

5-8-2020

Acoustic and Perceptual Characteristics of Prosody in Autism Spectrum Disorder: The Impact of Language, Motor Speech, and Auditory Processing

Colleen Gargan
University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/dissertations_2



Part of the [Speech Pathology and Audiology Commons](#)

Recommended Citation

Gargan, Colleen, "Acoustic and Perceptual Characteristics of Prosody in Autism Spectrum Disorder: The Impact of Language, Motor Speech, and Auditory Processing" (2020). *Doctoral Dissertations*. 1964.
<https://doi.org/10.7275/mfm1-gm46> https://scholarworks.umass.edu/dissertations_2/1964

This Open Access Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

**ACOUSTIC AND PERCEPTUAL CHARACTERISTICS OF PROSODY IN
AUTISM SPECTRUM DISORDER: THE IMPACT OF LANGUAGE, MOTOR
SPEECH, AND AUDITORY PROCESSING**

A Dissertation Presented

by

COLLEEN E. GARGAN

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2020

Communication Disorders

© Copyright by Colleen E. Gargan 2020

All Rights Reserved

**ACOUSTIC AND PERCEPTUAL CHARACTERISTICS OF PROSODY IN
AUTISM SPECTRUM DISORDER: THE IMPACT OF LANGUAGE, MOTOR
SPEECH, AND AUDITORY PROCESSING**

A Dissertation Presented

by

COLLEEN E. GARGAN

Approved as to style and content by:

Mary V. Andrianopoulos, Chair

Nathanial A. Whitmal, Member

William Matthews, Member

Shelley Velleman, Member

Sarah Poissant, Acting Department Chair
Department of Communication Disorders

DEDICATION

This dissertation is dedicated to my parents. Thank you for your love, support and encouragement as I pursued my Ph.D.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Dr. Mary Andrianopoulos, who expertly guided me through my master's and doctoral level education for the past eight years. Her passion and enthusiasm for science, research, and teaching is contagious, and was the driving force throughout my graduate career at the University of Massachusetts Amherst. This dissertation was inspired by my conversations with Dr. Andrianopoulos and improved in quality with her advice, ideas, and constructive feedback. It has been a great privilege and honor to work under her guidance while growing into the academician and clinical researcher I am today. I will attribute my ongoing professional success to this strong mentorship.

I also wish to thank my dissertation committee members: Dr. Shelley Velleman, for encouraging me to think critically and providing detailed feedback on this manuscript; Dr. William Matthews, for teaching me about research design, providing valuable feedback on my research questions, and expressing confidence in my abilities; and Dr. Nathaniel Whitmal, for his thoughtful suggestions related to the auditory processing and statistical components of this project. In addition, many thanks to Dr. Craig Wells and the Research, Educational Measures, and Psychometrics (REMP) graduate students for statistical consulting.

I gratefully acknowledge the U.S. Department of Education, Office of Special Education Programs for funding my master's and doctoral studies [U.S. DOE, H325K120327, "Training Speech Language Pathologists in the Public Schools to Deliver Evidence-based Models of Technology Effectively" (Project REMOTE)] and [U.S. DOE, H325K180163, "Interdisciplinary Preparation of Related-Service Professionals in

Evidence-Based Practices: High-Risk Students/Families and Remote Technologies” (Project iPREP)]. This research was partially funded by the Organization for Autism Research (OAR) and the University of Massachusetts Graduate School Dissertation Research Grant. In addition, thank you to the School of Public Health and Health Sciences Dean’s Ph.D. Fellowship for supporting my doctoral studies during the summer of 2018.

Last but not least, I am incredibly blessed with a strong support network of family and friends. Thank you for always being there for me, encouraging me to keep working hard, and for expressing genuine interest in my work.

ABSTRACT

ACOUSTIC AND PERCEPTUAL CHARACTERISTICS OF PROSODY IN AUTISM SPECTRUM DISORDER: THE IMPACT OF LANGUAGE, MOTOR SPEECH, AND AUDITORY PROCESSING

MAY 2020

COLLEEN E. GARGAN, B.S., UNIVERSITY OF MASSACHUSETTS AMHERST

M.A., UNIVERSITY OF MASSACHUSETTS AMHERST

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Mary V. Andrianopoulos

Autism Spectrum Disorder (ASD) is a lifelong neurodevelopmental condition that affects an individual's social communication, social interaction, and behavior (American Psychological Association, 2013). A striking feature that distinguishes some individuals with autism from their peers without autism is “atypical” prosody. A between-group study was conducted to investigate prosody, speech motor control, auditory processing of pitch, and trained listener ratings of prosodic “naturalness” among adolescents with ASD (n=17) compared to TD controls (n=17) matched for age, gender and language. The specific aims of this study were to: (1) determine if individuals with ASD have significant acoustic-perceptual differences in their receptive and expressive prosody; (2) identify the interrelationship between prosody and language, motor speech, and pitch processing abilities; and (3) investigate if there is an association between group membership (ASD vs. TD) and trained listener ratings of overall “naturalness”

(*natural* versus *unnatural*) of the speakers' speech. The findings of this study support that some individuals with autism perform with significantly less accuracy on receptive and expressive prosody tasks and had significantly longer duration of utterances in comparison to TD controls. There was a significant positive relationship between receptive vocabulary and expressive prosody in the ASD group, supporting the "Theoretical Interaction Model", while expressive vocabulary and speech motor control did not explain variability in expressive prosody above and beyond receptive vocabulary.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	Page v
ABSTRACT	vii
LIST OF TABLES	xiii
LIST OF FIGURES	xiv
CHAPTER	
1. INTRODUCTION	1
1.1 Autism Spectrum Disorder	1
1.2. Atypical Prosody in ASD.....	4
1.2.1 Language Ability	5
1.2.2 Speech Motor Control.....	7
1.2.3 Auditory Processing.....	8
1.3. Statement of the Problem.....	9
1.4. Purpose of the Study	10
1.5. Research Questions and Hypotheses	11
2. REVIEW OF THE LITERATURE	14
2.1. Autism Spectrum Disorder	14
2.2 Theoretical Links Between Autism and Atypical Speech, Voice and Prosody	19
2.2.1 The Social Feedback Loop.....	19
2.2.2 Social Reciprocal Interaction Impairment	20
2.2.3 Speech Attunement Framework.....	20
2.2.4 Theory of Mind	21

2.2.5 Speech Motor Control.....	23
2.2.6 Neurological Differences	25
2.3 Theoretical Links between Autism and Atypical Sensory Processing	27
2.3.1 Weak Central Coherence	27
2.3.2 Enhanced Perceptual Functioning.....	28
2.3.3 Allocation of Attention	28
2.4 Speech Differences in Autism Spectrum Disorder	29
2.5 Prosody in ASD	31
2.5.1 Pragmatic and Affective Prosody in ASD	34
2.5.2 Grammatical Prosody in ASD	35
2.5.3 Atypical Prosody in the Context of Theories.....	43
2.6 Auditory Processing in ASD.....	43
2.6.1 Enhanced Pitch Processing	44
2.6.2 Enhanced Pitch Memory and Labeling.....	59
2.6.3 Pitch Perception Summary.....	62
2.6.4 Pitch Processing in the Context of Cognitive Theories	66
2.6.5 Implications of Auditory Processing Differences in ASD.....	66
2.6.6 Limitations to Auditory Processing in ASD Findings	68
2.7 Implications of Atypical Speech, Voice, and Auditory Processing.....	69
3. METHODOLOGY	71
3.1 Research Design.....	71
3.2 Participants.....	72
3.3 Perceptual Judges.....	74

3.4 Measures	75
3.4.1 Expressive and Receptive Vocabulary.....	75
3.4.2 Profiling Elements of Prosody in Speech-Communication	78
3.4.3 Acoustic Measurements of Pitch and Duration	80
3.4.4 Oral Diadochokinetic Rate.....	81
3.4.5 Frequency Pattern Test	81
3.4.6 Naturalness Perceptual Rating Scale	83
3.5 Procedures.....	83
3.5.1 Standardized Language Testing.....	86
3.5.2 Assessment of Prosody	86
3.5.3 Assessment of Speech Motor Control.....	88
3.5.4 Assessment of Pitch Processing.....	89
3.5.5 Assessment of Speech, Voice, and Prosodic Naturalness.....	90
3.6 Statistical Procedures	92
4. RESULTS	95
4.1 Expressive and Receptive Vocabulary.....	95
4.2 Prosody	96
4.2.1 PEPS-C Results.....	96
4.2.2 Acoustic Measurements of Pitch and Duration	100
4.3 Predictors of Expressive Prosody	101
4.4 Speech, Voice, and Prosodic Naturalness.....	104
4.4.1 Reliability of the Naturalness Perceptual Rating Scale	105
4.4.2 Association Between Group Membership and Naturalness Ratings	105

5. CONCLUSIONS	110
5.1 Research Questions, Hypotheses, and Participant Demographics.....	110
5.2 Prosody	112
5.2.1 PEPS-C Findings	113
5.2.2 Prosody: Acoustic Measurements of Pitch and Duration	119
5.3 Predictors of Expressive Prosody	120
5.4 Speech, Voice, and Prosodic Naturalness.....	126
5.5 Summary	129
5.6 Limitations and Future Directions	130
APPENDICES	
A. DEMOGRAPHIC SCREENING QUESTIONS.....	147
B. SPEECH MOTOR CONTROL TASK	151
C. SPONTANEOUS SPEECH SAMPLE	152
D. NATURALNESS PERCEPTUAL RATING SCALE.....	153
E. SUMMARY OF RESULTS WITHIN AND BETWEEN GROUPS.....	154
F. BOXPLOT COMPARING PEPS-C EXPRESSIVE PROSODY COMPOSITE SCORES BETWEEN GROUPS.....	155
G. BOXPLOT COMPARING PEPS-C RECEPTIVE PROSODY COMPOSITE SCORES BETWEEN GROUPS.....	156
BIBLIOGRAPHY	157

LIST OF TABLES

Table	Page
Table 1. Demographic Characteristics of the Participants.....	133
Table 2. PEPS-C 2015 Results	134
Table 3. Inter-Judge Reliability for the PEPS-C 2015 Expressive Prosody Tasks.....	135
Table 4. Intra-Judge Reliability for the PEPS-C 2015 Expressive Prosody Tasks.....	136
Table 5. Acoustic Measurements of Pitch & Duration.....	137
Table 6. Frequency Pattern Test Right Ear Results, separated by group and rating.....	138
Table 7. Frequency Pattern Test Left Ear Results, separated by group and rating.....	139
Table 8. Frequency Pattern Test Results	140
Table 9. Inter-Judge Reliability for the Naturalness Perceptual Rating Scale.....	141
Table 10. Chi-Square Analysis Results for the Naturalness Perceptual Rating Scale	142
Table 11. Systematic Literature Review of PEPS-C Results	143
Table 12. Descriptive Analysis for the Naturalness Perceptual Rating Scale.....	144

LIST OF FIGURES

Figure	Page
Figure 1. Lexical Stress Duration Boxplot.....	145
Figure 2. Interaction of Group and PPVT Scores on Expressive Prosody.....	146

CHAPTER 1

INTRODUCTION

1.1 Autism Spectrum Disorder

Autism Spectrum Disorder (ASD) is a lifelong neurodevelopmental condition characterized by differences in social communication and social interaction, as well as restricted, repetitive patterns of behavior, interests, or activities (American Psychiatric Association, 2013). The estimated prevalence of ASD is one in 59 children in the United States, demonstrating a 150% increase since the year 2000 (Baio et al., 2018). Individuals are diagnosed with ASD according to the diagnostic criteria set forth by the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) (American Psychiatric Association [APA], 2013). The severity of autism ranges from mild to severe (APA, 1994; 2013).

One of the features that distinguish the communication skills of some individuals with ASD is atypical prosody (e.g., McCann & Peppe, 2003; Shriberg, Paul, McSweeney, Klin, Cohen, & Volkmar, 2001). Prosody concerns the suprasegmental properties of speech and the voice (e.g., fundamental frequency/pitch, duration, intensity/loudness, vocal quality, and stress) that enhance or change the pragmatic, affective, or grammatical meaning of an oral utterance (Shriberg, Kwiatkowski, Rasmussen, Lof, & Miller, 1992; Shriberg et al., 2001). Prosodic skills are critical for the understanding and production of pragmatic, affective, and grammatical information (DePape, Chen, Hall, & Trainor, 2012; Peppe, McCann, Gibbon, O'Hare, & Rutherford, 2006). Listeners' ability to use the prosodic features of oral language and decode it plays a pivotal role in social and language development (Chiew & Kjeldgaard, 2017). Some individuals may have a

prosodic disorder that negatively impacts the development or mastery of prosodic skills (Velleman, 2015). Prosodic disorders can be caused by differences in the acoustic-perceptual features of the voice (i.e., form) or the use of prosody to convey pragmatic, affective, and grammatical meaning (i.e., function).

Although atypical prosody is not a universal characteristic of ASD, research supports that 33% (Kargas, Lopez, Morris, & Ready, 2016) to 60% (Nadig & Shaw, 2015) of research participants with ASD have differences in receptive or expressive prosodic abilities as compared to control groups. It is widely reported that at least some individuals on the autism spectrum have differences in their prosodic abilities (e.g., Chiew & Kjelgaard, 2017; DePape et al., 2012; Diehl & Paul, 2012; Diehl & Paul, 2013; Gebauer, Skewes, Hørlyck, & Vuust, 2014; Kargas et al., 2016; McCann, Peppe, Gibbon, O'Hare, & Rutherford, 2007; Paul, Augustyn, Klin, & Volkmar, 2005; Peppe, McCann, Gibbon, O'Hare, Rutherford, 2007).

With respect to expressive prosody, individuals with ASD have been described in the literature as sounding “monotonic”, “machine-like”, “sing-song”, “awkward”, “odd”, “labored”, and “different” (Andrianopoulos, Zaretsky, McGuigan, & Warsaw, 2015; Filipe, Frota, Castro, & Vicente, 2014; Grossman, 2015; Grossman, Edelson, & Tager-Flusberg, 2013; Kanner, 1971; Shriberg et al., 2001). Empirical research supports that human listeners can perceive speech differences under controlled conditions that distinguish children with ASD from their typically developing peers (TD) during the production of oral narratives (Andrianopoulos et al., 2015). Andrianopoulos et al. (2015) developed an acoustic perceptual rating instrument comprised of seven linguistic variables (e.g., story sequencing, topic organization, story details) and six speech or

acoustic variables (e.g., articulation, fluency, intonation, rate, pitch). Speech Language Pathologists (SLPs; n=20) and naïve listeners (non-SLPs; n=20) applied the instrument to rate the oral narratives produced by a group of children with ASD (n=24; age range 8;0-11;0 years) and TD peers (n=23) matched for age, gender, and receptive language. For the main effect diagnosis, Andrianopoulos et al. (2015) found significant between-group differences for four (4) out of 13 variables: story sequencing ($p=0.008$); articulation ($p=0.001$); fluency ($p=0.005$); and rate ($p=0.02$). Trained and naïve listeners rated the group with ASD as sounding “different” than their TD peers based on these variables. In line with these results, Grossman (2015) reported that naïve listeners rated individuals with high-functioning autism as “socially awkward” at a significantly higher rate than control groups after attending to 1- to 3-second visual and/or auditory clips that were captured during a story-retelling task.

In addition to the perceptual descriptions of expressive prosody, acoustic findings related to prosodic abilities range from longer duration of utterances (Bonneh, Levanon, Dean-Pardo, Lossos, & Adini, 2011; Boucher, 2013; Filipe et al., 2014; Grossman, Bemis, Skwerer, & Tager-Flusberg, 2010; Velleman, Andrianopoulos, Boucher, Perkins, Averbach, Currier, Marsello, & Van Emmerik, 2009) to increased pitch variability (Diehl, Watson, Bennetto, McDonough, & Gunlogson, 2009), exaggerated or decreased pitch (DePape et al., 2012); and inappropriate phrasing, stress and resonance (Boucher, 2013; Shriberg et al., 2001; Velleman et al., 2009).

Regarding prosodic function, it has been suggested that individuals with autism have impaired pragmatic and affective prosody while grammatical prosodic abilities are relatively spared (McCann & Peppe, 2003; Shriberg et al., 2001). Other research has

demonstrated that individuals with autism do have impaired grammatical prosody. Specific differences in the acoustic or perceptual aspects of grammatical prosody have been identified by empirical evidence, such as: phrasing errors, increased repetitions and revisions, and reduced rate of speech during conversation (Shriberg et al., 2001); significantly poorer performance when “chunking” speech using prosody to disambiguate syntactically ambiguous phrases (Peppé, Cleland, Gibbon, O’Hare, & Martínez Castilla, 2011); and using pitch contours to understand (Diehl & Paul, 2013; Peppe et al., 2011) or produce questions versus statements (Diehl & Paul, 2013; Filipe et al., 2014). The range of perceptual and acoustic characteristics of atypical prosody across domains suggests that a spectrum of prosodic abilities exists among individuals with autism.

Studies have also demonstrated that individuals on the autism spectrum are significantly less accurate than control groups on receptive prosody tasks requiring them to disambiguate a question from a statement, identify *like* versus *dislike*, perceive emphasis on one word versus another in a sentence, and identify if muffled vocalizations sound the *same* or *different* with respect to intonation (e.g., Diehl & Paul, 2013; Gargan and Andrianopoulos, in progress). Other studies show contradictory findings, with no significant differences between groups on identifying questions versus statements (e.g., Filipe et al., 2014; Peppe et al., 2007) or perceiving emphasis on one word versus another (e.g., Peppe et al., 2007).

1.2. Atypical Prosody in ASD

There are divergent findings related to the characteristics of atypical prosody among individuals with autism. There is also a lack of consensus with respect to the underlying cause of these differences. It has been hypothesized that speech and prosodic

differences in ASD could be due to: (1) one's ability to 'tune in' to the speech environment but failure to 'tune up' their speech behaviors (Diehl and Paul, 2013; Diehl & Paul, 2012; Shriberg et al., 2011); (2) language deficits (Peppe et al., 2007; Peppe et al., 2011; Lyons, Simmons, & Paul, 2014); (3) social reciprocal interaction impairment (Nakai, Takashima, Takiguchi, & Takada, 2014); (4) the nature of the research task (Diehl and Paul, 2013; Peppe et al., 2007); (5) impaired auditory processing (Bonneh et al., 2011; Peppe et al., 2006; Peppe et al., 2007); and (6) speech motor control deficits (Peppe et al., 2007; Diehl & Paul, 2012; Velleman, Andrianopoulos et al., 2009; Andrianopoulos et al., 2015).

It is unlikely that a universal definition of atypical prosody with a single underlying cause will apply to all individuals on the autism spectrum. It is reasonable to hypothesize that communication difficulties such as atypical prosody have different origins, including language, speech motor control, and auditory processing differences. Since prosody, speech, language, and auditory processing are intertwined, it is important to systematically study and quantify each of these variables in one empirical investigation to determine the inter-relationships of these processes and their impacts on prosody. A brief overview of language ability, speech motor control, and auditory processing in ASD is provided below (please see Chapter 2 for an in-depth discussion and findings as they relate to prosody for each topic area).

1.2.1 Language Ability

Communication and language abilities among individuals with ASD range from significantly impaired to above average. Some individuals with ASD do not develop functional communication, while others have advanced linguistic knowledge but

difficulty with the use of language in social communication (Tager-Flusberg, 1981, 1995; Landa, 2000; cited in Paul et al., 2005). Some studies show that there is a strong connection between language ability and prosodic skills. DePape et al. (2012) reported differences in pitch excursions to mark information structure among individuals with ASD and high language abilities (ASD-high) compared to those with moderate language abilities (ASD-moderate). The individuals in the ASD-high group used larger pitch ranges, whereas the individuals in the ASD-moderate group used smaller pitch ranges, corresponding with “sing-song” or “monotone” descriptors of prosody, respectively. Grossman et al. (2010) reported that the preserved language skills among the participants with HFA in their study might have supported their performance on a lexical stress perception task.

