Article]. *Psychiatry Research, 112*(1), 41-50.

Hooker, C., & Park, S. (2002b). Emotion processing and its relationship to social functioning in schizophrenia patients. *Psychiatry Research, 112*(1), 41-50.

Hooker, C. I., Verosky, S. C., Germine, L. T., Knight, R. T., & D'Esposito, M. (2008). Mentalizing about emotion and its relationship to empathy. *Social Cognitive and Affective Neuroscience, 3*(3), 204-217.

Hu, S. W., Olulade, O., Castillo, J. G., Santos, J., Kim, S., Tamer, G. G., Luh, W.M. & Talavage, T. M. (2010). Modeling hemodynamic responses in auditory cortex at 1.5 T using variable duration imaging acoustic noise. *Neuroimage, 49*(4), 3027-3038.

Hubl, D., Koenig, T., Strik, W., Federspiel, A., Kreis, R., Boesch, C., Maier, S.E. Schroth G., Lvblad, K. & Dierks, T. (2004). Pathways that make voices: white matter changes in auditory hallucinations. *Archives of General Psychiatry, 61*(7), 658-668.

Hugdahl, K. (2009). "Hearing voices": auditory hallucinations as failure of top-down control of bottom-up perceptual processes. *Scandinavian Journal of Psychology, 50*(6), 553-560.

Hugdahl, K., Loberg, E. M., Specht, K., Steen, V. M., van Wageningen, H., & Jorgensen, H. A. (2007). Auditory hallucinations in schizophrenia: the role of cognitive, brain structural and genetic disturbances in the left temporal lobe. *Frontiers in Humam Neuroscience, 1*, 6.

Hugdahl, K., Westerhausen, R., Alho, K., Medvedev, S., & Hamalainen, H. (2008). The effect of stimulus intensity on the right ear advantage in dichotic listening. *Neuroscience Letters, 431*(1), 90-94.

Hunter, M. D., Eickhoff, S. B., Miller, T. W., Farrow, T. F., Wilkinson, I. D., & Woodruff, P. W. (2006). Neural activity in speech-sensitive auditory cortex during silence. *Proceedings of the Nattional Acaddemy of Science U S A, 103*(1), 189-194.

Hunter, M. D., & Woodruff, P. W. (2004). Characteristics of functional auditory hallucinations. *American Journal of Psychiatry, 161*(5), 923.

Ischebeck, A. K., Friederici, A. D., & Alter, K. (2008). Processing prosodic boundaries in natural and hummed speech: An fMRI study. *Cerebral Cortex, 18*(3), 541-552.

Jabbi, M., Swart, M., & Keysers, C. (2007). Empathy for positive and negative emotions in the gustatory cortex. *Neuroimage, 34*(4), 1744-1753

Jahshan, C. S., & Sergi, M. J. (2007). Theory of mind, neurocognition, and functional status in schizotypy. *schizophrenia research, 89*(1-3), 278-286.

Jardri, R., Pouchet, A., Pins, D., & Thomas, P. (2011). Cortical activations during auditory verbal hallucinations in schizophrenia: a coordinate-based meta-analysis. *American Journal of Psychiatry, 168*(1), 73-81.

Javitt, D. C. (2009). When doors of perception close: bottom-up models of disrupted cognition in schizophrenia. *Annual Review of Clinical Psychology, 5*, 249-275.

Jenner, J. A., Rutten, S., Beuckens, J., Boonstra, N., & Sytema, S. (2008). Positive and useful auditory vocal hallucinations: prevalence, characteristics, attributions, and implications for treatment. *Acta Psychiatrica Scandinavica, 118*(3), 238-245.

Jobard, G., Vigneau, M., Mazoyer, B., & Tzourio-Mazoyer, N. (2007). Impact of modality and linguistic complexity during reading and listening tasks. *Neuroimage, 34*(2), 784-800.

Johns, L. C., Rossell, S., Frith, C., Ahmad, F., Hemsley, D., Kuipers, E., & McGuire, P. K. (2001). Verbal self-monitoring and auditory verbal hallucinations in patients with schizophrenia. *Psychological Medicine, 31*(4), 705-715.

Johnson-Laird, P. N., & Oatley, K. (1989). The language of emotions: An analysis of a semantic field. *Cognition and Emotion*(3), 81-123.

Jones, S. R., & Fernyhough, C. (2007). Neural correlates of inner speech and auditory verbal hallucinations: a critical review and theoretical integration. *Clinical Psychology Reviews, 27*(2), 140-154.

Kang, J. I., Kim, J. J., Seok, J. H., Chun, J. W., Lee, S. K., & Park, H. J. (2008). Abnormal brain response during the auditory emotional processing in schizophrenic patients with chronic auditory hallucinations. *Schizophrenia Research*. 107(1):83-91

Kee, K. S., Green, M. F., Mintz, J., & Brekke, J. S. (2003). Is emotion processing a predictor of functional outcome in schizophrenia? *Schizophrenia Bulletin, 29*(3), 487-497.

