NEUROSTRUCTURAL CORRELATES OF PROSODY IN SCHIZOPHRENIA

by

DAVID I. LEITMAN

.

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy,
The City University of New York
2006

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirements for the degree of Doctor of Philosophy.

July 20th 2006	Daniel C. Javitt, M.D., PhD.
Date	Chair of Examining Committee
August 8th 2006	Joseph Glick, PhD.
Date	Executive Officer

Charles Schroeder, PhD

John Smiley, PhD.

Dolores Malaspina, MD, MSPH

Elliot Ross MD

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

NEUROSTRUCTURAL CORRELATES OF PROSODY IN SCHIZOPHRENIA

By

David I. Leitman

Advisor: Daniel C Javitt, MD, PhD

Social cognitive impairment is an enduring and debilitating aspect of schizophrenia, an illness that affects roughly one percent of the population worldwide. This dysfunction hinders those it afflicts in their ability to integrate into society, form interpersonal relationships, and remain gainfully employed. Over the past fifty years, researchers have attempted to study these deficits by assessing emotion perception of facial and vocal gestures.

Recent attention to sensory deficits within the visual and aural modalities has raised the possibility that social cognitive impairment in schizophrenia may result from both *misperception* and *misinterpretation*. This may reflect low-level and feed-forward, as well as higher-order and "top down" dysfunction. We conducted a series of studies with the aim of relating elemental audio-sensory processing deficits to the perception of vocal affect (prosody) in patients with schizophrenia. These studies revealed the following results. First, we observed large effect size (1.6 sd's) deficits in affective prosodic perception that were associated with poor global functioning. Second, we found that within-modality sensory disturbance (specifically-pitch perception) performance, as well

as executive processing, predicted affective prosodic dysfunction. Third, we observed that both pitch perception deficits and dysprosodia in patients were associated with white matter integrity estimates within fiber pathways of auditory processing regions in the brain. Fourth, we observed that prosodic deficits extended to the perception of sarcasm and counterfactual intent in interpersonal communicatory discourse. In addition to these affect-related prosodic deficits, we also found large effect size deficits in the perception of non-affective prosody distinctions, such as decoding interrogative versus declarative intent, and recognition of differential stress patterns within speech ("stress prosody"). Like their affective counterparts, audio-sensory processing and executive processing disturbance also significantly predicted non-affective prosodic deficits.

Taken together, these findings suggest that schizophrenia dysprosodia reflects a social communicatory dysfunction that is caused by neural deficits across multiple levels of cognition. This dysfunction begins with a deficiency in perception of core elemental acoustic cues and compounds with higher cognitive dysfunction. This results in perceptual deficits of affective, as well as non–affective vocal intent in the illness.

ACKNOWLEDGEMENTS

I would like to acknowledge all my past and current teachers, who have shared their knowledge with me. In particular, I would like to thank the following few individuals, without whom I would not be writing this thesis.

I would first like to thank Rabbi Yonassan David, of Yeshivat Pachad Yitzchak, Jerusalem, Israel, whose teachings have profoundly influenced my life intellectually, morally and spiritually. Studying under him has been one of my greatest privileges, and I am grateful to him for the knowledge and wisdom he has imparted to me.

I would also like to thank Professor Israel Abramov of Brooklyn College, CUNY my undergraduate mentor, who first exposed me to science, and who, as a scientist continues to be the benchmark against which I try to measure my own progress.

I would like to thank Dr. Dolores Malaspina, of Columbia University–NYSPI, who took me as her research assistant, exposed me to schizophrenia research, trusted me with her data, and taught me many of the research and clinical skills I now employ.

Finally, I would like to thank my collaborators at the Nathan Kline Institute (NKI), I would like to thank Dr. John Foxe (also of the program in cognitive neuroscience CCNY-CUNY) for the invaluable help and support and the considerable contribution that he provided to the research contained in this thesis. I would like to thank Drs. Pamela Butler, Matthew Hoptman, Pejman Sehatpour and Manuel Gomez-Ramirez for their invaluable collaboration, as well as the clinical research staff: Gail Silipo, Nadine Revheim.

My four years at NKI have been quite productive, and this is due, in large measure to the support staff. I am greatly indebted to the assistance of Alice Saperstien, Roey

Pasternak, Marina Shpaner, Jeannette Piesco, and Beth Higgins Rachel Ziwich. I would like to especially thank Erica Saccente, for her extensive help.

I would like to thank my advisor Dr. Dan Javitt, who took a chance on a research assistant with a vague idea for research on emotion in schizophrenia, and gave him wide latitude and consistent support. I am greatly indebted to him for the patience and countless hours he spent with me on this research that is as much his as it is mine. While I am saddened that my time as his graduate student is over, I look forward to continuing to collaborate with him in the future.

Lastly, I would like to thank my parents, my first teachers, from whom I continue to learn. Their support has indispensable and unwavering.

***

This work was supported in part by National Institute of Health (NIH), National Research Service Award (NRSA) F1-MH067339 (DIL).

"…if man realized that the universe, like him, can love and suffer,

he would be reconciled."

The Myth of Sisyphus, Albert Camus (1955)

"As emotions are described in novels, they interest us, for we are made to share them"

The Principles of Psychology, William James (1890)

"…whereas the beauty of all truly scientific work is to get to ever deeper levels. Is there no way from this level of individual description in the case of emotions? I believe there is a way out, but fear that few will take it."

The Principles of Psychology, William James [on emotion] (1890)

"Physically, we transmit signals or signs-audible, visual, tactual. They may be spoken or written words, or numbers, or pictures, or many other forms of physical expression that are said to be *meaningful* or *significant*. Again, we do not send signs we *share* them, for if I tell you something, I have still got that something in my head. We now both have it-*shared*. Whereas goods are sent or exchanges messages are always *shared*.**"**

On Human Communication, E Colin Cherry (1978)

"In the fundamental notion of symbolization — mystical, practical, or mathematical, it makes no difference — we have the keynote of all humanistic problems. In it lies a new conception of 'mentality,' that may illumine questions of life and consciousness, instead of obscuring them as traditional 'scientific methods' has done."

Philosophy in a New Key, Susan Langer (1953)

TABLE OF CONTENTS

1.0 INTRODUCTION

1.0 INTRODUCTION

Schizophrenia is a debilitating disease, which includes pathological characteristics of cognitive dysfunction and impaired social functioning. The former is a primary predictor of poor global outcome and chronic disability in schizophrenia (Harvey *et al.*, 1990; Kay & Murrill, 1990). The latter, social dysfunction, has been linked with emotion recognition abnormalities (Hooker & Park, 2002). The nature of these abnormalities has received considerable attention of late, but their exact etiology remains highly enigmatic. A common way these deficits are studied is by observing schizophrenia patients' ability to produce and perceive affect using facial gestures and vocal intonation.

During the last fifteen years, emotional and cognitive deficits in schizophrenia, have been linked to "higher order" information processing and fronto-cortical and temporo-cortical abnormalities in affected individuals (Phillips, 2003; Phillips *et al.*, 2003). Researchers have also found pervasive sensory dysfunction in both the visual (P. D. Butler *et al.*, 2001; M. F. Green & Nuechterlein, 1999) and aural modalities (Javitt *et al.*, 1992; E. F. Rabinowicz *et al.*, 2000). These findings suggest that patient perception of emotional intent, for example, could result from *misperception* or distortions at the sensory level, as well as higher order *misinterpretation*, stemming from the compounded effect of sensory, as well as cognitive neurocircuitry dysfunction.

To test this hypothesis in audition, we hypothesized that pitch perception deficits would have an increasingly cumulative effect downstream on more complex cognitive tasks, such as affective appraisal in speech-based vocal intonation changes (prosody).

Four empirical studies and a preliminary study (contained in **Appendix I)** comprise the body of this thesis. These studies are as follows:

Study I: Sensory Contributions to Impaired Prosodic Processing in Schizophrenia. In which we tested affective prosody decoding in patients and healthy controls and the relationship between prosody, pitch, facial affect measures, as well as clinical symptoms and global functioning.

Study II: Neural Substrates of Impaired Prosodic Detection in Schizophrenia. In which we tested whether associations between pitch perception and vocal affect perception reflected common neurostructural abnormalities in white matter fiber pathways.

Study III: Non-Affective Prosodic Deficits in Schizophrenia and Their Relationship to Sensorial Disturbance. In which we tested whether patients have non-affective prosodic deficits in addition to affective prosodic deficits.

Study IV: Theory of Mind (ToM) and Counterfactuality Deficits in Schizophrenia: Misperception or Misinterpretation. In which we tested whether prosodic deficits extend to the perception of sarcasm and whether such deficits are predicted by sensorial disturbance.

In order to provide the proper theoretical and empirical context, it is necessary to preface these studies with a summary of current and prior research on emotion and prosody, as well as prior research on both audio-sensory processing and emotion deficits in schizophrenia.

1.1 EMOTION AND PROSODY

Human beings are social creatures. As such, they have developed complex, parallel, and redundant mechanisms for the communication of information. This information can be quite varied; from factual information like football scores to subjective experiences or feelings, like satisfaction, joy, or loss. The richest form of this communication takes place through language. The variation of information that is conveyed through language is not limited to the number of nouns and verbs or the organization of these segments into phrases or sentences, but is further enriched in the *way* in which we say it. Changes in speech rate, word emphasis or stress, and the raising or lowering of pitch enable us to color our intent with enumerable shades of meaning. Such aspects of speech are called prosody, and, given their variation and complexities, have proven difficult to study systematically (Scherer, 1986).

Recent prosody research has progressed on two fronts: theoretical and applied cognitive neuroscience. On the theoretical front, Klaus Scherer (1984) has proposed a cognitive model for affective appraisal in speech. This model is based largely on psychophysical studies conducted during the sixties and seventies by researchers such as Davitz (1964) and Scherer himself, and includes early studies by researchers such as Knowner (1941) and Skinner (1935). On the cognitive neuroscience front, the advent of neuroimaging techniques, in concert with lesion studies, has advanced our knowledge of how the brain itself processes vocal emotion.

1.1.1 Definition of Emotion

Emotion is commonly characterized as a set of specific physiological and psychological response patterns to highly prototypical stimuli or situations. These response patterns have the additional features of being quick, and not cognitively demanding. Emotions are also presumed to be heavily hard-wired, and part of our core sensory motor and cognitive genetic and evolutionary inheritance (Darwin, 1865; Ekman 1972, also see Scherer 1993, 1986).

Scherer's (1993) theory for emotion builds on Darwin's emotion theory, and further breaks emotion down into "a sequence of state changes" that comprises five-functionally defined "organismic subsystems." These systems are: 1) cognitive system (appraisal), 2) the autonomic nervous system (arousal), 3) the motor (expression) system, 4) the motivational system (action tendencies), and 5) the monitor system (feeling). This thesis examines vocal affect perception, thus it is the "appraisal" component of emotion that is of most immediate interest to us.

Within the Scherer theoretical framework, "appraisal" refers to a recursive scanning evaluation and reevaluation of one's external environment. This cyclical process is affected by synergistic changes in the external environment and internal changes within the organism. Therefore, the appraisal of a given speech utterance depends upon online scanning and evaluation of the utterance, as well as its environmental context in conjunction with internal emoto-cogntive mental dynamics. Scherer outlines (1986, 1993) the nature of this appraisal in his component process model as involving a sequence of stimulus evaluation checks.

1.1.2 Scherer's (1984-2005) Component Process Model (CPM) and Stimulus Evaluation Checks (SEC)

Scherer's (1986) CPM model assumes that emotional decoding is based on the *sequential* appraisal of an external stimulus' 1) novelty, 2) intrinsic pleasantness, 3) conductivity to goal realization, 4) coping potential or control of or adjustment potential in the particular situation and 5) congruency with internal or external (social) norms.

The great utility of Scherer's SEC model is that it makes detailed predictions about the vocal patterns associated with differing emotional states (Scherer, 1986). These predictions are based on the neurophysiological changes initiated during affective appraisal. Specifically, these changes involve the coordination of the autonomic (ANS) and sympathetic nervous system (SNS). For example, as Darwin (1865) first observed, the appraisal of a threatening stimulus such as a snake will lead to a SNS increase, causing vocal musculature constriction. This constriction results in a higher pitch (fundamental frequency) utterance.

From an information processing and cognitive neuroscience perspective, correct decoding of this raised pitch utterance as being one of fear, requires that sensory and cognitive brain regions function properly, as well as act in concert. Thus, faulty vocal affect decoding can result from the *misperception* of acoustic cues within sensory systems as well as cognitive *misinterpretation* of these cues, i.e. the inability to link the presence of these cues with the underlying emotional state.

1.1.3 Acoustic Cues Salient to Prosodic Comprehension

Much like music, speech prosody relies on the recognition and production of structured patterns of pitch, duration and intensity (Juslin & Laukka, 2003; A. D. Patel & Daniele, 2003; A.D. Patel *et al.*, 1998). These prosodic elements of speech reflect latent content in human communication, often conveying meaning as well as emotional valence not explicitly reflected in words themselves. However, due to its inherent complexity, the physiological origin of prosodic recognition remains obscure.

One of the better-studied aspects of prosody that has a clear musical analogy is the influence of the pitch contour. Somewhat akin to a melodic contour, the pitch contour is the trajectory of fundamental frequency (F_0) over time, otherwise called "intonation" (Majewski & Blasdell, 1969; Pell, 1999). For example, Majewsky and Blasdell (1969) found that the determination of a sentence as declarative or interrogative could be made solely on the basis of whether or not there was a rise in pitch at the end of the sentence. Other studies have shown that intonation modulation (F_0 variability) can be used to convey affect (Banse & Scherer, 1996; B. L. Davis *et al.*, 2000; Davitz, 1964; Juslin & Laukka, 2001; Ladd *et al.*, 1985; Pell, 1999).

Other key acoustical cues include voice intensity, high frequency energy (commonly measured as the proportion of energy above 500 Hz), stress patterns, speech rate, and attack (Juslin & Laukka, 2001; Ladd *et al*., 1985; Scherer, 1986). As Scherer (1986) has summarized, differing constellations of these cues enable one to express a wide spectrum of emotion.

1.1.4 Neural Processing of Prosody

The history of the neurophysiological investigation of prosody is inextricably linked to observations of a right hemisphere (RH) dominance for emotion processing (J.C. Borod *et al.*, 1998; J. C. Borod *et al.*, 1990; Heilman *et al.*, 1984; Heilman & Gilmore, 1998; A. Schirmer & Kotz, 2006). However, recent research has offered more nuance to this contention. This research has drawn on similarities between prosody and music to illustrate a bi-hemispheric pattern of processing.

Evidence from neuroimaging (George *et al.*, 1996; Koelsch *et al.*, 2001; Maess *et al.*, 2001; Zatorre *et al.*, 1994; Zatorre *et al.*, 1992), lesion studies (T. W. Buchanan *et al.*, 2000; Peretz, 1993; Wildgruber *et al.*, 2005; Zatorre, 1985; Zatorre *et al*., 1992; Zatorre & Halpern, 1979, 1993), and studies of neurological conditions such as Parkinson's disease (Adolphs *et al.*, 1998; L.X. Blonder *et al.*, 1989; Breitenstein *et al.*, 2001; Caekebeke *et al.*, 1991; Critchley, 1981; Darkins *et al.*, 1988; Pell *et al.*, 2006; Pell & Leonard, 2003) reveal that the same regions in the cortex involved in music comprehension are also involved in processing certain auditory components of prosody, such as intonation. These areas include the right temporal lobe, Broca's area, RH, and, in particular, Broca's right homologue (Ethofer *et al.*, 2006; Friederici & Alter, 2004; Grandjean *et al.*, 2005; A.D. Patel *et al*., 1998; A. Schirmer & Kotz, 2006; Steinhauer, 2003; Wildgruber *et al*., 2005; Zatorre *et al*., 1992).

Further research has shown that the cognitive processing of prosody is distributed between both hemispheres. Studies by Pell (1998; Pell & Baum, 1997) and others (Friederici & Alter, 2004; E. D. Ross *et al.*, 1997) demonstrate this quite convincingly. Lexical tasks and timing aspects are handled in the left frontal lobe, particularly in

Broca's area (Liegeois-Chauvel *et al.*, 1999), while pitch and intonation are handled by its right homologue (Zatorre *et al*., 1992). This increased lateralization is seen most starkly in emotional prosody. Pell (1999) and others (Juslin & Laukka, 2001) have shown that prosodic cues that convey emotion rely heavily on change in F_0.

Consequently, deficits in affective prosody in schizophrenia, rather than reflecting impairment in some generalized or multi-modal emotional circuit, may represent aberrant processing of tonal or "music-like" aspects of speech resulting from RH impairment. Further, this RH specialization for music and prosody may have early cortical antecedents; studies by Liegeois-Chauvel, *et al.* (2001) have shown neurotopical differences between hemispheres at the level of auditory cortex. While the RH displays clear spectrally tonotopic maps, these maps are less present in its left hemisphere counterpart (Liegeois-Chauvel *et al*., 2001). Such early differentiation in hemispheric structure underscores the importance of understanding the timing of vocal emotion processing.

1.1.4.1 Temporal Dynamics of Prosody

William James (1890), in his seminal theory of emotion, sees emotion as an extension of instinct and links it closely with his description of movement. He argues that emotional expression and our subsequent physiological response, "…follow directly the perception of the exciting fact, and that our feeling of the same changes as they occur is emotion…" (James, 1890). Furthermore, when describing emotion and perception he states:

> "…no reader of the last two chapters (Instinct and Movement) will be inclined to doubt the fact that *objects do excite bodily changes by a preorganized mechanism, or the farther fact that changes are so indefinitely numerous and subtle that the entire organism may be called a*

> *sounding-board*, which every change of consciousness, however slight, may make reverberate. The various permutations combinations of which these organic activities are susceptible make it abstractly possible for that no shade of emotion however slight should be without bodily reverberation as unique when taken in its totality, as is mental mood itself." (James, 1890), (pg. 450)

By linking emotion to instinct, James emphasizes that emotional information is processed quickly and by a processing system that has high response fidelity. This system, like a "sounding board" is capable of unique responses to shades of affective stimuli. Influenced by Darwin's theories (1865), James suggests that this ability is an innate function, or a "preorganized mechanism."

This description is remarkably prescient of subsequent findings in affective communication research. Extensive investigations in non-human primates have revealed that affective stimuli initiate "fixed action patterns," or instinctual-like responses. For example, macaque monkeys, when sighting a snake on the ground, will elicit a unique vocalization that, when heard by other members of the troop, will lead to heightened arousal and a fixed orientation on the ground before them (Darwin, 1865). Quick responses like these may reflect emotion-dedicated cortical pathways that serve such affective processing (Ledoux, 2000). In humans, neurophysiologists have shown that affective recognition of facial affect is processed as early as 80 milliseconds post-stimulus onset (Kawasaki *et al.*, 2001), suggesting that before one can recognize the identity of a stimulus, the subject already has qualified its affective nature.

Affective cues also play an important and early role in language acquisition. Developmental researchers, such as Fernald (1990, 1993) and colleagues, have argued that

prosodic affect comprehension precedes formal language acquisition and remains a part of speech perception throughout life .

The trajectory of these findings suggests that a substantial amount of affective processing may be automatic and pre-attentive. Additional support of this contention comes from psychophysical studies that indicate a categorical perception of both visual (facial) and vocal affect. The presence of categorical perception adds nuance to evolutionary descriptions of the temporal dynamics of emotion and prosody decoding, in that, categorical perception is evidence that affective communication, like phonemic boundaries, is learned shorthand and not strictly a hard-wired subroutine.

Scherer's (CPM) model of vocal emotion explicitly appreciates pre-attentive processing of affective stimuli. Recently, he has proposed that the first stimulus evaluation check (SEC)--the "novelty check," likely occurs pre-attentively (Sander *et al.*, 2003).

An ideal tool to test this hypothesis electrophysiologically, is by using the event-related potentials (ERP), which give real-time estimates of cortical temporal dynamics. In particular, the mismatch negativity (MMN) has proven useful for indexing auditory scene novelty detection that is not dependent upon attention.

1.1.4.2 MMN as a Mechanism for Examining Pre-Attentive Prosodic Perception

MMN is a pre-attentive auditory ERP elicited most commonly in the context of an auditory "oddball" paradigm, in which a series of standard stimuli is interrupted by an infrequent deviant stimulus. In such a paradigm, the brain automatically organizes a template reflecting invariant features of the repetitive standards. MMN then reflects the outcome of a local process that compares each stimulus to the locally maintained template.

MMN usually occurs with a latency of approximately 100 ms, which may vary based upon degree of stimulus deviance. One of the characteristic features of MMN is that latency is locked to the timing of feature deviance, rather than to stimulus onset. Thus, MMN to duration deviance is delayed relative to pitch deviance, because pitch deviance can be directed at stimulus onset, whereas duration deviance (for longer deviant stimuli) cannot be detected until the normal offset of the standard stimulus. A primary role of MMN is to direct attention toward potentially significant alterations to the surrounding acoustic environment (Naatanen & Michie, 1979).

Complex changes and violations of abstract patterns can also elicit mismatch negativity. When series of tones, both of differing pitch and intensity, are uniquely paired so that no two tones are identical in acoustical features, mismatches are elicited. Violations of abstract patterns in which the deviance detection occurs, not because of any physical change in stimulus, but rather, because of the perceived relationship between stimuli, also elicit MMN. For example, MMNs are generated when consistently ascending tone pairs (such that the second tone in a pair is higher than the first) are infrequently punctuated by a descending deviant pair. Further, within language, phonemic changes also elicit MMNs pre-attentively. These findings suggest that a great deal of feature analysis within the auditory domain may in fact be pre-attentive in nature[*].

In an attempt to gauge the sensitivity of MMN generators to suprasegmental information, we constructed artificial stimuli that approximated the F_0 modulations that distinguish interrogative and declarative intent (see **Appendix I**). This was done by string-

[*] MMN mechanisms are further elaborated upon below in our discussion (Section 1.2) of audio-sensory memory disturbance in schizophrenia.

ing together pure tones with the frequency and duration properties of the annunciated syllables of a semantically neutral sentence, as used by Pell (1998).

Our analysis revealed that MMNs were elicited whether repetitive interrogative stimuli were presented against a background of declarative stimuli, or whether declarative stimuli were presented against a background of interrogative stimuli. In both cases, comparisons were made between the same stimulus (i.e., interrogative/declarative) presented as a standard and the same stimulus presented as a deviant in a separate run. Thus, the spectral content of standards and deviants within each condition was identical. Further, because of the symmetrical experimental design, the onset of deviance within each run (i.e. interrogative std/declarative deviant vs. declarative std/interrogative deviant) was identical, with the contours being identical up to 328 ms and diverging thereafter.

Despite the exact symmetry in the experimental design, significant differences were observed in both the timing and amplitude of contour-elicited MMNs, such that interrogative contours presented against a background of declarative contours elicited a larger MMN than declarative contours presented against a background of interrogative standards. This finding suggests that the MMN generators are responding not just to contour per se, but also to the ecological significance of the suprasegmental information contained within the contour. This is consistent with observations in social situations, in which individuals are able to detect when they are being asked a question, even after they have stopped paying close attention to the verbal material directed at them.

In the sole MMN study examining affective prosody, Kujala *et al.* (2005), using actual affective tokens, recently demonstrated MMNs to deviant presentations of "com-

manding", "sad", and "scornful" prosodic presentations of single words when contrasted with neutral standard presentations.

This research, while admittedly preliminary, suggests that MMN generators may be responsive to alterations in prosody, suggesting that prosodic information may be decoded, at least in part, in an attention-independent fashion, perhaps as a mechanism to trigger higher-level attention.

In summary, the neural dynamics of affective prosodic comprehension can be conceptualized as involving a "three-stage processing chain" (A Schirmer & Kotz, in press) that begins with sensation (stage 1) in primary auditory cortex . It then continues with integration (stage 2) within ventral aspects of temporal cortex and superior temporal sulcus (STS). There, aspects of the acoustical information are tagged as affective. Processing finally proceeds to the cognitive stage (stage 3) in inferior frontal regions, where this information is evaluated both semantically and contextually (reviewed in (A Schirmer & Kotz, in press)).

1.2 AUDIO-SENSORY DISTURBANCE IN SCHIZOPHRENIA

Since the advent of Carlsson's (1977) dopamine hypothesis for schizophrenia, clinical researchers have tended focus on cognitive disturbances, as reflected in frontal lobe dysfunction and executive control processes (Weinberger & Gallhofer, 1997). However, in the early sixties, observations of perceptual distortions within the visual and auditory modality by researchers like Chapman (1966), Lang (1965), and Ornitz (1969), led some to posit that patients have modulatory deficits in their ability to consistently amplify sensory signals (Ornitz, 1969). This hypothesis has been recently reasserted and given

greater credence by research on NMDA-related glutamate dysfunction (Javitt, 1987; Javitt & Coyle, 2004). Further support for this hypothesis can be garnered from recent behavioral, as well as electrophysiological research in the aural domain, which supports this view (see below).

1.2.1 Behavioral Studies

The best-studied auditory sensory deficits are those of pitch and duration. Deficits in basic pitch perception in schizophrenia manifest themselves behaviorally in simple tasks of tone matching, in which subjects discern whether pairs of pure tones separated by a short inter-stimulus interval differ in pitch. ASM stores representations of the simple physical properties of presented stimuli for periods of seconds to tens of seconds (Cowan, 1984). Using such paradigms, schizophrenic subjects have been shown to perform extremely poorly on tests of auditory discrimination (Javitt *et al.*, 1997; R.D. Strous *et al.*, 1995a). However, when task difficulty was decreased by increasing pitch separation between stimuli (Δf), retention of information appeared to be unimpaired in schizophrenia (Javitt *et al*., 1997). This strongly suggests that ASM in schizophrenia functions similarly to that of controls, except that a reduced precision of the pitch information is retained or utilized. Further, this deficit does not seem to reflect working memory dysfunction, as Rabinowicz *et al*. have shown that patients and controls show a similar rate of performance (accuracy) fall off when tone pair sequences are interrupted by distractor noise (E. F. Rabinowicz *et al*., 2000).

1.2.2 Electrophysiology Studies

Alternative electrophysiological methods utilizing cognitive event-related potentials (ERP) have used the "oddball" paradigm, in which a sequence of repetitive standard stimuli is interrupted infrequently by physically deviant "oddballs" of either pitch or duration. Results have shown that mismatch negativity (MMN), an ERP occurring at about 100 milliseconds after deviance onset, reflects pre-attentive cognitive processing of stimulus deviation. When the difficulty of task discrimination is increased by decreasing Δf, latencies of the MMN and behavioral responses increase in parallel, indicating that pre-attentive processes indexed by MMN govern attentive novelty detection (Novak *et al.*, 1992a, 1992b). Numerous studies have demonstrated that generation of MMN is decreased in schizophrenia (Javitt *et al.*, 1993; Javitt *et al.*, 1995b; Oades *et al.*, 1997; Umbricht & Krljes, 2005), although contrary results have also been reported (O'Donnell *et al.*, 1994). MMN in the "oddball" paradigm appears to give an electrophysiological index of ASM functioning (Cowan *et al.*, 1993; Ritter *et al.*, 1995), as both MMN and ASM appear to function in a fashion that is relatively independent of attention (Cowan *et al.*, 1993). The critical structures for ASM and MMN generation reside in the superior temporal lobe, specifically auditory cortex (Winkler *et al.*, 1995).

The aforementioned studies suggest the following. Impaired representational precision of tones in schizophrenic subjects, as measured by performance in auditory discrimination tasks, indicates primitive cortical processing abnormalities. Similar deficits in representational precision have been observed in studies of visual perceptual organization (E.F. Rabinowicz *et al.*, 1996) and visio-spatial processing (Fleming *et al.*, 1997), which may reflect a general principle underlying cognitive dysfunction in schizophrenia. If this

is the case, primitive disturbances within the primary auditory cortex, such as the generation of ASM, may be the seeds for higher-order cognitive dysfunction. Thus, inability to reliably generate precise ASM may lead to higher-order deficits in processing more complex auditory stimuli, such as prosodic cues in speech.

1.3 AFFECTIVE DISTURBANCE IN SCHIZOPHRENIA

Close to a century ago, Eugen Bleuler noted that "In the outspoken forms of schizophrenia, the emotional deterioration stand is the forefront of the clinical picture" (Bleuler, German-1911: Trans-1950). Yet, the nature of this symptom has remained enigmatic.

Clinically, these symptoms are seen as a cessation to display or to recognize affect in response to affective triggers or within the context of social communication. Numerous studies using clinical assessments have described emotional blunting in patients (Berenbaum *et al.*, 1987; Boeringa & Castellani, 1982; de Leon *et al.*, 1993; Fenton & McGlashan, 1992; Kukla & Gold, 1991; Shaheen & Ibrahim, 1979). Principally, these studies have indicated that, using a variety of scales, patients' premorbid functioning was strongly associated with poverty of affect and speech, as well as measurers of "indifference", "anhedonia", "avolition" and "asociality" (de Leon *et al*., 1993). These symptoms tend to cluster into a constellation that has been termed "negative symptoms" (N.C. Andreasen, 1984; N. C. Andreasen *et al.*, 1995), which seems to describe a distinct dimension of the illness. However, it is worthwhile to note that blunted affect or avolition within the negative symptom assessment measures, like the Scale for the Assessment of Negative Symptoms (SANS) (N.C. Andreasen, 1984), and ratings of "anhedonia" and

"avolition" do not fully capture the magnitude of affective disturbance in the illness. Exaggerated or improper affective responses are also commonly manifested, both in terms of facial and vocal gestures, yet these are overlooked by the SANS (de Leon *et al*., 1993).

Despite the evidence of less affective reactivity in patients, one may still ask whether patients with schizophrenia feel or conceptualize emotions. Extensive study of these questions by Kring and colleagues (Kring *et al.*, 1993; Kring & Neale, 1996) has persuasively shown that schizophrenia patients have intact conceptions of emotion and do indeed feel them. However, unlike their healthy counterparts, they are less likely to express them spontaneously (Kring & Neale, 1996).

1.4 AFFECTIVE PROSODY IN SCHIZOPHRENIA

1.4.1 Is There a Deficit?

Unlike facial affect recognition, prosodic recognition of affect in schizophrenia has received comparatively less attention. A recent review of this literature by Edwards (Edwards *et al.*, 2002) found seventeen studies with mixed results that, overall, support the assertion of modest, but enduring deficits in affective prosodic recognition, and more robust expressive deficits. However, our studies (see studies 1, 3 and 4) have consistently shown large effect size (sd ≈ 1.6) deficits in chronic patients with severe negative symptoms. A recent comparison of chronic and early stage patients found more severe prosodic deficits in chronic than early stage patients (Kucharska-Pietura *et al.*, 2005). This between-group study, while possibly confounded by medication effects, nevertheless suggests that prosodic deficits may have a developmental time course (Kucharska-Pietura *et al*., 2005).

1.4.2. Dysprosodia as an Illness Trait

A review of the affective prosody literature considerably suggests that dysprosodia may be a trait of schizophrenia. In addition to studies showing prosodic deficits in chronically ill patients (Kee *et al.*, 2003; Kee *et al.*, 1998; Kerr & Neale, 1993; Leitman *et al.*, 2005), prosodic deficits have been found in first-episode schizophrenics (Edwards *et al.*, 2001; Haskins *et al.*, 1995), as well as in psychotic children. Baltaxe and Simmons (1995) found that, of 47 children with schizophrenia referred to speech therapists, 81% had significant prosodic deficits.

Aprosodia within schizophrenia has also been shown to be unrelated to medication dose or illness length (Kerr & Neale, 1993; Leitman *et al.*, 2005), although contrary results have also been found (Kucharska-Pietura *et al*., 2005).

The preponderance of prosodic deficits in both child-onset schizophrenia (Baltaxe & Simmons, 1995), adult-onset schizophrenia (Kerr & Neale, 1993), and first episode psychosis (Edwards *et al*., 2002), suggest that prosodic impairment may indeed be an intrinsic illness trait (Kucharska-Pietura *et al*., 2005; Leitman *et al*., 2005). This suggestion is even more provocative, given evidence that prosodic comprehension is developmental in nature (Fernald; Kuhl, 2004) (also see above in Emotion and Prosody subsection 1.1). Given the convergence between prosodic impairment in schizophrenia and social developmental nature of prosody acquisition, a better understanding of prosodic deficits in schizophrenia and development should prove valuable to an etiological understanding of social dysfunction in the illness. As E Colin Cherry (1978) once quipped, when stressing the social nature of speech communication, "…a child learns to imitate the speech sounds

of its mother; it does not learn to make sounds like bells or frying bacon" (ibid, pg. 265). More recently, prominent speech scientists, like Patricia Kuhl (2005) and Michael Goldstein (2003), have proposed social feedback loops that drive sensory and motor speech acquisition, in which prosody serves as a salient cue. This social component to language acquisition has already received considerable attention in the autism field (e.g., (Kuhl *et al*., 2005)) While studies of prosody are underway in prodromal adolescents, retrospective evidence on child development of those who later develop schizophrenia is, to our knowledge, unstudied.[1]

1.4.3 Dysprosodia and Functional Outcome

Functioning within society requires one to communicate socially, whether it be buying groceries or interacting with coworkers. Thus, an inability to perceive emotion could likely lead to social isolation and poor functioning within society. Within schizophrenia, studies such as those by Kee and colleagues (2003) and others (J. Brekke *et al.*, 2005a; J. S. Brekke *et al.*, 2005b) have linked both facial and vocal emotion perception to functional outcome. This was done using the Strauss and Carpenter Outcome Scale and Role Functioning Scale, which measures patients' ability to gain and hold employment, as well as to live independently. In a slightly different vein, a study by Hooker and Park (2002) found that both face and vocal emotion perception, unlike comparative non-affective tasks, correlated with measures of social functioning. Finally, in Study 1 of this thesis, we found that dysprosodia alone, is directly related to functional outcome using the Independent Living Scale (ILS).