Although some authors state that language ability impacts prosody, results from Gargan and Andrianopoulos (in progress) do not fully support this hypothesis. Six individuals with ASD who had average to moderately high language scores scored below competence level on at least one prosody task. Only one individual with ASD, who had average language, scored above competency level on all tasks. Additionally, one TD participant with average language skills scored below competence on a phrase stress understanding task and one TD participant with average to moderately high language scored below competence on a lexical stress understanding task. The mixed presentation of findings, including strong receptive prosody skills and weak expressive prosody skills, or vice versa, regardless of language ability, suggests that there might be other underlying mechanisms impacting prosodic ability, such as speech motor control involvement

(Andrianopoulos et al., 2015; Diehl & Paul, 2012), auditory processing deficits (Peppe et al., 2006; Peppe et al., 2007), or a combination of the two (Bonneh et al., 2011).

1.2.2 Speech Motor Control

Generalized motor impairments are widely observed among children with ASD and motor disturbances are outlined in the DSM-5 as a feature of ASD. Individuals with autism demonstrate gross motor deficits (Pusponegoro et al., 2016), dyspraxia (Dziuk, Larson, Apostu, Mahone, Denckla, & Mostofsky, 2007); gait abnormalities (Shetreat-Klein, Shinnar, & Rapin, 2014), hypotonia, and apraxia (Harris, 2017; Kern et al., 2013). Overall, there is agreement in the literature that motor impairments are prevalent among individuals with autism (Ming, Brimacombe, and Wagner, 2007; Fulceri et al., 2019; Gernsbacher et al., 2008).

Based on the findings of generalized motor impairments in ASD, some researchers suggest that a subset of individuals with ASD may also have underlying motor speech impairments (Adams, 1998; Velleman, Andrianopoulos et al., 2009; Gargan & Andrianopoulos, in progress). The following characteristics of speech and voice in ASD appear to be consistent with a deficit in speech-motor control abilities: residual articulation distortion errors, uncodable utterances, and inappropriate utterances in terms of phrasing, stress, and resonance (Shriberg et al., 2001); increased loudness (Shriberg et al., 2001); deficits in focal oral motor skills and sequencing oromotor gestures (Velleman et al., 2010); low maximum phonation times for [a] and [f] prolongation; less varied syllable durations (Velleman et al., 2010); and “slurred” or “imprecise” sequencing of speech sounds (Gargan & Andrianopoulos, in progress).

1.2.3 Auditory Processing

It has been estimated that 96% of individuals with ASD have atypical sensory reactivity in visual, tactile, or auditory modalities (Mayer, Hamment, & Heaton, 2016; O’Riordan & Passetti, 2006; Karhson & Golob, 2016). With respect to the auditory domain, some studies have demonstrated that enhanced auditory processing occurs in approximately 1 in 5 children with autism (Jones, Happe, Baird, Simonoff, Marsden, Tregay et al., 2009). An enhanced perception of detail and reduced attention to the “big picture” results in an impaired ability to use context to interpret meaning (Jolliffe & Baron-Cohen, 1999). Individuals with autism who focus on local versus global auditory details are likely to have compromised perceptual, cognitive, language, and social development (Jones et al., 2009).

Some individuals with autism have difficulty understanding pitch-mediated cues such as prosody, despite enhanced abilities with pitch processing on a local level (McCann & Peppe, 2003; O’Connor, 2012). For example, Jarvinen-Pasley, Wallace, Ramus, Happe, and Heaton (2008) demonstrated that children with autism exhibited superior perceptual processing of speech and musical stimuli (i.e., they pointed to visuals of pitch contours that matched what they heard more accurately than controls), but performed with significantly less accuracy on a comprehension task (i.e., they answered comprehension questions about the speech stimuli less accurately than controls). Some findings support that individuals with ASD focus more on local level information, such as pitch, rather than linguistic information (e.g., prosody).

1.3. Statement of the Problem

Judgments of social awkwardness based on facial and vocal cues are formed rapidly, sometimes within 1 to 3 seconds of exposure (Grossman, 2015). Atypical prosody can result in many communication breakdowns, such as: not understanding the main point of an utterance; producing a misleading utterance; difficulty disambiguating between noun phrases or compound nouns; difficulty disambiguating a verb from a noun; difficulty understanding or expressing one's own feelings; or being perceived as "socially awkward" or "different" (Andrianopoulos et al., 2015; Peppe, 2006; Grossman, 2015). These subtle and very salient differences in the prosody of some individuals with autism can have significant negative effects on social interactions and vocational success (DePape et al., 2012; Grossman, 2015; Shriberg et al., 2001).

As previously mentioned, there is a lack of consensus regarding the characteristics and underlying cause of atypical prosody among some individuals with autism. There are several limitations to the current empirical research, such as: small sample sizes; the examination of one domain of prosody at a time; the use of either acoustic or perceptual measurements (not both); examination of one underlying mechanism at a time (e.g., language ability); the nature of the research tasks; and a lack of a formal training protocol for the assessment of social awkwardness based on spontaneous speech samples. To date, researchers have not investigated the inter-relationships among language, motor speech, and auditory processing, and the impact of these variables on prosody in a well-controlled sample of individuals with ASD compared to a TD group.

1.4. Purpose of the Study

The aims of this study are to: (1) determine if individuals with ASD have significant differences in their receptive prosody, expressive prosody, and duration (seconds) and pitch of their expressive Lexical Stress utterances; (2) identify the interrelationships between prosody and language, motor speech, and pitch processing abilities; and (3) investigate whether there is an association between group membership (ASD vs. TD) and trained listener ratings of overall “naturalness” of their speech, voice, and prosody based on a 20-second connected speech sample.

Researchers and clinicians need to be able to reliably and validly quantify, identify, and distinguish prosodic disorders using operational metrics comprised of acoustical, perceptual, and motoric behaviors observed on exam. This will enable researchers and clinicians to establish evidence-based interventions for prosodic disorders. An understanding of processing of auditory information in ASD will shed light on how spoken information is processed and in turn, how it affects the production of speech and prosody (e.g., Kargas et al., 2016). Research related to perceptual ratings of “social awkwardness” has gained recent attention, but empirical findings in this area are sparse. It is important to determine how listeners perceive individuals with ASD based on acoustic samples of their speech, voice, and prosody. This could lead to the development and implementation of programs (e.g., anti-bullying programs) to increase awareness of these differences and promote social communication success for individuals with autism. Identifying the features of speech, voice, and prosody that can be reliably perceived by the human ear in order to differentiate individuals with ASD from those without ASD will also be useful in a clinical or evaluation setting.

1.5. Research Questions and Hypotheses

It is hypothesized that individuals with ASD will have significant differences in at least one prosodic domain based on perceptual and acoustic measures; a subgroup of individuals with ASD who have atypical prosody will have impaired language, motor speech, and/or auditory processing abilities; and some individuals with ASD will be judged by human listeners as sounding “awkward” as compared to their TD peers. Due to the heterogeneous nature of this disorder, the findings will characterize phenotypes of ASD. The research questions and hypotheses are as follows:

1. Do children with ASD between the ages of 7;10-19;0 years perform with significantly less accuracy on receptive and expressive prosody tasks based on operational metrics using perceptual and acoustic measures as compared to a TD group matched for age, gender, and language?
 - a. H0: Children with ASD between the ages of 7;10-19;0 years do not perform with significantly less accuracy on receptive and expressive prosody tasks based on perceptual and acoustic measures as compared to a TD group matched for age, gender, and language.
 - b. H1: Children with ASD between the ages of 7;10-19;0 years do perform with significantly less accuracy on receptive and expressive prosody tasks based on perceptual and acoustic measures as compared to a TD group matched for age, gender, and language.

The investigator expects to reject the null hypothesis and support the alternative hypothesis. It is predicted that the outcomes of this study will

support the presence of acoustic and perceptual differences in prosody among those with ASD during elicited prosody tasks as compared to TD peers.

2. Is there a significant linear relationship between expressive prosodic abilities and language, motor speech, and pitch processing scores?
 - a. H0: There is not a significant linear relationship between expressive prosodic abilities and language, motor speech, and pitch processing scores.
 - b. H1: There is a significant linear relationship between expressive prosodic abilities and language, motor speech, and pitch processing scores.

The author expects to reject the null hypothesis and accept the alternative hypothesis. It is expected that individuals who perform with atypical precision or low accuracy on expressive prosody tasks will have below average language abilities, impaired motor speech, and/or enhanced pitch processing.

3. Is there an association between group membership (ASD vs. TD) and ratings of naturalness (“natural” vs. “unnatural”) based on two trained listeners’ perceptual ratings of speech, voice, and prosody based on 20-second connected speech samples?
 - a. H0: There is no association between group membership (ASD vs. TD) and ratings of naturalness (“natural” vs. “unnatural”) based on two trained listeners’ perceptual ratings of speech, voice, and prosody based on 20-second connected speech samples.

- b. H1: There is an association between group membership (ASD vs. TD) and ratings of naturalness (“natural” vs. “unnatural”) based on two trained listeners’ perceptual ratings of speech, voice, and prosody based on 20-second connected speech samples.

The author expects to reject the null hypothesis and support the alternative hypothesis (Kanner, 1971; Filipe et al., 2014; Grossman et al., 2013; Nadig & Shaw, 2012; and Bonnef et al., 2011).

CHAPTER 2

REVIEW OF THE LITERATURE

2.1. Autism Spectrum Disorder

The prevalence of Autism Spectrum Disorder (ASD) in the United States is one in 59 children (Baio et al., 2018). ASD primarily affects an individual's social communication, social interaction, and behavior (American Psychological Association, 2013). Individuals are diagnosed with ASD according to the diagnostic criteria set forth by the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) (APA, 2013).

Autism is diagnosed in males and females, and it occurs in all racial/ethnic communities, yet disparities exist across these groups with respect to prevalence and intellectual disability. For instance, the disorder is four times more common in males versus females. In addition, prevalence estimates are 7% higher for non-Hispanic white children in comparison to non-Hispanic black children and 22% higher among white children in comparison to Hispanic children. With respect to intellectual abilities, the Autism and Developmental Disabilities Monitoring (ADDM) Network reported that among the 8-year-old children from nine communities in the 2014 surveillance year, 31% of children with ASD had intellectual disability, 25% were classified as having borderline IQ (i.e., 71-85), and 44% had average to above average intellectual abilities (i.e., IQ >85). Interestingly, females were more likely than males to have an intellectual disability and males were more likely than females to have average or above average IQ. With respect to race/ethnicity, only 22% of white children were classified as having intellectual

disability, while 35% of Hispanic children and 44% of black children had intellectual disability (Baio et al., 2018).

Autism is a heterogeneous disorder with a lack of a single diagnostic marker (Baio et al., 2018). One identified risk factor of autism is genetics. Younger siblings of a child with ASD have a 1:5 (20%) chance of receiving a diagnosis of autism. Additionally, a baby brother of a child with ASD has a 1:4 (25%) chance, a baby sister of a child with ASD has a 1:9 (11%) chance, and an infant who has more than one older sibling with ASD has a 1:3 (33%) chance of receiving a diagnosis of autism (Ozonoff et al. (2011). There is also research to support structural and functional differences in the brains of individuals with autism, which may be linked with the impaired cognitive, social, emotional, and language functions associated with the disorder (Courchesne et al., 2004). For instance, it has been reported that children with ASD between the ages of 2 to 4 years have deviant growth patterns in the cerebral, cerebellar, and limbic structures, which are critical for the normal development of higher-order cognitive functions (Courchesne, 2004; Sparks et al., 2002; Waldie & Saunders, 2014).

Individuals can be reliably and validly diagnosed with ASD at 2 years of age (Lord et al., 2006; Kleinman et al., 2008), yet most are not diagnosed until four years of age (Baio et al., 2008). According the ADDM network report from the 2014 surveillance year, 85% of children had documented developmental concerns by 36 months of age, but only 42% had a comprehensive evaluation on record by this age. Furthermore, 39% did not receive a comprehensive evaluation until at least 4 years of age. The median age of earliest diagnosis in 2014 was 52 months (Baio et al., 2018). Parent studies are consistent in that the area of first concern among children later diagnosed with ASD is impaired

communication. Social skills, challenging behaviors, and motor skills are also among the top areas of early parental concern (Kozlowski, Matson, Horovitz, Worley, & Neal, 2011).

Although the severity and presence of symptoms vary across the spectrum, a notable feature of ASD is qualitative impairment in communication. Language abilities among individuals with ASD range from significantly impaired to above average. Moreover, some individuals with ASD do not develop functional communication while others have rather sophisticated linguistic knowledge but concomitant difficulty with the use of language in social contexts (American Speech Language Hearing Association, 2015). Among the individuals with ASD who develop spoken language, their speech is often delayed or atypical on several levels, including motor speech abilities, prosody, and acoustic-perceptual vocal features. Differences in the production of speech, prosody, and voice are a common clinical feature of some individuals with ASD and evidence shows they are observed during infancy. Thus, they are one of the earliest characteristics of the disorder to appear (Schoen, Paul, and Chawarska, 2011).

Speech abnormalities are not present in all individuals with ASD, but those who do have differences in their speech have been described as sounding “odd” and “awkward” (Filipe et al., 2014; Grossman et al., 2013; Nadig & Shaw, 2012; and Bonneht et al., 2011). Evidence shows that when these speech differences are present they remain static, even when other aspects of language improve (Kanner, 1971). The characteristics of the speech, voice, and prosody among some individuals with ASD often present a significant obstacle to social integration and vocational acceptance (Shriberg et al., 2001).

There is also agreement that individuals with ASD have generalized motor impairments. In Kanner's early observational reports, he documented an absence of crawling, clumsy gait, and impaired gross motor movements among some of his patients. Consistent with Kanner's early reports, individuals with autism continue to demonstrate gross motor deficits or delays (Pusponegoro et al., 2016; Ming et al., 2007), dyspraxia (Dziuk et al., 2007); gait abnormalities (Shetreat-Klein et al., 2014), hypotonia, and apraxia (Harris, 2017; Kern et al., 2013; Ming et al., 2007). Motor difficulties in autism are present early in life and may serve as a risk indicator for ASD. Some empirical research suggests that a disruption in early motor development at 6 months predicts expressive language abilities at 30 and 36 months of age (LeBarton & Landa, 2019). Overall, there is agreement in the literature that motor impairments are prevalent among some individuals with autism (Ming et al., 2007; Fulceri et al., 2019).

A third prominent characteristic of ASD is hyper- or hyposensitivity to sensory information. In 1943, Leo Kanner's description of autism highlighted an increased attention to detail among individuals with autism and an 'inability to experience wholes without full attention to the constituent parts' (Kanner, 1943, p. 247). He also stated that one of the children in his case studies could 'hum and sing many tunes accurately' at 1 year of age (Kanner, 1943, p. 217). Since Kanner's original description of autism, there is a growing body of evidence suggesting that individuals with ASD have exceptional perceptual abilities. For instance, some individuals with autism can identify the brand of a vacuum cleaner based on the sound alone (Happé & Frith, 2006).

It has been estimated that 96% of individuals with ASD have atypical sensory reactivity in one or more of the following sensory modalities: (1) the visual modality

(e.g., focusing on local versus global details); (2) the tactile modality (e.g., seeking or avoiding certain textures); and/or the auditory modality (e.g., enhanced pitch processing or sensitivity to loud sounds) (Mayer et al., 2016; O’Riordan & Passetti, 2006; Karhson & Golob, 2016).

With respect to the auditory domain, some studies have demonstrated that enhanced auditory processing is not a universal characteristic in individuals with autism (Jarvinen-Pasley & Heaton, 2007). It has been estimated that 1 in 5 children with autism exhibit enhanced pitch discrimination abilities (Jones et al., 2009). The symptoms may persist into adulthood and have an impact on quality of life (Mayer et al., 2016), in that they may cause distress to some individuals. For instance, one case study indicates that an individual with autism reportedly began to associate all sounds with musical notes as a child. When travelling on the Paris Metro, he became distressed when he noticed the opening-door signal on the train had a different pitch frequency than the one used on the London Underground (Heaton, Davis & Happe, 2008a).

The causes and consequences of atypical sensory processing are not well understood. There is increasing consensus, however, that atypical sensory processing may result in avoidance of social stimuli. In turn, the development of language, social, and cognitive abilities may be constrained (Mayer et al., 2016). Processing sensory information (e.g., auditory stimuli) is critical in the typical neurodevelopmental trajectory. Individuals with autism who focus on local versus global auditory details, or place their hands over their ears to protect them from sound, are likely to compromise perceptual, cognitive, language, and social development (Jones et al., 2009).

2.2 Theoretical Links Between Autism and Atypical Speech, Voice and Prosody

Several theories have been considered to explain the underlying cause of speech, voice, and prosodic impairments among some individuals with ASD. A brief review of the following theories will be discussed in the sections that follow: the social feedback loop; social reciprocal interaction impairment; speech attunement framework; Theory of Mind; speech motor control; and neurological differences.

2.2.1 The Social Feedback Loop

According to the social feedback loop framework, a child who produces speech or speech-like vocalizations is more likely to receive immediate, positive feedback from an adult than if the child produced a non-speech vocalization. Vocal interactions between adults and children are dependent on each other, in that the adults respond contingently to the child's speech-like vocalizations and the characteristics of the child's future vocal productions are contingent on the adult's previous response (Warlaumont et al., 2014). These interactions are hypothesized to play a role in the development of speech over time in individuals with and without autism. The feedback loop may be impaired, however, in children with autism due to: (1) fewer speech-like vocalizations among children with ASD and among toddlers at-risk for autism (Warlaumont, 2014; Plumb & Wetherby, 2013; Chenausky Nelson & Tager-Flusberg, 2017; Schoen et al., 2011); (2) a higher proportion of atypical vocalizations and distress vocalizations in toddlers with ASD (Plumb and Wetherby, 2013); (3) significantly fewer consonant types and canonical shapes in some infants with autism (Paul et al., 2011); (4) different responses to vocalizations among parents of those with ASD; and (5) impaired ability among infants with ASD to learn from adult responses (Warlaumont et al., 2014; Chenausky et al.,

2017). Paul et al., (2011) states that infants with autism may be less interested in the turn-taking exchanges (i.e., “conversations”) with communication partners, resulting in less back-and-forth babbling between infant and caregiver. A diminished social feedback loop negatively impacts an infant’s ability to learn about the social effects of their vocal productions and in turn negatively effects speech development (Warlaumont et al. 2014).

2.2.2 Social Reciprocal Interaction Impairment

Some researchers have suggested that delayed and continued impairments in speech production are due to intellectual and social impairments. As outlined in the DSM-5, individuals with autism have deficits in the area of social communication and interaction, including deficits in social-emotional reciprocity. Thus, individuals with autism have difficulties in a variety of areas such as: adjusting communication in various contexts (e.g., talking to a baby versus talking to an adult), difficulties with conversational turn-taking, and initiating or responding to social interactions. Studies have demonstrated that impaired social communication skills affect prosody, voice, and speech development.

2.2.3 Speech Attunement Framework

The speech attunement framework posits that a child must ‘tune in’ to the speech produced by others in their environment and ‘tune up’ their own speech in order to acquire articulate speech as well as prosody and vocal productions that match the native language. Some researchers suggest that individuals with autism have enhanced auditory perceptual abilities allowing them to “tune in” to vocal productions in the community, but they do not “tune up” their speech to sound like others in their community possibly

due to difficulty with pragmatics and social reciprocity. (Shriberg, Paul, Black, & van Santen, 2011).

2.2.4 Theory of Mind

Theory of Mind (ToM) refers to the ability to infer the mental states of others and make an inference about what someone knows, wants, feels, or believes. It is a crucial component of communicative and social competence (Baron-Cohen, Leslie, & Frith, 1985; Kaland, Krahmer, & Swerts, 2012). Research supports that typically developing children demonstrate the ability to attribute the mental states of others during the second year of life (e.g., Shantz, 1983; cited in Baron-Cohen et al., 1985) and it is well developed by four years of age (Wimmer & Perner, 1983; cited in Baron-Cohen et al., 1985).

It is widely agreed upon that ToM is a core deficit underlying impaired verbal and nonverbal communication among individuals with ASD (Baron-Cohen et al., 1985; Rutherford, Baron-Cohen, & Wheelwright, 2002; Chevallier, Noveck, Happe, & Wilson, 2011). Deficits in ToM have been frequently demonstrated using the “false belief” task (Baron-Cohen et al., 1985). Impaired ToM has also been demonstrated using tasks that require individuals to infer mental states by looking at someone’s eye region (e.g., Baron-Cohen et al., 1997) or facial expression (Grossman & Tager-Flusberg, 2012). More recently, the ability to produce and understand prosody has been used to test the ToM account in this population. To date, there are limited and contradictory findings related to the link between ToM and: (1) expressive prosodic abilities; and (2) emotion recognition abilities based on vocal cues (Grossman & Tager-Flusberg, 2012).