Keltner, D., & Haidt, J. (1999). Social functions of emotions at four levels of analysis. *Cognition & Emotion, 13*(5), 505-521.

Kerns, J. G. (2005). Positive schizotypy and emotion processing. [Article]. *Journal of Abnormal Psychology, 114*(3), 392-401.

Keshavan, M. S., Eack, S. M., Mermon, D. E., Montrose, D. M., Miewald, J., Gur, R. E., Gur, R.C, Sweeney, J.A. & Keshavan, M.S. (2010). Social Cognition Deficits Among Individuals at Familial High Risk for Schizophrenia. *Schizophrenia bulletin, 36*(6), 1081-1088.

Keysers, C., & Gazzola, V. (2006). Towards a unifying neural theory of social cognition. *Understanding Emotions, 156*, 379-401.

Kington, J. M., Jones, L. A., Watt, A. A., Hopkin, E. J., & Williams, J. (2000). Impaired eye expression recognition in schizophrenia. *Journal of Psychiatric Research, 34*(4-5), 341-347.

Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J. P., Thompson, W. L., Ganis, G., Sukel, K.E. & Alpert, N. M. (1999). The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science, 284*(5411), 167-170.

Kotz, S. A., Meyer, M., Alter, K., Besson, M., von Cramon, D. Y., & Friederici, A. D. (2003). On the lateralization of emotional prosody: an event-related functional MR investigation. *Brain and Language, 86*(3), 366-376.

Kraepelin, E. (1971 [1919]). Dementia praecox and paraphrenia. *Huntington, NY: R. E. Krieger*.

Kuperberg, G. R., McGuire, P. K., Bullmore, E. T., Brammer, M. J., Rabe-Hesketh, S., Wright, I. C., Lythgoe, D.J., Williams, S.C. & David, A. S. (2000). Common and distinct neural substrates for pragmatic, semantic, and syntactic processing of spoken sentences: an fMRI study. *Journal of Cognitive Neuroscience, 12*(2), 321-341.

Langdon, R., & Coltheart, M. (2004). Recognition of metaphor and irony in young adults: the impact of schizotypal personality traits. *Psychiatry research, 125*(1), 9-20.

Laroi, F., & Woodward, T. S. (2007). Hallucinations from a cognitive perspective. *Harvard Review of Psychiatry, 15*(3), 109-117.

Lattner, S., Meyer, M. E., & Friederici, A. D. (2005). Voice perception: Sex, pitch, and the right hemisphere. *Human Brain Mapping, 24*(1), 11-20.

Lee, L., Harkness, K. L., Sabbagh, M. A., & Jacobson, J. A. (2005). Mental state decoding abilities in clinical depression. *Journal of affective disorders, 86*(2-3), 247-258.

Lee, T. M., Chan, S. C., & Raine, A. (2009). Hyperresponsivity to threat stimuli in domestic violence offenders: a functional magnetic resonance imaging study. *Journal of Clinical Psychiatry, 70*(1), 36-45.

Leentjens, A. F., Wielaert, S. M., van Harskamp, F., & Wilmink, F. W. (1998). Disturbances of affective prosody in patients with schizophrenia; a cross sectional study. *Journal of neurology, neurosurgery, and psychiatry, 64*(3), 375-378.

Leitman, D. I., Foxe, J. J., Butler, P. D., Saperstein, A., Revheim, N., & Javitt, D. C. (2005). Sensory contributions to impaired prosodic processing in schizophrenia. *Biological Psychiatry, 58*(1), 56-61.

Leitman, D. I., Hoptman, M. J., Foxe, J. J., Saccente, E., Wylie, G. R., Nierenberg, J., Jalbrzikowski, M., Lim, K.O., & Javitt, D. C. (2007). The neural substrates of impaired prosodic detection in schizophrenia and its sensorial antecedents. *American Journal of Psychiatry, 164*(3), 474-482

Leitman, D. I., Loughead, J., Wolf, D. H., Ruparel, K., Kohler, C. G., Elliott, M. A., Biker, W.B., Gur, R.E., & Gur, R. C. (2008). Abnormal superior temporal connectivity during fear perception in schizophrenia. *Schizophrenia Bulletin, 34*(4), 673-678.

Leitman, D. I., Wolf, D. H., Ragland, J. D., Laukka, P., Loughead, J., Valdez, J. N., Javitt, D.C., Turetski, B.I., & Gur, R. C. (2010). "It's Not What You Say, But How You Say it": A Reciprocal Temporo-frontal Network for Affective Prosody. *Frontiers in Human Neuroscience, 4*, 19.

Leitman, D. I., Ziwich, R., Pasternak, R., & Javitt, D. C. (2006). Theory of Mind (ToM) and counterfactuality deficits in schizophrenia: misperception or misinterpretation? *Psychological Medicine, 36*(8), 1075-1083.