[1] The only study known to us that touches upon this issue is Walker and Lewine's (1990) classical study of autobiographical media recordings, in which observations of poorer affective reactivity, as well as gross and fine motor control, seemed to foreshadow subsequent diagnosis.

1.4.4 Expressive and Receptive Prosody

The connection between expressive and receptive prosodic abilities in schizophrenia is unclear. Alpert *et al.* (2000) showed greater impairment in prosodic expression (rather than comprehension) using acoustic analysis of select prosodic cues of patients' speech, when describing a pleasant or unpleasant experience. These studies examining spontaneous prosody echo other spontaneous prosody findings (N.C. Andreasen *et al.*, 1981; Fricchione *et al.*, 1986; Haskins *et al.,* 1995), all of which found significant expressive deficits that were unrelated to medication dose. Interestingly, studies (Shaw, 1999; Whittaker, 1994) that have compared spontaneous and receptive affective prosody, within the same sample of patients, have found no correlation between expressive and receptive prosody. The Haskins (1995) study, in which patients were additionally asked to repeat prosodic utterances, found only a minimal correlation between raters' judgments of prosodic similarity and receptive prosodic ability. This supports the conjecture that affective vocal perception may not be a rate-limiting factor in prosodic expression, and may have more prominent motor-related dysfunction contributions (M. Alpert *et al.*, 2000; Haskins *et al*., 1995).

1.4.5 Neural Locus of Prosodic Deficits

Studies that have suggested neural correlates for prosodic deficits have fallen into two categories: those that emphasize RH dysfunction, and those that suggest more widespread temporo-limbic disturbance. Based on the combined findings of affect recognition deficits across multiple channels, and in conjunction with data from lesion studies, Borod

(Joan C. Borod, 2000; J.C. Borod *et al*., 1998; J. C. Borod *et al.*, 1993; J. C. Borod *et al*., 1990; Martin *et al.*, 1990) has attributed schizophrenia dysprosodia to right hemisphere dysfunction, particularly, right dorsolateral (DLPFC) and inferior frontal cortex (IFC) dysfunction. Other researchers (Phillips, 2003; Phillips *et al*., 2003) acknowledge these brain regions, but emphasize fronto-temporo-limbic connections (Abdi & Sharma, 2004; R. W. Buchanan *et al.*, 1993; Kosaka *et al.*, 2002; Kucharska-Pietura *et al.*, 2003), particularly implicating amygdala and medial temporal function.

The large majority of these studies that used facial affect paradigms (Exner *et al.*, 2004; Gur *et al.*, 2002; Holt *et al.*, 2005; Johnston *et al.*, 2005; Kosaka *et al*., 2002; Phillips *et al.*, 1999) found hemodynamic abnormalities within these regions. The sole study to use vocal affect in schizophrenia (Mitchell *et al.,* 2005), found no significant differences in bold responses between patients and healthy controls. However, these results must be qualified by the lack of significant differences in behavioral prosody performance between groups (Mitchell *et al.*, 2004), which, given the general literature finding, suggests that either the patient group was non-typical or that the task was not adequately challenging.

Whether emphasizing right frontal cortex and/or temporo-limbic regions, these studies implicitly treat the presence of vocal and facial affect decoding deficits as symptoms of unitary emotion dysfunction. This approach is problematic, as investigations of cross-modal affective processing have revealed only modest correlations (Kerr & Neale, 1993), or no correlations (Leitman *et al.,* 2005) between aural and visual modalities. Moreover, they ignore the potential contribution of significant perceptual deficits within each individual modality (Johnston *et al*., 2005; Kee *et al*., 1998).

1.5 CONTRIBUTION OF THIS THESIS

A century of concentrated study of schizophrenia has revealed that afflicted individuals have a myriad of both cognitive and emotional disabilities. However, the link between these twin dysfunctions remains elusive.

This thesis rests on clinical observations that patients did not seem to perceive affect when conveyed by vocal inflection and a strong body of literature indicating elemental audio-sensory processing deficits in schizophrenia audition. We hypothesized that sensory dysfunction could upwardly generalize into cognitive and emotional deficits. The implications for such upward generalization would further efforts toward providing a unifying theory for cognitive and emotional deficits in the illness, by suggesting that dysfunction within both of these domains stems from deficiencies in the perceptual system employed by both.

1.5.1 Limitations of Previous Studies

While upward generalization of sensory deficits has been studied in the visual modality (J. Brekke *et al*., 2005a; P. D. Butler *et al*., 2001; M. Green & Walker, 1986; M.F. Green *et al.*, 1994; Kee *et al*., 1998), this approach has not been extensively employed in the auditory domain. However, decades ago, researchers such as Turner (Turner, 1964) and Jonsson (Jonsson & Sjostedt, 1973) suggested that such deficiencies might underlie deficits like prosody. Previous studies have also largely ignored non-affective prosody; but a sensory approach suggests that prosodic deficits may not be limited to affect.

1.5.2 Research Questions

In light of these limitations, these current studies examine the size and magnitude of affective prosodic deficits and their relationship to measures of executive and audio-sensory processing, clinical symptomatology, and functional outcome. We further sought to examine the extent of prosodic defects and whether they extend to sarcasm and non-affective prosodic distinctions like interrogative intent. Additionally, we sought to examine the neurostructural correlates of schizophrenia dysprosodia, by examining correlations between prosodic detection and white matter pathway integrity.

2.0 STUDY I

Sensory Contributions to Impaired Prosodic Processing in Schizophrenia

2.1 ABSTRACT

Background: Deficits in affect recognition are prominent features of schizophrenia. Within the auditory domain, patients show difficulty in interpreting vocal emotional cues based on intonation (prosody). The relationship of these symptoms to deficits in basic sensory processing has not been previously evaluated.

Methods: Forty-three patients and 34 healthy comparison subjects were tested on two affective prosody measures: voice emotion identification and voice emotion discrimination. Basic auditory sensory processing was measured using a tone-matching paradigm and the Distorted Tunes Test (DTT). A subset of subjects was also tested on facial affect identification and discrimination tasks.

Results: Patients showed significantly impaired performance on all emotion processing tasks. Within the patient group, a principal components analysis demonstrated significant intercorrelations between basic pitch perception and affective prosodic performance. In contrast, facial affect recognition deficits represented a distinct second component. Prosodic affect measures correlated significantly with severity of negative symptoms and impaired global outcome.

Conclusions: These results demonstrate significant relationships between basic auditory processing deficits and impaired receptive prosody in schizophrenia. The separate loading of auditory and visual affective recognition measures suggests that within-modality factors may be more significant than cross-modality factors in the etiology of affect recognition deficits in schizophrenia.

2.2 INTRODUCTION

Schizophrenia is associated with deficits in both higher order cognition (Braus *et al.*, 2002) and early sensory processing (P. D. Butler *et al.*, 2001; Foxe *et al.*, 2001; Javitt *et al.*, 2000). Although these two types of deficits are most often studied in isolation, the possibility exists that deficits in early sensory processing may contribute greatly to deficits in higher order cognition. In schizophrenia, one of the most significantly impaired functions is the ability to decode emotion based on either facial expression or speech intonation (M. Alpert *et al*., 2000; Edwards & McGorry, 2002; Edwards *et al*., 2001; Gur *et al*., 2002; Haskins *et al*., 1995; Kerr & Neale, 1993; E. D. Ross *et al.*, 2001a). These deficits have been attributed to deficits in emotion processing brain regions (Edwards *et al*., 2001; Gur *et al*., 2002), although the basis for the dysfunction remains to be determined. The present study evaluates the degree to which deficits in ability to decode emotion, especially in the auditory modality, depend upon more basic deficits in early sensory processing.

Prosody refers to our ability to recognize, comprehend, and produce affect as well as semantic meaning based on the intonation, stress, and rhythm patterns of vocal utterances. Emotional prosody refers to the ability to detect affect and infer emotion based on prosodic information, while semantic prosody refers to the ability to differentiate meaning, for example, differentiating questions from answers. Receptive prosody refers to the ability to decode prosodic information in statements made by others, while expressive prosody refers to the ability to express emotion or other prosodic information in our own utterances. Jonsson and Sjostedt (1973) first demonstrated deficits in receptive emotional prosody in schizophrenia. Subsequent studies demonstrated deficits in semantic, emo-

tional, and expressive prosody (e.g., (M. Alpert & Anderson, 1977; Fricchione & Howanitz, 1985; Kerr & Neale, 1993). Deficits in receptive prosody are seen in both chronic and first episode patients (Edwards *et al*., 2001; Haskins *et al*., 1995), as well as in children with schizophrenia (Baltaxe & Simmons, 1995). These deficits occur independently of medication (Kerr & Neale, 1993; E. D. Ross *et al*., 2001a), suggesting that these deficits represent a trait aspect of the illness (Edwards & McGorry, 2002) and not its treatment.

Within the realm of sensory perception, multiple studies have shown basic pitch perception and auditory sensory memory (i.e., tone matching) deficits in individuals with schizophrenia. These sensory deficits may have an important influence on higher-order processes such as prosodic comprehension, since, much like music, speech relies on the production and recognition of structured patterns of pitch, duration, and intensity. Further, much like elemental pitch processing, prosodic functioning relies substantially on right hemispheric function (A.D. Patel *et al*., 1998). We hypothesize that the presence of elemental pitch deficits in schizophrenia, which we have demonstrated previously (Javitt *et al*., 1999; E. F. Rabinowicz *et al*., 2000; R.D. Strous *et al*., 1995a), may significantly contribute to the deficits in affective prosodic functioning that are seen in patients with schizophrenia.

To test this hypothesis, we collected behavioral measures of affective prosodic comprehension in patients and comparison subjects. Within the patient group, we collected psychophysical measures of pitch perception, to both music and pure tones. To further illustrate the specific contribution of pitch perception to affective prosodic comprehension, measures of affective facial comprehension were also collected as a control con-

dition.

2.3 METHODS

2.3.1 Participants

Forty-three stable patients meeting DSM-IV criteria for either schizophrenia or schizoaffective disorder and 34 healthy control subjects volunteered to serve in this experiment. The Institutional Review Board of the Nathan Kline Institute for Psychiatric Research approved all experimental procedures, and all patients were recruited from facilities associated with the Institute. Written, informed consent was provided by all subjects after the procedures of the experiment were fully explained. Both patients and healthy comparison subjects received $10/hour for participation.

Diagnoses were obtained using the Structured Clinical Interview for DSM-IV (SCID), performed by Masters or Doctoral level psychologists, psychiatrists, or trained diagnostic technicians, using all available clinical information. Thirty-three patients met criteria for schizophrenia and 10 met criteria for schizoaffective disorder. Twenty-four of the patients were receiving only second-generation antipsychotics, primarily risperidone or olanzapine, 3 patients were receiving (haloperidol), and 16 patients were receiving a combination of antipsychotics. The mean daily antipsychotic dose was 1068.3 ± 423.4 mg chlorpromazine equivalents, using Hyman and Arana (Hyman & Arana, 1987) and "best estimate" conversion factors for new antipsychotic medications. Mean illness duration was 17.4 ± 9.6 years. The healthy control group consisted of staff volunteers as well as individuals who responded to local advertisements. Handedness of all subjects was

assessed using the Edinburgh Handedness Inventory (Oldfield, 1971). See **Table 1** for further demographic information.

2.3.2 Auditory Sensory Processing

Two tests of auditory sensory processing were employed: a tone matching task, which reflects processing within primary auditory regions (Liegeois-Chauvel *et al.*, 2001; Zatorre & Samson, 1991), and the distorted tunes task (Drayna *et al.*, 2001), which reflects processing within the unimodal sensory association cortex, especially the right hemisphere (Zatorre, 1985; Zatorre *et al.*, 2002). These measures were obtained for patients only.

Tone Matching Task (TMT)**.** A simple tone-matching paradigm was employed, as described previously (R.D. Strous *et al*., 1995a). Tones were generated on a Pentium personal computer (PC; Winbook, Hilliard, Ohio) with SoundBlaster Pro audio card (Creative Sound Systems, Milipitas, California) using the Neuroscan Stim (Neurosoft; Compumedics USA, El Paso, Texas) software, and were presented binaurally through headphones at nominal intensity level of 70 dB sound pressure level (SPL). Subjects were presented with pairs of 100-millisecond tones in series, with a 500-millisecond intertone interval. Within each pair, tones were either identical or differed in frequency by specified amounts in each block (2.5%, 5%, 10%, 20%, or 50%). In each block, half the tones were identical and half were dissimilar. Subjects responded by pressing one of two keys to answer whether the pitch was the same or different. Tones were derived from three reference base frequencies (500, 1000, and 2000 Hz), to avoid learning effects. In all, the test consisted of five sets of 26 pairs of tones and took approximately 20 minutes

to complete. Participant performance across these five levels was averaged and this score was used for analysis.

The Distorted Tunes Task (DTT). (Drayna *et al*., 2001) Consists of 26 familiar tunes ranging in length from 12 to 26 notes. Seventeen of the tunes are rendered melodically incorrect by changing the pitch of two to nine notes within the tune. Subjects respond "yes" or "no," whether they think that the melody is correct. Subject scores are calculated based on the percentage of correctly categorized melodies.

2.3.3 Affective Prosody

Two basic tests of affective prosodic processing were employed: 1) the Voice Emotion Identification Test (VOICE-ID), and 2) the Voice Emotion Discrimination Test (VOICE-DISCRIM) (Kerr & Neale, 1993).

Voice Emotion Identification Test. This test consists of 21 sentences of neutral content on audiotape (e.g., "He tossed the bread to the pigeons", "The boy went to the store"). The sentences are spoken aloud by male and female speakers to convey one of six different emotions (happiness, anger, fear, sadness, surprise, or shame). Participants are given a piece of paper with the six emotions listed. They are asked to listen to each sentence and to tell the experimenter which of the six emotions best describes the speaker's tone of voice. Participants are advised to guess if unsure.

Voice Emotion Discrimination Test. This test contains 35 separate pairs of sentences. Each pair contains sentences that consist of the same words (i.e. "The game ended at 4 o'clock", "The game ended at 4 o'clock") or different words (i.e., "The boy went to the store", "He tossed the bread to the pigeons"), and the second sentence is read with the

same or different prosody (emotion) as the first. Participants are asked to focus on the mood rather than the content of the sentences and to tell the experimenter whether the sentences are said in the same or in a different emotion. Participants are advised to guess if unsure**.**

2.3.4 Visual Affective Processing

Two basic tests of face emotion processing were employed: 1) the Face Emotion Identification Test (FACE-ID), and 2) the Face Emotion Discrimination Test (FACE-DISCRIM), (Kerr & Neale, 1993). These measures, which were added partway through the study and were available for only a subset of participants, use black and white photographs of faces displaying different emotions created by Izard (1971) and Ekman (1976).

Face Emotion Identification Test. This test consists of 19 photographs of facial emotions presented on videotape for approximately 15 seconds with a blank screen of approximately 10 seconds between pictures. Participants are asked to look at each face and tell the experimenter which of six emotions best describes the emotion in the photograph. The same six emotions are used as in the VOICE-ID test, and participants are again given a list of the emotions. Participants are asked to guess if unsure.

Face Emotion Discrimination Test. This test consists of 30 pairs of photographs. Each pair of photographs is presented simultaneously. Participants are asked to tell the experimenter whether the two people in each pair are displaying the same or different emotions on their faces. Participants are asked to guess if unsure.

2.3.5 Clinical Measures

Clinical measures included the Brief Psychiatric Rating Scale (BPRS) which included the anxiety/depression (items: somatic concern, anxiety, guilt feelings, depression) and positive symptom (items: disorganization, suspiciousness, hallucinatory behavior, unusual thought content) subscales (Overall & Gorham, 1961), as well as the Scale for the Assessment of Negative Symptoms (SANS), which used total score without globals as a summary measure (N.C. Andreasen, 1982) . In addition, the Independent Living Scales problem-solving factor subscale (ILS-PB) (Loeb, 1996) was employed as a measure of the patients' global functioning. The ILS-PB is comprised of 33 items related to money management, home/transportation, and social adjustment. Questions such as, "Tell me two reasons why it is important to pay your bills," "What might you do if both your lights and your TV went off at the same time," and "Why is it important to know about the side effects of the medicine you are taking" are used to elucidate effective strategies and ways to negotiate life on a day-to-day basis. The ILS-PB scores have been shown to predict a patient's functional living status (inpatient vs. outpatient) more strongly than verbal memory or other measure (Revheim & Medalia, 2004). Raw scores on the ILS-PB were converted into standard scores, per the instructions found in the manual, and were then used as variables in the present study.

2.3.6 Statistical Analysis

Primary analyses were performed on all measures using analysis of variance (ANOVA) using the variables gender (M/F) and group (patient/control). In addition, the relative magnitude of deficits between tests was assessed using a task x-group analysis. Effect sizes were expressed in SD units and interpreted according to criteria by Cohen

(1988). Correlation between measures was assessed using Spearman correlations. In addition, within the patient group, principal components analysis (PCA) was used to investigate the interrelationship between sensory processing and prosodic measures. The principal components analysis was restricted by the following criteria: 1) using the Kaiser criterion, eigenvalues that were ≥|1| were selected as components for rotation, and 2) component loading was considered significant only if it was ≥.6. All statistical analyses were performed using the JMP statistical software package (Academic Version 4.0.4; SAS Institute, Inc., Cary, North Carolina).

Two-tailed tests were used throughout, with preset α level of significance of $p \leq .05$. Data in text reflect mean ± sd.

2.4 RESULTS

2.4.1 Between-Group Analyses

Tone Matching Task and DTT performance were obtained for patients only. Tone Matching Task performance is shown in Table 2 and is similar to that obtained in previous studies in similar patient populations (Javitt *et al*., 1999; E. F. Rabinowicz *et al*., 2000; R.D. Strous *et al*., 1995a). Mean performance across all five levels was 77 ± 14%. In a prior study, we tested performance on the three middle levels (5%, 10%, and 20%) of the five used in the present study (R.D. Strous *et al*., 1995a). In this study, patients scored a mean of 72% correct, versus 93% correct for control subjects, with the between-group difference being highly significant and the effect size being on the order of 1.2 sd units. Mean performance on the DTT task was 69 ± 22% correct, with 66% of patients scoring below 88% correct, which is considered the threshold for "tune deafness" (Drayna *et al*.,

2001). In a prior study of healthy monozygotic and dizygotic twins, only 39.6% met this criterion ($z = 2.72$, $p < .006$, vs. present sample) (Drayna *et al*., 2001). Affective prosodic processing and visual affective processing measures were obtained for both patients and control subjects. As expected, patients' performances were significantly poorer than that of control subjects on all four affect discrimination and identification measures (**Table 3**). Prosodic measures yielded the most substantial deficits, with VOICE-DISCRIM scores 23 percentage points lower and VOICE-ID scores 33 percentage points lower for patients than control subjects. Less pronounced but similar results were observed in the visual modality, where FACE-DISCRIM and FACE-ID scores were reduced by 14 and 25 percentage points, respectively. There were no significant main effects for gender (all $p \geq .05$).

To assess relative magnitude of auditory and facial affect recognition deficits in patients, a repeated measures ANOVA was conducted across visual and auditory measures. Significant differences between patients and control subjects were observed for both the VOICE-ID/FACE-ID ($F = 6.84$, $df = 1,42$ $p < .01$), and the VOICE-DISCRIM/FACE-DISCRIM ($F = 26.56$, $df = 1,43$, $p < .001$) paired measures. However, there were no significant modality-by-group interactions for either the DISCRIM ($F = .87$, $df = 1,43$, $p = .32$), or the ID ($F = .99$, $df = 1,42$, $p = .38$) measures, thus indicating that both modalities were affected to a similar degree. Effect sizes for both effects were large (sd = 0.8–1.7), using Cohen's criteria (1988).

2.4.2 Within-Group Analyses

2.4.2.1 Correlations Among Measures and Clinical Ratings

Correlation analyses were performed to analyze the relationship between auditory and visual affective processing measures and clinical symptoms. VOICE-ID performance correlated significantly with scores on the BPRS conceptual disorganization factor [r_s (39) = .48, p < .01], SANS total score [r_s (39) = .41, p = .01], and ILS-PB [r_s (38) = .43, p < .01], with the latter correlation indicating a likely significant contribution to global outcome in schizophrenia. In addition, both VOICE-DISCRIM [r_s (39) = .53, p< .01] and FACE-DISCRIM [r_s (28) = .46, p < .01] correlated significantly with BPRS conceptual disorganization scores, reflecting a significant relationship to overall cognitive symptom levels. Finally, only FACE-DISCRIM [r_s (28) = .37, p < .05] correlated significantly with BPRS positive symptom scores. No other significant correlations were found between TMT or DTT and ratings of clinical symptoms.

2.4.2.2 Principal Components Analysis (PCA)

To evaluate interrelationships among prosody measures, a PCA was performed (Table 4). The PCA yielded only two components with eigenvalues ≥ |1|. These two components accounted for 66% of the variance in the data and seemed to adequately describe the data, based on screen plots and the rotation sums of squared loading. Selection of these two factors for rotation revealed the following pattern. Auditory sensory processing (TMT and DTT) and auditory affective measures (VOICE-DISCRIM and VOICE-ID) all loaded onto factor 1, which had an eigenvalue of 2.64 and explained 44% of the variance. Visual affective measures (FACE-ID and FACE-DISCRIM) loaded exclusively on factor 2, which had an eigenvalue of 1.30 and explained an additional 21.6% of the variance. There was no significant loading of auditory measures onto factor 2 of the PCA

model, or reciprocally, of visual measures onto factor 1. No other single factor explained more than 15% of the variance.

Correlational analyses were performed to further probe these relationships (**Figure 1**). A hierarchical pattern of results was observed, in which performance on the most basic measures of auditory discrimination (TMT) correlated significantly with performance on the DTT and VOICE-DISCRIM, but not with VOICE-ID. In contrast, DTT performance correlated significantly with both VOICE-DISCRIM and VOICE-ID performance, both of which also correlated significantly with each other. Neither the TMT nor the DTT correlated significantly with either of the visual affective measures. Scores on the two visual affective measures, FACE-ID and FACE-DISCRIM, correlated with each other but not with corresponding auditory affective measures.

2.5 DISCUSSION

The ability to decode other people's emotional states by analyzing either their vocal intonations or facial expression is an integral part of human existence, leading to a significant recent interest in this process in schizophrenia (e.g., (Edwards *et al.*, 2002; Gur *et al*., 2002; Suslow *et al.*, 2003). The present study demonstrates that patients with schizophrenia show significant impairments in the ability to decode affect based on either auditory vocal or visual facial cues, thus replicating previous work in this population (e.g., (M. Alpert *et al*., 2000; Edwards *et al*., 2001; Haskins *et al*., 1995; Kerr & Neale, 1993; E. D. Ross *et al*., 2001a)). The main objective of the current study, however, was to assess whether affective (emotional) prosodic dysfunction in patients was related to more fundamental deficits in early auditory sensory processing. This hypothesis is based on

recent work showing deficits in simple auditory processing in schizophrenia (Holcomb *et al.*, 1995; Javitt *et al*., 1997; E. F. Rabinowicz *et al*., 2000; R. D. Strous *et al.*, 1995b; Wexler *et al.*, 1998), given that auditory affective recognition depends heavily on the ability to decode changes in pitch and intonation. However, to our knowledge, no studies have previously examined affective recognition ability, relative to more basic components of sensory processing.

To assess potential relationships between sensory competence and affect discrimination, patients were evaluated on both their ability to perceive pitch changes of pure tones (TMT), as well as their ability to recognize complex pitch abnormalities within short musical sequences (DTT). Two corresponding tasks were used for characterization of auditory and visual affective processing. In the discrimination tasks, subjects had to indicate whether two stimuli (sentences or faces), expressed the same or different emotions. In the corresponding identification tasks, subjects were required to name the emotion. A limitation of the present task is that TMT measures were obtained for patients only, but deficits in TMT performance have been repeatedly documented in schizophrenia (Holcomb *et al*., 1995; Javitt *et al*., 1997; E. F. Rabinowicz *et al*., 2000; R. D. Strous *et al*., 1995b; Wexler *et al*., 1998). Control subjects, in general, show a threshold of 3% Δf for >90% correct performance versus 20% to 50% Δf observed in this and prior studies of schizophrenia (Holcomb *et al*., 1995; Javitt *et al*., 1997; E. F. Rabinowicz *et al*., 2000; R. D. Strous *et al*., 1995b; Wexler *et al*., 1998).

Results of the experiment confirmed our *a priori* hypothesis. First, patients were equally impaired in auditory and visual affective judgments. We conducted a principal components analysis, which showed that auditory affective measures and auditory sensory measures loaded exclusively onto one component. In contrast, all the visual tasks loaded onto a second, independent component. Critically, and in line with our main hypothesis, we found robust correlations between tone-matching ability and affective prosodic discrimination. Further, in these correlation analyses, a hierarchical pattern was observed, with deficits in basic measures predicting deficits in affective processing measures. As such, while showing only correlational relationships, our data suggest that prosodic processing deficits, ubiquitously seen in this population, may well be a consequence of more basic early sensory dysfunction in the auditory system. Thus, although schizophrenia is definitively associated with reduced limbic activation during affective processing (Gur *et al*., 2002), our results suggest that the abnormal limbic activation may result, at least in part, from loss of normal "bottom-up" input.

The hypothesis that "bottom-up" deficits drive impaired limbic function in prosodic tasks is also consistent with studies that show that schizophrenic patients do not have universal difficulty understanding the concept of emotion, only in detecting or expressing affect. For example, while watching movies, schizophrenic patients self-report levels of happiness and sadness similar to those reported by normal volunteers (Kring *et al.*, 2003). Similarly, when asked to categorize the emotional valence of words, schizophrenia patients and control subjects showed similar word valence patterns (Kring *et al*., 2003). This disjunction between emotional self-experience and outward emotional perception and expression mirrors that of Bleuler's original 1911 conjecture (Bleuler, Ger-

man-1911: Trans-1950). He noted that, while patients do not amplify their emotional responses in the same manner that healthy individuals do, they nevertheless show an intact conception of emotion.

In addition to showing relationships between early sensory processing dysfunction and higher order deficits in schizophrenia, the present study has implications for both the functional anatomy of schizophrenia, and the nature of underlying emotional processing disturbances. On the level of functional anatomy, schizophrenia is frequently considered to be associated with predominant left hemispheric disturbance, particularly with regards to temporal lobe processing. An issue with this literature is that many tasks used in schizophrenia rely on phonetic analysis or other left-lateralized abilities, and do not critically assess right hemisphere processing. When tasks, such as emotional prosody tasks, are used that stress right hemispheric function, right-sided deficits are observed as well (Mitchell *et al*., 2004; E. D. Ross *et al*., 2001a). In the present study, the prosody task that was selected is known to depend heavily upon right hemisphere processing (Lakshminarayanan *et al.*, 2003; Pell, 1998; Zatorre *et al*., 2002). Deficits were observed, consistent with prior literature, which document right-sided, as well as left-sided temporal lobe dysfunction in schizophrenia.

On a functional level, when patients had to compare affect between stimuli but did not have to name the affect, patients showed equivalently severe deficits in affect discrimination as they did in affect identification. Thus, the present study argues against a specific deficit in the ability to name, rather than simply to recognize emotion. There were significant correlations between basic auditory processing measures and auditory affective processing measures, and there were not significant intercorrelations between

auditory and visual affective processing measures, which argues against a generalized deficit in emotion processing, and for separate sensory-driven deficits in the decoding of emotion within both the auditory and visual systems. In the auditory system, the deficits reflect in part, more basic limitations in decoding the pitch changes that encode affective prosody during normal conversation.

Although the present study concentrated primarily on the auditory system, deficits in basic sensory processing have also been demonstrated within the visual system as well (e.g. (P. D. Butler *et al*., 2001; Foxe *et al*., 2001). Future studies will, therefore, have to determine the degree to which early sensory processing deficits contribute to visual, as well as auditory, affect discrimination in schizophrenia.

In the present study, deficits in prosodic processing correlated significantly with severity of conceptual disorganization and negative symptoms, as well as with problem-solving ability, as reflected in the ILS-PB. The ability to decode other people's emotions based on their tone of voice is a critical component of human interaction. Impaired emotion recognition ability may be a significant mediating variable between both basic disturbances in sensory processing and poor global outcome. It has also been observed that basic visual processing deficits contribute to poor global outcome in schizophrenia (M.F. Green *et al.*, 2000). Deficits in visual affect recognition, such as those observed here, may also be a mediating variable between basic impairments in visual processing and poor global outcome in schizophrenia. Taken together, such results would fit nicely within frameworks such as Braff's (1993), in which information processing deficits cascade upward into neuropsychological deficit symptoms, trait-related factors, and clinical outcome.

Finally, although the present study evaluated only affective prosody in schizophrenia, a prediction of the present study is that patients should have deficits in decoding non-affective prosody as well. Whether patients have deficits in decoding non-affective prosody, as well as affective prosody, has only been studied to a limited degree. For example, Kerr and Neale (1993) found that individuals with schizophrenia differed from comparison subjects on emotion perception tasks only to the same extent as they did on control tasks of basic speech perception. In contrast, Murphy and Cutting (Murphy & Cutting, 1990) reported that patients' performance on stress prosody tests was not significantly different from that of comparison subjects, whereas performance on emotional prosody tasks was significantly worse. We have previously observed decreased ability of patients to discriminate ambiguous speech sounds, suggesting at least that phonemic processing may be impaired at the sensory level (Cienfuegos *et al.*, 1999). Further studies, however, are required to fully resolve this issue.

In summary, our data suggest that deficiencies in elementary pitch perception, a fundamental building block in both melodic and prosodic comprehension, can significantly affect auditory affect recognition in schizophrenia. These findings further suggest that patients' inability to correctly infer other people's emotions from speech may be related more to deficits within sensory modality than to deficits in comprehending the concept of emotion. Future studies are required to further delineate the basis of the relationship between early auditory processing deficits and deficits in emotion recognition in schizophrenia and to determine the degree to which similar relationships can be observed in other sensory domains.

2.6 REFERENCES

Alpert, M., & Anderson, L. T. (1977). Imagery mediation of vocal emphasis in flat affect. *Arch Gen Psychiatry, 34*, 208-212.

Alpert, M., Rosenberg, S. D., Pouget, E. R., & Shaw, R. J. (2000). Prosody and lexical accuracy in flat affect schizophrenia. *Psychiatry Res, 97*(2-3), 107-118.

Andreasen, N. C. (1982). Negative symptoms in schizophrenia. *Archives of General Psychiatry, 39*, 784-788.

Baltaxe, C. A., & Simmons, J. Q., III. (1995). Speech and language disorders in children and adolescents with schizophrenia. *Schizophr Bull, 21*(4), 677-692.

Bleuler, E. (German-1911: Trans-1950). *Dementia praecox or the group of schizophrenias* (H. Zinkin, Trans.). New York: International Universities press.

Braff, D. L. (1993). Information processing and attention dysfunctions in schizophrenia. *Schizophrenia Bulletin, 19*, 233-259.

Braus, D. F., Weber-Fahr, W., Tost, H., Ruf, M., & Henn, F. A. (2002). Sensory information processing in neuroleptic-naive first-episode schizophrenic patients: A functional magnetic resonance imaging study. *Arch Gen Psychiatry, 59*(8), 696-701.

Butler, P. D., Schechter, I., Zemon, V., Schwartz, S. G., Greenstein, V. C., Gordon, J., *et al*. (2001). Dysfunction of early-stage visual processing in schizophrenia. *Am J Psychiatry, 158*(7), 1126-1133.

Cienfuegos, A., March, L., Shelley, A. M., & Javitt, D. C. (1999). Impaired categorical perception of synthetic speech sounds in schizophrenia. *Biol Psychiatry, 45*(1), 82-88.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences, 2nd edition*. Hillsdale, NJ: Lawrence Erlbaum Assoc.

Drayna, D., Manichaikul, A., de Lange, M., Snieder, H., & Spector, T. (2001). Genetic correlates of musical pitch recognition in humans. *Science, 291*(5510), 1969-1972.

Edwards, J., Jackson, H. J., & Pattison, P. E. (2002). Emotion recognition via facial expression and affective prosody in schizophrenia: A methodological review. *Clin Psychol Rev, 22*(6), 789-832.

Edwards, J., & McGorry, P. D. (2002). *Implementing early intervention in psychosis: A guide to establishing early psychosis services.* London: Martin Dunitz.

Edwards, J., Pattison, P. E., Jackson, H. J., & Wales, R. J. (2001). Facial affect and affective prosody recognition in first-episode schizophrenia. *Schizophr.Res, 48*(2-3), 235-253.

Ekman, P. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Foxe, J. J., Doniger, G. M., & Javitt, D. C. (2001). Early visual processing deficits in schizophrenia: Impaired p1 generation revealed by high-density electrical mapping. *Neuroreport, 12*, 3815-3820.

Fricchione, G., & Howanitz, E. (1985). Aprosodia and alexithymia--a case report. *Psychother Psychosom, 43*(3), 156-160.

Green, M. F., Kern, R. S., Braff, D. L., & Mintz, J. (2000). Neurocognitive deficits and functional outcome in schizophrenia: Are we measuring the "right stuff"? *Schizophr.Bull., 26*(1), 119-136.

Gur, R. E., McGrath, C., Chan, R. M., Schroeder, L., Turner, T., Turetsky, B. I., *et al*. (2002). An fmri study of facial emotion processing in patients with schizophrenia. *Am J Psychiatry, 159*(12), 1992-1999.

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

Holcomb, H. H., Ritzl, E. K., Medoff, D. R., Nevitt, J., Gordon, B., & Tamminga, C. A. (1995). Tone discrimination performance in schizophrenic patients and normal volunteers: Impact of stimulus presentation levels and frequency differences. *Psychiatry Res., 57*(1), 75-82.

Hyman, S. E., & Arana, G. W. (1987). *Handbook of psychiatric drug therapy* (Vol. 7). New York: Little, Brown.

