PERCEPTION AND PRODUCTION IN AUTISM SPECTRUM DISORDER

PERCEPTION AND PRODUCTION IN AUTISM SPECTRUM DISORDER

By

ANNE-MARIE R. DEPAPE, B.A., M.A.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the

Requirements for the Degree

Doctor of Philosophy

McMaster University

McMaster University DOCTOR OF PHILOSOPHY (2012) Hamilton, Ontario
(Psychology)

TITLE: Perception and Production in Autism Spectrum Disorder

AUTHOR: Anne-Marie R. DePape, B.A., M.A. (McMaster University)

SUPERVISOR: Dr. Laurel J. Trainor

NUMBER OF PAGES: xv, 177

ABSTRACT

Many behavioural theories describe Autism Spectrum Disorder (ASD) as having roots as a social disorder. However, our research adds to previous studies showing that those with ASD with normal intelligence have perceptual problems that affect their social functioning. We report that those with ASD perform worse than controls in all of our speech tasks that measure the ability to filter speech in noise, specialization for native speech sound categories, and audio-visual integration of speech sounds. Those with ASD also performed worse on tasks measuring specialization for native musical meters, but not on tasks measuring the use of absolute pitch or knowledge of tonal harmony. This research is important because little is known about some of these areas of auditory processing. Thus, our battery forms a profile in which to understand speech and musical processing in ASD. This research also provides some explanation for why perceptual areas that develop early instead of late are most impaired in ASD, which can have implications for remediation. Besides perceptual problems, we found in other research that those with ASD have abnormal prosody, which varies according to language ability. We report that those with Autism Moderate Language Functioning (A-moderateL) use a restricted pitch range relative to those with Autism High Language Functioning (A-highL) and controls, whereas those with A-highL use a larger pitch range relative to those with A-moderateL and controls. We also found that A-moderateL speakers and controls, but not A-highL speakers vary acoustic features to mark words representing focus relative to topic. This research is important because identifying different ASD language subgroups might lead to more appropriate speech and language therapy. Overall, this thesis contributes to our understanding of auditory perception and production in ASD, which may be used to develop better remediation and early screening of this disorder.

PREFACE

This thesis consists of studies that have been published in peer-reviewed journals. Chapter 2 involves research that has been published in the journal *PloS ONE*. The author of this thesis is the primary author of this work, including experimental design, stimulus generation, participant recruitment, data collection, data analysis and manuscript preparation. The second author of this paper provided assistance with experimental design, stimulus generation and manuscript preparation. The third author of this paper provided assistance with stimulus generation, participant recruitment and manuscript preparation. The fourth author of this paper is the thesis supervisor of the primary author. Chapter 3 involves research that has been published in the journal *Frontiers in Psychology*. The author of this thesis is the primary author of this work, including experimental design and programming, stimulus generation, participant recruitment, data collection, data analysis and manuscript preparation. The second author of this paper provided help with experimental design, stimulus generation, data analysis and manuscript preparation. The third author of this paper provided help with experimental design, stimulus generation, participant recruitment and manuscript preparation. The fourth author of this paper is the thesis supervisor of the primary author.

ACKNOWLEDGMENTS

Many people have contributed to the completion of my dissertation. Thanks to my mom, Mary Anne DePape who provided unconditional love and support throughout this entire process. She helped me to stay focused on the task at hand, when often times, it was easier to give up. Thanks to my grandparents, Lucy and Harvey Levac who instilled in me, at a young age, the value of education. Thanks to my brother, Robbie DePape and sister-in-law, Aryan DePape as well as my niece, Soleen DePape and nephew, Dylan DePape, for their love and inspiration.

Thanks to my boyfriend, Ryan Winter for encouraging me to follow my dreams, no matter how challenging the times. His no nonsense approach to life has helped to keep things in perspective. Thanks to Ryan’s family, including Bonita and Tom Winter as well as Natalie and Paul Postiglione for their love and support.

Thanks to my mentor, Elizabeth Christie who has watched me develop into the person who I am today. I am fortunate for all of the love and support she has provided throughout the years.

Thanks to my best friend, Christine Quaglia for being there during the best and worst of times. She helped me to stay true to myself. Thanks to Christine’s parents, Louie and Wilma Quaglia for being two of my biggest fans.

Thanks to my friends, Ritu Bedi and Marissa Barnes from York University. They appreciate the hard work that has gone into my schooling. I am fortunate for all of the love and support they have provided throughout the years.

Thanks to the girls in the Hamilton chapter of the Canadian Association for Girls In Science. They inspired me with their enthusiasm and passion to learn.

Thanks to the Social Sciences and Humanities Research Council of Canada for allowing me to pursue my graduate degree and seek out opportunities that would help with my career.

Thanks to my thesis supervisor, Dr. Laurel Trainor for letting me pursue an area of research that I had not considered previously. She encouraged me to develop my own voice in my research. She also supported me in seeking out opportunities that would further my understanding of my research topic. I will continue to use the skills that I have learned in her lab in my career.

Thanks to Dr. Geoffrey Hall for agreeing to be on my PhD committee and allowing me to complete a research practicum in the fMRI lab. He encouraged me to ask good questions about my topic and provided an atmosphere of learning, creativity and collaboration.

Thanks to Dr. Aoju Chen for allowing me to complete an internship one summer and asking me to return the following summer as a Junior Visiting Scientist at the Max Planck Institute for Psycholinguistics. I will always be grateful for the kindness that she provided during my first visit to Europe. She encouraged me to test my research ideas and to expand my research network. I have formed friendships from this experience that will last a lifetime.

Thanks to Dr. Dan Bosnyak for his feedback on this thesis. I am also grateful to him for allowing me to complete a statistics module in his lab so that I could complete my last degree requirement. I appreciate the time and guidance he provided with this course.

Thanks to Lisa Schumph, Deanna Finch-Smith and Justin Dewaard at the Salvation Army-Lawson Ministries. They helped me to get the word out about my research. They also included me in many ASD projects that were taking place in the Hamilton area, which allowed me to get more hands-on experience with my research topic. I admire their enthusiasm and dedication to helping those with ASD.

Finally, thanks to all the families who participated in my research. They invited me into their homes, sometimes living as far as Toronto or Peterborough. Others drove to the campus and gave up their weekends to participate in my research. They shared their stories with me. They told me about school, growing up, and what it was like to have ASD. I know that I am a better person and researcher because of the first-hand experience I had working with these families. Thank you!

TABLE OF CONTENTS

Chapter 2: Auditory Perception in Autism

Chapter 3: Speech Production in Autism

LISTS OF FIGURES

LISTS OF TABLES

LIST OF ABBREVIATIONS

ASD = Autism Spectrum Disorder

AS = Asperger's syndrome

HFA = High-Functioning Autism

PPVT = Peabody Picture Vocabulary Test

Leiter = Leiter International Performance Scale

A-moderateL = Autism Moderate Language Functioning

A-highL = Autism High Language Functioning

DECLARATION OF ACADEMIC ACHIEVEMENT

Anne-Marie DePape is the first author of this thesis. She is responsible for helping to create a battery of tests, pilot testing these measures with normal controls, recruiting adolescent participants with ASD and controls, testing these participants, analyzing the data and preparing the manuscript that is presented in chapter 2. She is also responsible for creating a prosody test, pilot testing this measure with normal controls, recruiting adult participants with ASD and controls, testing these participants, analyzing the data and preparing the manuscript that is presented in chapter 3.

Dr. Laurel Trainor is the thesis supervisor of the primary author. She stimulated the first author's interest in ASD and encouraged her to seek out opportunities that would help her to better understand her work (e.g., internship at Max Planck Institute). Dr. Trainor played a critical role with respect to the experimental design, stimulus generation, participant recruitment, data analysis and manuscript preparation.

Dr. Geoffrey Hall is an author on both of the papers presented in this thesis and was a member of the first author's PhD committee. He encouraged the first author to bring in brain research in order to provide a more comprehensive look of ASD perception and production. Dr. Hall played a critical role with respect to the experimental design, stimulus generation, participant recruitment and manuscript preparation.

Dr. Barbara Tillmann is an author on one paper (Chapter 2). She provided stimuli for one of the tasks used in the auditory battery to measure harmonic priming. She also provided assistance with respect to data analysis and manuscript preparation.

Dr. Aoju Chen is an author on one paper (Chapter 3) and was the supervisor of the first author when she completed an internship at the Max Planck Institute for Psycholinguistics. Dr. Chen provided guidance to the first author with respect to how to annotate and analyze acoustic features from speech recordings of adults with ASD. Dr. Chen encouraged the first author to present this research at international conferences, including the Workshop on Prosody and Meaning in Barcelona. Dr. Chen played a critical role with respect to the experimental design, stimulus generation, data analysis and manuscript preparation.

Dr. Dan Bosynak is not an author on this thesis, but was involved in this research as a member of the primary author's PhD committee. He provided feedback about stimulus design and data analysis via progress reports submitted each year by the first author to her committee.

CHAPTER 1

GENERAL INTRODUCTION

Autism Spectrum Disorder (ASD) is characterized by repetitive and restrictive interests, problems with communication, and impaired social functioning (American Psychiatric Association, 1994). ASD affects 9 in 10,000 individuals (Fombonne, 2003) and is nearly three times more likely to occur in males than females (Burd, Fisher, & Kerbeshian, 1987). Although no cause has been identified, many behavioural theories of ASD describe it as a social disorder. For example, research shows that those with ASD have poor Theory of Mind (ToM) (see review by Baron-Cohen, 2000). ToM is important for being able to understand the thoughts and feelings of others and normally develops between the ages of 3 and 5 years old (Wellman, 1993). Poor ToM may explain why those with ASD have problems identifying the emotions conveyed by other people's facial expressions and why they find it difficult to interact in meaningful ways with others.

Although impaired social functioning is a critical feature of ASD, a broader perspective might be necessary for understanding the full range of behaviours seen in this disorder, especially in light of evidence from perception research. Our research is not the first to show that abnormal perception can impact social functioning and communication (e.g., Peppé, McCann, Gibbon, O'Hare, & Rutherford, 2007), but it adds to a growing body of literature. In the first part of this dissertation (chapter 2), we show that those with ASD have perceptual abnormalities, such that they tend to focus on certain types of information in a stimulus that those who are typically developing have learned to ignore because it is not important for extracting the meaning of that stimulus. In the context of speech and music, learning to differentially process the sound features important for the particular language or musical system in the environment is known as enculturation. Here we explore various aspects of speech and music perception in order to understand when those with ASD perform worse than, similarly or better than controls. In the second part of this dissertation (chapter 3), we show that those with ASD typically have abnormal prosody, which might be due to the fact that their impaired perception sets them on a different trajectory with respect to language and communication. This impacts their ability to emphasize the important words in an utterance and makes it more difficult for individuals with this disorder to achieve their desired communicative intent. Taken together, this research provides a broad context in which to understand the full range of behaviours seen in ASD, including those of the autistic triad (impaired social functioning, repetitive and restrictive interests, and poor communication, APA, 1994). This research also contributes to our understanding of auditory perception and production in ASD, which may be used in the future to develop more appropriate remediation strategies and early screening practices for this group.

Perception in ASD

Those with ASD show a local processing bias where the parts of a stimulus are favoured over the whole (see review by Happé & Frith, 2006). This processing bias makes those with ASD well suited to extract the simple details of a stimulus, which, often times, typically developing individuals find irrelevant because they do not have meaning on their own. As an example, participants are asked in the Embedded Figures Task to find a simple shape that is located in a complex pattern (Shah & Frith, 1983). Research shows that those with ASD are faster and more accurate at locating the target shape than those who are typically developing (Shah & Frith, 1983). Similar results have been reported with the Navon Task where participants are presented with a large letter (e.g., "A") that has smaller letters printed inside of it that are either congruent ("a-a-a") or incongruent ("b-b-b") with the large letter. Research shows that those with ASD make fewer errors than controls when responding to the small letters over the large one (Plaisted, Swettenham, & Rees, 1999). The reason for this performance, according to the Weak Central Coherence (WCC) theory, is that the tendency to bring local elements together as a coherent whole is impaired among individuals with this disorder relative to those who are typically developing (Frith, 1989). However, although the results from the Navon task suggest that local processing is typically very good in ASD, they do not indicate that global processing is necessarily impaired in these individuals.

The Enhanced Perceptual Functioning (EPF) theory postulates that global processing is intact in those with ASD, in addition to enhanced low level processing (Mottron & Burack, 2001). In another study that used the Navon task, participants were instructed to respond to the global letter and ignore the local letter (Plaisted et al., 1999). The results of this work showed that those with ASD processed the large letter similarly as controls and, therefore, they showed normal global processing (Plaisted et al., 1999). This pattern of intact global processing has also been found in other research. For example, those with ASD had similar reaction times as controls when asked to combine visual details in a series of tasks, such as identifying the letter "s" when made up of a segmented line, indicating typical configural processing (Mottron, Burack, Iarocci, Belleville, & Enns, 2003). In the musical domain, even though those with autism have been shown to have intact or superior local processing in the form of pitch interval discrimination (e.g., Bonnel et al., 2003; Mottron, Peretz, & Ménard, 2000), they also show relatively normal global processing for melodic contours (Heaton, 2005; Mottron et al., 2000) and for chord sequences (Heaton, 2003; Heaton et al., 2007). This research indicates the need to have a battery that measures local and global processing across a variety of tasks in order to provide an accurate profile of auditory perception in ASD.

In addition to a local processing bias, research shows that those with ASD have problems with the way in which they attend to information in their environment and this can impact perception. For example, those with ASD are described as having “spotlight” attention because the aspects that they attend to are located in a narrow field of focus (Townsend & Courchesne, 1994). This suggests that performance among those with ASD is similar to those who are typically developing if the task requires attention to a narrow field of focus, but worse than controls if the task requires attention outside of this narrow field of focus. Other research indicates that once those with ASD fixate on an object inside their field of focus, they have problems shifting their attention away from that object (Mann & Walker, 2003). Thus, those with ASD may attend to aspects of a stimulus that those who are typically developing find irrelevant. Given that they have problems shifting their attention between objects in their environment, they may gain a different type of expertise than controls. This cognitive profile would make complex environments particularly challenging for those with ASD, where they need to attend to, and integrate information from, multiple objects at once and where there is a high demand placed on being able to shift attention flexibly between objects. Interestingly, these difficulties of attention in ASD may help to explain the insistence of those with ASD on sameness as well as the repetitive behaviours and restrictive interests that are typical of individuals with this disorder (APA, 1994). Critically, research indicates that problems with shifting attention can be detected in the infancy period (Zwaigenbaum, Bryson, Roberts, Brian, & Szatmari, 2005) as is the case for perceptual abnormalities (Pierce, Conant, Hazin, Stoner, & Desmond, 2010). Thus, these research areas involving perception and attention are likely to play important roles in the future with respect to the early screening of ASD.

Early Brain Development

We have described how information may be processed in ASD, but it is also important to understand the potential mechanism that might be driving abnormal perception in this group. There are a number of reports that musical processing might be more spared compared to language processing in ASD (e.g., Jarvinen-Pasley & Heaton, 2007), but none of the behavioural theories of ASD, whether they focus on the social or the perceptual causes of this disorder, have attempted to explain why music might be a more natural way than speech for those with ASD to communicate. Research by Heaton and colleagues has demonstrated that those with ASD experience less interference than controls when presented with speech and asked to respond to the pitch and ignore the word meaning (Heaton, Davis, & Happé, 2008; Järvinen-Pasley & Heaton, 2007; Järvinen-Pasley, Wallace, Ramus, Happé, & Heaton, 2008). This tendency might place those with ASD at a disadvantage with respect to typical speech and language development and conversely, make musical stimuli particularly attractive for this clinical group. It could also promote enhanced perceptual processing of musical stimuli as shown by better pitch discrimination in ASD than controls (Bonnel et al., 2003). Advances in neuroimaging techniques have helped to promote the idea of ASD as a neurodevelopmental disorder, even though this disorder is currently diagnosed using behavioural criteria. Research shows that those with ASD have a normal or smaller than normal brain size at birth (Courchesne et al., 2003; Gillberg & de Souza, 2002). After this period, those with ASD experience accelerated brain growth such that by age 5, they have reached their maximum brain size, which is up to 10 years earlier than those who are typically developing (Courchesne, 2004; Courchesne et al., 2001). Research shows that as brain size fluctuates so does its short-distance and long-distance neural connectivity (Lewis & Elman, 2007). In ASD, there are reports of both decreased long-distance neural connectivity (Barnea-Goraly et al., 2004; Bird, Catmur, Silani, Frith, & Frith, 2006; Castelli, Frith, Happé, & Frith, 2002; Just, Cherkassky, Keller, & Minshew, 2004; Just, Cherkassky, Keller, Kana, & Minshew, 2007; Skukla, Keehn, & Müller, 2010) and increased short-distance neural connectivity (Casanova, 2004; Casanova, Buzhoeveden, Switala, & Roy, 2002; Courchesne & Pierce, 2005). Reduced long-distance neural connectivity makes it difficult for brain regions to communicate and so processes such as coordination and integration are impaired (Just et al., 2004; Just et al., 2007). This connectivity issue is consistent with the Neural Complexity Hypothesis, that suggests that brain activity is more likely to be impaired when complex auditory tasks that require processing in associative auditory areas are used over simple auditory tasks that require processing in primary auditory areas (see review by Samson, Mottron, Jemel, Belin, & Ciocca, 2006). Thus, although research finds that both local and global processing are possible among those with ASD, it is likely that the type of task used and the degree of coordination between brain areas (i.e. neural complexity) are important in order to understand perceptual processing in this group. Evidence of increased short-distance connectivity comes from studies of minicolumn structure, which has been found to be smaller and more densely packed among those with ASD (Casanova, 2004; Casanova et al., 2002). From a metabolic standpoint, smaller neurons favour short-distance over long-distance neural connections and so this might help to explain the abnormal pattern of short-range connectivity seen in this group (Casanova, 2004; Casanova et al., 2002). Other research indicates that there might be abnormal pruning among those with ASD (Frith, 1991), which might give individuals with this disorder access to low level information via preserved rudimentary connections. The result is that increased short-distance connectivity might allow for greater fine-tuning of information that is not seen in controls (Casanova et al., 2002; Casanova, Switala, Trippe, & Fitzgerald, 2007), but it also could mean that those with ASD are more susceptible to problems with sensory modulation because brain activation spreads less selectively. Interestingly, sensory problems were included in the DSM-III (APA, 1980), but removed from the DSM-IV (APA, 1994) because they were not considered to be core symptoms of this disorder. These problems can involve any sensory domain and can manifest in those with ASD being insensitive to sensory input (hypo-reactivity), being too sensitive to sensory input (hyper-reactivity) or both, relative to typical controls (O'Neill & Jones, 1997; Watling, Deitz, & White, 2001). However, sensory problems will likely be included in the upcoming DSM-V (APA, 2013), as issues related to sensory processing continue to be a concern for those with ASD. Particularly relevant to the study of adolescents with ASD, sensory problems appear to increase between the ages of 6 and 9 years old (see review by Ben-Sasson, Hen, Fluss, Cermak, Engel-Yeger, & Gal, 2009), although the research is mixed about whether symptom severity decreases thereafter (Crane, Goddard, & Pring, 2009; Kern et al., 2006; Kern et al., 2007). Problems with sensory modulation are shown in a description given by a very high functioning individual with ASD: "The nerve endings on my skin were supersensitive. Stimuli that were insignificant to most people were like Chinese water torture" (Grandin, 1992). It is possible that such sensory-based problems could arise from the pattern of early brain development, along with abnormal neural connectivity and that this abnormal neural development might help to explain why those with ASD perceive the world differently than those who are typically developing.

Processing in the Auditory Domain

Most of the research concerning perceptual processing has been in the visual domain (e.g., Behrmann, Thomas, & Humphreys, 2006; Frith, 1989; Plaisted et al., 1999) even though it might be more difficult to avoid aversive sounds than sights in the environment as we cannot close our ears. Among those with ASD, there are reports of both hypersensitivity to sound (Baranek, Boyd, Poe, David, & Watson, 2007; Dahlgren & Gillberg, 1989; Gomes, Rotta, Pedroso, Sleifer, & Danesi, 2004; Kern et al., 2006; Rimland & Edelson, 1995; Rosenhall, Nordin, Sandstrom, Ahlsen, & Gillberg, 1999) and hyposensitivity to sound (Baranek, 1999; Dawson et al., 2004; Kern et al., 2006; Osterling & Dawson, 1994; Werner, Dawson, Osterling, & Dinno, 2000). However, research shows that hyposensitivity rather than hypersensitivity to sound differentiates individuals with ASD at an early age from those who are typically developing (Rogers & Ozonoff, 2005). Hyposensitivity to sound is described in the following: “James was typical of many autistic children. A sharp noise near his ears, for example, sharp hand clapping, produced only eye blinking. There was no other body or facial adjustment and no verbal indication that the sound was recognized” (Goldfarb, 1956). Indeed, those with ASD are often misdiagnosed as being deaf (Rabin & Katzman, 1998). Even though hyposensitivity to sound is documented in some of the earliest cases of ASD, it appears that this way of responding is still misunderstood. Importantly, individuals can have problems modulating and integrating sensory information, despite the fact that their hearing is normal (Baranek, 2002), leading to behaviours suggesting both hyposensitivity and hypersensitivity. Thus, individuals with ASD who cover their ears when they are in a crowded room may do so, not because they are hearing these noises louder than the average listener, but because they cannot filter out these sounds. And once a threshold is reached, it is possible that those with ASD simply switch off and do not meaningfully process any sound input for a period of time. From a clinical perspective, understanding the root causes of abnormal auditory processing is important in determining the way in which other people respond to those with ASD and, critically, the nature of effective interventions and the more effective ages at which to administer the interventions.

Speech Processing

With respect to speech processing, an important issue is whether those with ASD recognize these sounds as special and distinct from other sounds in their environment. Research with children shows that speech sounds elicit bilateral activation of speech specific brain regions among those with ASD, while the same sounds elicit more left hemisphere than right hemisphere activation among those who are typically developing (Boddaert et al., 2004). Such a pattern of brain activation has been found when typically developing individuals listen to foreign speech, as in one study where American-English participants listened to sounds from a South African language (Best & Avery, 1999). Thus, one issue is that those with ASD may have problems processing speech sounds because they attend to the full range of sounds in their environment instead of focusing more selectively on speech. Other research shows that although speech activates speech-specific regions in ASD, it also produces activation in additional regions, such as the parietal lobes, brainstem and cerebellum (Boddaert et al., 2004). Music may therefore serve as an alternative way for those with ASD to communicate because they can extract the emotional content of music. Research shows that those with ASD associate major musical and minor musical modes with happy and sad faces, respectively (e.g., Heaton, Hermelin, & Pring, 1999). Music might also provide a medium in which to remediate abnormal speech and language patterns. That is, it may be possible that the training that occurs in a musical context may transfer to another context thereby leading to improvements, such as with emotional deficits and problems with social communication (Allen & Heaton, 2010).

Although humans have the capacity to develop speech and language, realization of this capacity may be abnormal if the individual is not motivated to attend to speech. Research among those with ASD shows that they have a reduced preference for human speech sounds relative to those who are typically developing (Paul, Chawarska, Fowler, Cicchetti, & Volkmar, 2007). This reduced social interest can impact the normal development of processes such as enculturation, which are essential for efficient speech and language processing. Also, the fact that those with ASD show atypical auditory processing including a reduced orienting response (e.g., not responding when their name is called) indicates abnormal sound processing that likely manifests at behavioural as well as neural levels (see review by O'Connor, 2012). In this dissertation, we argue that those with ASD will be impaired in three key areas of speech processing: filtering, categorization, and audio-visual integration. This is partly because of the impaired social functioning that is seen in this group and partly because of the accelerated brain growth that they experience, particularly in the first few years of life. The effect is that speech will develop in less constrained ways among those with ASD thereby, resulting in a focus on the low level information in these sounds.

1. Auditory Filtering. Filtering involves the ability to attend to one source of auditory information, despite the fact that there are other sources of information that are actively competing for processing. Being able to filter means that you can carry a conversation in a noisy environment, which is something that we do on a regular basis. Research shows that those with ASD have problems with filtering, as evidenced by questionnaire data (Ashburner et al., 2008; Baker, Lane, Angley, & Young, 2008; Lane, Young, Baker, & Angley, 2010; Rogers, Hepburn, & Wehner, 2003; Schoen, Miller, Brett-Green, & Nielsen, 2009). These data suggest that those with ASD have filtering problems as reflected by scores from parent raters on questionnaire items. For example, in the Short Sensory Profile, parents tend to rate statements “doesn’t respond when name is called but you know the child’s hearing is okay” and “distracted or has trouble functioning if there is a lot of noise around” as describing their child with ASD (McIntosh, Miller, Shyu, & Dunn, 1999). Other measures include behavioural tasks (Alcántara, Weisblatt, Moore, & Bolton, 2004; Groen et al., 2009), auditory cortical event-related potentials (ERPs) and auditory brainstem responses (ABRs), (Lepistö et al., 2009; Russo, Nicol, Trommer, Zecker, & Kraus, 2009; Russo, Zecker, Trommer, Chen, & Kraus, 2009; Teder-Sälejärvi et al., 2005), all showing that those with ASD require a higher signal-to-noise ratio than controls in order to perceive speech in noise. However, little is known from these studies about how performance may be affected when irrelevant information comes from a different sound location in a filtering task. Thus, we measure the signal-to-noise ratio needed to perceive sentences presented to one ear while ignoring simultaneous sentences to the other ear. We argue that such a task is a more ecologically valid way of measuring filtering in noise as it reflects the typical situations of different sound sources in naturalistic environments.

2. Phoneme Categorization. Infants are able to discriminate between speech sound categories in native and foreign languages at 6 months of age (e.g., Werker & Lalonde, 1988). This means that they do not differentiate between speech sound differences that are part of their first language phonological system and those that are not. As an example, infants who are raised in Japanese speaking homes can discriminate between English /r/ and /l/ speech sound categories, despite the fact that this contrast does not occur in their native language (Goto, 1971). After 12 months of age, infants are like adults in that they have learned to ignore those contrasts that occur in foreign languages in favour for those that occur in their native language (Best, McRoberts, & Goodell, 2001). This specialization for their native language helps to promote future learning about speech and language by constraining the amount of information that individuals are required to process in detail. Given the impaired perception and socialization seen in ASD, it is possible that individuals with this disorder will show less specialization for their native language. We found one study that examined how those with ASD process speech sounds (Constantino et al., 2007), but it examined this processing only in foreign speech. Thus, this research cannot address the issue of specialization for native speech sound categories among those with ASD. Here we examine the issue of specialization by measuring the difference in performance across foreign speech sound categories that map onto one speech sound category (foreign) or two speech sound categories (native) in English. Performance in typical controls has been found to be at chance for the one category mapping condition and at ceiling for the two category mapping condition (Best et al., 2001). In the case of those with ASD, we expect that they will show reduced specialization for native speech sound categories relative to controls. We argue that this area of speech processing is important because it impacts the ability of those with ASD to communicate effectively in their environment. Furthermore, if those with ASD show less specialization for their native language then it could point to the need for very early speech and language remediation.

3. Audio-Visual Integration. Understanding speech requires that listeners integrate cues from multiple sources, such as the lips, eyes and hands, in addition to what the speaker is actually saying. Research in this area shows that those with ASD are less likely than typically developing individuals to integrate audio-visual information and are, instead, more likely to attend only to what they hear. Most of the research involving audio-visual integration of speech sounds is based on the McGurk effect where participants are asked to report what they hear after they are presented with audio and video inputs that are incongruent (audio "ba" and visual "ga") (McGurk & MacDonald, 1976). If participants integrate, the result is that they "hear" a third percept ("da"), which represents a fused response of what they see and what they hear (McGurk & MacDonald, 1976). Research involving those with ASD shows that these individuals integrate audio-visual information less than those who are typically developing and so they are less susceptible to this illusion (de Gelder, Vroomen, & van der Heide, 1991; Iarocci, Rombough, Yager, Weeks, & Chua, 2010; Williams, Massaro, Peel, Bosseler, & Suddendorf, 2004). There are conflicting reports, however, about whether those with ASD are impaired at lip-reading, and whether this affects audio-visual integration (Iarocci et al., 2010; Smith & Bennetto, 2007; Williams et al., 2004). Thus, we measure audio-visual integration in order to determine if the problem stems from when the audio and the visual information come together or when one of these modalities is presented on its own. We argue that integration is important, particularly in noisy environments where speech may be degraded or when the audio and the visual inputs are incongruent, such as in the case of sarcasm. These results will therefore further our understanding of the language and communication abilities of those with ASD.

Music Processing

The way that speech is processed among those with ASD has prompted some researchers to question whether music might be a more natural way for these individuals to communicate. Anecdotally, parents often report how their sons or daughters with ASD seem to be intrinsically motivated to listen to or play music, despite the fact that they do not have any formal musical training. There is evidence to motivate this research question given that the typical ASD brain shows a reversed or absent asymmetry for areas that are important for speech and language processing. In musical pitch processing, the right hemisphere tends to be more strongly activated than the left hemisphere (e.g., Johnsrude, Penhune, & Zatorre, 2000). Research shows that those with ASD have a larger right than left hemisphere for certain brain regions, including Broca's area (inferior frontal) and Wernicke's area (superior temporal) (Herbert et al., 2002) which, on the right, are important for musical processing. These brain abnormalities may therefore, hinder some aspects of speech and language processing, but critically, enhance music perception. This is consistent with behavioural and electrophysiological research showing intact or enhanced pitch processing of pure tones, complex tones and speech sounds, perhaps due to atypical right hemisphere dominance in this clinical group (see review by Haesen, Boets, & Wagemans, 2011). Interestingly, when enhanced pitch processing is found in older individuals with ASD, they tend to also have language impairments (Bonnel et al., 2010; Jones et al., 2009). Thus, enhanced pitch perception may make linguistic information less salient or conversely, impaired language may contribute to those with ASD being overly focused on the low level features of these sounds. Thus, we explore musical processing among those with ASD in three key areas: pitch memory, meter categorization and harmonic priming. We chose these areas in order to examine processing in tasks that rely on pitch as well as rhythm. Research shows that rhythm might support an important social function in music, such as through dancing or clapping (Hannon & Trainor, 2007; Kirschner & Tomasello, 2009) and so we might expect a different pattern of performance with respect to rhythm and pitch processing among those with ASD.

1. Absolute Pitch. For most adults and even infants who are as young as 6 months of age, pitch is processed using a relative pitch code (Plantinga & Trainor, 2005; Trainor & Trehub, 1992; Trehub, 2001; Trehub, Bull, & Thorpe, 1984). This means that we recognize a familiar song, such as “Mary Had a Little Lamb”, even if the piece is sung on a higher or lower starting pitch, so long as the distances between these pitches is preserved. For less than 5 out of every 10,000 individuals, pitch is processed using an absolute pitch code (Bachem, 1955; Brown et al., 2003). In these cases, an individual is able to name or produce a pitch that is heard, without relying on an external standard (Takeuchi & Hulse, 1993). Research in this area shows that those with ASD might have access to absolute pitch information (Brenton et al., 2008; Heaton, Pring, & Hermelin, 1999; Mottron, Peretz, Belleville, & Rouleau, 1999; Young & Nettelbeck, 1995). However, many of these studies are poorly controlled as they do not include control groups or they base their findings on case studies. Many of these studies also require that participants be able to read music, which may not be typical of the entire population with ASD. One exception is a collection of studies by Heaton and colleagues that examined pitch memory. In these studies, those with ASD and controls were presented with tones and asked to associate each tone with a picture or to disembed associated tones from musical chords (Heaton, 2003; Heaton et al., 1999; Heaton, Hermelin, & Pring, 1998). The results show that those with ASD performed better than controls, suggesting that pitch memory is better retained over time and, therefore, is less susceptible to decay in this clinical group (Heaton, 2003; Heaton et al., 1999; Heaton, et al., 1998). In other research where visuo-spatial locations were associated with tones, it was found that a subgroup with ASD had better pitch memory than controls as shown in both accuracy and reaction time data, but this was not typical of the entire clinical sample (Heaton, Williams, Cummins, & Happé, 2008). We measure the prevalence of absolute pitch in order to understand its rate among the entire sample of those with ASD using a task that does not require note naming, note association or formal musical training. We argue that the use of an absolute pitch code can lead to a different appreciation of music by taking attention away from the musical intervals that are most important to a melody. Thus, this area will further our understanding of musical processing in ASD.

2. Meter Categorization. Meter involves the perceptual extraction of an underlying pulse that can be broken down into different hierarchical beat patterns (Hannon & Trehub, 2005). In simple meters, which are common in Western music, the strong and weak beat durations form simple ratios (e.g., 2:1) (Hannon & Trehub, 2005). In complex meters, which are common in Eastern Europe, the alternation between beat durations form complex ratios (e.g., 3:2) (Hannon & Trehub, 2005). Similar to the specialization that occurs in speech, infants after 12 months of age learn to ignore the metrical categories of foreign music in favour for the metrical categories of their native music. We know of no studies on this topic involving those with ASD. Thus, we examine the issue of specialization for native metrical categories by measuring the difference in performance across native and foreign meter conditions. We argue that this aspect of musical processing is important in order to determine if there is an uneven profile in the way in which rhythm and pitch are perceived among those with ASD given the strong social component involved with rhythm. Given the social impairments found in ASD it is expected that rhythm will be more impaired than pitch processing.