Leumann, L., Feldon, J., Vollenweider, F. X., & Ludewig, K. (2002). Effects of typical and atypical antipsychotics on prepulse inhibition and latent inhibition in chronic schizophrenia. *Biological Psychiatry, 52*(7), 729-739.

Levitan, C., Ward, P. B., & Catts, S. V. (1999). Superior temporal gyral volumes and laterality correlates of auditory hallucinations in schizophrenia. *Biological Psychiatry, 46*(7), 955-962.

Liddle, P. F. (1992). Syndromes of Schizophrenia on Factor-Analysis. *British Journal of Psychiatry, 161*, 861-861.

Lipp, O. V., & Waters, A. M. (2007). When danger lurks in the background: Attentional capture by animal fear-relevant distractors is specific and selectively enhanced by animal fear. *Emotion, 7*(1), 192-200.

Litter, J., & Walker, E. (1993). Interpersonal behavior of preschizophrenic children: a study of home-movies. *Child Psychiatry and Human Development, 23*(4), 283-295.

Loberg, E. M., Jorgensen, H. A., & Hugdahl, K. (2004). Dichotic listening in schizophrenic patients: effects of previous vs. ongoing auditory hallucinations. *Psychiatry Reseach, 128*(2), 167-174.

Machii, K., Cohen, D., Ramos-Estebanez, C., & Pascual-Leone, A. (2006). Safety of rTMS to non-motor cortical areas in healthy participants and patients. *Clinical Neurophysiology, 117*(2), 455-471.

Mah, L., Arnold, M. C., & Grafman, J. (2004). Impairment of social perception associated with lesions of the prefrontal cortex. *American Journ al of Psychiatry, 161*(7), 1247-1255.

Mason, O., Claridge, G., & Jackson, M. (1995). New scales for the assessment of schizotypy. [Article]. *Personality and Individual Differences, 18*(1), 7-13.

Mason, R. A., & Just, M. A. (2011). Differentiable cortical networks for inferences concerning people's intentions versus physical causality. *Human Brain Mapping, 32*(2), 313-329.

Matsumoto, K., Samson, G. T., O'Daly, O. D., Tracy, D. K., Patel, A. D., & Shergill, S. S. (2006). Prosodic discrimination in patients with schizophrenia. *The British journal of psychiatry : the journal of mental science, 189*, 180-181.

Maylor, E. A. (1995). REMEMBERING VERSUS KNOWING TELEVISION THEME TUNES IN MIDDLE-AGED AND ELDERLY ADULTS. [Article]. *British Journal of Psychology, 86*, 21-25.

McCreery, C., & Claridge, G. (2002). Healthy schizotypy: the case of out-of-the-body experiences. [Article]. *Personality and Individual Differences, 32*(1), 141-154.

McGuire, P. K., Shah, G. M., & Murray, R. M. (1993). Increased blood flow in Broca's area during auditory hallucinations in schizophrenia. *The Lancet, 342*(8873), 703-706.

Mechelli, A., Allen, P., Amaro, E., Fu, C. H. Y., Williams, S. C. R., Brammer, M. J., Johns, L.C., & McGuire, P. K. (2007). Misattribution of speech and impaired connectivity in patients with auditory verbal hallucinations. *Human Brain Mapping, 28*(11), 1213-1222.

Mitchell. (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*(12), 3611-3618.

Mitchell, & Crow, T. J. (2005). Right hemisphere language functions and schizophrenia: the forgotten hemisphere? *Brain, 128*(Pt 5), 963-978.

Mitchell, Elliott, R., Barry, M., Cruttenden, A., & Woodruff, P. W. R. (2003). The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia, 41*(10), 1410-1421.

Mitchell, Elliott, R., Barry, M., Cruttenden, A., & Woodruff, P. W. R. (2004). Neural response to emotional prosody in schizophrenia and in bipolar affective disorder. *British Journal of Psychiatry, 184*, 223-230.

Mitchell, & Ross, E. D. (2008). fMRI evidence for the effect of verbal complexity on lateralisation of the neural response associated with decoding prosodic emotion. *Neuropsychologia, 46*(12), 2880-2887.

Mohanty, A., Heller, W., Koven, N. S., Fisher, J. E., Herrington, J. D., & Miller, G. A. (2008). Specificity of emotion-related effects on attentional processing in schizotypy. *Schizophrenia research, 103*(1-3), 129-137.

Morrison, G., Sharkey, V., Allardyce, J., Kelly, R. C., & McCreadie, R. G. (2000). Nithsdale schizophrenia surveys 21: a longitudinal study of National Adult Reading Test stability. *Psychological Medicine, 30*(3), 717-720.

Mueser, K. T., Bellack, A. S., & Brady, E. U. (1990). Hallucinations in schizophrenia. *Acta Psychiatrica Scandinavica, 82*(1), 26-29.