There is substantial support that individuals with ASD have difficulty with the production of pragmatic prosody, such as contrastive stress (Peppe et al., 2007). In order to produce contrastive stress appropriately, the speaker must know what information is new to the listener and use prosody to highlight certain aspects of an utterance. Previous findings indicate that individuals with ASD exhibit errors in contrastive stress production, such as misplaced stress on words (Shriberg et al., 2001) or stress on the wrong words (McCaleb & Prizant, 1985). It has been argued that these difficulties may be associated with decreased perspective-taking skills (Shriberg et al., 2001). Findings from other studies, however, do not support a link between the production of contrastive stress and ToM in autism. For example, Kaland, Swerts, and Krahmer (2013) found that individuals with autism used perspective-taking skills appropriately and had accuracy levels comparable to the control group. Differences were noted, however, in the acoustic features of their productions. For instance, the individuals with ASD produced smaller pitch ranges and were described as sounding monotonous. Thus, there was no support for the link between impaired ToM and impaired functional use of contrastive stress. Rather, acoustic findings revealed deficits in prosodic form (i.e., pitch range). The authors state that the findings could have been limited by the participant characteristics and the experimental task (Kaland et al., 2013).

ToM has also been considered as a core feature underlying the impaired ability to infer the mental states of others based on vocal information. Some findings indicate that individuals with ASD perform worse than controls on experimental tasks requiring them to label a speaker's mental state while listening to audio recordings (Rutherford et al., 2002; Golan, Baron-Cohen, Hill, & Rutherford, 2007). Other studies, however, show that

individuals with ASD perform comparably to controls during receptive pragmatic or affective prosody tasks (e.g., Grossman et al., 2010). Additionally, Chevallier et al. (2011) found that individuals with ASD had similar performance to controls when labeling basic emotions (e.g., happy), physical states (e.g., cold), and social emotions (e.g., guilt) based on vocal cues. The individuals with ASD did not demonstrate an impairment specific to stimuli involving advanced ToM processing (e.g., social emotions). Rather, they had slower reaction times and had subtle difficulties when increased cognitive demands were placed on them. The findings reflect decreased identification of emotions based on vocal cues in general when cognitive demands are increased, rather than a deficit due to the extent of mindreading abilities required. This finding sheds light on why some individuals with ASD have difficulty interpreting or using information conveyed through the voice during complex social interactions, but may perform accurately on simple and explicit experimental tasks (Chevallier et al., 2011).

Thus, there are findings to both support and refute the idea that impaired ToM underlies atypical pragmatic prosody abilities and the ability to infer the mental states of others based on vocal cues. Differences in methodology and participant characteristics across studies may contribute to the contradictory findings. Some research groups hypothesize that deficits in the ability to recognize emotions through vocal cues are better explained by diminished social motivation and social attention (Chevallier et al., 2011).

2.2.5 Speech Motor Control

Some researchers suggest that a subset of individuals with ASD may also have underlying motor speech impairments (Adams, 1998; Velleman et al., 2009 Gargan &

Andrianopoulos, in progress). Chenausky, Brignell, Morgan, and Tager-Flusberg (2019) reported that four subgroups emerged among 54 minimally verbal children with ASD who participated in a speech praxis test as follows: (1) speech within normal limits (n=12), (2) non-childhood apraxia speech impairment (n=16), (3) suspected childhood apraxia of speech (n=13), and (4) insufficient speech to rate (n=13). The authors reported that receptive vocabulary predicted the number of different words produced during a semi-structured language sample for the first two groups, while speech production ability predicted the number of different words produced by the latter two groups. It remains unknown if speech motor control is the primary contributor to limited spoken language among some children with autism.

There is still limited agreement and a lack of support with respect to speech-related motor abilities underlying prosodic impairments in ASD. It is hypothesized that a failure to acquire spoken language in some individuals with ASD, despite having adequate cognitive abilities and communicative intent, could be due to praxis deficits in speech, though motivation to communicate and joint attention abilities must also be considered as factors that could impact an individuals' ability to imitate speech sounds on request (Chenausky et al., 2019). Nonetheless, impairments at the level of speech-motor control would affect the prosodic and vocal qualities of speech in this population. The speech characteristics in individuals with ASD that have been documented in published literature may be suggestive of an underlying motor speech impairment due to motor speech programming and planning impairments and/or motor speech execution impairments (Adams, 1998; Seal & Bonvillian, 1997; Velleman et al., 2009).

The following characteristics of speech and voice in ASD appear to be consistent with a deficit in speech-motor control abilities: residual articulation distortion errors, uncodable utterances, and inappropriate utterances in terms of phrasing, stress, and resonance (Shriberg et al., 2001); increased loudness (Shriberg et al., 2001); deficits in focal oral motor skills and sequencing oromotor gestures (Velleman et al., 2009); low maximum phonation times for [a] and [f] prolongation; less varied syllable durations (Velleman et al., 2009); and “slurred” or “imprecise” sequencing of speech sounds (Gargan & Andrianopoulos, 2018, in progress). In addition, Parish-Morris et al. (2018) reported that adult speakers with ASD exhibited less diverse mouth movements during a 3-minute spontaneous conversation in comparison to TD peers. The authors hypothesize that subtle oral-motor impairments and/or reduced phonological diversity could contribute to restricted mouth movements.

2.2.6 Neurological Differences

There are significant differences in the neural underpinnings of prosodic comprehension and production in individuals with ASD compared to comparison groups. Although there is no consensus across studies on the differences in brain activation in ASD during the perception and production of prosody, there appear to be four hypotheses to explain the differences: (1) hypoactivation in some brain regions (Hesling et al., 2010; Watanabe et al., 2012; Loveland et al., 2010; and Korpilahiti et al., 2007); (2) different or increased attentional demands (Eigsti et al., 2012; Gebauer et al., 2014); (3) increased processing effort (Wang et al., 2006; Wang et al., 2007; Colich et al., 2012); and (4) enhanced pitch perception with concomitant difficulty in understanding the meaning of the contours (Minagawa-Kawai et al., 2009). The differences across studies are likely due

to differences in the samples, experimental tasks, and the type of prosody examined, among other methodological qualities of the studies. Additionally, it should be noted that the majority of studies included individuals with high-functioning ASD and only one study included lower-functioning individuals with autism. Thus, the generalizability of findings is limited. Nonetheless, despite differences across studies, there does appear to be some agreement that individuals with ASD have differences in brain activation during the perception and production of prosody, and these differences may underlie the communication, social-emotional, and language processing impairments in this population (Hesling et al., 2010; Gebauer et al., 2014; Eigsti et al., 2012). Continued research in the area of receptive and expressive prosodic abilities in ASD (considering both behavioral performance and neural underpinnings) is needed to better understand the neural and behavioral differences.

In summary, there is support in the literature that the factors hypothesized in the theories mentioned previously may contribute to the atypical speech, voice, and prosody differences observed among individuals with autism. Due to the heterogeneous nature of this disorder and differences in methodology across studies, it is unlikely that there is one theory to account for the range of speech atypicalities in this population. It is crucial that researchers and clinicians understand all possible contributing factors causing differences in speech, voice, and prosody. This will ensure that interventions contain appropriate content that is tailored to the individual.

Next, theories to explain atypical sensory processing will be discussed.

2.3 Theoretical Links between Autism and Atypical Sensory Processing

There are three prominent theories to explain perceptual processing of social and nonsocial information among individuals with ASD: (1) the Weak Central Coherence Theory (WCC); (2) the Enhanced Perceptual Functioning theory (EPF); and (3) the Neural Complexity Hypothesis. Relevant to this review are the WCC and EPF theories, which offer explanations regarding why some individuals with autism perform more accurately on tasks that require local processing as compared to their typically developing peers (O'Connor, 2012). More recently, Mayer et al. (2016) proposed a fourth theoretical framework addressing the allocation of attentional resources while interpreting auditory information. Findings from the literature will be discussed within the context of these cognitive theories.

2.3.1 Weak Central Coherence

The weak central coherence (WCC) theory, first proposed by Frith in 1989, suggests that individuals with autism have a detail-focused processing style with reduced global processing abilities. Thus, some individuals with autism attend to small pieces of information rather than large, globally coherent patterns of information. This enhanced perception of detail and reduced attention to the “big picture” results in an impaired ability to use context to interpret meaning (Jolliffe & Baron-Cohen, 1999). This hypothesis has been supported in studies that have examined cognitive flexibility and executive dysfunction in autism. In 2001, Rinehart, Bradshaw, Moss, Brereton, and Tonge demonstrated that individuals with autism were slower than control groups with respect to shifting from processing information at a local level (i.e., a detail, such as individual digits) to a more global level (i.e., a large number comprised of smaller digits).

This finding is consistent with the argument that individuals with autism have difficulty inhibiting the processing of local information. Additionally, it demonstrates that they have difficulty shifting their attention away from the details to process information on a more global level (Plaisted et al., 1999; cited in Rinehart et al., 2001). The WCC theory states that individuals with ASD have enhanced local-level processing abilities and weakened global processing.

2.3.2 Enhanced Perceptual Functioning

Frith's original coherence account has been modified in several ways, however, to emphasize superior local processing alongside intact global processing (Happé & Frith, 2006). The Enhanced Perceptual Functioning framework, an alternative account to the WCC theory, supports locally oriented and enhanced perceptual functioning in individuals with autism. This hypothesis is similar to the WCC theory in that the local bias is retained. The local bias is not, however, viewed as a result of a deficit in global processing (Mottron, Dawson, Soulières, Hubert, & Burack, 2006). This theory proposes that locally oriented processing is over-specialized, but higher-order functions are unimpaired.

2.3.3 Allocation of Attention

The model proposed by Mayer et al. (2016) postulates that there may be an initial difference in the allocation of attentional resources among individuals with ASD compared to control groups. Individuals with ASD are hypothesized to allocate attentional resources to pitch rather than higher-order language processing.

In summary, there are several prominent theories that may explain differences in speech, voice, prosody, and auditory processing among individuals with ASD. A review of the literature will be discussed in the context of these theories in the sections that follow.

2.4 Speech Differences in Autism Spectrum Disorder

Unusual vocalizations are one of the earliest behavioral characteristics of autism to appear (Schoen et al., 2011). For instance, studies have demonstrated that toddlers (ages 18-36 months) with autism produce significantly more atypical non-speech vocalizations as compared to age-matched and language-matched controls (Schoen et al., 2011). Additionally, children (mean age 44 months) with autism who are preverbal (i.e., less than 5 words) display impaired vocal quality during the production of canonical syllables (Sheinkopf, Mundy, Oller, & Steffens, 2000). Moreover, toddlers age 15 months who were later diagnosed with ASD produced atypical patterns of distress, such as cries with a higher fundamental frequency (F0) and shorter duration, as compared to infants who were at low risk for ASD (Esposito, Rostagno, Venuti, Haltigan, Messinger, 2014).

There is a paucity of research to account for why some individuals with autism develop minimal expressive spoken language, despite having access to intervention. Due to the heterogeneous nature of this disorder, it is unlikely that one single underlying mechanism is preventing individuals with ASD from developing spoken language (Tager-Flusberg & Kasari, 2013). Factors that have been identified to play a role in the development of spoken language include the following: (1) oral motor skills during the first year of life (Gernsbacher et al., 2008); (2) social motivation (Tager-Flusberg &

Kasari, 2013); and (3) the social feedback loop (Warlaumont et al., 2014). The following predictors of the development of expressive language, albeit not limited to spoken language, have also been outlined in the literature: (1) the ability to imitate sounds; and (2) nonverbal cognitive abilities (Luyster et al., 2008; cited in Tager-Flusberg & Kasari, 2013).

Several studies have examined oral motor abilities among individuals with ASD and the relation between oral motor abilities and speech production. Page and Boucher (1998) sought to find if oral motor and manual motor impairments co-exist in some children with autism. Their study was designed to assess oral motor, manual, and gross motor skills in 33 school-aged children with autism. Oral motor and manual motor impairments were prevalent among participants, with 70% of the sample demonstrating impaired tongue movements, 55% demonstrating manual motor impairments, and 16-17% demonstrating gross motor impairments. These findings are consistent with other published evidence, which suggests that many children with autism have oral motor impairments (Page & Boucher, 1998, Rapin et al., 1996 as cited in Page & Boucher, 1998; Seal and Bonvillion, 1997). It has been hypothesized that oral motor and manual motor impairments are due to dyspraxic motor impairments or neuromuscular deficits such as hypotonia (Page & Boucher 1998, Rapin et al., 1996 as cited in Page & Boucher, 1998). Findings from this study suggest that the observed oromotor impairments among these individuals with ASDs were dyspraxic in nature. However, it is important to note that motor impairments may also have been due to hypotonia. A limitation to this study was that the formal tests used required imitation. Thus, impaired performance on these

tasks may have been due to impaired imitation skills rather than dyspraxia (Page & Boucher, 1998).

Similarly, findings from a study conducted in 2002 support the presence of verbal dyspraxia in some children with autism. Gernsbacher, Hill Goldsmith, O'Reilly, Sauer, DeRuyter & Blanc (2002) sought to find if difficulties in oral- and manual- motor behaviors distinguish children with ASD from TD children. Oral-motor and manual-motor markers significantly differentiated the children with ASD from the TD children in this study. Dyspraxic oral-motor and manual-motor markers were present in 15% of children in this subgroup of participants with ASD. Interestingly, the majority of the children with dyspraxic features were nonverbal. This suggests that children with ASD with concomitant developmental verbal dyspraxia are likely to have limited speech output (Gernsbacher et al., 2002).

Adams (1998) used perceptual measures to examine the oral-motor and motor speech characteristics of children with ASD as compared to typically developing (TD) peers. The participants included four children with autism and four TD children, matched for age, gender, and ethnicity. The children with autism demonstrated poorer performance on all tasks, which included oral movement, simple phonemic/syllabic productions, and complex phonemic/syllabic productions. The children with ASD had greater difficulty with speech and nonspeech oral-motor movements as compared to TD peers (Adams, 1998).

2.5 Prosody in ASD

Prosodic aspects of speech encompass several acoustic features that correlate with subjective impressions of an individual's vocal productions. The acoustic and perceptual

features of prosody include frequency/pitch, intensity/loudness, duration, intonation, stress, and rhythm. The ability to comprehend and produce prosody plays a pivotal role in social and language development (Chiew & Kjelgaard, 2017). More specifically, prosodic skills are critical for the understanding and production of pragmatic, affective, and grammatical information (Peppe, McCann, Gibbon, O'Hare, & Rutherford, 2006; DePape et. al, 2012). Some individuals, such as those with ASD, may have a prosodic disorder that negatively impacts the development or mastery of prosodic skills (Velleman, 2015). Prosodic disorders can occur at the form or function level, caused by differences in the acoustic features of the voice or the use of prosody to convey pragmatic, affective, and grammatical meaning.

Although atypical prosody is not a universal characteristic of ASD, research supports that 33% (Kargas et al., 2016) to 60% (Nadig & Shaw, 2015) of participants on the autism spectrum exhibit differences in their prosodic abilities at the form or function level (e.g., Chiew & Kjelgaard, 2017; Diehl & Paul, 2013; Kargas et al., 2016; DePape et. al., 2012; Peppe, McCann, Gibbon, O'Hare, Rutherford, 2007; Paul et al., 2005). However, empirical research findings are inconclusive in describing the characteristics of atypical prosody. There is a lack of consensus regarding the perceptual and acoustic features that contribute to atypical production of prosody, as well as which domain of prosody is most affected in some speakers on the autism spectrum.

Individuals with ASD have been subjectively described in the literature as sounding “monotonic”, “robotic”, “odd”, “exaggerated”, “socially awkward”, “labored”, or “atypical” (Kanner, 1971; Filipe et al., 2014; Grossman et al., 2013; Nadig & Shaw, 2012; and Bonneh et al., 2011; Grossman et al., 2010). Moreover, acoustic findings

related to expressive prosodic abilities range from longer duration of utterances (Bonneh et al., 2011; Grossman et al., 2010; Filipe et al., 2014) to increased pitch variability (Diehl et al., 2009), and inappropriate phrasing, stress and resonance (Shriberg et al., 2001).

It has been purported that individuals with autism have impaired pragmatic and affective prosody while grammatical prosodic abilities are relatively spared (Shriberg et al., 2001; McCann & Peppe, 2003). In contrast, other research has demonstrated that individuals with autism have impaired grammatical prosody, though maybe to a lesser extent than pragmatic and affective impairments (e.g., Paul et al., 2005). In the sections that follow, a brief definition of each domain of prosody is discussed.

Pragmatic prosody is used to convey information beyond what is provided by the syntax of the sentence. For instance, the pragmatic use of stress can be used to make a specific word within a sentence the focus of attention (e.g., ‘I want the BLACK socks’ to indicate that the speaker the socks that are black rather than a previously mentioned color) (Shriberg et al., 2001; McCann & Peppe, 2003). Affective prosody refers to changes in register to express emotion and/or to meet the communicative expectations of a given social interaction (e.g., talking to peers versus a baby or sounding happy versus sad) (Shriberg et al., 2001). Lastly, grammatical prosody is used to signal syntactic information within sentences. For instance, stress can be used to indicate whether a word is a noun (e.g., PREsent) or a verb (e.g., preSENT) and a compound noun (e.g., HOTdog) or a noun phrase (e.g., hot DOG). Pitch contours can be used to indicate if an utterance is a question (rising pitch) or a statement (falling pitch) (Shriberg et al., 2001). The presence or absence of a pause can indicate whether an utterance contains three nouns

(e.g., chocolate, cake, and fruit) or two nouns (e.g., chocolate-cake and fruit) (McCann & Peppe, 2003).

2.5.1 Pragmatic and Affective Prosody in ASD

A relatively large body of research focuses on pragmatic and affective prosody among individuals with autism (Shriberg et al., 2001; McCann & Peppe, 2003; Kargas et al., 2016). Although findings are not entirely consistent across studies, some individuals with autism have difficulty with the following areas of receptive and expressive pragmatic/affective prosody: 1) labeling emotions; 2) contrastive stress; and 3) understanding or expressing like vs. dislike.

Findings from previous research have demonstrated that individuals with ASD who have difficulty with the comprehension of affective prosody have impairments in labeling emotions or matching emotions to facial expressions (Boucher, Lewis, & Collis, 1998, cited in Eigsti et al., 2012; Schultz, 2005, cited in Eigsti et al., 2012).

Additionally, a series of studies have demonstrated that individuals with HFA or ASD perform significantly worse on tasks requiring the production of contrastive stress (Peppe et al., 2007; Peppe et al., 2011). It has been hypothesized that language abilities play a role in the production of contrastive stress. For instance, a group with ASD with high language functioning demonstrated higher pitch ranges, but did not mark contrastive stress, whereas those with low language skills had a smaller pitch range but produced contrastive stress (DePape et al., 2012). Furthermore, some individuals with HFA produce contrastive stress comparably to their TD peers, but use a smaller pitch range and sound less dynamic when producing contrastive stress. This results in smaller prominence differences when producing words that are the focus of the sentence versus

words that are not the focus (Kaland et al., 2013). Studies that failed to demonstrate behavioral differences on expressive contrastive stress demonstrated differences between groups with respect to average fundamental frequency (f0), f0 range, and standard deviation (SD) of f0 (Diehl & Paul, 2013). Thus, although their productions may be perceived as categorically accurate, acoustic differences exist. Findings related to the perception of contrastive stress are contradictory, in that two studies did not find significant differences between individuals with HFA and ASD compared to TD controls (Peppe et al., 2007; Diehl & Paul, 2013), while one study found that groups with ASD and LD performed significantly worse than a TD group (Diehl & Paul, 2013).

Lastly, several studies demonstrate that individuals with ASD have impaired affective prosody on tasks where they are required to comprehend or produce one-word statements expressing like versus dislike (e.g., Peppe et al., 2011; Peppe et al., 2007; Diehl & Paul, 2013). In contrast, Diehl & Paul (2013) did not find significant differences on the expressive task, but acoustic findings revealed that the individuals with ASD produced significantly longer utterances than the other groups when they were expressing dislike for the food item.