Izard, C. E. (1971). *The face of emotion*. New York: Appelton-Century-Crofts.

Javitt, D. C., Liederman, E., Cienfuegos, A., & Shelley, A. M. (1999). Panmodal processing imprecision as a basis for dysfunction of transient memory storage systems in schizophrenia. *Schiz. Bull., 25*, 763-775.

Javitt, D. C., Shelley, A., & Ritter, W. (2000). Associated deficits in mismatch negativity generation and tone matching in schizophrenia. *Clin.Neurophysiol., 111*(10), 1733-1737.

Javitt, D. C., Strous, R. D., Grochowski, S., Ritter, W., & Cowan, N. (1997). Impaired precision, but normal retention, of auditory sensory ("echoic") memory information in schizophrenia. *J Abnorm Psychol, 106*(2), 315-324.

Jonsson, C. O., & Sjostedt, A. (1973). Auditory perception in schizophrenia: A second study of the intonation test. *Acta Psychiatr.Scand., 49*(5), 588-600.

Kerr, S. L., & Neale, J. M. (1993). Emotion perception in schizophrenia: Specific deficit or further evidence of generalized poor performance? *J Abnorm.Psychol., 102*(2), 312-318.

Kring, A. M., Barrett, L. F., & Gard, D. E. (2003). On the broad applicability of the affective circumplex: Representations of affective knowledge among schizophrenia patients. *Psychol Sci, 14*(3), 207-214.

Lakshminarayanan, K., Ben Shalom, D., van Wassenhove, V., Orbelo, D., Houde, J., & Poeppel, D. (2003). The effect of spectral manipulations on the identification of affective and linguistic prosody. *Brain Lang, 84*(2), 250-263.

Liegeois-Chauvel, C., Giraud, K., Badier, J. M., Marquis, P., & Chauvel, P. (2001). Intracerebral evoked potentials in pitch perception reveal a functional asymmetry of the human auditory cortex. *Ann NY Acad Sci, 930*, 117-132.

Loeb, P. A. (1996). *Ils: Independent living scales manual*. San Antonio: The Psychological Corporation, Harcourt Race Jovanovich, Inc.

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

Murphy, D., & Cutting, J. (1990). Prosodic comprehension and expression in schizophrenia. *J Neurol Neurosurg.Psychiatry, 53*(9), 727-730.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Overall, J. E., & Gorham, D. E. (1961). The brief psychiatric rating scale. *Psychol Reports, 10*, 799-812.

Patel, A. D., Peretz, I., Tramo, M., & Labreque, R. (1998). Processing prosodic and musical patterns: A neuropsychological investigation. *Brain Lang, 61*(1), 123-144.

Pell, M. D. (1998). Recognition of prosody following unilateral brain lesion: Influence of functional and structural attributes of prosodic contours. *Neuropsychologia, 36*(8), 701-715.

Rabinowicz, E. F., Silipo, G., Goldman, R., & Javitt, D. C. (2000). Auditory sensory dysfunction in schizophrenia: Imprecision or distractibility? *Arch Gen Psychiatry, 57*(12), 1149-1155.

Revheim, N., & Medalia, A. (2004). Verbal memory, problem-solving skills and community status in schizophrenia. *Schizophr Res, 68*(2-3), 149-158.

Ross, E. D., Orbelo, D. M., Cartwright, J., Hansel, S., Burgard, M., Testa, J. A., *et al*. (2001). Affective-prosodic deficits in schizophrenia: Comparison to patients with brain damage and relation to schizophrenic symptoms [corrected]. *J Neurol Neurosurg Psychiatry, 70*(5), 597-604.

Strous, R. D., Cowan, N., Ritter, W., & Javitt, D. C. (1995a). Auditory sensory ("echoic") memory dysfunction in schizophrenia. *Am.J.Psychiatry, 152*(10), 1517-1519.

Strous, R. D., Grochowski, S., Cowan, N., & Javitt, D. C. (1995b). Dysfunctional encoding of auditory information in schizophrenia. *Schizophrenia Research, 15*, 135.

Suslow, T., Roestel, C., Droste, T., & Arolt, V. (2003). Automatic processing of verbal emotion stimuli in schizophrenia. *Psychiatry Res, 120*, 131-144.

Wexler, B. E., Stevens, A. A., Bowers, A. A., Sernyak, M. J., & Goldman-Rakic, P. S. (1998). Word and tone working memory deficits in schizophrenia. *Arch Gen Psychiatry, 55*(12), 1093-1096.

Zatorre, R. J. (1985). Discrimination and recognition of tonal melodies after unilateral cerebral excisions. *Neuropsychologia, 23*(1), 31-41.

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends Cogn Sci, 6*, 37-46.

Zatorre, R. J., & Samson, S. (1991). Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain, 114 (Pt 6)*, 2403-2417.

2.7 TABLES AND FIGURES

Table 1. Demographic and Clinical Characteristics of Healthy Control and Patient Populations

Demographic/Clinical Criteria	Control ($n = 34$)	Schizophrenia ($n = 43$)
Age	36 ± 9	39 ± 12
Gender (M/F)	14/20	33/10
Handedness R/L	30/4	40/3
Education	15.6 ± 2.3	10.6 ± 3.2
BPRS Total ($n = 39$)	N/A	34.3 ± 8.4
BPRS Conceptual Disorganization	N/A	2.1 ± 1.3
BPRS Anxiety/Depression	N/A	6.8 ± 3.1
BPRS Positive Symptoms	N/A	8.6 ± 4.4
SANS Total ($n = 39$)	N/A	29.2 ± 12.5
ILS-PB ($n = 38$)	N/A	32.8 ± 13

Values are mean ± SD.

M, male; F, female; R, right; L, left; BPRS, Brief Psychiatric Rating Scale; SANS, Schedule for Assessment of Negative Symptoms; ILS-PB, Independent Living Scales-Problem Solving Subscale.

Table 2. Tone Matching Test Performance in Patients with Schizophrenia

Level (% ΔF)	Performance (% Correct)	SD
50	92	10
20	85	14
10	76	18
5	71	20
2.5	61	15
Average Across All Levels	77	14

Table 3. Auditory and Visual Affective Processing

Measure	Group	*n*	Mean	SD	F^a	*df*	*p*	E.S. (d)
VOICE-DISCRIM	Schizophrenia	43	64.4	14.0	36.0	1,73	<.0001	1.6
	Comparison	34	83.3	7.2				
VOICE-ID	Schizophrenia	43	44.5	15.4	37.3	1,73	<.0001	1.7
	Comparison	34	66.5	10.0				
FACE-DISCRIM	Schizophrenia	31	78.5	11.1	15.9	1,41	<.001	1.6
	Comparison	14	92.0	4.0				
FACE-ID	Schizophrenia	30	59.6	17.0	4.7	1,40	<.04	.8
	Comparison	14	72.9	16.6				

E.S., Effect size in standard deviation (d) units; VOICE-DISCRIM, Voice emotion discrimination task; VOICE-ID, Voice emotion identification task; FACE-DISCRIM, Face emotion discrimination task; FACE-ID, Face emotion identification task.

[a]Analysis of variance was performed with factors of group and gender. Values represent the main efect of group. There were no significant gender or gender by group findings.

Table 4. Principal Components Analysis of Auditory and Visual Sensory and Affective Processing Measures

Rotated Factor Pattern	Factor 1 % of Total Variance	Factor 2 % of Total Variance
TMT	.82[a]	−.01
DTT	.83[a]	.12
VOICE-DISCRIM	.75[a]	.14
VOICE-ID	.60[a]	.31
FACE-DISCRIM	.09	.89[b]
FACE-ID	.15	.84[b]

TMT, Tone matching task; DTT, Distorted tunes task; VOICE-DISCRIM, Voice emotion discrimination task; VOICE-ID, Voice emotion identification task; FACE-DISCRIM, Face emotion discrimination task; FACE-ID, Face emotion identification task.

[a]Significant loading on Factor 1.

[b]Significant loading on Factor 2.

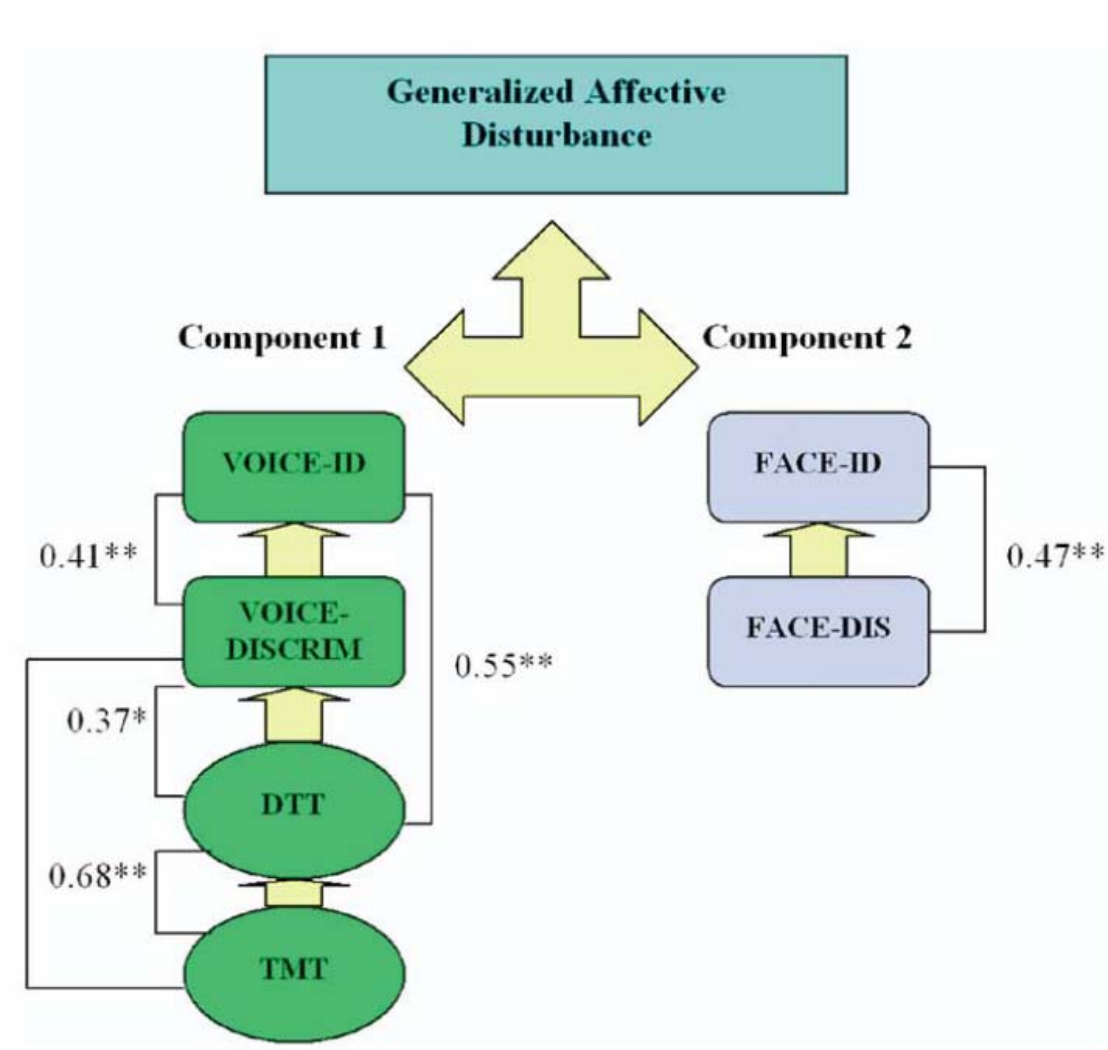


Figure 1. Schematic diagram of interrelationship between sensory and affective measures. Values shown represent Spearman correlation coefficients between indicated measures.*$p < .05$, **$p \leq .01$. TMT, Tone Matching Task; DTT, Distorted Tunes Task; VOICE-DISCRIM, Voice Emotion Discrimination Test; VOICE-ID, Voice Emotion Identification Test; FACE-DISCRIM, Face Emotion Discrimination Test; FACE-ID, Face Emotion Identification Test.

3.0 STUDY II

Neural Substrates of Impaired Prosodic Detection in Schizophrenia

3.1 ABSTRACT

Background: Individuals with schizophrenia show severe deficits in the ability to infer other people's emotions based upon vocal inflection (affective prosody), leading to reduced social competence. This study investigates neural substrates of impaired prosodic processing in schizophrenia using voxelwise and region of interest (ROI)-based diffusion tensor-magnetic resonance imaging (DTI) approaches.

Methods: 19 patients with schizophrenia and 19 healthy comparison subjects participated in both behavioral and DTI studies. Prosodic performance was evaluated using the Voice Emotion Identification (VOICE-ID) task. Basic perceptual functioning and executive processing were evaluated using the Distorted Tunes (DTT) and Wisconsin Card Sorting (WCST) tests, respectively. DTI-based fractional anisotropy (FA) scores were used for correlation analyses.

Results: As expected, patients showed significant and correlated deficits in both VOICE-ID (sd = 1.6) and DTT (sd = 1.2) performance. Impaired VOICE-ID performance correlated significantly with reduced FA values within primary and secondary auditory pathways, as well as in orbitofrontal cortex, corpus callosum, and amygdala bilaterally. Impaired DTT performance showed a similar pattern of correlation involving auditory and amygdalar pathways, but not prefrontal cortex. WCST performance in schizophrenia correlated primarily with prefrontal (cingulate fasciculus) FA. Voxelwise analyses were confirmed with subsequent ROI-based approaches.

Conclusions: Impaired prosodic processing in schizophrenia is associated with reduced white matter integrity within sensory, as well as amygdalar and prefrontal brain regions. Substrates of impaired emotion detection overlapped extensively with substrates of impaired basic pitch processing, but not of executive processing. These findings highlight the importance of impaired sensory processing to the pathophysiology of perceptual deficits in schizophrenia.

3.2 INTRODUCTION

Schizophrenia is associated with deficits in the ability to decode emotion based upon modulation of intonation (affective prosody). Such deficits, which serve as a proxy for social cognition, contribute substantially to impaired global outcome in schizophrenia (J. Brekke *et al*., 2005a). Traditionally, such deficits have been attributed to generalized neurocognitive dysfunction, particularly involving processes such as executive function and working memory (MF Green, 1996) that are based in dorsolateral prefrontal cortex (DLPFC), as well as limbic dysfunction (Phillips, 2003). More recently however, specific contributions of auditory processing deficits have been noted as well, with deficits in simple tone matching predicting deficits in prosodic identification (Leitman *et al.*, 2005). These findings predict both sensory-level and cognitive-level contributions to impaired prosodic processing. The present study investigates neural substrates of impaired auditory emotion detection using a combined behavioral and structural imaging approach.

Prosodic processing is a crucial but understudied area in schizophrenia. As early as the beginning of the last century, it was noted by Bleuler, "the intonation in patient's speech is often peculiar. In particular, there is often an absence, exaggeration or misplacement of modulation (Bleuler, German-1911: Trans-1950). Subsequent studies demonstrated, that patients have reduced ability to infer other's emotions based upon vocal modulation, reflecting impaired receptive, as well as productive, prosody (Kerr & Neale, 1993). Nevertheless, in schizophrenia, the internal representation of emotion appears unaffected in that patients express subjective happiness when discussing happy events and sadness when discussing sad events (Kring & Neale, 1996). Thus, the apparent inability of patients to process emotion does not seem to reflect inability to understand the concept

of emotion, and may instead reflect primary difficulty in either expressing or decoding the modulations needed to convey prosodic intent. Over recent years, sensory-level disturbances have been extensively documented with both the auditory (E. F. Rabinowicz *et al.*, 2000) and visual (P. D. Butler *et al.*, 2001) systems. Neural correlates of sensory contributions to impaired prosodic processing have not, however, been investigated.

In order to evaluate neural correlates of impaired auditory emotion detection in schizophrenia, correlative analyses were performed between a widely used behavioral measure of auditory emotion recognition ability, and structural brain organization in schizophrenia determined using magnetic resonance-diffusion tensor imaging (DTI). Neuroanatomical abnormalities in schizophrenia have been extensively documented and shown to involve both white and gray matter structures (K. O. Lim *et al.*, 1999; Shenton *et al.*, 2001).

DTI is analyzed using the parameter of fractional anisotropy (FA), which reflects the relative diffusion of water parallel to the long axis of a white matter tract relative to diffusion across the axonal membrane (Lim, 1998; Kubicki, 2005). Reduced FA in schizophrenia is thought to reflect either axonal or myelin-related pathology (Kubicki, 2005), both of which have been documented in schizophrenia (K. L. Davis *et al.*, 2003). Reduction in FA may also reflect disorganization of fiber bundles ("disconnectivity"), although in areas of crossing fibers, such "disonnectivity" may lead to paradoxical increases in FA (Pomara *et al.*, 2001).

Regardless of underlying etiology, FA has proved effective for evaluating structure/function relationships. For example, increased impulsivity has been found to correlate selectively with reduced FA inferior frontal white matter (Hoptman *et al.*, 2004;

Hoptman *et al.*, 2002), while impairments in executive processing have been found to correlate selectively with reduced FA in anterior prefrontal regions (Kubicki *et al.*, 2003; K. Lim *et al.*, 2005) Impairments in visual processing have been found to correlate selectively with reduced FA in optic radiations (P. Butler *et al.*, 2005).

For this study, we analyzed relationships between regional FA in schizophrenia patients and healthy comparison subjects and performance on two separate tasks: the Distorted Tune Task (DTT) (Drayna *et al.*, 2001), which measures the ability to detect incorrect notes within common melodies, and the Voice Emotion Identification Task (VOICE-ID) (Kerr & Neale, 1993), which measures the ability to decode emotions based upon tone of voice. The DTT was originally developed to assess genetic contributions to musical pitch perception abilities, and shows high heritability within families (Drayna *et al*., 2001). The VOICE-ID task has been used by ourselves (Leitman *et al*., 2005) and others (J. Brekke *et al*., 2005a; Kerr & Neale, 1993), and is highly sensitive to affective prosodic deficits in schizophrenia.

In the brain, auditory projection paths begin at the level of medial geniculate nucleus (MGN) and project to primary auditory cortex (Heschl's gyrus, A1) via superolaterally projecting thalamocortical (acoustic) radiations (**Figure 1A**). From auditory cortex, fibers project to higher brain regions along both ventral and dorsal divisions of the arcuate fasciculus (Parker *et al.*, 2005). The ventral stream is primarily involved in acoustic feature analysis (Arnott *et al.*, 2004), while the dorsal stream is thought to processes spatial and spectral motion (Belin & Zatorre, 2000), including speech (Arnott *et al*., 2004).

Affective prosodic comprehension can be conceptualized as involving a "three stage processing chain" that begins with sensation (stage 1) in primary auditory cortex and continues with integration (stage 2) within ventral aspects of temporal cortex and superior temporal sulcus (STS). There aspects of the acoustical information are tagged as affective. Processing proceeds finally to the cognitive stage (stage 3) in inferior frontal regions, where this information is evaluated both semantically and contextually (A Schirmer & Kotz, in press).

In the present study, two specific tasks, DTT and VOICE-ID, were used to evaluate functioning of sensation and integration phases of affective prosodic comprehension. Patients with schizophrenia have been found previously to show deficits in sensory-level performance, as reflected in impaired tone matching ability (E. F. Rabinowicz *et al*., 2000) and reduced auditory event-related potential (ERP) generation (Javitt *et al.*, 1995a). For the present study, we hypothesized that joint impairments in DTT and VOICE-ID performance in patients would correlate significantly with reduced FA primarily in basic auditory brain regions (i.e., acoustic radiations) that subserve sensation, and that VOICE-ID impairments would show additional correlations in ventral and dorsal stream projection regions that subserve integration and cognitive evaluation.

As a control condition, we evaluated FA correlations with performance levels on the Wisconsin Card Sorting Test (WCST), a widely used, visually-based test of executive/prefrontal performance that would not be expected to show correlations with auditory sensory regions. In a prior study, increased perseverative error rate on the WCST was found to correlate with reduced FA in specific regions of cingulum (Kubicki *et al*.,

2003). In the present study, patterns of voxelwise correlations to the WCST were compared to pattern observed for DTT/VOICE-ID.

3.3 METHODS

3.3.1 Participants

Nineteen patients (1 female) meeting DSM-IV criteria for either schizophrenia (N=17) or schizoaffective disorder (N=2) participated in this study. Patients had a mean age of 35 ± 11 years and a mean education level of 11 ± 3 grades achieved. Diagnoses were based upon Structured Clinical Interview for DSM-IV (SCID), using all available clinical information. Twelve patients were receiving only second-generation antipsychotics (primarily risperidone or olanzapine), two patients were receiving clozapine, two patients were receiving only traditional antipsychotics (haloperidol), and three patients were receiving combination treatment. Mean chlorpromazine equivalency (CPZ) dose was 1298.3 ± 780.4 mg. Mean illness duration was 15.7 ± 8.7 years. Clinical ratings of patients followed methods described previously (Leitman *et al*., 2005). Patients had a mean rating of 35.5 ± 7.1 on the Brief Psychiatric Rating Scale, and a mean rating of 32.5 ± 12.8 on the Schedule for Assessment of Negative Symptoms.

The healthy comparison group of nineteen subjects (6 female) had a mean age of 36 ± 9 years and a mean education level of 16 ± 2 grades achieved. This group consisted of staff volunteers as well as individuals who responded to local advertisements.

The local institutional review boards approved all experimental procedures and all subjects provided written informed consent after study procedures were fully explained. Participants received $10/hour for participation.

All subjects save one patient were right-handed as assessed using methods described previously (Leitman *et al*., 2005). Within this sample, twelve of nineteen patients and seven of nineteen healthy comparison subjects had been in a prior study (Leitman *et al*., 2005). WCST was conducted solely on patients and two subjects' data were not included due to methodological considerations.

3.3.2 Behavioral Measures

The following behavioral measures were used:

Voice Emotion Identification Test (VOICE-ID) (Kerr & Neale, 1993): This test consists of twenty one spoken sentences conveying one of six different emotions (happiness, anger, fear, sadness, surprise, or shame). Participants choose one of the six emotions for identification. Performance is calculated based upon the percentage of correctly identified sentences.

Distorted Tunes Task (DTT): The Distorted Tunes Task (Drayna *et al*., 2001) consists of twenty six popular tunes ranging in length from twelve to twenty six notes. Seventeen of the tunes are rendered melodically incorrect by changing the pitch of two to nine notes within the tune. Subjects respond "yes" or "no" as to whether the melody is correct and are asked to report whether the melody was familiar or not.

The Wisconsin Card Sorting Task (WSCT) (Heaton, 1993): (Patients only) This task has been well described elsewhere (Heaton, 1993). For this study, we used perseverative error rate (PSV) as our primary dependent measure.

3.3.3 Magnetic Resonance Imaging (MRI)

Scanning was performed on a 1.5T Siemens Vision system (Erlangen, Germany) at the NKI Center for Advanced Brain Imaging. Three main sequences were acquired: a magnetization prepared rapidly acquired gradient echo scan (MPRAGE; TR/TE = 11.4/4.9ms, matrix = 256x256, FOV = 300mm, NEX = 1, 1.17mm slice thickness, 172 slices, no gap), a turbo spin echo scan (TSE; TR/TE = 5000/22,90ms, matrix = 256x256, FOV = 224mm, NEX = 1, 5mm slice thickness, 26 slices, no gap), and a DTI sequence. The DTI sequence has been described elsewhere (K. O. Lim & Helpern, 2002) (TR/TE = 6000/100ms, matrix = 128x128 (interpolated to 256x256), FOV = 240mm, 5mm slice thickness, 20 slices, no gap) and employs a double echo pulse to minimize eddy current effects (Reese TG, 2003). The sequence entailed four acquisitions of six diffusion-weighted images ($\underline{b}$ = 1000 s/mm^2) for 20 slices. In addition, two acquisitions without diffusion weighting ($\underline{b}$ = 0 s/mm^2) were acquired

3.3.4 Neuroimaging Data

FA was calculated using custom software. The $\underline{b}$ = 0 images were corrected for susceptibility induced distortion and were transformed into MNI space using methods described elsewhere (Ardekani *et al.*, 2003; Hoptman *et al.*, 2004). Images were matched to a template in MNI space, and the final voxel size was 2x2x2mm^3. A white matter mask was computed from the mean normalized patient FA image using a nonparametric image segmentation algorithm (Otsu, 1979) and was applied to all of the standardized images.

This approach limited the voxels to white matter and resulted in fewer statistical comparisons, thereby lowering the probability of false positive tests.

Following transformation into Talairach space, images were masked such that only voxels with data present for all subjects were included in the analyses. This ensured that missing data, which would have zero values, would not drive correlations.

3.3.5 Statistical Analysis

Between-group comparisons of prosodic (VOICE-ID) and pitch (DTT) performance were performed using repeated-measures analysis of variance (rmANOVA). Spearman correlation coefficients were used to measure the relationship between task performances within groups.

For neuroimaging data, a voxelwise correlation approach was used similar to that of Baudewig *et al.* (2003), with thresholds as described previously (Hoptman *et al*., 2004; K. Lim *et al*., 2005). The approach is protected against false positive correlations using voxels significant $p \leq .05$ that are grown from a seed voxel with a significance value of $p < .005$. To supplement this criterion, we only considered clusters with more than 11 contiguous voxels. To assess areas of shared correlation across tasks, maps of each Task-FA correlation clusters were overlaid, and a new map representing overlap regions (identically thresholded for each task) was generated.

Our voxelwise correlation analysis was two-tailed; nevertheless we focused *a priori* only on correlations in which worse performance predicted FA reductions. This analysis was performed on patient data, and FA levels in regions in which there were sig-

nificant Task x FA correlation clusters were used to create regions of interest (ROI) for between-group comparisons.

Between-group comparisons of the effect of FA on performance was conducted by means of separate repeated measure ANCOVAs, which assessed between-group performance on DTT and VOICE-ID covarying for FA levels sampled within the seed voxel for each ROI. Strength of correlation with covariate was used to assess contribution of FA to performance across groups.

Additionally, potential hemispheric differences were analyzed using laterality indices, which were computed following a method adopted previously (Javitt *et al*., 1995a). Separate laterality analyses were performed for dorsal and ventral stream and analyzed across groups using rmANOVA.

3.4 RESULTS

3.4.1 Neuropsychological Results

Prosodic (VOICE-ID) and pitch (DTT) processing measures were obtained for all subjects. Patients showed significantly impaired performance across both tests ($F_{1,36}$ = 56.6, $p < 0.001$) (**Figure 2**). The group X task interaction was non-significant ($F_{1,36} = 2.5$, $p < 0.13$). DTT and VOICE-ID scores were significantly correlated both across groups (r_s = 0.55, N = 38, $p < .001$) and within the patient group alone (r_s = 0.54, N = 19, $p < .02$), but were not significantly correlated for comparison subjects ($r_s = 0.25$, N = 19, $p > .3$)**.** For patients, worse WCST performance (increased perseverative errors) predicted worse VOICE-ID ($r_s = -0.55$, N=17, $p < 0.02$), but not DTT ($r_s = -0.23$, N = 17, $p < 0.34$) per-

formance. Additionally, within patients, medication dosage did not correlate with neuropsychological measures (all p's > 0.22).

3.4.2 Structural correlations

VOICE-ID: As predicted, impaired VOICE-ID performance was significantly correlated with FA in regions lying between auditory thalamus (MGN) to primary auditory cortex (acoustic radiations, **Figure 1A**). These regions are known to contain auditory radiations from MGN to A1 (**Figure 1B**). In addition to these regions, significant correlation clusters were observed bilaterally along the ventral and dorsal auditory pathway in temporal and frontal cortex (**Figure 3A**) Other areas in which correlation clusters were observed include the corpus callosum (CC) splenium and body, as well as the posterior commissure (PC) and right cingulum (**Figure 3D**). Clusters were also observed adjacent to both left and right amygdala medial-laterally (**Figure 4A**). Areas of significant correlation also included white matter in Brodmann's regions 44, 45, 46 and orbitofrontal cortex (**Figures 3A & 3D**).

DTT: Also as predicted, the pattern of correlations with DTT closely resembled the pattern of correlation with VOICE-ID (**Figure 3B**). Regions of overlap included primary auditory radiations, dorsal and ventral stream auditory projections (**Figure 3A, 3B**), and amygdala (**Figure 4A**). No correlations, however, were observed in regions 44, 45 or 46, or in orbitofrontal cortex.

WCST: As opposed to DTT and VOICE-ID, no significant correlations of WCST were observed in regions of the acoustic radiations or along either dorsal or ventral auditory radiations (**Figure 3C**). Further, there were no significant areas of overlap between

WCST and VOICE-ID (**Figure 3A, 3C**). Significant correlation clusters were observed between WCST perseverative error scores and FA levels in white matter in the regions of right anterior cingulate gyrus (ACG) (**Figure 3F**). Even in frontal regions, however, little overlap was observed between correlation clusters for VOICE-ID and WCST. Finally, in contrast to VOICE-ID and DTT, no correlations in the vicinity of the amygdala were observed (**Figure 4B**).

3.4.3 ROI-based analyses

In order to further confirm correlations between impaired DTT and VOICE-ID performance in patients, mean FA values were extracted for ROIs located within dorsal and ventral stream auditory regions. In order to assess the relative contributions of group membership (comparison subjects vs. patients) and FA levels in the dorsal and ventral streams a multivariate ANCOVA was conducted in which group membership and FA were entered as covariates, with task performance as the dependent measure (**Table 1**). This approach revealed significant effects for both group and FA on DTT performance (all p's < 0.05) and significant effects for both FA and group on VOICE-ID performance across all ROIs (all p's < 0.05) except for left ventral stream where there was a group effect, but no significant FA effect (**Table 1**). There was a tendency toward relatively greater left sided FA in comparison subjects than in patients ($F_{1,36} = 3.2$, $p = .08$), although neither group showed significant laterality (all p's > .15).

3.5 DISCUSSION

Deficits in auditory emotion recognition are among the strongest predictors of poor social outcome in schizophrenia (J. Brekke *et al*., 2005a), yet neural correlates have been investigated to only a limited degree. The present study utilizes a combined behavioral, voxelwise, and ROI-based investigation in order to localize areas of potential relevance to impaired prosodic processing. Significant correlations were observed between prosodic processing deficits and regions (e.g., prefrontal, periamygdalar) that are classically associated with neurocognitive dysfunction (Phillips, 2003; Phillips *et al.*, 2003). However, prominent deficits were observed with reduced FA in regions such as primary auditory radiations and dorsal and ventral auditory streams, suggesting that deficits in voice emotion recognition arise from sensory-level disturbance in schizophrenia as well. These findings thus support our prior observations of processing deficits within, rather than across, sensory modalities (Leitman *et al*., 2005) and suggest that functional and structural deficits within early sensory regions contribute to the overall pattern of cognitive dysfunction in schizophrenia.

For the present study, structure/function relationships were assessed using voxelwise DTI analysis. As opposed to volumetric approaches targeting gray matter regions, DTI provides a measure of integrity of white matter tracts in brain, which in turn may serve as a measure of disconnectivity or dismyelination within specific brain pathways (K. O. Lim *et al*., 1999; Shenton *et al*., 2001). FA reductions in schizophrenia have been observed across brain regions, consistent with underlying reductions in oligodendrocytic markers (K. L. Davis *et al*., 2003). Further, regionally specific correlations have already been observed for several well-validated tasks (Hoptman *et al.,* 2004; Hoptman *et al*.,

2002). Within this study, for example, reduced WCST performance in schizophrenia correlated with reduced FA within cingulate fasciculus, consistent with prior investigations in the field (Kubicki *et al.,* 2003) as well as functional brain imaging studies (Buchsbaum *et al.*, 2005). In contrast, no significant correlations were observed in auditory regions. Using the present approach, we (K. Lim *et al.,* 2005) have also previously demonstrated significant associations between verbal declarative memory, attention and FA in task-relevant regions attesting to the regional specificity of the current analysis approach.

The primary finding of this study was that reduced performance on the DTT and VOICE-ID tasks correlated independently with reduced FA in brain regions containing primary auditory radiations from MGN of thalamus to Heschl's gyrus (HG) and subsequent dorsal and ventral stream auditory projections. Additional areas of commonality included the genu and splenium of corpus callosum (CC) and middle cingulate gyrus (CG), consistent with lesion and neuroimaging studies implicating these regions in musical pitch and affective prosodic processing (Arnott *et al*., 2004; Blood & Zatorre, 2001; E. D. Ross *et al*., 1997). Correlation clusters adjacent laterally to the amygdala were also observed in both tasks. As such, these findings indicate significant contributions of low-level auditory processing deficits to higher order failures of neurocognition in schizophrenia.

In addition to areas of commonality, we also observed differences in correlation patterns between the pitch and prosodic tasks, particularly in frontal cortex. Here, prosodic correlations extended more anteriorly to Brodmann's areas 44, 45, and 46 – implicated particularly in speech perception (Arnott *et al*., 2004) (**Figure 2B**). Other areas involved in the affective evaluation of speech such as prefrontal and orbitofrontal cor-

tex (A Schirmer & Kotz, in press) also showed significant prosody-FA, but not pitch-FA, correlations (**Figures 2B, 3B**). The somewhat greater severity of prosodic vs. pitch identification deficits observed in our patient sample may reflect the greater extent of brain involvement engaged by the prosodic identification task.

The finding of sensory level correlations in patients with schizophrenia is consistent with well-replicated deficits in auditory processing that have been demonstrated using both electrophysiological (e.g., (Javitt *et al*., 1995b)) and behavioral (e.g. E. F. Rabinowicz *et al*., 2000) approaches. Structural imaging studies of auditory primary cortex conflict with some studies(Hirayasu *et al.*, 2000), but not all (Kulynych *et al.*, 1996), and show reduced volume of superior temporal auditory regions studies. However, postmortem changes analogous to those observed in prefrontal cortex have been demonstrated in auditory cortex as well (Sweet *et al.*, 2004), supporting its involvement in the pathophysiology of schizophrenia.