Murphy, D., & Cutting, J. (1990). Prosodic comprehension and expression in schizophrenia. *Journal of Neurology, Neurosurgery and Psychiatry, 53*(9), 727-730.

Nayani, T. H., & David, A. S. (1996). The auditory hallucination: a phenomenological survey. *Psychological Medicine, 26*(1), 177-189.

Nelson, & Willson, J. (1991). National Adult Reading Test (NART). Retrieved from

Nelson, A. L., Combs, D. R., Penn, D. L., & Basso, M. R. (2007). Subtypes of social perception deficits in schizophrenia. [Article]. *Schizophrenia Research, 94*(1-3), 139-147.

Nenadic, I., Smesny, S., Schlosser, R. G., Sauer, H., & Gaser, C. (2010). Auditory hallucinations and brain structure in schizophrenia: voxel-based morphometric study. *British Journal of Psychiatry, 196*(5), 412-413.

Ngan, E. T. C., Vouloumanos, A., Cairo, T. A., Laurens, K. R., Bates, A. T., Anderson, C. M., Werker, J.S., & Liddle, P. F. (2003). Abnormal processing of speech during oddball target detection in schizophrenia. *Neuroimage, 20*(2), 889-897.

Niendam, T. A., Jalbrzikowski, M., & Bearden, C. E. (2009). Exploring predictors of outcome in the psychosis prodrome: implications for early identification and intervention. *Neuropsychol Rev, 19*(3), 280-293. doi: 10.1007/s11065-009-9108-z

Northoff, G., & Bermpohl, F. (2004). Cortical midline structures and the self. *Trends in Cognitive Sciences, 8*(3), 102-107.

Olin, S. C., & Mednick, S. A. (1996). Risk factors of psychosis: identifying vulnerable populations premorbidly. *Schizophrenia Bulletin, 22*(2), 223-240.

Olson, I. R., Plotzker, A., & Ezzyat, Y. (2007). The Enigmatic temporal pole: a review of findings on social and emotional processing. *Brain, 130*(Pt 7), 1718-1731.

Orchard, T. L., & Yarmey, A. D. (1995). The Effects of Whispers, Voice-Sample Duration, and Voice Distinctiveness on Criminal Speaker Identification. *Applied Cognitive Psychology, 9*(3), 249-260.

Pascual-Leone, A., Bartres-Faz, D., & Keenan, J. P. (1999). Transcranial magnetic stimulation: studying the brain-behaviour relationship by induction of 'virtual lesions'. *Achives of Philosophical Transaction of the Royal Society London B Biological Science, 354*(1387), 1229-1238.

Pascual-Leone, A., Tormos, J. M., Keenan, J., Tarazona, F., Canete, C., & Catala, M. D. (1998). Study and modulation of human cortical excitability with

transcranial magnetic stimulation. *Journal of Clinical Neurophysiology, 15*(4), 333-343.

Pascual-Leone, A., Valls-Sole, J., Wassermann, E. M., & Hallett, M. (1994). Responses to rapid-rate transcranial magnetic stimulation of the human motor cortex. *Brain, 117 (Pt 4)*, 847-858.

Pascual-Leone, A., Walsh, V., & Rothwell, J. (2000). Transcranial magnetic stimulation in cognitive neuroscience--virtual lesion, chronometry, and functional connectivity. *Current Opinion in Neurobiology, 10*(2), 232-237.

Pattyn, N., Neyt, X., Henderickx, D., & Soetens, E. (2008). Psychophysiological investigation of vigilance decrement: boredom or cognitive fatigue? *Physiology and Behavior, 93*(1-2), 369-378.

Paus, T. (2005). Inferring causality in brain images: a perturbation approach. *Archives of Philosophical Transaction of the Royal Society London Biological Science, 360*(1457), 1109-1114.

Peelen, M. V., Atkinson, A. P., & Vuilleumier, P. (2010). Supramodal Representations of Perceived Emotions in the Human Brain. *Journal of Neuroscience, 30*(30), 10127-10134.

Pell, M. D. (2006). Cerebral mechanisms for understanding emotional prosody in speech. *Brain and Language, 96*(2), 221-234.

Pell, M. D., Monetta, L., Paulmann, S., & Kotz, S. A. (2009). Recognizing Emotions in a Foreign Language. *Journal of Nonverbal Behavior, 33*(2), 107-120.

Penfield, W., & Perot, P. (1963). The Brain's Record of Auditory and Visual Experience. A Final Summary and Discussion. *Brain, 86*, 595-696.

Perrot, P., Aversano, G., & Chollet, G. (2007). Voice disguise and automatic detection: Review and perspectives. *Progress in Nonlinear Speech Processing, 4391*, 101-117 269.

Pessoa, L., Kastner, S., & Ungerleider, L. G. (2002). Attentional control of the processing of neural and emotional stimuli. *Brain Research: Cognitive Brain Research, 15*(1), 31-45.