3. Harmonic Priming. As with meter, experience with Western music during development leads to perceptual specialization for the rules of tonal harmony in that musical system. Between 4 and 7 years of age, typically developing children have acquired some implicit knowledge of Western harmonic structure (Corrigall & Trainor, 2010; Schellenberg, Bigand, Poulin-Charronnat, Garnier, & Stevens, 2005; Trainor & Corrigall, 2010; Trainor & Trehub, 1994). Chord sequences follow preference rules such that a sequence sets up expectations in enculturated listeners for which chords are likely to come next, and these expectations can be measured implicitly with reaction times (e.g., Bigand, Madurell, Tillmann, & Pineau, 1999). In a typical implicit paradigm, chord sequences are presented, half of which have expected final chords and half unexpected. Reaction times on an irrelevant variable, such as to indicate the timbre of the final chord, are compared for sequences with expected and unexpected endings. For example, ending a sequence with a dominant to tonic chord progression produces faster reaction times than endings with a tonic to subdominant chord progression (Bigand et al., 1999; Schellenberg et al., 2005; Tillmann, Bigand, Escoffier, & Lalitte, 2006; Tillmann, Peretz, Bigand, & Gosselin, 2007). Those who have a weaker representation of these rules, perhaps because they failed to internalize rules of Western tonality, may instead attend to the surface features in chord progressions, such as whether some chords repeat. One study asked participants with ASD to report aloud if chord sequences sounded complete or not (Heaton, Williams, Cummins, & Happé, 2007). Those with ASD were found to process chord sequences similarly to controls in this respect (Heaton et al., 2007). This study measured accuracy, but reaction time might be a more sensitive measure, and thus, may reveal differences between those with ASD and controls. Therefore, we examine harmonic priming by measuring both accuracy and reaction time. We argue that this aspect of musical processing is important in order to understand whether those with ASD have formed expectations about music simply by listening to it in their environment, which in turn, can impact processes that are important for music learning and memory.

Speech Production

Thus far, we have discussed that those with ASD have abnormal perception and, critically, that this impacts the way in which language and communication develops. In the second part of this dissertation, we examine speech production in ASD. Here we show that those with ASD are impaired in the way that they vary an acoustic feature, such as pitch in their speech and how they emphasize important information by varying this acoustic feature in their utterances. This may be due to the fact that abnormal perception sets up a cascade of events such that those with ASD are placed on a different developmental trajectory, which impacts their motivation and ability to communicate effectively.

Social Development

An early sign of ASD is an inattention to faces (Osterling, Dawson, & Munson, 2002), which can impact the ability to achieve certain developmental milestones. For example, joint attention requires that infants attend to adult faces in order to share an experience, something that is important for later language development (Tomasello & Todd, 1983). Attending to faces is also important for the development of Theory of Mind (ToM), whereby children learn that others have thoughts and feelings, an area that is known to be impaired in ASD (Baron-Cohen, Leslie, & Frith, 1985). Although it is difficult to untangle the effects of these experiences on biology, research shows that the brains of those with ASD respond very differently to faces. In one study, faces produced less activation in the fusiform gyrus for those with ASD than those who are typically developing (Dalton et al., 2005; Pierce, Muller, Ambrose, Allen, & Courchesne, 2001; Schultz et al., 2000). Research also shows that when those with ASD fixate on the eyes of the human face, it produced activity in the amygdala, which is a brain area that is important for signaling threatening social stimuli (Dalton et al., 2005). If the amygdala is activated during face processing then those with ASD might be less motivated to attend to these stimuli in the future, thereby producing a cascade of developmental effects that ultimately impact their motivation and ability to communicate effectively.

Prosody and General Language Ability

There are conflicting reports describing the speech of those with ASD including: “robotic”, “wooden”, “stilted”, “bizarre”, “over precise”, “monotone” and “singsong” (Baltaxe & Simmons, 1985; Baron-Cohen & Staunton, 1994; Fay & Schuler, 1980; Frith, 1991). Past work in this area is primarily based on subjective ratings, rather than robust acoustic analysis on how those with ASD vary prosodic features in their speech. Of the studies involving acoustic measurements, most concern the prosody of children with ASD rather than adults (Baltaxe, Simmons, & Zee, 1984; Bonneh, Levanon, Dean-Pardo, Lossos, & Adini, 2011; Diehl et al., 2009; Fosnot & Jun, 1999; Green & Tobin, 2009; Grossman, Bemis, Skwerer, & Tager-Flusberg, 2010; Hubbard & Trauner, 2007; Nadig & Shaw, 2011; Paccia & Curcio, 1982; Paul, Bianchi, Augustyn, Klin, & Volkmar, 2008; Sharda et al., 2010). However, there is a need for research on the production patterns of adults with ASD as this is an area that can have lifelong consequences, affecting their ability to make friends and achieve meaningful employment (Paul, Augustyn, Klin, & Volkmar, 2005). Research also indicates that abnormal prosody is resistant to change, probably because it is rarely targeted directly in speech and language therapy (Bellon-Harn, Harn, & Watson, 2007).

Studies employing acoustic analyses have generally found increased pitch variability in children with ASD, whether the corpus analysed consisted of isolated words (Bonneh et al., 2011), conversations (Green & Tobin, 2009; Nadig & Shaw, 2011; Sharda et al., 2010), narratives (Diehl et al., 2009) or reading aloud (Green & Tobin, 2009). However, there appear to be individual differences. Baltaxe et al. (1984) found that children with ASD had either very narrow or very wide pitch ranges, suggesting heterogeneity among children. Similarly, Green and Tobin (2009) found that although children with ASD as a group showed larger pitch ranges and larger pitch variability compared to typically developing children, those with ASD could be divided into three distinct groups, consisting of those with narrow, typical or wide pitch ranges. Similar variance across individuals might also exist for prosodic use of duration, although there is less research on this question. Nadig and Shaw (2011) reported no difference in overall speech rate between children with and without ASD. In other studies, adults with ASD were found to produce less lengthening than controls on stressed syllables in imitative speech (Paul et al., 2008), but children with ASD were found to produce more lengthening than controls on stressed syllables in spontaneous speech (Grossman et al., 2010). Clearly, more research is needed in order to understand the prosodic use of duration in ASD.

Part of the variability in production patterns seen among those with ASD may be explained by their general language skills. Language ability is an important indicator in ASD, as language is highly predictive of the general prognosis for a child (see Kjelgaard & Tager-Flusberg, 2001). Language is also related to a number of specific abilities. For example, of children with ASD, only those with poor language skills show a low ability to suppress word meanings that are not consistent within a context; those with language skills in the normal range show normal context-dependent suppression (Brock, Norbury, Einav, & Nation, 2008; Norbury, 2005). Similarly, language ability predicts whether children with ASD use the appropriate amount of information in descriptions of objects according to the knowledge of their communication partner (Nadig, Vivanti, & Ozonoff, 2009). In one study, Norbury and colleagues (2009) used eye tracking while participants watched videos of peers interacting in familiar situations. Interestingly, they found that those with ASD and poor language skills were similar to normally developing controls in their viewing patterns of the eyes and mouths of their peers, whereas those with ASD and normal language ability spent less time than the other groups viewing the eyes. This suggests that language skills may not necessarily be connected with better communication skills, and indicates that the origins and nature of communication problems in ASD may differ between children with higher and lower language functioning. In the current study, we investigate the general and communicative use of prosody in high-functioning adults with ASD who score above or below the mean of the normal population on vocabulary, which is highly related to general language skills in ASD (e.g., see Kjelgaard & Tager-Flusberg, 2001).

Prosody and Information Structure

Good communication depends on varying acoustic features in speech, but it also involves doing so in a way that is useful to listeners. Utterances typically involve two meaningful aspects: the focus and the topic (Vallduví & Engdahl, 1996). The focus involves new information in an utterance, while the topic involves old information from the previous utterance (Vallduví & Engdahl, 1996). Given that focus words have a higher information value than topic words, the former is made more prominent than the latter, such as with larger pitch ranges and longer word durations (Chen, 2009). Speakers who use acoustic features such as these are able to capture the listener's attention and convey that they are tuned to their conversational partner. Those who do not do this may frustrate their partner and give the impression that they do not care whether good communication occurs. With respect to those with ASD, some research using subjective judgments of recorded utterances indicates that those with ASD make focus and topic words equally prominent (McCaleb & Prizant, 1985), whereas other research suggests that they make the beginning of a sentence prominent irrespective of its information value (e.g., Baltaxe et al., 1984; Baltaxe & Guthrie, 1987; Peppé, Cleland, Gibbon, O'Hare, & Martínez-Castilla, 2011; Peppé, McCann, Gibbon, O'Hare, & Rutherford, 2006, 2007; Shriberg et al., 2001). Most of these studies examined contrastive stress, where correct prominence is placed on the contrastive focus. For example, when presented with an informationally incorrect sentence such as "The green sheep has the ball" participants might respond, "No, the green *COW* has the ball" (Peppé et al., 2006), accenting the word correcting the information. The literature on speakers who are typically developing shows that focus information structure is marked using pitch to a lesser extent in the sentence-final position (object) than sentence-initial position (subject), and that such marking does not become adult-like until age 7 (Chen, 2011). Thus, in the present study we examine the marking of (non-contrastive) focus and topic in both sentence-initial and sentence-final positions for sentences that followed a subject-verb-object (SVO) format.

Research Purpose

Chapter 2 is a published paper in *PLoS ONE* showing the results of an auditory battery of tests that was administered to adolescents with ASD. This battery measures processing in speech (filtering, phoneme categorization and audio-visual integration) and music (absolute pitch, meter categorization and harmonic priming), with the goal of developing a perceptual profile for individuals with ASD. This profile informs readers about the nature of ASD, including how the behaviours involved in the autistic triad may be affected by perceptual problems. In the future, this research may provide some foundation in which to develop more appropriate remediation strategies and screening practices for ASD. Chapter 3 is a published paper in *Frontiers in Psychology* whose research purpose is to examine how adults with ASD vary their prosody with the use of pitch and duration, generally in their speech and as a function of the information structure. Its purpose is also to examine whether the level of current language ability of ASD speakers (which in our sample also reflected whether or not there had been an early language delay and whether a diagnosis of high-functioning autism [HFA] or Asperger's syndrome [AS] had been given) is associated in a predicable way with prosody use. There are inconsistent reports about how individuals with ASD sound when they are speaking ranging from "monotone" to "singsong", probably because much of the literature is based on subjective impressions rather than acoustic measurements. Thus, we examine speech production in order to understand these speech patterns among those with ASD and their potential effect on communication. This research will help to determine if there are critical periods in which early speech and language interventions should be in place for those with ASD. Taken together, this dissertation informs us about the perception and production abilities of those with ASD in order to help individuals with this disorder achieve their potential in both speech and music.

CHAPTER 2: INTRODUCTION

Impaired social functioning is a critical feature of ASD, but it cannot account for the full range of behaviours seen in this disorder, especially in light of evidence from perception research. For example, those with ASD show a local processing bias where the parts of a stimulus are favoured over the whole (see review by Happé & Frith, 2006). This processing bias makes those with ASD well suited to extract the simple details of a stimulus, which, often times, typically developing individuals find irrelevant because they do not have meaning on their own. One explanation for why those with ASD retain access to low-level information of a stimulus is because of decreased long-distance neural connectivity (e.g., Barnea-Goraly et al., 2004) and increased short-distance neural connectivity (e.g., Casanova, 2004), both of which could impact perception.

Here we measure processing in key areas of speech (auditory filtering, phoneme categorization and multisensory integration) and music (absolute pitch, meter categorization and harmonic priming). We examine these areas in order to determine if music might provide a more natural way for those with ASD to communicate, especially given the impairments found in speech (e.g., Boddaert et al., 2004). This profile will inform readers about the nature of ASD, including the range of behaviours involved in the autistic triad. It will also provide a foundation in which to develop more appropriate remediation strategies and screening practices for individuals with this disorder.

Auditory Processing in High-Functioning Adolescents with Autism Spectrum Disorder

Anne-Marie R. DePape[a], Geoffrey B. C. Hall[b], Barbara Tillmann[d], & Laurel J. Trainor[a,c]

a Psychology, Neuroscience & Behaviour, McMaster University

b Psychiatry and Behavioural Neurosciences, McMaster University

c Rotman Research Institute, Baycrest Hospital

d Lyon Neuroscience Research Center, CNRS-UMR 5292, INSERM U1028, Université Lyon 1

Running head: AUDITORY PROCESSING IN AUTISM

Corresponding author: Laurel J. Trainor, Department of Psychology, Neuroscience & Behaviour, McMaster University, 1280 Main Street West, Hamilton, ON, L8S 4K1, Canada, E-mail: ljt@mcmaster.ca, Phone: 1-905-525-9140 Ext. 23007, Fax: 1-905-529-6225

Keywords: Autism, auditory processing, perception, speech, music, adolescents

Acknowledgements: This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada to LJT and a graduate fellowship from the Social Sciences and Humanities Research Council of Canada to ARD.

Reprinted from DePape, A.M., Hall, G.B.C., Tillmann, B., & Trainor, L.J. (2012). Auditory processing in high-functioning adolescents with autism spectrum disorder. *PloS ONE*, *7,* 1-17. doi:10.1371/journal.pone.0044084

Abstract

Autism Spectrum Disorder (ASD) is a pervasive developmental disorder including abnormalities in perceptual processing. We measure perception in a battery of tests across speech (filtering, phoneme categorization, multisensory integration) and music (pitch memory, meter categorization, harmonic priming). We found that compared to controls, the ASD group showed poorer filtering, less audio-visual integration, less specialization for native phonemic and metrical categories, and a higher instance of absolute pitch. No group differences were found in harmonic priming. Our results are discussed in a developmental framework where culture-specific knowledge acquired early compared to late in development is most impaired, perhaps because of early-accelerated brain growth in ASD. These results suggest that early auditory remediation is needed for good communication and social functioning.

Auditory Processing in Adolescents with Autism Spectrum Disorder

Autism Spectrum Disorder (ASD) is a pervasive developmental disorder that includes abnormalities in perceptual processing [1], language and communication [2], and social interaction [3]. Although a diagnosis on the basis of social behavior and language delay is often not possible until a child is at least 3 years old, recent evidence suggests that perceptual processing differences are apparent in the infancy period [4-7]. Indeed the early perceptual capacities of those with ASD may set up a cascade of developments that contribute to the poor social skills and perseveration seen at older ages. ASD is associated with a particular processing style in which local stimuli details are very well processed, sometimes at the expense of global processing [8]. For example, those with ASD tend to perform better than those without ASD on tasks such as finding visual embedded figures [9-11]. Auditory processing is of particular interest in this regard as there are reports of both hypersensitivity to sound [12-17] and hyposensitivity to sound [15,18-21], both of which could interfere with the quality of communicative exchanges and thereby interrupt language and communication development. In this paper, we measure several aspects of auditory processing in speech and music with the purpose of developing an auditory profile that characterizes high-functioning ASD.

From a developmental standpoint, we might expect that aspects of speech and music learning that typically occur early in development, such as perceptual reorganization for native phonemic categories and musical metrical structure, might be particularly affected in ASD. There is evidence that brain growth is accelerated in ASD early in development (particularly 6 to 24 months of age) and slows sooner compared to normal development [22-24]. There are also reports of both an overdevelopment of short-distance neural connectivity [25-27] and reduced long-distance neural connectivity [27-33]. Such irregular patterns of connectivity would be expected to contribute, among other things, to abnormal auditory perceptual processing in those with ASD [34]. Although the precise implications of these neurodevelopmental abnormalities for perception are not known, they might lead to less categorical perception of speech and musical sounds, more attention to less relevant sound features, a focus on local compared to global features, and less specialization for the particular language or musical system in one's native environment.

With respect to speech processing, research shows that those with ASD activate the middle and inferior temporal gyri bilaterally when listening to speech sounds, whereas controls show more left hemisphere activation [35]. Furthermore, listening to speech sounds produces activation outside of speech-specific areas, such as the brainstem, cerebellum, cingulum and posterior parietal that is not seen in controls [35]. Thus, those with ASD produce abnormal brain activation patterns that involve recruiting suboptimal neural networks for speech sounds. Other research shows that those with ASD show a reversal of the typical left-right brain size asymmetry for areas important for speech and language processing, including the left inferior frontal gyrus (or Broca’s area) and the posterior left superior temporal gyrus (or Wernicke’s area) [36]. Taken together, those with ASD appear to respond differently to speech sounds than controls. It may be that there is a lesser degree of differentiation in people with ASD between the neural pathways that they use to process speech versus environmental sounds, compared with typically developing individuals.

We have created a battery of tests to examine: (1) ability to filter out sounds that are irrelevant to a task and focus on those that are relevant, (2) sensitivity to phonemic categories relevant to the language spoken, (3) multisensory integration of auditory and visual information in speech, (4) propensity to use an absolute pitch code, (5) development of specialization for the metrical categories used in the musical system in the native environment, and (6) internalization of the rules of tonal harmony used in the musical system in the native environment. Here we measure each of these abilities in high-functioning adolescents with ASD in comparison to controls and examine whether there are correlations between these abilities that could reflect general auditory processing styles in ASD. The rationale for including each of these specific tests is outlined in the following paragraphs.

Test 1: The ability to filter out sounds that are irrelevant to a task and focus on those that are relevant is critical for being able to follow a conversation in a noisy environment, as most environments contain several objects emitting sounds that overlap in time. Questionnaire-based research on this topic suggests that those with ASD score high on items that tap into auditory filtering problems [37-41]. For example, in the Short Sensory Profile, parents tend to rate statements “doesn’t respond when name is called but you know the child’s hearing is okay” and “distracted or has trouble functioning if there is a lot of noise around” as describing their child with ASD [42]. Behavioral tasks [43,44], auditory cortical event-related potentials (ERPs), and auditory brainstem responses (ABRs), the latter two derived from electroencephalogram (EEG) recordings [45,46-48], indicate filtering problems in ASD. Those with ASD require a higher signal-to-noise ratio than controls in order to perceive speech in pink noise, noise from a competing talker or noise with the long-term spectral shape of speech [43,44]. Furthermore, those with ASD show less evidence of segregating incoming sounds into the auditory objects that compose them [46], and greater difficulty ignoring distracting sounds in peripheral spatial locations [45] compared to normal controls. In the present paper, we measure the ability to ignore one speech stream while attending to another, a task that adults need to perform virtually every day. Specifically, we measure the signal-to-noise ratio needed to perceive sentences presented to one ear while ignoring simultaneous sentences presented to the other ear.

Test 2: Efficient processing of speech relies on perceiving speech sounds according to the phonemic categories of the language spoken. Typically developing infants are able to discriminate between all possible speech contrasts at 6 months of age, but by 12 months, infants have already become specialized for categorical contrasts used in their native language and have difficulty discriminating contrasts used in foreign languages but not their native language [49-51]. Synaptic pruning appears to underlie perceptual specialization [52]. Given the evidence of abnormal neural connectivity in development in ASD, there is reason to suspect that phoneme perception may develop to be less language-specific in ASD than it is in controls. Interestingly, in typically developing infants, a context involving human social interaction is much more effective than an equal amount of exposure in a non-interactive context for phonemic learning to occur [53]. Thus, early social deficits in ASD might also be hypothesized to lead to poorer specialization for the native language. People with ASD also find faces less salient than do people without ASD [54-56], so infants who go on to develop ASD might have impoverished visual input during the process of learning native phoneme categories. The only experimental evidence of phoneme categorization in ASD comes from a study that examined foreign speech contrasts [57]. Although this study found that there were no group differences in performance, it cannot address the issue of specialization because it did not compare perception of foreign and native speech sound categories. Here we measure specialization by comparing perception of sounds from a foreign language that map onto one versus two sound categories in the native language.

Test 3: Typical listeners integrate information about speech from different sensory modalities. In particular, visual information from the eyes and mouth is combined with auditory information to produce a single percept of the sounds produced by a speaker. Experimental evidence of multisensory integration in speech comes from studies on the McGurk effect, in which participants are asked to report what they hear when presented with audio and visual inputs that are incongruent. For example, when presented with a visual “ga” and an audio “ba”, people report hearing a third percept “da” which represents a fusion between what they see and what they hear [58]. A few studies show that those with ASD tend to be less susceptible to this illusion than those who are typically developing, as evidenced by fewer fused responses [59-61]. In an experimental task where the auditory and visual information was not in conflict and so no third percept was produced, Smith and Bennetto [62] examined multisensory integration by comparing speech perception performance with the audio alone and with the audio and visual information together. The results showed that both groups performed better in the bimodal than unimodal condition, but the ASD group benefited less than the control group from the addition of visual information [62]. Additional research suggests that those with ASD are less accurate lip-readers than controls [60-62], although one study did not find any deficits in this area [59]. Critically, some of these studies indicate that these deficits might contribute to problems with the audio-visual integration of speech [60,61]. Here we compare those with ASD and controls on their integration of audio and visual information in speech when the information from these two modalities is in conflict. We additionally examine the relative contribution of lip-reading to this integration, or lack thereof, in ASD by including a visual only (lip-reading) condition.

Test 4: Pitch information is crucial for processing both prosody in speech and melody in music. In humans, pitch is processed in two basic ways. One way is to use a relative code where the pitch distances between tones are encoded, such that a melody is recognized regardless of the starting note or pitch register. By at least as young as 6 months of age, infants process pitch using this type of code [63-66]. The second way is to use an absolute code in which individual tones are recognized without relying on an external reference. The ability to name absolute pitches in isolation is extremely rare in adults, being found in less than 5 out of every 10,000 individuals [67,68]. Although sometimes considered a gift, absolute pitch may hinder melodic and prosodic perception because it focuses attention on single tones instead of on the entire melody, word or phrase. A number of studies indicate that absolute pitch processing may be more prevalent in ASD [69-72]. However, these studies primarily tested participants who were explicitly familiar with music reading and Western musical nomenclature, as the tasks required naming notes in Western notation. Here we measure the prevalence of absolute pitch processing in ASD using a task that does not require explicit knowledge of musical structure and can therefore be used in non-musicians with and without ASD.

Test 5: Just as exposure to a language early in development leads to perceptual processing specialized for that language, exposure to a musical system, such as Western tonality, results in specialized perceptual processing for that musical system [73,74]. Such specialization occurs for both the rhythm (meter) structure [75-77] and the pitch (tonal) structure [64,78], and this musical enculturation is essential for appreciation of the music in one's culture. Meter involves the perceptual extraction of an underlying pulse that can be broken down into different hierarchical beat patterns [75]. Simple meters involve strong and weak beat durations that form simple ratios (e.g., 2:1), whereas complex meters involve strong and weak beat durations that form complex ratios (e.g., 3:2) [73,79]. Typically developing 6-month-old infants are able to detect changes in simple meters that are common in Western music as well as changes in complex meters that are rare in Western music but common in Eastern European music [75]. Similar to the enculturation that occurs in language, infants exposed to Western music lose the ability to perceive complex meters in favour of simple meters that are common in their environment by 12 months of age [75,80]. For the same reasons that we suspect less specialization for native phonemic categories in individuals with ASD, namely early developmental differences in brain development and social interaction, we expect that metrical perception would be less Western-specific in those with ASD compared to controls. To our knowledge, no experimental data to date has addressed how those with ASD perceive metrical categories. Here we measure specialization by comparing perception of native and foreign metrical structures that were implemented in short musical sequences.

Test 6: As with meter, experience with Western music during development leads to perceptual specialization for the rules of tonal harmony in that musical system. Between 4 and 7 years of age, typically developing children have acquired some implicit knowledge of Western harmonic structure [74,78,81,82]. Chord sequences follow preference rules such that a sequence sets up expectations in enculturated listeners for which chords are likely to come next, and these expectations can be measured implicitly with reaction times [83-86]. In a typical implicit paradigm, chord sequences are presented, half of which end with expected chords and half with unexpected chords. Response times on an indirect task, which does not require judging the sequence itself, but rather just making a speeded judgment on the final chord (the target), such as a timbre identification task, are compared for sequences with expected and unexpected endings. For example, when the final chord of the sequence functions as the tonally related (supposed to be expected) tonic chord, response times on this chord were faster than when the final chord functions as the less-related (supposed to be less-expected) subdominant chord [82,83,85,86]. The construction of these experimental stimuli allows the conclusion that listeners have acquired knowledge about the regularities of the Western tonal system, which provides the basis to develop expectations for the final chord type (favoring the tonic over the subdominant). This cognitive priming interpretation contrasts with a sensory priming interpretation. Sensory priming would predict faster processing for the subdominant chord based on the advantages of repetition priming (the repeated presentation of the subdominant inside the sequence) which does not require tonal knowledge. One study asked participants with ASD to report aloud if chord sequences sounded complete or not [87]. Those with ASD were found to process chord sequences similarly to controls in this respect [87]. This study measured accuracy, but reaction times might be a more sensitive measure to reveal group differences. Here we examined harmonic priming by measuring both accuracy and reaction time, and by using an indirect task (i.e., participants were not asked to explicitly judge the sequences' endings, but were instead asked to quickly discriminate between two target timbres).

In sum, we have developed a battery of tests that measures auditory perceptual processing across the domains of speech (filtering, phoneme categorization and multisensory integration) and music (absolute pitch, meter categorization and harmonic priming). We expect that relative to those who are typically developing, those with ASD would focus more on surface details and less on relative or categorical aspects that are often most important. Our goal is to produce an auditory profile that characterizes high-functioning ASD that could help to inform remediation programs related to auditory processing in communication and social functioning.

Method

Participants

This research was approved by the McMaster University Research Ethics Board and conforms to the Canadian Tri-Council Policy Statement on Ethical Conduct for Research Involving Humans. Written informed consent was obtained from all participants as well as their parents. A total of 54 adolescent male participants (*M* age = 14.8 years, *range* = 11 to 18 years) were tested, 27 with ASD (15 Asperger's syndrome and 12 High-Functioning Autism) and 27 showing typical development (controls). Among the ASD group, 16 participants had diagnoses (ADOS and ADI) [88-89] that were completed at the Offord Centre in Hamilton and 11 confirmed through a letter from their family doctor (diagnosis outside of Hamilton) because ADOS and ADI scores were not available. None of the participants in the control group had a family member with ASD, but 11 out of 27 participants in the ASD group (41%) reported a family member with this disorder. All participants were monolingual English speakers, who had similar chronological age and years of musical experience (see Table 1). Among participants who had some musical experience (19 controls and 16 ASD), their experience was similar across the groups. The control group collectively represented 8 instruments (drums, guitar, piano, saxophone, trumpet, trombone, violin, and recorder), while the ASD group collectively represented 11 instruments (French horn, recorder, trumpet, cello, piano, guitar, harmonica, keyboard, glockenspiel, viola and drums). Participants in both groups were more likely to report learning how to play these instruments in the context of a music class at school or from a family member at home instead of through private lessons. Interestingly, the reported estimate of absolute pitch was higher in the ASD group (22%) than the control group (4%), despite the fact that none of the participants had been tested previously for this ability. With respect to family background, both groups reported on average having one sibling, although there was a range of 0 to 4 siblings per household. Finally, most participants identified themselves as being right-handed (74% controls and 85% ASD) rather than left-handed.

Table 1. *Demographic and background information by group*

	Group	
	Control	ASD
Age, in years		
Mean	14.6	15.0
SD	2.0	1.7
Music, in years		
Mean	2.7	2.3
SD	2.9	2.8
Forward Digits		
Mean	10.4	9.6
SD	2.3	2.1
Backward Digits		
Mean	5.7	5.6
SD	1.9	2.2
PPVT, standard		
Mean	112.3	107.2
SD	9.7	16.4
Leiter, standard		
Mean	106.4	99.4
SD	15.4	17.0

Note. Music = Years of Musical Experience; PPVT = Peabody Picture Vocabulary Test; Leiter = Leiter International Performance Scale

Procedure and Measures

All participants in the control group were tested at McMaster University. Participants with ASD were either tested at McMaster (*n* = 16) or in a quiet room in their own home (*n* = 11) if they lived outside Hamilton. All efforts were made to ensure consistency between testing locations, such that testing in the home was free from distractions. Participants were told that they would be asked to play a series of games using paper and pencil or a laptop computer (Acer Notebook) that was connected to a set of headphones (Sennheiser HAD 200). They were asked to perform to the best of their ability and were assured that they would receive practice trials before starting each game to ensure that they understood the instructions.

After obtaining informed consent, all participants were tested in the same order on the following tasks: Pitch Discrimination (based on [90]), Absolute Pitch (based on [63,91]), Harmonic Priming (based on [85]), Digit Span Subtest of the Wechsler Memory Scale-III [92], Peabody Picture Vocabulary Test-III [93], McGurk Auditory-Visual Integration Task (based on [58]), Phoneme Categorization (based on [49]), Metrical Categorization (based on [75]), Hearing Thresholds, Competing Sentences Test (based on [94]), and Leiter International Performance Scale [95]. We did not measure full scale intelligence as we were interested in particular skills, such as digit span (as some of our tasks had memory demands) and receptive vocabulary (given the linguistic components involved in our tasks). While participants completed these measures, caregivers were asked to complete a Background Information Form. Those with ASD took approximately

4 hours (2 x 2-hour sessions) to complete the auditory battery, whereas those in the control group took approximately 3 hours. Participants were compensated $10 for each hour of their time, and received a debriefing statement at the end of the session.

Background and Baseline Measures

Background Information Form. This parental report contained 16 questions across four areas: demographic information, language exposure, family background and musical training.

Leiter International Performance Scale [95]. This standardized test measures non-verbal intelligence through the use of visualization and reasoning. The four subscales took approximately 30 minutes to complete.

Peabody Picture Vocabulary Test-III [93]. This standardized test measures receptive vocabulary through the use of pictures and took approximately 25 minutes to complete.

Digit Span Subtest of the Wechsler Memory Scale-III [92]. The forward digit span portion of this standardized test measures short-term memory and involves repeating back sequences of 1 to 9 digits. The backward digit span portion measures working memory and involves repeating back sequences of 1 to 9 digits in the opposite order to that presented. This subtest took approximately 5 minutes to complete.

Hearing Thresholds. Thresholds were measured in the right and left ears at 500, 1000, 2000, 4000, and 8000 Hz. Each tone was first presented at 30 dB SPL and adjusted for intensity using a programmable attenuator [96]. Participants were instructed to raise their hand whenever they heard the tone. Following standard audiological assessment procedures, the signal was increased or decreased in amplitude by 2 dB from the previous trial depending on whether the participant was able to detect the tone on the previous trials. The stopping rule for each frequency was three consecutive missed trials. Threshold was measured as the intensity at which a tone for a particular frequency was detected 50 percent of the time. Normal hearing involves an absolute threshold between 0 dB and 20 dB (specifically, 9.5 dB at 500 Hz, 5.3 dB at 1000 Hz, 4.3 dB at 2000 Hz, 8.0 dB at 4000 Hz, and 18.7 dB at 8000 Hz) [97]. This test took approximately 10 minutes to complete.

Pitch Discrimination [90]. On each of 40 trials, two pure tones were presented separated by 1 sec using Presentation 11.0 [98]. The first tone was always 524 Hz while the second tone was higher or lower in pitch by .25 (8 Hz), .50 (15 Hz), 1.00 (30 Hz) or 2.00 (61 Hz) semitones. Participants were instructed to press the “up” arrow on the keyboard (“A” key) if the second tone was higher in pitch than the first tone and press the “down” arrow on the keyboard (“L” key) if the second tone was lower. The order of trials was randomized across participants. This task took approximately 10 minutes to complete.

Speech Perception Measures

1. Competing Sentences Test (based on [94]). This test measured the signal to noise ratio needed to repeat back a simple sentence (5 to 6 words) spoken by a male speaker in one ear while ignoring a semantically related sentence in the other ear. The distracting sentence (the "noise") was presented at 50 dB above each individual's threshold at 1000 Hz. A response was counted as correct if it contained at least two words from the target sentence and no words from the distracting sentence. The test was programmed using Microsoft Visual Basic and presented on the laptop computer, connected to a programmable attenuator [96]. The signal was initially presented to all participants at 30 dB above their hearing threshold for 1000 Hz. A Bayesian adaptive psychometric procedure [99] was used such that the test ended when the standard deviation of the signal threshold estimate reached 1.5 dB or less. Performance was indicated by the signal to noise ratio (SNR) (signal in dB – noise in dB) at which performance was 50 percent correct. This test took approximately 15 minutes to complete.

2. Phoneme Categorization (based on [49]). On each trial, participants heard three 300 msec phonemes in an ABB or AAB format and determined whether the first or last phoneme was different from the other two. All of the speech sounds were from a South African language (Zulu) and spoken by the same female native speaker. Two sets of two phoneme categories were used such that one contrast (24 trials) mapped onto distinct phonemic categories in English (specifically, voiced lateral fricative vs. voiceless lateral fricative) and should therefore be easy for English speakers to discriminate, whereas the other contrast (24 trials) mapped onto the same phonemic category in English (specifically, plosive bilabial stop vs. implosive bilabial stop) and should therefore be more difficult for English speakers to discriminate. On each trial, for each of the three phonemes, one of 6 possible tokens for a given phoneme (matched in duration and fundamental frequency) was chosen randomly with the constraint that the same token could not be used twice in the same trial. The measure of interest was the comparison between accuracy on the one- and two-category mapping conditions. Before starting each test condition, participants received 6 practice trials where they were asked to discriminate between contrasts in the one- and the two-category mapping conditions. We used stimuli from both test blocks in the experimental phase to ensure that participants in the practice phase understood the task instructions. Participants were told to respond as accurately as they could. The task was programmed in Presentation 11.0 [98]. The order of stimulus presentation was randomized across participants. This task took approximately 15 minutes to complete.