Our voxelwise analysis revealed unequivocal shared correlations between both pitch and prosodic performance and FA levels within auditory processing regions. However, in our study the ROI-based correlations between prosodic and pitch performance and FA levels that we observed in patients and across groups was not seen in comparison subjects alone. The reason for this difference is unclear, and may reflect near ceiling performance observed in some of the comparison subjects. Alternatively, however, it is possible that for comparison subjects, sensory aspects of prosodic detection may simply not be rate limiting. Decoding of emotion depends upon a synergistic processing chain that begins with sensation and ends at evaluation and semantic comprehension (A Schirmer & Kotz, in press). Physical features of speech, such as voice intensity or degree of pitch

modulation, are arbitrarily set by society and, it must be presumed, are not meant to be rate-limiting for the average listener. However, just as the absolute intensity of speech may become rate-limiting for an individual with an overall hearing impairment, modulation of speech may become rate-limiting for individuals with schizophrenia who show overall increased threshold for detecting such modulations (March *et al.*, 1999)

The use of a voxelwise FA approach in this study has both specific strengths and specific limitations. The strength of the approach is that all brain regions had the same *a priori* opportunity to correlate with prosodic performance. Thus, the fact that we observed a specific pattern of correlation involving sensory, frontal and limbic regions is not an artifact of our *a priori* hypotheses. Sensory contributions to cognitive dysfunction in schizophrenia may be missed in ROI-based studies because few investigators entertain *a priori* sensory-based hypotheses. The study was protected against type I error using well-established significance thresholds (Baudewig *et al*., 2003; Hoptman *et al*., 2004; K. Lim *et al*., 2005) Nevertheless, these findings should be confirmed in an independent cohort using both voxelwise and ROI-based analyses. Another limitation of this study is that all subjects were receiving antipsychotic medication at the time of testing. However, no correlation was observed between medication dose and either behavioral performance or FA levels. Further, the primary findings in this study were observed within, rather than across, patient groups, making a primary treatment effect unlikely.

In summary, patients with schizophrenia show severe deficits in the ability to process pitch and the ability to decode emotion based upon vocal intonation. The present findings suggest that both deficits may reflect failures in auditory processing starting at the most basic levels of auditory cortical processing.

3.5 REFERENCES

Ardekani, B. A., Nierenberg, J., Hoptman, M. J., Javitt, D. C., & Lim, K. O. (2003). Mri study of white matter diffusion anisotropy in schizophrenia. *Neuroreport, 14*(16), 2025-2029.

Arnott, S. R., Binns, M. A., Grady, C. L., & Alain, C. (2004). Assessing the auditory dual-pathway model in humans. *Neuroimage, 22*(1), 401-408.

Baudewig, J., Dechent, P., Merboldt, K. D., & Frahm, J. (2003). Thresholding in correlation analyses of magnetic resonance functional neuroimaging. *Magn Reson Imaging, 21*(10), 1121-1130.

Belin, P., & Zatorre, R. J. (2000). 'what', 'where' and 'how' in auditory cortex. *Nat Neurosci, 3*(10), 965-966.

Bleuler, E. (German-1911: Trans-1950). *Dementia praecox or the group of schizophrenias* (H. Zinkin, Trans.). New York: International Universities press.

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc Natl Acad Sci U S A, 98*(20), 11818-11823.

Brekke, J., Kay, D. D., Lee, K. S., & Green, M. F. (2005). Biosocial pathways to functional outcome in schizophrenia. *Schizophr Res, 80*(2-3), 213-225.

Buchsbaum, B. R., Greer, S., Chang, W. L., & Berman, K. F. (2005). Meta-analysis of neuroimaging studies of the wisconsin card-sorting task and component processes. *Hum Brain Mapp, 25*(1), 35-45.

Butler, P., Schechter, I., Saperstein, A., Revheim, N., Silipo, G., Zemon, V., *et al*. (2005). Deficits in contrast gain underly early visual processing dysfunction in schizophrenia. *Arch Gen Psychiatry, in press*.

Butler, P. D., Schechter, I., Zemon, V., Schwartz, S. G., Greenstein, V. C., Gordon, J., *et al*. (2001). Dysfunction of early-stage visual processing in schizophrenia. *Am J Psychiatry, 158*(7), 1126-1133.

Davis, K. L., Stewart, D. G., Friedman, J. I., Buchsbaum, M., Harvey, P. D., Hof, P. R., *et al*. (2003). White matter changes in schizophrenia: Evidence for myelin-related dysfunction. *Arch Gen Psychiatry, 60*(5), 443-456.

Drayna, D., Manichaikul, A., de Lange, M., Snieder, H., & Spector, T. (2001). Genetic correlates of musical pitch recognition in humans. *Science, 291*(5510), 1969-1972.

Green, M. (1996). What are the functional consequences of neurocognitive deficits in schizophrenia. *Am. J. Psychiatry, 153*, 321-330.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtis, G. (1993). The wisconsin card sorting task.

Hirayasu, Y., McCarley, R. W., Salisbury, D. F., Tanaka, S., Kwon, J. S., Frumin, M., *et al*. (2000). Planum temporale and heschl gyrus volume reduction in schizophrenia: A magnetic resonance imaging study of first-episode patients. *Arch Gen Psychiatry, 57*(7), 692-699.

Hoptman, M. J., Ardekani, B. A., Butler, P. D., Nierenberg, J., Javitt, D. C., & Lim, K. O. (2004). Dti and impulsivity in schizophrenia: A first voxelwise correlational analysis. *Neuroreport, 15*(16), 2467-2470.

Hoptman, M. J., Volavka, J., Johnson, G., Weiss, E., Bilder, R. M., & Lim, K. O. (2002). Frontal white matter microstructure, aggression, and impulsivity in men with schizophrenia: A preliminary study. *Biol Psychiatry, 52*(1), 9-14.

Javitt, D. C., Schroeder, C. E., Steinschneider, M., Arezzo, J. C., Ritter, W., & Vaughan, H. G., Jr. (1995a). Cognitive event-related potentials in human and non-human primates: Implications for the pcp/nmda model of schizophrenia. *Electroencephalogr.Clin.Neurophysiol.Suppl, 44*, 161-175.

Javitt, D. C., Shelley, A. M., Grochowski, S., & Ritter, W. (1995b). Mismatch negativity (mmn) as an index of impaired auditory sensory memory in schizophrenia. *Schizophrenia Research, 15*, 179.

Kerr, S. L., & Neale, J. M. (1993). Emotion perception in schizophrenia: Specific deficit or further evidence of generalized poor performance? *J Abnorm.Psychol., 102*(2), 312-318.

Kring, A. M., & Neale, J. M. (1996). Do schizophrenic patients show a disjunctive relationship among expressive, experiential, and psychophysiological components of emotion? *J Abnorm Psychol, 105*(2), 249-257.

Kubicki, M., Westin, C. F., Nestor, P. G., Wible, C. G., Frumin, M., Maier, S. E., *et al*. (2003). Cingulate fasciculus integrity disruption in schizophrenia: A magnetic resonance diffusion tensor imaging study. *Biol Psychiatry, 54*(11), 1171-1180.

Kulynych, J. J., Vladar, K., Jones, D. W., & Weinberger, D. R. (1996). Superior temporal gyrus volume in schizophrenia: A study using mri morphometry assisted by surface rendering. *American Journal of Psychiatry, 153*, 50-56.

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. *Biol Psychiatry, 58*(1), 56-61.

Lim, K., Ardekani, B., Nierenberg, J., Butler, P., Javitt, D., & Hoptman, M. (2005). Neurocognitive correlates of white matter integrity in schizophrenia. *American Journal of Psychiatry*, in press.

Lim, K. O., Hedehus, M., Moseley, M., de Crespigny, A., Sullivan, E. V., & Pfefferbaum, A. (1999). Compromised white matter tract integrity in schizophrenia inferred from diffusion tensor imaging. *Arch Gen Psychiatry, 56*(4), 367-374.

Lim, K. O., & Helpern, J. A. (2002). Neuropsychiatric applications of dti - a review. *NMR Biomed, 15*(7-8), 587-593.

March, L., Cienfuegos, A., Goldbloom, L., Ritter, W., Cowan, N., & Javitt, D. C. (1999). Normal time course of auditory recognition in schizophrenia, despite impaired precision of the auditory sensory ("echoic") memory code. *J. Abnorm. Psychol., 108*, 69-75.

Otsu, N. (1979). A threshold selection model from gray-level histograms. *IEEE Trans (SMC), 9*, 63-66.

Parker, G. J., Luzzi, S., Alexander, D. C., Wheeler-Kingshott, C. A., Ciccarelli, O., & Lambon Ralph, M. A. (2005). Lateralization of ventral and dorsal auditory-language pathways in the human brain. *Neuroimage, 24*(3), 656-666.

Phillips, M. L. (2003). Understanding the neurobiology of emotion perception: Implications for psychiatry. *Br J Psychiatry, 182*, 190-192.

Phillips, M. L., Drevets, W. C., Rauch, S. L., & Lane, R. (2003). Neurobiology of emotion perception ii: Implications for major psychiatric disorders. *Biol Psychiatry, 54*(5), 515-528.

Pomara, N., Crandall, D. T., Choi, S. J., Johnson, G., & Lim, K. O. (2001). White matter abnormalities in hiv-1 infection: A diffusion tensor imaging study. *Psychiatry Res, 106*(1), 15-24.

Rabinowicz, E. F., Silipo, G., Goldman, R., & Javitt, D. C. (2000). Auditory sensory dysfunction in schizophrenia: Imprecision or distractibility? *Arch Gen Psychiatry, 57*(12), 1149-1155.

Reese TG, H. O., Weisskopf RM and Wedeen VJ. (2003). Reduction of eddy-current-induced distortion in diffusion mri using a twice-refocused spin echo. *Magn Reson Med, 49*, 177-182.

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

Schirmer, A., & Kotz, S. (in press). Beyond the right hemisphere: Brain mechanism mediating emotional processing. *Trends in Cogntive Neuroscience*.

Shenton, M. E., Dickey, C. C., Frumin, M., & McCarley, R. W. (2001). A review of mri findings in schizophrenia. *Schizophr Res, 49*(1-2), 1-52.

Sweet, R. A., Bergen, S. E., Sun, Z., Sampson, A. R., Pierri, J. N., & Lewis, D. A. (2004). Pyramidal cell size reduction in schizophrenia: Evidence for involvement of auditory feedforward circuits. *Biol Psychiatry, 55*(12), 1128-1137.

Talairach, J. T., P. (1988). *Co-planar steriotaxic atlas of the human brain, 3 dimensional proportional system: An approuch to cerebral imaging*. New York: Thieme Medical Publishers.

3.6 TABLES AND FIGURES

Table 1: Multivariate Analysis of Musical Pitch (DTT) and Affective Prosody (VOICE-ID) Performance Co-Varying for Group and FA Levels Within Dorsal and Ventral Auditory Streams

ROI	FACTOR	MEASURE	*DF*	*F*	*p*
Right Ventral Stream	FA	DTT	1,37	11.9	0.001
(x=36, y=-21, z=-7[1])		VOICE-ID	1,37	14.6	0.001
	GROUP	DTT	1,37	11.9	0.02
		VOICE-ID	1,37	42.3	<0.0001
Right Dorsal Stream	FA	DTT	1,37	6.2	0.02
(x=44, y=-4, z=15[1])		VOICE-ID	1,37	4.9	0.03
	GROUP	DTT	1.37	7.7	0.01
		VOICE-ID	1,37	28.6	<0.001
Left Ventral Stream	FA	DTT	1,37	7.9	0.01
(x=36, y=-21, z=-7[1])		VOICE-ID	1,37	4.1	0.051
	GROUP	DTT	1.37	6.7	0.01
		VOICE-ID	1,37	26.4	<0.0001
Left Dorsal Stream	FA	DTT	1,37	7.4	0.01
(x=-47, y=-11, z=19[1])		VOICE-ID	1,37	1.2	0.28
	GROUP	DTT	1.37	4.0	0.053
		VOICE-ID	1,37	21.7	<0.0001

[1]Talairarch (Talairach, 1988) coordinates of ROI Locations for Between-Groups Analysis

Figure 1 (A-C):

(A) Fractional anisotropy (FA) map at the level of medial geniculate nucleus (MGN) and primary auditory cortex (Heschl's gyrus, HG). Medial-lateral fibers are shown in red, superior-inferior in blue and anterior posterior in green. *Inset*: Magnified view showing acoustic radiations (*), which project superolaterally from MGN to HG.

(B) View of fiber pathways at the level of ventral stream. *Inset*: Magnified view showing medial lateral (red*) radiations to Broca's area branching off of the superior arcuate fibers (green).

(C) Three-dimensional voxelwise correlation map for VOICE-ID performance. Inset, within-patient scatter plot and correlation values between FA levels from auditory radiations and VOICE-ID performance.

Figure 2: Behavioral task performance between groups: [1] Distorted Tunes Task. [2] Voice Emotion Identification. Effect sizes (d) for DTT and VOICE-ID were 1.2 and 1.6 sd units, respectively. Dashed lines through bars represent chance performance levels.

*$p < 0.01$, ** $p < 0.001$

Figure 3 (A-F): Voxelwise correlation maps for dorsal (blue) and ventral (green) auditory pathways for (A) Voice Emotion Identification (VOICE-ID), (B) Distorted Tunes Test (DTT) and (C) Wisconsin Card Sorting Test (WCST). (D-F) Voxelwise correlation maps for cingulate fasciculus (pink).

Figure 4 (A, B): Voxelwise correlations maps for (A) DTT/VOICE-ID and (B) WCST at the level of amygdala. Arrows indicate periamygdala correlations.

Figure 1 (A-C)

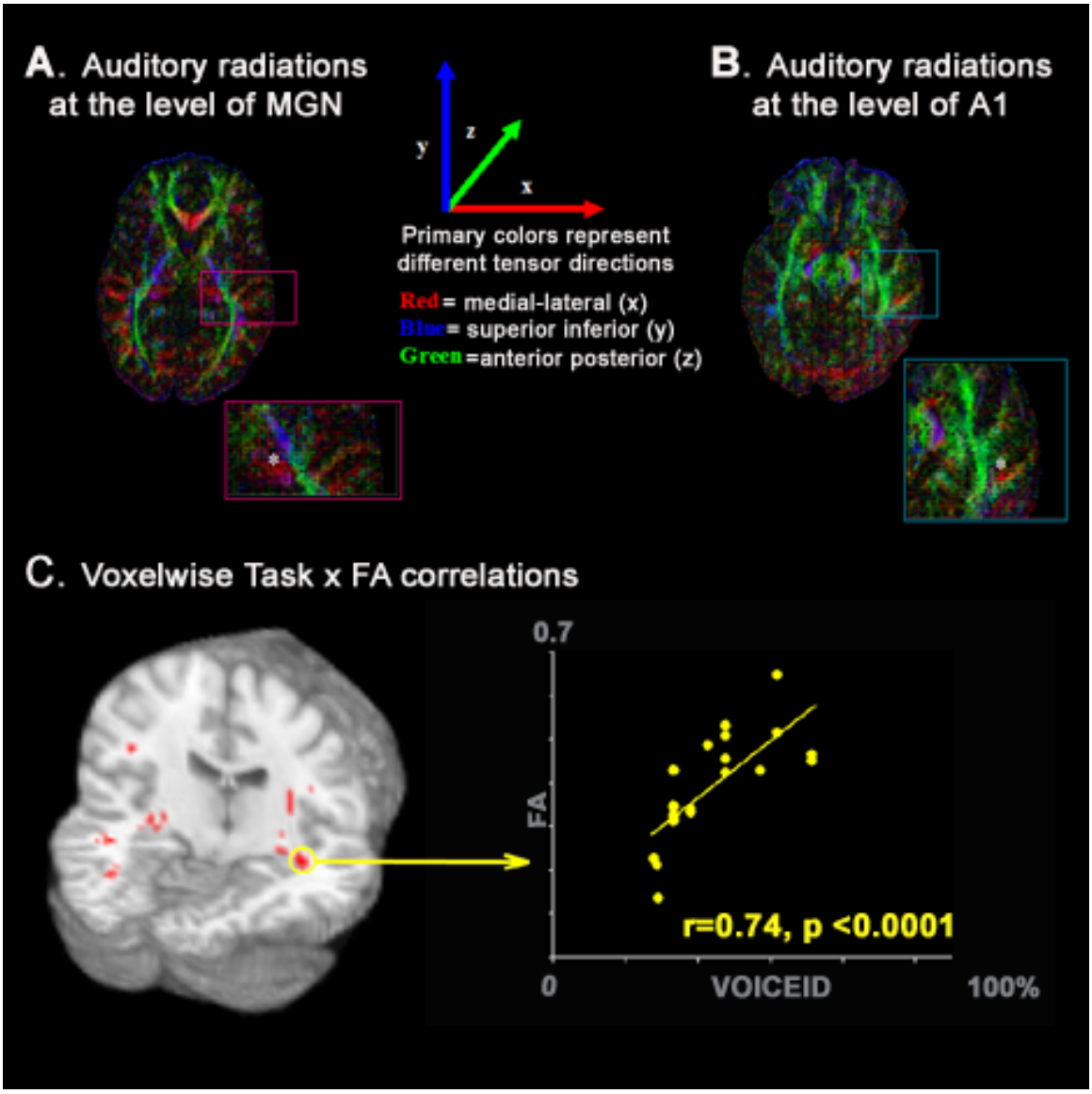

A. Auditory radiations at the level of MGN
y
z
x
Primary colors represent different tensor directions
Red = medial-lateral (x)
Blue = superior inferior (y)
Green =anterior posterior (z)
B. Auditory radiations at the level of A1
C. Voxelwise Task x FA correlations
0.7
FA
r=0.74, p <0.0001
0
VOICEID
100%

Figure 2

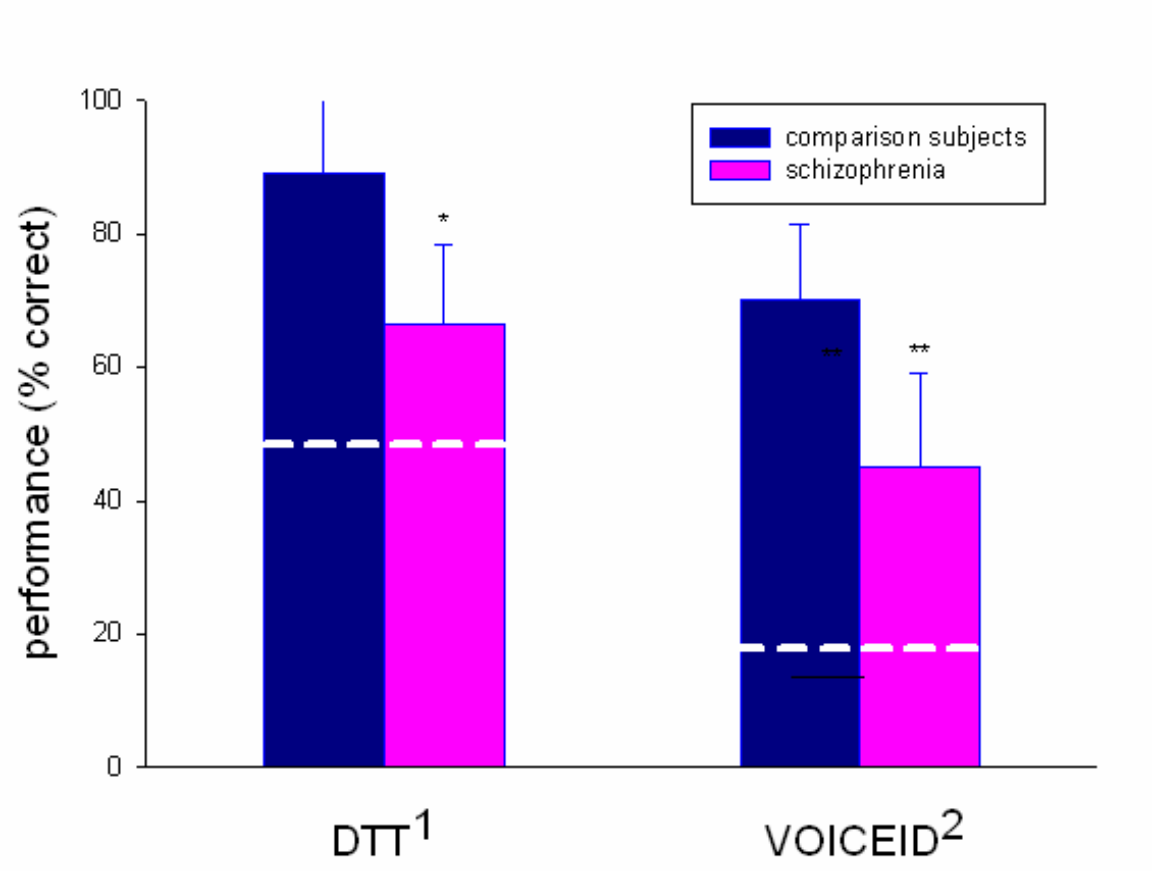

100
80
60
40
20
0
performance (% correct)
comparison subjects
schizophrenia
*
**
**
DTT1
VOICEID2

Figure 3 (A-F)

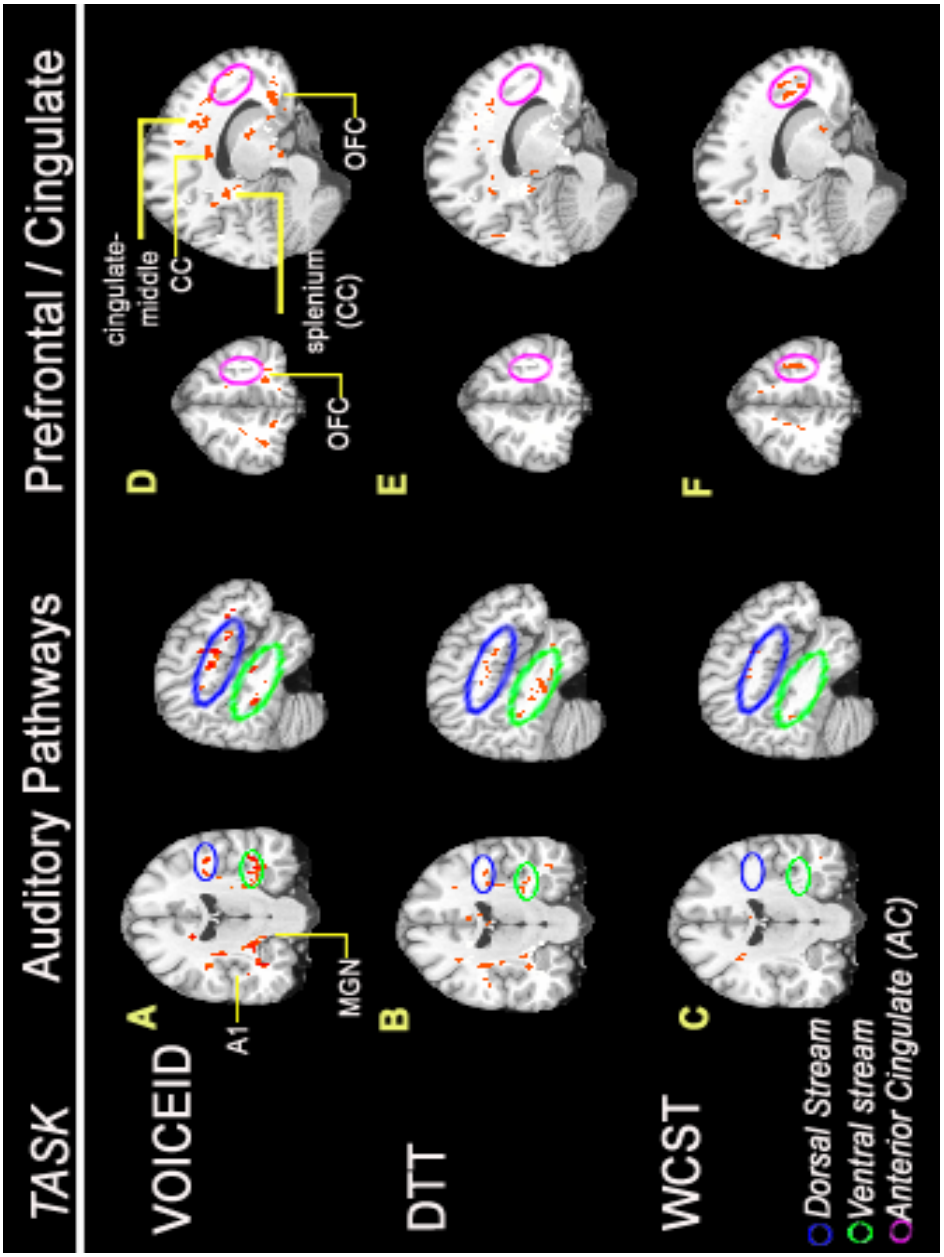
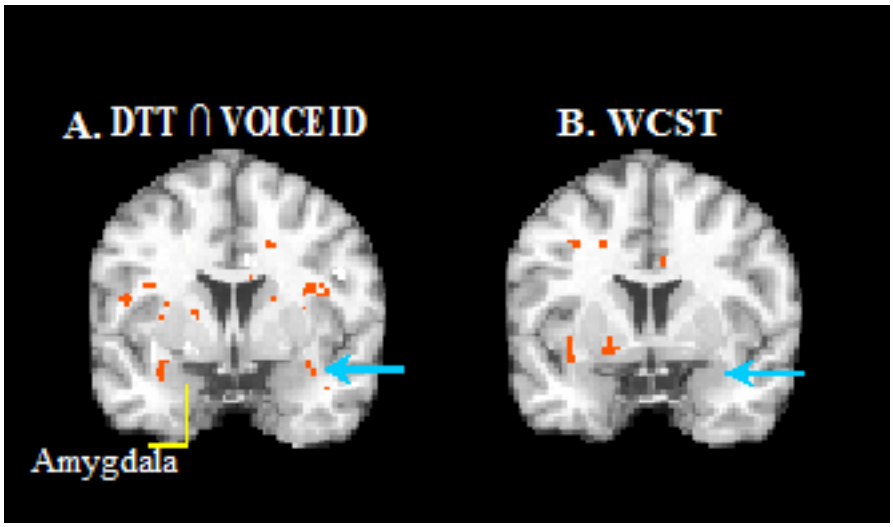

TASK
Auditory Pathways
Prefrontal / Cingulate
VOICEID
DTT
WCST
A
B
C
D
E
F
A1
MGN
cingulate-middle
CC
splenium (CC)
OFC
OFC
Dorsal Stream
Ventral stream
Anterior Cingulate (AC)

Figure 4



4.0 STUDY III

Non-Affective Prosodic Deficits in Schizophrenia and Their Relationship to Sensorial Disturbance

4.1 ABSTRACT

Background: Patients with schizophrenia show deficits in the ability to decode emotion based upon vocal intonation (affective prosody), leading to reduced social competence. Whether or not such deficits are limited to affective communication, however, has not been comprehensively explored.

Methods: Non-affective prosody was evaluated in 24 schizophrenia patients and 17 comparison subjects, relative to deficits in affective prosody and basic perceptual performance.

Results: Large effect size deficits (d > 1.0 sd's) were found across all prosodic measures. Within patients, poor performance on both affective and non-affective measures significantly correlated with pitch perception.

Conclusions: Schizophrenia patients show deficits in non-affective, as well as affective prosody, with both forms of prosodic deficit relating to underlying deficits in perceptual processing. The pattern of deficits suggests left-, as well as right-, hemispheric prosodic dysfunction.

4.2 INTRODUCTION

Poor social communication has long been considered one of the hallmarks of schizophrenia (Bleuler, German-1911: Trans-1950), yet its etiology has remained obscure. A key element in social communication is the ability to decode emotion based upon vocal intonation, a process termed affective prosody. Deficits in affective prosodic performance were first demonstrated in schizophrenia over nine decades ago. Since that time, such deficits have been shown to contribute directly to impaired social competence (J. Brekke *et al*., 2005a) and to stem, at least in part, from underlying impairments in basic perceptual processing (Leitman *et al*., 2005). In addition to conveying emotion, vocal intonation is also used to denote whether a sentence is a question or a statement (interrogative/declarative, and even which element of a sentence is of primary interest ("stress" prosody). Despite the long-standing interest in affective prosody in schizophrenia, the degree to which prosodic deficits extend even to non-affective aspects of communication (like declarative/interrogative intent) has not been previously evaluated.

Here our objectives were threefold: First, to assess whether schizophrenia patients have significant non-affective receptive prosodic deficits; second, to determine the degree to which such deficits correlate with impairments in pitch or executive processing; and third, to compare the magnitudes of deficits observed in non-affective prosody with those of affective prosody. To this end, we compared between-groups performance on non-affective prosodic tasks that have been used in lesion studies (Weintraub *et al.*, 1981) with affective prosody tasks used previously in schizophrenia (Kerr & Neale, 1993). Interrelationships between measures were assessed using correlational and principal component (PCA) analyses.

4.3 METHODS

4.3.1 Participants

Seventeen healthy volunteers (3 females; age = 32.5 ± 10.6; verbal IQ = 109.7 ± 10.4) who were staff or had responded to local advertisement and twenty-four chronically ill (illness duration = 18.8 ± 8.4 years) patients (3 females; age = 37.8 ± 10.2; verbal IQ = 94.1 ± 7.5) meeting DSM-IV criteria for either schizophrenia (N = 21) or schizoaffective disorder (N = 3) and receiving both typical and/or atypical antipsychotic medications (CPZ equivalent dose = 1373 ± 829 mg) took part in this study. Trained clinicians conducted screening, diagnoses, clinical assessment, and symptom ratings (BPRS total score = 137.9 ± 9.5), with diagnosis based on the Structured Clinical Interview (SCID) for the DSM-IV, and chart review.

The procedures conducted were under local Institutional Review Board supervision. All subjects had the procedure explained to them verbally before giving their written informed consent. All subjects spoke English as their first language and were right handed.

4.3.2 Materials

Non-affective prosody was assessed using Weintraub's Sentence Discrimination (SD) and Semantic Comprehension Tasks (SC) (Weintraub *et al*., 1981): Twenty-five pairs of semantically neutral sentences, such as "Jack climbed the mountain", were repeated after a brief delay. Seventeen of the pairs differed due to either stress (where stressed emphasis shifted between the subject and object of the sentence) or declara-

tive/interrogative differences. Eight pairs were identical. Subjects were asked whether the sentences were said in a same or different manner. Score reflected percent correct. Additionally, scores were broken down into percent correct for stress and declarative/interrogative distinctions.

The SC task consisted solely of sixteen utterances, expressing either declarative (eight utterances) or interrogative (eight utterances) intent. Subjects were asked whether the speaker posed a question or a statement. Score reflected percent correct.

Affective prosody was assessed using the Voice Emotion Identification (VOICE-ID) and Discrimination (VOICE-DIS) tasks (Kerr & Neale, 1993). The Distorted tunes task (DTT) total correct score (Drayna *et al*., 2001) and the Wisconsin Card Sort Task (WCST) perseverative error rate were used as comparison measures of sensory and executive contributions, respectively.

All auditory tasks required forced choice response and were presented on a CD player in a sound-attenuated room. Due to constraints on subject availability, all measures were not collected on all subjects.

4.3.3 Statistical Analyses

Between-groups effects across all auditory measures were assessed using MANOVA, with *post hoc* contrasts for specific measures (*t* tests). Non-parametric signal detection measures of sensitivity used A` after Snodgrass & Corwin (Snodgrass, 1988 #24), and bias B`` after Grier (Grier, 1971).

Correlations between non-affective and affective prosody, as well as pitch perception, were calculated within the patient group only, using Pearson's correlation coeffi-

cient (*r*) with Bonferroni corrections for multiple comparisons. In terms of PCA, factor selection and rotation were conducted on eigenvalues ≥ 1 (see (Leitman *et al*., 2005)).

Given that a prior study (Murphy & Cutting, 1990) has suggested that IQ may influence non-affective prosodic performance, *post hoc* assessments of the independence of prosodic deficits from premorbid IQ differences between groups were examined using a MANCOVA for all prosodic measures with group and IQ as fixed and covarying factors respectively. All statistical tests were two-tailed, with $\alpha \leq 0.05$, and computed in JMP software (SAS Institute Inc. Cary, NC, USA).

4.4 RESULTS

Patients performed significantly worse than controls across all prosodic measures (**Table 1**) with no significant group x task interaction ($p > .5$). On the SD task, patients showed significant decrements in performance on interrogative/declarative items, as well as stress items (all *p's* < .01).

Within non-affective prosody measures, patients were significantly less sensitive (A`) on both SD (0.86 ± 0.19 vs. 0.98 ± 0.32) and SC tasks (0.83 ± 0.17 vs. 0.98 ± 0.03) in detecting differing prosody or interrogative intent respectively ($p < .001$). However, there were no significant differences in terms of bias (B``) in the SD task (0.54 ± 0.14 vs. 0.68 ± 0.11) ($p > .5$) or in the SC task (0.37 ± 0.60 vs. 0.72 ± 0.67) ($p > .08$).

A correlation matrix between all neuropsychological measures revealed significant inter-measure correlations: Both non-affective prosody measures significantly correlated with their affective counterparts [SC x VOICE-ID ($r = 0.63$, N = 24, $p < .001$), SD x VOICE-DIS ($r = 0.61$, N = 21, $p < .003$)]. A similar pattern was seen in correlations with

pitch perception (SC x DTT: $r = 0.43$, N = 24, $p < .03$), although with the SD task, significance was only at trend (SD x DTT: $r = 0.39$, N = 24, $p < .058$). Further, analysis of the "stress" subset of the SD task found a significant correlation with VOICE-DIS ($r = 0.53$, N = 24, $p < .01$) but not DTT ($r = 0.18$, N = 24, $p > .4$). All correlations remained significant after Bonferroni correction except the SC x DTT comparison, which was at trend.