Phillips, M. L., Senior, C., & David, A. S. (2000). Perception of threat in schizophrenics with persecutory delusions: an investigation using visual scan paths. *Psychological Medicine, 30*(1), 157-167.

Pickett, C. L., Gardner, W. L., & Knowles, M. (2004). Getting a cue: the need to belong and enhanced sensitivity to social cues. *Personality and social psychology bulletin, 30*(9), 1095-1107.

Pickup, G. J. (2006). Theory of mind and its relation to schizotypy. *Cognitive neuropsychiatry, 11*(2), 177-192.

Price, C. J., Mummery, C. J., Moore, C. J., Frakowiak, R. S., & Friston, K. J. (1999). Delineating necessary and sufficient neural systems with functional imaging studies of neuropsychological patients. *Journal of Cognitive Neuroscience, 11*(4), 371-382.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Science U S A, 98*(2), 676-682.

Ramachandra, V. (2009). On whether mirror neurons play a significant role in processing affective prosody. *Perceptual and Motor Skills, 108*(1), 30-36.

Rema, V., & Ebner, F. F. (2003). Lesions of mature barrel field cortex interfere with sensory processing and plasticity in connected areas of the contralateral hemisphere. *Journal of Neuroscience, 23*(32), 10378-10387.

Ribolsi, M., Koch, G., Magni, V., Di Lorenzo, G., Rubino, I. A., Siracusano, A., & Centonze, D. (2009). Abnormal Brain Lateralization and Connectivity in Schizophrenia. *Reviews in the Neurosciences, 20*(1), 61-70.

Rosenkranz, J. A., & Grace, A. A. (2001). Dopamine attenuates prefrontal cortical suppression of sensory inputs to the basolateral amygdala of rats. *Journal of Neuroscience, 21*(11), 4090-4103.

Roskies, A. L., Fiez, J. A., Balota, D. A., Raichle, M. E., & Petersen, S. E. (2001). Task-dependent modulation of regions in the left inferior frontal cortex during semantic processing. *Journal of Cognitive Neuroscience, 13*(6), 829-843.

Ross, E. D. (1981). The Aprosodias - Functional-Anatomic Organization of the Affective Components of Language in the Right-Hemisphere. *Archives of Neurology, 38*(9), 561-569.

Ross, E. D. (2010). Cerebral Localization of Functions and the Neurology of Language: Fact versus Fiction or Is It Something Else? [Article]. *Neuroscientist, 16*(3), 222-243.

Ross, E. D., Freeman, D., Dunn, G., & Garety, P. (2009). A Randomized Experimental Investigation of Reasoning Training for People With Delusions. *Schizophrenia bulletin*.

Ross, E. D., & Mesulam, M. M. (1979). Dominant language functions of the right hemisphere? Prosody and emotional gesturing. *Archives of Neurology, 36*(3), 144-148.

Ross, E. D., & Monnot, M. (2008a). Neurology of affective prosody and its functional-anatomic organization in right hemisphere. *Brain and Language, 104*(1), 51-74.

Ross, E. D., & Monnot, M. (2008b). Neurology of affective prosody and its functional-anatomic organization in right hemisphere. *Brain and Language, 104*(1), 51-74.

Ross, E. D., & Monnot, M. (2011). Affective prosody: What do comprehension errors tell us about hemispheric lateralization of emotions, sex and aging effects, and the role of cognitive appraisal. *Neuropsychologia, 49*(5), 866-877.

Ross, E. D., Orbelo, D. M., Cartwright, J., Hansel, S., Burgard, M., Testa, J. A., & Buck, R. (2001). Affective-prosodic deficits in schizophrenia: comparison to patients with brain damage and relation to schizophrenic symptoms [corrected]. *Journal of neurology, neurosurgery, and psychiatry, 70*(5), 597-604.

Ross, E. D., Prodan, C. I., & Monnot, M. (2007). Human facial expressions are organized functionally across the upper-lower facial axis. *Neuroscientist, 13*(5), 433-446.

Ross, E. D., Thompson, R. D., & Yenkosky, J. (1997). Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions. *Brain and Language, 56*(1), 27-54.

Rossell, S. L., & Boundy, C. L. (2005). Are auditory-verbal hallucinations associated with auditory affective processing deficits? *Schizophrenia research, 78*(1), 95-106.

Rossell, S. L., & David, A. S. (2006). Are semantic deficits in schizophrenia due to problems with access or storage? *Schizophrenia Research, 82*(2-3), 121-134.

Rossi, S., Hallett, M., Rossini, P. M., & Pascual-Leone, A. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical Neurophysiology, 120*(12), 2008-2039.

Rotarska-Jagiela, A., Oertel-Knoechel, V., DeMartino, F., van de Ven, V., Formisano, E., Roebroeck, A., . . . Linden, D. E. (2009). Anatomical brain connectivity and positive symptoms of schizophrenia: a diffusion tensor imaging study. *Psychiatry Research, 174*(1), 9-16.