2.5.2 Grammatical Prosody in ASD

Research related to the expressive and receptive abilities of grammatical prosody among individuals with ASD reveals contradictory findings across studies. Shriberg et al. (2001) used the PVSP to investigate the prosody-voice profiles of individuals with HFA and Asperger Syndrome (AS) (n=30; age range, 10-49 years) compared to typically developing controls (n=53; age range, 10-30 years). Conversational speech samples were collected through standardized ADOS interviews (Lord et al., 1996) and used for narrow

phonetic transcription and prosody-voice coding. These researchers found that the HFA and AS groups had residual articulation distortion errors, uncodable utterances, and inappropriate utterances in terms of prosody and voice (i.e., phrasing, stress, and resonance). With respect to prosody, the phrasing errors included increased repetitions and revisions or reduced rate of speech. These errors were consistent with a deficit in the grammatical use of prosody. Conversely, the inappropriate stress among those with HFA and AS reflected a deficit in the pragmatic or emphatic domains, while grammatical use of stress was intact. Voice and resonance differences were also found, in that the individuals with HFA and AS had significantly more utterances rated as too nasal.

Shriberg et al. (2001) hypothesized that deficits in phrasing may reflect formulation difficulties, lexical complexity, or perceptual-motor deficits. The findings from their study indicate that higher levels of grammatical complexity were associated with increased errors. Deficits in the use of stress, as well as vocal and resonance behaviors, may also reflect impaired social cognition or perceptual-motor deficits. It should be noted, however, that deficits in perceptual-motor skills would be reflected in both grammatical and pragmatic contexts. The use of stress among participants with ASD in this study was significantly different in the pragmatic context only, thus appearing to be most consistent with a deficit in social cognition or ToM. Lastly, the authors argue that the nasal speech among the individuals with HFA and AS in this study reflected a deficit in social-adaptive behavior. Limitations to Shriberg et al. (2001) include the use of conversational sampling and a lack of structured tasks, as well as the wide age range of participants.

Paul et al. (2005) investigated the perception and production of grammatical and pragmatic/affective prosody among individuals with high functioning autism (n=27; mean age, 16.8 years) compared to typically developing controls (n=13; mean age 16.7 years). The following aspects of prosody were assessed for both grammatical and pragmatic/affective functions: stress, intonation, and phrasing. On the receptive tasks, the participants listened to a sentence and indicated a response by selecting the written sentence that represented the correct way of saying the word they heard, or by pointing to a picture. The expressive task required the participants to read written stimuli aloud. Their written responses were recorded on a score sheet and spoken responses were audio recorded for analysis by a rater who was blind to participant diagnosis.

The findings support those reported by Shriberg et al. (2001), which suggest that individuals with autism have difficulty with the production and perception of stress. While Shriberg et al. (2001) only found significant differences in the affective/pragmatic use of stress, Paul et al. (2005) observed impairments in the production and perception of stress for both grammatical and affective/pragmatic functions. Thus, individuals with ASD had difficulty on tasks assessing all aspects of stress.

It is important to note that the results from Paul et al. (2005) revealed ceiling effects on some tasks. All participants scored with greater than 90% accuracy on five out of the twelve tasks, which limited interpretation of findings regarding the perception and production of intonation and phrasing for all prosodic functions. Moreover, a lack of significant differences on tasks with ceiling effects could be explained by the following: (1) some skills were well-established in both groups; (2) individuals with ASD developed a strategy to allow them to perform accurately; and/or (3) some tasks were difficult for

both groups, such as the grammatical phrasing task. It should be noted that while there were differences in both grammatical and pragmatic/affective stress, the differences between groups were larger on the pragmatic/affective tasks. In addition to ceiling effects, other limitations of this study include the use of perceptual judgments without acoustic analysis, and the use of natural, uncontrolled speech.

Paul, Bianchi, Augustyn, Klin, and Volkmar (2008) used perceptual and acoustic analysis to examine the ability to reproduce stress in a nonsense syllable imitation task among individuals with autism (n=46; age range, 7;4 – 28;7) compared to a group of TD peers (n=20; age range, 7;11 – 27;5). The group with ASD consisted of children with HFA (n=22), AS (n=19), and PDD-NOS (n=5) who were required to have a verbal IQ of 70 or greater to participate in this study. The stimuli consisted of pre-recordings of the syllable [ma] with varied rhythm and intonation. The participant listened to strings of 2-6 syllables and was expected to imitate the syllables when the tape was stopped. A rater who was blind to both participant diagnosis and stimuli rated the utterance and assigned it to a stressed or unstressed category. Point-to-point agreement between perceptual judgments by two blind raters on 15% of the samples revealed 96% agreement. The participants' vocal productions were also analyzed acoustically. Acoustic analysis included measurement of pitch range and duration of each syllable produced. Perceptual ratings revealed that the individuals with autism were significantly less likely to be scored as correct on their stressed and unstressed productions. There were no significant differences in pitch range (i.e., stressed syllables had greater pitch ranges among all groups). Nonetheless, although both groups showed longer durations for stressed versus unstressed syllables, the group with ASD showed significantly less difference in duration

between syllable categories. These acoustic differences may explain why individuals with ASD were less likely to be scored as correct. Consistent with Shriberg et al. (2001), the authors suggest that the individuals with ASD may have difficulty in perceptual and/or motor skills needed for stress production, or impaired social abilities, such as decreased motivation to complete the task. The authors argue that the differences in performance among the group with ASD indicate that a deficit in the ability to manage the conversational function of stress cannot be the only explanation for the difficulties, as the task did not require linguistic or pragmatic processing.

Grossman et al. (2010) investigated the perception and production of lexical stress among individuals with HFA (n=16; age range, 7;6-17;0) compared to TD peers (n=15; age range, 7;6-18;0) who were matched for age, nonverbal IQ, and receptive vocabulary ability. The stimuli in the receptive task consisted of picture stimuli for two-syllable words with stress patterns in American English (e.g., a picture displaying the target word “HOTdog” with a picture of a hotdog; and a picture of a hot dog displaying “hot DOG”). The participants were expected to choose the picture that matched the meaning of the target word. There were no significant differences between groups on the receptive task, suggesting that the individuals with HFA performed comparably to their peers. Additionally, both groups were more accurate at discriminating words with first-syllable stress, which was expected because that is the primary stress pattern of American English.

The authors hypothesized that the relatively preserved language skills among the group with HFA contributed to the ability to successfully disambiguate compound nouns from noun phrases. An alternative explanation for the lack of group differences is that the

TD group's performance was poorer than expected. The low performance among the TD group could be explained by the wide age-range of participants. Additionally, some participants were below the age of competence for this task (i.e., 10 years).

The expressive task required participants to listen to a sentence and fill in the missing word at the end of a sentence while looking at a corresponding picture. There were no significant group differences in the participants' ability to disambiguate word pairs through the production of first-stress and last-syllable stress patterns. Additionally, both groups produced significantly shorter first-syllable stress items, as expected in American English. Although behavioral findings demonstrate similarities in the ability to correctly produce lexical stress to disambiguate word pairs, acoustic analysis of pitch, intensity, and duration revealed that the group with HFA produced significantly longer productions and exaggerated pauses between syllables. The group with HFA was categorically accurate in their productions, but quantifiably different from the control group. This study did not include subjective ratings, but the authors informally observed the productions in the HFA group to sound more labored and awkward. These findings are consistent with previous studies (e.g., Paul et al., 2005) that demonstrated a deficit in the production of stress among those with ASD, despite relatively intact perceptual abilities (Grossman et al., 2010).

Filipe et al. (2014) used acoustic and perceptual measures to assess how individuals with Asperger Syndrome (AS) use prosody to differentiate statements versus questions. Participants included children with AS (n=12; age range, 8-9 years) and TD peers (n=17) matched for age, gender and nonverbal intelligence. The ability to receptively and expressively distinguish questions versus statements was assessed using

the Turn-End subtest from the Profiling Elements of Prosody in Speech-Communication, consisting of one-word utterances with rising or falling intonation. There were no significant differences between groups on the perception of questions versus statements. This findings contradicts Diehl and Paul (2013), who demonstrated that a group with ASD performed significantly worse than a TD control group. With respect to the production of questions versus statements, there were no significant differences as determined by perceptual judgments. However, the participants with AS were described as sounding significantly more odd than the TD group. Furthermore, acoustic measurements revealed that the AS participants had longer durations of utterances, higher pitch ranges, significantly higher mean pitches, and significant differences in maximum pitch. This finding is consistent with previous studies demonstrating that individuals with autism may be categorically accurate in their perception or production of prosody, but there are acoustic and perceptual differences (e.g., Grossman et al., 2010).

Kargas et al. (2016) conducted a study to investigate syllable stress perception sensitivity among individuals with high-functioning ASD (n=21; mean age 30 years, 3 months) compared to a TD group (n=21, mean age 29 years, 5 months). The relationship among perception of syllable stress, speech production, and communicative ability in those with ASD was explored. The stimuli consisted of 10 four-syllable words that had first syllable stress (e.g., AUditory) and 10 four-syllable words that had second syllable stress (e.g., caPAcity). The participants made same-different judgments about the location of the syllable stress when presented with word pairs. The stimuli were similar to the stimuli used in Paul et al. (2008) in that the meaning of the words was not important for making same-different decisions. The findings revealed that the group with ASD was

significantly less sensitive in the perception of syllable stress as compared to controls. Additionally, poorer perceptual sensitivity of syllable stress was associated with atypical speech production, but not communicative ability. It is interesting to note that the performance within the ASD group was variable, and a subgroup of individuals with ASD (33%) demonstrated poor perception of syllable stress and atypical quality of speech production. This finding is in agreement with the argument that not all individuals with ASD have atypical receptive and expressive prosodic abilities. One limitation to this study is that a composite score for vocal atypicalities was used and the individual features of speech were not examined. The authors also suggest that further research is needed to understand how differences in low-level auditory discrimination abilities play a role in receptive prosody among individuals with ASD.

To summarize, grammatical prosodic abilities have been investigated using acoustic (Paul et al., 2008; Grossman et al., 2010; and Filipe et al., 2014) and perceptual (Shriberg et al., 2001; Paul et al., 2005; Paul et al., 2008; and Filipe et al., 2014) measures. To date, the findings related to grammatical prosody reveal that some individuals with ASD demonstrate: (1) poor perception of syllable stress, which is associated with atypical speech production; (2) impaired production of stress at the word level in sentence completion tasks (Paul et al., 2005) and during nonsense syllable tasks (Paul et al., 2008); (3) phrasing errors during connected speech, such as increased repetitions and revisions with increased lexical complexity (Shriberg et al., 2001); (4) longer durations when producing two-syllable words due to exaggerated pauses between syllables (Grossman et al., 2010); and (5) longer durations when producing one-word questions and statements (Filipe et al., 2014). The participants with ASD were also

described as sounding more labored (Grossman et al., 2010) and odd/awkward in their productions (Grossman et al., 2010; Filipe et al., 2014).

2.5.3 Atypical Prosody in the Context of Theories

In line with the theoretical contributions mentioned previously, six primary hypotheses have been outlined in the literature to account for atypical prosodic abilities among individuals with autism. Some authors hypothesize that speech differences in ASD could be due to: (1) a lack of one's ability to 'tune up' their speech behaviors (Diehl and Paul, 2013; Diehl & Paul, 2012; Shriberg et al., 2011); (2) language deficits (Peppe et al., 2007; Peppe et al., 2011); (3) social reciprocal interaction impairment (Nakai et al., 2014); (4) the nature of the research task (Diehl and Paul, 2013; Peppe et al., 2007); (5) impaired auditory memory (Peppe et al., 2006; Peppe et al., 2007); and (6) motor deficits (Peppe et al., 2007; Diehl & Paul, 2012; Velleman, Andrianopoulos, Boucher, Perkins, Currier, Marsello, Lippe, & Van Emmerik, 2009; Andrianopoulos et al., 2015).

2.6 Auditory Processing in ASD

Published literature supports that individuals with ASD have differences in their auditory processing abilities on behavioral and neural levels. With respect to findings on the behavioral level, the differences in auditory processing abilities outlined are as follows: (1) enhanced pitch processing; (2) increased sensitivity to loud sounds; (3) atypical orientation to auditory stimuli; (3) atypical processing of affective and grammatical prosody; and (4) impaired auditory processing in background noise (O'Connor, 2012).

The following sections will address behavioral findings supporting atypical pitch perception among some individuals with ASD in the context of the EPF, WCC, and attentional theories.

2.6.1 Enhanced Pitch Processing

Empirical investigations support that children with autism exhibit the following:

(1) enhanced pitch processing for simple and complex auditory information (Bonnell, Mottron, Peretz, Trudel, Gallun, and Bonnell, 2003; Eigsti & Fein, 2013; Heaton et al., 2008c; Heaton, Hudry, Ludlow, & Hill 2008b; Stanutz, 2014; Jarvinen-Pasley & Heaton, 2007; Jarvinen-Pasley et al., 2008; Mayer et al., 2016; O’Riordan & Passetti, 2006; Jones et al., 2009; and Chodury, Sharda, Foster, Germain, Tryfon, & Doyle-Thomas et al., 2017); (2) enhanced pitch memory for musical tones (Heaton et al., 2008c; Heaton et al., 2003; and Stanutz et al., 2014); and (3) enhanced pitch labeling of musical tones and speech (Heaton et al., 2003; and Heaton et al., 2008a).

2.6.1.1 Speech and Non-speech Stimuli

Several studies have demonstrated that individuals with autism exhibit enhanced pitch discrimination of speech and non-speech stimuli. Heaton et al. (2008b) conducted a study to investigate pitch discrimination abilities, comparing words, non-words, and non-speech pitch contours that were systematically varied in pitch. The participants included 14 children with ASD (age range: 6;11 to 14;9 years) and 14 controls matched for age and receptive vocabulary. The children listened to three blocks of stimulus pairs. Each block was separated by type and counterbalanced across the participants within each

group. The children pressed a button to indicate whether the two sounds were the same or different.

The findings from the Heaton et al. (2008b) study revealed the following: (1) the children with autism performed more accurately than the control group on pitch discrimination tasks; (2) both groups demonstrated lower performance on stimuli containing linguistic information; and (3) there was not a negative relationship between pitch processing and language abilities in ASD (i.e., enhanced discrimination abilities were exhibited by individuals with autism who did not have significant language impairment ($n=2$), as well as those who had very low receptive language scores ($n=2$)). One limitation to these findings is that children with low chronological age and verbal mental age were excluded due to task difficulty. The authors suggest that future research could develop a task that is appropriate for younger children with autism who have lower verbal mental ages.

Mayer et al. (2016) conducted two experiments to: (1) investigate the developmental trajectory of pitch discrimination abilities across child, adolescent, and adult individuals with ASD compared to control groups without ASD; (2) determine if pitch discrimination abilities are correlated with receptive vocabulary; and (3) identify the cognitive, clinical, and behavioral correlates of enhanced pitch discrimination abilities in adults. Data from the child cohorts was obtained from Heaton et al. (2008b) and consisted of 14 children with ASD (age range: 6;11 to 14;9 years) and 14 control children matched for age and receptive vocabulary. The adolescent cohort data were obtained from a previous study conducted by Mayer (2009) and consisted of 14 adolescents with ASD (age range: 9;8 to 16;5 years) and 14 control children matched for age and

nonverbal IQ. The adult cohort in the current study consisted of 19 adults with ASD (age range: 23;9 to 59;8 years) and 19 typically developing adults matched for age and receptive vocabulary.

Experiment 1 used stimuli that were developed and used in a previous study conducted by Heaton et al. (2008b) and consisted of speech (i.e., monosyllabic words) and non-speech stimuli (i.e., pitch contours analogous to the words in the speech condition). The participants indicated if the stimuli were the same or different.

The adult cohorts with and without autism in this study did not perform significantly differently on the pitch discrimination experimental task. These findings contradict previous reports of enhanced pitch discrimination abilities among some adults with ASD (Heaton et al., 2008a). The findings can be explained by the results from the analysis of developmental trajectory of pitch discrimination abilities across groups as follows: (1) as age increases, pitch discrimination abilities increase in the TD group and this is associated with language ability (linear trend); and (2) as age increases, pitch discrimination abilities remain stable among those with ASD and are not aligned with language skills (no trend). The failure to find a significant difference between the adult cohorts may be due to an increase in pitch discrimination abilities from childhood and adolescence into adulthood for the TD group (i.e., closing the gap in performance over time). Conversely, the performance among the individuals with autism is enhanced at a young age and remains stable. The findings from experiment 1 were consistent with Heaton et al. (2008b) in that the participants with and without ASD had more difficulty discriminating stimuli that contained semantic content. This illustrates that both groups, in both studies, were focused on the semantic content of the auditory information.

Experiment 2 sought to identify cognitive, behavioral, and clinical factors that are correlated with pitch discrimination abilities. A positive relationship was revealed between receptive vocabulary and pitch discrimination abilities in the control group. Contrary to other studies that argue that language deficits are linked to enhanced pitch perception abilities (Jarvinen-Pasley & Heaton 2007; Jarvinen-Pasley et al., 2008), this study did not find a negative correlation between receptive language and pitch discrimination abilities in the ASD group. This indicates that enhanced pitch discrimination abilities were not associated with impaired language development in the group of individuals with ASD, but were positively associated with language skills in the typically developing group.

Although the findings from Mayer et al. (2016) can be interpreted in the context of the EPF model, the authors propose a new theoretical framework related to attention. The results from this study indicate that this characteristic remains stable over time among individuals with autism. In contrast, individuals who do not have autism initially allocate their attentional resources to interpret meaning from auditory information rather than focusing on psychoacoustic cues. As they develop, they have increased attentional resources, which can allow for simultaneous processing of information, resulting in an increase in pitch discrimination abilities as adults.

It is important to note that Mayer et al. (2016) was the first to examine the developmental trajectory for pitch discrimination among individuals with autism. One limitation of the Mayer et al. (2016) findings is that cross-sectional data from three separate studies was used. The matching criteria were consistent within each cohort, but not across cohorts. For example, the child and adult cohorts were matched on age and

receptive language, whereas the adolescent cohorts were matched on age and nonverbal intelligence. Future research could address the same research questions in a longitudinal study with cohorts matched closely for age, diagnosis, and verbal and nonverbal intelligence (Mayer et al., 2016).

Jarvinen-Pasley & Heaton (2007) tested the hypothesis that auditory processing is less domain-specific among individuals with autism or Asperger syndrome. This study included a group of participants with Autism or Asperger syndrome (n=19; mean age 11;5) and a group of typically developing controls (TD) (n=19; mean age 11;6) who were matched for chronological age as well as nonverbal and verbal intelligence. The procedure required the participants to listen to 72 different pitch sequence stimulus pairs that either shared four-pitch sequences or differed in two pitches. Participants were expected to determine if the tone sequences were the same or different. Speech and nonspeech stimuli were used to assess auditory processing across speech and music domains. The experiment consisted of three conditions that increased in complexity: music-music, speech-speech, and speech-music stimulus pairs.

The results of this study revealed that there were no significant differences between stimulus conditions within the autism group. There were, however, significant differences between all conditions for the control group. Between-group comparisons revealed: (1) no significant differences between groups in the music-music condition; (2) significant differences between groups in the speech-speech condition; and (3) significant differences between groups in the speech-music condition. The children with autism detected pitch differences in both speech and musical stimuli. The control group demonstrated a comparable ability to differentiate pitch in the music-music condition.

This demonstrates that the perceptual processing abilities in general were comparable between groups. Contrary to the group with autism, however, the control group's performance deteriorated when the condition required them to discriminate pitch between stimuli containing linguistic content (i.e., speech-speech or speech-music conditions (Jarvinen-Pasley et al., 2007)).