An examination of the interrelationship between prosody measures, pitch perception and executive processing using PCA (**Figure 1A**) yielded only two criteria-meeting components, which, when rotated, revealed that DTT and SD loaded exclusively onto the first component (0.77 and 0.82, respectively) and WCST onto the second (0.95). SC however, loaded significantly on both components (component 1 = 0.59, component 2 = 0.61). Correlations between VOICE-ID and SC were highly significant (**Figure 1B**). Finally, patient performance on non-affective prosody measures did not significantly correlate with illness duration or medication dosage (all p's > .2). Further, a *post hoc* MANCOVA for all prosody measures, covarying for both group membership and IQ, revealed a significant main effect (all p's < .01) for all measures with no significant effect for IQ (all p's > 2), but significant effects for group on SC and VOICE-ID but not SD performance (all p's < .04).

4.5 DISCUSSION

The primary finding of this study is that patients have significantly reduced sensitivity to non-affective, as well as affective, prosodic distinctions in speech. Specifically, patients have substantial detriments in semantic intent (interrogative/declarative), as well as stress pattern perception. These deficits are significantly interrelated with deficits in affective prosodic performance as well as pitch perception, but are distinct from the variance attributable to executive processing as reflected in WCST (**Figure 1**). The deficit, moreover, is independent of verbal IQ, and unrelated to either medication dosage or illness duration; supporting our *a priori* hypothesis that prosodic dysfunction in the illness has significant sensory, as well as cognitive antecedents. Finally, deficits in non-affective prosodic performance were of similar magnitude to deficits in prosodic domains (all d's > 1.0), although not as large as those reported in sarcasm perception in a similar population (Leitman *et al.*, 2006).

The present study does not include specific measures of real-world interaction in patients. However, it can be inferred that the inability to correctly interpret whether the subject or object of a sentence is the focus of stress, or whether an utterance is a statement or question, must interfere with everyday communication and contribute to poor social communication skills in schizophrenia. Additionally, while "stress" prosody has been evaluated previously with conflicting results (Murphy & Cutting, 1990; E.D. Ross *et al.*, 2001b), this study is the first to evaluate semantic prosody in schizophrenia. The present study suggests interrelated deficits in semantic and stress cues commonly used in everyday interaction.

The present findings are also significant on a neurological level, as schizophrenia aprosodia has generally been attributed to right hemisphere dysfunction (J. C. Borod *et al*., 1990; E.D. Ross *et al*., 2001b). As opposed to affective prosody, non-affective prosody is subserved primarily by the left hemisphere (E.D. Ross *et al*., 2001b), thus suggesting a generalized, bihemispheric pattern of prosodic disturbance in schizophrenia.

In conclusion, this study indicates that patients have substantially impaired non-affective, as well as affective prosody, suggesting that schizophrenia aprosodia, rather than being a proxy of impaired affect appraisal, reflects a more general social communicatory disturbance. This contention has both etiological, as well as real world clinical implications.

4.6 REFERENCES

Bleuler, E. (German-1911: Trans-1950). *Dementia praecox or the group of schizophrenias* (H. Zinkin, Trans.). New York: International Universities press.

Borod, J. C., Welkowitz, J., Alpert, M., Brozgold, A. Z., Martin, C., Peselow, E., *et al*. (1990). Parameters of emotional processing in neuropsychiatric disorders: Conceptual issues and a battery of tests. *J Commun Disord, 23*(4-5), 247-271.

Brekke, J., Kay, D. D., Lee, K. S., & Green, M. F. (2005). Biosocial pathways to functional outcome in schizophrenia. *Schizophr Res, 80*(2-3), 213-225.

Drayna, D., Manichaikul, A., de Lange, M., Snieder, H., & Spector, T. (2001). Genetic correlates of musical pitch recognition in humans. *Science, 291*(5510), 1969-1972.

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychol Bull, 75*(6), 424-429.

Kerr, S. L., & Neale, J. M. (1993). Emotion perception in schizophrenia: Specific deficit or further evidence of generalized poor performance? *J Abnorm.Psychol., 102*(2), 312-318.

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. *Biol Psychiatry, 58*(1), 56-61.

Leitman, D. I., Ziwich, R., Pasternak, R., & Javitt, D. C. (2006). Theory of mind (tom) and counterfactuality deficits in schizophrenia: Misperception or misinterpretation? *Psychol Med*, 1-9.

Murphy, D., & Cutting, J. (1990). Prosodic comprehension and expression in schizophrenia. *J Neurol Neurosurg.Psychiatry, 53*(9), 727-730.

Ross, E. D., Orbelo, D. M., Cartwright, J., Hansel, S., Burgard, M., Testa, J. A., *et al*. (2001). Affective-prosodic deficits in schizophrenia: Profiles of patients with brain damage and comparison with relation to schizophrenic symptoms. *J Neurol Neurosurg.Psychiatry, 70*(5), 597-604.

Weintraub, S., Mesulam, M. M., & Kramer, L. (1981). Disturbances in prosody. A right-hemisphere contribution to language. *Arch.Neurol, 38*(12), 742-744.

4.7 TABLES AND FIGURES

Table 1: Performance (% correct) on Indicated Tasks for Schizophrenia (N = 24) and Healthy Comparison Subjects (N = 17)

Measure[1]	*Schizophrenia (Mean/SD)*	*Comparison (Mean/SD)*	*t*	*Effect size (d)*
Non-affective prosody				
Sentence discrimination (SD)	79.8±20.6	98.1±2.5	-3.8***	1.4
SD – stress subfactor	79.1±26.5	99.5±2.0	-3.5***	1.1
Semantic comprehension (SC) (interrogative/declarative)	79.2±16.3	96.3±5.9	-3.9***	1.4
Affective prosody				
Voice emotion identification (VOICE-ID)	41.3±14.7	64.1±13.5	-5.0**	1.6
Voice emotion discrimination (VOICE-DIS)	70.0±13.8	84.9±5.3	-4.0**	1.4
Perceptual processing				
Distorted tunes task (DTT)	88.7±17.1	69.7±19.0	3.1*	1.1

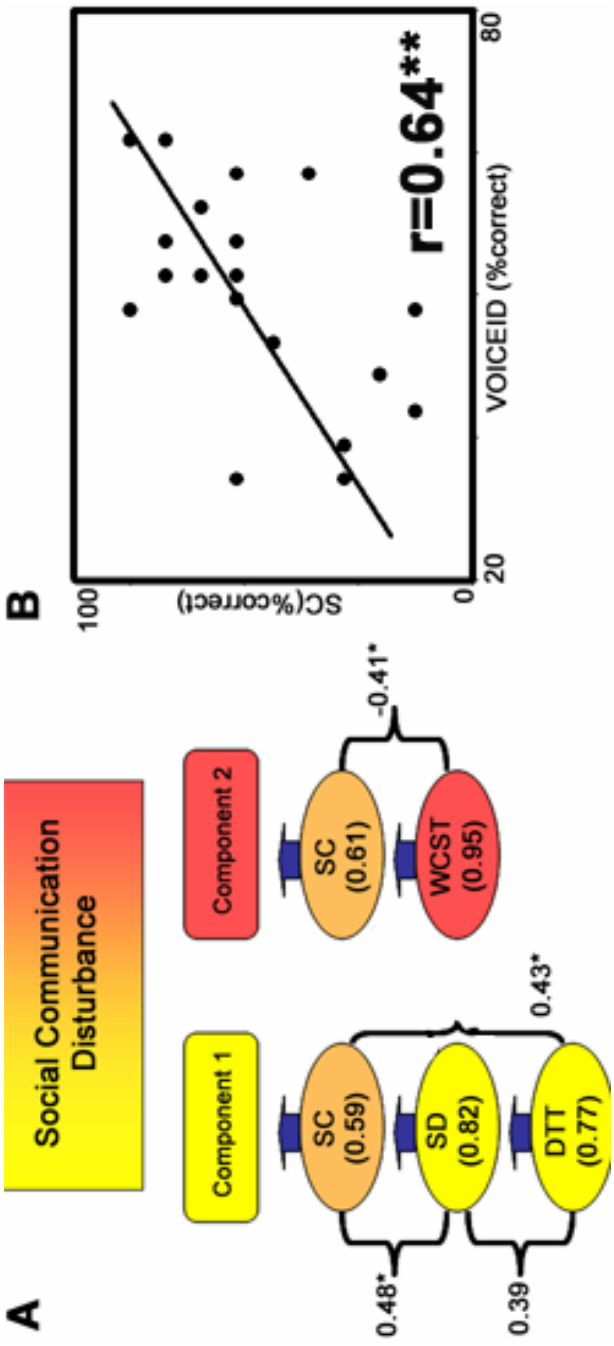
[1] (Between-group MANOVA across measures: $F_{4,31} = 46.6$, $p < .0001$)
*p < .04, ** p < .001,*** Significant using Man-Whityey non-parametrice tests p < .001

Figure 1: PCA on Correlations Between Pitch Perception and Non-Affective prosody performance.

Caption:

A: Schematic diagram of interrelationship between non-affective prosody measures, pitch perception and executive processing. Values inside circles represent factor rotation, while values outside circles represent Pearson correlation coefficients between indicated measures. *$p < .05$. SD and DTT rely exclusively on component 1 (sensory processing), WCST relies exclusively on component 2 (executive processing) and SC relies on both components.

B: Semantic comprehension (SC) by affective prosody (VOICE-ID).


A
Social Communication Disturbance
Component 1
Component 2
SC (0.59)
SD (0.82)
DTT (0.77)
SC (0.61)
WCST (0.95)
0.48*
0.39
0.43*
-0.41*
B
SC(%correct)
VOICEID (%correct)
100
0
20
80
r=0.64**

5.0 STUDY IV

Theory of Mind (ToM) and Counterfactuality Deficits in Schizophrenia: Misperception or Misinterpretation?

5.1 ABSTRACT

Background: Theory of Mind (ToM) refers to the ability to infer another person's mental state based upon interactional information. ToM deficits have been suggested to underlie critical aspects of social interaction failure in disorders such as autism and schizophrenia, although developing paradigms for demonstration of such deficits remains an ongoing area of research. Recent studies have explored the use of sarcasm perception, in which subjects must infer an individual's sincerity or lack thereof, as a "real life" index of ToM ability, and as an index of functioning of specific right hemispheric structures. Sarcasm detection ability has not previously been studied in schizophrenia, although patients have been shown to have deficits in the ability to decode emotional information from speech ("affective prosody").

Method: Twenty-two schizophrenia patients and seventeen control subjects were tested on their ability to detect sarcasm from spoken speech as well as measures of affective prosody and basic pitch perception.

Results: Despite normal overall intelligence, patients performed substantially worse than controls in the ability to detect sarcasm (d = 2.2), showing both decreased sensitivity (A`) in detection of sincerity vs. sarcasm and an increased bias (B``) toward sincerity. Correlations across groups revealed significant relationships between impairments in sarcasm recognition, affective prosody and basic pitch perception.

Conclusions: These findings demonstrate substantial deficits in ability to infer an internal subjective state based upon vocal modulation among subjects with schizophrenia. Deficits were related to, but were significantly more severe than, more general forms of prosodic and sensorial misperception, and are consistent with both right hemispheric and "bottom up" theories of the disorder.

5.2 INTRODUCTION

Cardinal to schizophrenia psychopathology is communicatory disturbance (Mitchell & Crow, 2005). This disturbance is multifaceted and can take the form of language disorganization as well as difficulty in decoding affect based upon either visual (i.e., facial), or auditory input. C.D. Frith (1992) and others have sought to tie these deficits to impairment in Theory of Mind (ToM), or the ability to infer one's own or another's internal mental state based upon behavioral interaction. In Frith's (1992) conceptualization, ToM-based impairments stem from a confabulation between one's subjective cognitive representations of reality and objective reality. Thus leading to false beliefs, delusions, and, within social interaction the negation of communicatory, cues vital for attitudinal and/or affective signals (C.D. Frith, 1992). Others see ToM deficits as reflecting more general impairment of core executive functioning ability and negative symptoms (Hardy-Bayle *et al.*, 1994), although such deficits typically remain significant even when controlling for such factors (Langdon *et al.*, 2002). ToM has become an important focus of current research as it provides a conceptual framework for unification of otherwise diverse elements of the schizophrenia symptomatology.

Classical methods for testing ToM involve the use of scenarios, either in the form of short stories, or graphically, through sequential picture sets in which social situations, jokes or ironic intent are conveyed. In these tasks, patients' comprehensions of the scenario, as well as their insight into the character's beliefs and/or intentions are assessed. Such studies have shown significant deficits in patients compared to controls in both childhood (Pilowsky *et al.*, 2000) and adult-onset schizophrenia (Mazza *et al.*, 2001) (see

review (Brune, 2005b)), although the statistical magnitude of such deficits is frequently limited by the complexity of the tasks involved.

Current research in social cognitive processing and schizophrenia has begun to link ToM ability to social dysfunction and global outcome in the illness: Roncone *et al*. (2002) found that ToM ability, using tasks that consisted of false belief scenarios presented via cartoons, correlated with global social functioning. However, Brune (Brune, 2005a) found that ToM defects were a predictor of social behavioral abnormalities using the social behavioral scale (Wykes, 1986). Similarly, a retrospective study found that ToM dysfunction in patients has also been associated with poor childhood socialization (Schenkel *et al.*, 2005). These studies suggest that ToM processes may significantly mediate social cognition and ability within schizophrenia, and that the failure to develop ToM during childhood may significantly impair outcome in later life.

More recently, there has been increased interest in attitudinal communication in speech as a more "real life" measure of ToM ability (Channon *et al.*, 2005; McDonald, 1999; Shamay-Tsoory *et al.*, 2005), especially communication of sarcastic vs. literal intent. Within everyday social interaction, attitudinal communications are frequently made in which the literal content of a series of words is negated by the modulation of stress and intonation. Thus, sentences such as "Now that was a good idea," when stated with primary stress and typical downward going, sentence-ending intonation on the last word typically convey their literal meaning, whereas the same sentence stated with stress on the second word and absence of usual intonation on the final word convey counterfactual intent. Communicating counterfactual intent in this manner is commonly referred to as "sarcasm".

Sarcasm can be distinguished from "irony" and general pragmatics, in that it communicates criticism, often of a particular target, and its prosody also conveys attitudes such as scorn or contempt (Jorgensen, 1996). The comprehension of sarcasm has been directly related to mentalizing ability (Winner & Leekam, 1991), in that comprehension of counterfactual intent is predicated on one's ability to cognitively represent the mental state of others as well as oneself (McDonald, 1999). More explicitly, Channon *et al*. have shown that sarcasm comprehension ability correlates strongly with non-linguistic ToM tasks when presented as written dialogue within a social scenario (Channon *et al*., 2005).

While sarcasm perception deficits have not previously been studied in schizophrenia, patients have been shown to have significant difficulty in decoding other interoceptive information, such as affect, from vocal intonation (prosody) (Kee *et al*., 2003; Kerr & Neale, 1993; E.D. Ross *et al*., 2001b). Given these findings, the present study investigates the degree to which patients with schizophrenia can infer counterfactuality of communication. Both affective prosody and sarcasm perception are thought to depend upon right hemisphere functioning (Kaplan *et al.*, 1990; Mitchell & Crow, 2005; Ozonoff & Miller, 1996; Shamay-Tsoory *et al*., 2005), with lesions studies suggesting that right prefrontal cortex (rPFC) integrates affective sensory information crucial for sarcasm perception (Shamay-Tsoory *et al*., 2005). Similarly, ToM function has been linked to medial PFC, the superior temporal sulcus and the temporo-parietal junction (Apperly *et al.*, 2004; C. D. Frith & Frith, 1999). These findings provide a potential neuroanatomical underpinning for the present results.

In prior studies, we have also demonstrated that affective dysprosodia in patients correlates with deficits in basic pitch perception, such as the ability to match tones following a brief delay, suggesting significant "bottom-up" contributions to higher order dysfunction (Leitman *et al*., 2005). Given that sarcastic intent utilizes tonal modulation, we tested the ability of subjects to perceive sarcasm in short vocal utterances, in concert with pitch and musical perception. We predicted that patients would show a significant impairment in the ability to detect sarcasm, with the impairment being worst in subjects with the poorest sensory performance, suggesting that misperception may contribute significantly to ToM deficits in schizophrenia.

5.3 METHODS

5.3.1 Participants

Twenty-two patients, meeting DSM-IV criteria for either schizophrenia or schizoaffective disorder, took part in this study (**Table 1**). Our patient sample was comprised of both inpatients (N = 15) and outpatients (N = 7). All patients were receiving antipsychotics (conventional or atypical) with a mean CPZ dose of 1461 ± 657 mg. Clinical assessment using the Brief Psychiatric Rating Scale (BPRS) (Overall & Gorham, 1961) and the Scale for the Assessment of Negative Symptoms (SANS) (N.C. Andreasen, 1984), indicated that our patients are severely ill (see **Table 1**). Both diagnoses and symptom ratings were conducted by trained clinicians, with diagnosis based on the Structured Clinical Interview for DSM-IV (SCID) and chart review (First *et al.*, 1997).

The control group consisted of seventeen healthy volunteers. These subjects were either volunteers who responded to a local advertisement or hospital staff and were all screened using the SCID. The procedures conducted were under the supervision of the local institutional review board and all subjects had the procedure explained to them verbally before signing informed consent and all subjects spoke English as their primary language.

5.3.2 Stimuli

Sarcasm perception was assessed using the attitudinal subtest of the Aprosodia Battery (APT) (Orbelo *et al.*, 2005). This battery consists of ten semantically neutral sentences such as “This looks like a safe boat” that were recorded by a female speaker in both a sincere or sarcastic manner for a total of twenty distinct utterances. These utter-

ances were repeated twice for a total forty stimuli. The stimuli were then randomly played via CD player in a sound attenuated room at a conversational hearing level. Subjects were instructed to answer after each sentence whether the speaker was being sincere or sarcastic. If subjects were confused by the instructions, further elaborations of the task instructions, using more commonplace synonyms was provided. Subjects score reflected the percent correct (Orbelo *et al.,* 2005). Additionally, scores were divided into the percent correct of sincere and sarcastic items.

Thus far, the APT has been used to assess attitudinal prosody in geriatric populations and is also currently used (Orbelo *et al.,* 2005) in neurological investigations of individuals with brain damage. In these studies, patient groups tend to perform worse than comparison subjects on both affective prosody as well as the APT. This finding, coupled with the fact that our subjects performed above chance and had good within and between group variance with no floor or ceiling effects (see results, **Figure 1**), suggests that this test is appropriate for examining individual and group differences in schizophrenia.

Affective prosodic ability was assessed using the Voice Emotion Identification Task (VOICE-ID) (Kerr & Neale, 1993). This test consists of 21 sentences of neutral content on audiotape. The sentences are spoken by male and female speakers, conveying one of six different emotions (happiness, anger, fear, sadness, surprise, or shame). Participants in a forced choice manner chose 1 of the 6 emotions for identification. Performance is calculated based upon the percentage of correctly identified sentences.

Sensory processing was assessed using two tasks: Pure tone matching threshold was obtained using the Tone Matching Task (TMT) (Leitman *et al*., 2005; E. F. Rabinowicz *et al*., 2000). This task uses an adaptive "up-down" transfer staircase method. Three rotating base frequencies of 5000, 1000 and 2000 hertz were used, with tone duration set at 100 milliseconds and an inter-stimulus interval between tones of 300 milliseconds. Threshold was assessed as the minimum reliable distinguishable difference between tones. Additionally, the Distorted Tunes Task (DTT) (Drayna *et al*., 2001), a task which has previously shown genetic contributions to pitch perception was used. This task consists of twenty popular tunes ranging in length from 12 to 26 notes. Seventeen of the tunes are rendered melodically incorrect by changing the pitch of two to nine notes within the tune. Subjects respond "yes" or "no" as to whether the melody is correct and are asked to report whether the melody was familiar or not. Patient score reflects the percentage of correctly categorized melodies.

Verbal (premorbid) IQ was collected using the quick IQ test (Ammons & Ammons, 1962). Due to constraints on patient availability, not all measures were collected on all patients. VOICE-ID and TMT scores have been published previously (Leitman *et al*., 2005).

5.3.3 Statistical Analysis

Between-group comparisons of sarcasm perception were assessed using an independent sample *t*-test and effect size calculation and conventions followed that of Cohen with a d value of 0.2, 0.5 and 0.8 reflecting the cutoffs for small medium and large effect sizes (Cohen, 1988). Additionally non-parametric signal detection measures of sensitivity

using A` after Snodgrass and Corwin (Snodgrass & Corwin, 1988):

$$A' = 0.5 + \left\{ \frac{(y-x)(1+y-x)}{4y(1-x)} \right\} \text{ when } y \geq x, \text{ or } 0.5 + \left\{ \frac{(x-y)(1+x-y)}{4x(1-y)} \right\} \text{ when } y \leq x.$$

and bias B`` using Grier's calculation (Grier, 1971):

$$B'' = \left\{ \frac{y(1-y)-x(1-x)}{y(1-y)+(1-x)} \right\} \text{ when } y \geq x, \text{ or } \left\{ \frac{x(1-x)-y(1-y)}{x(1-x)+(1-y)} \right\} \text{ when } y \leq x.$$

Where x is the probability of a falsely identifying a sincere item as sarcastic (False Alarms), and y is the probability of a correctly identifying a sarcastic utterance (hits) and hit >FA.

ANCOVA was used to assess the independence of attitudinal performance for factors of group membership IQ and gender. Logistic regression analysis was used to assess the significance and degree of between-group separation using our combined behavioral measures. In this way, we hoped to illustrate the degree of group differentiation possible, based on task performance. Correlation analysis of attitudinal perception, prosody and pitch measures was conducted across groups using Pearson correlation coefficient (r). Correlation analyses of behavioral measures were collected within the patient group using Pearson correlation coefficient (r) and Bonferroni correction for multiple comparisons.

Post hoc analysis examined correlations between symptom ratings and attitudinal prosody. For the BPRS and SANS ratings correlation analysis, we used global scores of subscales and total scores. All tests were two-tailed with Type I error ≤ 0.05, and statis-

tics were computed in JMP (SAS) software (2001).

5.4 RESULTS

5.4.1 Between-Group Analyses

As predicted, patients showed extremely robust deficits in their ability to perceive sarcasm, which was reflected in total score ($t_{37} = -6.72$, $p < .0001$), as well as percent correct of sarcastic items ($t_{37} = -5.22$, $p < .0001$) recognized and percent correct of sincere items ($t_{37} = -3.61$, $p < .0001$) (**Table 2**). Nevertheless, performance was significantly above chance, indicating ability to understand general task demands (**Figure 1**). Significant, but smaller effect size differences between patients and controls for VOICE-ID ($t_{30} = -4.12$, $p < .0003$), TMT ($t_{30} = 3.32$, $p < .002$) and DTT ($t_{29} = -2.97$, $p < .006$) were also found. Using logistic regression, a metric combining attitudinal prosodic performance and TMT threshold correctly identified 21/22 (95.4%) individuals with schizophrenia, as well as 11/12 (91.7%) controls for whom full data were available.

In order to further analyze the basis for the deficit, signal detection measures were computed for both sensitivity (A`) and bias (B``). Patients were significantly less sensitive than controls in detecting sarcasm ($t_{37} = 5.14$, $p < .001$). Further, patients showed significantly greater bias than controls toward identifying statements as being sincere even when they were not ($t_{37} = -2.45$, $p < .02$). An ANCOVA of attitudinal perception ability controlling for factors of group and IQ and gender revealed that patients sarcasm deficits were significant overall ($F_{1, 28} = 16.49$, $p < .0001$), but that there was a significant IQ effect after controlling for gender differences between groups ($F_{1, 28} = 4.36$, $p = .046$) however, there was no effect of gender ($F_{1, 28} = 0.35$, $p < .56$) (**Figure 2**).

5.4.2 Correlational Analysis

Poor performance on the APT (sarcasm) was significantly correlated with elevated tone matching thresholds ($r = -0.39$, $df = 27$, $p < .04$), DTT performance ($r = 0.40$, $df = 26$, $p < .05$) and poor affective prosodic performance ($r = 0.55$, $df = 27$, $p < .003$) across groups (**Figure 3**). Similar patterns of correlation were observed for A` [TMT ($r = -0.38$, $df = 27$ $p < .04$), DTT ($r = 0.43$, $df = 26$, $p < .02$), VOICE-ID ($r = 0.56$, $df = 27$, $p < .002$)], but not B``. No correlations were found within either group individually (all p's > 0.2). Of the across groups correlations, only the correlations between attitudinal perception and sensitivity (A`) and affective prosody remained significant following Bonferroni correction

Finally, sarcasm perception scores did not correlate with medication (CPZ equivalent) dosage or with positive or negative symptoms, as measured by the total BPRS and SANS scores, respectively (all p's > .1). However, there was a significant correlation between sarcasm perception and the SANS avolition factor ($r = -0.44$, $df = 21$, $p < .03$).

5.5 DISCUSSION

The primary finding of the present study is that patients show profound deficits in the ability to decode sarcasm based upon tone of voice. The effect size of this deficit ($d = 2.2$ sd units) was larger than that of affective prosodic deficits observed by ourselves (Leitman *et al*., 2005) and others (Kerr & Neale, 1993; E.D. Ross *et al*., 2001b), and larger than the "general" deficit of approximately 1.5 sd observed across cognitive domains in schizophrenia (Bilder *et al.*, 1991; Saykin *et al.*, 1991). Based upon signal detection analyses, this deficit consists of two components–a marked insensitivity to whether

statements are sarcastic or sincere ($d = 1.9$, $p < .001$), and a less pronounced bias toward accepting statements as sincere even when they were not ($d = 0.8$, $p < .02$). Additionally, the results of our ANCOVA indicated that patients have significant sarcasm perception deficits that could not be accounted for solely by IQ or gender differences, although higher IQ did significantly predict better sarcasm perception scores. Thus, our results are somewhat consistent with prior ToM findings (Mazza *et al*., 2001; Pickup & Frith, 2001), patients' ability to perceive sarcasm was independent of general intelligence, as reflected by verbal IQ, and may thus represent a specific feature of the disorder. Further, within patients, there were no significant correlations between sarcasm performance and medication dose, suggesting that sarcasm perception is unrelated to antipsychotic treatment.

A significant correlation was observed between sarcasm, affective prosody and musical pitch measures. While no significant correlation was found within the patient group itself, the scatter plot of the across groups correlation (**Figure 2**) reveals that patient and control performance tended to lie along a continua with control performance markedly better than patients on both measures. This finding is consistent with neuropsychological observations which suggest that emotion perception may mediate one's ability to infer counterfactual intent upon which sarcasm comprehension is predicated (Shamay-Tsoory *et al*., 2005).

The present findings of sarcasm deficits are limited by the fact that they represent a relatively small sample of patients who were primarily male and chronically ill. Furthermore, correlations between our measures of pitch and affective prosody were only significant across groups. Future research must assess whether sarcasm perception im-

pairment is detectable in recent onset and prodromal schizophrenia and whether sarcasm deficits represent a trait of schizophrenia.

Within the framework of ToM, these results can be interpreted in one of several ways. First, deficits in ToM have been linked to right prefrontal cortex (rPFC) dysfunction, based upon both evaluation of brain lesion (Shamay-Tsoory *et al*., 2005) and functional imaging (Vollm *et al.*, 2006) approaches. Alternatively, more general right hemisphere (RH) dysfunction has also been postulated to play a key role in schizophrenia in general (Kaplan *et al*., 1990; Mitchell & Crow, 2005). Finally, the ability to decode ToM information in the present task depends heavily upon the ability to perceive the relevant intonational cues that communicate counterfactuality. Support for all three possibilities can be found within the body of prior schizophrenia literature, suggesting that the severity of dysfunction in ToM observed using the present paradigm may reflect interactions among multiple overlapping and distributed systems (Vollm *et al*., 2006).

To the extent that the present deficits may reflect RH dysfunction, they are supportive of several recent theories proposing that RH dysfunction plays a central role in schizophrenia (Mitchell & Crow, 2005). In one study of RH damaged individuals (Kaplan *et al*., 1990), subjects were significantly impaired in detecting sarcastic, but not sincere utterances. This was similar to that observed in the present study, although signal-detection measures such as sensitivity (A`) and bias (B``) were not reported. Disturbances in RH function would be expected to give rise to the disturbances in both verbal (Leitman *et al*., 2005; E.D. Ross *et al*., 2001b) and facial affect recognition (J. C. Borod *et al*., 1993; David & Cutting, 1990; Heimberg *et al.*, 1992) that have been reported in schizophrenia. Indeed, behavioral studies in schizophrenia (Brune, 2005a), as well as fMRI re-

sults (Vollm *et al*., 2006), suggest substantial overlap between general affective cortical networks and specific ToM processing brain regions (Vollm *et al*., 2006).

Lastly, over recent years, deficits in basic auditory and visual processing have been extensively documented using both behavioral and neurophysiological paradigms. Patients with schizophrenia show significant deficits in ability to detect even substantial changes in pitch, as measured using both tone matching (E. F. Rabinowicz *et al*., 2000) and mismatch negativity (MMN) paradigms, (Javitt *et al*., 1995b). Within the visual system, patients show significant impairments in functioning of the magnocellular visual pathways, which is critical for decoding low contrast, low spatial frequency (P. D. Butler *et al*., 2001) and motion information (Kim *et al.*, 2005). Deficits in auditory processing have been shown previously to contribute significantly in a "bottom-up" fashion to impairments in affective identification (Leitman *et al*., 2005).

In the present study, a similar relationship between pitch perception, affective prosody and sarcasm performance is observed, suggesting a substantial "bottom-up" contribution. Interestingly, a similar relationship has been observed within the visual system, such that impairment in motion detection ability predicts impaired ability to decode ToM information from visually presented ("eyes test") stimuli (Kelemen *et al.*, 2005). More generally, in mentalizing experiments testing empathy, individuals with greater sensitivity to perceiving affective prosodic cues are also those with greater empathic self report (Pickett *et al.*, 2004). Thus, while there may be significant dysfunction in the evaluative systems of the brain such as the frontal cortex, a substantial source of ToM and sarcasm performance deficits may be related to the misperception of the cues themselves.

Finally, the correlations between patient clinical ratings of negative symptom subscales like "avolition" and our ToM measure replicate prior ToM schizophrenia findings and further underscore the theoretical implications for the role of affect in ToM processing mentioned above. Furthermore, prior research has shown strong correlations between basic sensory deficits and negative symptoms as well as global outcome measures, yet interpretation of these results remained enigmatic.

Like affective prosody (Leitman *et al*., 2005), impairments in perception of sarcasm may result from significant "bottom-up" antecedents. Such social communicatory impairment as the inability to recognize affect and attitudinal changes may have profound effects on global outcome in a society in which social communication is a precondition for working and living well with others. In this case, inability to recognize sarcasm may particularly lead individuals to feel like they have been misled following a counterfactual communication, rather than just belittled. The connotation of different inflections, such as those used to communicate emotional or attitudinal prosody, are not innate and must be learned during childhood and adolescence. To the extent that sensory dysfunction interferes with such learned associations, it may also undermine the ability to develop ToM concepts, such as counterfactual communication.

In addition to theoretical implications for schizophrenia, the present study underscores practical issues in clinical communication. Although use of sarcasm is never recommended in clinical communication, it is nevertheless a common feature of societal interaction. The present findings, more so than other ToM studies in schizophrenia, underscore the great difficulty that patients have in decoding counterfactual intent, and in addition, demonstrate a bias toward trust that may, in itself, contribute significantly to poor

social functioning in schizophrenia. Caregivers in general, however, may be unaware of the profundity of such deficits, given the ease with which most unaffected individuals detect sarcasm. The present findings thus not only weigh against use of intonation to convey counterfactuality of information during therapeutic communications in patients with schizophrenia, but also suggest that deficits in this real world ability should be appreciated by family members and other caregivers in order to minimize potential miscommunications that may lead to adverse outcomes in schizophrenia.

5.6 REFERENCES

Ammons, R. B., & Ammons, C. H. (1962). The quick test (qt): Provisional manual. *Psychological Reports, 11*, 111-162.

Andreasen, N. C. (1984). *The scale for the assessment of negative symptoms (sans)*. Iowa City: The University of Iowa.

Apperly, I. A., Samson, D., Chiavarino, C., & Humphreys, G. W. (2004). Frontal and temporo-parietal lobe contributions to theory of mind: Neuropsychological evidence from a false-belief task with reduced language and executive demands. *J Cogn Neurosci, 16*(10), 1773-1784.

Bilder, R. M., Lipschutz-Broch, L., Reiter, G., Geisler, S., Mayerhoff, D., & Lieberman, J. A. (1991). Neuropsychological deficits in the early course of first episode schizophrenia. *Schizophr. Res., 5*(3), 198-199.

Borod, J. C., Martin, C. C., Alpert, M., Brozgold, A., & Welkowitz, J. (1993). Perception of facial emotion in schizophrenic and right brain-damaged patients. *J Nerv Ment Dis, 181*(8), 494-502.

Brune, M. (2005a). Emotion recognition, 'theory of mind,' and social behavior in schizophrenia. *Psychiatry Res, 133*(2-3), 135-147.

Brune, M. (2005b). "theory of mind" in schizophrenia: A review of the literature. *Schizophr Bull, 31*(1), 21-42.

Butler, P. D., Schechter, I., Zemon, V., Schwartz, S. G., Greenstein, V. C., Gordon, J., *et al*. (2001). Dysfunction of early-stage visual processing in schizophrenia. *Am J Psychiatry, 158*(7), 1126-1133.

Channon, S., Pellijeff, A., & Rule, A. (2005). Social cognition after head injury: Sarcasm and theory of mind. *Brain Lang, 93*(2), 123-134.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences, 2nd edition*. Hillsdale, NJ: Lawrence Erlbaum Assoc.

David, A. S., & Cutting, J. C. (1990). Affect, affective disorder and schizophrenia. A neuropsychological investigation of right hemisphere function. *Br J Psychiatry, 156*, 491-495.