Roux, P., Christophe, A., & Passerieux, C. (2010). The emotional paradox: dissociation between explicit and implicit processing of emotional prosody in schizophrenia. *Neuropsychologia, 48*(12), 3642-3649.

Rubinow, D. R., & Schmidt, P. J. (2006). Gonadal steroid regulation of mood: the lessons of premenstrual syndrome. *Frontiers in Neuroendocrinology, 27*(2), 210-216.

Russell, J. A., Bachorowski, J. A., & Fernandez-Dols, J. M. (2003). Facial and vocal expressions of emotion. *Annual Review of Psychology, 54*, 329-349.

Rymarczyk, K., & Grabowska, A. (2007). Sex differences in brain control of prosody. *Neuropsychologia, 45*(5), 921-930.

Sack, A. T., & Linden, D. E. (2003). Combining transcranial magnetic stimulation and functional imaging in cognitive brain research: possibilities and limitations. *Brain Research: Brain Research Reviews, 43*(1), 41-56.

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*(4), 848-858.

Sander, K., & Scheich, H. (2005). Left auditory cortex and amygdala, but right insula dominance for human laughing and crying. *Journal of Cognitive Neuroscience, 17*(10), 1519-1531.

Sanjuan, J., Lull, J. J., Aguilar, E. J., Marti-Bonmati, L., Moratal, D., Gonzalez, J. C., . . . Keshavan, M. S. (2007). Emotional words induce enhanced brain activity in schizophrenic patients with auditory hallucinations. *Psychiatry Research, 154*(1), 21-29.

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Science, 10*(1), 24-30.

Schirmer, A., Zysset, S., Kotz, S. A., & Yves von Cramon, D. (2004). Gender differences in the activation of inferior frontal cortex during emotional speech perception. *Neuroimage, 21*(3), 1114-1123.

Schneider, F., Gur, R. C., Gur, R. E., & Shtasel, D. L. (1995). Emotional processing in schizophrenia: neurobehavioral probes in relation to psychopathology. *Schizophrenia research, 17*(1), 67-75.

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain, 123 Pt 12*, 2400-2406.

Seitz, R. J., Schafer, R., Scherfeld, D., Friederichs, S., Popp, K., Wittsack, H. J., Azari, N.P., & Franz, M. (2008). Valuating other people's emotional face expression: a combined functional magnetic resonance imaging and electroencephalography study. *Neuroscience, 152*(3), 713-722.

Seok, J. H., Park, H. J., Chun, J. W., Lee, S. K., Cho, H. S., Kwon, J. S., & Kim, J. J. (2007). White matter abnormalities associated with auditory hallucinations in schizophrenia: a combined study of voxel-based analyses of diffusion tensor imaging and structural magnetic resonance imaging. *Psychiatry Research, 156*(2), 93-104.

Shaw, P., Bramham, J., Lawrence, E. J., Morris, R., Baron-Cohen, S., & David, A. S. (2005). Differential effects of lesions of the amygdala and prefrontal cortex on recognizing facial expressions of complex emotions. *Journal of cognitive neuroscience, 17*(9), 1410-1419.

Shea, T. L., Sergejew, A. A., Burnham, D., Jones, C., Rossell, S. L., Copolov, D. L., & Egan, G. F. (2007). Emotional prosodic processing in auditory hallucinations. *Schizophrenia Research, 90*(1-3), 214-220.

Shean, G., Bell, E., & Cameron, C. D. (2007). Recognition of nonverbal affect and schizotypy. *The Journal of psychology, 141*(3), 281-291.

Shergill, S. S., Kanaan, R. A., Chitnis, X. A., O'Daly, O., Jones, D. K., Frangou, S., Williams, S.C., Howard, R.J., Barker, G.J., Murray, R.M. & McGuire, P. (2007). A diffusion tensor imaging study of fasciculi in schizophrenia. *American Journal of Psychiatry, 164*(3), 467-473.

Siever, L. J. (2008). Neurobiology of aggression and violence. *American Journal of Psychiatry, 165*(4), 429-442.

Siever, L. J., & Davis, K. L. (2004). The pathophysiology of schizophrenia disorders: Perspectives from the spectrum. *American Journal of Psychiatry, 161*(3), 398-413.

Simosky, J. K., Freedman, R., & Stevens, K. E. (2008). Olanzapine improves deficient sensory inhibition in DBA/2 mice. *Brain Research, 1233*, 129-136.

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: review of literature and implications for future research. *Neuroscience and Biobehavioral Review, 30*(6), 855-863.

Slotnick, S. D., & Schacter, D. L. (2004). A sensory signature that distinguishes true from false memories. *Nature Neuroscience, 7*(6), 664-672.