The finding that only the control group's performance was poorer in the conditions that contained speech (but not that of the ASD group) contradicts findings from previous studies. For example, Mayer et al. (2016) and Heaton et al. (2008b) demonstrated increased performance among individuals with and without autism on stimuli containing non-speech versus speech stimuli. It is important to note that the contradictory findings may be due to the specific characteristics of the participants and differences in the experimental tasks. For instance, Mayer et al. (2016) and Heaton et al. (2008b) separated the tasks in their study by stimulus type, since pilot testing demonstrated that the participants experienced confusion when the stimulus types were presented together (i.e., words, non-words, and pitch contours presented randomly within the blocks). Jarvinen-Pasley et al. (2007) created stimuli that examined skills within and across auditory conditions (e.g., speech-speech versus speech-music). Furthermore, the stimuli in the current experiment consisted of 4-syllable words and 4-tone musical stimuli, whereas Mayer et al.'s (2016) and Heaton et al.'s (2008b) stimuli included monosyllabic words and non-speech sounds. It is possible that the increased semantic complexity in Jarvinen-Pasley et al., (2007) contributed to the semantic capture exhibited by the TD group. Nonetheless, the control group's performance in this study is consistent with a domain-specificity account of auditory processing among individuals who are

typically developing (i.e., the mechanisms responsible for processing speech hindered their performance in the conditions that contained linguistic information). Conversely, the individuals with autism demonstrated domain-general processing (i.e., they exhibited similar pitch sensitivity across stimulus conditions) (Jarvinen-Pasley et al., 2007).

The authors state that the outcomes of this study regarding enhanced perceptual processing ability among individuals with autism, particularly in the speech conditions, are consistent with the WCC hypothesis. Similarly, the authors also state that their empirical findings can be interpreted within the context of the EPF theory because the individuals with autism were observed to have enhanced processing of low-level perceptual information in speech in that participants focused on acoustic-perceptual information versus linguistically relevant information. The authors conclude that enhanced processing of low-level acoustic-perceptual information may account for the deficits in prosodic and semantic processing among individuals with autism. As previously reported, approximately 25-30% of individuals who are diagnosed with autism fail to develop spoken language (Tager-Flusberg & Kasari, 2013). Of note with respect to the proposed attentional model is that the authors postulate that overly selective attention toward low-level perceptual information in speech may hinder the development of language and/or language processing in individuals with autism (Jarvinen-Pasley et al., 2007).

Jarvinen-Pasley et al. (2008) conducted two experiments to test processing biases towards perceptual (e.g., pitch contour in experiment 1; temporal patterns in experiment 2) and semantic information among individuals with autism (n=20) compared to a control group matched for age, gender, and intelligence (n=20). Although both experiments

included balanced groups of 20 participants, it is important to note that only 60% of the same participants from experiment 1 were included in experiment 2. The mean age for both experiments ranged from 12;0 years to 12;8 years of age.

Experiment 1 used speech and musical samples with simple pitch contours to examine processing biases toward pitch information. Experiment 2 used speech and musical samples ranging from monosyllabic to tri-syllabic words or perceptually analogous pitch contours to examine processing biases toward temporal information. Each experiment contained three conditions: (1) a linguistic perceptual condition; (2) a linguistic comprehension condition; and (3) a non-linguistic perceptual condition. The child participants were expected to: (1) point to the shape that matched the utterance they heard; (2) answer questions regarding the speech samples; and (3) match a melody to a visual display, respectively.

The results from experiment 1 suggest the following: (1) the children with autism exhibited superior perceptual processing on both speech and musical stimuli; (2) the control group demonstrated superior comprehension skills; (3) neither group showed evidence of semantic interference, as performance differences between linguistic perceptual and nonlinguistic perceptual stimuli were not statistically significant within each group. This finding is inconsistent with findings from Mayer et al. (2016) and Heaton et al. (2008b).

The results from experiment 2 demonstrated the following: (1) children with ASD performed significantly better in the linguistic perceptual condition compared to controls; (2) the control group showed superior comprehension; (3) children with ASD did not perform significantly differently on linguistic perceptual versus non-linguistic perceptual

stimuli whereas the controls performed significantly higher in the nonlinguistic perceptual condition. The participants with autism demonstrated enhanced perceptual performance in both experiments compared to controls in speech and musical stimuli. Contrary to experiment 1, the control group demonstrated significantly better performance when the stimuli did not include semantic information (i.e., they experienced semantic capture). This is consistent with Mayer et al. (2016) and Heaton et al. (2008b).

The group of individuals with ASD showed enhanced perceptual processing. Their performance on the sentence comprehension task, however, was inferior to that of controls, despite being matched for age, nonverbal intelligence, and receptive vocabulary. The results in this study are consistent with EPF and WCC theories in that the individuals with ASD processed local versus global information. Additionally, semantic interference was not evident when the stimuli required the participants to process pitch contour information. When the task involved detecting differences in temporal processing, however, the control participants were impacted by semantic interference. This indicates that some individuals with autism may exhibit enhanced perceptual processing of certain components of speech.

Similar to Jarvinen-Pasley and Heaton (2007), Jarvinen-Pasley et al. (2008) suggest that overly focused auditory processing of speech stimuli may inhibit the development of higher-level language skills, such as prosodic processing and sentence comprehension. These findings are contradictory with respect to a lack of semantic capture among individuals with autism or the control groups in the pitch condition: a

decrease in performance in the semantically meaningful stimuli in the temporal condition can be explained by variations in task demands within and across studies.

2.6.1.2 Pure Tone Stimuli

Bonnel et al. (2003) tested the hypothesis that individuals with autism perform better than control groups on tasks involving: (1) memorizing picture-pitch associations; and (2) detecting pitch changes in melodies. They also examined whether absolute pitch is due to abnormally high sensitivity to local processing of sounds in the autism group. This investigation included a group of individuals with High-Functioning Autism (HFA), mean age 17 years, and a group of typically developing (TD) peers matched for chronological age, full-scale IQ, and laterality. All participants included in the experiments and data analysis had normal hearing.

The authors conducted two experiments. In experiment 1, the participants were required to judge the pitch of pure tones in a “same-different” pitch discrimination task. Forty trials consisting of pairs of tones presented at the same intensity level with varied frequency were administered. Within each trial, the first tone was randomly presented at 500, 700, 1000, or 1500 Hz. For 20 of the trials (50%), the tones were the same pitch. For the other 50% of the trials, the second tone was 1% (hard), 2% (medium), or 3% (easy) higher in pitch. The ability to discriminate between the two sounds per trial required the participants to use “trace” memory, which only lasts a few seconds. The participants indicated whether the two consecutive sounds were the same or different by selecting a button on a response box. The response box consisted of six (6) buttons, which allowed the participant to also indicate how confident they were in their answer (e.g., for “same”, there was a red, orange, and yellow button representing high confidence, moderate

confidence, and low confidence, respectively). The individuals with HFA performed better than controls in all three conditions, although significant differences were only revealed in the 2% and 3% conditions.

The second experiment also consisted of 40 trials in total. In this experiment, the participants were instructed to judge if a tone was higher or lower than a previously presented tone. Participants were required to use context coding by referring to previous trials in order to make categorical judgments. The tones were presented at the same intensity level (58-60 dB SPL) with 1000 Hz pure tone as the low reference tone and a high tone that varied from 1%-3% higher. The participants were also required to indicate whether a tone was high or low using the same response method as experiment 1. The individuals with HFA demonstrated advanced categorization abilities as compared to the controls (Bonnell et al., 2003).

The findings from this study revealed that pitch processing was enhanced in the HFA group in a discrimination and categorization task. The group of individuals with autism performed the same on both tasks, whereas the control participants' performance deteriorated in the categorization tasks. These findings support that the individuals with autism were less sensitive to the task requirements. Additionally, the group with autism may have used the same type of memory on both tasks (i.e., trace memory), while the control participants may have switched to a context-coding type of memory during the categorization task. The results of this study can be interpreted within the context of the Enhanced Perceptual Functioning Model (Bonnell et al., 2003).

Eigsti and Fein (2013) conducted a study to investigate pitch discrimination in children with ASD (n=29), children with a history of ASD who were classified by the

authors as having optimal outcomes (OO; n=26), and children who are typically developing (TD; n=20). Furthermore, the authors examined how pitch discrimination is related to current symptomatology, early-language milestones, and current language skills. The participants ranged in age from 8-21 years and were matched on full-scale IQ chronological age. The participants completed diagnostic and standardized tasks (i.e., the ADOS, ADI-R, and SCQ) as well as a tone discrimination task used in a previous study by Bonnel et al. (2003).

The results of this investigation suggest that the ASD group demonstrated the best pitch discrimination. There were significant group differences between the TD group and ASD group (i.e., TD group had lower scores). The OO group did not perform significantly differently than the TD or ASD group. A correlation between early-language milestones and auditory perceptual abilities was revealed. The enhanced pitch discrimination among individuals with autism compared to typically developing controls in the same/different tasks involving pure tones is consistent with findings from Bonnel et al. (2003) (Eigsti & Fein, 2013).

This study demonstrated that difficulty with learning first words was associated with enhanced performance on the tone discrimination task. In agreement with previous studies, there was no association between pitch discrimination abilities and current language skills (Heaton et al., 2008b; Mayer et al., 2016). Lastly, a significant relationship was evidenced between symptom severity and pitch discrimination ability.

The authors postulate that a heightened sensitivity to acoustic discrimination negatively impacts word learning during the first two years of life. The results of this study are consistent with previous studies that have demonstrated a link between

enhanced auditory discrimination and delayed language development (Jones et al., 2009). The authors used tones at 500-1,500Hz to assess pitch discrimination, which were described as being distinct from those that are key to human speech. The authors state that attention to human speech (whether typical or atypical) would not explain the enhanced pitch discrimination abilities among individuals with autism in this study, yet this conclusion somewhat contradicts the new theoretical framework proposed by Mayer et al. (2016), which addresses how the allocation of attention may impact auditory processing (Eigsti & Fein, 2013).

O’Riordan and Passetti (2006) conducted an investigation to examine auditory and tactile discrimination abilities among individuals with autism (n=12) and typically developing children (n=12). The mean age of each group was 8;7 years and the two groups were not significantly different with respect to general IQ. Although three experiments were conducted in this study, the details from only experiment 1 (pitch discrimination) will be described, followed by the conclusion of all three experiments. The stimuli in experiment 1 consisted of two tones (tone A and tone B) presented in a three-tone sequence (ABA). Tone A was fixed in frequency, whereas tone B decreased in frequency in each pair until the two tones were identical in frequency. The participants listened to two alternating tones and were told that one tone would stay the same while the other got lower on each trial. The participants needed to determine whether the two tones eventually became identical.

The results of O’Riordan and Passetti’s (2006) study indicate that the children with autism decided that the two tones were identical significantly later in the sequence as compared to the control group. They performed better than the control group and

demonstrated enhanced pitch discrimination abilities. Experiments 2 and 3 examined tactile discrimination and revealed no significant differences between groups. The results of this study support that enhanced processing in the visual domain extends to the auditory, but not to the tactile, modality in individuals with autism (O’Riordan & Passetti, 2006).

In another study, Chowdhury et al. (2017) examined the relationship of low- and high-level auditory tasks (i.e., pitch processing) to current verbal and nonverbal abilities among children with ASD compared to TD controls. The aims of the study were as follows: (1) examine between-group differences in the area of low- and high-level auditory tasks and verbal and nonverbal cognitive abilities; (2) determine if there is a relationship between verbal abilities and performance on the low-level (pitch direction; PD) and high-level (global-local; GL) tasks; and (3) examine the relationship between nonverbal abilities on auditory PD and GL tasks. The participants consisted of children with ASD (n=17) and children who were TD (n=19) matched on age (mean age: 13 years). Participants’ verbal and nonverbal cognitive abilities were assessed using subtests from the Weschler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999, cited in Chowdhury et al., 2017).

The low level PD task required the participants to listen to pairs of tones that differed in pitch and decide if the second tone was higher or lower in pitch than the first one. The stimuli ranged in difficulty for both pitch and temporal parameters. In the high-level GL task, the participants heard three-tone triplet sequences combined to form a sequence of nine harmonic tones. The local level was the pattern within each three-tone triplet, whereas the global level was defined as the pattern created by the first tone of

each three-tone triplet. Two blocks of stimuli were completed within the GL task: (1) participants decided whether the pattern within a three-tone triplet was going up, down, or if it remained neutral (local); and (2) the participants directed their attention to the first tone of each triplet and ignored the local information (global).

With respect to aim 1, no differences in performance were found between groups on auditory tasks or cognitive measures. Regarding aim 2, there was no effect of verbal intelligence on low- and high-level auditory tasks. A significant relationship between nonverbal abilities and auditory perception in both groups was observed in that nonverbal abilities predicted performance on the low-level PD task and the local (but not global) pitch processing during the more complex GL task (aim 3).

Chowdhury et al.'s (2017) findings contradict previous literature that reports a preference for local processing among individuals with ASD and enhanced pitch discrimination abilities among children with autism (e.g., Bonnel et al., 2003; Eigsti et al., 2013; Heaton et al., 2008b). Furthermore, the findings do not support the EPF or WCC frameworks. The results suggest that some individuals with ASD have preserved perceptual-cognitive abilities. Therefore, enhanced perceptual processing is not a universal characteristic and perceptual strengths may only be demonstrated by subgroups of individuals with ASD. The lack of support for the EPF and WCC theoretical frameworks in this study may be due to: (1) differences in the tasks used; (2) differences in sample characteristics; and (3) heterogeneity of the autism spectrum. The results do indicate, however, that enhanced local processing in the visuospatial domain is related to enhanced performance in local auditory pitch discrimination tasks in individuals with ASD and those who are TD. More specifically, individuals who have enhanced local

processing in the visuospatial domain also show better performance on the auditory pitch-processing tasks.

Jones et al. (2009) conducted a study including participants with ASD (n=72) and typically developing controls (n=57) matched for IQ and age (mean age: 15;6). The auditory discrimination tasks in this study assessed the discrimination of pitch, intensity, and loudness. The results revealed the following: (1) there was no difference in auditory discrimination abilities at the group level; (2) a subgroup (20%) of individuals with autism demonstrated enhanced pitch discrimination abilities, average intelligence and delayed language development (consistent with Eigsti & Fein, 2013); and (3) differences in auditory discrimination may be related to the behaviors exhibited in response to auditory input (i.e., advanced auditory discrimination abilities could lead to atypical physiological and/or psychological responses).

2.6.2 Enhanced Pitch Memory and Labeling

Heaton, Davis, and Happe (2008a) investigated pitch-labeling abilities among an individual with high functioning autism (AC) (n=1) compared to a group of individuals with absolute pitch (AP) (n=9; age range: 21-60 years). Absolute pitch is the ability to name isolated tones out of context (Heaton et al., 2008a; Heaton et al., 2008c). It is a rare form of memory, estimated to have a prevalence of 1 in 10,000 among individuals who are typically developing. The incidence of AP among individuals with autism, obtained from a parental survey, is 1 in 20 (Heaton et al., 2008a).

The stimuli consisted of complex tones, sine tones, and words. The results indicated that the individuals with AP performed better on tasks including complex tones and sine tones as compared to speech. Consistent with previous findings among control

participants, this demonstrates that their performance was hindered by semantic information. Although both groups demonstrated an ability to name pitch tones in words, the individual with autism performed 5.92 standard deviations higher than the control group. The authors hypothesize that AC's strikingly superior performance as compared to the control group may be explained by: (1) a reduced preference and orientation to social stimuli, including speech and language; and/or (2) increased interest and attention to musical stimuli.

Stanutz, Wapnick and Burack (2014) conducted a study to examine pitch memory in children with autism (n=25; mean age 7;10 years) compared to typically developing children (n=25; mean age 10;5 years) matched for age and intelligence. The first task required the participants to determine if two isolated pure tones were the same or different. Next, they discriminated between two melodies to determine if the pair were the same or different. Additionally, they were assessed 1 week later with respect to their long-term memory for melody.

The results indicated that individuals with autism had enhanced pitch discrimination abilities in both discrimination tasks as compared to controls. Interestingly, the individuals with autism performed more accurately on the pitch discrimination task in the melodic context, which was thought to be the more difficult task. This finding supports the EPF hypothesis, in that the individuals with autism had to discriminate the pitch difference of the leading note within the whole melodic context. Thus, they had enhanced local processing with intact global processing. Furthermore, the group with autism also exhibited superior long-term memory for melody. It is hypothesized that memory for music among individuals with autism may be comparable

to the memory of those with absolute pitch (i.e., the labeling of a tone is instant and does not require working memory). The authors suggest that pitch memory may predict enhanced nonverbal reasoning skills and delayed language development (Stanutz et al., 2014).

Heaton (2003) examined pitch memory, labeling, and chord disembedding among a group of children with autism (n=14; mean age 10;9) compared to typically developing controls matched for age and intelligence. Three experiments were conducted to assess: (1) pitch memory and labeling; (2) disembedding labeled tones from musical chords; and (3) disembedding unlabeled tones from musical chords.

In experiment 1, the child was exposed to four (4) musical tones in conjunction with a picture of 4 different animals. The child was told that each animal had a favorite note, and they were told to name each animal before hearing the associated musical tone. Next, the animal pictures were removed and the child was told what animal liked the tone before hearing the tone. This served as a familiarization and sound association phase. Lastly, the four animal pictures were placed in front of the child. As each tone was presented, the child indicated which animal liked that tone the best (pitch identification). In experiment 2, following a familiarization block, three tones were presented and the child had to point to the animal that was missing. Lastly, in experiment 3, following training, the task was completed using musical stimuli. The children had to listen to music followed by individual tones and determine if the individual tone was in the music that preceded it. Next, individual tones were played before a chord. Again, the child had to determine if the individual tone was part of the chord. The results from this study indicated that the individuals with autism had enhanced pitch memory, labeling, and

chord segmentation. There were no significant differences, however, in experiment 3, which required long-term memory. The findings demonstrate that there are qualitative differences in how children with autism process pitch information compared to control participants.

Heaton, Williams, Cummins, and Happe (2008c) conducted two experiments to assess pitch discrimination and pitch memory in a visuo-spatial format among children with ASD (n=32; mean age 14;6 years) compared to controls (n=35; mean age 14;0 years). Experiment 1 required the participants to identify the distance between two tones. The visual stimulus display included an image of a staircase with 8 steps presented on a laptop screen. Each step represented an octave (i.e., each step had a corresponding tone). A boy moved up and down the stairs based on the sound that was presented. Following a training phase, the participants put the boy on the step that corresponded with a tone they heard. Experiment 2 was completed one week after experiment 1 to assess long-term memory representations for individual musical pitches. Again, the child indicated which step the man would stand on based on the tone presented. Consistent with previous studies, enhanced pitch processing skills were only demonstrated in a subgroup of individuals with ASD. Therefore, homogeneity within samples of children with autism with respect to pitch processing skills cannot be assumed (Heaton et al., 2008c).

2.6.3 Pitch Perception Summary

In summary, approximately 96% of individuals with ASD have atypical sensory processing (Mayer et al., 2016). With respect to the auditory domain, there is substantial evidence on a behavioral level demonstrating enhanced pitch processing, heightened sensitivity to loud noises, atypical orientation to auditory stimuli, atypical processing of

affective and grammatical prosody, and impaired auditory processing in background noise (O'Connor, 2012).

Empirical investigations related to pitch processing support that some children with autism exhibit the following: (1) enhanced pitch processing for simple and complex auditory information (Bonnell, 2003; Eigsti & Fein, 2013; Heaton et al., 2008c; Heaton, Hudry, Ludlow, & Hill 2008b; Stanutz, 2014; Jarvinen-Pasley & Heaton, 2007; Jarvinen-Pasley et al., 2008; Mayer, Hannent, & Heaton, 2016; O'Riordan & Passetti, 2006; Jones et al., 2009; and Chodury, Sharda, Foster, Germain, Tryfon, & Doyle-Thomas et al., 2017); (2) enhanced pitch memory for musical tones (Heaton et al., 2008c; Heaton et al., 2003; and Stanutz et al., 2014); and (3) enhanced pitch labeling of musical tones (Heaton et al., 2003; and Heaton et al., 2008a).

Although there is evidence demonstrating enhanced pitch processing, pitch memory, and pitch labeling at the group level in some studies, some evidence shows that enhanced auditory processing is not a universal characteristic among individuals with autism. Specifically, it has been estimated that 1 in 5 children with autism exhibit enhanced pitch discrimination abilities (Jones et al., 2009). The subgroup of individuals with enhanced auditory processing, such as enhanced pitch perception, may represent a specific phenotype (Heaton et al., 2008c; Jones et al., 2009).