Drayna, D., Manichaikul, A., de Lange, M., Snieder, H., & Spector, T. (2001). Genetic correlates of musical pitch recognition in humans. *Science, 291*(5510), 1969-1972.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. (1997). *Structural clinical interview for dsm-iv axis i disorders (scid-iv)*. New York: Biometrics Research Department, New York State Psychiatric Institute.

Frith, C. D. (1992). *The cognitive neuropsychology of schizophrenia*. Hove UK: Laurence Earlbaum Associates.

Frith, C. D., & Frith, U. (1999). Interacting minds--a biological basis. *Science, 286*(5445), 1692-1695.

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychol Bull, 75*(6), 424-429.

Hardy-Bayle, M. C., Passerieux, C., Claudel, B., Olivier, V., & Chevalier, J. F. (1994). [communication disorders in schizophrenic patients. Cognitive explanation and clinical reconsideration]. *Encephale, 20*(4), 393-400.

Heimberg, C., Gur, R. E., Erwin, R. J., Shtasel, D. L., & Gur, R. C. (1992). Facial emotion discrimination: Iii. Behavioral findings in schizophrenia. *Psychiatry Res, 42*(3), 253-265.

Javitt, D. C., Shelley, A. M., Grochowski, S., & Ritter, W. (1995). Mismatch negativity (mmn) as an index of impaired auditory sensory memory in schizophrenia. *Schizophrenia Research, 15*, 179.

Jorgensen, J. (1996). The functions of sarcastic irony in speech. *Journal of Pragmatics, 26*, 613-634.

Kaplan, J. A., Brownell, H. H., Jacobs, J. R., & Gardner, H. (1990). The effects of right hemisphere damage on the pragmatic interpretation of conversational remarks. *Brain Lang, 38*(2), 315-333.

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

Kelemen, O., Erdelyi, R., Pataki, I., Benedek, G., Janka, Z., & Keri, S. (2005). Theory of mind and motion perception in schizophrenia. *Neuropsychology, 19*(4), 494-500.

Kerr, S. L., & Neale, J. M. (1993). Emotion perception in schizophrenia: Specific deficit or further evidence of generalized poor performance? *J Abnorm.Psychol., 102*(2), 312-318.

Kim, D., Zemon, V., Saperstein, A., Butler, P. D., & Javitt, D. C. (2005). Dysfunction of early-stage visual processing in schizophrenia: Harmonic analysis. *Schizophr Res, 76*(1), 55-65.

Langdon, R., Coltheart, M., Ward, P. B., & Catts, S. V. (2002). Disturbed communication in schizophrenia: The role of poor pragmatics and poor mind-reading. *Psychol Med, 32*(7), 1273-1284.

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. *Biol Psychiatry, 58*(1), 56-61.

Mazza, M., De Risio, A., Surian, L., Roncone, R., & Casacchia, M. (2001). Selective impairments of theory of mind in people with schizophrenia. *Schizophr Res, 47*(2-3), 299-308.

McDonald, S. (1999). Exploring the process of inference generation in sarcasm: A review of normal and clinical studies. *Brain Lang, 68*(3), 486-506.

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

Orbelo, D. M., Grim, M. A., Talbott, R. E., & Ross, E. D. (2005). Impaired comprehension of affective prosody in elderly subjects is not predicted by age-related hearing loss or age-related cognitive decline. *J Geriatr Psychiatry Neurol, 18*(1), 25-32.

Overall, J. E., & Gorham, D. E. (1961). The brief psychiatric rating scale. *Psychol Reports, 10*, 799-812.

Ozonoff, S., & Miller, J. N. (1996). An exploration of right-hemisphere contributions to the pragmatic impairments of autism. *Brain Lang, 52*(3), 411-434.

Pickett, C. L., Gardner, W. L., & Knowles, M. (2004). Getting a cue: The need to belong and enhanced sensitivity to social cues. *Pers Soc Psychol Bull, 30*(9), 1095-1107.

Pickup, G. J., & Frith, C. D. (2001). Theory of mind impairments in schizophrenia: Symptomatology, severity and specificity. *Psychol Med, 31*(2), 207-220.

Pilowsky, T., Yirmiya, N., Arbelle, S., & Mozes, T. (2000). Theory of mind abilities of children with schizophrenia, children with autism, and normally developing children. *Schizophr Res, 42*(2), 145-155.

Rabinowicz, E. F., Silipo, G., Goldman, R., & Javitt, D. C. (2000). Auditory sensory dysfunction in schizophrenia: Imprecision or distractibility? *Arch Gen Psychiatry, 57*(12), 1149-1155.

Roncone, R., Falloon, I. R., Mazza, M., De Risio, A., Pollice, R., Necozione, S., *et al*. (2002). Is theory of mind in schizophrenia more strongly associated with clinical and social functioning than with neurocognitive deficits? *Psychopathology, 35*(5), 280-288.

Ross, E. D., Orbelo, D. M., Cartwright, J., Hansel, S., Burgard, M., Testa, J. A., *et al*. (2001). Affective-prosodic deficits in schizophrenia: Profiles of patients with brain damage and comparison with relation to schizophrenic symptoms. *J Neurol Neurosurg.Psychiatry, 70*(5), 597-604.

Saykin, A. J., Gur, R. C., Gur, R. E., Mozley, P. D., Mozley, L. H., Resnick, S. M., *et al*. (1991). Neuropsychological function in schizophrenia. Selective impairment in memory and learning. *Arch.Gen.Psychiatry, 48*(7), 618-624.

Schenkel, L. S., Spaulding, W. D., & Silverstein, S. M. (2005). Poor premorbid social functioning and theory of mind deficit in schizophrenia: Evidence of reduced context processing? *J Psychiatr Res, 39*(5), 499-508.

Shamay-Tsoory, S. G., Tomer, R., & Aharon-Peretz, J. (2005). The neuroanatomical basis of understanding sarcasm and its relationship to social cognition. *Neuropsychology, 19*(3), 288-300.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experemental Psychology: General, 117*, 34-50.

Vollm, B. A., Taylor, A. N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., *et al*. (2006). Neuronal correlates of theory of mind and empathy: A functional magnetic resonance imaging study in a nonverbal task. *Neuroimage, 29*(1), 90-98.

Winner, E., & Leekam, S. (1991). Distinguishing irony from deception: Understanding the speakers's second order intention. *British Journal of Developmental Psychology, 9*, 257-270.

5.7 TABLES AND FIGURES

Table 1. *Demographic and clinical characteristics of healthy control and patient populations (values are mean ± S.D.)*

Demographic/clinical criteria	Control ($n=17$)	Schizophrenia ($n=22$)
Age (years)	34·8 ± 8·9	37.5 ± 11·2
Gender (M/F)	9/8	20/2
Verbal IQ	111·5 ± 10·4	96·0 ± 10·7
Illness duration	N.A.	11·4 ± 2·3
BPRS total scores	N.A.	43·6 ± 10·9
SANS total scores	N.A.	31·2 ± 9·1
SANS SS: affective flattening	N.A.	2·0 ± 0·9
SANS SS: ALOGIA	N.A.	1·6 ± 0·9
SANS SS: avolition-apathy	N.A.	2·1 ± 0·7
SANS SS: anhedonia-asociality	N.A.	2·4 ± 0·7

BPRS, Brief Psychiatric Rating Scale (Overall & Gorham, 1961); SANS, Scale for the Assessment of Negative Symptoms (Andreasen, 1984); SS, subscales; N.A., not applicable.

Table 2. *Summary of control and patient performance*

Measure	Group	*n*	Mean	s.d.	Effect size (*d*)
Sarcasm	Schizophrenia	22	65·6	12·2	2·2
	Control	17	88·9	8·5	
Sarcasm-*A′*	Schizophrenia	22	0·73	0·05	1·9
	Control	17	0·94	0·48	
Sarcasm-*B″*	Schizophrenia	22	0·35	0·39	0·8
	Control	17	0·68	0·45	
VOICEID	Schizophrenia	21	48·3	16·8	1·6
	Control	9	71·4	10·4	
TMT	Schizophrenia	21	1·01	0·57	1·3
	Control	9	0·41	0·28	
DTT	Schizophrenia	20	74·7	17·2	1·2
	Control	8	92·7	8·0	

VOICEID, Voice Emotion Identification Test (Kerr & Neale, 1993); TMT, Tone Matching Threshold (Leitman *et al*. 2005); DTT, Distorted Tunes Test (Drayna *et al*. 2001).

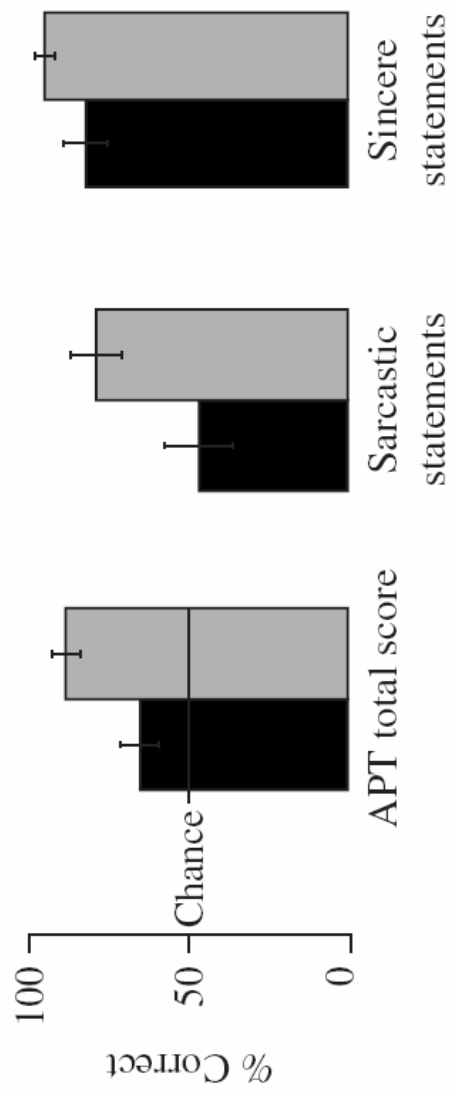


FIG. 1. Per cent correct of sincere (20) and sarcastic (20) statements ($p<0{\cdot}0001$). Error bars represent S.D. for the mean scores represented in the graph. ■, Schizophrenia patients; □, healthy controls.

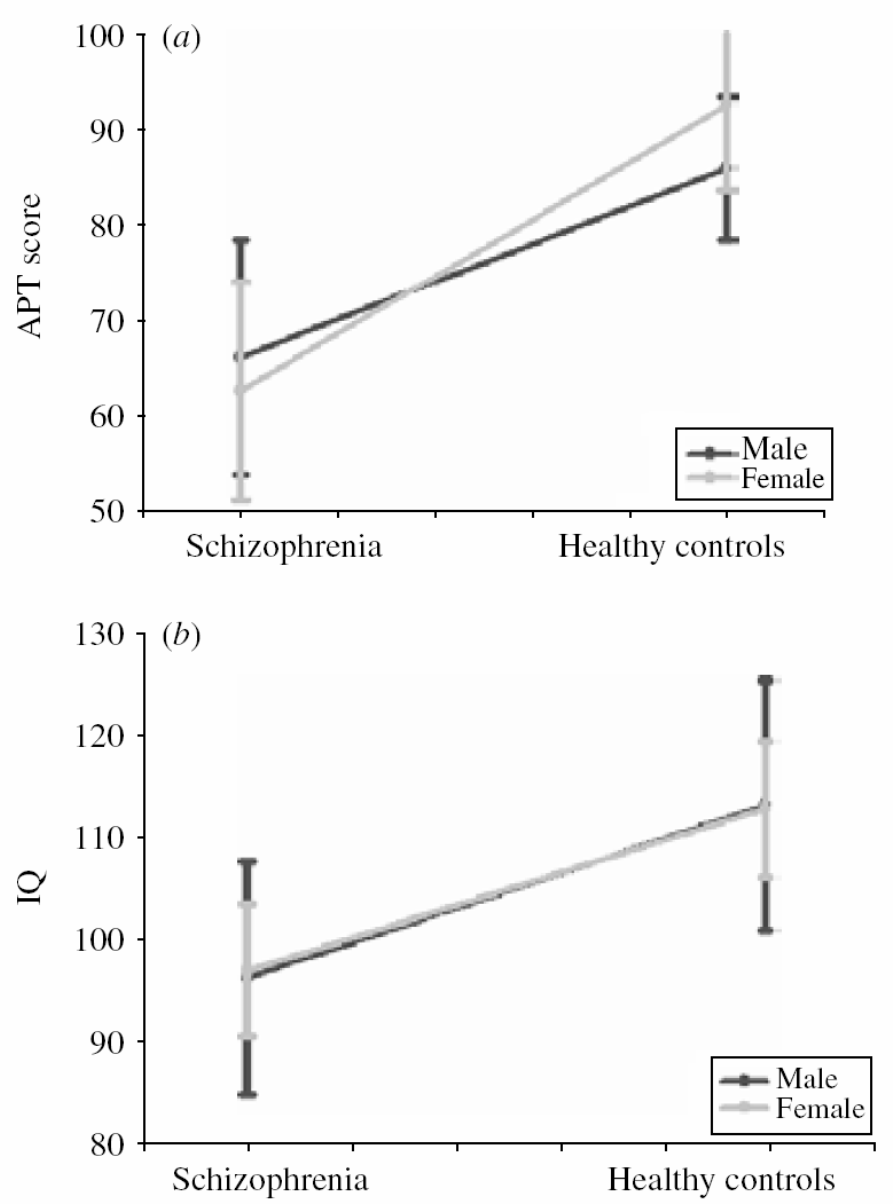


FIG. 2. Gender effects on (*a*) sarcasm perception and (*b*) IQ score. Error bars represent S.D. for the mean scores represented in the graph.

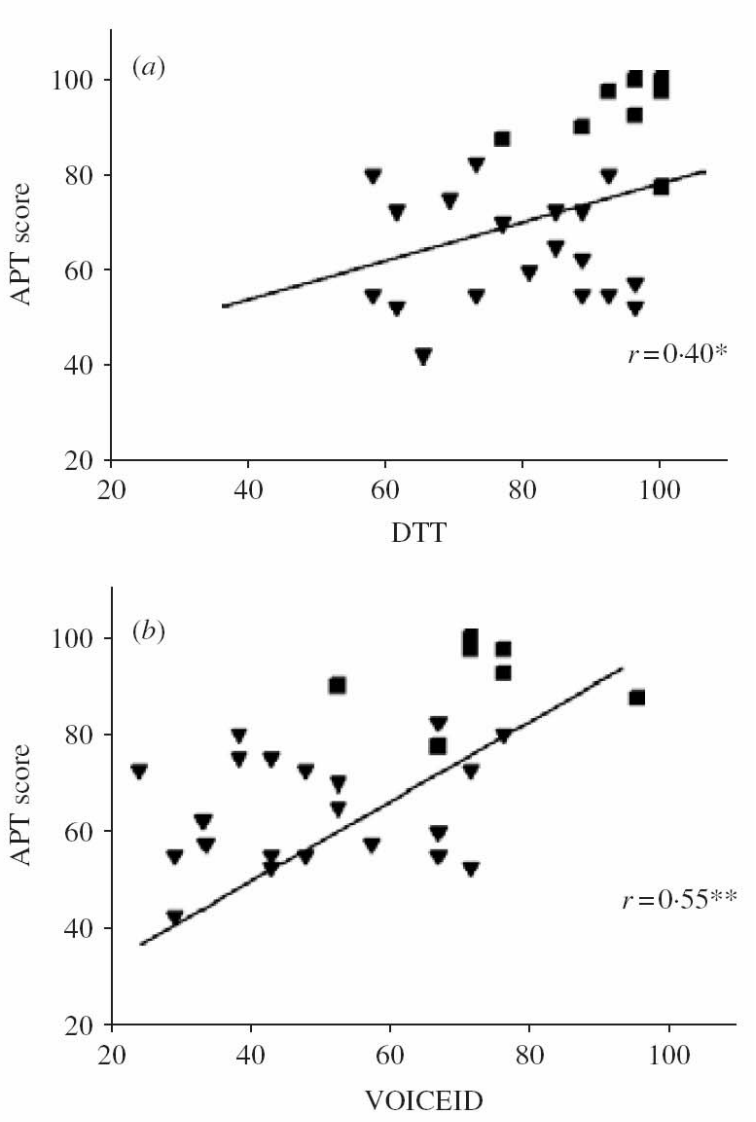


FIG. 3. Scatterplots of (*a*) Distorted Tunes Test (DTT) and (*b*) affective prosody Voice Emotion Identification Task (VOICEID) by attitudinal perception score (APT) across groups ($p<0{\cdot}05$). These figures illustrate in general good separation in performance for both DTT *versus* APT and VOICEID *versus* APT. ■, Healthy controls; ▼, schizophrenia patients.

6.0 GENERAL DISCUSSION

6.0 GENERAL DISCUSSION

6.1 Main Findings

This thesis examined the nature of prosodic deficits in schizophrenia. The goal was to examine the relationship between receptive dysprosodia, and sensory, as well as, cognitive disturbance. We then asked whether prosodic dysfunction and sensory disturbance had shared neural correlates. Our findings in this study suggest that affective prosodic dysfunction stems, at least in part, from dysfunction in auditory pathway neurocircuitry. We then examined the parameters of respective dysprosodia in the illness. In particular, we sought to relate affective dysprosodia to the perception of counterfactual intent in sarcasm. We also hoped to demonstrate that dysprosodia in schizophrenia extends to non-affective prosodic communication, such as indicating interrogative or declarative intent, or pragmatics like stressing the subject or object within a statement.

The results of these four studies indicate large effect size deficits in chronic patients across all domains of prosodic communication. The most severe deficits were found in the communication of sarcasm (2.2 sd), followed by affective prosodic deficits (1.6-1.8 sd) and lastly non-affective prosodic deficits (1.1 sd). These effect size differences across prosodic domains could stem from the degree of cognitive complexity and/or the extent of the neural systems involved in the evaluation of these differing prosodic forms. For example, sarcasm perception, which is thought to rely, in part, on ToM ability (McDonald, 1999) and medial inferior frontal functioning (Shamay-Tsoory *et al*., 2005), can be thought of as being more cognitively intense and/or involving more extensive neural resources, than say, distinguishing questions from statements.

Alternatively, a reductionist perspective would suggest that differential performance across prosodic domains might stem from the acoustical analysis demands present in the stimuli themselves. As the spectrographic aanalysis illustrates, declarative-interrogative distinctions rely on a relatively isolated change in pitch rise at the end of the speech token. Acoustical analysis of affective prosodic distinctions are much more complex, relying on a constellation of changes in pitch, voice intensity, speech rate and high frequency energy (Juslin & Laukka, 2001; Scherer, 1986). This is true for sarcastic–sincere prosodic distinctions as well. These two explanations of cognitive processing load or acoustical analysis complexity are likely two sides of the same coin, in that the nature of the acoustic signal drives the recruitments of greater neuronal systems and arrays.

6.2 Study Limitations, Methodological Issues and Future Approaches

The limitations of these studies fall into two categories. The first concern is the population studied and our ability to generalize to the schizophrenia population as a whole. This is a serious concern; our patients were chronically ill, on heavy does of antipsychotic medication and predominantly male. While our studies found no correlation between medication dosages or gender, medication effects cannot be ruled out and gender differences need to be more fully explored. Thus, it is not clear if the full range of observed prosodic deficits in our patient population were present before illness onset, and/or whether medication exacerbates them. Prior research (see introduction) by Kerr and Neale (Kerr & Neale, 1993) and Edwards (Edwards *et al*., 2002; Edwards *et al*., 2001) and others (Kucharska-Pietura *et al*., 2005) has suggested that prosody is a trait present in

early psychosis, if not before, and is not attributable to medication, but the current studies presented in this thesis cannot rule out this possibility.

After four years of collecting prosodic data, we now have the power to examine gender issues, and we have plans to examine prodromal, as well as first episode patients as well as schizophrenia and schizoaffective subtypes, using the prosodic tasks presented in this thesis.

An extension of these concerns, and one that is germane to studies of behavior in schizophrenia in general, is the heterogeneous nature of the schizophrenia diagnosis. Extensive research over the last half century has shown that schizophrenia symptomatology can have many etiopathic origins, including neurological insults(Kelly *et al.*, 2003), obstetric complications(Marcus *et al.*, 1993), prenatal diet,(Susser & Lin, 1992) advanced paternal age(Malaspina *et al.*, 2001) and early life events(Corcoran *et al.*, 2001). Further, study of prosodic deficits should be conducted to see if aprosodia is related more specifically to illness stemming from any particular etiopathic origin.

A second concern is the tasks themselves, and the ecological validity for presenting posed prosodic utterances in the absence of context. This is a well known concern in vocal and facial affect research, and has been addressed at length by Scherer (Scherer, 1986). In terms of ecological validity, the absence of context calls into question whether judgments of affect under such artificial circumstances, realistically approximate prosodic dysfunction in the illness. However, studies (Pickett *et al*., 2004) have shown correlations between prosodic ability and social competence in healthy individuals. Further, measures such as the Abrams emotional blunting scale (Abrams & Taylor, 1978) have correlated with prosodic deficits in Parkinson's Disease (L.X. Blonder *et al*., 1989; Wein-

traub *et al*., 1981). Similarly, within schizophrenia, studies have shown significant associations between prosodic measures and clinical ratings, such as the schedule for affective and negative symptoms (SANS) (J. S. Brekke *et al.*, 2002; Hooker & Park, 2002; Leitman *et al*., 2005).

It is also important to note that current prosody tasks (with the exception of the Ross Aprosodia battery) explore primary emotions with clear boundaries such as happiness, sadness, or anger. Within actual day-to-day communication, however, prosodic cues may be more subtle, conveying irony or sarcasm, for example. Thus, consistent deficits on prosodic tasks examining basic and starkly contrasting emotions may reflect more profound ecological emotion dysfunction.

Our studies consistently found high performance variability within patients across prosodic measures. In this respect, our findings are ubiquitous of patient performance on a wide array of neuropsychological tests (Heinrichs & Zakzanis, 1998) as well as on sensory measures (E. F. Rabinowicz *et al*., 2000). This variability is no doubt reflective of the heterogeneous nature of the illness. Given the increasing evidence of dysprosodia as a trait, further characterization of these deficits, in relation to negative symptoms, may provide clues about the etiology of schizophrenia, as well as new directions for clinical intervention.

As with many tests of affect, it is sometimes difficult to fractionate arousal and valence aspects of affect recognition. Arousal differences between groups are of particular concern, as measures of Hypothalamic Pituitary Axis (HPA) function, such as cortisol (Corcoran *et al.*, 2001; Corcoran *et al.*, 2002), heart rate variability (HRV) (Malaspina *et al.*, 2002), as well as amygdala structure (Bogerts *et al.*, 1990) and function have been

shown to be abnormal in schizophrenic patients, and associated with emotion perception in the illness (Exner *et al*., 2004; Fahim *et al.*, 2005). Future research should examine prosodic perception in conjunction with measures of arousal.

Future research should also focus on using functional neuroimaging approaches such as fMRI and ERP to examine the cortical dynamics of prosodic disturbance(s?). Clinically, these studies suggest that prosody may be a worthy target for cognitive remediation. We are currently proceeding along these two fronts, and have high expectations that the resulting data will prove clinically beneficial, as well as, etiologically informative.

One final comment concerns motor theories of prosodic dysfunction. This thesis has not been adequately explored in schizophrenia. Studies in Parkinsonism (L. X. Blonder *et al.*, 1991; L.X. Blonder *et al*., 1989; L.X. Blonder *et al.*, 1995; Weintraub *et al*., 1981) have clearly demonstrated that motor and basal ganglia dysfunction can contribute to receptive, as well as, expressive prosody deficits. Given the known abnormalities in basal ganglia (Heimer, 2000) functioning and dopaminergic transmission in schizophrenia, motor contributions to aprosodia seem likely. Prior researchers (M. Alpert *et al*., 2000; J. C. Borod *et al.*, 1989; Fricchione *et al*., 1986) have indeed raised the question of motor dysfunction. However, a clear connection between motor impairment and dysprosodia in schizophrenia has not yet been shown. Future schizophrenia studies should examine prosody in patients in conjunction with indexes of motor functioning.

7.0 CONCLUSION

The results of this thesis indicate that patients with schizophrenia have severe deficits in the ability to perceive affective as well as non-affective prosody. These deficits correlated with both sensorial disturbance and executive processing dysfunction. These findings are consistent with Scherer's (1986, 2003) SEC model, in that this thesis suggests that impaired perception of vocal cues, like pitch change hinder patient appraisal of vocal affect. This leads to a cognitive inability to "mentalize" or simulate the emotion itself and to discern the sender's intent, as patients must rely on executive processes that are already compromised.

These findings further suggest that rather than reflecting emotional impairment, affective prosody and its non-affective counterpart, reflects more generalized social communicatory dysfunction. This distinction is not mere semantics, but has important conceptual implications. E. Colin Cherry (Cherry, 1978), in his seminal treatise-"On Human Communication," wrote that all communication is not the sending of signs or symbols but the *sharing* of them (ibid, pg-305). Thus, communication is a form of "communion", or social bonding whereby information, feelings, and isolation are transferred and overcome.

Within schizophrenia, dysprosodia means that not only, is the sharing of meaning impaired, but also, that the sharing of others' intent is also inhibited. At first glance, such deficits seem removed or distinct from the general neurocognitive dysfunction commonly observed in patients. Yet, as Susan Langer[*] points out, cognition as a concept basically measures one's ability to process 'symbols.' Some are *discursive* like mathematical equa-

[*] I am indebted to Michael Beldoch for his insight from a paper entitled, "The Sensitivity to Expression of Emotion in Three Modalities," found in (Davitz, 1964).

tions or words that have objective definitions. *Non-discursive* symbols, however, like art or the emotion communication are referential – that is, they communicate the intent of the person expressing them. To Langer, these symbols are no more mysterious or impervious to systemic analysis, but just rely on a different set of logic. In our opinion, this perspective provides a unifying approach to the plethora of cognitive and social deficits that complicate effective treatment of schizophrenia. This myriad of deficits, together, can be seen as a general form of symbol processing dysfunction, in which social communication, due to its non-discursive (i.e. self-referential) nature requires greater abstraction, and is, thus, differentially impaired.

8.0 Appendix I

Mismatch Negativity to Tonal Contours Suggests

Pre-Attentive Perception of Prosody

8.1 ABSTRACT

Modulation of speech conveys information that is decoded within audio-sensory structures. For example, statements and questions are distinguished by pitch contours that encode semantic prosody. This study evaluated the sensitivity of early auditory structures to semantic prosody using mismatch negativity (MMN), an auditory event-related potential (ERP) sensitive to preattentive stimulus deviance. High-density ERP to pitch contour stimuli were collected in a passive listening oddball paradigm from 11 healthy subjects. Voltage analysis revealed significant MMN responses to declarative and interrogative oddball stimuli. Further, MMN was significantly larger to interrogative, than declarative, deviants, indicating non-symmetric brain processing. The elicitation of the MMN demonstrates that tonal-contour patterns that are ecologically valid abstractions of semantic prosody can be represented in pre-attentive auditory sensory memory.

8.2 INTRODUCTION

The majority of information conveyed by speech is encapsulated within individual segments that are decoded successively into phonemes, words, and sentences. However, additional information is conveyed not only in what is said, but also in how it is stated. This suprasegmental modulation of speech contains information regarding various forms of prosodies, including, for example, whether the speaker is happy or sad (emotional prosody), or whether the utterance is a statement or a question, which is a form of semantic prosody.

Although much is known about mechanisms by which segmental information in the brain is decoded into phonemes and words, relatively less is known about suprasegmental decoding mechanisms. In particular, whereas it is known that phonemes are processed even by preattentive auditory mechanisms (Aaltonen *et al.*, 1987), only one study (Kujala *et al.*, 2005), to our knowledge, has been performed with respect to prosodic information. There is a good reason to think that prosodic cues might be processed preattentively. For example, developmental studies have suggested that prosodic cues may be detected even by prelinguistic babies, and have high interactive salience (Fernald, 1985). We hypothesized, therefore, that suprasegmental information such as declarative/interrogative distinctions would be processed automatically within low-level auditory cortical regions.

One of the most effective methods for analysis of the locus of processing of auditory information is through the analysis of mismatch negativity (MMN). MMN is an auditory ERP elicited most commonly in the context of an auditory oddball paradigm, in which a series of standard stimuli is interrupted by an infrequent deviant stimulus. In

such a paradigm, the brain automatically organizes a template reflecting invariant features of the repetitive standards. MMN then reflects the outcome of a local process that compares each stimulus to the locally maintained template. MMN usually occurs with a latency of approximately 100-200 ms, which varies based upon degree of deviance. One of the characteristic features of MMN is that latency is locked to the timing of feature deviance, rather than to stimulus onset. Thus, for example, MMN to duration deviance is delayed relative to pitch deviance, because pitch deviance can be detected very quickly after stimulus onset whereas duration deviance cannot be detected until after the offset of the deviant stimulus or the normal offset of the standard stimulus (whichever is shorter). MMN is thought to reflect the operation of neural mechanisms within auditory cortex for directing attention toward potentially significant alterations within the surrounding acoustic environment (Naatanen, 1995). MMN deficits in pitch and duration perception have been found to index aberrant audio-sensory processing in clinical populations most notably in schizophrenia (Javitt, 2000).

One key acoustical cue of speech that is used to decode prosodic intent is the contour, or trajectory of fundamental frequency (F_0), across the individual segments. In general, terminally ascending contours indicate "interrogative" intent while contours with a flat or slightly downward trajectory indicate declarative intent (Majewski & Blasdell, 1969). Suprasegmental F_0 contours alone have been shown to be sufficient for discerning interrogative or declarative intent (Majewski & Blasdell, 1969) as well as for emotional prosodic comprehension (Lakshminarayanan *et al.*, 2003; Majewski & Blasdell, 1969), even if underlying speech segments are masked. Further, in studies of dysprosodia resulting from brain lesions, Van Lancker and colleagues have shown that individuals with

prosodic dysfunction were also poor at using F_0 cues, suggesting that F_0 contour recognition may be among the skills necessary for proper prosodic comprehension (Van Lancker & Sidtis, 1992).

For the present study, the sensitivity of MMN generators to suprasegmental information was evaluated by construction of artificial stimuli that approximated the F_0 modulations signifying interrogative and declarative intent, as previously reported by Pell (Pell, 1998). The duration and the frequency of the tones that formed each contour were matched save for the terminal tone, which pitched either upward or downward to approximate the contours of interrogative versus declarative utterances, respectively. Thus, the interrogative and declarative stimuli, each of which had a total duration of 428 ms, were identical for the first 328 ms, but deviated in frequency during the final 100 ms. Analyses of suprasegmental MMN, therefore, were performed relative to onset of stimulus deviance at 328 ms.

Interrogative and declarative stimuli served as deviants in alternate runs. Because the interrogative and declarative stimuli differed in overall stimulus energy, MMN waveforms were derived by comparing ERP responses to deviant stimuli (interrogative or declarative) in one run to ERP responses to the same stimulus type in the alternate run. Thus, all MMNs in the present study were derived by subtracting like-from-like stimuli, in that the same stimulus served as both deviant and standard in alternating blocks.

8.3 METHODS

8.3.1 Participants

Informed consent was obtained from 14 (6 female) healthy control subjects with a mean age of 33 ± 11 yrs. All subjects reported that they were right handed, had normal hearing, and were medication free at the time of testing. Three subjects were excluded from analysis due to high levels of noise within their data. All procedures conducted were under the supervision of the local institutional review board.

8.3.2 Stimuli and task

Subjects were presented with two-tonal contours: an interrogative contour and a declarative contour. Each contour consisted of 5 sequential sinusoidal tones each of which had a 10ms envelope created using a Hanning window. The first four of the five tones that made up both the declarative and interrogative tonal contours were identical, matched for both duration and frequency (**Table 1**). Using an inter-stimulus interval (ISI) of 500 ms between tonal contours across all presentations, two types of oddball blocks were presented: an "interrogative" deviant block and a "declarative" deviant block. In the interrogative block, three declarative contours were followed by an interrogative contour in a fixed manner. In the declarative deviant block the contour types were reversed with three interrogative standards followed by a declarative deviant. For the first four subjects, runs included an additional deviant in which the entire stimulus was shifted upward by 100 Hz (pitch deviance). However, this deviant was subsequently omitted in the interest of time.

Four blocks of each type (declarative and interrogative) were presented each of which contained 240 standard contours and 80 deviant contours, for a total of 960 standards and 320 deviants. All comparisons were made across blocks, with the response to the deviant stimulus in one block, being compared to the response to the same stimulus in the opposite block. All tonal contours were presented binaurally at 75db (SPL) through Sennheiser HD 600 headphones. Subjects were instructed that the experiment was designed to test their passive auditory responses to tonal sequences to which they need not attend. Subjects watched a silent movie during the course of stimulus presentation.

8.3.3 Data Collection

High-density event-related potentials (ERP) were recorded continuously from 128 scalp electrodes referenced to nose with bandwidth of 0.5 to 100 Hz and digitized at a sampling rate of 500 Hz. Impedances were kept to < 5 kΩ.

Epochs (–200 to 700 ms relative to stimulus onset) were constructed off-line. Trials with blinks and large eye movements were rejected off-line on the basis of horizontal (HEOG) and vertical (VEOG) electro-oculogram. No systematic differences in HEOG or VEOG were seen across conditions (artifact rejection window of ± 100 μV). An artifact criterion of ± 100 μV was used at all other electrode sites to reject trials with excessive EMG or other noise transients from –100 ms pre-stimulus to 450 ms post-stimulus.

Accepted trials were averaged for each subject. The average number of accepted sweeps for deviant contours per condition was 250 ± 39. For average files, baselines were corrected to zero over the –100 to 0 ms latency range. For source analysis, average files

were filtered using a 0.5-45 Hz zero-phase-shift band-pass digital filter with roll-off of 24 db/oct.