Smeets, F., Lataster, T., Dominguez, M. D., Hommes, J., Lieb, R., Wittchen, H. U., & van Os, J. (2010). Evidence That Onset of Psychosis in the Population Reflects Early Hallucinatory Experiences That Through Environmental Risks and Affective Dysregulation Become Complicated by Delusions. *Schizophrenia Bulletin*.

Snodgrass, J. G., & Corwin, J. (1988). Perceptual identification thresholds for 150 fragmented pictures from the Snodgrass and Vanderwart picture set. *Perceptual and motor Skills, 67*(1), 3-36.

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*(3), 572-578.

Sommer, I. E., Ramsey, N. F., & Kahn, R. S. (2001). Language lateralization in schizophrenia, an fMRI study. *Schizophrenia Research, 52*(1-2), 57-67.

Spreckelmeyer, K. N., Kutas, M., Urbach, T., Altenmuller, E., & Munte, T. F. (2009). Neural processing of vocal emotion and identity. *Brain and Cognition, 69*(1), 121-126.

Stephane, M., Barton, S., & Boutros, N. N. (2001). Auditory verbal hallucinations and dysfunction of the neural substrates of speech. *Schizophrenia Research, 50*(1-2), 61-78.

Stephane, M., Thuras, P., Nasrallah, H., & Georgopoulos, A. P. (2003). The internal structure of the phenomenology of auditory verbal hallucinations. *Schizophrenia Research, 61*(2-3), 185-193.

Takahashi, H., Yahata, N., Koeda, M., Matsuda, T., Asai, K., & Okubo, Y. (2004). Brain activation associated with evaluative processes of guilt and embarrassment: an fMRI study. *Neuroimage, 23*(3), 967-974.

Tamietto, M., Latini Corazzini, L., de Gelder, B., & Geminiani, G. (2006). Functional asymmetry and interhemispheric cooperation in the perception of emotions from facial expressions. *Experimental brain research, 171*(3), 389-404.

Todd, J., Michie, P. T., & Jablensky, A. (2003). Association between reduced duration mismatch negativity (MMN) and raised temporal discrimination thresholds in schizophrenia. *Clinical Neurophysiology, 114*(11), 2061-2070.

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*(3), 361-367.

Tsoi, D. T., Lee, K. H., Khokhar, W. A., Mir, N. U., Swalli, J. S., Gee, K. A., Pluck, G., & Woodruff, P. W. (2008). Is facial emotion recognition impairment in schizophrenia identical for different emotions? A signal detection analysis. *Schizophrenia Research, 99*(1-3), 263-269.

van 't Wout, M., Aleman, A., Kessels, R. P., Cahn, W., de Haan, E. H., & Kahn, R. S. (2007). Exploring the nature of facial affect processing deficits in schizophrenia. *Psychiatry Research, 150*(3), 227-235.

van 't Wout, M., Aleman, A., Kessels, R. P., Laroi, F., & Kahn, R. S. (2004). Emotional processing in a non-clinical psychosis-prone sample. *Schizophrenia Research, 68*(2-3), 271-281.

van Hoof, J. J. M. (2002). The abnormal development of drive and guidance mechanisms in the brain: the pathogenesis of schizophrenia. *Acta Neuropsychiatrica, 14*(3), 134-146.

van Hooren, S., Versmissen, D., Janssen, I., Myin-Germeys, I., a Campo, J., Mengelers, R., vas Os, J., & Krabbendam, L. (2008). Social cognition and neurocognition as independent domains in psychosis. *Schizophrenia research, 103*(1-3), 257-265.

van Rijn, S., Aleman, A., van Diessen, E., Berckmoes, C., Vingerhoets, G., & Kahn, R. S. (2005). What is said or how it is said makes a difference: role of the right fronto-parietal operculum in emotional prosody as revealed by repetitive TMS. *European Journal of Neuroscience, 21*(11), 3195-3200

van Veelen, N. M., Vink, M., Ramsey, N. F., Sommer, I. E., van Buuren, M., Hoogendam, J. M., & Kahn, R. S. (2011). Reduced language lateralization in first-episode medication-naive schizophrenia. *Schizophrenia Research*. 127(1-3):195-201

Vanlancker, D. R., Cummings, J. L., Kreiman, J., & Dobkin, B. H. (1988). Phoonagnosia: a dissociation between famliar and unfamiliar voices. [Article]. *Cortex, 24*(2), 195-209.

Velez Feijo, A., Rieder, C. R., & Chaves, M. L. (2008). Did depressive symptoms affect recognition of emotional prosody in Parkinson's disease? *Neuropsychiatric disease and treatment, 4*(3), 669-674.

Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends in Cognitive Sciences, 9*(12), 585-594.

Walker, J. P., Pelletier, R., & Reif, L. (2004). The production of linguistic prosodic structures in subjects with right hemisphere damage. *Clinical Linguistics and Phonetics, 18*(2), 85-106.

Warren, J. D., Scott, S. K., Price, C. J., & Griffiths, T. D. (2006). Human brain mechanisms for the early analysis of voices. *Neuroimage, 31*(3), 1389-1397.