2.6.3.1 Developmental Trajectory of Pitch Processing

With respect to the developmental trajectory of enhanced pitch processing, this characteristic is reported to be exhibited early, but not later, in development among individuals with ASD. This could be due to the significant increase in pitch discrimination abilities among the typically developing groups over time and stable

performance among those with ASD (Mayer et al., 2016). This might explain why, although previous studies have demonstrated that adults with autism demonstrate enhanced pitch discrimination abilities (e.g., Heaton et al., 2005), some studies failed to find a significant difference in performance between adult cohorts with and without autism (Mayer et al., 2016).

2.6.3.2 Speech and Non-Speech Stimuli

Findings from several studies support the hypothesis that children with autism exhibit enhanced pitch processing of speech and non-speech stimuli compared to control groups (Heaton et al., 2008b; Jarvinen-Pasley et al., 2007; Jarvinen-Pasley et al., 2008b).

The findings are inconsistent with respect to how individuals with and without autism perform on pitch discrimination tasks that involve speech versus non-speech stimuli. Some studies show that individuals with and without autism perform more accurately on pitch discrimination tasks involving non-speech stimuli as compared to tasks that assess their ability to discriminate pitch in speech stimuli (Mayer et al., 2016; Heaton et al., 2008; Jarvinen-Pasley et al., 2008). Conversely, another study demonstrates that individuals with and without autism exhibit enhanced pitch processing abilities across different classes of auditory stimuli and do not appear to be impacted by semantic information (Jarvinen-Pasley and Heaton 2007). The control participants in the latter study, however, did demonstrate deteriorated performance in a condition where they were expected to detect differences in temporal patterning of linguistic stimuli as compared to non-linguistic stimuli. Thus, in this study, control participants exhibited semantic capture in a condition requiring temporal processing, but not in the condition requiring pitch processing.

2.6.3.3 Pure Tone Stimuli

Several studies used pure tone stimuli to examine pitch-processing abilities among individuals with autism (Bonnell et al., 2003; Eigsti & Fein, 2013; Chowdury et al., 2017; O’Riordan & Passetti, 2006; and Jones et al., 2009).

While some studies have demonstrated enhanced pitch processing of pure tone stimuli among individuals with autism as compared to control groups (Bonnell et al., 2003; Eigsti & Fein, 2013; O’Riordan & Passetti, 2006), these findings have not been supported by others. For example, Jones et al. (2009) did not find significant differences in pitch discrimination abilities at the group level in their study. A subgroup of children with autism, however, did demonstrate enhanced pitch processing abilities. In accordance with these findings, Chowdhury et al. (2017) did not find significant differences in performance between groups. These authors did not examine their data at the subgroup level. Nonetheless, this supports the hypothesis that enhanced pitch processing abilities are not a universal characteristic among those with ASD (Mayer et al., 2016).

2.6.3.4 Enhanced Pitch Memory and Labeling

The evidence for enhanced pitch memory and labeling among individuals with autism is less abundant than research addressing pitch discrimination. The findings from current studies demonstrate that individuals with autism: (1) perform more accurately on pitch labeling tasks as compared to control groups (Heaton et al., 2008a; Heaton 2003); exhibit superior long-term memory for melody (Stanutz et al., 2014); and have enhanced pitch memory and chord segmentation (Heaton 2003). One study demonstrated that only a subgroup of the individuals with autism demonstrated enhanced pitch discrimination

and memory. Thus, enhanced pitch memory may not be a universal characteristic of individuals with autism (Heaton et al., 2008c).

2.6.4 Pitch Processing in the Context of Cognitive Theories

Among the articles reviewed, support for the cognitive theories was as follows:

(1) two studies supported the WCC theory (Jarvinen-Pasley et al., 2007; Jarvinen-Pasley et al., 2008); (2) ten studies supported the EPF hypothesis (Bonnell et al., 2003; Eigsti & Fein, 2013; Heaton et al., 2008c; Heaton et al., 2008b; Heaton 2003; Stanutz et al., 2014; Jarvinen-Pasley et al., 2007; Jarvinen-Pasley et al., 2008b; Mayer et al., 2016; and Jones et al., 2009); (3) one study failed to support the WCC and EPC hypotheses (Chowdhury et al., 2017); (4) one study interpreted their findings in the context of a hypothesis they titled the Enhanced Discrimination Hypothesis, which appears to be consistent with the EPF model (O’Riordan & Passetti, 2006); and (5) one study interpreted their findings in the context of a developmental model in addition to the EPF model (Mayer et al., 2016). Thus, the majority of studies support the EPF cognitive theory with respect to pitch processing abilities among those on the autism spectrum.

2.6.5 Implications of Auditory Processing Differences in ASD

Understanding auditory processing abilities in individuals with autism has theoretical and clinical implications. Several implications of enhanced pitch processing among individuals with ASD have been outlined in the literature. The implications will be discussed in the sections that follow.

2.6.5.1 Pitch Processing and Language Impairment

Some individuals with autism who demonstrate enhanced auditory abilities also display early-language delays (Jones et al., 2009; Eigsti & Fein, 2013; Heaton et al., 2008a; Stanutz et al., 2014). It has been hypothesized that enhanced pitch processing could result in overly detailed representations of phonological information. Thus, focusing on fine acoustic differences (e.g., pitch) rather than developing phonological categories may impede first word learning. It has been postulated by some authors that enhanced pitch processing of speech stimuli contributes to language impairments in some individuals with autism (Jarvinen-Pasley and Heaton, 2007; Jarvinen-Pasley et al., 2008). Interestingly, most studies find no association between pitch perception and current language skills between the ages of 6 to 15 years of age (Heaton et al., 2008a, Heaton et al., 2008b; Mayer et al., 2016; Eigsti & Fein, 2013). Eigsti & Fein (2013) suggest that this may indicate that: (1) pitch discrimination may be most noticeable during early language acquisition, but not in later fluent language use; or (2) standardized language assessments may be limited in sensitivity. Nonetheless, it is hypothesized that auditory discrimination abilities play a unique role in the development of language.

2.6.5.2 Pitch Processing and Attention

While some studies suggest that attentional abilities may play a role in the discrimination of pitch (Mayer et al., 2016; Heaton et al., 2008a), other authors suggest that enhanced pitch discrimination abilities are not due to attention to human speech. For example, Eigsti and Fein (2013) used tones that were described as being distinct from those that are key to human speech.(500 - 1,500 Hz). Thus, attention to human speech,

whether typical or atypical, cannot explain the enhanced pitch discrimination abilities among individuals with autism in this study.

2.6.5.3 Pitch Processing, Nonverbal Reasoning, and Symptom Severity

Some authors suggest that: (1) pitch memory may predict nonverbal reasoning abilities (Stanutz et al., 2014); and (2) there is a significant relationship between nonverbal abilities and performance on tasks assessing auditory perception (Chowdhury et al., 2017). Additionally, many individuals with autism exhibit atypical behaviors and aversive responses to auditory information in response to everyday sounds (Bonnell et al., 2003). Some authors suggest that enhanced discrimination abilities are associated with symptom severity (Eigsti & Fein, 2013). More specifically, some individuals who have greater symptom severity have better pitch discrimination abilities (Eigsti & Fein, 2013). A better understanding of the behavioral correlates of atypical pitch processing will enable clinicians and researchers to understand and address these treatments through evidence-based assessment and intervention.

2.6.6 Limitations to Auditory Processing in ASD Findings

Although there is some consistency in findings across studies, the results must be interpreted with caution for the following reasons: (1) the heterogeneity of individuals with autism; (2) characteristics of the participants across studies; (3) the inclusion of mostly school-age children with high-functioning autism; (4) differences in matching criteria for control groups (e.g., matched on receptive language, intelligence, or both); (5) small sample sizes; (6) the use of cross-sectional data (Mayer et al., 2016); and (7) varying degrees of task difficulty within and across studies.

2.7 Implications of Atypical Speech, Voice, and Auditory Processing

The prevalence of speech, voice, and prosodic differences among individuals with ASD for grammatical and pragmatic/affective purposes have clinical implications. These differences may hinder one's social and vocational success (e.g., Shriberg et al., 2001). Researchers and clinicians need to be able to identify and differentiate prosodic disorders in both grammatical and pragmatic/affective domains. The ability to characterize different profiles of prosodic disorders will enable professionals to: (1) determine the underlying mechanism most consistent with the disorder (i.e., is it due to social deficits or perceptual-motor deficits, or both?); and (2) develop interventions with appropriate content and form (Shriberg et al., 2001), such as whether interventions should focus on perceptual-motor details, social cognition, or both (Paul et al., 2008). Several research groups have discussed how an understanding of basic speech perception or processing of low-level auditory information can assist researchers and clinicians in understanding how spoken information is processed and in turn, how it affects the production of speech (e.g., Kargas et al., 2016). We know that subgroups of individuals with ASD have atypical prosody (e.g., Kargas et al., 2016) and/or enhanced pitch discrimination abilities (e.g., Jones et al., 2009). Some authors suggest that overly focused auditory processing of speech stimuli may inhibit the development of higher-level language skills, such as prosodic processing and sentence comprehension (Jarvinen-Pasley and Heaton 2007; Jarvinen-Pasley et al., 2008). Other empirical investigations have revealed speech motor control deficits among subgroups of children with ASD (Adams, 1998; Velleman, Andrianopoulos et al., 2009; Gargan & Andrianopoulos, 2018). Although it has been hypothesized that atypical prosody may be impaired due to atypical auditory processing

abilities, motor speech abilities, or a combination of the two (e.g., Bonnef et al., 2011), this has yet to be directly investigated within the same investigation.

CHAPTER 3

METHODOLOGY

3.1 Research Design

A between-groups research design was used in this study. For research question one, the independent variable is group membership (ASD vs. TD) and the independent variables are receptive and expressive percent accuracy scores on a prosody assessment tool, and duration (seconds) and fundamental frequency/pitch (Hz) of lexical stress expressive prosody utterances. For the second research question, the dependent variable is prosody and the independent variables include: group membership; expressive and receptive vocabulary scores; syllables per second on AMR tasks; and percent accuracy on a frequency pattern test. For the third research question, the independent variable is group membership (ASD vs. TD) and the dependent variable is listener perceptions of “naturalness”.. In addition, for the third research question, qualitative perceptual ratings of speech, voice, and prosody across five domains (i.e., speech sound errors, rate of speech, pitch, prosody, and resonance) were analyzed using a 4-point Likert severity scale for each category and speaker.

A power analysis for independent t-tests was calculated with power at .8, an alpha of .05, and a medium effect size of .5 (Cohen, 1988). The results indicate that a sample size of 64 participants was needed to determine between-group differences, should they exist. Previous research in the Autism Spectrum Disorder / Motor Speech Disorder (ASD/MSD) lab at the University of Massachusetts Amherst revealed that a sample size of 64 participants is not feasible within a 1-year time frame (e.g., Boucher, 2013; Gargan

& Andrianopoulos, in progress). Previous literature on the current research topic has used paired sample t-tests or repeated measures ANOVA to examine between-group difference among 15 individuals with ASD and 16 TD controls (Grossman et al., 2010). A balanced sample of 34 participants for the current study was deemed feasible and appropriate for calculation of nonparametric or parametric statistics.

3.2 Participants

A total of 34 participants who met a set of predetermined inclusionary and exclusionary criteria were included in this study. The study consisted of two distinct groups: seventeen (17) individuals with a formal diagnosis of ASD (group 1) between the ages of 7;10 and 19;0 years and seventeen (17) typically developing (TD) individuals matched for age, gender, and language (group 2).

Inclusion criteria required that all participants were between the ages of 7;10 to 19;0 years of age, were monolingual English speakers who lived in Massachusetts, communicated verbally using connected speech, had no uncorrected hearing or visual impairments, no craniofacial or structural differences, no history of self-injury, injury to others, or damage to property in the past six months, and were enrolled in the general curriculum at school, even if modified under an Individualized Education Program (IEP). Additionally, all participants in the group with ASD were previously diagnosed with ASD by a qualified medical professional according to the criteria set forth in the Diagnostic and Statistical Manual of Mental Disorders (4th edition; APA, 1994). Individuals who were TD had no history of autism, speech impairment (mild articulation errors acceptable), language impairment, or other learning disability (Attention Deficit Disorder/Attention Deficit Hyperactivity Disorder [ADD/ADHD] acceptable).

Among the participants in Group 1, 17 individuals (100%) had a definitive diagnosis of ASD. A total of 12 males and 5 females participated. The mean age was 13;4 years, with a range of 7;10 to 19;0 years of age. Co-morbid diagnoses included the following: ADHD (n = 2); anxiety (n = 4); sensory processing disorder (n = 3); post-traumatic stress disorder (n = 2); mood disorder – NOS (n = 2); oppositional defiant disorder (n = 1); and dysgraphia (n = 1). Group 2 was comprised of 11 males and 6 females. The mean age was 13;4 years, with a range of 8;1 to 17;11 years. Co-morbid diagnoses included ADD/ADHD to match group 1 to the extent possible (n = 3) and a mild articulation disorder (n = 1). Four out of five participants with ADD/ADHD were on medication at the time of testing. One participant's mother felt her child no longer had symptoms of ADD/ADHD and they were not on medication. **Table 1** contains the demographic information of the participants.

Participants were recruited through local organizations and school systems. The investigator contacted the superintendent and special education administrators at the local public school districts. In addition, the investigator contacted local organizations, such as Community Resources for People of Autism in Easthampton, MA for dissemination of the recruitment flyer. Interested parents/guardians of individuals under 18 years of age provided verbal consent before discussing their child's eligibility for this study with the principal investigator over the telephone. Individuals who met the predetermined inclusionary and exclusionary criteria specified above were entered into an electronic and encrypted database and coded by an ID number to ensure confidentiality. The electronic database is in a HIPAA secure database approved by the UMass IRB (UMass Box). Paper-based intake forms are stored in a locked file in the ASD/MSD locked laboratory

in 234 Arnold House, to which only the principal investigators have access. Due to slow participant recruitment across an 8-month timeframe, the investigator included 5 participants per group from a previous study conducted by this research group. The current study (2019-5487), as well as the previous study (2016-3216), was approved by the Human Subjects Review Committee at the University of Massachusetts Amherst. The investigator also received IRB approval to combine data from study 2016-3216 and 2019-5487 to increase the current sample size. For ease of reporting, the participants will be discussed as having taken part in the same study, as recruitment, inclusion/exclusion criteria, methods and procedures were nearly identical across the two studies, with the exception of the administration of one additional task in the 2019-5487 study (i.e., the Frequency Pattern Test). Only ten participants per group from the current study (2019-5487) completed the Frequency Pattern Test (FPT).

3.3 Perceptual Judges

A total of two first-year students enrolled in the Speech and Language Pathology master's degree program at the University of Massachusetts Amherst served as perceptual judges (Listener 1 and Listener 2). The students were trained during a two-hour session led by the principle investigator. The training consisted of reviewing objective definitions of speech, voice, and prosodic descriptors, review of the Naturalness Perceptual Rating Scale, and listening to ten connected speech samples that were approximately 15 seconds in duration, produced by speakers with and without autism. At the conclusion of the training session, the two listeners were able to reliably and consistently describe “typical” and “atypical” features of the voice across the 5 categories

(i.e., speech sound errors, rate of speech, pitch, prosody, and resonance), in agreement with the principal investigator and Dr. Andrianopoulos' subjective descriptions.

3.4 Measures

3.4.1 Expressive and Receptive Vocabulary

The EVT-2 is a norm-referenced, standardized assessment tool that was administered to examine expressive vocabulary and word retrieval for children and adults. The EVT-2 takes approximately 10 to 20 minutes to administer on average. The test battery includes two parallel (alternate, but equivalent) forms (form A and form B). Each form consists of example items and 190 test items (representing 20 content areas and parts of speech) that are arranged in order from least to most difficult. For each test item, the examiner presents a picture and stimulus question. The participant responds with one word to orally label the item, provide a synonym, or answer a question (Williams, 2007).

The EVT-2 is normed based on age and grade level by season (e.g., Fall and Spring). The standardization sample consisted of 3,540 examinees ranging in age from 2;6 years to 90 years and older. According to the EVT-2 manual, the normative sample was representative of the English speaking U.S. population (ages of 2;6 years to 81;0 years) and was stratified by race/ethnicity, self or primary caregiver education level, and geographic region. This norm-referenced measure of expressive vocabulary provides highly reliable and valid scores (Williams, 2007).

The reliability of the EVT-2 was estimated using internal consistency (i.e., the correlation of each examinee's total score on the odd numbered items with his/her total

score on the even numbered items), alternate form reliability (calculated using data from 507 examinees who took both form A and form B), and test-retest stability (calculated using 348 of examinees scores who took the same form of the test twice). Split-half reliabilities range from .88 to .97 (i.e., good to excellent) for 28 age groups in the population sample. Alternate form reliabilities are also good to excellent, ranging from .83 to .91. Lastly, test-retest correlations range from .94 to .97, indicating that EVT-2 performance is highly stable over time (Williams, 2007).

The EVT-2 also demonstrates content validity and evidence of strong correlations with other tests, and has undergone studies with special populations. With respect to content validity, the stimulus words are of high or moderately high frequency in American English and could be acquired through common life experiences. The EVT-2 is highly correlated with the Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4; Dunn & Dunn, 2007), ranging from .80 to .84 across age groups. The EVT-2 also correlates strongly with the Comprehensive Assessment of Spoken Language (CASL; Carrow-Woolfolk, 1999) Lexical/Semantic Composite (.80) and Antonyms subtest (.84). In addition, the EVT-2 correlates with the Clinical Evaluation of Language Fundamentals, Fourth Edition (CELF-4; Semel, Wiig, & Secord, 2003) expressive language scores (.77 and .79), but slightly less with CELF-4 receptive language scores (.68 and .69). Correlations with the Group Reading Assessment and Diagnostic Evaluation (GRADE; Williams, 2001) total score range from .60 to .70 (Williams, 2007).

The EVT-2 was also administered to 12 groups that represent special populations (e.g., speech impairments, language delays, language disorders, ADHD). The difference in standardized scores compared to the general population average was significant at the

.001 level for all groups, though performance varied across groups (Williams, 2007). It is important to note that some individuals, such as those with autism, were excluded from these special population samples. There is no available information in the manual regarding how individuals with autism score in comparison to the general population.

The PPVT-4 is a standardized instrument that measures receptive vocabulary among children and adults. It takes approximately 10 to 15 minutes to administer and contains 228 items arranged from least to most difficult. For each item, the examiner says a word and the examinee responds by selecting a picture from an array of four that best illustrates the meaning of the test item. The PPVT-4 and EVT-2 were 100% co-normed using an identical standardization sample. Reliability and validity of the PPVT-4 were conducted by using similar procedures as the EVT-2.

Internal consistency reliability of the PPVT-4 was measured using split-half reliabilities and coefficient alpha of each form. The split-half reliabilities are very high across age and grade ranges, with a mean of .94 or .95 on each test form. Coefficient alpha averaged .97 and .96 for forms A and B, respectively. The reliability for the alternate (but equivalent) forms is strong and ranges between .87 and .93 (mean = .89). Lastly, the test-retest correlation was .93 (calculated using 340 examinees scores who were retested using the same test form 4 weeks after the initial testing). The test-retest correlation indicates that individuals' performance remained stable regardless of factors that could influence performance at different times (e.g., physical or emotional states) (Dunn & Dunn, 2007).

Validity of the PPVT-4 was established by examining content validity, correlations with other tests, and studies with special populations. The PPVT-4 content

was selected from a list of words across 20 content areas that could be illustrated and are a measure of receptive vocabulary for standard American English. The PPVT-4 is correlated with the EVT-2 (.80 to .84 across age groups). Therefore, the PPVT-4 and EVT-2 instruments provide a comprehensive evaluation of receptive and expressive vocabulary. It is also moderately to highly correlated with the CASL and CELF-4 for examinees of elementary school age (mid-.60s to high .70s). According to the manual, the CASL Lexical/Semantic Composite is most closely aligned with the PPVT-4 (correlation = .79). The PPVT-4 is also correlated with the GRADE (.63 on average). Similar to the EVT-2, the special clinical samples performed significantly different (.001) than the general population, but individuals with autism were excluded from the special population samples (Dunn & Dunn, 2007).