3. McGurk Task (based on [58]). This computerized task measured the audio-visual integration of speech. On each trial, participants heard and/or saw a face making mouth movements for a consonant-vowel pair (“ba”, “ga”, or “da”) that was produced 6 times at a normal speaking rate. There were three types of audiovisual trials: matched auditory “ba” plus visual “ba” (12 trials); matched auditory “ga” plus visual “ga” (12 trials); and mismatched auditory “ba” plus visual “ga” (24 trials). There were four types of single modality control trials presented after the audiovisual trials: auditory only “ba”

(6 trials), auditory only “ga” (6 trials), visual only “ba” (6 trials), and visual only “ga” (6 trials). There were five tokens each of “ba” and “ga”. On each trial, one token was chosen randomly. Participants were asked to indicate what they heard (or saw in the case of visual only trials) by pressing “1” if they heard “ba”, “2” if they heard “ga”, and “3” if they heard “da”. If participants were integrating the audio-visual information in the mismatched trials then they should report hearing “da” (McGurk illusion). The measures of interest were the susceptibility to the McGurk illusion and the relative contribution of lip-reading to multisensory integration. Participants were instructed not to stop looking at the face in the video until she stopped talking, and their behavior was monitored in this regard throughout the task. The task was programmed in Presentation 11.0 [98]. The order of stimulus presentation was randomized across participants. This task took approximately 20 minutes to complete.

Music Perception Measures

1. Absolute Pitch (based on [63,91]). In the initial control condition, on each trial participants heard a 500 msec piano tone, followed by 16 seconds of silence, followed by a second tone that was either at the same pitch (6 trials) or one semitone higher (3 trials) or lower (3 trials), and indicated whether the two pitches were the same or different. The experimental condition was identical except that the silent period contained a 500 msec pause, then 15 interfering tones (randomly chosen on each trial but ranged within an octave such as A2/110 Hz to F3/175 Hz), followed by a pause of 8 seconds and then the final tone. The same random order of trials was used for each participant. The stimuli were presented in Windows Media Player 10.0. A perfect score on the experimental condition indicated absolute pitch processing. This task took approximately 15 minutes to complete.

2. Metrical Perception. Using the stimuli of [75], on each trial participants were first familiarized for 15 seconds with a melody based on traditional Eastern European folk music that either (4 trials) had a Western-typical simple meter (8 note measure subdivided into 2+2+2+2 250 msec beats) or (4 trials) had a Western-atypical complex meter (7 note measure subdivided into 3+2+2 250 msec beats). Following each familiarization, a 30 second test melody was presented that contained an extra note that either preserved the metrical structure of 8 or 7 beats, or added one beat to it, transforming the simple meter into a complex 9-beat meter or the complex meter into a simple 8-beat meter. Before starting the test condition, participants received 5 practice trials using a familiar melody ("Mary Had a Little Lamb") to ensure that they understood the instructions. Participants were asked to indicate if the two melodies had the same beat or not. Using a response box, participants pressed "1" if the two melodies were "very well" matched, "2" if the two melodies were "somewhat well" matched, "3" if the two melodies were "somewhat poorly" matched, and "4" if the two melodies were "very poorly" matched. The task was programmed in Presentation 11.0 [98]. The order of stimulus presentation was randomized across participants. This task took approximately 25 minutes to complete.

3. Harmonic Priming (from [85]). The task was programmed in Presentation 11.0

[98]. In this implicit task participants heard an eight-chord sequence (with the first seven chords, each sounding for 620 msec, defining the prime context, played in piano) on each trial, and indicated whether the 8th chord (the target, duration of 2000 msec) was played in "piano" or "harp" timbre. Unlike in the original experiment by [85], we changed the way that participants made their response by using the terms "piano" and "harp" instead of Timbre A and Timbre B. Importantly, in half the chord sequences, the target chord functioned as the tonic chord (preceded by the dominant chord; 12 trials), whereas in half, it functioned as the subdominant chord (preceded by the tonic chord; 12 trials)[1]. If participants process these chords according to the regularities of Western tonal music, faster response times are predicted for the tonic targets, which are supposed to be the appropriate, expected ending, than for the subdominant targets, which are also part of the tonality, but are less expected (Figure 1). If participants do not show this cognitive priming effect, they might be influenced by sensory (repetition) priming, leading to faster processing of the subdominant targets, which also occurred in the prime context, than of the tonic targets, which did not occur in the prime context. Note that both predictions require processing the sequences globally as the last two chords are kept constant between the two conditions across the sequence set. Participants initially received 12 practice trials involving single chords to ensure that they understood the instructions of the speeded timbre discrimination task. They were instructed to respond as fast and as

---

[1] There are seven chords that define a key in Western tonal music. Each of these chords is based on a different degree of the scale that forms a hierarchy of stability, depending on the currently installed key. The most stable chord is the tonic (I), and all other chords are perceived in relation to this chord. The next most stable chord is the dominant (V), followed by the subdominant (IV) chord, etc.

accurately as possible and both speed and accuracy were analyzed. The order of stimulus presentation was randomized. This task took approximately 15 minutes to complete.

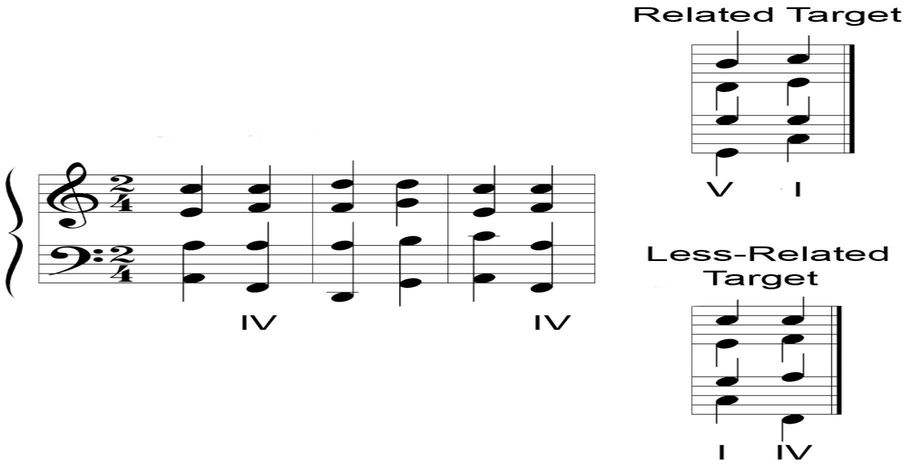

Figure 1. *Example stimuli for Harmonic Priming Task* (from [85]). The 6 chords of the prime context are shown on the left. These are followed by either a dominant (V) to tonic (I, expected) progression or a tonic (I) to subdominant (IV, less expected) progression. Note that the target chord repeats in the prime context for the less-expected subdominant target chord (IV chord), but not for the expected tonic target chord (I chord), ruling out sensory priming explanations of observed processing differences (adapted from [130], Figure 2).

Results

Background and Baseline Measures

The groups were matched in chronological age, (*p* = .43, *M dif* = -.407, 95% CI [-1.435, .620]), in non-verbal intelligence, (*p* = .12, *M dif* = 6.963, 95% CI [-1.884, 15.809]), in receptive vocabulary, (*p* = .18, *M dif* = 5.074, 95% CI [-2.344, 12.492]), and in musical experience, (p = .60, *M dif* = .407, 95% CI [-1.128, 1.943]) (Table 1, see page 39).

Digit Span. An ANOVA revealed a main effect of condition, $F(1, 52) = 243.04$, $p < .001$, $\eta^2 = .82$ (*M dif* = 4.296, 95% CI [3.743, 4.849]), with better performance for forward than backward digit span, but no main effect of group (p = .39, *M dif* = .444, 95% CI [-.591, 1.480]), or interaction involving group (p = .23) (Table 1, see page 39). In sum, there were no significant differences between control and ASD groups in short-term or working memory for digits.

Hearing Thresholds. An ANOVA conducted on hearing thresholds revealed a main effect of condition, $F(1, 52) = 12.24$, $p = .001$, $\eta^2 = .20$ (*M dif* = .98, 95% CI [.416, 1.54]), with better performance in the right than the left ear, and a main effect of frequency, $F(4, 52) = 14.90$, $p < .001$, $\eta^2 = .23$ with better performance for lower than higher frequencies tested. However, there was no main effect of group and no interactions (all $ps > .63$). In sum, there were no significant differences in hearing thresholds between control and ASD groups.

Pitch Discrimination. An ANOVA conducted on accuracy scores revealed a main effect of direction, $F(1, 52) = 6.45$, $p = .01$, $\eta^2 = .11$ (*M dif* = .29, 95% CI [.06, .52]), with better performance for rising than falling pitch, and a main effect of size, $F(1, 52) = 45.24$, $p < .001$, $\eta^2 = .47$ (*M dif* = .83, 95% CI [.58, 1.07]), with better performance for large than small pitch changes, but no main effect of group or interactions involving group (all $ps > .05$). In sum, control and ASD groups exhibited similar pitch discrimination thresholds.

Speech Perception Measures

1. Competing Sentences Task. An ANOVA conducted on mean signal to noise ratio (SNR) (signal in dB – noise in dB) revealed a main effect of ear, $F(1, 52) = 21.01$, $p < .001$, $\eta^2 = .31$ ($M\ dif = 2.23$, 95% CI [1.25, 3.21]; Cohen's $d = .84$[2]), with better performance when the target was presented to the right than the left ear, and a main effect of group, $F(1, 52) = 52.88$, $p = .000$, $\eta^2 = .54$ ($M\ dif = 8.31$, 95% CI [6.01, 10.61]; $d = 1.95$), with better performance for the control than ASD group (Figure 2). There was no significant interaction between ear and group ($p = .47$). In sum, those with ASD perform worse than controls when required to filter a spatially segregated stream of information.[3]

---

[2] Guideline for interpreting effect sizes for Cohen's d: $0 < d < 0.2$ (small), $.2 < d < .8$ (medium), and $d > .8$ (large). *Source*: Essentials of Statistics for the Behavioral Sciences (4th edition) by F Gravetter and LB Wallnau (page 235).

[3] To determine whether this performance difference for the competing sentences task is explained by receptive vocabulary, we performed an ANCOVA on mean SNR, with receptive vocabulary as our covariate. The ANCOVA revealed a main effect of group, $F(1, 51) = 47.46$, $p = .000$, $\eta^2 = .51$, but not of ear, $F(1, 51) = 2.95$, $p = .09$, and no main effect of receptive vocabulary ($p = .37$). There was a significant interaction between ear and receptive vocabulary, $F(1, 51) = 5.44$, $p = .024$, $\eta^2 = .11$, with better performance when the target sentence was delivered to the right than left ear, but no significant interaction between ear and group, $F(1, 51) = 1.56$, $p = .22$.

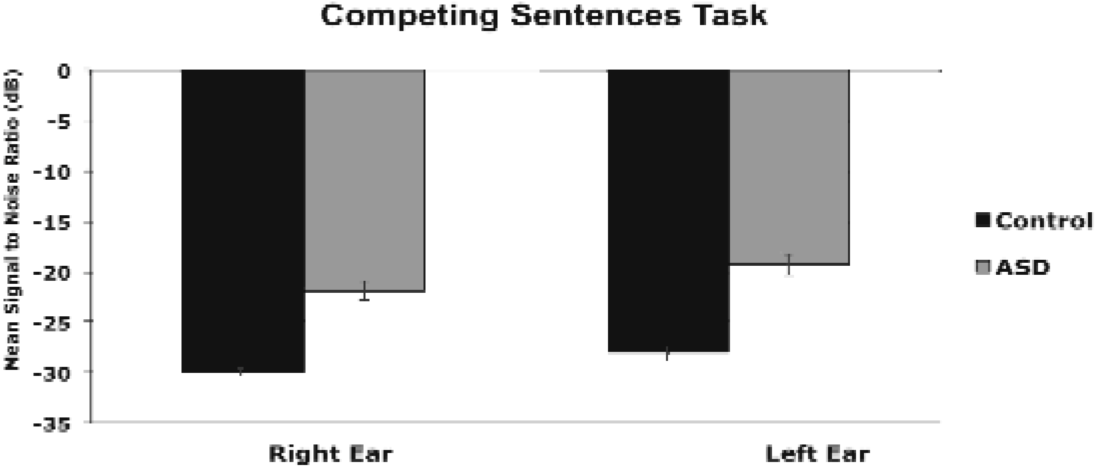

Figure 2. *Mean signal to noise ratios for the Competing Sentences Task by group.* The signal-to-noise ratio needed to detect the sentence in one ear in the presence of a competing sentence in the other ear are shown on the y-axis. Those with ASD performed significantly worse than controls as evidenced by higher signal to noise ratios.

2. Phoneme Categorization. An ANOVA conducted on accuracy scores revealed a main effect of condition, $F(1, 52) = 138.02$, $p < .001$, $\eta^2 = .73$ (*M dif* = 4.50, 95% CI [3.73, 5.27]; $d = 2.85$), with better performance for two-category mapping (native categories) than one-category mapping (foreign categories). The main effect of group was not significant ($p = .58$, (*M dif* = .50, 95% CI [-1.28, 2.28]; $d = 0.12$), but there was an interaction between condition and group, $F(1, 52) = 4.73$, $p = .03$, $\eta^2 = .083$ (Figure 3) such that those with ASD showed a smaller difference between the two-category and one-category mapping conditions than did controls. Simple main effects using independent samples t-tests revealed no significant difference in performance between groups in the native categories [$t\ (52) = 1.56$, $p = .13$, $d = .43$] or foreign categories [$t < 1$, *d* = .09] conditions.[4] In sum, those with ASD showed less specialization for native speech sound categories than controls.

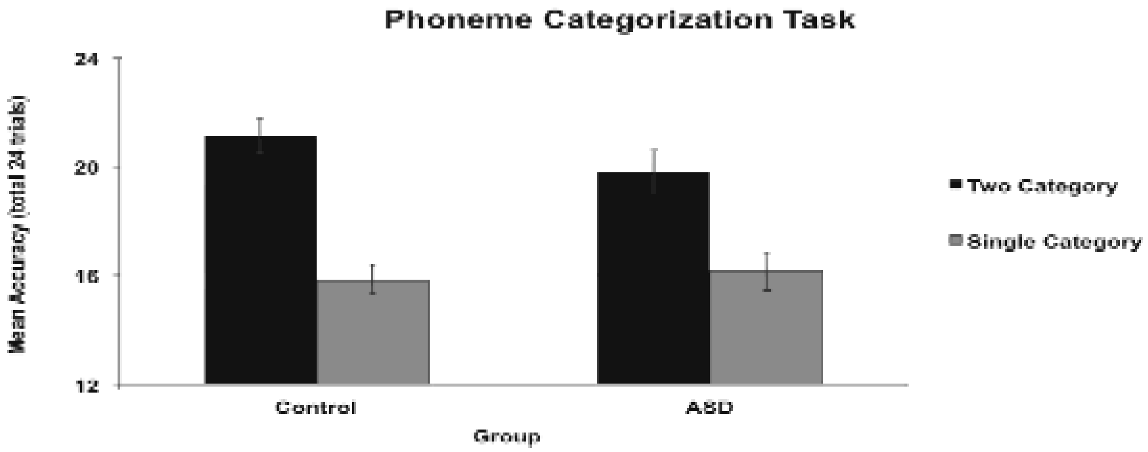

Figure 3. *Mean accuracy for the Phoneme Categorization Task by group.* In a 3-interval forced choice design, subjects heard three phonemes that fell into two categories in either the pattern ABB or AAB and had to determine whether the middle sound was most similar to the first or last sound. The number correct out of 24 is shown on the y-axis for the cases where the speech sounds fell into one or into two phonemic categories in the native language. Error bars represent standard error of the mean. Group differences were found such that those with ASD showed a significantly smaller difference than controls between the two-category and one-category mapping conditions.

---

[4] To determine how much of this performance difference for the phoneme categorization task is explained by receptive vocabulary, given that the task involved verbal instructions, we performed an ANCOVA on mean accuracy scores, with receptive vocabulary as our covariate. The ANCOVA revealed no main effect of condition, *F*(1, 51) = 2.70, *p* = .11 (*M dif* = 4.50, 95% CI [3.72, 5.28]), no significance for the main effects of group (*p* = .71, (*M dif* = .34, 95% CI [-1.48, 2.15]) and receptive vocabulary (*p* = .34), and no significant interaction between condition and receptive vocabulary (*p* = .81). However, the important interaction between condition and group remained significant, *F*(1, 51) = 4.63, *p* = .04, η^2 = .084.

3. McGurk Task (Matched Trials). Performance on matched audiovisual trials was close to or at ceiling for both groups (Figure 4, upper panel, see page 55). For BA trials, performance was 97.5% correct for the control group and 98.4% correct for the ASD group. Planned independent samples t-tests revealed no significant difference in performance between groups in *ba* responses [$t < 1$, $d = .12$], *ga* responses [$t < 1$, $d = .18$] or *da* responses [$t < 1$, $d = .10$]. For GA trials, performance was 100% correct for both groups. Thus, planned independent samples t-tests could not be performed on these trials because the standard deviation was equal to 0 for both groups. In sum, there were no significant differences between groups in the matched trials.

McGurk Task (Mismatched Trials). Performance was 19.9% correct (i.e., the response was *ba* when presented with the auditory /ba/ and visual /ga/) in the control group and 44.5% correct in the ASD group. Planned independent samples t-tests revealed a significant difference in performance between groups for *ba* responses [$t\ (52) = 2.89$, $p = .006$, $d = .79$] as well as for *da* responses [$t\ (52) = 3.01$, $p = .004$, $d = .82$], although not for *ga* responses [$t < 1$, $d = .03$]. Those with ASD were less likely to integrate audio-visual speech sounds (i.e., less likely to experience the McGurk illusion) as evidenced by fewer *da* responses and more *ba* responses than controls.

McGurk Task (Audio Trials). Performance was high for both groups on auditory alone trials. For BA trials, overall performance was 95.0% correct for the control group and 97.5% correct for the ASD group. Planned independent samples t-tests revealed no significant differences in performance between groups for *ba* responses [$t < 1$, $d = .26$], *da* responses [$t < 1$, $d = .20$], or *ga* responses [$t\ (52) = 1.00$, $p = .32$ $d = .30$]. For GA trials, overall performance was 100% correct for the control group and 98.2% correct for the ASD group, precluding performance of t-tests on *ba*, *da* and *ga* responses, but indicating very high performance. In sum, performance was at or near ceiling for both groups and there were no measurable significant differences between groups in performance for the audio only trials.

McGurk Task (Visual Trials). For visual only BA trials, performance was 100% correct for the control group and 92.7% correct for the ASD group. Thus planned independent samples t-test could not be performed on *ba, da and ga* responses, but performance was at or close to ceiling for both groups. For GA trials (more difficult task than BA trials because the place of articulation is at the back of the mouth for /ga/) the groups did not differ in the number of correct (*ga*) responses, 67.8% for the control group and 65.5% for the ASD group [$t < 1$, $d = .08$]. Furthermore, the groups did not differ significantly in the number of correct *da* responses [$t < 1$, $d = .20$] although there was a significant difference in the number of *ba* responses [$t\ (52) = 2.95$, $p = .005$, $d = .80$]. Thus, performance was similar across groups in lip reading, although the distribution of errors differed somewhat in the case of GA trials.

In sum, the McGurk Task results indicate that those with ASD were less susceptible to the McGurk illusion, but no group differences were found for auditory alone or visual alone (lip reading) speech sound discrimination.

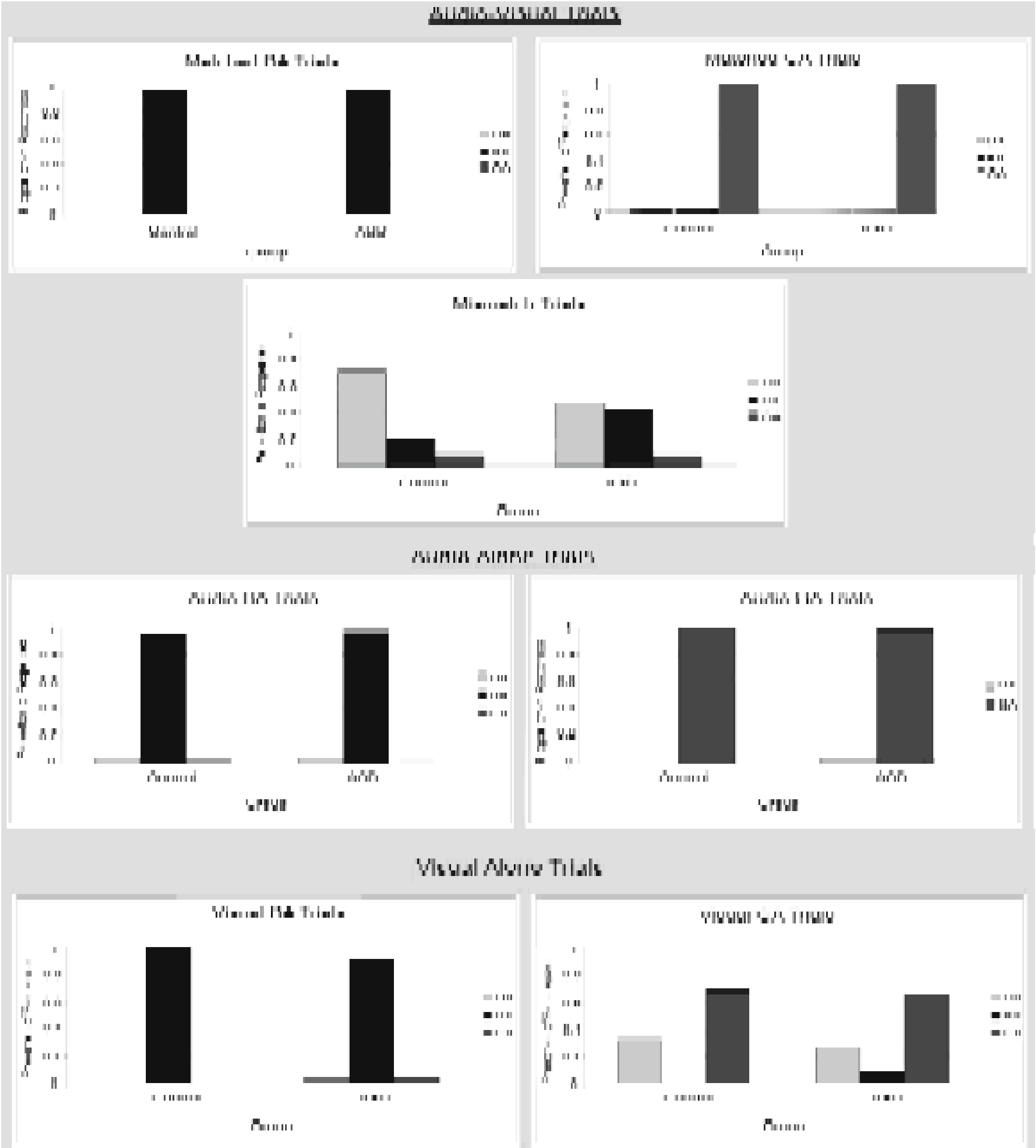

Figure 4. *Performance in the McGurk Task by group.* Proportion of “ba”, “da” and “ga” responses are shown for each stimulus type. Audio-Visual Trials. No group differences were found for the matched audio-visual trials. However, group differences were found for the mismatched audio-visual trials with those with ASD being less likely than controls to report hearing “da” (McGurk illusion) than “ba”. Audio Alone Trials. No group differences were found for audio-alone trials. Visual Alone Trials. No group differences were found for visual-alone trials.

Music Perception Measures

1. Absolute Pitch Test. Three participants in the ASD group (3 out of 27 or 11%), but none in the control group (0 out of 27 or 0%), showed perfect performance in this task, indicating absolute pitch processing. To determine whether the ASD sample differed from the normal population, we used the binomial distribution and set the probability of absolute pitch to 5/10,000 [67,68]. The probability of obtaining 3 or more individuals with absolute pitch from a sample of 27 given $p = 5/10000$ is .0000004. We can therefore robustly reject the null hypothesis that our ASD sample was drawn from the normal population (Figure 5, see next page over). Next, we conducted an ANOVA to determine if there was a significant difference in pitch memory between controls and those with ASD who did not demonstrate absolute pitch (i.e., $n = 24$ ASD, $n = 27$ controls). The results revealed a main effect of condition, $F(1, 49) = 142.82$, $p < .001$, $\eta^2 = .75$ (*M dif* = 3.96, 95% CI [3.30, 4.63]; $d = 4.84$), with worse performance in the presence of interference tones, but no main effect of group ($p = .50$, (*M dif* = .27, 95% CI [-.53, 1.07]; $d = 0.10$), or interaction ($p = .91$) (Figure 6, see next page over). Together, these results indicate that the prevalence of absolute pitch is higher among those with ASD than in the normal population. However, when those with absolute pitch were removed from the sample, no difference in pitch memory between the ASD and control groups was apparent.

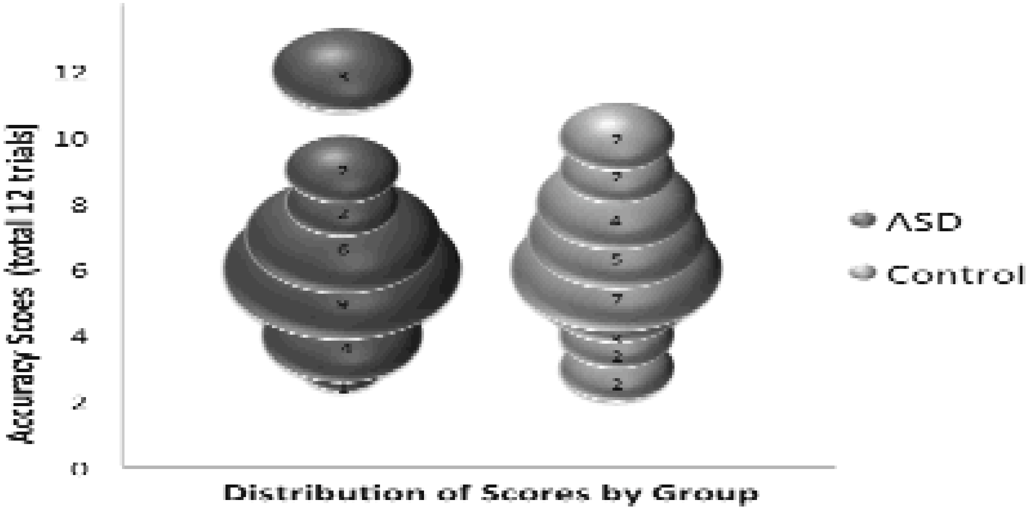

Figure 5. *Distribution of absolute pitch scores by group.* The size of the bubbles (and the number in each bubble) indicate the number of subjects who obtained each score (number correct out of 12 trials). The column of bubbles on the left represents the ASD data and the column on the right the control data. Three participants with ASD but no controls showed perfect performance, indicative of absolute pitch processing.

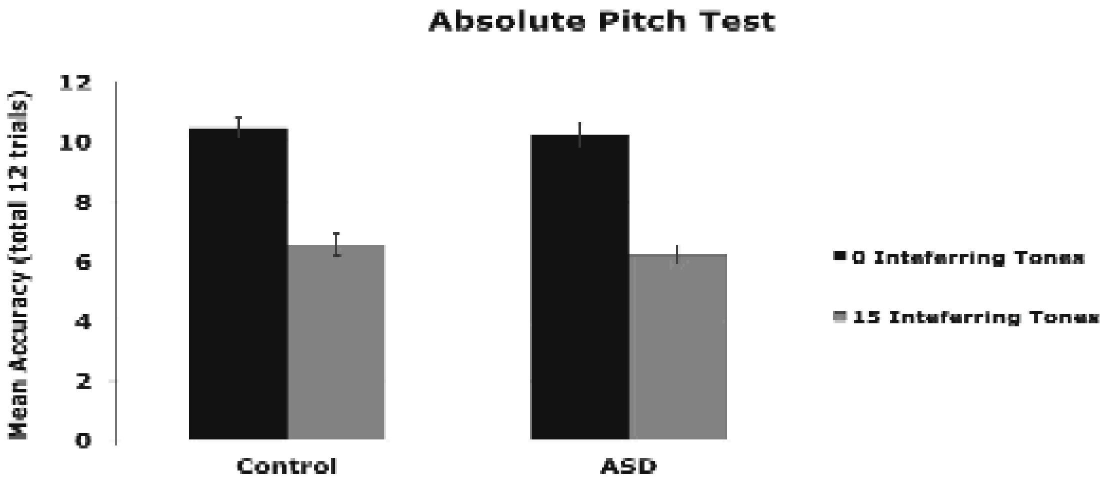

Figure 6. *Mean accuracy for the Absolute Pitch Test by group.* Mean number correct out of 12 for determining whether two tones had the same or different pitches when there were 0 or 15 interference tones, after removing the three ASD subjects with perfect scores in the 15-tone interference condition indicative of perfect pitch. Error bars represent standard error of the mean. There were no significant differences in performance between groups after removing these participants.

2. *Meter Perception.* An ANOVA conducted on accuracy scores revealed a main effect of condition, $F(1, 52) = 40.23$, $p < .001$, $\eta^2 = .44$ (*M dif* = 1.70, 95% CI [1.17, 2.24]; d =2.61), with better performance for simple (native) than complex meter (foreign) meters, but no main effect of group ($F < 1$, *M dif* = .11, 95% CI [-.47, .69]; $d = 0.07$). There was, however, an interaction between condition and group, $F(1, 52) = 4.28$, $p = .04$, $\eta^2 = .08$ (Figure 7), such that those with ASD showed a smaller performance difference between simple and complex meter conditions than controls. Simple main effects using independent samples t-tests revealed no significant difference in performance between groups in the simple meter [t (52) = 1.13, $p = .27$, $d = .31$] or complex meter [t (52) = 1.69, $p = .10$, $d = .46$] conditions. In sum, those with ASD showed less specialization for simple meters than controls.[5] Together, these results indicate that even after accounting for receptive vocabulary, those with ASD still show a smaller difference in performance between simple (native) and complex (foreign) meter compared to controls.

---

[5] To determine how much of this performance difference for the meter categorization task is explained by receptive vocabulary, given that the task involved verbal instructions, we performed an ANCOVA on mean accuracy scores, with receptive vocabulary as our covariate. The ANCOVA revealed no significant effects of condition ($p = .62$, *M dif* = 1.70, 95% CI [1.16, 2.25]), group ($p = .53$, *M dif* = .18, 95% CI [-.40, .77]), or receptive vocabulary, $F(1, 51) = 1.75$, $p = .19$, and the interaction between condition and receptive vocabulary was also not significant. Importantly, the interaction between condition and group remained significant, $F(1, 51) = 3.86$, $p = .05$, $\eta^2 = .070$.

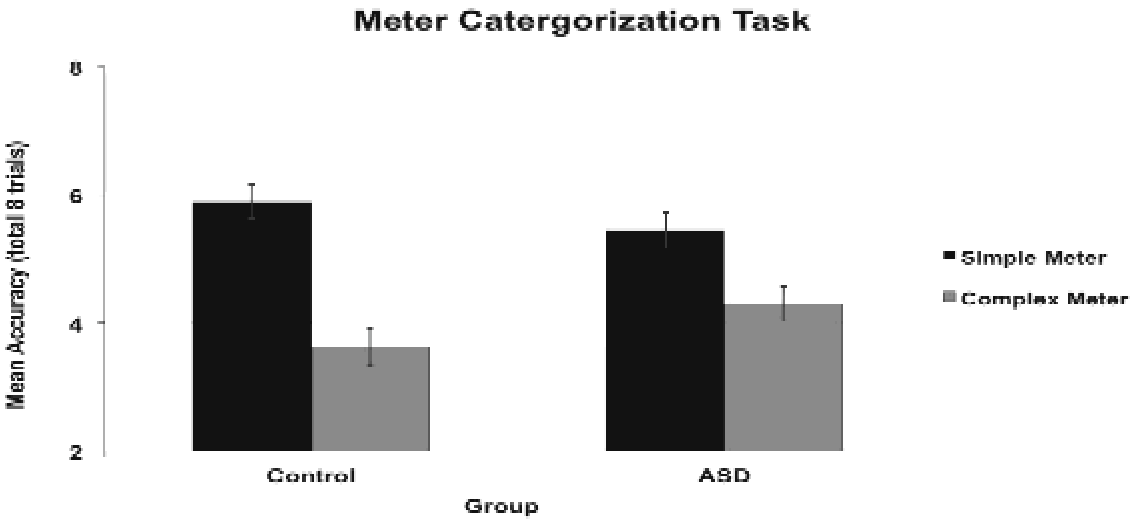

Figure 7. *Mean accuracy for the Meter Categorization Task by group.* Number correct out of 8 trials is shown on the y-axis. On each trial, it was to be determined whether an excerpt had the same or a different meter compared to a standard excerpt. Error bars represent standard error of the mean. Group differences were found such that those with ASD showed a significantly smaller difference than controls between the simple meter (typical in Western music) and complex meter (rare in Western music) conditions.