8.3.4 Statistical Analyses

Separate statistical analyses were performed for the interrogative and declarative deviance conditions. For each deviance type, point-wise ("running") paired t-tests (2-tailed) were calculated to detect significant differences between responses to stimuli presented as deviants vs. the same stimulus (interrogative/declarative) presented as a standard. In order to protect against type I error due to multiple comparisons, we employed a significance criterion requiring at least 10 consecutive data points (= 20 ms at a 500 Hz digitization rate) to meet an 0.05 alpha criterion threshold (Guthrie & Buchwald, 1991). MMN onset and offset were defined respectively as the first and last points at which there was a statistically significant difference between standard and deviant conditions at electrode Fz, provided that this difference continued for 10 consecutive data points.

Peak amplitudes for each deviance type were determined for each subject within the overall running-t significance window. Amplitudes were defined as the most negative value occurring within the running-t significance window. Topographical analyses of MMN distribution to interrogative vs. declarative contours were conducted using ANOVA with factors of deviance type and electrode location. For lateralization analyses, lateralized amplitudes were determined by summing across 3 electrode pairs located to the left and right of Fz (approx. equivalent to locations F3 and F4). For anterior posterior analyses, 11 frontocentral electrode locations ranging from AFz to Cz were used, and

analyses were corrected for non-sphericity using the Greenhouse–Geisser method. All significance levels in text are 2-tailed with preset α-level for significance of $p < 0.05$.

Source localizations were estimated using Local Auto-Regressive Average (LAURA) modeling of the surface voltage topography (Grave de Peralta & Gonzalez, 2002), to provide a distributed inverse solution estimate of the location of underlying generator regions.

8.4 RESULTS

Statistical comparisons between conditions were performed using running t-tests relative to the point at which the interrogative and declarative stimuli diverged. A comparison of the interrogative contour in the deviant versus standard position revealed a significant negative difference (MMN waveform), with an onset latency of 398 ms post stimulus onset, corresponding to 70 ms post deviance onset (PDO). This negative difference (MMN) had a duration of 80 milliseconds using the running t-test method described above and a peak amplitude of -3.5 ± 1.2 μV ($t = 3.34$, $p < .0001$) at 112ms PDO. A comparison of the declarative contour in the deviant versus standard position also revealed a significant negative difference waveform with an onset latency at 432 ms post stimulus onset, corresponding to 104 milliseconds post deviance onset. The MMN waveform had a duration of 102 ms, and a peak amplitude of -1.2 ± 0.9 μV ($t = 1.20$, $p = .001$) at 140 ms PDO (**Figure 1**). Statistical comparison of onset latencies demonstrated a significantly earlier onset time for interrogative, than declarative, MMN (Mean difference = 38 ± 11.0 ms, $t_{1,10.} = 12.9$, $p = .0001$).

Scalp voltage distributions of MMN are shown in **Figure 2**. Separate ANOVAs were conducted to compare distributions of interrogative vs. declarative MMN across lateral and anterior/posterior dimensions. For lateralization analyses, MMN values were averaged for 3 pairs of electrodes centered around Fz (**Table 2**). Analysis was conducted using ANOVA with factors of stimulus type and hemisphere. This analysis indicated no significant hemispheric difference between deviant types ($F_{1,10} = 0.87$, $p = .37$), but did indicate that the interrogative contour mismatch was significantly larger than its declarative counterpart ($F_{1,10.} = 11.94$, $p < .01$). Anterior posterior analyses were performed using an ANOVA for factors of deviant type and electrode location. Electrodes that straddled the midline between from posterior locations (Fz) to anterior locations (AFz) were used for this analysis.

This analysis revealed a significant interaction between electrode location and stimulus type ($F_{10,100.} = 8.89$, $p = .01$, $\varepsilon = 0.26$), reflecting a more anterior topography for MMN to the declarative vs. interrogative deviance.

In general, MMN waveforms are thought to arise from auditory sensory cortex. Source localization (LAURA) analyses performed at the point of maximum amplitude for each MMN waveform revealed bilateral distributed sources that were located within superior temporal gyrus for each MMN type (**Figure 3**), confirming *a priori* expectations.

8.5 DISCUSSION

MMN shows well-known sensitivity to segmental aspects of speech, such as phonemic structure (Aaltonen et al., 1987). This study examined whether MMN shows similar sensitivity to suprasegmental aspects of speech, using contours constructed to ap-

proximate tonal contours associated with declarative vs. interrogative utterances. A comparison of the subtraction waves across conditions found significant MMN-like activity in both conditions time-locked to the declarative/interrogative contour. These findings support the concept that, like other elements of speech, the suprasegmental contour information used to infer interrogative prosody can be decoded preattentively within low-level auditory cortical regions. This result is consistent with prior studies showing sensitivity of MMN to, for example, changes in stimulus pattern (e.g., (Saarinen *et al.*, 1992)). However, this study is the first to utilize tonal contours resembling those of normal declarative and interrogative utterances, and the first to show that such information is decoded against a background of more complex spectral information.

MMN was elicited whether repetitive interrogative stimuli were presented against a background of declarative stimuli, or whether declarative stimuli were presented against a background of interrogatives. In both cases, comparisons were made between the same stimulus (i.e., interrogative/declarative) presented as a standard and the same stimulus presented as a deviant in a separate run. Thus, the spectral content of standards and deviants within each comparison was identical. Further, because of the symmetrical experimental design, the onset of deviance within each run (i.e. interrogative std/declarative deviant vs. declarative std/interrogative deviant) was identical, with the contours being identical up to 328 ms and diverging thereafter.

Despite the exact symmetry in the experimental design, significant differences were observed in both the timing and amplitude of contour-elicited MMN, such that interrogative contours presented against a background of declarative contours elicited a larger MMN than declarative contours presented against a background of interrogative stan-

dards. This finding suggests that the MMN generators are responding not just to contour per se, but also to the ecological significance of the suprasegmental information contained within the contour. Commonly, in both educational and social situations, individuals are able to detect when they are being asked a question, even after they have stopped paying close attention to the verbal material directed at them. The large MMN elicited by interrogative stimuli presented against a background of declarative statements may underlie the ability of questions to automatically capture attention even when the preceding declarative information has been ignored.

An alternative explanation for the larger MMN to interrogative vs. declarative stimuli is based on Western music theory, in which termination of a musical sequence on an ascending note does not provide the closure found on ending on a descending note. MMN amplitude differences between the interrogative and declarative contour may thus be simply due to the fact that frequency declension elicits a smaller MMN than frequency ascension. In such case, the greater automatic salience of ascending notes may have led to its adoption to signify the interrogative within most Western societies.

Although we are not aware of other MMN studies performed using interrogative prosodic contours, Kujala *et al*. (Kujala et al., 2005) recently demonstrated MMN to affective prosody. In that study, MMNs were elicited to deviant presentations of "commanding", "sad" and "scornful" prosodic presentations of single words when contrasted with neutral standard presentations. MMN occurred with a latencies ranging from 178 ms to 312 ms, which given the deviance onset between standard and deviant presentations, is not inconsistent with our findings here. Thus, MMN generators may be responsive, in general, to alterations in prosody over and above sensitivity simply to the underlying

spectral properties of the stimuli, suggesting that prosodic information may be decoded, at least in part, in a preattentive, attention-independent fashion.

Deficits in MMN generation to simple frequency and duration deviances have been reported in pathological conditions, such as schizophrenia, and have been shown to correlated with impaired discrimination of basic tonal deviances (Javitt, 2000). In addition, patients with schizophrenia show impaired decoding of auditory prosodic information (Leitman *et al.*, 2005). The present findings suggest that MMN may be useful for assessing the locus of dysfunction in conditions that are associated with impaired detection of prosodic information. such as schizophrenia, autism, and hemi-parkinsonism.

8.6 REFERENCES

Aaltonen, O., Niemi, P., Nyrke, T., & Tuhkanen, M. (1987). Event-related brain potentials and the perception of a phonetic continuum. *Biol Psychol, 24*(3), 197-207.

Fernald, A. (1985). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20*(1), Jan-113.

Grave de Peralta, R., & Gonzalez, A. (2002). Comparison of algorithms for the localization of focal sources: Evaluation with simulated data and analysis of experimental data. *Int. J. Bioelectromagn, 4*(1), online journal.

Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology, 28*(2), 240-244.

Javitt, D. C. (2000). Intracortical mechanisms of mismatch negativity dysfunction in schizophrenia. *Audiol Neurootol, 5*(3-4), 207-215.

Kujala, T., Lepisto, T., Nieminen-von Wendt, T., Naatanen, P., & Naatanen, R. (2005). Neurophysiological evidence for cortical discrimination impairment of prosody in asperger syndrome. *Neurosci Lett, 383*(3), 260-265.

Lakshminarayanan, K., Ben Shalom, D., van Wassenhove, V., Orbelo, D., Houde, J., & Poeppel, D. (2003). The effect of spectral manipulations on the identification of affective and linguistic prosody. *Brain Lang, 84*(2), 250-263.

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. *Biol Psychiatry, 58*(1), 56-61.

Majewski, W., & Blasdell, R. (1969). Influence of fundamental frequency cues on the perception of some synthetic intonation contours. *J Acoust Soc Am, 45*(2), 450-457.

Naatanen, R. (1995). The mismatch negativity: A powerful tool for cognitive neuroscience. *Ear and Hearing, 16*, 6-18.

Pell, M. D. (1998). Recognition of prosody following unilateral brain lesion: Influence of functional and structural attributes of prosodic contours. *Neuropsychologia, 36*(8), 701-715.

Saarinen, J., Paavilainen, P., Schoger, E., Tervaniemi, M., & Naatanen, R. (1992). Representation of abstract attributes of auditory stimuli in the human brain. *Neuroreport, 3*(12), 1149-1151.

Van Lancker, D., & Sidtis, J. J. (1992). The identification of affective-prosodic stimuli by left- and right-hemisphere-damaged subjects: All errors are not created equal. *J Speech Hear Res, 35*(5), 963-970.

8.7 TABLES AND FIGURES

Table 1. Contour characteristics of stimuli used in this study (Pell, 1998)

Contour	**Hz**	**msec**	**Hz**	**msec**	**Hz**	**msec**	**Hz**	**msec**	**Total Duration**
Interrogative	227	104	211	107	183	117	**320**	100	428
Declarative	227	104	211	107	183	117	**162**	100	428

Table 2: Peak MMN amplitude by hemisphere (n=11)

	Peak Amplitude(μV) – mean (sd)	
Variable	Left hemisphere	Right hemisphere
Declarative contour (LH)	1.8 (.7)	1.6 (.9)
Interrogative contour (LH)	3.5 (1.8)	3.3 (1.1)

Figure 1. Scalp waveforms: The first two columns illustrate the waveforms elicited by to the interrogative (left) and declarative (right) contours when presented either as standard (blue trace) or deviant (red trace) stimuli within the MMN stimulation paradigm at indicated electrodes. The final column illustrates the deviant minus standard subtraction waveforms for the interrogative (black) and declarative (gray) stimulus types. The horizontal yellow bars lying on the x-axis in the first two columns reflect the period of significant difference between the standard and deviant presentations of the contours. The black and gray arrows under the subtraction waveforms reflect the maximum peak of the interrogative and declarative contour subtractions respectively.

Figure 2. Topography of MMN activity**.** Voltage maps illustrating distribution of activity for interrogative (top) and declarative (bottom) difference waveforms at latency of peak MMN stimulus onset.

Figure 3. Source analysis of MMN activity. Distributed inverse solution for MMN generators showing bilateral sources within superior temporal plane (primary and secondary auditory cortex). Solutions were determined by Local Auto-Regressive Average (LAURA) (Grave de Peralta & Gonzalez, 2002)modeling A graph of the global field power for each of the deviant conditions are presented to the left the source analysis figures.

Figure 1

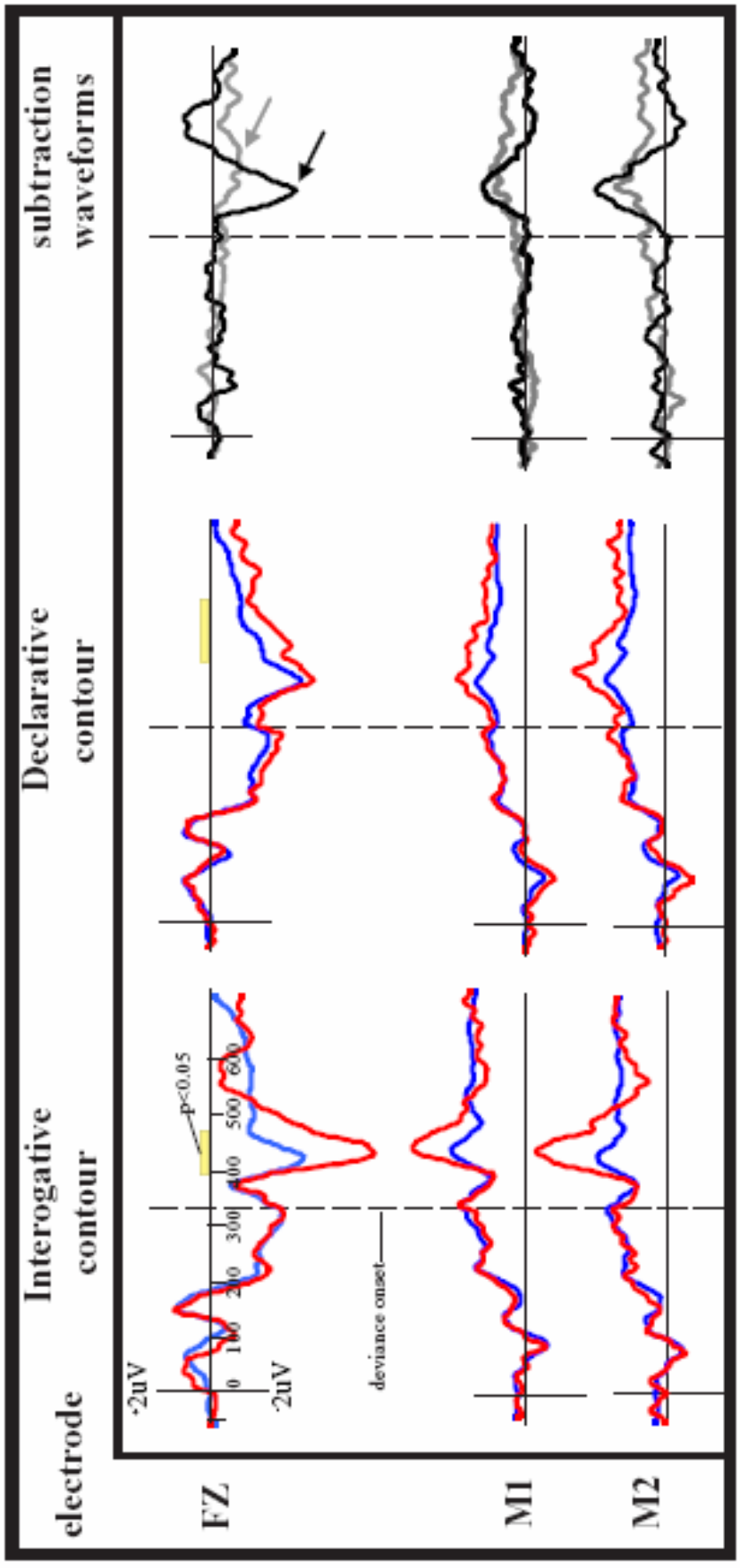

electrode
Interogative contour
Declarative contour
subtraction waveforms
FZ
M1
M2
-2uV
2uV
0
100
200
300
400
500
600
p<0.05
deviance onset

Figure 2

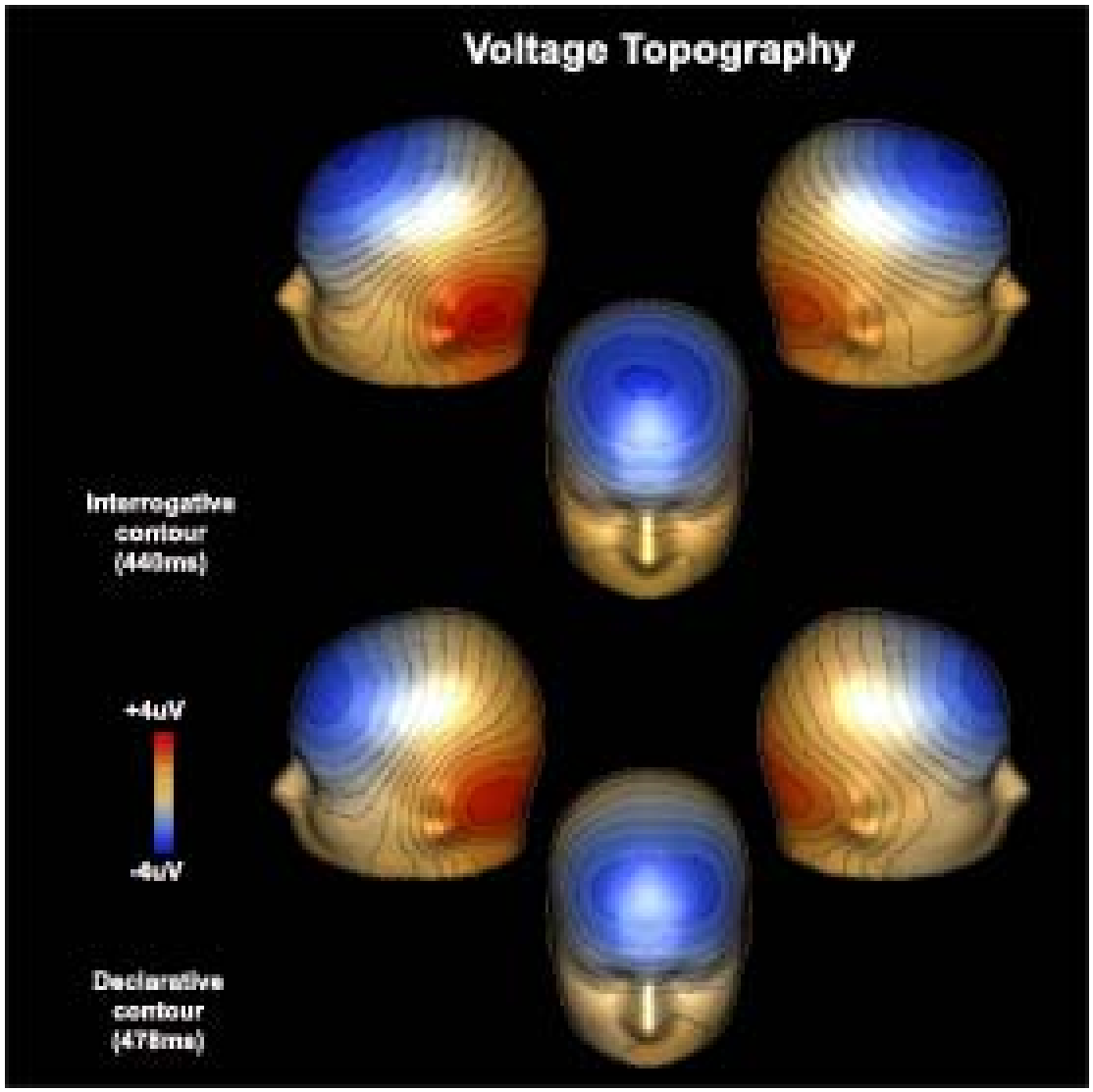
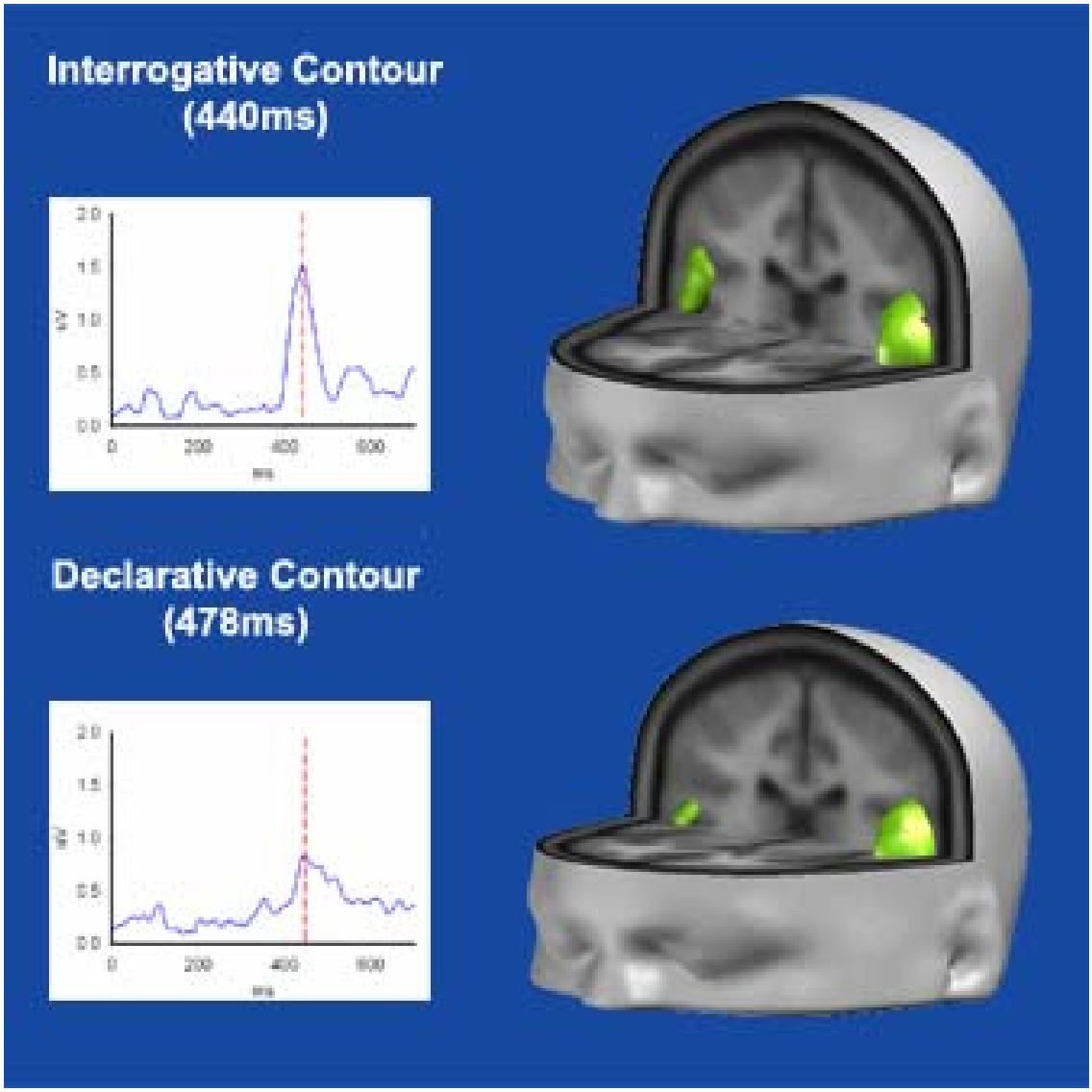

Voltage Topography
Interrogative contour (448ms)
+4uV
-4uV
Declarative contour (478ms)

Figure 3



8.0 BIBLIOGRAPHY

8.0 BIBLIOGRAPHY

Aaltonen, O., Niemi, P., Nyrke, T., & Tuhkanen, M. (1987). Event-related brain potentials and the perception of a phonetic continuum. *Biol Psychol, 24*(3), 197-207.

Abdi, Z., & Sharma, T. (2004). Social cognition and its neural correlates in schizophrenia and autism. *CNS Spectr, 9*(5), 335-343.

Abrams, R., & Taylor, M. A. (1978). A rating scale for emotional blunting. *Am.J Psychiatry, 135*(2), 226-229.

Adolphs, R., Schul, R., & Tranel, D. (1998). Intact recognition of facial emotion in parkinson's disease. *Neuropsychology, 12*(2), 253-258.

Alpert, M., & Anderson, L. T. (1977). Imagery mediation of vocal emphasis in flat affect. *Arch Gen Psychiatry, 34*, 208-212.

Alpert, M., Rosenberg, S. D., Pouget, E. R., & Shaw, R. J. (2000). Prosody and lexical accuracy in flat affect schizophrenia. *Psychiatry Res, 97*(2-3), 107-118.

Ammons, R. B., & Ammons, C. H. (1962). The quick test (qt): Provisional manual. *Psychological Reports, 11*, 111-162.

Andreasen, N. C. (1982). Negative symptoms in schizophrenia. *Archives of General Psychiatry, 39*, 784-788.

Andreasen, N. C. (1984). *The scale for the assessment of negative symptoms (sans)*. Iowa City: The University of Iowa.

Andreasen, N. C., Alpert, M., & Martz, M. J. (1981). Acoustic analysis: An objective measure of affective flattening. *Arch.Gen.Psychiatry, 38*(3), 281-285.

Andreasen, N. C., Arndt, S., Alliger, R., Miller, D., & Flaum, M. (1995). Symptoms of schizophrenia. Methods, meanings, and mechanisms. *Arch Gen Psychiatry, 52*(5), 341-351.

Apperly, I. A., Samson, D., Chiavarino, C., & Humphreys, G. W. (2004). Frontal and temporo-parietal lobe contributions to theory of mind: Neuropsychological evidence from a false-belief task with reduced language and executive demands. *J Cogn Neurosci, 16*(10), 1773-1784.

Ardekani, B. A., Nierenberg, J., Hoptman, M. J., Javitt, D. C., & Lim, K. O. (2003). Mri study of white matter diffusion anisotropy in schizophrenia. *Neuroreport, 14*(16), 2025-2029.

Arnott, S. R., Binns, M. A., Grady, C. L., & Alain, C. (2004). Assessing the auditory dual-pathway model in humans. *Neuroimage, 22*(1), 401-408.

Baltaxe, C. A., & Simmons, J. Q., III. (1995). Speech and language disorders in children and adolescents with schizophrenia. *Schizophr Bull, 21*(4), 677-692.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *J Pers Soc Psychol, 70*(3), 614-636.

Baudewig, J., Dechent, P., Merboldt, K. D., & Frahm, J. (2003). Thresholding in correlation analyses of magnetic resonance functional neuroimaging. *Magn Reson Imaging, 21*(10), 1121-1130.

Belin, P., & Zatorre, R. J. (2000). 'what', 'where' and 'how' in auditory cortex. *Nat Neurosci, 3*(10), 965-966.

Berenbaum, S. A., Abrams, R., Rosenberg, S., & Taylor, M. A. (1987). The nature of emotional blunting: A factor-analytic study. *Psychiatry Res, 20*(1), 57-67.

Bilder, R. M., Lipschutz-Broch, L., Reiter, G., Geisler, S., Mayerhoff, D., & Lieberman, J. A. (1991). Neuropsychological deficits in the early course of first episode schizophrenia. *Schizophr. Res., 5*(3), 198-199.

Bleuler, E. (German-1911: Trans-1950). *Dementia praecox or the group of schizophrenias* (H. Zinkin, Trans.). New York: International Universities press.

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

Blonder, L. X., Gur, R. E., & Gur, R. C. (1989). The effects of right and left hemiparkinsonism on prosody. *Brain Lang, 36*(2), 193-207.

Blonder, L. X., Pickering, J. E., Heath, R. L., Smith, C. D., & Butler, S. M. (1995). Prosodic characteristics of speech pre- and post-right hemisphere stroke. *Brain Lang, 51*(2), 318-335.

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc Natl Acad Sci U S A, 98*(20), 11818-11823.

Boeringa, J. A., & Castellani, S. (1982). Reliability and validity of emotional blunting as a criterion for diagnosis of schizophrenia. *Am.J Psychiatry, 139*(9), 1131-1135.

Bogerts, B., Ashtari, M., Degreef, G., Alvir, J. M., Bilder, R. M., & Lieberman, J. A. (1990). Reduced temporal limbic structure volumes on magnetic resonance images in first episode schizophrenia. *Psychiatry. Res., 35*(1), 1-13.

Borod, J. C. (2000). *The neuropsychology of emotion*. Oxford: Oxford University Press.

Borod, J. C., Alpert, M., Brozgold, A., Martin, C., Welkowitz, J., Diller, L., et al. (1989). A preliminary comparison of flat affect schizophrenics and brain-damaged patients on measures of affective processing. *J Commun Disord, 22*(2), 93-104.

Borod, J. C., Cicero, B. A., Obler, L. K., Welkowitz, J., Erhan, H. M., Santschi, C., et al. (1998). Right hemisphere emotional perception: Evidence across multiple channels. *Neuropsychology., 12*(3), 446-458.

Borod, J. C., Martin, C. C., Alpert, M., Brozgold, A., & Welkowitz, J. (1993). Perception of facial emotion in schizophrenic and right brain-damaged patients. *J Nerv Ment Dis, 181*(8), 494-502.

Borod, J. C., Welkowitz, J., Alpert, M., Brozgold, A. Z., Martin, C., Peselow, E., et al. (1990). Parameters of emotional processing in neuropsychiatric disorders: Conceptual issues and a battery of tests. *J Commun Disord, 23*(4-5), 247-271.

Braff, D. L. (1993). Information processing and attention dysfunctions in schizophrenia. *Schizophrenia Bulletin, 19*, 233-259.

Braus, D. F., Weber-Fahr, W., Tost, H., Ruf, M., & Henn, F. A. (2002). Sensory information processing in neuroleptic-naive first-episode schizophrenic patients: A functional magnetic resonance imaging study. *Arch Gen Psychiatry, 59*(8), 696-701.

Breitenstein, C., Van Lancker, D., Daum, I., & Waters, C. H. (2001). Impaired perception of vocal emotions in parkinson's disease: Influence of speech time processing and executive functioning. *Brain Cogn, 45*(2), 277-314.

Brekke, J., Kay, D. D., Lee, K. S., & Green, M. F. (2005a). Biosocial pathways to functional outcome in schizophrenia. *Schizophr Res, 80*(2-3), 213-225.

Brekke, J. S., Long, J. D., & Kay, D. D. (2002). The structure and invariance of a model of social functioning in schizophrenia. *J Nerv Ment Dis, 190*(2), 63-72.

Brekke, J. S., Nakagami, E., Kee, K. S., & Green, M. F. (2005b). Cross-ethnic differences in perception of emotion in schizophrenia. *Schizophr Res, 77*(2-3), 289-298.

Brune, M. (2005a). Emotion recognition, 'theory of mind,' and social behavior in schizophrenia. *Psychiatry Res, 133*(2-3), 135-147.

Brune, M. (2005b). "theory of mind" in schizophrenia: A review of the literature. *Schizophr Bull, 31*(1), 21-42.

Buchanan, R. W., Breier, A., Kirkpatrick, B., Elkashef, A., Munson, R. C., Gellad, F., et al. (1993). Structural abnormalities in deficit and nondeficit schizophrenia [see comments]. *Am J Psychiatry, 150*(1), 59-65.

Buchsbaum, B. R., Greer, S., Chang, W. L., & Berman, K. F. (2005). Meta-analysis of neuroimaging studies of the wisconsin card-sorting task and component processes. *Hum Brain Mapp, 25*(1), 35-45.

Butler, P., Schechter, I., Saperstein, A., Revheim, N., Silipo, G., Zemon, V., et al. (2005). Deficits in contrast gain underly early visual processing dysfunction in schizophrenia. *Arch Gen Psychiatry, in press*.

Butler, P. D., Schechter, I., Zemon, V., Schwartz, S. G., Greenstein, V. C., Gordon, J., et al. (2001). Dysfunction of early-stage visual processing in schizophrenia. *Am J Psychiatry, 158*(7), 1126-1133.

Caekebeke, J. F., Jennekens-Schinkel, A., van der Linden, M. E., Buruma, O. J., & Roos, R. A. (1991). The interpretation of dysprosody in patients with parkinson's disease. *J Neurol Neurosurg Psychiatry, 54*(2), 145-148.

Carlsson, A. (1977). Does dopamine play a role in schizophrenia? *Psychol Med, 7*(4), 583-597.

Channon, S., Pellijeff, A., & Rule, A. (2005). Social cognition after head injury: Sarcasm and theory of mind. *Brain Lang, 93*(2), 123-134.

Chapman, J. (1966). The early symptoms of schizophrenia. *Br J Psychiatry, 112*(484), 225-251.

Cherry, C. (1978). *On human communication: A review, a survey and a criticism* (3rd ed ed.). Cambridge, Mass.: MIT Press.

Cienfuegos, A., March, L., Shelley, A. M., & Javitt, D. C. (1999). Impaired categorical perception of synthetic speech sounds in schizophrenia. *Biol Psychiatry, 45*(1), 82-88.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences, 2nd edition*. Hillsdale, NJ: Lawrence Erlbaum Assoc.

Corcoran, C., Gallitano, A., Leitman, D., & Malaspina, D. (2001). The neurobiology of the stress cascade and its potential relevance for schizophrenia. *J Psychiatr Pract, 7*(1), 3-14.

Corcoran, C., Mujica-Parodi, L., Yale, S., Leitman, D., & Malaspina, D. (2002). Could stress cause psychosis in individuals vulnerable to schizophrenia? *CNS Spectr, 7*(1), 33-38, 41-32.

Cowan, N. (1984). On short and long auditory stores. *Psychol.Bull., 96*(2), 341-370.

Cowan, N., Winkler, I., Teder, W., & Naatanen, R. (1993). Memory prerequisites of mismatch negativity in the auditory event- related potential (erp). *J Exp Psychol.Learn.Mem.Cogn, 19*(4), 909-921.

Critchley, E. M. (1981). Speech disorders of parkinsonism: A review. *J Neurol Neurosurg Psychiatry, 44*(9), 751-758.

Darkins, A. W., Fromkin, V. A., & Benson, D. F. (1988). A characterization of the prosodic loss in parkinson's disease. *Brain Lang, 34*(2), 315-327.