3.4.2 Profiling Elements of Prosody in Speech-Communication

The Profiling Elements of Prosody in Speech-Communication 2015 (PEPS-C; Peppe, 2015) is a semi-automated test battery that examines receptive and expressive prosody at the form and function level in adults and children. It includes a receptive and expressive task for six linguistic functions that are conveyed by prosody: (1) turn end: indicating whether an utterance requires an answer or not; (2) affect: indicating mood/emotions/options; (3) lexical stress: disambiguating two-syllable word pairs (e.g., IMport versus imPORT); (3) phrase stress: disambiguating two nouns from a compound noun (e.g., hot dog versus HOTdog); (4) chunking: prosodic phrase boundaries indicating how speech can be verbally ‘chunked’, as in the difference between ‘fruit, salad, and milk’ and ‘fruit-salad and milk’; and (5) contrastive stress or focus: emphasizing one word in an utterance to focus attention on it. The PEPS-C also contains four (4) prosodic

form tasks: (1) two auditory discrimination tasks that require the participant to decide if the prosody/intonation is the same or different in pairs of wordless stimuli; and (2) two imitation tasks that require the production of prosodic variation in 3-syllable words or 6-7-syllable phrases.

Although the PEPS-C is not standardized, it has been used empirically by other researchers with children who have one of the following disorders: autism, pragmatic impairment, specific language impairment, speech impairment, stuttering, and aided hearing loss (Peppe & McCann, 2003). The PEPS-C was normed on 120 southern British English children ages 5-14 years (Wells, Peppe, and Goulendaris, 1995; cited in Peppe & McCann, 2003). It is important to note that the small amount of normative data for the PEPS-C is more than 10 years old and it may not be appropriate for use with different ethnic groups or individuals from special populations.

The PEPS-C has been administered to approximately 191 individuals with High-Functioning Autism or ASD between 2002 and 2017 (Peppe et al., 2006; Peppe et al., 2007; McCann et al., 2007; Peppe et al., 2011; Diehl & Paul, 2012; Diehl & Paul, 2013; Filipe et al., 2014; Gargan and Andrianopoulos (in progress); Velleman, 2002). Of those participants with HFA or ASD, approximately 18% were administered the Standard English version of the test (Diehl & Paul, 2012; 2013; Gargan and Andrianopoulos, in review; Velleman, 2002). It is unlikely that these studies are representative of the larger population of individuals with ASD who are monolingual speakers of American English in the United States, as sample sizes in each study were small. Nonetheless, the available data from these studies related to reliability of scoring the PEPS-C expressive subtests is worth considering.

Empirical studies show good inter-judge reliability of scoring the expressive PEPS-C tasks. For example, Diehl and Paul (2013) reported point-to-point reliability for correct/incorrect ratings on each subtest ranging from .84 to .96 across participants (mean = .88 across subtests). In addition, Gargan and Andrianopoulos (in progress) calculated inter-judge reliability using Cohen's kappa coefficient for the Lexical Stress and Phrase Stress subtests of the PEPS-C (2015). When taking into account agreement by chance, inter-rater reliability on the Lexical Stress task was found to be Kappa = .55. Inter-rater reliability on the Phrase Stress task was found to be Kappa = .65. According to Landis and Koch (1977), there was moderate to substantial agreement between two trained examiners on the PEPS-C 2015 Lexical Stress and Phrase Stress expressive ratings (Gargan & Andrianopoulos, in review).

3.4.3 Acoustic Measurements of Pitch and Duration

Acoustic measurements included analysis of sixteen audio-recorded one-word Lexical Stress utterances from the PEPS-C test per participant using the Multi-Speech acoustic software program, including the Multidimensional Voice Profile Program and Motor Speech Program (Kay Pentax, Model 3700) to measure speech and voice quality on the basis of duration (seconds) and fundamental frequency (Hz). Published norms are available based on age and gender for each of these variables. For example, Wilson (1987) reported acceptable limits of fundamental frequency for males and females based on age. Additionally, Campisi, Tewfik, Manoukian, Schloss, Pelland-Blais, and Sadeghi (2002) also provide normative data on acoustic variables for children who are four to 18 years old.

3.4.4 Oral Diadochokinetic Rate

The rate of diadochokinesis is a widely used index of motor speech skill (McClean, 2000; Williams et al., 2000; cited in Devadiga & Bhat, 2012). Oral diadochokinetic rate is obtained through rapid repetitions of syllables in two tasks: alternate motion rates (AMRs) and sequential motor rates (SMRs). The AMR tasks require an individual to produce rapid repetitions of single syllables (e.g., [pa], [ta], or [ka]) whereas sequential motor rates (SMRs) require an individual to produce rapid repetitions of syllable sequences (e.g., [pa-ta-ka]). AMRs and SMRs are a valid probe of speaking rate and articulatory performance (Devadiga & Bhat, 2012). Normative data for school-age children between the ages of 6 and 13 years are provided in Fletcher (1972).

3.4.5 Frequency Pattern Test

The Frequency Pattern Test (FPT) is a behavioral, criterion-referenced, central auditory processing disorder test battery. It is composed of three 150-msec tones separated by 200-msec intervals. Each tone is either low frequency (L; 880Hz) or high frequency (H; 1122 Hz). There are six combinations of three-tone sequences (LLH, LHL, LHH, HLH, HLL, and HHL). The FPT is available on compact disc (CD) and contains 60 frequency patterns (i.e., 6 combinations by 10 randomizations) with approximately 6 seconds between each pattern. Participants are expected to indicate the pattern of tones they heard (Musiek, 1994).

The FPT has some published findings related to reliability. According to Musiek and Pinheiro (1987), the FPT has good sensitivity (.80) and specificity (.88), indicating acceptable levels of diagnostic accuracy (Musiek & Pinheiro, 1987). In addition, Humes, Coughlin, and Talley (1996) demonstrated excellent test-retest reliability. However, the

available sensitivity, specificity, and test-retest findings are limited in that they are based on studies consisting of elderly subjects, or individuals with cerebral, brainstem, and cochlear lesions (Musiek & Pinheiro, 1987; Humes et al, 1996).

There are some published normative data available for the FPT. Musiek (2002) provides the following norms based on age (years; months): 8;0 to 8;11, 40% accuracy; 9;0 to 9;11, 65% accuracy; 10;0 to 10;11, 72% accuracy; and 11;0 through adulthood, 75% accuracy. These accuracy levels were identical for the left and right ear. Additional normative data for the FPT can be found for children between the ages of 7 and 12 years of age in Kelly (2007) and McDermott et al. (2016). Kelly's (2007) findings are comparable in each ear. Mean performance for the right ear revealed the following per age group: 7;0-8;11, 73%; 9;0-10;11, 85%; and 11;0-12;11, 91%. Mean performance levels for the left ear reveal the following per age group: 7;0-8;11, 71%; 9;0-10;11, 87%; and 11;0-12;11, 93%). Based on the available findings, there is an improvement in scores across age groups, indicating that discrimination of frequencies matures over time and becomes more accurate around 11 to 12 years of age (e.g., Kelly, 2007).

With respect to cut-off scores, Musiek (1994) reported a cutoff score of 78%, but this finding is based on a study including adults. Bellis (2003) reports a wider range of cutoff scores for children (i.e., 35% for 7 year olds to 80% for 12 year olds). It should be noted that the children in Bellis (2003) were allowed to hum their response or use visual cues, rather than a verbal response. Furthermore, the cut-off scores provided by Bellis (2003) are not ear-specific.

3.4.6 Naturalness Perceptual Rating Scale

The Naturalness Perceptual Rating Scale (adapted from Andrianopoulos et al., 2015; and Gargan & Andrianopoulos, in review) consists of two parts. Part one requires trained listeners to rate connected speech samples across five categories (i.e., speech sound errors, rate of speech, pitch, prosody, and resonance) with respect to presence and severity of the descriptors using a 4-point Likert scale, where: 0 = not present, 1 = mild/negligible, 2 = moderate, and 3 = severe. Part two requires trained listeners to rate the same connected speech samples on overall “naturalness” of speech, voice, and prosody using a binary scale, where 1 = unnatural and 2 = natural. This instrument is not norm-referenced. However, pilot data from Gargan and Andrianopoulos (in review) revealed moderate agreement between two raters ($Kappa = .52$; Landis & Koch, 1977) who completed the rating scale after listening to vocal recordings of the expressive PEPS-C (2015) lexical stress and phrase stress tasks produced by adolescents with ASD ($n=11$) and TD individuals ($n=11$).

3.5 Procedures

The assessment protocol was designed to examine expressive and receptive language abilities, expressive and receptive prosodic abilities, motor speech, programming, and planning abilities (i.e., performance on AMR and SMR tasks), auditory processing skills (i.e., pitch processing), perceptual ratings of severity of connected speech samples across 5 categories (i.e., speech sound errors, rate of speech, pitch, prosody, and resonance), and overall “naturalness” of prosody during connected speech.

The study was conducted in Dr. Andrianopoulos' ASD/MSD Laboratory and at the Speech, Language, and Hearing Clinic at the University of Massachusetts Amherst. Participants completed all experimental tasks in one session, which lasted approximately 2 hours. Breaks and snacks were provided as necessary. The PI directly supervised 100% of the data collection sessions.

A total of five graduate student Research Assistants (RAs) enrolled in the master's program in Speech Language Pathology at the University of Massachusetts Amherst assisted with the collection of data. The PI also collected a portion of the data. The RA's were trained by the PI in the administration of the protocol and were deemed reliable and accurate in their administration and scoring of tests prior to the data collection phase. The RAs were blind to group membership and the purpose of the study. All investigators directly involved in research of human subjects were certified in the Collaborative IRB Training Initiative (CITI). The specific breakdown of the five RAs' assistance per task in the protocol is provided next.

Four RAs were trained to administer and score the EVT-2 and PPVT-4. Training consisted of reviewing the assessment tool manual and practice administering the test at least twice prior to the data collection phase. One graduate student RA in SLP (Examiner 1) was trained to administer and subjectively rate the participants' expressive prosody on the PEPS-C according to the scoring procedures outlined in the PEPS-C manual. Training consisted of review of operational metrics (e.g., definition of prosody, definition of stress, definition of an accurate versus inaccurate response on the PEPS-C tasks) prior to the data collection phase. Examiner 1 reviewed all objective definitions, learned the scoring procedures per task, completed the PEPS-C training program, and scored the PEPS-C

expressive subtests for five practice participants. Examiner 1 also practiced administering the test at least three times prior to the data collection phase. The PI and Examiner 1 achieved “moderate” to “substantial” inter-judge agreement on the five practice participants according to Landis & Koch’s (1977) interpretation of Cohen’s Kappa on the practice items. This ensured that the subjective measures would be obtained consistently during the data collection phase of this project (Kazdin, 2011).

Examiner 1 was also trained to administer and score the AMRs, SMRs, and the FPT. Training consisted of reviewing the manuals, task instructions, and practice administering and scoring the test at least twice prior to the data collection phase.

Language testing was completed in a quiet room free of distraction. Prosody, AMR, SMR, frequency pattern testing, and spontaneous speech tasks were administered in a 1:1 setting in a sound-treated, double-walled chamber (Industrial Acoustics Company, Inc., Bronx, NY, Model 1604) with an ambient noise level of 25 dBA. All expressive prosody, AMR, SMR, and connected speech tasks were acoustically recorded. A miniature head-mounted professional grade condenser microphone [AKG C-410] was positioned on each participant with a 45° off-axis with a mouth-to-microphone distance held constant one inch to the right of each participant’s mouth. The speech samples were recorded using the Tascam portable multi-track digital recorder (DR-680) at a sampling rate of 44.1 kHz. The input level control was adjusted to prevent clipping and digital distortion. At the beginning of testing, the microphone was adjusted to allow headroom of 10 dB on the digital recording, and this setting remained constant at this level without adjustment throughout the remainder of testing. The recorded acoustic data was subjected to post-hoc analyses for the expressive Lexical Stress utterances on the PEPS-C, to obtain

trained listener perceptual ratings of “naturalness” based on a 20-second connected speech sample, and inter-rater reliability calculations. The acoustic samples were edited and analyzed using the Multi-Speech acoustic software program, including the Multidimensional Voice Profile Program (MDVP; Kay Pentax, Model 3700).

3.5.1 Standardized Language Testing

The Expressive Vocabulary Test, Second Edition (EVT-2; Williams, 2007) and the Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4; Dunn & Dunn, 2007) were administered to assess expressive and receptive vocabulary, respectively.

3.5.2 Assessment of Prosody

The general American English version of the PEPS-C (Peppe, 2015) was administered to assess grammatical, pragmatic, and affective prosody. The participants were seated in a chair in front of a Dell Latitude laptop and next to the examiner to complete the receptive and expressive versions of these PEPS-C tasks.

Prior to administering the PEPS-C, a vocabulary check was completed with each participant. The participants were encouraged to ask for a definition of the words in the PEPS-C that were unfamiliar to them. The instructions for each subtest were read to each participant. Two sample items were administered prior to the 16 experimental trials per subtest to confirm that participants understand the expectations of the task. Each subtest score ranged from 0 to 16 points and each participant completed the experimental tasks and stimulus items in the same order, consistent with the PEPS-C (2015) instructions. The raw scores were converted into percent accuracy scores ranging from 0%-100% accuracy per subtest. Participants “pass” the subtest if they achieve a score of 12 out of

16 correct (i.e., 75% accurate or higher), which indicates a degree of competence according to the PEPS-C manual (Peppe, 2015).

In the receptive tasks, a pre-recorded auditory stimulus produced by a female American English speaker was presented through the laptop speakers with two corresponding visual stimuli as response choices on a laptop screen. The participants were told that they would hear the speaker say something, and they were expected to press on the left or right side of the laptop screen to select the written text that corresponded with what they heard. The participants responses per item were automatically scored by the PEPS-C (2015) software as 1 (correct) or 2 (incorrect) based on where the stress was actually produced in the recording in comparison to the participants selected response. The scores were transferred into an Excel Spreadsheet, which is part of the PEPS-C software package. For the expressive tasks, the same stimuli per subtest appeared on the laptop screen one at a time and the participants were told to read what they saw aloud. The trained examiner was blind to the stimulus items and rated each utterance in real-time on a separate keypad. The score was automatically transferred into an Excel Spreadsheet by the PEPS-C software (Peppe, 2015).

The PEPS-C expressive task recordings from one subtest (Lexical Stress) were analyzed for duration (seconds) and fundamental frequency/pitch (Hz) of utterances. The investigator trimmed each recorded expressive utterance per participant using the MDVP software program by moving a cursor precisely to the beginning and end of the waveform per utterance. The trimmed file was checked three times by the PI to make sure it was trimmed as close to the waveform as possible, while also making sure part of the utterance was not accidentally trimmed off. A total of 544 Lexical Stress utterances were

trimmed and analyzed by the PI using this procedure. The duration, pitch, minimum pitch, and maximum pitch measurements were manually entered into an Excel Spreadsheet for each utterance.

3.5.3 Assessment of Speech Motor Control

Each participant was read instructions outlined in Fletcher (1972) before producing Alternate Motion Rates (AMRs) by repeating each of [pΛ], [tΛ] and [kΛ] as many times as possible to measure speed and coordination of the articulators; and Sequential Motor Rates of [pΛtΛkΛ] to assess sequencing, motor programming and planning and coordination of the articulators and speech mechanism. Three tokens for each AMR and SMR task per participant were averaged together to provide a stable sample of the participant's overall motor speech abilities. All productions were acoustically recorded for analysis. The participants' average syllables per second were calculated by the PI and entered into an Excel spreadsheet and compared to published norms (Kent, Kent & Rosenbeck, 1987). In addition, the first 3 to 6 seconds of each SMR production was coded by the investigator and entered into the Excel spreadsheet as: 2 = 75% to 100% correct; 1 = 50% to 75% correct; and 0 = less than 50% correct. The scores were averaged together across the three tokens.

Examples of productions for each code are provided below:

- 2: /p^t^k^/ for 100% of the participants' production
- 1: /p^t^k^/, /p^t^k^/, /k^p^t^k^/, /p^t^k^/, /k^/, /p^t^k^/, /p^t^k^/, /p^k^t^/, /p^k^/
- 0: /p^t^k^/, /p^k^t^/, /p^k^t^/, /k^p^t^k^/, /k^p^t^/, /k^p^t^/

3.5.4 Assessment of Pitch Processing

The FPT was administered according to the author's instructions. The participants were told that they would hear three consecutive tones that would either be high or low in pitch. According to Musiek (2002) verbal responses, humming the patterns, or pointing to high and low visual displays are all acceptable modes of responses. The meaning of "high" and "low" pitch was explained to each participant, with visuals and examples. Two sample patterns were produced by voice by the examiner, with a corresponding visual, and the examiner said the pattern and check for understanding. Next, two sample patterns were produced by voice by the examiner without a visual and the examiner and participant agreed on the pattern produced. One additional pattern was produced by voice by the examiner and the participant was expected to verbally state the pattern. Finally, the examiner administered three practice items from the FPT starting at 1:00 minute on the audio recording. Musiek (2002) states that if an individual performs poorly when responding verbally, it is worthwhile to ask the participant hum their response. Therefore, participants in the current study who had difficulty during practice with a verbal response were asked to hum their response instead. Once the examiner was comfortable with the participants' responses on the practice items, formal testing began, starting with the right ear (Musiek, 2002).

The test was administered at 50 dB HL and 30 patterns were presented to each ear. The FPT CD has 6-7 seconds between sound patterns; however, the participants were provided as much time as they need to respond to each pattern (i.e., the examiner paused the CD as needed to allow for a response). Patterns were not repeated. Scoring guidelines provided by Musiek (2002) are outlined below:

- Patterns were not repeated, unless the participant did not have enough time to respond or did not hear the pattern.
- If the participant responded with a pattern of more than 3 tones (e.g., HLLL for HLL), that was scored as incorrect.
- Pattern reversals (e.g., HLH instead of LHL) were scored as incorrect.
- If the participant performed inaccurately on the first 14 or 15 items, or got 14 or 15 correct, testing was terminated.

The percent accuracy scores per ear, per participant, were manually entered into an Excel spreadsheet.

3.5.5 Assessment of Speech, Voice, and Prosodic Naturalness

Participants engaged in three, 1-3 minute connected speech samples with the examiner. The recording with sufficient spontaneous language (i.e., multiple segments of at least 3-second utterances of connected speech per sample) was selected for analysis.

To collect a spontaneous speech sample, the examiner engaged the participant in a conversation about school, hobbies, or a favorite game, book, or activity. Sample questions and prompt items were generated for the examiner to use. The prompt questions ensured that the spontaneous speech samples were controlled for content. A list of possible questions or prompts is provided below:

- Tell me about your favorite game.
- What is the goal of the game?
- Who is your favorite character? Why? Can you describe them?
- What is your favorite book? What was it about?

- What is your favorite thing to do when you're not at school?

Next, a connected speech sample was obtained through a story-telling task. The participant was instructed to tell a story while looking at the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2; Lord, Luyster, Gotham, & Guthrie, 2012) cartoon story sequence consisting of 5 sequenced pictures. Lastly, the participant completed a picture description task using the ADOS-2 picture description materials.

As mentioned previously, the PI selected a connected speech sample from one of the tasks above that contained at least 20 seconds of connected speech. The majority of connected speech samples were spontaneous (i.e., 15 in the ASD group and 12 in the TD group), while a small proportion of the connected speech samples selected for the listening study in each group were obtained through the picture description or story-telling task (i.e., 2 in the ASD group and 5 in the TD group). Two trained listeners (Listener 1 and Listener 2) applied the Naturalness Perceptual Rating Scale (Gargan & Andrianopoulos, in review) to rate the acoustically recorded, 20-second connected speech sample for each participant with respect to speech sound errors, rate of speech, pitch, prosody, resonance, and overall “naturalness”. The listeners completed the Naturalness Perceptual Rating Scales individually, within one week of training, during one two-hour listening session. The following guidelines were provided on the day of the listening study:

- Use the Apple desktop computer in the ASD/MSD laboratory
- Use the provided noise cancelling headphones (Sennheiser closed dynamic headphones)

- Set the volume at a comfortable level, which will remain constant for the entire listening session
- Listen to seven (7) “anchor items” prior to starting to re-orient you to the descriptor items and rating scale
 - The anchor items consisted of sample acoustic recordings of speakers with and without autism. The anchor items represented certain descriptors on the Naturalness Perceptual Rating Scale to orient the listeners’ ears before starting the study. For example, one anchor item was labeled “severe atypical prosody; mild articulation distortions on /s/; moderate nasal resonance” while another anchor item was labeled, “typical / no concerns”, and another anchor item was labeled, “moderately slow rate, moderate atypical prosody, moderate nasal resonance, and moderately monotone”.
- Only repeat a sample one time (i.e., listen to a sample 2 times maximum per participant in its entirety)

The ratings were entered into an Excel spreadsheet by the PI for each participant, per judge.