3. Harmonic Priming Task. Due to technical problems, the data are missing for 1 participant in the control group and 5 participants in the ASD group. To ensure that the groups were matched on accuracy, we conducted a 2x2x2 ANOVA on accuracy scores with harmonic target type and timbre as within-participant factors and group as a between-participants factor. The only significant effect was the main effect of timbre, $F(1, 45) = 10.62$, $p = .002$, $\eta^2 = .19$, $M\ dif = .26$, 95% CI [.10, .42]; $d = .66$, with both groups performing more accurately for piano chords than for harp chords (Figure 8a, see page 62).

When calculating the response time performance for each participant, only correct responses were included. Response times that were less than 250 msec or greater than 2,500 msec were excluded from the analyses, which are considered to be conservative cutoffs for outliers for reaction time data [100]. These reaction times occurred infrequently (13 out of 2,256 responses) and accounted for less than 1% of total responses. Additionally, one participant with ASD was removed from the final sample because of mean reaction times that were 4 standard deviations slower than the group means.

A 2x2x2 ANOVA conducted on correct response times revealed a main effect of harmonic target type, $F(1, 45) = 9.68$, $p = .003$, $\eta^2 = .18$, *M dif* = 21.42, 95% CI [7.55, 35.28]; $d = .36$, with faster performance for expected tonic chords than unexpected subdominant chords, and a main effect of timbre, $F(1, 45) = 79.242$, $p < .001$, $\eta^2 = .64$, *M dif* = 101.45, 95% CI [78.49, 124.40]; $d = 1.69$, with faster performance for harp than piano chords, but no main effect of group ($F < 1$). There were also no interactions involving group ($p > .21$; Figure 8b, see page 62). There was an interaction between timbre and harmonic target type, $F(1, 45) = 3.95$, $p = .05$, $\eta^2 = .08$. Thus two 2x2 ANOVAs were conducted by group on reaction times for piano and harp endings separately. The ANOVA on piano endings revealed a main effect of harmonic target type, $F(1, 45) = 11.86$, $p = .001$, $\eta^2 = .21$ *M dif* = 36.56, 95% CI [15.18, 57.95]; $d = .58$, with faster performance for expected targets than unexpected target chords, but no main effect of group ($F < 1$, *M dif* = 2.23, 95% CI [-77.45, 81.90]; $d = .02$), and no interaction ($F < 1$). An ANOVA conducted on reaction times for harp endings revealed no significant main effects or interactions (all $F < 1$). Thus, the harmonic priming effect occurred only for piano endings. This pattern of results was also found previously [85] and may reflect the fact that the change to harp timbre is very salient such that processing the timbre change occurs faster than processing of the harmonic information, so no (or less) effect of chord type (tonic/subdominant) is typically seen. It should be noted that the two groups in our sample were similar in showing the priming effect for expected harmonic endings when sequences (composed of piano chords) ended with a piano chord but not when they ended with a harp chord. In sum, the piano chords produced harmonic priming with faster reaction times for the expected than unexpected chord endings, but this performance was the same across control and ASD groups.

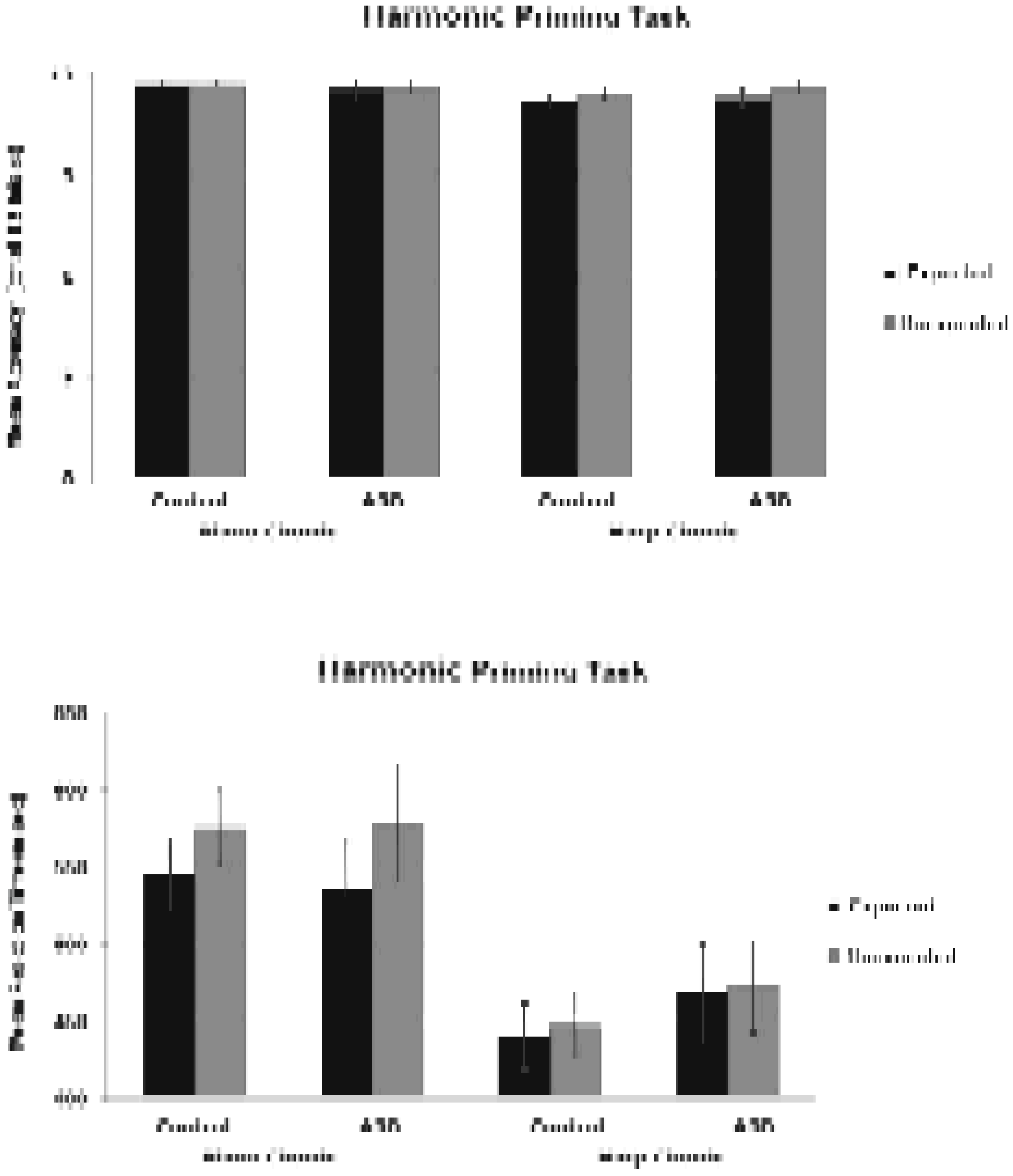

Figure 8. *Performance in the Harmonic Priming Task by group.* In this implicit task, subjects determined whether the last chord in a sequence was in piano or harp timbre. A. Performance on the 12 trials was very high for chords that were expected and for chords that were not expected, with no significant group differences. B. No significant group differences were found in reaction time performance. Both groups responded faster to the expected tonic target chords than to the less expected subdominant target chords.

Correlations Between Tasks

We examined how receptive vocabulary, non-verbal intelligence, and processing small pitch changes were related to the tasks in the battery using Pearson correlations across our entire sample ($N = 54$). We found that receptive vocabulary was only related to performance on the McGurk task. Specifically, it was related to visual alone BA trials ($r = .40$, $p = .003$) and to *ga responses* on mismatched trials ($r = -.34$, $p = .012$). Non-verbal intelligence was significantly related to simple meter processing ($r = .29$, $p = .03$), processing small pitch changes ($r = .38$, $p = .005$), and lip-reading for the visual alone GA trials ($r = -.30$, $p = .03$). Pitch change processing showed the highest number correlations with other battery tasks. Specifically, being able to detect small pitch changes was related to native speech sound processing ($r = .29$, $p = .03$), simple meter processing ($r = .29$, $p = .03$), absolute pitch processing ($r = .36$, $p = .009$) and working memory for digits ($r = .37$, $p = .007$). Detecting small pitch changes was also related to reaction times for expected ($r = -.51$, $p = .000$) and unexpected ($r = -.31$, $p = .03$) piano chord endings, and to reaction times for expected ($r = -.41$, $p = .004$) and unexpected ($r = -.55$, $p = .000$) harp chord endings. In sum, these correlations show that pitch processing in particular may underlie performance on several of the tasks in the auditory battery.

We also performed Pearson correlations on the same variables using only data from the ASD group ($n = 27$). Here we found a similar pattern of results with the exception that the correlation between receptive vocabulary and responding “ga” on mismatched trials was no longer significant (*r* = -.29, *p* = .14), along with the correlation between detecting small pitch changes and reaction times for unexpected piano chord endings (*r* = -.34, *p* = .14). However, receptive vocabulary was still related to visual alone BA trials (*r* = .42, *p* = .03) for the ASD group. Non-verbal intelligence was also significantly related to simple meter processing (*r* = .39, *p* = .04), processing small pitch changes (*r* = .41, *p* = .04), and lip-reading in the visual alone GA trials (*r* = -.50, *p* = .007). Being able to detect small pitch changes was related to native speech sound processing (*r* = .46, *p* = .02), simple meter processing (*r* = .39, *p* = .05), absolute pitch processing (*r* = .49, *p* = .01) and working memory for digits (*r* = .42, *p* = .03). Finally, detecting small pitch changes was related to reaction times for expected piano chord endings (*r* = -.60, *p* = .004), and for reaction times for expected (*r* = -.45, *p* = .04) and unexpected (*r* = -.59, *p* = .005) harp chord endings. In sum, a similar pattern of correlations was found for the ASD group as was found for the entire sample.

Discussion

Relative to typically developing adolescents, we found that adolescents with ASD were impaired on some auditory tasks but not on others, forming a profile by which we can further our understanding of auditory processing in this disorder. In general, the two groups were similar in terms of thresholds for sound detection, short-term memory, working memory, receptive vocabulary and non-verbal intelligence. However, compared to controls, the ASD group showed evidence of filtering problems (our competing sentences task showed that filtering problems persist at the level of speech sentence processing), less integration of auditory and visual information in speech, less enculturation to the phonemic categories of their language, and less enculturation to the metrical categories of the musical system in their environment. In general, those with ASD tended to be more impaired on tasks involving speech than on tasks involving musical sounds. Interestingly, with respect to music, although those with ASD showed less metric enculturation than controls, both groups showed similar enculturation to the harmonic pitch structure of Western tonal music. Although it is possible that the group difference on the meter enculturation task reflects difficulties in understanding the explicit task requirements, this is unlikely as the groups did not differ on receptive vocabulary scores, and no significant relation was found between performance on the explicit meter task and receptive vocabulary.

As discussed in the introduction, neural development in ASD appears to be particularly disrupted early in development, with early accelerated brain growth and disrupted patterns of neuronal connectivity [22,23,25-27,30-32]. Our results are generally consistent with the idea that skills acquired early in development are more disrupted in ASD. Efficient processing of speech relies on perceiving speech sounds according to the phonemic categories of the language spoken. By 12 months of age, normally developing infants, like adults, have become specialized for the language in their environment, and they have difficulty discriminating foreign phonemic categories that map onto a single category in their native language [49-51]. Interestingly, ours was the first study to show less specialization for native-language phonemic categories in adolescents with ASD compared to typically developing controls, and this difference between groups persisted even after accounting for individual differences in receptive vocabulary. These results suggest that native language learning may develop more slowly because perception is less constrained among those with ASD than controls. This finding is also consistent with the idea that those with ASD focus on low-level characteristics of sounds whether or not they are relevant to the task. We did not have access to whether individuals with ASD in our study showed early language delay or not, but it would be interesting for future studies to examine whether the development of native phonemic categories is affected by whether or not language delay is present. Interestingly, early social communication may also play a role in the diminution of specialization for native phonemic categories in ASD. In one study, the ability to discriminate foreign speech categories in infancy was maintained when infants interacted with a live person speaking that foreign language, but not when infants were exposed to recordings of that language [53]. Thus, there may be multiple reasons for less phonemic specialization in ASD. In any case, these results suggest that very early remediation may be needed in order to promote development of optimal speech circuits for language in ASD.

Similar to the acquisition of sensitivity to one's native phonemic categories, specialization for the metrical rhythm structure of the music system in one's environment is also seen by 12 months of age. We found, in this first study of metrical enculturation in ASD, that those with ASD were less specialized than controls for processing rhythms with simple meters typical of Western music compared to complex meters. This difference could not be explained by amount of musical experience in terms of formal music training as the groups did not differ on this variable. Interestingly, poor socialization early in development may also impair native rhythmic acquisition. Few species can entrain to an auditory beat, and all those who do so appear to be vocal learners [101,102]. Furthermore, rhythmic entrainment between people during music making has been shown to increase social bonds and promote prosocial behavior [103,104]. Those with ASD are certainly able to process musical rhythms, but a lack of cultural specialization and social motivation may mean that they experience music somewhat differently from typically developing individuals. At the same time, it should be pointed out that children and adults with ASD appear to perceive emotion in music similarly to normal controls [70,105].

Everyday experience with Western music during normal development also leads to perceptual specialization for tonal pitch structure [73,74]. Sensitivity to harmonic structure develops rather late. Some implicit knowledge of harmonic structure can be seen by ages 4 to 7 [74,78,81,82], and explicit judgments emerge between 6 and 12 years of age [106]. Interestingly, we found that those with ASD were similar to controls in showing faster responses to expected tonic chords than to unexpected subdominant chords. Importantly, accuracy was also equivalent in both groups. Thus, the later developing skill related to harmonic structures appears to be relatively spared in ASD, consistent with the idea that the brain is most abnormal early in development. This finding is also generally consistent with the research of Heaton and colleagues [107-109], which suggests that musical pitch processing is more spared than speech processing in those with ASD. We extended these previous studies, however, by using an implicit task and examining both accuracy and reaction times.

Consistent with previous literature [69-72] we found a high instance of absolute pitch processing in our population with ASD (11% compared to 0% in the control group). Unlike tests for absolute pitch used in previous studies, which required participants to have musical training because the tasks involved naming notes according to Western notational conventions, we used a task that did not require formal musical training. Thus, we show that the previous findings also extend to those without musical training. Although absolute pitch is sometimes considered to be a gift, the more complex, but very common, ability to process relative pitch (comparing the pitch distance between two tones) is more important for both music and speech processing because it enables recognition of melodies and prosodic patterns across high or low pitch registers. It is also interesting that relative pitch typically develops early, with evidence that infants at least as young as 6 months recognize melodies transposed to higher or lower pitch registers [63-66]. The prevalence of absolute pitch in ASD, then, is consistent with early abnormalities in brain development. It is also consistent with the general prevalence of savant syndromes in ASD, which is about 10% [110]. In the typical population, the presence of absolute pitch is associated with early experience on a fixed-pitch instrument, leading researchers to speculate that it develops when there is a genetic predisposition combined with a particular environment [111]. In our ASD population, there was no evidence of greater musical experience in those with absolute pitch, suggesting that ASD may involve a genetic propensity for absolute pitch.

Heaton (2009) argues that absolute pitch in individuals with ASD is acquired differently than in the rest of the population with absolute pitch, and that anatomical features associated with absolute pitch in the normal population, such as the relative size of the planum temporale, are not present in those with ASD. Recent research suggests that in the general population, those with absolute pitch ability show local hyperconnectivity between the posterior superior and middle temporal gyri [112]. Because those with ASD have been shown to have greater short-range connectivity [25,26], it would be interesting to determine whether the brains of those with ASD with absolute pitch also show this feature. Although the number of those with ASD who have absolute pitch in our sample is small, they appear to show similar enculturation to Western tonal harmony as those who do not have absolute pitch, suggesting that absolute pitch and harmonic enculturation are separate abilities in ASD. Given that harmonic enculturation likely relies on relative pitch processing, this suggests that individuals with ASD may use absolute pitch processing in one task (pitch memory) and relative pitch processing in another task (harmonic priming), consistent with previous research [71].

Interestingly, when those with absolute pitch were eliminated from the sample, pitch memory was similar in those with and without ASD as measured by the ability to hold the pitch of one tone in mind and compare it to that of a second tone, whether or not there were interference tones in between. Thus we add to the literature on absolute pitch processing in ASD by showing that although it is more prevalent than in the general population, the majority of those with ASD show similar pitch memory performance as those without ASD. This similar performance across groups on memory stands in contrast to the decrements shown by the ASD group in ignoring one speech stream while attending to a second simultaneous speech stream. This latter difficulty persisted in the

ASD group even after accounting for receptive vocabulary, which is consistent with other reports of difficulty filtering non-speech stimuli [46]. The ability to group sounds into different perceptual streams develops early, with evidence for segregation of both simultaneous sounds [113] and sequential sounds [114-118] during infancy, although this ability continues to improve until 9 to 11 years of age [119]. Poor filtering in ASD in the context of deciphering speech signals, therefore, may have its origins early in development when brain growth and connectivity are abnormal.

Audiovisual integration is also present during the infancy period, both for speech and non-speech stimuli [120,121]. Consistent with previous literature [59,62], and the notion that early developing abilities will be particularly impaired in ASD, we found that those with ASD were less susceptible to the McGurk effect, integrating face and sound information to a much lesser extent than those who were typically developing. On the other hand, we found similar performance across groups on auditory alone and visual alone (lip-reading) conditions. We also found a significant correlation between performance on visual alone BA trials and receptive vocabulary, suggesting a link between lip-reading and general language abilities. Interestingly, higher receptive vocabulary scores were also related to a reduced likelihood of responding "ga" on the mismatched trials. Our results suggest that the audiovisual integration deficit found in our and previous studies cannot be entirely accounted for by differences in lip-reading ability in the absence of sound. It is possible that the reduction of long-range connectivity in ASD results in inadequate integration of auditory and visual information and perhaps a lack of top-down modulation of activity in sensory regions [29,30,122]. The ventral bank of the superior temporal sulcus of the left hemisphere has been implicated in audio-visual speech integration [123] so it would be interesting to examine activation patterns in those with ASD in this region using functional imaging techniques.

It remains for future research to determine how reported abnormalities in brain development [22-33] relate to the auditory processing profile for ASD revealed in the present paper. However, hypotheses to explore include whether reduction in long-range connectivity leads to less top-down modulation of perceptual processes, which would affect the ability to filter out irrelevant information and the ability to decipher speech in noisy environments. Reduced long-distance connectivity might also be expected to make integration and synchronization between sensory regions difficult for those with ASD, consistent with decreased auditory-visual integration. Increased local connectivity might relate to the propensity of those with ASD to focus on details to a greater extent that for normally developing individuals, leading to categorical perception that is less specialized for sounds in the native environment and to the increased frequency of absolute pitch in ASD.

It is noteworthy that musical processing appears to be relatively preserved among those with ASD. Interestingly, the ability to detect small pitch changes was preserved in those with ASD, and this ability was positively related to pitch memory, metrical processing, and harmonic processing as well as native phoneme processing. Overall, our results suggests that music might be a powerful remediation tool. Indeed there are suggestions that individuals with autism are more attracted to music than to speech

[109,124,125]. Perhaps the regular structure of music can provide a scaffold for the organization of sensory input [126,127,128]. Music also has added benefits for social development in that group music making can increase prosocial behavior, including co-operation and eye contact [129].

Finally, it is worth noting that there was considerable variability among those with ASD on some of the tasks, adding to the evidence that ASD manifests differently from individual to individual. For example, compared to controls, a high proportion of those with ASD (11%) had absolute pitch, but the other 89% appeared to process pitch similarly as controls. In the McGurk task, 9 out of 27 of those with ASD appeared to have normal auditory-visual speech integration, but the rest showed marked difference from the norm. It is important to understand these individual differences and how they develop. It is possible that more typical outcomes such as these are in part the result of particular experiences or training programs early in development, but it is impossible to determine this from the present data. That many of the impairments found in the ASD group were dependent upon abilities normally acquired during infancy suggests that there might be sensitive periods for the development of these abilities. Thus, future research is needed to determine whether there are sensitive periods during which these perceptual abnormalities can be best ameliorated through specific training.

References

1. Frith U (1989) Autism: Explaining the enigma. Oxford: Blackwell.

2. Peppé S, McCann J, Gibbon F, O'Hare A, Rutherford M (2007) Receptive and expressive prosodic ability in children with high-functioning autism. Journal of Speech, Language, and Hearing Research 50: 1015-1028.

3. Dawson G, Meltzoff A, Osterling J, Rinaldi J, Brown, E (1998) Children with autism fail to orient to naturally-occurring social stimuli. *Journal of Autism and Developmental Disorders 28*: 479-485.

4. Klin A, Lin DJ, Gorrindo P, Ramsay G, Jones W (2009) Literature review: Preference for physical compared to biological motion: Genome-wide association study. Autism Research 2: 178-179.

5. Pierce K, Conant D, Hazin R, Stoner R, Desmond J (2010) Preference for geometric patterns early in life as a risk factor for autism. Archives of General Psychiatry 68: 101-109.

6. Rutherford MD (2005) A retrospective journal-based case study of the development of a child with autism and his twin. Neurocase 11: 1-9.

7. Zwaigenbaum L, Bryson S, Lord C, Rogers S, Carter A, et al. (2009) Clinical assessment and management of toddlers with suspected autism spectrum disorder: Insights from studies of high-risk infants. Pediatrics 123: 1383-1391.

8. Happé F, Frith U (2006) The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. Journal of Autism and Developmental Disorders 35: 5-25.

9. Baron-Cohen S, Hammer J (1997) Parents of children with asperger syndrome: What is the cognitive phenotype? Journal of Cognitive Neuroscience 9: 548-554.

10. Plaisted K, Swettenham J, Rees L (1999) Children with autism show local precedence in a divided attention task and global precedence in a selective attention task. Journal of Child Psychology and Psychiatry 40: 733-742.

11. Shah A, Frith U (1983) An islet of ability in autistic children. A research note. Journal of Child Psychology and Psychiatry and Applied Disciplines 24: 613-620.

12. Baranek GT, Boyd BA, Poe MD, David FJ, Watson LR (2007) Hyperresponsive sensory patterns in young children with autism, developmental delay, and development. American Journal of Mental Retardation 112: 233-245.

13. Dahlgren SV, Gillberg C (1989) Arousal, attention, and the socioemotional impairments of individuals with autism. In: Dawson G, editor. Autism: Nature, diagnosis and treatment. New York: Guilford. pp. 49-74.

14. Gomes E, Rotta NT, Pedroso FS, Sleifer P, Danesi MC (2004) Auditory hypersensitivity in children and teenagers with autistic spectrum disorder. Arq Neuropsiquiatr 62: 797-801.

15. Kern JK, Trivedi MH, Garver CR, Grannemann BD, Andrews AA, et al. (2006) The pattern of sensory processing abnormalities in autism. Autism 10: 480-490.

16. Rimland B, Edelson SM (1995) Brief report: A pilot study of auditory integration training in autism. Journal of Autism and Developmental Disabilities 25: 61-70.

17. Rosenhall U, Nordin V, Sandstrom M, Ahlsen G, Gillberg C (1999) Autism and hearing loss. Journal of Autism and Developmental Disorders 29: 349-357.

18. Baranek GT (1999) Autism during infancy: A retrospective video analysis of sensory-motor and social behaviors at 9-12 months of age. Journal of Autism and Developmental Disorders 29: 213-224.

19. Dawson G, Toth K, Abbott R, Osterling J, Munson J, et al. (2004) Early social attention impairments in autism: Social orienting, joint attention, and attention to distress. Developmental Psychology 40: 271-83.

20. Osterling J, Dawson G (1994) Early recognition of children with autism: A study of first birthday home video tapes. Journal of Autism and Developmental Disorders 24: 247-257.

21. Werner E, Dawson G, Osterling J, Dinno N (2000) Brief report: Recognition of autism spectrum disorder before one year of age: A retrospective study based on home videotapes. Journal of Autism and Developmental Disorders 30: 157-162.

22. Courchesne E, Karns CM, Davis HR, Ziccardi R, Carper RA, et al. (2001) Unusual brain growth patterns in early life in patients with autistic disorder: An MRI study. Neurology 57: 245-254.

23. Courchesne E, Carper R, Akshoomoff N (2003) Evidence of brain overgrowth in the first year of life in autism. The Journal of the American Medical Association 290: 337-344.

24. Courchesne E (2004) Brain development in autism: Early overgrowth followed by premature arrest of growth. Mental Retardation and Developmental Disability Research Review 10: 106-111.

25. Casanova MF (2004) Intracortical circuitry: One of Psychiatry's missing assumptions. European Archives of Psychiatry and Clinical Neuroscience 254: 148-151.

26. Casanova MF, Buzhoeveden DP, Switala AE, Roy E (2002) Minicolumnar pathology in autism. American Academy of Neurology 58: 428-432.

27. Courchesne E, Pierce K (2005) Why the frontal cortex in autism might be talking only to itself: Local over-connectivity but long-distance disconnection. Current Opinion in Neurobiology 15: 225-230.

28. Barnea-Goraly N, Kwon H, Menon V, Eliez S, Lotspeich L, et al. (2004) White matter structure in autism: Preliminary evidence from diffusion tensor imaging. Biological Psychiatry 55: 323-326.

29. Bird G, Catmur, C, Silani, G, Frith, C, Frith, U (2006) Attention does not modulate neural responses to social stimuli in autism spectrum disorders. Neuroimage 31: 1614-1624.

30. Castelli F, Frith C, Happé F, Frith U (2002) Autism, asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. Brain 125: 1839-1849.

31. Just MA, Cherkassky VL, Keller TA, Kana RK, Minshew NJ (2007) Functional and anatomical cortical underconnectivity in autism: Evidence from an fMRI study of an executive function task and corpus callosum morphometry. Cerebral Cortex 17: 951-961.

32. Just MA, Cherkassky VL, Keller TA, Minshew NJ (2004) Cortical activation and synchronization during sentence comprehension in high-functioning autism: Evidence of underconnectivity. Brain 127: 1811-1821.

33. Skukla DK, Keehn B, Müller R (2010) Tract-specific analyses of diffusion tensor imaging show widespread white matter compromise in autism spectrum disorder. Journal of Child Psychology and Psychiatry 52: 286-295.

34. Griffiths TD (2007) Sorting out Sound. Neuron 56: 580-581.

35. Boddaert N, Chabane N, Belin P, Bourgeois M, Royer V, et al. (2004) Perception of complex sounds in autism: Abnormal auditory cortical processing in children. American Journal of Psychiatry 161: 2117-2120.

36. Herbert MR, Harris GJ, Adrien KT, Ziegler DA, Makris N, et al. (2002) Abnormal asymmetry in language association cortex in autism. Annals of Neurology 52: 588-596.

37. Ashburner J, Ziviani J, Rodger S (2008) Sensory processing and classroom emotional

behavioural, and education outcomes in children with autism spectrum disorder. American Journal of Occupational Therapy 62: 564-573.

38. Baker AE, Lane A, Angley MT, Young RL (2008) The relationship between sensory processing patterns and behavioural responsiveness in autistic disorder: A pilot study. Journal of Autism and Developmental Disorders 38: 867-875.

39. Lane A, Young RL, Baker AEZ, Angley M (2010) Sensory processing subtypes in autism: Association with adaptive behavior. Journal of Autism and Developmental Disorders 40: 112-122.

40. Rogers SJ, Hepburn S, Wehner E (2003) Parent reports of sensory symptoms in toddlers with autism and those with other developmental disorders. Journal of Autism and Developmental Disorders 33: 631-42.

41. Schoen SA, Miller LJ, Brett-Green BA, Nielsen DM (2009) Physiological and behavioral differences in sensory processing: A comparison of children with autism spectrum disorder and sensory modulation disorder. Frontiers in Integrative Neuroscience 3: 1-11.

42. McIntosh DN, Miller LJ, Shyu V, Dunn W (1999) Overview of the Short Sensory Profile (SSP). In: Dunn W, editor. The sensory profile: Examiner's manual. San Antonio, TX: The Psychological Corporation. pp. 59-73.

43. Alcántara JI, Weisblatt EJL, Moore BCJ, Bolton PF (2004) Speech-in-noise perception in high-functioning individuals with autism or asperger's syndrome. Journal of Child Psychology and Psychiatry 45: 1107-1114.

44. Groen WB, van Orsouw L, ter Huurne N, Swinkels S, van der Gaag RJ, et al. (2009) Intact spectrum but abnormal temporal processing of auditory stimuli in autism. Journal of Autism and Developmental Disorders 39: 742-750.

45. Teder-Sälejärvi WA, Pierce KL, Courchesne E, Hillyard SA (2005) Auditory spatial localization and attention deficits in autistic adults. Cognitive Brain Research 23: 221-234.

46. Lepistö T, Kuitunen A, Sussman E, Saalasti S, Jansson-Verkasalo E, et al. (2009) Auditory stream segregation in children with asperger syndrome. Biological Psychology 82: 301-307.

47. Russo N, Nicol T, Trommer B, Zecker S, Kraus N (2009) Brainstem transcription of speech is disrupted in children with autism spectrum disorders. Developmental Science 12: 557-567.

48. Russo NM, Zecker S, Trommer B, Chen J, Kraus N (2009) Effects of background noise on cortical encoding of speech in autism spectrum disorders. Journal of Autism and Developmental Disorders 39: 1185-1196.

49. Best CT, McRoberts GW, Goodell E (2001) Discrimination of non-native consonant contrasts varying in perception assimilation to the listener's native phonological system. Journal of Acoustic Society of America 109: 775-794.

50. Werker JF, Tees RC (2005) Speech perception as a window for understanding plasticity and commitment in language systems of the brain. Developmental Psychobiology 46: 233-251.

51. Werker JF, Lalonde CE (1988) Cross-language speech perception: Initial capabilities and developmental change. Developmental Psychology 24: 672-683.

52. Gopnik A, Meltzoff AN, Kuhl PK (1999) The scientist in the crib: Minds, brains and how children learn. New York: Harper Collins.

53. Kuhl PK, Tsao FM, Liu, HM (2003) Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. Proceedings of the National Academy of Sciences 100: 9096-9101.

54. Dawson G, Webb SJ, McPartland J (2005) Understanding the nature of face processing impairment in autism: Insights from behavioral and electrophysiological studies. Developmental Neuropsychology 27: 403-424.

55.Osterling JA, Dawson G, Munson JA (2002) Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation. Developmental Psychopathology 14: 239-251.

56. Schultz RT, Gauthier I, Klin A, Fulbright RK, Anderson AW, et al. (2000) Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and Asperger syndrome. Archives of General Psychiatry 57: 331-340.

57. Constantino JN, Yang D, Gray TL, Gross MM, Abbacchi AM, et al. (2007) Clarifying the association between language and social development in autism: A study of non-native phoneme recognition. Journal of Autism and Developmental Disorders 37: 1256-1263.

58. McGurk H, MacDonald J (1976) Hearing lips and seeing voices. Nature 264: 746-748.

59. De Gelder B, Vroomen J, van der Heide L (1991) Face recognition and lip-reading in autism. European Journal of Cognitive Psychology 3: 69-86.

60. Iarocci G, Rombough A, Yager J, Weeks DJ, Chua R (2010) Visual influences on speech perception in children with autism. Autism 14: 305-320.

61. Williams JHG, Massaro DW, Peel NJ, Bosseler A, Suddendorf T (2004) Visual auditory integration during speech imitation in autism. Research in Developmental Disabilities 25: 559-575.

62. Smith EG, Bennetto L (2007) Audiovisual speech integration and lipreading in autism. Journal of Child Psychology and Psychiatry 48: 813-821.

63. Plantinga J, Trainor LJ (2008) Infants' memory for isolated tones and the effects of interference. Music Perception 26: 121-128.

64. Trainor LJ, Trehub SE (1992) A comparison of infants’ and adults’ sensitivity to Western musical structure. Journal of Experimental Psychology: Human Perception and Performance 18: 394-402.

65. Trehub SE (2001) Musical predispositions in infancy. Annals of the New York Academy of Sciences 930: 1-16.

66. Trehub SE, Bull D, Thorpe LA (1984) Infants' perception of melodies: The role of melodic contour. Child Development 55: 821-830.

67. Bachem A (1955) Absolute pitch. Journal of the Acoustical Society of America 27: 1180-1185.

68. Brown WA, Cammuso K, Sachs H, Winklosky B, Mullane J, et al. (2003) Autism-related language, personality, and cognition in people with absolute pitch: Results of a preliminary study. Journal of Autism and Developmental Disorders 33: 163-167.

69. Brenton JN, Devries SP, Barton C, Minnich H, Sokol D (2008) Absolute pitch in a four-year-old boy with autism. Pediatric Neurology 39: 137-138.

70. Heaton P, Pring L, Hermelin B (1999) A pseudo-savant: A case of exceptional musical splinter skills. Neurocase 5: 503-509.

71. Mottron L, Peretz I, Belleville S, Rouleau N (1999) Absolute pitch in autism: A case-study. Neurocase 5: 485-501.

72. Young RL, Nettelbeck T (1995) The abilities of a musical savant and his family. Journal of Autism and Developmental Disorders 25: 231-248.