Darwin, C. (1865). *The expression of emotion in man and animals* (1948 ed.). London, England: Watts and Co.

David, A. S., & Cutting, J. C. (1990). Affect, affective disorder and schizophrenia. A neuropsychological investigation of right hemisphere function. *Br J Psychiatry, 156*, 491-495.

Davis, B. L., MacNeilage, P. F., Matyear, C. L., & Powell, J. K. (2000). Prosodic correlates of stress in babbling: An acoustical study. *Child Dev., 71*(5), 1258-1270.

Davis, K. L., Stewart, D. G., Friedman, J. I., Buchsbaum, M., Harvey, P. D., Hof, P. R., et al. (2003). White matter changes in schizophrenia: Evidence for myelin-related dysfunction. *Arch Gen Psychiatry, 60*(5), 443-456.

Davitz, J. (1964). *The communication of emotional meaning*. Westport CN: Greenwood Press.

de Leon, J., Peralta, V., & Cuesta, M. J. (1993). Negative symptoms and emotional blunting in schizophrenic patients. *J Clin.Psychiatry, 54*(3), 103-108.

Drayna, D., Manichaikul, A., de Lange, M., Snieder, H., & Spector, T. (2001). Genetic correlates of musical pitch recognition in humans. *Science, 291*(5510), 1969-1972.

Edwards, J., Jackson, H. J., & Pattison, P. E. (2002). Emotion recognition via facial expression and affective prosody in schizophrenia: A methodological review. *Clin Psychol Rev, 22*(6), 789-832.

Edwards, J., & McGorry, P. D. (2002). *Implementing early intervention in psychosis: A guide to establishing early psychosis services.* London: Martin Dunitz.

Edwards, J., Pattison, P. E., Jackson, H. J., & Wales, R. J. (2001). Facial affect and affective prosody recognition in first-episode schizophrenia. *Schizophr.Res, 48*(2-3), 235-253.

Ekman, P. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Ethofer, T., Anders, S., Erb, M., Herbert, C., Wiethoff, S., Kissler, J., et al. (2006). Cerebral pathways in processing of affective prosody: A dynamic causal modeling study. *Neuroimage, 30*(2), 580-587.

Exner, C., Boucsein, K., Degner, D., Irle, E., & Weniger, G. (2004). Impaired emotional learning and reduced amygdala size in schizophrenia: A 3-month follow-up. *Schizophr Res, 71*(2-3), 493-503.

Fahim, C., Stip, E., Mancini-Marie, A., Mensour, B., Boulay, L. J., Leroux, J. M., et al. (2005). Brain activity during emotionally negative pictures in schizophrenia with and without flat affect: An fmri study. *Psychiatry Res, 140*(1), 1-15.

Fenton, W. S., & McGlashan, T. H. (1992). Testing systems for assessment of negative symptoms in schizophrenia. *Arch.Gen.Psychiatry, 49*(3), 179-184.

Fernald, A. (1985). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20*(1), Jan-113.

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development, 60*(6), 1497-1510.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. (1997). *Structural clinical interview for dsm-iv axis i disorders (scid-iv)*. New York: Biometrics Research Department, New York State Psychiatric Institute.

Fleming, K., Goldberg, T. E., Binks, S., Randolph, C., Gold, J. M., & Weinberger, D. R. (1997). Visuospatial working memory in patients with schizophrenia. *Biol.Psychiatry, 41*(1), 43-49.

Foxe, J. J., Doniger, G. M., & Javitt, D. C. (2001). Early visual processing deficits in schizophrenia: Impaired p1 generation revealed by high-density electrical mapping. *Neuroreport, 12*, 3815-3820.

Fricchione, G., & Howanitz, E. (1985). Aprosodia and alexithymia--a case report. *Psychother Psychosom, 43*(3), 156-160.

Fricchione, G., Sedler, M. J., & Shukla, S. (1986). Aprosodia in eight schizophrenic patients. *Am J Psychiatry, 143*(11), 1457-1459.

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

Frith, C. D. (1992). *The cognitive neuropsychology of schizophrenia*. Hove UK: Laurence Earlbaum Associates.

Frith, C. D., & Frith, U. (1999). Interacting minds--a biological basis. *Science, 286*(5445), 1692-1695.

George, M. S., Parekh, P. I., Rosinsky, N., Ketter, T. A., Kimbrell, T. A., Heilman, K. M., et al. (1996). Understanding emotional prosody activates right hemisphere regions. *Arch.Neurol., 53*(7), 665-670.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proc Natl Acad Sci U S A, 100*(13), 8030-8035.

Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., et al. (2005). The voices of wrath: Brain responses to angry prosody in meaningless speech. *Nat Neurosci, 8*(2), 145-146.

Grave de Peralta, R., & Gonzalez, A. (2002). Comparison of algorithms for the localization of focal sources: Evaluation with simulated data and analysis of experimental data. *Int. J. Bioelectromagn, 4*(1), online journal.

Green, M. (1996). What are the functional consequences of neurocognitive deficits in schizophrenia. *Am. J. Psychiatry, 153*, 321-330.

Green, M., & Walker, E. (1986). Symptom correlates of vulnerability to backward masking in schizophrenia. *Am J Psychiatry, 143*(2), 181-186.

Green, M. F., Kern, R. S., Braff, D. L., & Mintz, J. (2000). Neurocognitive deficits and functional outcome in schizophrenia: Are we measuring the "right stuff"? *Schizophr.Bull., 26*(1), 119-136.

Green, M. F., & Nuechterlein, K. H. (1999). Backward masking performance as an indicator of vulnerability to schizophrenia. *Acta Psychiatr Scand Suppl, 395*, 34-40.

Green, M. F., Nuechterlein, K. H., & Mintz, J. (1994). Backward masking in schizophrenia and mania. I. Specifying a mechanism. *Arch.Gen.Psychiatry, 51*(12), 939-944.

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychol Bull, 75*(6), 424-429.

Gur, R. E., McGrath, C., Chan, R. M., Schroeder, L., Turner, T., Turetsky, B. I., et al. (2002). An fmri study of facial emotion processing in patients with schizophrenia. *Am J Psychiatry, 159*(12), 1992-1999.

Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology, 28*(2), 240-244.

Hardy-Bayle, M. C., Passerieux, C., Claudel, B., Olivier, V., & Chevalier, J. F. (1994). [communication disorders in schizophrenic patients. Cognitive explanation and clinical reconsideration]. *Encephale, 20*(4), 393-400.

Harvey, P. D., Docherty, N. M., Serper, M. R., & Rasmussen, M. (1990). Cognitive deficits and thought disorder: Ii. An 8-month followup study. *Schizophr Bull, 16*(1), 147-156.

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

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtis, G. (1993). The wisconsin card sorting task.

Heilman, K. M., Bowers, D., Speedie, L., & Coslett, H. B. (1984). Comprehension of affective and nonaffective prosody. *Neurology, 34*(7), 917-921.

Heilman, K. M., & Gilmore, R. L. (1998). Cortical influences in emotion. *J Clin Neurophysiol, 15*(5), 409-423.

Heimberg, C., Gur, R. E., Erwin, R. J., Shtasel, D. L., & Gur, R. C. (1992). Facial emotion discrimination: Iii. Behavioral findings in schizophrenia. *Psychiatry Res, 42*(3), 253-265.

Heinrichs, R. W., & Zakzanis, K. K. (1998). Neurocognitive deficit in schizophrenia: A quantitative review of the evidence. *Neuropsychology, 12*(3), 426-445.

Hirayasu, Y., McCarley, R. W., Salisbury, D. F., Tanaka, S., Kwon, J. S., Frumin, M., et al. (2000). Planum temporale and heschl gyrus volume reduction in schizophrenia: A magnetic resonance imaging study of first-episode patients. *Arch Gen Psychiatry, 57*(7), 692-699.

Holcomb, H. H., Ritzl, E. K., Medoff, D. R., Nevitt, J., Gordon, B., & Tamminga, C. A. (1995). Tone discrimination performance in schizophrenic patients and normal volunteers: Impact of stimulus presentation levels and frequency differences. *Psychiatry Res., 57*(1), 75-82.

Holt, D. J., Weiss, A. P., Rauch, S. L., Wright, C. I., Zalesak, M., Goff, D. C., et al. (2005). Sustained activation of the hippocampus in response to fearful faces in schizophrenia. *Biol Psychiatry, 57*(9), 1011-1019.

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

Hoptman, M. J., Ardekani, B. A., Butler, P. D., Nierenberg, J., Javitt, D. C., & Lim, K. O. (2004). Dti and impulsivity in schizophrenia: A first voxelwise correlational analysis. *Neuroreport, 15*(16), 2467-2470.

Hoptman, M. J., Volavka, J., Johnson, G., Weiss, E., Bilder, R. M., & Lim, K. O. (2002). Frontal white matter microstructure, aggression, and impulsivity in men with schizophrenia: A preliminary study. *Biol Psychiatry, 52*(1), 9-14.

Hyman, S. E., & Arana, G. W. (1987). *Handbook of psychiatric drug therapy* (Vol. 7). New York: Little, Brown.

Izard, C. E. (1971). *The face of emotion*. New York: Appelton-Century-Crofts.

James, W. (1890). *The principles of psychology* (1950 ed. Vol. 2). New York: Dover Publications.

Javitt, D. C. (1987). Negative schizophrenic symptomatology and the pcp (phencyclidine) model of schizophrenia. *Hillside J Clin Psychiatry, 9*(1), 12-35.

Javitt, D. C. (2000). Intracortical mechanisms of mismatch negativity dysfunction in schizophrenia. *Audiol Neurootol, 5*(3-4), 207-215.

Javitt, D. C., & Coyle, J. T. (2004). Decoding schizophrenia. *Sci Am, 290*(1), 48-55.

Javitt, D. C., Doneshka, P., Steen, M. E., Nussenzweig, I., Ritter, W., & Vaughan, H. G., Jr. (1992). Abnormalities of early cortical processing in schizophrenia revealed by auditory event-related potentials. *Electroencephalography and clinical neurophysiology, 82*, 102P.

Javitt, D. C., Doneshka, P., Zylberman, I., Ritter, W., & Vaughan, H. G., Jr. (1993). Impairment of early cortical processing in schizophrenia: An event- related potential confirmation study. *Biol.Psychiatry, 33*(7), 513-519.

Javitt, D. C., Liederman, E., Cienfuegos, A., & Shelley, A. M. (1999). Panmodal processing imprecision as a basis for dysfunction of transient memory storage systems in schizophrenia. *Schiz. Bull., 25*, 763-775.

Javitt, D. C., Schroeder, C. E., Steinschneider, M., Arezzo, J. C., Ritter, W., & Vaughan, H. G., Jr. (1995a). Cognitive event-related potentials in human and non-human primates: Implications for the pcp/nmda model of schizophrenia. *Electroencephalogr.Clin.Neurophysiol.Suppl, 44*, 161-175.

Javitt, D. C., Shelley, A., & Ritter, W. (2000). Associated deficits in mismatch negativity generation and tone matching in schizophrenia. *Clin.Neurophysiol., 111*(10), 1733-1737.

Javitt, D. C., Shelley, A. M., Grochowski, S., & Ritter, W. (1995b). Mismatch negativity (mmn) as an index of impaired auditory sensory memory in schizophrenia. *Schizophrenia Research, 15*, 179.

Javitt, D. C., Strous, R. D., Grochowski, S., Ritter, W., & Cowan, N. (1997). Impaired precision, but normal retention, of auditory sensory ("echoic") memory information in schizophrenia. *J Abnorm Psychol, 106*(2), 315-324.

Johnston, P. J., Stojanov, W., Devir, H., & Schall, U. (2005). Functional mri of facial emotion recognition deficits in schizophrenia and their electrophysiological correlates. *Eur J Neurosci, 22*(5), 1221-1232.

Jonsson, C. O., & Sjostedt, A. (1973). Auditory perception in schizophrenia: A second study of the intonation test. *Acta Psychiatr.Scand., 49*(5), 588-600.

Jorgensen, J. (1996). The functions of sarcastic irony in speech. *Journal of Pragmatics, 26*, 613-634.

Juslin, P. N., & Laukka, P. (2001). Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion, 1*(4), 381-412.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychol Bull, 129*(5), 770-814.

Kaplan, J. A., Brownell, H. H., Jacobs, J. R., & Gardner, H. (1990). The effects of right hemisphere damage on the pragmatic interpretation of conversational remarks. *Brain Lang, 38*(2), 315-333.

Kawasaki, H., Kaufman, O., Damasio, H., Damasio, A. R., Granner, M., Bakken, H., et al. (2001). Single-neuron responses to emotional visual stimuli recorded in human ventral prefrontal cortex. *Nat Neurosci, 4*(1), 15-16.

Kay, S. R., & Murrill, L. M. (1990). Predicting outcome of schizophrenia: Significance of symptom profiles and outcome dimensions. *Compr. Psychiatry, 31*, 91-102.

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

Kee, K. S., Kern, R. S., & Green, M. F. (1998). Perception of emotion and neurocognitive functioning in schizophrenia: What's the link? *Psychiatry Res, 81*(1), 57-65.

Kelemen, O., Erdelyi, R., Pataki, I., Benedek, G., Janka, Z., & Keri, S. (2005). Theory of mind and motion perception in schizophrenia. *Neuropsychology, 19*(4), 494-500.

Kelly, B. D., O'Callaghan, E., Lane, A., & Larkin, C. (2003). Schizophrenia: Solving the puzzle. *Ir J Med Sci, 172*(1), 37-40.

Kerr, S. L., & Neale, J. M. (1993). Emotion perception in schizophrenia: Specific deficit or further evidence of generalized poor performance? *J Abnorm.Psychol., 102*(2), 312-318.

Kim, D., Zemon, V., Saperstein, A., Butler, P. D., & Javitt, D. C. (2005). Dysfunction of early-stage visual processing in schizophrenia: Harmonic analysis. *Schizophr Res, 76*(1), 55-65.

Knowner, F. (1941). Analysis of some experimental variations of simulated vocal expressions and actions. *Journal of Social Psychology, 14*, 369-372.

Koelsch, S., Maess, B., Gunter, T. C., & Friederici, A. D. (2001). Neapolitan chords activate the area of broca. A magnetoencephalographic study. *Ann.N.Y.Acad.Sci, 930*, 420-421.

Kosaka, H., Omori, M., Murata, T., Iidaka, T., Yamada, H., Okada, T., et al. (2002). Differential amygdala response during facial recognition in patients with schizophrenia: An fmri study. *Schizophr Res, 57*(1), 87-95.

Kring, A. M., Barrett, L. F., & Gard, D. E. (2003). On the broad applicability of the affective circumplex: Representations of affective knowledge among schizophrenia patients. *Psychol Sci, 14*(3), 207-214.

Kring, A. M., Kerr, S. L., Smith, D. A., & Neale, J. M. (1993). Flat affect in schizophrenia does not reflect diminished subjective experience of emotion. *J Abnorm Psychol, 102*(4), 507-517.

Kring, A. M., & Neale, J. M. (1996). Do schizophrenic patients show a disjunctive relationship among expressive, experiential, and psychophysiological components of emotion? *J Abnorm Psychol, 105*(2), 249-257.

Kubicki, M., Westin, C. F., Nestor, P. G., Wible, C. G., Frumin, M., Maier, S. E., et al. (2003). Cingulate fasciculus integrity disruption in schizophrenia: A magnetic resonance diffusion tensor imaging study. *Biol Psychiatry, 54*(11), 1171-1180.

Kucharska-Pietura, K., David, A. S., Masiak, M., & Phillips, M. L. (2005). Perception of facial and vocal affect by people with schizophrenia in early and late stages of illness. *Br J Psychiatry, 187*, 523-528.

Kucharska-Pietura, K., Russell, T., & Masiak, M. (2003). Perception of negative affect in schizophrenia--functional and structural changes in the amygdala. Review. *Ann Univ Mariae Curie Sklodowska [Med], 58*(2), 453-458.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nat Rev Neurosci, 5*(11), 831-843.

Kuhl, P. K., Coffey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: Behavioral and electrophysiological measures. *Dev Sci, 8*(1), F1-F12.

Kujala, T., Lepisto, T., Nieminen-von Wendt, T., Naatanen, P., & Naatanen, R. (2005). Neurophysiological evidence for cortical discrimination impairment of prosody in asperger syndrome. *Neurosci Lett, 383*(3), 260-265.

Kukla, F., & Gold, R. (1991). [negative schizophrenic symptoms and their detection]. *Fortschr.Neurol Psychiatr., 59*(2), 60-66.

Kulynych, J. J., Vladar, K., Jones, D. W., & Weinberger, D. R. (1996). Superior temporal gyrus volume in schizophrenia: A study using mri morphometry assisted by surface rendering. *American Journal of Psychiatry, 153*, 50-56.

Ladd, D., Silverman, K., Tolkmitt, F., Bergmann, G., & Scherer, K. (1985). Evidence for the independent function of intonation,contour type, voice quality, and f0 range in signalling speaker affect. *Journal of the Acoustical Society of America, 78*, 435-444.

Lakshminarayanan, K., Ben Shalom, D., van Wassenhove, V., Orbelo, D., Houde, J., & Poeppel, D. (2003). The effect of spectral manipulations on the identification of affective and linguistic prosody. *Brain Lang, 84*(2), 250-263.

Lang, P. J., & Buss, A. H. (1965). Psychological deficit in schizophrenia. Ii. Interference and activation. *J Abnorm Psychol, 70*, 77-106.

Langdon, R., Coltheart, M., Ward, P. B., & Catts, S. V. (2002). Disturbed communication in schizophrenia: The role of poor pragmatics and poor mind-reading. *Psychol Med, 32*(7), 1273-1284.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annu Rev Neurosci, 23*, 155-184.

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. *Biol Psychiatry, 58*(1), 56-61.

Leitman, D. I., Ziwich, R., Pasternak, R., & Javitt, D. C. (2006). Theory of mind (tom) and counterfactuality deficits in schizophrenia: Misperception or misinterpretation? *Psychol Med*, 1-9.

Liegeois-Chauvel, C., de Graaf, J. B., Laguitton, V., & Chauvel, P. (1999). Specialization of left auditory cortex for speech perception in man depends on temporal coding. *Cereb Cortex, 9*(5), 484-496.

Liegeois-Chauvel, C., Giraud, K., Badier, J. M., Marquis, P., & Chauvel, P. (2001). Intracerebral evoked potentials in pitch perception reveal a functional asymmetry of the human auditory cortex. *Ann NY Acad Sci, 930*, 117-132.

Lim, K., Ardekani, B., Nierenberg, J., Butler, P., Javitt, D., & Hoptman, M. (2005). Neurocognitive correlates of white matter integrity in schizophrenia. *American Journal of Psychiatry*, in press.

Lim, K. O., Hedehus, M., Moseley, M., de Crespigny, A., Sullivan, E. V., & Pfefferbaum, A. (1999). Compromised white matter tract integrity in schizophrenia inferred from diffusion tensor imaging. *Arch Gen Psychiatry, 56*(4), 367-374.

Lim, K. O., & Helpern, J. A. (2002). Neuropsychiatric applications of dti - a review. *NMR Biomed, 15*(7-8), 587-593.

Loeb, P. A. (1996). *Ils: Independent living scales manual*. San Antonio: The Psychological Corporation, Harcourt Race Jovanovich, Inc.

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in broca's area: An meg study. *Nat.Neurosci., 4*(5), 540-545.

Majewski, W., & Blasdell, R. (1969). Influence of fundamental frequency cues on the perception of some synthetic intonation contours. *J Acoust Soc Am, 45*(2), 450-457.

Malaspina, D., Dalack, G., Leitman, D., Corcoran, C., Amador, X. F., Yale, S., et al. (2002). Low heart rate variability is not caused by typical neuroleptics in schizophrenia patients. *CNS Spectr, 7*(1), 53-57.

Malaspina, D., Harlap, S., Fennig, S., Heiman, D., Nahon, D., Feldman, D., et al. (2001). Advancing paternal age and the risk of schizophrenia. *Arch Gen Psychiatry, 58*(4), 361-367.

March, L., Cienfuegos, A., Goldbloom, L., Ritter, W., Cowan, N., & Javitt, D. C. (1999). Normal time course of auditory recognition in schizophrenia, despite impaired

precision of the auditory sensory ("echoic") memory code. *J. Abnorm. Psychol., 108*, 69-75.

Marcus, J., Hans, S. L., Auerbach, J. G., & Auerbach, A. G. (1993). Children at risk for schizophrenia: The jerusalem infant development study. Ii. Neurobehavioral deficits at school age. *Arch Gen Psychiatry, 50*(10), 797-809.

Martin, C. C., Borod, J. C., Alpert, M., Brozgold, A., & Welkowitz, J. (1990). Spontaneous expression of facial emotion in schizophrenic and right- brain-damaged patients. *J.Commun.Disord., 23*(4-5), 287-301.

Mazza, M., De Risio, A., Surian, L., Roncone, R., & Casacchia, M. (2001). Selective impairments of theory of mind in people with schizophrenia. *Schizophr Res, 47*(2-3), 299-308.

McDonald, S. (1999). Exploring the process of inference generation in sarcasm: A review of normal and clinical studies. *Brain Lang, 68*(3), 486-506.

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

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

Murphy, D., & Cutting, J. (1990). Prosodic comprehension and expression in schizophrenia. *J Neurol Neurosurg.Psychiatry, 53*(9), 727-730.

Naatanen, R. (1995). The mismatch negativity: A powerful tool for cognitive neuroscience. *Ear and Hearing, 16*, 6-18.

Naatanen, R., & Michie, P. T. (1979). Early selective-attention effects on the evoked potential: A critical review and reinterpretation. *Biol Psychol, 8*(2), 81-136.

Novak, G., Ritter, W., & Vaughan, H. G., Jr. (1992a). The chronometry of attention-modulated processing and automatic mismatch detection. *Psychophysiology, 29*(4), 412-430.

Novak, G., Ritter, W., & Vaughan, H. G., Jr. (1992b). Mismatch detection and the latency of temporal judgements. *Psychophysiology, 29*(4), 398-411.

O'Donnell, B. F., Hokama, H., McCarley, R. W., Smith, R. S., Salisbury, D. F., Mondrow, E., et al. (1994). Auditory erps to non-target stimuli in schizophrenia: Relationship to probability, task-demands, and target erps. *Int.J.Psychophysiol., 17*(3), 219-231.

Oades, R. D., Dittmann-Balcar, A., Zerbin, D., & Grzella, I. (1997). Impaired attention-dependent augmentation of mmn in nonparanoid vs paranoid schizophrenic patients: A comparison with obsessive-compulsive disorder and healthy subjects. *Biol.Psychiatry, 41*(12), 1196-1210.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Orbelo, D. M., Grim, M. A., Talbott, R. E., & Ross, E. D. (2005). Impaired comprehension of affective prosody in elderly subjects is not predicted by age-related hearing loss or age-related cognitive decline. *J Geriatr Psychiatry Neurol, 18*(1), 25-32.

Ornitz, E. M. (1969). Disorders of perception common to early infantile autism and schizophrenia. *Compr Psychiatry, 10*(4), 259-274.

Otsu, N. (1979). A threshold selection model from gray-level histograms. *IEEE Trans (SMC), 9*, 63-66.

Overall, J. E., & Gorham, D. E. (1961). The brief psychiatric rating scale. *Psychol Reports, 10*, 799-812.

Ozonoff, S., & Miller, J. N. (1996). An exploration of right-hemisphere contributions to the pragmatic impairments of autism. *Brain Lang, 52*(3), 411-434.

Parker, G. J., Luzzi, S., Alexander, D. C., Wheeler-Kingshott, C. A., Ciccarelli, O., & Lambon Ralph, M. A. (2005). Lateralization of ventral and dorsal auditory-language pathways in the human brain. *Neuroimage, 24*(3), 656-666.

Patel, A. D., & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition, 87*(1), B35-45.

Patel, A. D., Peretz, I., Tramo, M., & Labreque, R. (1998). Processing prosodic and musical patterns: A neuropsychological investigation. *Brain Lang, 61*(1), 123-144.

Pell, M. D. (1998). Recognition of prosody following unilateral brain lesion: Influence of functional and structural attributes of prosodic contours. *Neuropsychologia, 36*(8), 701-715.

Pell, M. D. (1999). Fundamental frequency encoding of linguistic and emotional prosody by right hemisphere-damaged speakers. *Brain Lang, 69*(2), 161-192.

Pell, M. D., & Baum, S. R. (1997). The ability to perceive and comprehend intonation in linguistic and affective contexts by brain-damaged adults. *Brain Lang, 57*(1), 80-99.

Pell, M. D., Cheang, H. S., & Leonard, C. L. (2006). The impact of parkinson's disease on vocal-prosodic communication from the perspective of listeners. *Brain Lang, 97*(2), 123-134.

Pell, M. D., & Leonard, C. L. (2003). Processing emotional tone from speech in parkinson's disease: A role for the basal ganglia. *Cogn Affect Behav Neurosci, 3*(4), 275-288.

Phillips, M. L. (2003). Understanding the neurobiology of emotion perception: Implications for psychiatry. *Br J Psychiatry, 182*, 190-192.

Phillips, M. L., Drevets, W. C., Rauch, S. L., & Lane, R. (2003). Neurobiology of emotion perception ii: Implications for major psychiatric disorders. *Biol Psychiatry, 54*(5), 515-528.

Phillips, M. L., Williams, L., Senior, C., Bullmore, E. T., Brammer, M. J., Andrew, C., et al. (1999). A differential neural response to threatening and non-threatening negative facial expressions in paranoid and non-paranoid schizophrenics. *Psychiatry Res, 92*(1), 11-31.

Pickett, C. L., Gardner, W. L., & Knowles, M. (2004). Getting a cue: The need to belong and enhanced sensitivity to social cues. *Pers Soc Psychol Bull, 30*(9), 1095-1107.

Pickup, G. J., & Frith, C. D. (2001). Theory of mind impairments in schizophrenia: Symptomatology, severity and specificity. *Psychol Med, 31*(2), 207-220.

Pilowsky, T., Yirmiya, N., Arbelle, S., & Mozes, T. (2000). Theory of mind abilities of children with schizophrenia, children with autism, and normally developing children. *Schizophr Res, 42*(2), 145-155.

Pomara, N., Crandall, D. T., Choi, S. J., Johnson, G., & Lim, K. O. (2001). White matter abnormalities in hiv-1 infection: A diffusion tensor imaging study. *Psychiatry Res, 106*(1), 15-24.

Rabinowicz, E. F., Opler, L. A., Owen, D. R., & Knight, R. A. (1996). Dot enumeration perceptual organization task (depot): Evidence for a short-term visual memory deficit in schizophrenia. *J Abnorm.Psychol., 105*(3), 336-348.

Rabinowicz, E. F., Silipo, G., Goldman, R., & Javitt, D. C. (2000). Auditory sensory dysfunction in schizophrenia: Imprecision or distractibility? *Arch Gen Psychiatry, 57*(12), 1149-1155.

Reese TG, H. O., Weisskopf RM and Wedeen VJ. (2003). Reduction of eddy-current-induced distortion in diffusion mri using a twice-refocused spin echo. *Magn Reson Med, 49*, 177-182.

Revheim, N., & Medalia, A. (2004). Verbal memory, problem-solving skills and community status in schizophrenia. *Schizophr Res, 68*(2-3), 149-158.

Ritter, W., Deacon, D., Gomes, H., Javitt, D. C., & Vaughan, H. G., Jr. (1995). The mismatch negativity of event-related potentials as a probe of transient auditory memory: A review. *Ear Hear., 16*(1), 52-67.

Roncone, R., Falloon, I. R., Mazza, M., De Risio, A., Pollice, R., Necozione, S., et al. (2002). Is theory of mind in schizophrenia more strongly associated with clinical and social functioning than with neurocognitive deficits? *Psychopathology, 35*(5), 280-288.

Ross, E. D., Orbelo, D. M., Cartwright, J., Hansel, S., Burgard, M., Testa, J. A., et al. (2001a). Affective-prosodic deficits in schizophrenia: Comparison to patients

with brain damage and relation to schizophrenic symptoms [corrected]. *J Neurol Neurosurg Psychiatry, 70*(5), 597-604.

Ross, E. D., Orbelo, D. M., Cartwright, J., Hansel, S., Burgard, M., Testa, J. A., et al. (2001b). Affective-prosodic deficits in schizophrenia: Profiles of patients with brain damage and comparison with relation to schizophrenic symptoms. *J Neurol Neurosurg.Psychiatry, 70*(5), 597-604.

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

Saarinen, J., Paavilainen, P., Schoger, E., Tervaniemi, M., & Naatanen, R. (1992). Representation of abstract attributes of auditory stimuli in the human brain. *Neuroreport, 3*(12), 1149-1151.

Saykin, A. J., Gur, R. C., Gur, R. E., Mozley, P. D., Mozley, L. H., Resnick, S. M., et al. (1991). Neuropsychological function in schizophrenia. Selective impairment in memory and learning. *Arch.Gen.Psychiatry, 48*(7), 618-624.

Schenkel, L. S., Spaulding, W. D., & Silverstein, S. M. (2005). Poor premorbid social functioning and theory of mind deficit in schizophrenia: Evidence of reduced context processing? *J Psychiatr Res, 39*(5), 499-508.

Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychol Bull, 99*(2), 143-165.

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends Cogn Sci, 10*(1), 24-30.

Schirmer, A., & Kotz, S. A. (in press). Beyond the right hemisphere: Brain mechanism mediating emotional processing. *Trends in Cogntive Neuroscience*.

Shaheen, O., & Ibrahim, S. (1979). The phenomenon of emotional blunting: A wholistic quantitative qualitative approach. *Egypt.J Psychiatry, 2*(1), 27-40.

Shamay-Tsoory, S. G., Tomer, R., & Aharon-Peretz, J. (2005). The neuroanatomical basis of understanding sarcasm and its relationship to social cognition. *Neuropsychology, 19*(3), 288-300.

Shenton, M. E., Dickey, C. C., Frumin, M., & McCarley, R. W. (2001). A review of mri findings in schizophrenia. *Schizophr Res, 49*(1-2), 1-52.

Skinner, E. (1935). A calibrated recording and analysis of pitch, force and quality of vocal tones expressing happiness and saddness; and a determination of the ptich and forceof the subjective concepts of ordinary, soft and loud tones. *Speech Monographs, 2*(81-137).

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experemental Psychology: General, 117*, 34-50.

Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain Lang, 86*(1), 142-164.

Strous, R. D., Cowan, N., Ritter, W., & Javitt, D. C. (1995a). Auditory sensory ("echoic") memory dysfunction in schizophrenia. *Am.J.Psychiatry, 152*(10), 1517-1519.

Strous, R. D., Grochowski, S., Cowan, N., & Javitt, D. C. (1995b). Dysfunctional encoding of auditory information in schizophrenia. *Schizophrenia Research, 15*, 135.

Suslow, T., Roestel, C., Droste, T., & Arolt, V. (2003). Automatic processing of verbal emotion stimuli in schizophrenia. *Psychiatry Res, 120*, 131-144.

Susser, E. S., & Lin, S. P. (1992). Schizophrenia after prenatal exposure to the dutch hunger winter of 1944-45. *Arch Gen Psychiatry, 49*, 983-988.

Sweet, R. A., Bergen, S. E., Sun, Z., Sampson, A. R., Pierri, J. N., & Lewis, D. A. (2004). Pyramidal cell size reduction in schizophrenia: Evidence for involvement of auditory feedforward circuits. *Biol Psychiatry, 55*(12), 1128-1137.

Talairach, J. T., P. (1988). *Co-planar steriotaxic atlas of the human brain, 3 dimensional proportional ssystem: An approach to cerebral imaging*. New York: Thieme Medical Publishers.

Turner, J. (1964). Schizophrenics as judges of vocal expressions of emotional meaning. In J. Davitz (Ed.), *The communication of emotional meaning* (pp. 129-142). {Davitz, 1964 #15}: Greenwood Press.

Umbricht, D., & Krljes, S. (2005). Mismatch negativity in schizophrenia: A meta-analysis. *Schizophr Res, 76*(1), 1-23.

Van Lancker, D., & Sidtis, J. J. (1992). The identification of affective-prosodic stimuli by left- and right-hemisphere-damaged subjects: All errors are not created equal. *J Speech Hear Res, 35*(5), 963-970.

Vollm, B. A., Taylor, A. N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., et al. (2006). Neuronal correlates of theory of mind and empathy: A functional mag-netic resonance imaging study in a nonverbal task. *Neuroimage, 29*(1), 90-98.

Weinberger, D. R., & Gallhofer, B. (1997). Cognitive function in schizophrenia. *Int.Clin.Psychopharmacol., 12 Suppl 4*, S29-S36.

Weintraub, S., Mesulam, M. M., & Kramer, L. (1981). Disturbances in prosody. A right-hemisphere contribution to language. *Arch.Neurol, 38*(12), 742-744.

Wexler, B. E., Stevens, A. A., Bowers, A. A., Sernyak, M. J., & Goldman-Rakic, P. S. (1998). Word and tone working memory deficits in schizophrenia. *Arch Gen Psychiatry, 55*(12), 1093-1096.

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

Winkler, I., Tervaniemi, M., Huotilainen, M., Ilmoniemi, R., Ahonen, A., Salonen, O., et al. (1995). From objective to subjective: Pitch representation in the human auditory cortex. *Neuroreport, 6*, 2317-2320.

Winner, E., & Leekam, S. (1991). Distinguishing irony from deception: Understanding the speakers's second order intention. *British Journal of Developmental Psychology, 9*, 257-270.

Zatorre, R. J. (1985). Discrimination and recognition of tonal melodies after unilateral cerebral excisions. *Neuropsychologia, 23*(1), 31-41.

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends Cogn Sci, 6*, 37-46.

Zatorre, R. J., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *J Neuroscience, 14*(4), 1908-1919.

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.

Zatorre, R. J., & Samson, S. (1991). Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain, 114 (Pt 6)*, 2403-2417.