Warren, J. E., Sauter, D. A., Eisner, F., Wiland, J., Dresner, M. A., Wise, R. J., Rosen, S., & Scott, S. K. (2006). Positive emotions preferentially engage an auditory-motor "mirror" system. *Journal of Neuroscience, 26*(50), 13067-13075.

Waters, F. A., & Badcock, J. C. (2009). Memory for speech and voice identity in schizophrenia. *Journal of Nervous and Mental Diseases, 197*(12), 887-891.

Waters, F. A., Badcock, J. C., & Maybery, M. T. (2006). Selective attention for negative information and depression in schizophrenia. *Psychological Medicine, 36*(4), 455-464.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scale. *Journal of Personality and Social Psychology, 54*, 1063-1070.

Weis, S., & Hausmann, M. (2010). Sex hormones: modulators of interhemispheric inhibition in the human brain. *Neuroscientist, 16*(2), 132-138.

Weis, S., Hausmann, M., Stoffers, B., & Sturm, W. (2010). Dynamic changes in functional cerebral connectivity of spatial cognition during the menstrual cycle. *Human Brain Mapping*. In press

Weiss, E. M., Hofer, A., Golaszewski, S., Siedentopf, C., Felber, S., & Fleischhacker, W. W. (2006). Language lateralization in unmedicated patients during an acute episode of schizophrenia: a functional MRI study. *Psychiatry Research, 146*(2), 185-190.

Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., & Ethofer, T. (2008). Cerebral processing of emotional prosody--influence of acoustic parameters and arousal. *Neuroimage, 39*(2), 885-893.

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

Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., & Ackermann, H. (2005). Identification of emotional intonation evaluated by fMRI. *Neuroimage, 24*(4), 1233-1241.

Willems, R. M., de Boer, M., de Ruiter, J. P., Noordzij, M. L., Hagoort, P., & Toni, I. (2010). A Dissociation Between Linguistic and Communicative Abilities in the Human Brain. *Psychological Science, 21*(1), 8-14.

Wittfoth, M., Schroder, C., Schardt, D. M., Dengler, R., Heinze, H. J., & Kotz, S. A. (2010). On emotional conflict: interference resolution of happy and angry prosody reveals valence-specific effects. *Cereb Cortex, 20*(2), 383-392.

Woodruff, P. W., Wright, I. C., Bullmore, E. T., Brammer, M., Howard, R. J., Williams, S. C., Shapleske, J., Rossell, S., David, A.S., McGuire, P.K. & Murray, R. M. (1997). Auditory hallucinations and the temporal cortical response to speech in schizophrenia: a functional magnetic resonance imaging study. *American Journal of Psychiatry, 154*(12), 1676-1682.

Youn, T., Park, H. J., Kim, J. J., Kim, M. S., & Kwon, J. S. (2003). Altered hemispheric asymmetry and positive symptoms in schizophrenia: equivalent current dipole of auditory mismatch negativity. *Schizophrenia Research, 59*(2-3), 253-260.

Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science, 256*(5058), 846-849.

Zhang, Z. J., Hao, G. F., Shi, J. B., Mou, X. D., Yao, Z. J., & Chen, N. (2008). Investigation of the neural substrates of voice recognition in Chinese schizophrenic patients with auditory verbal hallucinations: an event-related functional MRI study. *Acta Psychiatrica Scandinavica, 118*(4), 272-280

Zinck, A. (2008). Self-referential emotions. *Consciousness and Cognition, 17*(2), 496-5054

Zinck, A., & Newen, A. (2008). Classifying emotion: a developmental account. *Synthese, 161*(1), 1-25.

ADDENDUM

The articles presented in chapters 1, 2, 5 and 6 have been submitted to international peer-reviewed journals and are currently under review. The study presented in chapter 3 has been accepted for publication in *Brain Stimulation*.

Journal articles accepted:

Alba-Ferrara, L., Ellison, A., & Mitchell R.L. (2011). Decoding emotional prosody: Resolving differences in functional neuroanatomy from fMRI and lesion studies using TMS. *Brain Stimulation*, in press.

Articles currently under review:

Alba-Ferrara, L., Fernyhough, C., Weis, S., Mitchell, R. L., & Hausmann, M. Contributions of emotional prosody comprehension deficits to the formation of auditory verbal hallucinations in schizophrenia.

Alba-Ferrara, L., Hausmann, M., Mitchell, R. L., & Weis, S. The neural correlates of emotional prosody comprehension: Disentangling simple from complex emotion.

Alba-Ferrara, L., Weis, S., Damjanovic, L., Rowett, M., & Hausmann, M. Voice identity recognition failure in schizophrenia patients with auditory verbal hallucinations.

Alba-Ferrara, L., Hirnstein, M., Weis, S., Rowett, M., & Hausmann, M. Emotional prosody modulates attention in schizophrenia patients with hallucinations.