73. Hannon EE, Trainor LJ (2007) Music acquisition: Effects of enculturation and formal training on development. Trends in Cognitive Sciences 11: 466- 472.

74. Trainor LJ, Corrigall KA (2010) Music acquisition and effects of musical experience. In: Riess-Jones M, Fay RR, editors. Springer Handbook of Auditory Research: Music Perception. Heidelberg: Springer. pp. 89-128.

75. Hannon EE, Trehub SE (2005a) Metrical categories in infancy and adulthood. Psychological Science 16: 48-55.

76. Hannon EE, Trehub SE (2005b) Tuning in to musical rhythms: Infants learn more readily than adults. Proceedings of the National Academy of Sciences 102: 12639-12643.

77. Trehub SE, Hannon EE (2009) Conventional rhythms enhance infants' and adults' perception of music. Cortex 45: 110-118.

78. Trainor LJ, Trehub SE (1994) Key membership and implied harmony in Western tonal music: Developmental perspectives. Perception and Psychophysics 56: 125-132.

79. Clayton M (2000) Time in Indian music. New York: Oxford University Press.

80. Gerry DW, Faux AL, Trainor LJ (2010) Effects of Kindermusik training on infants' rhythmic enculturation. Developmental Science 13: 545-551.

81. Corrigall KA, Trainor LJ (2010) Musical enculturation in preschool children: Acquisition of key and harmonic knowledge. Music Perception 28: 195-200.

82. Schellenberg EG, Bigand E, Poulin-Charronnat B, Garnier C, Stevens C (2005) Children's implicit knowledge of harmony in Western music. Developmental Science 8: 551-566.

83. Bigand E, Madurell F, Tillmann B, Pineau M (1999) Effect of global structure and temporal organization on chord processing. Journal of Experimental Psychology: Perception and Performance 25: 184-197.

84. Poulin-Charronnat B, Bigand E, Madurell F, Peereman R (2004) Musical structure modulates semantic priming in vocal music. Cognition 94: B67-B78.

85. Tillmann B, Bigand E, Escoffier N, Lalitte P (2006) The influence of musical relatedness on timbre discrimination. European Journal of Cognitive Psychology 18: 343-358.

86. Tillmann B, Peretz I, Bigand E, Gosselin N (2007) Harmonic priming in an amusic patient: The power of implicit tasks. Cognitive Neuropsychology 24: 1-20.

87. Heaton P, Williams K, Cummins O, Happe F (2007) Beyond perception: Musical representation and on-line processing in autism. Journal of Autism and Developmental Disorders 37: 1355-1360.

88. Lord C, Rutter M, Goode S, Heemsbergen J, Jordan H, et al. (1989) Autism diagnostic observation schedule: A standardized observation of communicative and social behavior. Journal of Autism and Developmental Disorders 19: 185-212.

89. Lord C, Rutter M, Le Couteur A (1994) Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. Journal of Autism and Developmental Disorders 24: 659-685.

90. Pfordresher PQ, Brown S (2009) Poor-pitch singing in the absence of “tone deafness.” Music Perception 25: 95-115.

91. Ross DA, Olson IR, Marks LE, Gore JC (2004) A nonmusical paradigm for identifying absolute pitch possessors. Journal of the Acoustical Society of America 116: 1793-1799.

92. Wechsler D (1997) Digit Span Subtest of the Wechsler Memory Scale-III. San Antonio, TX: The Psychological Corporation.

93. Dunn LM, Dunn LM (1997) Peabody Picture Vocabulary Test (3rd ed) Circle Pines, MN: American Guidance Services.

94. Willeford JA (1977) Assessing central auditory behavior in children: A test battery approach. In: Keith R, editor. Central auditory dysfunction. New York: Grune & Stratton. pp. 43-72.

95. Roid G, Miller L (2005) Leiter International Performance Scale. GL Assessment: London.

96. Tucker-Davis Technologies (2007) Programmable attenuator. Alachua, FL: Tucker-Davis Technologies.

97. Han LA, Poulsen T (1998) Equivalent threshold sound pressure levels (ETSPL) for sennheiser HDA 200 earphone and etymotic research ER-2 insert earphone in the frequency range 125 Hz to 16 kHz. Scandinavian Audiology 27: 105-112.

98. Neurobehavioral Systems (2007) Presentation 11.0. Albany, CA: Neurobehavioral Systems, Incorporated.

99. Watson AB, Pelli DG (1983) QUEST: A Bayesian adaptive psychometric method. Perception & Psychophysics 33: 113-120.

100. Ratcliff R (1993) Methods for dealing with reaction time outliers. Psychological Bulletin, 114: 510-532.

101. Patel AD, Iversen JR, Bregman MR, Schulz I (2009) Experimental evidence for synchronization to a musical beat in a nonhuman animal. Current Biology 19: 827-830.

102. Schachner A, Brady TF, Pepperberg I, Hauser M (2009) Spontaneous motor entrainment to music in multiple vocal-mimicking species. Current Biology 19: 831-836.

103. Kirschner S, Tomasello M (2009) Joint drumming: Social context facilitates synchronization in preschool children. Journal of Experimental Child Psychology 102: 299-314.

104. Kirschner S, Tomesello M (2010) Joint music making promotes prosocial behaviour in 4-year-old children. Evolution and Human Behaviour 31: 354-364.

105. Allen R, Hill E, Heaton P (2009) 'Hath charms to soothe...': An exploratory study of how high-functioning adults with ASD experience music. Autism 13: 21-41.

106. Costa-Giomi E (2003) Young children's harmonic perception. Annals of the New York Academy of Sciences 999: 477-484.

107. Heaton P (2009) Assessing musical skills in autistic children who are not savants. Philosophical Transactions of the Royal Society B: Biological Sciences 364: 1443-1447.

108. Heaton P, Hudry K, Ludlow A, Hill EL (2008) Superior discrimination of speech pitch and its relationship to verbal ability in autism spectrum disorders. Cognitive Neuropsychology 25: 771-782.

109. Jarvinen-Pasley A, Heaton P (2007) Evidence for reduced domain-specificity in auditory processing in autism. Developmental Science 10: 786-793.

110. Rimland B, Hill AL (1984) Idiot savants. In: Wortis J, editor. Mental Retardation and Developmental Disabilities. New York: Plenum Press. pp. 55-169.

111. Baharloo S, Johnston PA, Service SK, Gitschier J, Freimer NB (1998) Absolute pitch: An approach for identification of genetic and nongenetic components. Journal of Human Genetics 62: 224-231.

112. Loui P, Li H, Hohmann A, Schlaug G (2011) Enhanced cortical connectivity in absolute pitch musicians: A model for local hyperconnectivity. Journal of Cognitive Neuroscience 23: 1015-1026.

113. Folland N, Butler BE, Smith NA, Trainor LJ (2012) Processing simultaneous auditory objects in infancy: Music and mistuned harmonics. Journal of the Acoustical Society of America 131: 993-997.

114. Demany L (1982) Auditory stream segregation in infancy. Infant Behavior & Development 48: 261-276.

115. Fassbender C (1993) Auditory grouping and segregation processes in infancy. Norderstedt: Kaste Verlag.

116. McAdams S, Bertoncini J (1997) Organization and discrimination of repeating sound sequences by newborn infants. Journal of the Acoustical Society of America 102: 2945-2953.

117. Smith NA, Trainor LJ (2011) Auditory stream segregation improves infants' selective attention to target tones amid distracters. Infancy 16: 1-14.

118. Winkler I, Kushnerenko E, Horvath J, Ceponiene R, Fellman V, et al. (2003) Newborn infants can organize the auditory world. Proceedings of the National Academy of Sciences of the United States of America 100: 11812-11815.

119. Sussman E, Wong R, Horváth J, Winkler I, Wang W (2007) The development of perceptual organization of sound by frequency separation in 5-11 year-old children. Hearing Research 225: 117-127.

120. Kawabe T, Shirai N, Wada Y, Miura K, Kanazawa S, et al. (2010) The audiovisual tau effect in infancy. PLoS ONE 5: 1-7.

121. Rosenblum LD, Schmuckler MA, Johnson JA (1997) The McGurk effect in infants. Perceptual Psychophysics 59: 347-357.

122. Iarocci G, McDonald J (2006) Sensory integration and the perceptual experience of persons with autism. Journal of Autism and Developmental Disorders 36: 77-90.

123. Calvert GA, Campbell R, Brammer MJ (2000) Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. Current Biology 10: 649-657.

124. Finnigan E, Starr E (2010) Increasing social responsiveness in a child with autism: A comparison of music and non-music interventions. Autism 14: 321-348.

125. Simpson K, Keen D (2010) Teaching young children with autism graphic symbols embedded within an interactive song. Journal of Developmental and Physical Disabilities 22: 165-177.

126. Allen R, Heaton P (2010) Autism, music, and the therapeutic potential of music in alexithymia. Music Perception 27: 251-261.

127. Wan CY, Demaine K, Zipse L, Norton A, Schlaug G (2010) From music making to speaking: Engaging the mirror neuron system in autism. Brain Research Bulletin 82: 161-168.

128. Wan CY, Schlaug G (2010) Neural pathways for language in autism: The potential for music-based treatments. Future Neurology 5: 797-805.

129. Overy K, Molnar-Szakacs I (2009) Being together in time: Musical experience and the mirror neuron system. Music Perception 26: 489-504.

130. Bigand E, Poulin B, Tillmann B, Madurell F, D'Adamo DA (2003) Sensory versus cognitive components in harmonic priming. Journal of Experimental Psychology 29: 159-171.

CHAPTER 3: INTRODUCTION

Those with ASD perceive faces differently than normal controls (Pierce, Conant, Hazin, Stoner, & Desmond, 2010), which can impact their ability to achieve certain social milestones that are important for language development (e.g., joint attention). The result may be that those with ASD are less motivated to communicate with others and this might be evident in aspects of their speech. There are reports of ASD speech sounding, “robotic”, “monotone” and even “singsong” (e.g., Frith, 1991). However, past work in this area is primarily based on subjective ratings of production patterns instead of robust acoustic analysis.

In addition to how acoustic features are varied, good communication depends on varying these features in a way that is useful to listeners. Research indicates that conversations typically involve new information (focus) and given information from previous utterances (topic) (Vallduví & Engdahl, 1996). Given that focus words have a higher information value than topic words, the former is made more prominent than the latter, such as with larger pitch ranges and longer word durations (Chen, 2009). Speakers who use acoustic features such as these are able to capture the listener’s attention and convey that they are tuned to their conversational partner.

Here we measure how adults with ASD relative to typical speakers vary acoustic features generally in their speech and as a function of the information structure. We also examine whether the level of current language ability of our ASD speakers is associated in a predicable way with prosody use.

Use of prosody and information structure in high functioning adults with Autism in relation to language ability

Anne-Marie R. DePape[a], Aoju Chen[b,c], Geoffrey B. C. Hall[a] and Laurel J. Trainor[a,d] *

[a] Psychology, Neuroscience & Behaviour, McMaster University

[b] Utrecht Institute of Linguistics, Utrecht University

[c] Max Planck Institute for Psycholinguistics

[d] Rotman Research Institute, Baycrest Hospital

Running head: PROSODY AND INFORMATION STRUCTURE IN AUTISM

Keywords: prosody, language ability, information structure, pitch, duration, Autism

Correspondence: Laurel J. Trainor, Department of Psychology, Neuroscience & Behaviour, McMaster University, 1280 Main Street West, Hamilton, ON, L8S 4K1, Canada, E-mail: ljt@mcmaster.ca, Phone: 1-905-525-9140 Ext 23007, Fax: 1-905-529-6225

Acknowledgements: This research was supported by a Natural Sciences and Engineering Research Council of Canada grant to LJT and by a Junior Visiting Fellowship from the Max Planck Institute for Psycholinguistics (MPI) to ARD. We thank Wolfgang Klein (MPI) for his support in this research and Tilman Harpe (MPI) for drawing the pictures for our experimental task.

Reprinted from DePape, A.M., Chen, A., Hall, G.B.C., & Trainor, L.J. (2012). Use of prosody and information structure in high functioning adults with Autism in relation to language ability. *Frontiers in Psychology*, 3:72. doi: 10.3389/fpsyg.2012.00072

Abstract

Abnormal prosody is a striking feature of the speech of those with Autism Spectrum Disorder (ASD), but previous reports suggest large variability among those with ASD. Here we show that part of this heterogeneity can be explained by level of language functioning. We recorded semi-spontaneous but controlled conversations in adults with and without Autism Spectrum Disorder and measured features related to pitch and duration to determine (1) general use of prosodic features, (2) prosodic use in relation to marking information structure, specifically, the emphasis of new information in a sentence (focus) as opposed to information already given in the conversational context (topic), and (3) the relation between prosodic use and level of language functioning. We found that, compared to typical adults, those with ASD with high language functioning generally used a larger pitch range than controls but did not mark information structure, whereas those with moderate language functioning generally used a smaller pitch range than controls but marked information structure appropriately to a large extent. Both impaired general prosodic use and impaired marking of information structure would be expected to seriously impact social communication and thereby lead to increased difficulty in personal domains, such as making and keeping friendships, and in professional domains, such as competing for employment opportunities.

1. Introduction

Autism Spectrum Disorder (ASD) involves impaired social interactions, repetitive and restrictive behaviors, and problems with communication (American Psychiatric Association, 1994). One striking feature of the speech of those with ASD is abnormal prosody (e.g., Baltaxe, Simmons, & Zee, 1984; Bonneh, Levanon, Dean-Pardo, Lossos, & Adini, 2011; Diehl, Watson, Bennetto, McDonough, & Gunlogson, 2009; Green & Tobin, 2009; McCann & Peppé, 2003; Nadig & Shaw, 2011; Paul, Augustyn, Klin, & Volkmar, 2005; Paul, Orlovski, Marchinko, & Volkmar, 2009; Sharda et al., 2010; Shriberg et al., 2001). Prosody (or intonation) refers to suprasegmental features of speech, including pitch, duration, and intensity. According to Roach (2000) prosody serves important communicative functions at the grammatical, pragmatic and affective levels. For example, prosody is used to distinguish speech acts such as questions, statements and imperatives; to convey what is old and new information, and other sorts of pragmatic cues; and, at the affective level, to convey information about a speaker's feeling state (e.g., Chun, 1988; Cruttenden, 1997; Gussenhoven, 2004; Halliday, 1967; Ladd, 1996; Nespor & Vogel, 1986). In the present paper, we examine prosody at the pragmatic level. Abnormal prosody was included in the early descriptions of ASD (Asperger, 1944; Kanner, 1943), but has not been considered a defining feature of ASD, likely because the abnormalities appear to manifest differently across individuals (Baltaxe et al., 1984; Bonneh et al., 2011; Diehl et al., 2009; Green & Tobin, 2009; Schreibman, Kohlenberg, & Britten, 1986; Van Lancker, Cornelius, & Kreiman, 1989). The prosody of ASD speech has been variously described as sounding "robotic", "wooden", "stilted", "monotone", "bizarre", "over precise", and even "singsong" (Baltaxe & Simmons, 1985; Baron-Cohen & Staunton, 1994; Fay & Schuler, 1980; Frith, 1991). Abnormalities appear to include both decreased and increased use of prosodic expression in ASD (Schreibman et al., 1986; Van Lanacker et al., 1989), and there is suggestive evidence of "prosodic disorganization" in that prosody is not necessarily used to highlight the intended meaning (e.g., see Green & Tobin, 2009).

Here we report detailed acoustic analyses of prosodic use in adults with and without ASD in sentences generated in semi-spontaneous conversations in which sentence structure and use of specific words were highly controlled. Furthermore, we examine whether level of current language ability (which in our sample also reflected whether or not there had been early language delay and whether a diagnosis of high-functioning autism [HFA] or Asperger's syndrome [AS] had been given) was associated in a predicable way with prosody use in adults with ASD. In contrast to communication deficits, language ability (encompassing articulation, phonological processing, vocabulary, grammatical and semantic skills) is highly variable in ASD, ranging from the high end of the normal distribution to completely non-verbal (e.g., Kjelgaard & Tager-Flusberg, 2001; Lord & Paul, 1997). Such variability is consistent with recent genetic studies that indicate that although ASD is strongly heritable, it is etiologically heterogenetic, with many loci that each contribute a small amount to genetic susceptibility (e.g., Geschwind, 2009).

Language ability is an important indicator in ASD, as language is highly predictive of the general prognosis for a child (see Kjelgaard & Tager-Flusberg, 2001).

Furthermore, language is related to a number of specific abilities. For example, of children with ASD, only those with poor language skills show a low ability to suppress word meanings that are not consistent within a context; those with language skills in the normal range show normal context-dependent suppression (Brock, Norbury, Einav, & Nation, 2008; Norbury, 2005). Similarly, language ability predicts whether children with ASD use the appropriate amount of information in descriptions of objects according to the knowledge of their communication partner (Nadig, Vivanti, & Ozonoff, 2009). In one study, Norbury and colleagues (2009) used eye tracking while participants watched videos of peers interacting in familiar situations. Interestingly, they found that those with ASD and poor language skills were similar to normally developing controls in their viewing patterns of the eyes and mouths of their peers, whereas those with ASD and normal language ability spent less time than the other groups viewing the eyes. This suggests that language skills may not necessarily be connected with better communication skills, and indicates that the origins and nature of communication problems in ASD may differ between children with higher and lower language functioning. In the present paper, we investigate the general and communicative use of prosody in high-functioning adults with ASD who score above or below the mean of the normal population on vocabulary, which is highly related to general language skills in ASD (e.g., see Kjelgaard & Tager-Flusberg, 2001).

Most studies of prosody in ASD have examined children rather than adults or even adolescents (e.g., Baltaxe et al., 1984; Bonneh et al., 2011; Diehl et al., 2009; Fosnot & Jun, 1999; Green & Tobin, 2009; Grossman, Bemis, Skwerer, & Tager-

Flusberg, 2010; Hubbard & Trauner, 2007; Nadig & Shaw, 2011; Paccia & Curcio, 1982; Paul, Bianchi, Augustyn, Klin, & Volkmar, 2008; Sharda et al., 2010). Despite descriptions of monotone speech, studies employing acoustic analyses have generally found increased pitch variability in children with ASD, whether the corpus analysed consisted of isolated words (Bonneh et al., 2011), conversations (Green & Tobin, 2009; Nadig & Shaw, 2011; Sharda et al., 2010), narratives (Diehl et al., 2009) or reading aloud (Green & Tobin, 2009). However, there appear to be individual differences. Baltaxe et al. (1984) found that children with ASD had either very narrow or very wide pitch ranges, suggesting heterogeneity among children. Similarly, Green and Tobin (2009) found that although children with ASD as a group showed larger pitch ranges and larger pitch variability compared to typically-developing children, those with ASD could be divided into three distinct groups, consisting of those with narrow, typical or wide pitch ranges. Similar variance across individuals might also exist for prosodic use of duration, although there is less research on this question. Nadig and Shaw (2011) reported no difference in overall speech rate between children with and without ASD. In other studies, adults with ASD were found to produce less lengthening than controls on stressed syllables in imitative speech (Paul et al., 2008), but children with ASD were found to produce more lengthening than controls on stressed syllables in spontaneous speech (Grossman et al., 2010). Clearly, more research is needed in order to understand the prosodic use of duration in ASD.

With respect to pitch, global measures of pitch range and variability do not entirely capture the abnormal nature of prosody in those with ASD. For example, experienced raters rated the prosody of those with ASD as more atypical than that of normally-developing children, even though they rated both populations as sounding similar in terms of amount of pitch variation (Nadig & Shaw, 2011). Prosodic use in ASD has been described as "disorganized", likely indicating that pitch and duration variation are not always used to enhance communication (see Green & Tobin, 2009). For example, those with ASD appear to use a restricted number of prosodic contours in their utterances (Green & Tobin, 2009), consistent with the idea that prosodic variation is not always optimized for communicative intent in those with ASD. Furthermore, it is also possible that this lack of utterance-level contour variation might contribute to a sense of overall monotony.

Critical to an understanding of prosodic abnormalities in ASD is the question of whether prosody is used to enhance communication. The present paper examines the use of prosody to mark *information structure* in individuals with ASD. In normal conversation, prosody is used to convey what is important in an utterance with respect to the talker's beliefs about the listener's knowledge state (Chafe, 1976; Clark & Haviland, 1977; Prince, 1986). Two of the most widely discussed information structural categories are (1) topic, which refers to what a sentence is about and typically represents given information, and (2) focus, which typically represents new information about the topic (Lambrecht, 1994; Vallduví & Engdahl, 1996). For example, "boy" is the topic and "apple" is the focus of the sentence "The boy is eating an apple" when uttered in response to the question "*What* is the boy eating?". However, "apple" is the topic and "boy" is the focus of the same sentence when uttered in response to "*Who* is eating the apple?".

Among typical speakers, focus words are produced with a larger pitch range and longer duration than topic words, all other acoustic features being equal (Chen, 2009). Making focal information more prominent can facilitate language comprehension whereas making the topical information more prominent can delay comprehension (e.g., Birch & Clifton 1995; Chen, 2010; Nooteboom & Terken, 1982). Inappropriate marking of information can lead to problems in achieving desired communicative intents and produce, among other things, confusion between conversational partners (Fine, Bartolucci, Ginsberg, & Szatmari, 1991).

Developmentally, the tendency to use a falling pitch contour across a sentence may sometimes override children's ability to mark intended meanings, for example, not using a rising contour when appropriate to ask a question (Wells, Peppé, & Goulandris, 2004). One study of Dutch-speaking children found that when answering a question, 7- to -8-year-old, but not 4- to 5-year-old, children emphasized focus words appropriately (Chen, 2011). In particular, the 4- to 5-year-olds accented focus words with several types of accents (e.g., rise, fall, downstepped fall – a fall with a lower peak than the preceding accent) and showed no adult-like preference for falling accents in the sentence-final (object) position, a problem that the author attributed to the children's need to check and seek confirmation (hence the final rise) and a lack of knowledge of the typical functions of downstepped fall. On the other hand, earlier work on English children and a study of German children suggested that when the focal information is contrastive, even 3- to 4-year-olds showed evidence of using prosody appropriately (Hornby & Hass, 1970; Müller, Höhle, Schmitz, & Jürgen, 2006).

Previous reports of abnormalities in topic and focus accentuation in ASD mainly used subjective judgments of accent rather than acoustic measurements of pitch or duration in focus marking. One study found that children with ASD accentuated focus and topic words equally (McCaleb & Prizant, 1985), whereas others, including one with adults, found that those with ASD accentuated the beginning of a sentence irrespective of its information value (e.g., Baltaxe, 1984; Baltaxe & Guthrie, 1987; Peppé, Cleland, Gibbon, O'Hare, & Martínez-Castilla, 2011; Peppé, McCann, Gibbon, O'Hare, & Rutherford, 2006, 2007; Shriberg et al., 2001). Most of these studies examined contrastive stress, where correct prominence is placed on the contrastive focus. For example, when presented with an informationally incorrect sentence such as "The green sheep has the ball" participants might respond, "No, the green *COW* has the ball" (Peppé et al., 2006), accenting the word correcting the information. The typically developing literature shows that focus information structure is marked to a lesser extent in the sentence-final (object) compared to sentence-initial (subject) position. Developmentally, sentence-final marking appears to develop later than sentence-initial marking. As mentioned above, Chen (2011) found that the marking of information structure in the sentence-final position in typically developing children was not adult-like until age 7. In the present study we examine the marking of (non-contrastive) focus and topic in both sentence-initial and sentence-final positions.

The small amount of research on prosody in adolescents and adults with ASD suggests that the abnormalities documented in children persist through late development and are resistant to change (Diehl et al., 2009; Paul et al., 2005; Shriberg et al., 2001).

Not surprisingly, atypical prosody in adults with ASD can have real-life consequences, such as affecting their ability to make friends and achieve meaningful employment (Paul et al., 2005; Van Bourgondien & Woods, 1992). Thus, a full understanding of the nature of the prosodic deficits is important.

We collected semi-spontaneous speech samples in adults in a controlled but interactive paradigm that enabled us to directly measure pitch and duration features of the same words in focus and topic conditions in sentence-initial and sentence-final positions. We had three main goals: (1) To compare the general use of prosodic pitch and duration in adults with and without ASD; (2) to examine the use of pitch and duration to convey information structure in adults with and without ASD in short, controlled conversations; and (3) to examine whether individual differences in use of prosody are related to level of language functioning.

2. Method

2.1 Participants

We tested 12 adult male participants (*M* = 25.4 years; *range* = 17 to 34 years) with a diagnosis of ASD (Table 1). Of these 6 had receptive vocabulary standard scores of 100 or greater and 6 had scores below 100 as measured by the standardized Peabody Picture Vocabulary Test-III (PPVT; Dunn & Dunn, 1997). ASD participants had been seen at clinic (Offord Centre), assessed using standard instrument batteries (ADOS and ADI) (Lord et al., 1989; Lord, Rutter, & Le Couteur, 1994), and all carried formal psychiatric diagnoses of either AS or HFA. Participants completed the PPVT and a questionnaire on languages spoken and family history of ASD. Previous research has found that scores on the PPVT are correlated with scores on the Clinical Evaluation of Language Fundamentals (CELF) test, which includes assessments of morphology, syntax, semantics and working memory for language (Kjelgaard & Tager-Flusberg, 2001). Thus, the PPVT can be used as a measure of general language functioning. The categorization by current language ability (PPVT) followed their diagnoses, such that all 6 with score of 100 or greater (Autism High Language Function, A-highL group) carried a diagnosis of AS and the others (Autism Moderate Language Functioning, A-moderateL group) a diagnosis of HFA. In addition, all 6 in the A-moderateL group experienced early language delay whereas none in the A-highL group experienced early language delay. Six subjects showing typical development (Normal Controls, NC, group) were also tested (*M* = 26.3 years; *range* = 23 to 34 years) to provide a standard for comparison purposes, as such detailed comparative acoustic analyses of topic and focus do not exist for English. None of the participants in the NC group had a family member diagnosed with ASD. All participants were monolingual English-speakers and the groups were matched in age ($F < 1$). The A-moderateL group performed significantly worse on receptive vocabulary than the NC group ($p = .003$) and A-highL ($p = .006$) groups. NC and A-highL groups did not differ ($p = .95$) by post-hoc Tukey's HSD tests (Table 1, see next page).

Table 1. *Demographic and background information by group*

	Control		A-highL		A-moderateL	
	Age (years)	PPVT	Age (years)	PPVT	Age (years)	PPVT
Individuals scores	23	111	24	100	24	80
	24	104	32	109	29	96
	25	111	17	104	18	82
	25	120	18	104	27	94
	27	94	18	114	29	98
	34	100	30	101	33	85
Mean (SD)	26.3 (4.0)	106.7 (9.2)	23.2 (6.6)	105.3 (5.3)	26.7 (5.2)	89.2 (7.8)

Note. PPVT = Peabody Picture Vocabulary Test (standard scores)

2.2 Materials and procedure

The research was approved by the McMaster University Research Ethics Board and conformed to the principles set out in the Canadian Tri-Council Ethics Policy. All participants gave informed consent. Testing lasted approximately one hour and took place in an acoustically treated room. Participants received a debriefing statement after completing the study.

Participants were tested individually playing the “Under the Shape” game (Chen, 2011), in which they were asked questions about pictures presented on a computer. Their verbal responses were recorded for offline acoustic analysis. This task measured how participants vary prosody according to two variables, *information structure* (topic/focus) and *sentence position* (initial/final), adapted from Chen (2011) for use with children and adults. This task was administered on an Acer Notebook using Microsoft Office PowerPoint. Responses were recorded in Sound Studio 3 (Felt Tip Incorporated, 2009) and saved as .wav files at a 44.1 kHz sampling rate with 16 bit resolution using a Mac iBook G4. A microphone (D770 Emotion AKG) was connected to the iBook using a US-122 USB Audio/MIDI Interface. Participants were seated about two inches away from the microphone.

During the *Familiarization Phase*, participants were told that they would see pictures of people, animals and objects performing different actions. They were asked to report aloud what they saw on the screen (e.g., “rabbit”), when they were shown a picture. This phase included 30 pictures presented in a fixed order and took about 2 minutes to complete. The purpose was to ensure that participants could identify and use a consistent label for each picture. Participants were asked to remember these labels as they would see the same pictures in the next phase of the game.

During the *Experimental Phase*, the "Under the Shape” game was played. Two referents, which could be people, animals or objects, were presented on the screen at the same time but one was covered by an opaque rectangle. The experimenter posed a *who* or a *what* question. When the experimenter pressed a button on the keyboard, the rectangle was removed and the participant was then able to answer the experimenter’s question (see Figure 1 A and B, see next page).

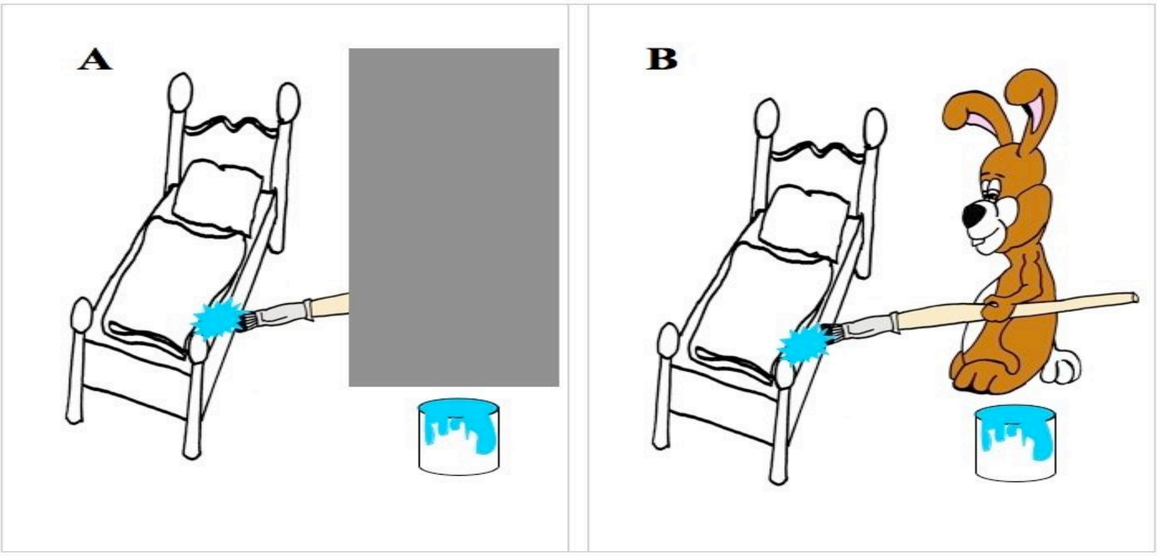

Figure 1. *Example trial of initial focus and final topic.* A. Experimenter: "Look! A bed. [shown picture of a bed with blue paint on it] It looks like someone is painting the bed. *Who* is painting the bed?" [shape disappears to reveal a picture of a rabbit holding a brush next to a paint can] B. Participant: "The rabbit is painting a bed."

This procedure measured how participants converse with a live speaker. The experimenter received training so that all questions were asked using the same prosody, with prominence placed on the first word, which was either *who* or *what*.

Responses to *who* and *what* question types differed in terms of whether the new information (focus) occurred in the sentence-initial position (subject) and the given information (topic) in the sentence-final (object) position or vice versa. Note, however, that the subject was always at the beginning of the sentence and the object at the end, regardless of which was the focus in terms of containing new information. For example, when "*WHO* is painting the bed?" (see Figure 1 A and B) was asked, the new information (focus) occurred in the initial position, "The *RABBIT* is painting the bed". Conversely, when "*WHAT* is the rabbit painting" (see Figure 2 A and B, see next page) was asked, the

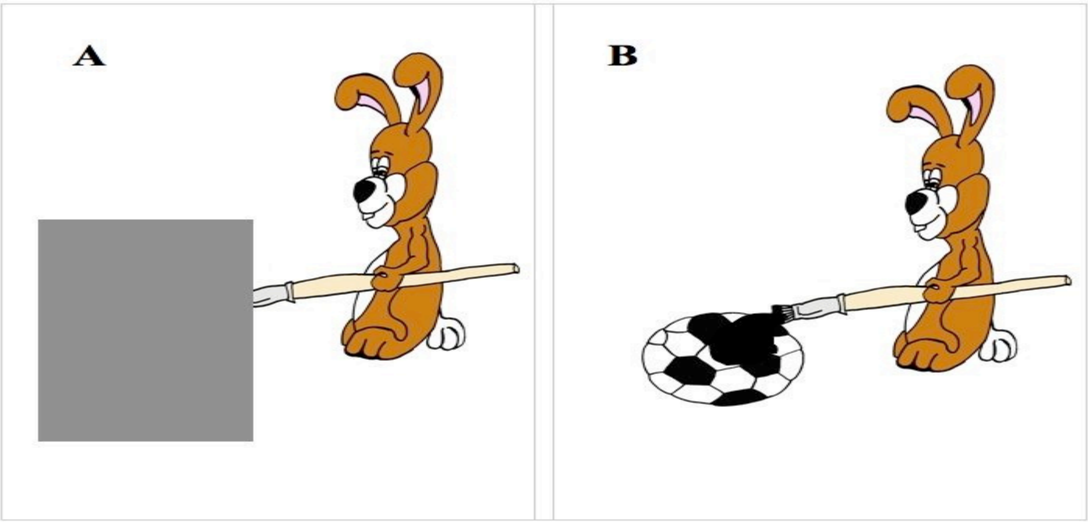

Figure 2. *Example trial of initial topic and final focus.* A. Experimenter: "Look! A rabbit. [shown picture of a rabbit holding a brush] It looks like the rabbit is painting something. *What* is the rabbit painting?" [shape disappears to reveal a picture of a ball] B. Participant: "The rabbit is painting a ball."

new information (focus) occurred in the final position, "The rabbit is painting the *BALL*". For each sentence position (initial/final), all nouns were used in topic and focus contexts in order to ensure that the acoustic analyses compared the same words across different contexts. To avoid boredom, every combination of subject and object nouns occurred only once during the experiment. Participants were required to respond to all questions using a full sentence. This response format ensured that each sentence contained a subject in the sentence-initial position and an object in the sentence-final position. Following four practice trials, participants completed 22 trials in the experimental phase, with equal numbers of *who* and *what* questions.

2.3 Acoustic annotation

Prior to acoustic analysis, we annotated the shape of the pitch contour in the subject and object words of the responses. Note that although strictly speaking we were interested in different emphasis between subject and object phrases, we analyzed the noun in each phrase, so we will refer to subject and object words. We found that these words were usually spoken with a rise-fall contour (84% of words), although they differed in the size (range) of the rise and fall. Thus, for the pitch analysis, we chose to examine range-rise (i.e., the difference between the peak and the preceding lowest pitch value) and range-fall (i.e., the difference between the peak and the proceeding lowest pitch value). In cases where there was only a fall with no preceding rise (7% of words), the rise was given a value that matched the fall (range-rise of zero). In cases where there was only a rise with no subsequent fall (9%), the fall was given a value that matched the rise (range-fall of zero). We also measured word duration.

The subject and object words were acoustically annotated by examining the waveform using the wide-band spectrum and pitch track in Praat 5.1.0.7 (Boersma & Weenink, 2009) and checked for octave errors by comparing visual displays of pitch tracks with auditory perceptions. The data were coded by the first author after receiving sufficient training from the second author. All data were checked independently by the second author for both accuracy and consistency and corrections were made by the two transcribers together. Three F0-related landmarks were labeled in each word, as illustrated in Figure 3 (see next page):

- Beginning F0 minimum: the initial lowest pitch in the subject noun (L1) and in the object noun (L4).

- F0 maximum: the highest pitch in the subject noun (H1) and in the object noun (H2) before the beginning of the pitch fall.
- Final F0 minimum: the lowest pitch reached following the F0 maximum in the subject noun (L2) and in the object noun (L5).

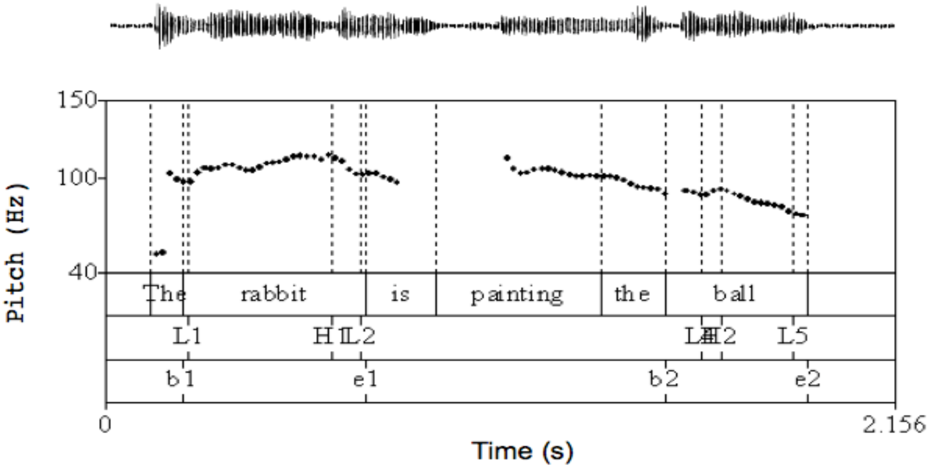

Figure 3: *Acoustic Analysis*. The sentence ‘The rabbit is painting the ball’ was produced as an answer to the question ‘what is the rabbit painting?’ by a speaker with A-highL . The landmarks in the subject noun ‘rabbit’ and the object noun ‘ball’ are the following: F0 minimum in the rising portion (L1/L4), F0 maximum (H1/H2), F0 minimum in the falling portion (L2/L5), beginning of the word (b1/b2), and end of the word (e1/e2).

When labeling the F0-related landmarks, we discarded micro-prosodic effects by searching for the highest F0 after the first three to five periods of the accented vowel and the lowest F0 before the voice started to fade out towards the end of the word. Octave errors were observed occasionally in the region where the F0 minimum was expected because of the transition from one phoneme to another and creaky voice. These errors were manually corrected after the F0 values at the H and L landmarks were automatically extracted.

Further, two segmental landmarks were labeled in each noun:

- The beginning of the word: b1 and b2 marking the onset of the first cycle in the waveform of the word-initial phoneme in the subject noun and in the object noun, respectively.
- The end of the word: e1 and e2 marking the offset of the last cycle in the waveform of the word-final phoneme in the subject noun and in the object noun, respectively.

Three measurements were then obtained for each noun:

- Range-rise: H1 - L1 for subject nouns and H2 - L4 for object nouns (measured in semitones or 1/12 octaves).
- Range-fall: H1 - L2 for subject nouns and H2 - L5 for object nouns (measured in semitones or 1/12 octaves).
- Word duration: Time_{e1} - Time_{b1} for subject nouns and Time_{e2} - Time_{b2} for object nouns (measured in seconds).

3. Statistical analysis and results

First, an Analysis of Variance (ANOVA) was conducted on absolute pitch to determine whether all groups used similar initial pitch levels across sentence position. The absolute pitch was operationalized as the lowest pitch preceding the pitch peak in each noun (L1 in the subject noun and L4 in the object noun). In the analysis, L1 of each subject noun and L4 of each object noun served as the dependent variable, sentence position (subject, object) as a within-subjects variable, and group (A-highL, A-moderateL, NC) as a between-subjects variable.

An ANOVA was also conducted with absolute duration to determine whether all groups used similar word durations across sentence position. In the analysis, word duration ($time_{e1}$ - $time_{b1}$ for subject nouns and $time_{e2}$ - $time_{b2}$ for object nouns) served as the dependent variable, group (A-highL, A-moderateL, NC) as a between-subjects variable, and sentence position (subject, object) as a within-subjects variable.

To examine information structure, ANOVAs were conducted with each of the following as the dependent measure: subject word range-rise, subject word range-fall, subject word duration, object word range-rise, object word range-fall, and object word duration. Each ANOVA was conducted with word (22 word pairs) and information structure (topic, focus) as within-subject variables and group (A-highL, A-moderateL, NC) as a between-subjects variable. We then conducted two types of planned pair-wise comparisons. We used non-parametric tests because of our relatively small sample size and fairly large within-group variability. First, we used Mann-Whitney U tests to compared between groups as to whether or not they differed in range-rise, range-fall, and duration for topic and focus separately. Second, and most importantly, we wanted to determine whether each group distinguished between topic and focus words. For this we conducted planned Wilcoxon signed-rank tests for each of our dependent measures. Finally, we tested whether there were significant Pearson correlations between our measure of language (PPVT) and each dependent variable for our entire sample ($n = 18$):

subject word range-rise, subject word range-fall, subject word duration, object word range-rise, object word range-fall, and object word duration.

When measuring how acoustic features are varied across topic and focus, it is important that the same words are compared. This is because the intrinsic pitch of vowels causes some words to have larger pitch ranges than others, and different segmental markup causes some words to be longer in duration than others. For the “Under the Shape” game (Chen, 2011), some participants occasionally used different labels on different trials for the same object (e.g., “bunny” and “rabbit”), an error that was made on a total of 19 out of 396 word pairs (4.8 %). These cells were replaced with the mean for that word for that particular group given that replacing up to 5% of data in this manner has been found to be acceptable (Rubin, Witkiewitz, St. Andre, & Reilly, 2007).

3.1 Pitch and duration

The ANOVA conducted on absolute pitch revealed a main effect of sentence position, $F(1, 15) = 42.634$, $p < .001$, $\eta^2 = .74$, with pitch falling from sentence-initial (M = 119.34 Hz, SEM = 4.55 Hz) to sentence-final words (declination) (M = 104.09 Hz, SEM = 3.74 Hz), but no main effect of group, $F(2, 15) = 2.870$, $p = .09$ ($\eta^2 = .28$). There was also no significant interaction between group and sentence position ($F < 1$), suggesting no overall differences in pitch range across the sentences.

The ANOVA conducted on absolute duration revealed a main effect of sentence position, $F(1, 15) = 6.287$, $p = .024$, $\eta^2 = .30$, with shorter durations for the sentence-initial (M = .328 sec, SEM = .012 sec) than for the sentence-final words (M =.355 sec, SEM = .015 sec), but no main effect of group, $F(2, 15) = 1.870$, $p = .20$. There was no significant interaction between group and sentence position ($F < 1$), indicating no overall differences between groups in duration and suggesting similar durational variation across the sentences.

3.2 Sentence-initial (subject)

3.2.1 Initial range-rise

In the initial (subject) position, the ANOVA on range-rise revealed no significant effects (Figure 4 A, see page 113). Planned Mann-Whitney tests revealed that for topic words, the A-moderateL group used a significantly smaller range rise than the NC ($U = 2.00$, $p = .010$, $r = .74$) and A-highL ($U = 2.00$, $p = .010$, $r = .74$) groups. There were no significant differences across groups for focus words.

Planned Wilcoxon signed-rank tests found no significant differences in range-rise for any group between topic and focus words (individual data is shown in Figure 4 B, see page 113).

In sum, although the A-moderateL group used a smaller range-rise for topic words, there was no significant difference in range-rise across groups with respect to use of information structure, with none of the groups using this initial range-rise to mark information structure.

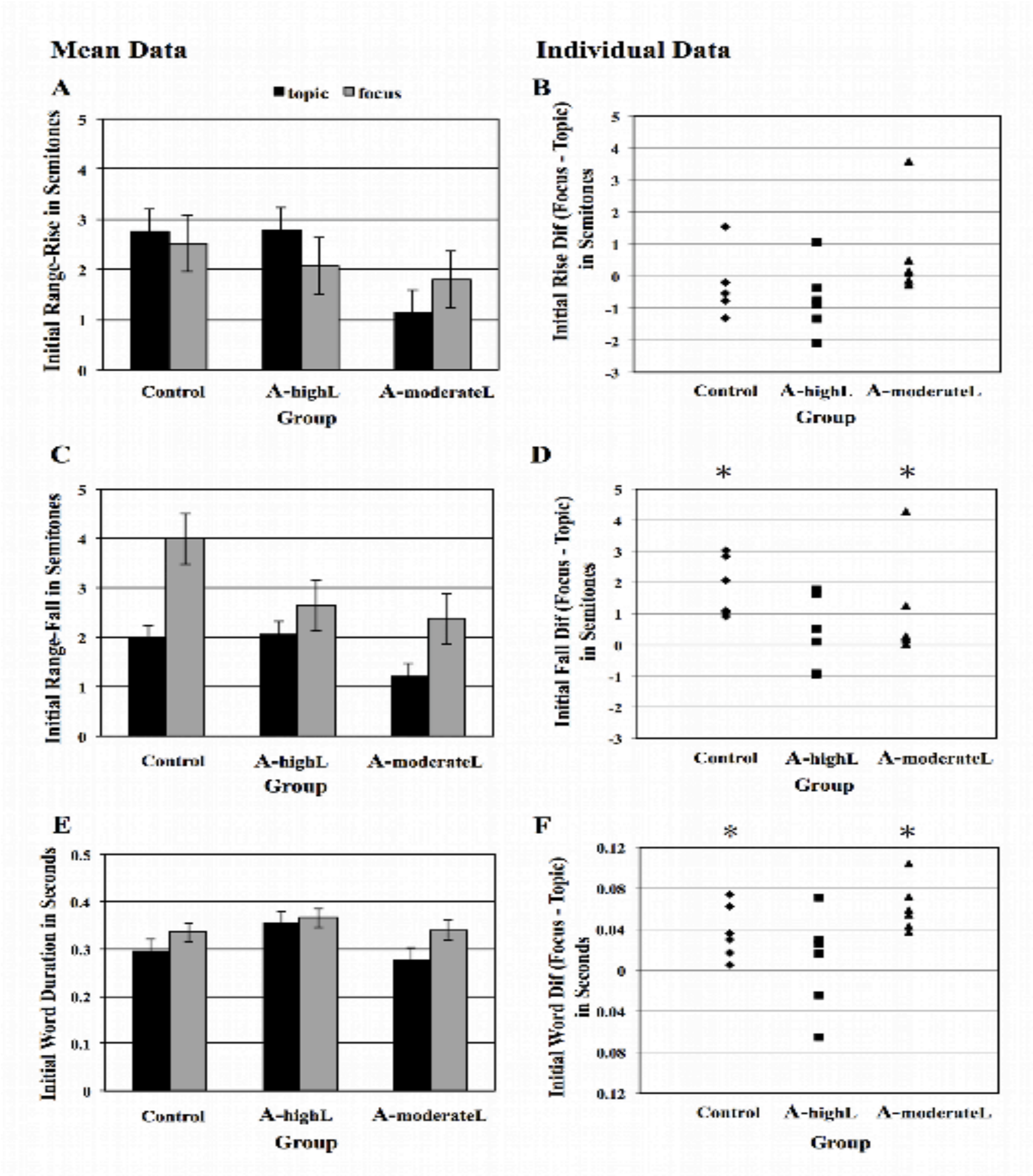

Figure 4. *Sentence-initial results.* A. Mean range-rise and standard error by group. B. Individual data for range-rise difference (focus - topic) by group. Note that no difference between topic and focus is represented by the zero line. C. Mean range-fall and standard error by group. D. Individual data for range-fall difference (focus - topic) by group. E. Mean word duration and standard error by group. F. Individual data for duration difference (focus - topic) by group. * $p < .05$

3.2.2 Initial range-fall

The ANOVA on range-fall revealed significant main effects of information structure, $F(1, 15) = 17.31$, $p = .001$, $\eta^2 =.54$ and of group, $F(2, 15) = 3.56$, $p = .05$, $\eta^2 = .32$ (see Figure 4 C, see page 113). Post-hoc tests using Tukey's HSD showed that the main effect for group was due to a significantly smaller range-fall overall in the A-moderateL compared to the NC group ($p = .04$).

Planned Mann-Whitney tests revealed that for topic words, the A-moderateL group used a significantly smaller range-fall than the NC ($U = 5.00$, $p = .04$, $r = .60$) and A-highL ($U = 5.00$, $p = .04$, $r = .60$) groups. For focus word, the NC group used a significantly larger range-fall than the A-highL group ($U = 4.00$, $p = .03$, $r = .65$) and there was a trend for the NC group to use a larger range-fall than the A-moderateL group ($U = 9.00$, $p = .15$, $r = .42$). This is consistent with the greatest marking of information structure by the control group.

Planned Wilcoxon signed-rank tests revealed significantly larger range-falls for focus than topic in the NC ($p = .03$, $d = .96$) and A-moderateL ($p = .03$, $d = .64$) groups, but not in the A-highL ($p = .46$) group (see Figure 4 D, see page 113).

In sum, the A-moderateL group used a smaller pitch range overall, and particularly for topic words, compared to the NC and A-highL groups. On the other hand, the NC and A-moderateL groups marked information structure by using larger range-falls for focus compared to topic words, whereas those in the A-highL group did not.

3.2.3 Initial duration

The ANOVA on word duration revealed a significant main effect of information structure, $F(1, 15) = 20.01$, $p < .001$, $\eta^2 = .57$, with longer word durations for focus than topic, no main effect of group, $F(2, 15) = 1.97$, $p = .17$, and a significant interaction between information structure and group, $F(2, 15) = 3.57$, $p = .05$, $\eta^2 = .32$ (see Figure 4 E, see page 113).

Planned Mann-Whitney tests revealed that the A-highL group used a longer duration for topic words than the A-moderateL group ($U = 5.00$, $p = .04$, $r = .60$) but there were no significant effects for focus words.

Planned Wilcoxon signed-rank tests revealed a significant difference between topic and focus for the NC ($p = .03$, $d = .43$) and A-moderateL ($p = .03$, $d = .70$), groups, but not for the A-highL ($p = .46$), group (see Figure 4 F, see page 113).

In sum, the NC and A-moderateL groups used word duration to mark information structure, but the A-highL group did not.

3.2.4 Initial correlations with PPVT

Finally, across the entire sample, there were significant (or approaching significant) Pearson correlations between PPVT and the size of the sentence-initial range-rise (subject), $r = .48$, $p = .04$, and range-fall, $r = .46$, $p = .06$, but not between PPVT and duration, $p > .23$ (Table 2, see next page), again suggesting that difference in language ability underlies the different prosodic strategies.

Table 2. *Pearson Correlations Between Receptive Vocabulary and Prosody (n = 18 speakers)*

	Receptive vocabulary (PPVT)
Sentence-initial range-rise	.48*
Sentence-initial range-fall	.46+
Sentence-initial duration	.30
Sentence-final range-rise	.01
Sentence-final range-fall	.48*
Sentence-final duration	.27

Note. * p <.05; + p <.06, PPVT = Peabody Picture Vocabulary Test (standard scores)

3.3 Sentence-final (object)

3.3.1 Final range-rise

In the final (object) position, the ANOVA on range-rise revealed no significant effect of, or interactions involving, group. However, there was a significant main effect of information structure, $F(1, 15) = 14.21$, $p = .002$, $\eta^2 = .49$, with a larger range-rise for focus than for topic (see Figure 5 A, see page 117).

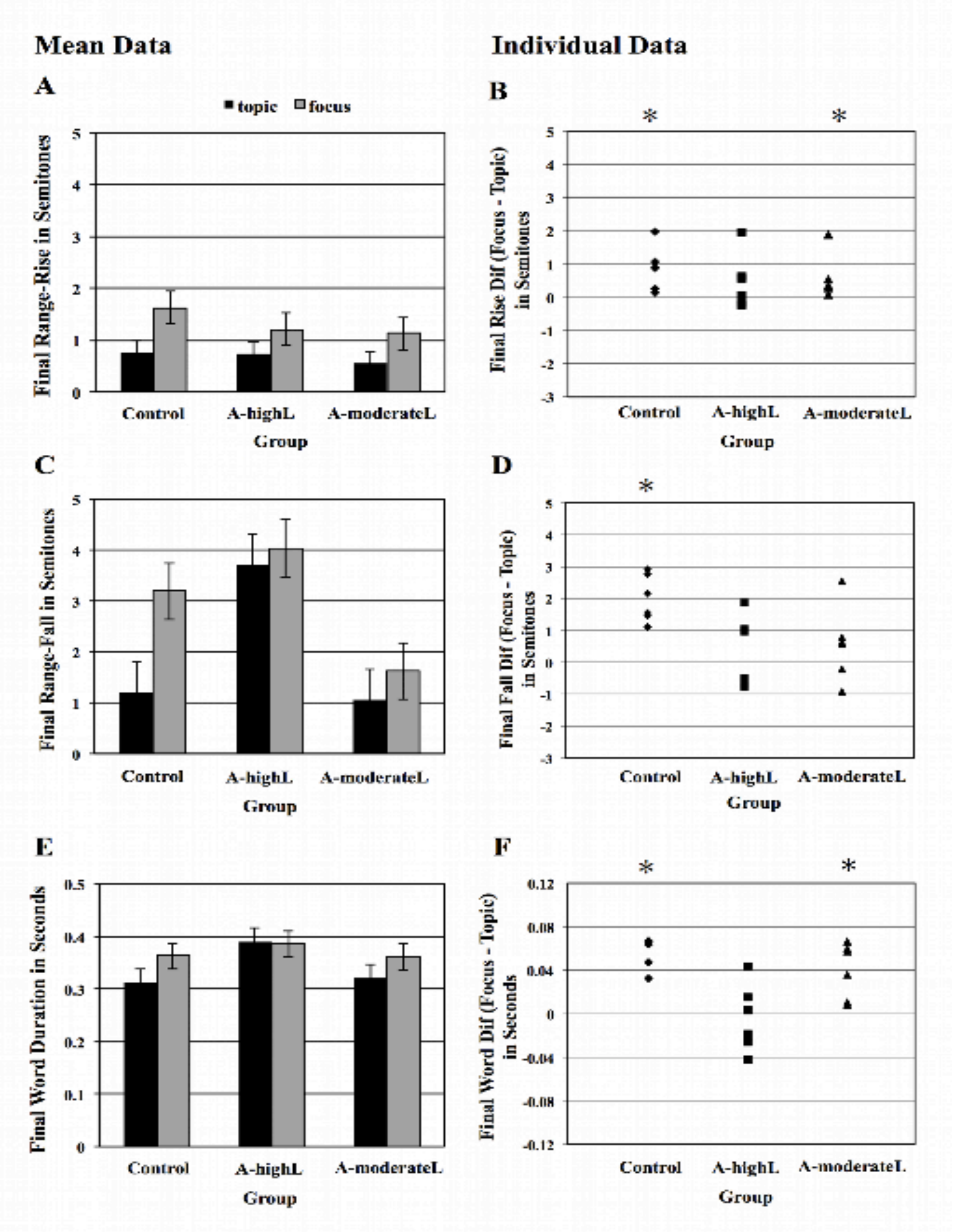

Figure 5. *Sentence-final results.* A. Mean range-rise and standard error by group. B. Individual data for range-rise difference (focus - topic) by group. Note that no difference between topic and focus is represented by the zero line. C. Mean range-fall and standard error by group. D. Individual data for range-fall difference (focus - topic) by group. E. Mean word duration and standard error by group. F. Individual data for duration difference (focus - topic) by group. * $p < .05$

Planned Mann-Whitney tests revealed no differences between groups. Planned Wilcoxon signed-rank test results revealed significant differences between topic and focus for the NC ($p = .03$, $d = .48$) and A-moderateL ($p = .03$, $d = .49$) groups, but not for the A-highL ($p = .25$) group (see Figure 5 B, see page 117).

In sum, the NC and A-moderateL groups used range-rise to mark information structure, using a larger range-rise for focus than for topic words, whereas those with A-highL did not.

3.3.2 Final range-fall

The ANOVA on range-fall revealed a significant main effect of information structure, $F(1, 15) = 15.83$, $p = .001$, $\eta^2 = .51$, with a larger range-fall for focus than for topic. There was also a significant main effect of group, $F(2, 15) = 5.75$, $p = .01$, $\eta^2 = .43$, and an interaction between information structure and group, $F(2, 15) = 4.67$, $p = .03$, $\eta^2 = .38$ (see Figure 5 C, see page 117). Post-hoc tests using Tukey's HSD revealed an overall larger pitch range in the A-highL compared to A-moderateL group ($p = .01$).

Planned Mann-Whitney tests revealed that for topic words, the A-highL group showed a significantly larger range-fall compared to the NC ($U = 0.00$, $p = .004$, $r = .83$) and A-moderateL group ($U = 0.00$, $p = .004$, $r = .83$), consistent with overall exaggerated pitch excursions in the A-highL group. The A-moderateL group showed a significantly smaller range-fall for focus words compared to the NC ($U = 5.00$, $p = .04$, $r = .60$) and A-highL ($U = 3.00$, $p = .016$, $r = .69$) groups, consistent with smaller pitch excursions in the A-moderateL group.

Planned Wilcoxon signed-rank tests revealed a significantly larger range-fall for focus than topic for the NC group (p = .03, d = 1.2), but not for the A-highL (p = .34) and A-moderateL (p = .34) groups (see Figure 5 D, see page 117).

In sum, the A-highL group used relatively large pitch ranges, consistent with a singsong quality, particularly for topic, which should be deemphasized in the final position, whereas the A-moderateL group used relatively small pitch ranges, consistent with a monotone quality. Importantly, the NC group used a larger range-fall to mark sentence-final focus compared to topic words, whereas the A-highL and A-moderateL groups did not.

3.3.3 Final duration

For word duration, there was a significant main effect of information structure, $F(1, 15) = 24.17, p < .001, \eta^2 =.62$, with longer word durations for focus than for topic (see Figure 5 E, see page 117). The main effect of group was not significant, $F(2, 15) = 1.23, p = .32$, but there was a significant interaction between information structure and group, $F(2, 15) = 8.17, p = .004, \eta^2 = .52$.

Planned Mann-Whitney test revealed no significant differences between groups. Planned Wilcoxon signed-rank tests revealed that the difference between topic and focus was significant for the NC (p = .03, d = .47) and A-moderateL (p = .03, d = .38) groups, but not for the A-highL group (p = .25) (see Figure 5 F, see page 117).

In sum, the NC and A-moderateL groups used word duration to mark information structure, but the A-highL group did not.

3.3.4 Final correlations with PPVT

Finally, across the entire sample (n = 18) there was a significant Pearson correlation between PPVT and the size of the sentence-final (object) range-fall, $r = .48$, $p = .04$, although not between PPVT and the size of the range-rise, $p > .97$, or duration, $p > .27$ (Table 2, see page 116), again suggesting that difference in language ability underlie the different prosodic strategies.

4. Discussion

Even with only six participants in each of the subgroups, with the detailed acoustic analyses we performed, we found robust and marked differences in performance between those with ASD with stronger language skills (A-highL group) compared to those with weaker language skills (A-moderateL group). Regardless of information structure, compared to controls, we found larger pitch ranges for those with ASD with strong language skills, and smaller pitch ranges for those with moderate language skills. It is worth noting that these differences cannot be explained by potential differences in overall pitch height as the three groups did not differ significantly in initial absolute (starting) pitch. The small pitch range of those with ASD and moderate language skills is consistent with a monotone quality to their speech, whereas the large pitch range of those with ASD and stronger language skills is consistent with a singsong quality. It would be interesting to test this notion further in future studies to see whether speech with these different prosodic pitch characteristics is indeed perceived as monotone and singsong, respectively. With respect to duration, we did not find any significant group differences in how this acoustic feature was varied in general when information structure was not considered. Thus, pitch appears to be the primary contributor to general abnormal prosody in ASD, a finding that could help to inform future remediation programs in speech and language. Our finding that individuals with ASD could be divided into subgroups who use either a smaller or a larger pitch range than normal is consistent with previous reports of heterogeneity in this regard (e.g., Baltaxe et al., 1984; Green & Tobin, 2009). Furthermore, our results extend previous studies by indicating that in ASD, use of a smaller pitch range is associated with moderate language skill, whereas use of a larger pitch range is associated with high language skill.

With respect to communication, an examination of the details of how information is marked is critical. We found that controls used pitch to mark information structure in both sentence positions, with larger pitch falls for focus than topic words in both sentence-initial (subject) and sentence-final (object) positions, and larger pitch rises for focus than topic words in sentence-final positions. To the extent that the A-moderateL group varied pitch, they tended to mark information structure similarly to controls, although their pitch excursions were smaller than those of controls (about one semitone, or 1/12 octave smaller on average) and they did not show significantly larger pitch falls for focus than topic words in sentence final positions. Marking of information in sentence-final positions does appear to develop later than in sentence-initial positions (Chen, 2011), perhaps because it goes against the natural tendency for sentences in English to stress the initial subject word more than the final object word, all else being equal. It is also possible that the failure of the A-moderateL group to use pitch to mark information structure in the sentence-final position reflects working memory constraints and difficulty in integrating acoustic and linguistic structure over a sentence. In any case, although those with ASD and moderate language skills marked information to a lesser extent than controls, they did mark information structure appropriately. On the other hand, those in the A-highL group did not vary pitch significantly as a function of information structure at any position in the sentence, despite their general use of large pitch variation. Given that the extent of pitch fall is an important marker of information structure in West Germanic languages (Chen, 2009; Hanssen et al., 2008), those with ASD with higher language skills are not using prosody well to communicate with their conversational partners.

With respect to the marking of information structure using duration, the control and A-moderateL groups used longer word durations for focus than for topic words in both sentence positions, but the A-highL group did not. We found considerable within-group variability in how speakers in the A-highL group used duration, although we could not find any characteristics that correlated with duration differences across topic and focus. In general, the results for duration are consistent with those for pitch in that those with ASD with better language skills demonstrate the least use of prosody to convey information structure.

Our finding of better communication in terms of marking information structure in those with ASD with moderate language skills, compared to in those with high language skills, is consistent with a previous report using eye tracking to determine communicative competence. Norbury et al. (2009) found that teenagers with ASD with poorer language skills were similar to typically-developing teenagers in spending an appropriate proportion of time viewing the eyes and mouths of peers interacting in video recordings, whereas those with ASD with better language skills spent less time viewing the eyes and were slower to fixate on the eyes than the other groups. Together, the present results and those of Norbury and colleagues intriguingly suggest that although those with ASD with higher language skills obviously have some advantages over those with poorer language skills, basic automatic communication strategies of where to look and how to vary pitch and duration in utterances may be defining characteristics of their communication impairments. On the other hand, the communication difficulties of those with ASD with poorer language skills might have a different origin. Individuals in this category appear relatively unimpaired in terms of the automatic strategies of where to look and how to use pitch and duration for communicative intent. Their communication difficulties may originate in poor language skills in general rather than specific difficulties in prosodic use related to information structure.

It is also of interest that those in the A-highL group had diagnoses of Asperger's whereas those in the A-moderateL had diagnoses of high functioning autism. However, the lack of consistent differences between those with Asperger’s and high functioning autism has led to the proposal to remove this distinction in the DSM-5. Of the research that finds differences between ASD subgroups, some have pointed out that there might be as many as 6 definitions currently being used for AS (Diehl et al., 2009). These definitions range from those with AS having milder symptoms of ASD to those with AS not experiencing an early language delay in contrast to those with HFA. These differences in definition can make comparison between studies difficult if not impossible. We argue that it is better to use a well-defined criterion, such as language ability, to distinguish the groups.

It is possible, nonetheless, that those in the A-highL group, who also had a diagnosis of Asperger's, had more explicit knowledge of language and that this may have actually impaired natural use of prosody. In thinking about alternative explanations for the results, it is also interesting to consider the question of whether or not there was an early language delay and, if so, whether it resulted in different early experiences. All of those in the A-moderateL group experienced early language delay whereas none of those in the A-highL group did so. Thus, those in the A-moderateL group were likely diagnosed early and likely received early speech intervention, whereas those in the A-highL group were likely diagnosed later and likely did not receive speech intervention (Foster & King, 2003; Howlin & Asgharian, 2007). It is therefore possible that the lack of early language delay in AS may make it harder to detect problems with language abilities early on, including the general use of prosody and marking of information structure that are often reported among those with HFA. Although speech intervention rarely targets prosody (Bellon-Harn, Harn, & Watson, 2007; McCann, Peppé, Gibbon, O'Hare, & Rutherford, 2007; Paul et al., 2005), it may provide experience with the systematic variation in acoustic cues related to listener comprehension. From the present data, it is not possible to determine to what extent the prosodic differences we observed between the A-highL and A-moderateL groups is due to different genetic etiologies or different experiences with developmental interventions. However, our research serves as an important starting point for understanding how different prosodic problems may arise in those with ASD.

Importantly, the present study also contributes to the finding that the prosodic abnormalities identified in children with ASD persist into adulthood (Diehl et al., 2009; Paul et al., 2005; Shriberg et al., 2001). Given that atypical prosody in adults with ASD impacts both their personal lives, in terms of making and keeping friends, and their professional lives, in terms of gaining and keeping employment (Paul et al., 2005; Van Bourgondien & Woods, 1992), further research on the extent to which appropriate information-marking can be trained in children and adults is critical.

The present study has some limitations. First, once subgroups were formed based on language ability, the sample size was not large and an outlier analysis was not possible. However, in the case of initial range-rise and initial range-fall, one subject in the A-moderateL group appears to show a larger difference between focus and topic than others in his group. Despite this, robust and consistent differences were found across groups in the use of pitch and duration both overall and in marking information structure, but a replication with a larger sample would be good. A second limitation is that semi-spontaneous speech was used rather than spontaneous speech. While this had the critical advantage of enabling us to compare the same words across topic and focus contexts and sentence-initial and sentence-final positions, replication of these results should be performed with a large sample of spontaneously speech. A third limitation is that we did not include an extensive assessment of language functioning, although our measure of vocabulary can be used as a proxy. Given the robust differences we found between those with ASD with high and those with more moderate language abilities, it would be interesting for future studies to replicate our findings and also to determine whether there are different relationships between prosodic use and different language skills, such as articulation, phonological processing, vocabulary, grammatical and semantic skills. It would also be of interest to examine speakers with ASD from languages in which information structure is primarily marked by overt syntactic operations.

Regardless of the origin of the differences, both the A-highL and A-moderateL groups used abnormal prosody, which would affect their ability to communicate effectively. Although those with moderate language skills used pitch and duration cues to mark information structure, they varied pitch to a lesser extent than controls, and this would likely give the impression that they were uninterested in conversation. Indeed, in real communicative contexts, such use of monotonous speech might override the fact that those in the A-moderateL group mark information structure appropriately for the most part. On the other hand, those with high language skills used more prosodic variation relative to control and those in the A-moderateL group (average size of range-fall across sentence positions was approximately 0.5 semitones and 1.5 semitones larger than control and A-moderateL groups, respectively), but the way that they did so with respect to information structure was not useful to listeners. This use of prosody is likely distracting because the indiscriminant use of large pitch excursions does not direct the listener's attention to focus words. It remains for future research to document the precise effects of different prosodic abnormalities related to information structure on typical listeners, but it is evident that abnormal prosody can have serious consequences for social communication (Peppé et al., 2006; Peppé et al., 2007; Wells, Peppé, & Goulandris, 2004).

In conclusion, we conducted detailed analyses of prosodic pitch and duration usage in adults with ASD and found that compared to controls, those with high language functioning used exaggerated prosody in general but did not use pitch and duration communicatively to convey information structure, whereas those with moderate language function varied prosody less in general compared to controls, but did use pitch and duration communicatively to convey information structure. These results suggest that at least some of the heterogeneity of prosodic use among adults with ASD is related to level of language functioning. Regardless of subgroup differences, because prosodic cues to information structure are largely processed without conscious awareness in typical listeners, inappropriate use of prosody may be interpreted at a conscious level by listeners as a lack of interest in being a good conversational partner. Such speakers will likely be judged as less engaged in communication, which could make it more difficult for them to compete in job interviews and form lasting friendships. It is therefore important to understand the details of prosodic use in different subgroups with ASD in order to inform remediation strategies.

References

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Asperger, H. (1944). Die autistischen Psychopathen im Kindesalter. *Archiv fur Psychiatrie und Nervenkrakheiten 117,* 76-136.

Baltaxe, C. (1984). Use of contrastive stress in normal, aphasic, and autistic children. *Journal of Speech and Hearing Research 24,* 97-105.

Baltaxe, C. A. M., and Guthrie, D. (1987). The use of primary sentence stress by normal, aphasic and autistic children. *Journal of Autism and Developmental Disorders 17,* 255-271.

Baltaxe, C. A. M., and Simmons, J. Q. (1985). "Prosodic development in normal and autistic children," in *Communication Problems in Autism,* ed. E. Schloper and G. Mesilov, (New York: Plenum), 95-125.

Baltaxe, C., Simmons, J. Q., and Zee, E. (1984). "Intonation patterns in normal, autistic, and aphasic children," in *Proceedings of the Tenth International Congress of Phonetic Sciences,* eds. A. Cohen and M. P. R. Broecke (Dortrecht-Holland: Foris Publications), 713-718.

Baron-Cohen, S., and Staunton, R. (1994). Do children with autism acquire the phonology of their peers? An examination of group identification through the window of bilingualism. *First Language, 14* 241-248.

Bellon-Harn, M. L., Harn, W. E., and Watson, G. D. (2007). Targeting prosody in an eight-year-old child with high-functioning autism during an interactive approach to therapy. *Child Language Teaching & Therapy, 23* 157-179.

Birch, S., and Clifton, C. (1995). Focus, accent and argument structure: Effects on language comprehension. *Language and Speech 38,* 365-391.

Boersma, P., and Weenink, D. (2009). *Praat: Doing phonetics by computer* [Computer program]. Version 5.1.07, retrieved 16 August 2009 from http://www.praat.org/

Bonneh, Y. S., Levanon, Y., Dean-Pardo, O., Lossos, L., and Adini ,Y. (2011). Abnormal speech spectrum and increased pitch variability in young autistic children. *Frontiers in Human Neuroscience 4,* 1-7.

Brock, J., Norbury, C. F., Einav, S., and Nation, K. (2008), Do individuals with autism process words in context? Evidence from language-mediated eye-movements. *Cognition 108,* 896-904.

Chafe, W. (1976). “Givenness, contrastiveness, subject, topic, and point of view,” in *Subject and Topic,* ed. C. Li (New York: Academic Press), 25-55.

Chen, A. (2009). “The phonetics of sentence-initial topic and focus in adult and child Dutch,” in *Phonetics and Phonology: Interactions and Interrelations,* eds. M. Vigário, S. Frota and M. J. Freitas (Amsterdam: Benjamins), 91-106.

Chen, A. (2010). Is there really an asymmetry in the acquisition of the focus-to-accentuation mapping? *Lingua 120,* 1926-1939.

Chen, A. (2011). Tuning information packaging: Intonational realisation of topic and focus in child Dutch. *Journal of Child Language 38,* 1055-1083.

Chun, D. M. (1988). The neglected role of intonation in communicative competence and proficiency. *The Modern Language Journal 72*, 295-303.

Clark, H. H., and Haviland, S. E. (1977). “Comprehension and the given-new contract,” in *Discourse production and comprehension,* ed. R. O. Freedle (Norwood, NJ: Ablex Publishing), 1-40.

Cruttenden, A. (1997). *Intonation.* Cambridge: Cambridge University Press.

Diehl, J., Watson, D., Bennetto, L., McDonough, J., and Gunlogson, C. (2009). An acoustic analysis of prosody in high-functioning autism. *Applied Psycholinguistics, 30*, 385-404.

Dunn, L. M., and Dunn, L. M. (1997). *Peabody Picture Vocabulary Test* (3rd ed.). Circle Pines, MN: American Guidance Services.

Fay, W., and Schuler, A. L. (1980). *Emerging language in autistic children.* Baltimore: University Park Press.

Felt Tip Incorporated. (2009). *Sound Studio 3.* New York, NY: Felt Tip Incorporated.

Fine, J., Bartolucci, G., Ginsberg, G., and Szatmari, P. (1991). The use of intonation to communicate in pervasive developmental disorders. *Journal of Child Psychology and Psychiatry 32,* 771-782.

Fosnot, S. M., and Jun, S. (1999, August). *Prosodic characteristics in children with stuttering or autism during reading and imitation*. In *Proceedings of the 14th International Congress of Phonetic Sciences,* eds. J. J. Ohala and Y. Hasegawa (Dordrecht, the Netherlands: Foris), 1925-1928.

Foster, B., and King, B., H. (2003). Asperger syndrome: to be or not to be? *Current Opinion in Pediatrics 15*, 491-494.

Frith, U. (1991). *Autism and Asperger Syndrome.* Cambridge: Cambridge University Press.

Geschwind, D. H. (2009). Advances in Autism. *Annual Review of Medicine 60,* 367-380.

Green, H., and Tobin, Y. (2009). Prosodic analysis is difficult…but worth it: A study in high functioning autism. *International Journal of Speech-Language Pathology 11,* 308-315.

Grossman, R. B., Bemis, R. H., Skwerer, D. P., and Tager-Flusberg, H. (2010). Lexical and affective prosody in children with high-functioning autism. *Journal of Speech, Language, and Hearing Research 53,* 778-793.

Gussenhoven, C. (2004). The phonology of tone and intonation. Cambridge: Cambridge University Press.

Halliday, M. K. K. (1967). Intonation and grammar in British English. The Hague: Mouton.

Hanssen, J., Peters, J., and Gussenhoven, C. (2008, May). *Prosodic Effects of Focus in Dutch Declaratives.* In *Proceedings of the 4th Conference on Speech Prosody*, eds. P. A. Barbosa, S. Madureira, and C. Reis (Campinas, Brazil: Editora RG/CNPq), 609-612.

Hornby, P. A. and Hass, W. A. (1970). Use of contrastive stress by preschool children. *Journal of Speech and Hearing Research 13,* 359-399.

Howlin, P., and Asgharian, A. (2007). The diagnosis of autism and Asperger syndrome: findings from a survey of 770 families. *Developmental Medicine & Child Neurology, 41,* 834-839.

Hubbard, K., and Trauner, D. A. (2007). Intonation and emotion in autistic spectrum disorders. *Journal of Psycholinguistic Research 36,* 159 173.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2*, 217-250.

Kjelgaard, M., and Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes 16,* 287-308.

Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.

Lambrecht, K. (1994). Information structure and sentence form: Topics, focus, and the representations of discourse referents. Cambridge: Cambridge University Press.

Lord, C., and Paul, R. (1997). "Language and communication in autism," in *Handbook of autism and pervasive development disorders*, eds. D. J. Cohen and F. R. Volkmar (New York: John Wiley), 195-225.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., and Schopler, E. (1989). Autism diagnostic observation schedule: A standardized observation of communicative and social behavior. *Journal of Autism and Developmental Disorders 19,* 185-212.

Lord, C., Rutter, M., and Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders 24,* 659-685.

McCaleb, P., and Prizant, B. M. (1985). Encoding of new versus old information by autistic children. *Journal of Speech and Hearing Disorders 50,* 230-240.

McCann, J., and Peppé, S. (2003). Prosody in autistic spectrum disorders: A critical review. *International Journal of Language and Communication Disorders 38,* 325-350.

McCann, J., Peppé, S., Gibbon, F. E., O'Hare, A., and Rutherford, M. (2007). Prosody and its relationship to language in school-aged children with high-functioning autism. *International Journal of Language and Communication Disorders 42,* 682-702.

Müller, A., Höhle, B., Schmitz, M., and Weissenborn, J. (2006). Focus-to-stress alignment in 4- to 5-year-old German-learning children. "Language Acquisition and Development," in *Proceedings of GALA 2005, eds.* A. Belletti, E. Bennati, C. Chesi, E. Di Domenico, and I. Ferrari (Cambridge: Cambridge Scholars Press), 393-407.

Nadig, A., and Shaw, H. (2011). Expressive prosody in high-functioning autism: Increased pitch range and what it means to listeners. *Journal of Autism and Developmental Disorders,* April 29, Epublication ahead of print.

Nadig, A., Vivanti, G., and Ozonoff, S. (2009). Adaptation of object descriptions to a partner under increasing communicative demands: A comparison of children with and without autism. *Autism Research 2,* 334-347.

Nespor, N., and Vogel, I. (1986). *Prosodic Phonology*. Dordrecht: Foris Publications.

Nooteboom S. G., and Terken J. M. B. (1982). What makes speakers omit pitch accents? An experiment. *Phonetica 39*, 317-336.

Norbury, C. F. (2005). Barking up the wrong tree? Lexical ambiguity resolution in children with language impairments and autistic spectrum disorders. *Journal of Experimental Child Psychology 90,* 142-171.

Norbury, C. F., Brock, J., Cragg, L., Einav, S., Griffiths, H., and Nation, K. (2009). Eye-movement patterns are associated with communicative competence in autistic spectrum disorders. *Journal of Child Psychology and Psychiatry 50,* 834-842.

Paccia, J., and Curcio, F. (1982). Language processing and forms of immediate echolalia in autistic children. *Journal of Speech and Hearing Research 25,* 42-47.

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

Paul, R., Bianchi, N., Augustyn, A., Klin, A., and Volkmar, F. (2008). Production of syllable stress in speakers with autism spectrum disorders. *Research in Autism Spectrum Disorders, 2* 110-124.

Paul, R., Orlovski, S., Marchinko, H. and Volkmar, F. (2009). Conversational behaviors in youth with high-functioning autism and asperger syndrome. *Journal of Autism and Developmental Disorders, 39,* 115-125.

Peppé, S. J. E., Cleland, J., Gibbon, F. E., O'Hare, A. E., and Martínez-Castilla, P. (2011). Expressive prosody in children with autism spectrum conditions, *Journal of Neurolinguistics*, *24,* 41-53.

Peppé, S., McCann, J., Gibbon, F., O'Hare. A, and Rutherford, M. (2007). Receptive and expressive prosodic ability in children with high-functioning autism. *Journal of Speech, Language, and Hearing Research 50,* 1015-1028.

Peppé, S., McCann, J., Gibbon, F., O'Hare, A., and Rutherford, M. (2006). Assessing prosodic and pragmatic ability in children with high-functioning autism. *Journal of Pragmatics 38,* 1776-1791.

Prince, E. F. (1986). "On the syntactic marking of presupposed open propositions," in *Papers from the Parasession on Pragmatics and Grammatical Theory*, eds. A. Farley, P. Farley, and K. E. McCullough (22nd Regional Meeting in Chicago, IL: Chicago Linguistic Society), 208-222.

Roach, P. (2000). *English phonetics and phonology.* Cambridge: Cambridge University Press.

Rubin, L. H., Witkiewitz, K., St. Andre, J. and Reilly, S. (2007). Methods for handling missing data in the behavioural neurosciences: Don't throw the baby rat out with the bath water. *The Journal of Undergraduate Neuroscience Education 5,* A71-A77.

Schreibman, L., Kohlenberg, B. S., and Britten, K. R. (1986). Differential responding to content and intonation components of a complex auditory stimulus by nonverbal and echolalic autistic children. *Analysis and Intervention in Developmental Disabilities 6,* 109-125.

Sharda, M., Subhadra, T. P., Sahay, S., Nagaraja, C., Singh, L., Mishra, R., Sen, A., Singhal, N., Erickson, D., and Singh, N. C. (2010). Sounds of melody—pitch patterns of speech in autism. *Neuroscience. Letters 478*, 42-45.

Shriberg, L., Paul, R., McSweeney, J., Klink, A., Cohen, D., and Volkmar, F. (2001). Speech and prosody characteristics of adolescents and adults with high functioning autism and Asperger syndrome. *Journal of Speech, Language and Hearing Research 44*, 1097-1115.

Vallduví, E., and Engdahl, E. (1996). The linguistic realisation of information packaging. *Linguistics 34*, 459-519.

Van Bourgondien, M. E., and Woods, A. V. (1992). "Vocational possibilities for high-functioning adults with autism," in *High-functioning individuals with autism,* eds. E. Schopler and G. B. Mesibov (New York: Plenum Press), 227-239.

Van Lancker, D., Cornelius, C., and Kreiman, J. (1989). Recognition of emotional-prosodic meanings in speech by autistic, schizophrenic, and normal children. *Developmental Neuropsychology 5,* 207-226.

Wells, B., Peppé, S., and Goulandris, A. (2004) Intonation development from five to thirteen. *Journal of Child Language 31,* 749-778.

CHAPTER 4

GENERAL DISCUSSION

The research presented in this dissertation revealed a number of important findings. In chapter 2, we showed that adolescents with ASD had problems ignoring one stream of input while attending to another, integrating audio-visual information in speech, and categorizing the phonemic and rhythmic categories that are important to the language and musical system in their environment. In general, those with ASD tended to be more impaired on tasks involving speech than on tasks involving music, forming a profile by which we can begin to understand auditory processing in this disorder. In chapter 3, we showed differences in speech production patterns between those with Autism High Language Functioning (A-highL) who carried a diagnosis of Asperger's syndrome (AS) and those with Autism Moderate Language Functioning (A-moderateL) who carried a diagnosis of High-Functioning Autism (HFA). In particular, the A-moderateL group used a smaller, and the A-highL group used a larger, pitch range than controls. We also found that the A-highL group did not vary pitch range or duration to mark topic and focus, whereas the A-moderateL group did differentiate topic and focus, although to a lesser extent than controls. Taken together, these results inform us about the perception and production abilities of individuals with ASD relative to those who are typically developing.

Speech Processing Results

Although humans have the capacity to develop speech and language, realization of this capacity may be abnormal among those with ASD if they are not motivated to attend to speech sounds (e.g., Paul, Chawarska, Fowler, Cicchetti, & Volkmar, 2007). Reduced motivation to interact with others can impact enculturation, which is important for the development of efficient speech and language processing. Early neural development in ASD also appears to be atypical (Barnea-Goraly et al., 2004; Casanova, 2004; Casanova, Buzhoeveden, Switala, & Roy, 2002; Castelli, Frith, Happé, & Frith, 2002; Courchesne et al., 2001, Courchesne, Carper, & Akshoomoff, 2003; Courchesne & Pierce, 2005; Just, Cherkassky, Keller, & Minshew, 2004; Just, Cherkassky, Keller, Kana, & Minshew, 2007), which could impact normal speech and language development. In this research, we measured three key areas of speech processing, ability to filter speech in noise, specialization for native speech sound categories, and audio-visual integration of speech sounds. We found group differences in performance in all of these areas among those with ASD relative to controls.

1. Auditory Filtering. Filtering involves the ability to attend to one source of auditory information, despite the fact that there are other sources of information that are actively competing for processing. Being able to filter means that you can carry a conversation in a noisy environment, which is something that we do on a regular basis. Here we measured the signal-to-noise ratio needed to perceive sentences presented to one ear while ignoring simultaneous sentences to the other ear. We found that the ASD group showed poorer filtering than the control group, which persisted even after accounting for individual differences in receptive vocabulary (see footnote on page 50). These results might add to the previous research in ASD about auditory stream analysis (e.g., Lepistö et al., 2009), namely that poor auditory filtering might stem from poor auditory stream analysis among individuals with ASD (Alcántara, Weisblatt, Moore, & Bolton, 2004). Research shows that typically developing infants are able to segregate sounds into two perceptual streams (Smith & Trainor, 2011), but this ability does not become adult-like until 9 to 11 years old (Sussman, Wong, Horváth, Winkler, & Wang, 2007). Among those with ASD, this development might be disrupted given the problems seen with early brain development (Courchesne et al., 2003; Gillberg & de Souza, 2002). Thus, those with ASD show poor filtering so that they might function normally in simple auditory environments, but a breakdown occurs in complex auditory environments that require more advanced processing.

2. Phoneme Categorization. Efficient processing of speech relies on perceiving speech sounds according to the phonemic categories of the language spoken. Young infants are equally able to discriminate speech sounds that fall into two different phonemic categories in their language of acquisition as they are to discriminate speech sounds that fall into one phonemic category. However, by 12 months of age, normally developing infants, like adults, have become specialized for the language in their environment, and they have difficulty discriminating foreign phonemic categories that map onto a single category in their native language (e.g., Best, McRoberts, & Goodell, 2001; Werker & Lalonde, 1998; Werker & Tees, 2005). Here we measured specialization by comparing the amount of decrement between the two category (native) and one category (foreign) mapping conditions. Interestingly, we found less specialization for native language phonemic categories in those with ASD compared to those who are typically developing, which persisted even after accounting for individual differences in receptive vocabulary (see footnote on page 52). This reduced specialization therefore, means that those with ASD attend to sounds (especially native speech) less selectively in their environment, which could impact the way in which they communicate with others in that environment[6]. It also means that language may develop more slowly because perception is less constrained among those with ASD compared to controls. An auditory system that is not as finely tuned as controls can be a major hindrance to effective communication because of the number of speech sounds that are available in a given environment and because these sounds decay rapidly in speech. Thus, very early speech and language remediation might be needed for those with ASD, but of course, such remediation depends on developing effective early screening practices for this disorder.

3. Audio-Visual Integration. Understanding speech requires that listeners integrate cues from multiple sources, such as the lips, eyes and hands, in addition to what the speaker is actually saying. Most of the research involving audio-visual integration of speech sounds is based on the McGurk effect where participants are asked to report what they hear after they are presented with audio and video inputs that are incongruent (audio "ba" and visual "ga") (McGurk & MacDonald, 1976). If participants integrate, the result is that they "hear" a third percept ("da"), which represents a fused response of what they see and what they hear (McGurk & MacDonald, 1976). Here we compared the ability of those with ASD and controls to integrate audio and visual information in speech in order

---

[6] The Pearson correlation between native speech sound processing and simple meter processing for the entire sample was approaching significance ($r = .24$, $p = .08$). When the same correlation was performed only for the ASD group, the correlation was not significant ($r = .27$, $p = .17$). However, our main results using ANOVA show that auditory processing is less domain-specific in the ASD than in the control group.

to determine both the susceptibility to the McGurk illusion and the relative contribution of lip-reading to this susceptibility in ASD. We found that those with ASD were less susceptible to the McGurk illusion as they were more likely than controls to report hearing “ba” than “da”. With respect to lip-reading, we did not find any significant impairment in ASD and so poor audio-visual integration in speech cannot be attributed to this deficit. It is likely that the poor long-distance connectivity in ASD makes communication between brain regions difficult (Just et al., 2004; Just et al., 2007), especially between the superior temporal sulcus and other brain areas, which are important for the integration of speech sounds (Calvert, Campbell, & Brammer, 2000). Thus, those with ASD might find it difficult to identify speech in noisy environments, where being able to combine information that is seen and heard would help comprehension. Given that the audio-visual integration of non-speech sounds seems to be less impaired (Foss-Feig et al., 2010; van der Smagt, van Engeland, & Kemner, 2007) than of speech sounds (e.g., de Gelder, Vroomen, & van der Heide, 1991), it might be possible to train those with ASD on these stimuli with the assumption that a transfer of learning may occur to speech sounds.

Music Processing Results

Research shows that those with ASD have reduced or absent brain asymmetry for certain regions, including Broca's area (inferior frontal) and Wernicke's area (superior temporal) (Herbert et al., 2002) which, on the right, are important for musical processing. These brain abnormalities may therefore, hinder some aspects of speech and language processing, but critically, enhance music perception. There is also research indicating that individuals with ASD are more attracted to music than speech (Finnigan & Starr, 2010; Jarvinen-Pasley & Heaton, 2007; Simpson & Keen, 2010). Thus, we measured three key areas of music processing (pitch memory, meter categorization and harmonic priming) and found impairments in tasks involving rhythm, but not involving pitch among those with ASD compared to those with typical development.

1. Absolute Pitch. For most adults and even infants who are as young as 6 months of age, pitch is processed using a relative pitch code (Plantinga & Trainor, 2005; Trainor & Trehub, 1992; Trehub, 2001; Trehub, Bull, & Thorpe, 1984). This means that we recognize a familiar song, such as “Mary Had a Little Lamb”, even if the piece is sung on a higher or lower starting pitch, so long as the distances between these pitches is preserved. For less than 5 out of every 10,000 individuals, pitch is processed using an absolute pitch code (Bachem, 1955; Brown et al., 2003). In these cases, an individual is able to name or produce a pitch that is heard, without relying on an external standard (Takeuchi & Hulse, 1993). Research in this area shows that those with ASD might have access to absolute pitch information (Brenton, Devries, Barton, Minnich, & Sokol, 2008; Heaton, Pring, & Hermelin, 1999; Mottron, Peretz, Belleville, & Rouleau, 1999; Young & Nettelbeck, 1995). Here we measured the prevalence of absolute pitch in order to understand its rate among the entire sample of those with ASD using a task that does not require note naming or formal musical training. This approach to pitch memory is similar to that used by Heaton and colleagues (Heaton, 2003; Heaton et al., 1999; Heaton, Pring, & Hermelin, 1998), although ours does not require associating a tone with a picture or location. We found a high instance of this ability in our sample with ASD, with 3 of our participants with ASD, but none of our controls scoring perfect on this task, with perfect scores indicating absolute pitch processing. Once those with absolute pitch were removed from the sample, however, there were no group differences in performance. Although absolute pitch is sometimes considered to be a gift, the more complex, but very common, ability of relative pitch is more important for both music and speech processing because it enables recognition of melodies and prosodic patterns across high or low pitch registers. Interestingly, relative pitch typically develops early (e.g., Trainor & Trehub, 1992). The same pattern of performance where skills that are based on early as opposed to late learning are impaired was seen in music as it was in speech. This developmental effect may be due to problems seen with early brain development (Courchesne et al., 2003; Gillberg & de Souza, 2002) and in turn, with the pattern of neural connectivity that follows (Courchesne et al., 2001, 2003; Courchesne & Pierce, 2005). Research with typical developing individuals shows that there is increased short-range neural connectivity among those with absolute pitch processing relative to those without this ability (Loui, Li, Hohmann, & Schlaug, 2011). It remains for future research to examine this pattern of neural connectivity among those with ASD who also possess absolute pitch.

2. Meter Categorization. Similar to the acquisition of sensitivity to one’s native phonemic categories, specialization for the metrical rhythm structure of the music system in one’s environment is also seen early in development. By 12 months of age, Western infants are like Western adults in being much better at detecting changes to rhythms whose metrical structures form simple ratio beat durations (e.g., 2:1) that are typical in

Western music, in comparison to metrical structures with complex ratios (e.g., 3:2) that are rare in Western music. However, 6-month-old infants are equally good at processing simple and complex rhythms, like adults who grew up in cultures whose folk music contains complex rhythms (Hannon & Trehub, 2005; Hannon & Trainor, 2007). Here we examined the issue of specialization for native metrical categories by measuring the difference in performance across the simple meter (native) and complex meter (foreign) conditions. Similar to our finding of less sensitivity to native phonemic categories, we found that those with ASD were less specialized than controls for processing rhythms with simple compared to complex meters. Thus, specialization for native metrical categories is atypical, pointing to the issue that skills that rely on early instead of late learning seem to be most abnormal. It also appears that rhythm processing is impaired, perhaps because this aspect of music, relative to other aspects of music, such as pitch, provides an important social function[7]. For example, research shows that being able to follow the rhythm, whether it be through dancing or clapping, can increase social bonding (Hannon & Trainor, 2007; Kirschner & Tomasello, 2009). Given the social impairments that are seen in ASD, it should not be surprising that these individuals perform poorly with respect to rhythm processing. Thus, it might be beneficial for those with ASD to participate in drumming circles where they can engage their musical interests as well as develop skills that are important for normal socialization.

3. Harmonic Priming. Experience with Western music during development leads

---

[7] It must be acknowledged that rhythmic activities do not always occur in a social context. An example of a rhythmic activity that can be engaged in on our own is rocking. Rocking is something that those with ASD are reported to engage in and is consistent with the repetitive and restrictive domain of the autistic triad.

to perceptual specialization for the rules of tonal harmony in that musical system. Between 4 and 7 years of age, typically developing children have acquired some implicit knowledge of Western harmonic structure (Corrigall & Trainor, 2010; Schellenberg, Bigand, Poulin-Charronnat, Garnier, & Stevens, 2005; Trainor & Corrigall, 2010; Trainor & Trehub, 1994). Chord sequences follow preference rules, such that a sequence sets up expectations in enculturated listeners for which chords are likely to come next, and these expectations can be measured implicitly with reaction times (e.g., Bigand, Madurell, Tillmann, & Pineau, 1999). In a typical implicit paradigm, chord sequences are presented, half of which have expected final chords and half unexpected. Reaction times on an irrelevant variable, such as to indicate the timbre of the final chord, are compared for sequences with expected and unexpected endings. For example, ending a sequence with a dominant to tonic chord progression produces faster reaction times than endings with a tonic to subdominant chord progression (Bigand et al., 1999; Schellenberg et al., 2005; Tillmann, Bigand, Escoffier, & Lalitte, 2006; Tillmann, Peretz, Bigand, & Gosselin, 2007). Here we measured the amount of harmonic priming in ASD using reaction time and accuracy measures. We found that those with ASD were similar to controls in that they were faster to judge the timbre of the final chord when the sequence ended with a typical dominant-to-tonic progression compared to when it ended with a less typical tonic-to-subdominant progression. Accuracy was also equivalent in both groups. Thus, this later developing skill appears to be relatively spared in ASD, consistent with the idea that the brain is most abnormal early in development. It appears that those with ASD have therefore, formed expectations about the way in which chords should be ordered in a sequence so that it is consistent with their experience with the music system in that environment. It also shows that pitch processing seems to be relatively intact among those with ASD, which is similar to the results that we found with pitch memory.

Speech Production Results

There are conflicting reports describing the speech of those with ASD including: “robotic”, “wooden”, “stilted”, “bizarre”, “over precise”, “monotone” and “singsong” (Baltaxe & Simmons, 1985; Baron-Cohen & Staunton, 1994; Fay & Schuler, 1980; Frith, 1991). Past work in this area is primarily based on subjective ratings, instead of robust acoustic analysis on how those with ASD vary prosodic features in their speech. Past work has also not examined how general language skills, such as receptive vocabulary may affect production patterns in ASD. Furthermore, few studies have examined the content of these utterances, namely the information structure. Research indicates that in conversations, utterances typically involve two meaningful aspects: the focus and the topic (Vallduví & Engdahl, 1996). The focus involves new information in an utterance, while the topic involves old information from the previous utterance (Vallduví & Engdahl, 1996). Given that focus words have a higher information value than topic words, the former is made more prominent than the latter, such as with larger pitch ranges and longer word durations (Chen, 2009). Of the studies that have examined information structure among those with ASD, one study found that they accentuate focus and topic words equally (McCaleb & Prizant, 1985), whereas other studies found that they accentuate the beginning of a sentence irrespective of its information value (e.g., Baltaxe, Simmons, & Zee, 1984). Here we measured how those with ASD varied acoustic features generally in their speech and as a function of the information structure. We also examined whether the level of current language ability of our ASD speakers is associated in a predicable way with prosody use.

The groups were formed by a median split, with the higher-performing half of the individuals in one group and the lower-performing individuals in the other. The categorization by current language ability (PPVT) also followed diagnoses, such that all 6 with scores of 100 or greater (Autism High Language Function, A-highL, group) carried a diagnosis of Asperger's syndrome and the others (Autism Moderate Language Functioning, A-moderateL, group) a diagnosis of High-Functioning Autism. In addition, all 6 in the A-moderateL group experienced an early language delay whereas none in the A-highL group experienced an early language delay.

We found that those with ASD vary prosody differently depending on whether they were in the A-moderateL or A-highL group. We found that the A-moderateL group sounds "monotone" consistent with a restricted pitch range, whereas the A-highL group sounds "singsong" consistent with a larger than normal pitch range. No group differences were found in the way that duration was generally varied in speech. With respect to the marking of information structure, we found that the A-moderateL group used larger pitch ranges and longer word durations for focus than topic, as did controls, whereas the A-highL group did not. Thus, the A-moderateL group varies acoustic features according to the information structure, but it could be that the "monotone" quality of their speech gives the impression to listeners that they are disinterested in conversation. In contrast, the “singsong” quality of speech among the A-highL group may be distracting to listeners, in addition to the fact that this group does not effectively draw listeners’ attention to focus over topic words. It is for future research to determine how listeners perceive the production patterns of those with A-moderateL or A-highL, and whether it may be possible with very early speech and language remediation to normalize this aspect of their speech. At the same time, these results need to be interpreted with caution as our samples involved a relatively small number of participants. This point is discussed further below.

Research Limitations

One limitation of the perception research is that we did not collect questionnaire data from caregivers. For example, we could have measured auditory filtering problems using the Short Sensory Profile (McIntosh, Miller, Shyu, & Dunn, 1999). This would have provided a different perspective from which to understand issues in sensory modulation, in addition to the psychoacoustic testing we conducted. It would also have allowed us to compare our sample of participants to other studies that have reported data from the Short Sensory Profile. Similarly, we could have administered a questionnaire on social and communicative functioning, such as the Vineland Adaptive Behavior Scales (Sparrow, Cicchetti, & Balla, 2005) or the Children's Communicative Checklist (Bishop, 2006). This would have enabled us to investigate the effects of the perceptual processing deficits we measure on everyday functioning.

Another limitation of the perception research is that we did not have background information about whether our participants with ASD had experienced an early language delay or not. This information would have been interesting to collect in order to determine whether there is a relationship between delayed language development and pitch perception or the amount of specialization for native speech sound categories.

A limitation of the production research is that we have limited background information about our adult participants with ASD. That is, we know that they received a diagnosis of either AS (no language delay) or HFA (language delay), but we do not know specific information about treatments and services that they may have received throughout their development. For example, with respect to speech production, we do not know if our adult participants received early speech remediation, particularly for those who experienced an early language delay. This information would have helped to provide a larger context in which to understand our research results.

Another limitation of the production research is that our production study was based on a small sample size of adults with moderate to high language functioning. Thus, we do not know if the pattern of production results can be generalized to all individuals with this disorder. In particular, it is possible that the differences we observed between groups would be exaggerated if lower-functioning individuals were included.

Finally, another limitation of the production research is that we used a semi-spontaneous speech task, which may not provide an entirely accurate understanding of how acoustic features and information structure are used naturally in speech. Although our research is an important first step, more research is needed that uses tasks measuring semi-spontaneous and spontaneous speech. Finally, we found that as the A-highL group used a larger than normal pitch range and the A-moderateL group used a smaller than normal pitch range, consistent with a monotone and singsong quality, respectively. However, we did not perform any perceptual tests to determine how these production patterns are perceived by typical listeners. Thus, we do not know whether listeners would perceive the speech of those with A-moderateL as “monotone” and those with A-highL as “singsong”. It is for future research to determine how the production patterns of those with ASD are perceived and critically, how these impressions may be affected by the marking of information structure.

Future Research Directions

Our perceptual research shows the aspects of auditory processing in which those with ASD are lagging in performance and the aspects in which they are performing better than or similar to controls. However, these results are based on our battery of tests that were administered by an experimenter in a laboratory setting. We did not collect self-report data, for example, and so we do not know how those with ASD think about their own abilities and in turn, the impact that perceptual problems may have on their environment. Collecting this information would provide a richer research context in which to understand perceptual abnormalities among those with ASD. It would also provide the participants with the opportunity to report their first-hand experiences, which might be appropriate for an adolescent sample. This information could help to develop self-awareness, namely Theory of Mind (ToM), which involves an understanding of the thoughts and feelings of others, an area that is known to be impaired in ASD (Baron-Cohen, Leslie, & Frith, 1985). Developing ToM may play an important role in the remediation of abnormal perception, particularly by making those with ASD aware of their behaviours and in turn, the impact that these behaviours can have on their environment. Thus, asking participants with ASD to report on their experiences would allow for a more comprehensive understanding of abnormal perception in this group.

We found that those with ASD had a higher prevalence of absolute pitch than in the normal population, but we did not do any testing to determine if these individuals fit the profile of musical savants. Although 11 percent of the ASD sample had absolute pitch processing, which is similar to the estimate of 10 percent of those with ASD who have savant skills (Rimland & Hill, 1984), performance in the musical domain would have to be formally tested. This testing in future research would help to determine what it is about individuals with ASD who have absolute pitch processing that differentiates them from others with this disorder or those in the control group. The idea of ASD subgroups is consistent with other research that finds superior perception (Jones et al., 2009) and memory (Heaton, Williams, Cummins, & Happé, 2008) for pitch among a subgroup with ASD (10 to 20% of the sample), which was not typical of the larger ASD group. This testing could include a brain imaging component in order to determine whether those with absolute pitch in the ASD group have similar patterns of neural connectivity as reported among those with the same ability in the typically developing literature (Loui, Hohmann, & Schlaug, 2011). Overall, this research would allow for a better understanding of how absolute pitch develops and what are the contributing factors responsible for this ability in ASD.

Our results are consistent with aspects of both the WCC and EPF theories. The observation that those with ASD showed poorer filtering, less audio-visual integration and reduced specialization for native speech sound categories and musical meters is consistent with the WCC theory that those with ASD have difficulty in bringing local elements together to form a coherent whole. However, consistent with the EPF theory, we also found that global processing was intact among those with ASD in some tasks, such as the harmonic priming task. This processing may be explained by the developmental age at which implicit knowledge of tonal harmony develops as well as the fact that musical processing seems to be generally more spared than speech processing. Interestingly, we found enhanced processing of low-level information in the use of an absolute pitch code among a subgroup with ASD. It is for future research to replicate and expand these results in order to better understand auditory processing in this clinical group.

General Conclusions

This dissertation provides a comprehensive profile of the perception abilities in language and music as well as the speech production patterns among those with ASD relative to controls. In chapter 2, we found that adolescents with ASD had problems ignoring one stream of input while attending to another, integrating audio-visual information in speech, and categorizing the phonemic and rhythmic categories that are important to the language and musical system in their environment. In general, those with ASD tended to be more impaired on tasks involving speech than on tasks involving music, forming a profile by which we can begin to understand auditory processing in this disorder. This profile informs readers about the nature of ASD, including the range of behaviours involved in the autistic triad, which involves impaired social functioning, repetitive and restrictive interests, and poor communication (APA, 1994). This research may ultimately provide a foundation in which to develop more appropriate remediation strategies and screening practices for ASD. In chapter 3, we found differences in speech production patterns between ASD language subgroups, such that those with A-moderateL used a smaller, and those with A-highL used a larger, pitch range than controls. We also found that the A-highL group did not vary pitch range or duration to mark topic and focus, whereas the A-moderateL group did differentiate topic and focus, although to a lesser extent than controls. This research will help to determine if there are critical periods in which early speech and language interventions should be in place for those with ASD in order to normalize these aspects of speech. Taken together, our results inform us about the perception and production abilities of individuals with ASD relative to those who are typically developing. Our results also provide a foundation in which to better help individuals with this disorder, whether it be through the use of speech or music.

CHAPTER 1 AND 4 REFERENCES

Alcántara, J. I., Weisblatt, E. J. L., Moore, B. C. J., & Bolton, P. F. (2004). Speech-in-noise perception in high-functioning individuals with autism or asperger's syndrome. *Journal of Child Psychology and Psychiatry,* 45, 1107-1114.

Allen, R., & Heaton, P. (2010). Autism, music, and the therapeutic potential of music in alexithymia. *Music Perception, 27,* 251-261.

American Psychiatric Association. (1980). *Diagnostic and Statistical Manual of Mental Disorders* (3th ed.). Washington, DC.

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC.

Ashburner, J., Ziviani, J., & Rodger, S. (2008). Sensory processing and classroom emotional behavioural, and education outcomes in children with autism spectrum disorder. *American Journal of Occupational Therapy,* 62, 564-573.

Bachem, A. (1955). Absolute pitch. *Journal of the Acoustical Society of America,* 27, 1180-1185.

Baker, A. E., Lane, A., Angley, M. T., & Young, R. L. (2008). The relationship between sensory processing patterns and behavioural responsiveness in autistic disorder: A pilot study. *Journal of Autism and Developmental Disorders*, *38,* 867-875.

Baltaxe, C. A. M., & Guthrie, D. (1987). The use of primary sentence stress by normal, aphasic and autistic children. *Journal of Autism and Developmental Disorders, 17,* 255-271.

Baltaxe, C. A. M., & Simmons, J. Q. (1985). Prosodic development in normal and autistic children. In E. Schloper & G. Mesilov (Eds.), *Issues in Autism (Vol. 3), Communication Problems in Autism.* New York: Plenum.

Baltaxe, C. A. M., Simmons, J. Q., & Zee, E. (1984). Intonation patterns in normal, autistic, and aphasic children. In A. Cohen & M. P. R. Broecke (Eds.), *Proceedings of the Tenth International Congress of Phonetic Sciences* (pp. 713-718). Dortrecht-Holland: Foris Publications.

Baranek, G. T. (1999). Autism during infancy: A retrospective video analysis of sensory-motor and social behaviors at 9-12 months of age. *Journal of Autism and Developmental Disorders,* 29, 213-224.

Baranek, G. T. (2002). Efficacy of sensory and motor interventions for children with autism. *Journal of Autism and Developmental Disorders,* 32, 397-422.

Baranek, G. T., Boyd, B. A., Poe, M. D., David, F. J., & Watson, L. R. (2007). Hyperresponsive sensory patterns in young children with autism, developmental delay, and development. *American Journal of Mental Retardation,* 112, 233-245.

Barnea-Goraly, N., Kwon, H., Menon, V., Eliez, S., Lotspeich, L., & Reiss, A. L. (2004). White matter structure in autism: Preliminary evidence from diffusion tensor imaging. *Biological Psychiatry,* 55, 323-326.

Baron-Cohen, S. (2000). Theory mind and autism: A fifteen year review. In S. Baron-Cohen, H. Tager-Flusberg, and D. J. Cohen (Ed.), *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience* (pp. 3-20). Oxford University Press: Oxford, UK.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a 'theory of mind'? *Cognition, 21,* 37-46.

Baron-Cohen, S., & Staunton, R. (1994). Do children with autism acquire the phonology of their peers? An examination of group identification through the window of bilingualism. *First Language, 14,* 241-248.

Behrmann, M., Thomas, C., & Humphreys, K. (2006). Autism: seeing it differently. *Trends in Cognitive Science, 10,* 258-264.

Bellon-Harn, M. L., Harn, W. E., & Watson, G. D. (2007). Targeting prosody in an eight-year-old child with high-functioning autism during an interactive approach to therapy. *Child Language Teaching & Therapy, 23,* 157-179.

Ben-Sasson, A, Hen, L., Fluss, R., Cermak, S. A., Engel-Yeger, B., & Gal, E. (2009). A meta-analysis of sensory modulation symptoms in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39,* 1-11.

Best, C. T., & Avery, R. A. (1999). Left-hemisphere advantage for click consonants is determined by linguistic significance and experience. *Psychological Science, 10,* 65-69.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perception assimilation to the listener's native phonological system. *Journal of Acoustic Society of America,* 109, 775-794.

Bigand, E., Madurell, F., Tillmann, B., & Pineau, M. (1999). Effect of global structure and temporal organization on chord processing. *Journal of Experimental Psychology: Perception and Performance*, 25, 184-197.

Bird, G., Catmur, C., Silani, G., Frith, C., & Frith, U. (2006). Attention does not modulate neural responses to social stimuli in autism spectrum disorders. *Neuroimage,* 31, 1614-1624.

Bishop, D. V. M. (2006). *Children's Communication Checklist-2 (CCC-2).* San Antonio, TX: Pearson.

Boddaert, N., Chabane, N., Belin, P., Bourgeois, M., Royer, V., Barthelemy, C., … Zilbovicius, M. (2004). Perception of complex sounds in autism: abnormal auditory cortical processing in children. *American Journal of Psychiatry,* 161, 2117-2120.

Bonneh, Y. S., Levanon, Y., Dean-Pardo, O., Lossos, L., & Adini ,Y. (2011). Abnormal speech spectrum and increased pitch variability in young autistic children. *Frontiers in Human Neuroscience, 4,* 1-7.

Bonnel, A., Mottron, L., Peretz, I., Trudel, M., Gallun, E., & Bonnel, A. (2003). Enhanced pitch sensitivity in individuals with autism: A signal detection analysis. *Journal of Cognitive Neuroscience, 15,* 226-235.

Bonnel, A., McAdams, S., Smith, B., Berthiaume, C., Bertone, A., Ciocca, V., … Mottron, L., (2010). Enhanced pure-tone pitch discrimination among persons with autism but not Asperger syndrome. *Neuropsychologia, 48*, 2465-2475.

Brenton, J. N., Devries, S. P., Barton, C., Minnich, H., & Sokol, D. (2008). Absolute pitch in a four-year-old boy with autism. *Pediatric Neurology,* 39, 137-138.

Brock, J., Norbury, C. F., Einav, S., & Nation, K. (2008). Do individuals with autism process words in context? Evidence from language-mediated eye-movements. *Cognition, 108,* 896-904.

Brown, W. A., Cammuso, K., Sachs, H., Winklosky, B., Mullane, J., Bernier, R., … Folstein, S. E. (2003). Autism-related language, personality, and cognition in people with absolute pitch: Results of a preliminary study. *Journal of Autism and Developmental Disorders,* 33, 163-167.

Burd, I., Fisher, W., & Kerbeshian, J. (1987). A prevalence study of pervasive developmental disorders in North Dakota. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 700-703.

Calvert, G.A., Campbell, R., & Brammer, M.J., 2000. Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology, 10,* 649-657.

Casanova, M. F. (2004). Intracortical circuitry: One of Psychiatry’s missing assumptions. *European Archives of Psychiatry and Clinical Neuroscience,* 254, 148-151.

Casanova, M. F., Buzhoeveden, D. P., Switala, A. E., & Roy, E. (2002). Minicolumnar pathology in autism. *American Academy of Neurology,* 58, 428-432.

Casanova, M. F., Switala, A., E. Trippe, J., & Fitzgerald, M. (2007). Comparative minicolumnar morphometry of three distinguished scientists, *Autism, 11,* 557-569.

Castelli, F., Frith, C., Happé, F., & Frith, U. (2002). Autism, asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain,* 125, 1839-1849.

Chen, A. (2009). The phonetics of sentence-initial topic and focus in adult and child Dutch. In M. Vigário, S. Frota and M. J. Freitas (Eds.), *Phonetics and Phonology: Interactions and Interrelations* (pp. 91-106). Amsterdam: Benjamins.

Chen, A. (2011). Tuning information packaging: Intonational realisation of topic and focus in child Dutch. *Journal of Child Language, 38,* 1055-1083.

Constantino, J. N., Yang, D., Gray, T. L., Gross, M. M., Abbacchi, A. M., Smith, S. C., … Kuhl P. K. (2007). Clarifying the association between language and social development in autism: A study of non-native phoneme. *Journal of Autism and Developmental Disorders,* 37, 1256-1263.

Corrigall, K. A., & Trainor, L. J. (2010). Musical enculturation in preschool children: Acquisition of key and harmonic knowledge. *Music Perception,* 28, 195-200.

Courchesne, E. (2004). Brain development in autism: Early overgrowth followed by premature arrest of growth. *Mental Retardation and Developmental Disability Research Review,* 10, 106-111.

Courchesne, E., Carper, R., & Akshoomoff, N. (2003). Evidence of brain overgrowth in the first year of life in autism. *The Journal of the American Medical Association,* 290, 337-344.

Courchesne, E., Karns, C. M., Davis, H. R., Ziccardi, R., Carper, R. A., Tigue, Z. D., … Courchesne, R. Y. (2001). Unusual brain growth patterns in early life in patients with autistic disorder: An MRI study. *Neurology,* 57, 245-254.

Courchesne, E., & Pierce, K. (2005). Why the frontal cortex in autism might be talking only to itself: Local over-connectivity but long-distance disconnection. *Current Opinion in Neurobiology,* 15, 225-230.

Crane, L., Goddard, L., & Pring, L. (2009). Sensory processing in adults with autism spectrum disorders. *Autism: The International Journal of Research and Practice, 13*, 215-228.

Dahlgren, S. V., & Gillberg, C. (1989). Arousal, attention, and the socioemotional impairments of individuals with autism. In G. Dawson (Ed.), *Autism: Nature, diagnosis and treatment* (pp. 44-74). New York: Guilford.

Dalton, K. M., Nacewicz, B. M., Johnstone, T., Schaefer, H. S., Gernsbacher, M. A., Goldsmith, H. H., … Davidson, R. J. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nature Neurocience,* 8, 519-526.

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism: Social orienting, joint attention, and attention to distress. *Developmental Psychology, 40,* 271-283.

de Gelder, B., Vroomen, J., & van der Heide, L. (1991). Face recognition and lip-reading in autism. *European Journal of Cognitive Psychology,* 3, 69-86.

Diehl, J., Watson, D., Bennetto, L., McDonough, J., & Gunlogson, C. (2009). An acoustic analysis of prosody in high-functioning autism. *Applied Psycholinguistics, 30*, 385-404.

Fay, W., & Schuler, A. L. (1980). *Emerging language in autistic children.* Baltimore: University Park Press.

Finnigan, E., & Starr, E. (2010). Increasing social responsiveness in a child with autism: A comparison of music and non-music interventions. *Autism, 14,* 321-348.

Fombonne, E. (2003). Editorial. *Journal of the American Medical Association, 289,* 87-89.

Fosnot, S. M., & Jun, S. (1999). *Prosodic characteristics in children with stuttering or autism during reading and imitation*. Paper presented at the 14th International Congress of Phonetic Sciences.

Foss-Feig, J., Kwakye, L., Cascio, C., Burnette, C., Kadivar, H., Stone, W., & Wallace, M. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Experimental Brain Research, 203,* 381-389.

Frith, U. (1989). *Autism: Explaining the enigma.* Oxford: Blackwell.

Frith, U. (1991). *Autism and Asperger Syndrome.* Cambridge: Cambridge University Press.

Gillberg, C., & de Souza, L. (2002). Head circumference in autism, asperger syndrome, and ADHD: A comparative study. *Developmental Medicine & Child Neurology,* 44, 296-300.

Goldfarb, W. (1956). Receptor preferences in schizophrenic children. *Archives of Neurology and Psychiatry, 76,* 643-652.

Gomes, E., Rotta, N. T., Pedroso, F. S., Sleifer, P., & Danesi, M. C. (2004). Auditory hypersensitivity in children and teenagers with autistic spectrum disorder. *Arq Neuropsiquiatr,* 62, 797-801.

Goto, H. (1971) Auditory perception by normal Japanese adults of the sounds "l" and "r". *Neuropsychologia 9*, 317-323.

Grandin, T. (1992). Calming effects of deep touch pressure in patients with autistic disorder, college students, and animals. *Journal of Child and Adolescent Psychopharmacology, 2,* 63-72.

Green, H., & Tobin, Y. (2009). Prosodic analysis is difficult…but worth it: A study in high functioning autism. *International Journal of Speech-Language Pathology, 11,* 308-315.

Groen, W. B., van Orsouw, L., ter Huurne, N., Swinkels, S., van der Gaag, R. J., Buitelaar, J. K., & Zwiers, M. P. (2009). Intact spectrum but abnormal temporal processing of auditory stimuli in autism. *Journal of Autism and Developmental Disorders,* 39, 742-750.

Grossman, R. B., Bemis, R. H., Skwerer, D. P., & Tager-Flusberg, H. (2010). Lexical and affective prosody in children with high-functioning autism. *Journal of Speech, Language, and Hearing Research, 53,* 778-793.

Haesen, B., Boets, B., & Wagemans, J. (2011). A review of behavioural and electrophysiological studies on auditory processing and speech perception in autism spectrum disorders. *Research in Autism Spectrum Disorders, 5,* 701-714.

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

Hannon, E. E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science,* 16, 48-55.

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders* 35, 5-25.

Heaton, P. (2003). Pitch memory, labeling and disembedding in autism. *Journal of Child Psychology and Psychiatry, 44,* 543-551.

Heaton, P. (2005). Interval and contour processing in autism. *Journal of Autism and Developmental Disorders, 35,* 787-793.

Heaton, P., Davis, R., & Happé, F. (2008). Exceptional Absolute Pitch Perception for spoken words in an able adult with autism. *Neuropsychologia,* 46, 2095-2098.

Heaton, P., Hermelin, B., & Pring, L. (1998). Autism and pitch processing: A precursor for savant musical ability? *Music Perception, 15,* 291-305.

Heaton, P., Hermelin, B., & Pring, L. (1999). Can children with autistic spectrum disorders perceive affect in music? An experimental investigation. *Psychological Medicine,* 29, 1405-1410.

Heaton, P., Pring, L., & Hermelin, B. (1999). A pseudo-savant: A case of exceptional musical splinter skills. *Neurocase,* 5, 503-509.

Heaton, P., Williams, K., Cummins, O., & Happé, F. (2007). Beyond perception: Musical representation and on-line processing in autism. *Journal of Autism and Developmental Disorders,* 37, 1355-1360.

Heaton, P., Williams, K., Cummins, O., & Happé, F. (2008). Autism and pitch processing splinter skills: A group and sub-group analysis. *Autism, 12,* 21-37.

Herbert, M. R., Harris, G. J., Adrien, K. T., Ziegler, D. A., Makris, N., & Kennedy, D. N. (2002). Abnormal asymmetry in language association cortex in autism. *Annals of Neurology,* 52, 588-596.

Hubbard, K., & Trauner, D. A. (2007). Intonation and emotion in autistic spectrum disorders. *Journal of Psycholinguistic Research, 36,* 159-173.

Iarocci, G., Rombough, A., Yager, J., Weeks, D. J., & Chua, R. (2010). Visual influences on speech perception in children with autism. *Autism,* 14, 305-320.

Järvinen-Pasley A., & Heaton, P. (2007). Evidence for reduced domain-specificity in auditory processing in autism. *Developmental Science, 10,* 786-793.

Järvinen-Pasley, A., Wallace, G. L., Ramus, F., Happé, F., & Heaton, P. (2008). Enhanced perceptual processing of speech in autism. *Developmental Science,* 11, 109-121.

Johnsrude, I. S., Penhune, V. B., & Zatorre, R. J. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain, 123,* 155-163.

Jones, C. R., Happé, F., Baird, G., Simonoff, E., Marsden, A. J., Tregay, J., … Charman, T. (2009). Auditory discrimination and auditory sensory behaviours in autism spectrum disorders. *Neuropsychologia, 47*, 2850-2858.

Just, M. A., Cherkassky, V. L., Keller, T. A., Kana, R. K., & Minshew, N. J. (2007). Functional and anatomical cortical underconnectivity in autism: Evidence from an fMRI study of an executive function task and corpus callosum morphometry. *Cerebral Cortex,* 17, 951-961.

Just, M. A., Cherkassky, V. L., Keller, T. A., & Minshew, N. J. (2004). Cortical activation and synchronization during sentence comprehension in high-functioning autism: Evidence of underconnectivity. *Brain,* 127, 1811-1821.

Kern, J. K., Trivedi, M. H., Garver, C. R., Grannemann, B. D., Andrews, A. A., Savla, J. S., … Schroeder, J. (2006). The pattern of sensory processing abnormalities in autism. *Autism,* 10, 480-490.

Kern, J. K., Trivedi, M. H., Grannemann, B. D., Garver, C. R., Johnson, D. G., Andrews, A. A., … (2007). Sensory correlations in autism. *Autism, 11,* 123-134.

Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology, 102,* 299-314.

Kjelgaard, M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes, 16,* 287-308.

Lane, A., Young, R. L., Baker, A. E. Z., & Angley, M. (2010). Sensory processing subtypes in autism: Association with adaptive behavior. *Journal of Autism and Developmental Disorders,* 40, 112-122.

Lepistö, T., Kuitunen, A., Sussman, E., Saalasti, S., Jansson-Verkasalo, E., Nieminen-von Wendt, T., & Kujala, T. (2009). Auditory stream segregation in children with asperger syndrome. *Biological Psychology,* 82, 301-307.

Lewis, J. D., & Elman, J. L. (2007) Growth-related neural reorganization and the autism phenotype: A test of the hypothesis that altered brain growth leads to altered connectivity. *Developmental Science,* 11, 135-155.

Loui, P., Li, H., Hohmann, A., & Schlaug, G. (2011). Enhanced cortical connectivity in absolute pitch musicians: A model for local hyperconnectivity. *Journal of Cognitive Neuroscience, 23,* 1015-1026.

Mann, T., & Walker, P. (2003). Autism and a deficit in broadening the spread of visual attention. *Journal of Child Psychology and Psychiatry*, *44,* 274-284.

McCaleb, P., & Prizant, B. M. (1985). Encoding of new versus old information by autistic children. *Journal of Speech and Hearing Disorders, 50,* 230-240.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature,* 264, 746-748.

McIntosh, D. N., Miller L. J., Shyu, V., & Dunn, W. (1999). Overview of the Short Sensory Profile. In W. Dunn (Ed.), *The Sensory Profile: Examiner's Manual* (pp. 59-73). San Antonio, TX: The Psychological Corporation.

Mottron, L., & Burack, J. (2001). Enhanced perceptual functioning in the development of autism. In J. A. Burack, T. Charman, N. Yirmiya, & P. R. Zelazo (Eds.), *The development of autism: Perspectives from theory and research* (pp. 131-148). Mahwah, NJ: Erlbaum.

Mottron, L., Burack, J., Iarocci, G., Belleville, S., & Enns, J. (2003). Locally oriented perception with intact global processing among adolescents with high functioning autism: Evidence from multiple paradigms. *Journal of Child Psychology and Psychiatry,* 44, 906-913.

Mottron, L., Peretz, I., Belleville, S., & Rouleau, N. (1999). Absolute pitch in autism: A case-study. *Neurocase,* 5, 485-501.

Mottron, L., Peretz, I., & Ménard, E. (2000). Local and global processing of music in high-functioning persons with autism: Beyond cerebral coherence? *Journal of Child Psychology and Psychiatry, 41,* 1057-1065.

Nadig, A., & Shaw, H. (2011). Expressive prosody in high-functioning autism: Increased pitch range and what it means to listeners. *Journal of Autism and Developmental Disorders,* April 29, Epublication ahead of print.

Nadig, A., Vivanti, G., & Ozonoff, S. (2009). Adaptation of object descriptions to a partner under increasing communicative demands: A comparison of children with and without autism. *Autism Research, 2,* 334-347.

Norbury, C. F. (2005). Barking up the wrong tree? Lexical ambiguity resolution in children with language impairments and autistic spectrum disorders. *Journal of Experimental Child Psychology, 90,* 142-171.

Norbury, C. F., Brock, J., Cragg, L., Einav, S., Griffiths, H., & Nation, K. (2009). Eye-movement patterns are associated with communicative competence in autistic spectrum disorders. *Journal of Child Psychology and Psychiatry, 50,* 834-842.

O'Connor, K. (2012). Auditory processing in autism spectrum disorder: A review. *Neuroscience & Biobehavioral Reviews, 36,* 836-854.

O'Neill, M., & Jones, R. S. (1997). Sensory-perceptual abnormalities in autism: A case for more research? *Journal of Autism & Developmental Disorders 27,* 283-293.

Osterling, J., & Dawson, G. (1994). Early recognition of children with autism: A study of first birthday home video tapes. *Journal of Autism and Developmental Disorders,* 24, 247-257.

Osterling, J., Dawson, G., & Munson, J. (2002). Early recognition of one year old infants with autism spectrum disorder versus mental retardation: A study of first birthday party home videotapes. *Development and Psychopathology, 14,* 239-51.

Paccia, J., & Curcio, F. (1982). Language processing and forms of immediate echolalia in autistic children. *Journal of Speech and Hearing Research, 25,* 42-47.

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

Paul, R., Bianchi, N., Augustyn, A., Klin, A., & Volkmar, F. (2008). Production of syllable stress in speakers with autism spectrum disorders. *Research in Autism Spectrum Disorders, 2,* 110-124.

Paul, R., Chawarska, K., Fowler, C., Cicchetti, D., & Volkmar, F.R. (2007). Listen my children and you shall hear: Auditory preferences in toddlers with autism spectrum disorders. *Journal of Speech Language and Hearing Research*, *50*, 1350-1364.

Peppé, S., Cleland, J., Gibbon, F. E., O'Hare, A. E., & Martínez-Castilla, P. (2011). Expressive prosody in children with autism spectrum conditions, *Journal of Neurolinguistics*, *24,* 41-53.

Peppé, S., McCann, J., Gibbon, F., O'Hare, A., & Rutherford, M. (2006). Assessing prosodic and pragmatic ability in children with high-functioning autism. *Journal of Pragmatics, 38,* 1776-1791.

Peppé, S., McCann, J., Gibbon, F., O'Hare. A., & Rutherford, M. (2007). Receptive and expressive prosodic ability in children with high-functioning autism. *Journal of Speech, Language, and Hearing Research, 50,* 1015-1028.

Pierce, K., Conant, D., Hazin, R., Stoner, R., & Desmond, J. (2010). Preference for geometric patterns early in life as a risk factor for autism. *Archives of General Psychiatry,* 68, 101-109.

Pierce, K., Muller, R. A., Ambrose, J., Allen, G., & Courchesne, E. (2001). Face processing occurs outside the fusiform `face area' in autism: Evidence from functional MRI. *Brain, 124,* 2059-2073.

Plaisted, K., Swettenham, J., & Rees, L. (1999). Children with autism show local precedence in a divided attention task and global precedence in a selective attention task. *Journal of Child Psychology and Psychiatry* 40, 733-742.

Plantinga, J., & Trainor, L. J. (2005). Memory for melody: Infants use a relative pitch code. *Cognition*, 98, 1-11.

Rabin, I., & Katzman, R. (1998). Neurobiology of autism. *Annual Neurology, 43,* 7-14.

Rimland, B., & Edelson, S. M. (1995). Brief report: A pilot study of auditory integration training in autism. *Journal of Autism and Developmental Disabilities, 25,* 61-70.

Rimland, B., & Hill, A. L. (1984). Idiot savants. In J. Wortis (Ed.), *Mental Retardation and Developmental Disabilities* (pp. 55-169). New York: Plenum Press.

Rogers, S. J., Hepburn, S., & Wehner, E. (2003). Parent reports of sensory symptoms in toddlers with autism and those with other developmental disorders. *Journal of Autism and Developmental Disorders,* 33, 631-42.

Rogers, S. J., & Ozonoff, S. (2005). Annotation: What do we know about sensory dysfunction in autism? A critical review of the empirical evidence. *Journal of Child Psychology and Psychiatry, 46,* 1255-1268.

Rosenhall, U., Nordin, V., Sandstrom, M., Ahlsen, G., & Gillberg, C. (1999). Autism and hearing loss. *Journal of Autism & Developmental Disorders,* 29, 349-357.

Russo, N. M., Nicol, T., Trommer, B., Zecker, S., & Kraus, N. (2009). Brainstem transcription of speech is disrupted in children with autism spectrum disorders. *Developmental Science,* 12, 557-567.

Russo, N. M., Zecker, S., Trommer, B., Chen, J., & Kraus, N. (2009). Effects of background noise on cortical encoding of speech in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 39, 1185-1196.

Samson, F., Mottron, L., Jemel, B., Belin, P., & Ciocca, V. (2006). Can spectro-temporal complexity explain the autistic pattern of performance on auditory tasks? *Journal of Autism & Developmental Disorders, 36,* 65-76.

Schellenberg, E. G., Bigand, E., Poulin-Charronnat, B., Garnier, C., & Stevens, C. (2005). Children's implicit knowledge of harmony in Western music. *Developmental Science,* 8, 551-566.

Schoen, S. A., Miller, L. J., Brett-Green, B. A., & Nielsen, D. M. (2009). Physiological and behavioral differences in sensory processing: A comparison of children with autism spectrum disorder and sensory modulation disorder. *Frontiers in Integrated Neuroscience,* 3, 1-11.

Schultz, R. T., Gauthier, I. T., Klin, A., Fulbright, R. K., Anderson, A. W., Volkmar, F., … Gore, J. C. (2000). Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and asperger syndrome. *Archives of General Psychiatry,* 57, 331-340.

Shah, A., & Frith, U. (1983). Why do autistic individuals show superior performance on the block design test? *Journal of Child Psychology and Psychiatry, 34,* 1351-1364.

Sharda, M., Subhadra, T. P., Sahay, S., Nagaraja, C., Singh, L., Mishra, R., … Singh, N. C. (2010). Sounds of melody—pitch patterns of speech in autism. *Neuroscience Letters, 478*, 42-45.

Shriberg, L., Paul, R., McSweeney, J., Klink, A., Cohen, D., & Volkmar, F. (2001). Speech and prosody characteristics of adolescents and adults with high functioning autism and Asperger syndrome. *Journal of Speech, Language and Hearing Research, 44*, 1097-1115.

Simpson, K., & Keen, D. (2010). Teaching young children with autism graphic symbols embedded within an interactive song. *Journal of Developmental and Physical Disabilities, 22,* 165-177.

Skukla, D. K., Keehn, B., & Müller, R. (2010). Tract-specific analyses of diffusion tensor imaging show widespread white matter compromise in autism spectrum disorder. *Journal of Child Psychology and Psychiatry,* 52, 286-295.

Smith, E. G., & Bennetto, L. (2007). Audiovisual speech integration and lipreading in autism. *Journal of Child Psychology and Psychiatry, 48,* 813-821.

Smith, N. A., & Trainor, L. J. (2011). Auditory stream segregation improves infants' selective attention to target tones amid distracters. *Infancy, 1,* 1-14.

Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland Adaptive Behavior Scales, Second Edition (Vineland-II).* San Antonio, TX: Pearson.

Sussman, E., Wong, R., Horváth, J., Winkler, I., & Wang, W. (2007). The development of perceptual organization of sound by frequency separation in 5-11 year-old children. *Hearing Research, 225,* 117-127.

Takeuchi, A. H., & Hulse, S. H. (1993). Absolute pitch. *Psychological Bulletin, 113,* 345-361.

Teder-Sälejärvi, W. A., Pierce, K. L., Courchesne, E., & Hillyard, S. A. (2005). Auditory spatial localization and attention deficits in autistic adults. *Cognitive Brain Research, 23,* 221-234.

Tillmann, B., Bigand, E., Escoffier, N., & Lalitte, P. (2006). The influence of musical relatedness on timbre discrimination. *European Journal of Cognitive Psychology,* 18, 343-358.

Tillmann, B., Peretz, I., Bigand, E., & Gosselin, N. (2007). Harmonic priming in an amusic patient: The power of implicit tasks. *Cognitive Neuropsychology,* 24, 1-20.

Tomasello, M., & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language, 4,* 197-212.

Townsend, J., & Courchesne, E. (1994). Parietal damage and narrow "spotlight" spatial attention. *Journal of Cognitive Neurosciences, 6,* 218-230.

Trainor, L. J., & Corrigall, K. A. (2010). Music acquisition and effects of musical experience. In M. Riess-Jones & R. R. Fay (Eds.), *Springer Handbook of Auditory Research: Music Perception* (pp. 89-128). Heidelberg: Springer.

Trainor, L. J., & Trehub, S. E. (1992). A comparison of infants' and adults' sensitivity to Western musical structure. *Journal of Experimental Psychology: Human Perception and Performance,* 18, 394-402.

Trainor, L. J., & Trehub, S. E. (1994). Key membership and implied harmony in Western tonal music: Developmental perspectives. *Perception and Psychophysics,* 56, 125-132.

Trehub, S. E. (2001). Musical predispositions in infancy. *Annals of the New York Academy of Sciences,* 930, 1-16.

Trehub, S. E., Bull, D., & Thorpe, L. A. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development*, 55, 821-830.

Vallduví, E., & Engdahl, E. (1996). The linguistic realisation of information packaging. *Linguistics* 34, 459-519.

Van der Smagt, M. J., van Engeland, H., & Kemner, C. (2007). Brief report: Can you see what is not there? Low-level auditory-visual integration in autism spectrum disorder. *Journal of Autism and Developmental Disabilities, 37,* 2014-2019.

Watling, R.I., Deitz, J., & White, O. (2001) Comparison of sensory profile scores of young children with and without autism spectrum disorders, *American Journal of Occupational Therapy, 55,* 416-423.

Wellman, H. (1993). Early understanding: The normal case. In: S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from autism* (pp. 10-39). Oxford: Oxford University Press.

Werker, J. F., & Lalonde, C. E. (1998). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology, 24*, 672-683.

Werker, J. F., & Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology, 46,* 233-251.

Werner, E., Dawson, G., Osterling, J., & Dinno, N. (2000). Brief report: Recognition of autism spectrum disorder before one year of age: A retrospective study based on home videotapes. *Journal of Autism and Developmental Disorders,* 30, 157-162.

Williams, J. H. G., Massaro, D. W., Peel, N. J., Bosseler, A., & Suddendorf, T. (2004). Visual auditory integration during speech imitation in autism. *Research in Developmental Disabilities,* 25, 559-575.

Young, R. L., & Nettelbeck, T. (1995). The abilities of a musical savant and his family. *Journal of Autism and Developmental Disorders,* 25, 231-248.

Zwaigenbaum, L., Bryson, S., Roberts, W., Brian, J., & Szatmari, P. (2005). Behavioral markers of autism in the first year of life. *International Journal of Develpmental Neuroscience, 23*, 143-152.