3.6 Statistical Procedures

All data were entered into an Excel Spreadsheet and analyzed using the R programming language (R Core Team, 2015). Two-sample t-tests were used to determine if there were differences between groups with respect to age and receptive and expressive language. To address the first research question, the PEPS-C overall percent accuracy scores, PEPS-C subtest percent accuracy scores, and the mean duration (seconds), mean

pitch, mean minimum pitch, and mean maximum pitch for the Lexical Stress utterances were subjected to the Welch-Aspin t-test to account for the unequal variances between the two groups on these tasks (Delacre, Lakens, & Leys, 2017). To address the second research question, a multiple linear regression was calculated to predict expressive prosody (DV) based on receptive vocabulary, expressive vocabulary, AMRs and group membership (IVs). Descriptive analysis was also completed for the same research question. Performance on the FPT is reported on using descriptive analysis, as only ten participants per group completed the pitch-processing task. Fischer Exact tests were conducted to determine if a significantly higher proportion of individuals with ASD failed the FPT per ear in comparison to TD peers. The speech motor control tasks are also reported on descriptively to examine meaningful differences between groups and identify characteristics consistent with a motor speech disorder, should they exist. Lastly, the proportion of individuals who performed sub-optimally on prosody, language, speech motor control, and/or pitch processing tasks is discussed to qualitatively compare the groups across all variables and identify meaningful differences between groups. To address the third research question, individual trained listener responses from Part 1 of the naturalness perceptual rating scale are discussed descriptively. In addition, Chi-Square analysis was completed for Part 2 of the rating scale responses, using the agreed-upon naturalness ratings between the two listeners to determine if group membership and naturalness ratings are associated. Confidence Intervals at 95%, degrees of freedom, and level of statistical significance were calculated. An alpha level of 0.05 was used for all statistical tests. Effect size was calculated using Cohen's d , where $d < .2$ = small; $d < .5$ = moderate; and $d < .8$ = large effect (Cohen, 1988). Inter-judge reliability will be

calculated using Cohen's kappa, where $\kappa < 0$ = poor; 0.0-0.20 = slight; 0.21-0.41 = fair; 0.41-0.60 = moderate; 0.61-0.80 = substantial; and 0.81-1.0 = almost perfect (Landis & Koch, 1977).

CHAPTER 4

RESULTS

4.1 Expressive and Receptive Vocabulary

Seventeen individuals with a formal diagnosis of ASD between the ages of 7;10 (years;months) and 19;0 (mean age: 13;4 years; SD: 3.43) and 17 TD individuals between the ages of 8;1 and 17;11 (mean age = 13;4 years; SD = 2.86) participated in this study.

To ensure the groups were similar in terms of vocabulary skills, the Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4; Dunn & Dunn, 2007) was administered to evaluate receptive vocabulary and the Expressive Vocabulary Test, Second Edition (EVT-2; Williams, 2007) was administered to assess expressive vocabulary. We did not find significant differences at the 0.05 alpha level between groups on the PPVT-4 ($t = -1.89$, $df = 24$, $p = 0.07$, 95% CI [-22.6, .95]) or on the EVT-2 ($t = -.9$, $df = 27$, $p = 0.37$, 95% CI [-18.99, 7.34]).

Two participants with ASD scored “extremely low” on the PPVT-4 and the EVT-2 and one participant with ASD scored “moderately low” on the PPVT-4. The investigator did not eliminate these participants, as it did not have an effect at the group level. In addition, individuals with autism who have low language abilities are often excluded from research studies examining prosody in autism. This investigator included these participants to better represent the larger, heterogeneous population of individuals with autism with respect to language ability.

4.2 Prosody

Prosodic abilities of individuals with ASD in comparison to their TD controls will be discussed on perceptual and acoustic levels. The perceptual ratings of prosody will be discussed in the context of the PEPS-C results. Acoustic features of prosody (i.e., duration and pitch) were obtained and quantified using the Multi-Speech and Multi-Dimensional Voice Profile (MDVP) programs from Kay Pentax.

4.2.1 PEPS-C Results

The mean receptive and expressive composite scores from the PEPS-C test will be discussed in this section. In addition, the receptive and expressive results per PEPS-C subtest will be provided to determine which tasks the groups scored significantly different on. The summary of results for the PEPS-C receptive and expressive tasks are provided in **Table 2**.

PEPS-C Composite Scores: Overall, the group with ASD performed with significantly less accuracy than the TD controls on the receptive prosody test composite score ($t = -2.803$, $df = 22$, $p = .01$, 95% CI $[-18.04, -2.7]$, Cohen's $d = .96$) and the expressive prosody test composite score ($t = -3.34$, $df = 17$, $p = .003$, 95% CI $[-21.65, -4.95]$, Cohen's $d = 1.14$). The individuals with ASD exhibited differences in prosodic abilities on receptive and expressive tasks in comparison to TD controls.

Auditory Discrimination: The auditory discrimination task examined the participants' ability to identify whether the intonation/prosody of muffled words (2-3 syllables) and phrases (6-7 syllables) were the same or different. There were no significant differences in performance among the participants with ASD and the TD controls ($t = -1.6921$, $df = 31$, $p = .1$, 95% CI $[-19.32, 1.79]$, Cohen's $d = .58$).

Imitation: The imitation task examined the participants' ability to imitate the intonation of 2-syllable words and 6-7-syllable phrases. The TD group performed significantly better than the ASD group on this task ($t = -2.52$, $df = 17$, $p = .021$, 95% CI $[-19.85, -1.78]$, Cohen's $d = .86$).

Turn End: The Turn End task assessed the ability to understand and produce questions versus statements. The TD group performed significantly better than the ASD group on the receptive task ($t = -2.96$, $df = 16$, $p = .009$, 95% CI $[-24.40, -4.06]$, Cohen's $d = 1.01$); however, no significant difference in performance between groups was observed on the expressive task ($t = -2.00$, $df = 19$, $p = .059$, 95% CI $[-21.16, .45]$, Cohen's $d = .68$).

Affect: The Affect task assesses the ability to understand and express affect (like versus dislike) in single words. There were no significant differences in performance between groups on the receptive task ($t = -.922$, $df = 28$, $p = .363$, 95% CI $[-10.59, 4.01]$, Cohen's $d = .31$) or on the expressive affect task ($t = -1.504$, $df = 32$, $p = .142$, 95% CI $[-21.60, 3.24]$, Cohen's $d = .51$).

Lexical Stress: The Lexical Stress task assesses the ability to perceive and produce grammatical stress in two-syllable words (e.g., 'IMprint' versus 'imPRINT'). There were no significant differences between groups on the receptive lexical stress task ($t = -1.35$, $df = 31$, $p = .184$, 95% CI $[-19.43, 3.90]$, Cohen's $d = .46$). The TD group performed significantly better than the group of participants with ASD on the expressive lexical expression task ($t = -2.95$, $df = 25$, $p = .006$, 95% CI $[-27.33, -4.90]$, Cohen's $d = 1.01$).

Phrase Stress: The Phrase Stress task assesses the ability to distinguish and produce two words versus a compound noun (e.g., ‘I saw a blue bird today’ versus ‘I saw a bluebird today’). The TD group of participants performed significantly better than the participants with ASD on the receptive phrase stress task ($t = -2.656$, $df = 31$, $p = .01$, 95% CI [-24.73, -3.26], Cohen’s $d = .91$) and on the expressive phrase stress task ($t = -3.34$, $df = 18$, $p = .003$, 95% CI [-21.65, -4.95], Cohen’s $d = 1.42$).

Boundary: The Boundary test assesses the ability to comprehend syntactically ambiguous phrases and produce syntactically ambiguous phrases unambiguously (e.g., ‘chicken, fingers, and fruit’ versus ‘chicken-fingers and fruit’). One participant with autism (A4) was eliminated from the statistical analysis for the receptive and expressive portions of this task, as he did not complete the task due to fatigue. The TD participants as a group performed significantly better than the group of participants with ASD on the receptive boundary task ($t = -2.70$, $df = 16.79$, $p = .015$, 95% CI [-23.43, -2.87], Cohen’s $d = .95$). However, there were no significant differences between groups on the expressive boundary task ($t = -1.017$, $df = 23$, $p = .322$, 95% CI [-15.71, 5.38], Cohen’s $d = .35$).

Contrastive Stress: The Contrastive Stress task assesses the ability to identify and produce contrastive stress (e.g., ‘The GREEN sheep has the ball’ (not the blue one)). The group of TD participants performed significantly better than the group with ASD on the receptive contrastive stress task ($t = -2.3786$, $df = 18$, $p = .028$, 95% CI [-28.80, -1.78], Cohen’s $d = .81$) and on the expressive contrastive stress task ($t = -2.92$, $df = 16$, $p = .009$, 95% CI [-36.32, -5.85], Cohen’s $d = 1.00$).

4.2.1.1 Intra- and Inter-judge Reliability of PEPS-C Expressive Scores

Inter-judge reliability of the PEPS-C expressive scores was calculated using four randomly selected participants (ASD $n = 3$; TD $n = 1$). A total of 256 expressive utterances across four expressive prosody tasks (i.e., Lexical Stress, Phrase Stress, Contrastive Stress, and the Affect task) produced by the four randomly selected participants were scored by the PI and Examiner 1 to calculate inter-judge reliability. These four subtests were selected for intra- and inter-judge reliability as they represent grammatical, affective, and pragmatic domains of expressive prosody. In addition, the PI scored 256 expressive prosody utterances on a second occasion, 4-months post data collection, to calculate intra-judge reliability (ASD $n = 2$; TD $n = 2$). Cohen's kappa coefficient was used, since this statistic takes into account agreement by chance to measure inter-rater reliability between two judges or intra-rater reliability for one judge. According to criteria outlined by Landis & Koch (1977), observer agreement for categorical data can be divided into six levels of strength based on the kappa value: poor (<0.00), slight ($0.00-0.20$), fair ($0.21-.40$), moderate ($0.41-0.60$), substantial ($0.61-0.81$), or almost perfect ($0.81-1.00$).

Inter-judge reliability for the PEPS-C expressive utterances selected for analysis yielded moderate (i.e. Kappa = 0.43) to almost perfect (i.e. Kappa = 1) agreement. Similarly, intra-rater reliability for the PEPS-C expressive utterances selected for analysis yielded substantial (i.e., Kappa = .71) to almost perfect (i.e., Kappa = 1) agreement. This level of agreement was deemed adequate to support the results in the current study. The inter- and intra-judge reliability findings can be found in **Table 3** and **Table 4**.

4.2.2 Acoustic Measurements of Pitch and Duration

The acoustic correlates of prosody include duration, fundamental frequency (i.e., pitch), and intensity/loudness. In the current study, duration (seconds) and pitch (Hz) measurements obtained from the MDVP for all Lexical Stress task expressive utterances were examined. The Lexical Stress prosody task consisted of elicited two-syllable words with first-syllable stress (nouns) or second-syllable stress (verbs). Group statistics for the acoustic measurements per group can be found in **Table 5**.

The average duration of utterances ranged between .76 – 1.31 seconds in the ASD group and between .61 – 1.49 seconds in the TD group ($M = 1.01$ vs. $.87$, $SD = .19$ vs. $.19$, respectively). As a group, participants with ASD demonstrated significantly longer duration of utterances on the Lexical Stress expressive prosody task ($t = 2.155$, $df = 32$, $p = .03$, 95% CI $[.007, .270]$, Cohen's $d = .73$). **Figure 1** shows boxplots for acoustic measurements of duration (seconds).

Closer inspection of the data revealed that the words with first-syllable stress were produced with shorter average durations at the group level in comparison to words with second-syllable stress. This finding is not surprising, as first-syllable stressed words are typically produced with shorter durations than second-syllable stressed words in American English (Grossman et al., 2010). A two-sample t-test comparing average difference scores between words with first- and second-syllable stress revealed that there were no significant differences between groups ($t = -.152$, $df = 32$, $p = .88$, 95% CI $[-.58, .50]$, Cohen's $d = .05$). In other words, the difference in stress conditions did not depend on group membership, as 13 participants per group (76%) exhibited shorter average durations on first-syllable words in comparison to second-syllable words. Although the

average durations for first syllable words were shorter than second-syllable words, this difference was not large enough to be considered statistically significant for the ASD group ($t = .851$, $df = 16$, $p = .407$) or for the TD group ($t = 1.49$, $df = 16$, $p = .155$).

By contrast, no significant differences between the ASD and TD groups were found with respect to mean pitch ($t = .526$, $df = 30$, $p = .6$, 95% CI [-20.58, 34.89], Cohen's $d = .18$), minimum pitch ($t = .779$, $df = 31$, $p = .44$, 95% CI [-12.87, 28.79], Cohen's $d = .26$), or maximum pitch ($t = .127$, $df = 32$, $p = .89$, 95% CI [-37.38, 42.37], Cohen's $d = .04$) of Lexical Stress expressive utterances. Although group differences in pitch did not reach statistical significance, it is worth discussing some observed patterns in the data. At the group level, the mean pitch was higher among ASD speakers compared to TD controls ($M = 182$ Hz vs. 174 Hz, $SD = 43$ Hz vs. 35 Hz, respectively) and the maximum pitch was higher in the ASD group in comparison to TD peers ($M = 244$ Hz vs. 242 Hz, $SD = 59$ Hz vs. 54 Hz, respectively). These patterns are consistent with previous literature demonstrating higher mean pitch and a higher maximum pitch among speakers with ASD compared to TD controls (Filipe et al., 2014). The minimum pitch for the ASD group and TD group was 139 Hz. and 131 Hz., respectively. The comparable minimum pitch level across groups is also consistent with previous findings demonstrating no significant differences in minimum pitch levels (e.g., Filipe et al., 2014).

4.3 Predictors of Expressive Prosody

A multiple linear regression analysis was calculated to predict prosody based on receptive vocabulary (PPVT), expressive vocabulary (EVT), speech motor control (AMRs), and group membership [i.e., the “Full Model” with all variables]. The results of

the Full Model revealed EVT and AMRs not to be statistically significant predictors to the model ($p > .05$). However, results of the multiple linear regression calculated to predict prosody based on PPVT and group membership [i.e., “Sub Model” with PPVT-4 and group] revealed a statistically significant regression equation ($F(2,31) = 19.54$, $p = .000$), with an adjusted R^2 of .53. An ANOVA was also conducted to determine if the Full Model explains more than the Sub Model. The ANOVA analysis revealed results that are not significant ($p > .05$), indicating that EVT and AMR scores do not account for a significant amount of variability in prosody, above and beyond PPVT and group membership. Participants’ predicted expressive prosody is equal to $41.208 - 8.53 + .44$, where group is coded as 1 = ASD, 0 = TD, and PPVT is measured in standard scores with 1-point increments. Participants expressive prosody increased .44 points for each 1-point increase on the PPVT and participants in the ASD group had a 9-point decrease in expressive prosody. Both PPVT scores and group membership were significant predictors of expressive prosody. Overall, PPVT and group membership accounted for 53% of the variability in expressive prosody scores in the Sub Model.

To take the analysis one step further, multiple linear regression analysis was conducted to determine if there is an interaction between PPVT-4 and group membership as predictor variables [i.e., “Theoretical Interaction Model”]. The results revealed a statistically significant regression equation ($F(3,30) = 20.03$, $p = .000$), with an adjusted R^2 of .63. Participants’ predicted expressive prosody is equal to $98.354 - 80.33 + .05$, where group is coded as 1 = ASD, 0 = TD. ANOVA analysis revealed results that were significant ($p = .003$), indicating that the Sub Model was not significantly different than the Theoretical Interaction Model. For the ASD group, the regression equation is Prosody

= .58 PPVT + 17.92. For the TD group, the regression equation = .06 PPVT + 98.35.

This indicates that there is an interaction between group and PPVT-4 as predictor variables. Overall, PPVT-4 and group membership accounted for 63% of the variability in expressive prosody scores, with a moderate positive relationship in the ASD group.

Please see **Table 2**.

The SMR and Frequency Pattern Test (FPT) data were not included in the multiple linear regression models, as the data for the SMRs was categorical and only ten participants per group completed the FPT. Nonetheless, quantitative and qualitative differences during the SMR and FPT were noted within and between groups. First, six participants with ASD (35.3%) were only 50%-75% accurate on average across the three SMR tokens and three participants with ASD (18%) were less than 50% accurate. As a group, nine participants with ASD (53%) scored with less than 75% accuracy on the SMR tasks using the scoring procedures created for this study (Please see Measures and Methods). In contrast, only three TD participants scored between 50%-75% accuracy on SMR tasks, while the remaining 14 TD participants (82%) scored with at least 75% accuracy on average.

With respect to the relationship between prosody and pitch processing, five participants with ASD (50%) did not pass the FPT task based on normative data for their age (Musiek, 2002) in at least one ear as follows: four individuals with ASD (24%) failed the FPT task in both ears and one individual with ASD (6%) failed the FPT task in the left ear. In contrast, only one TD participant (6%) failed the FPT in their left ear. The remaining participants in both groups (82%) passed the test in both ears. Although a higher number of individuals with ASD failed the FPT test in each ear, the results were

not significant ($p > .05$). Please see **Table 6**, **Table 7** and **Table 8** for descriptive results per group.

4.4 Speech, Voice, and Prosodic Naturalness

To address the third research question, audio-recorded connected speech samples were trimmed to 20 seconds in duration by the PI. The PI selected the best 20-second sample for each participant based on the quality of the recording and to capture at least 20 seconds of continuous connected speech for each speaker. In the ASD group, 14 of the connected speech samples were spontaneous, whereas 2 samples were selected from the picture description task (i.e., the Cookie Theft) and one sample was perceived by the PI to be scripting or delayed echolalia. In the TD group, 12 of the connected speech samples were spontaneous, four samples were obtained from the Cookie Theft task, and 1 sample was obtained from the ADOS Cartoon Story-Telling task, in which the participant told a story while looking at pictures. Although the type of connected speech sample varied within each group, the majority of samples included spontaneous speech and the picture description or story-telling recordings were subjectively judged by the PI to be representative of the speakers' spontaneous speech.

Mode Likert-scale ratings from the Naturalness Perceptual Rating Scale revealed that the two trained listeners rated the group of speakers with ASD as having mild/negligible differences in rate of speech (i.e., lengthened syllables, pauses between syllables, too fast, too slow) and moderate differences in pitch (i.e., too high or too low), prosody (e.g., “sing-song”, “different”), and resonance (i.e., hyper/hypo-nasality) during 20-second connected speech samples. In contrast, the mode Likert-scale severity ratings for the TD group indicated mild/negligible differences in resonance.

4.4.1 Reliability of the Naturalness Perceptual Rating Scale

Inter-judge reliability for ratings of naturalness yielded slight agreement (i.e. Kappa = 0.16) among the speakers with ASD and substantial agreement (i.e. Kappa = .76) agreement among the TD speakers. Differences in inter-judge agreement across the two groups may suggest that it is more challenging to perceptually analyze the speech, voice, and prosody of speakers with ASD versus speakers who are TD. Nonetheless, when considering inter-judge reliability across all 34 speakers (ASD and TD combined), there was moderate agreement between listeners (Kappa = .47). This level of agreement was deemed adequate to support the results in the current study. The inter-judge reliability findings for the Naturalness Perceptual Rating Scale can be found in **Table 9**.

4.4.2 Association Between Group Membership and Naturalness Ratings

The agreed-upon ratings with respect to “natural” or “unnatural” ratings between the two listeners were selected for Chi-Square analysis with the two variables being group membership (ASD vs. TD) and ratings of naturalness (natural vs. unnatural). The results reveal a significant association between group membership and ratings of naturalness. Please see **Table 10**. Among the agreed-upon ratings, the listeners agreed on the following: ASD speakers who sounded “unnatural” (n = 6; 35%); ASD speakers who sounded “natural” (n = 4; 24%); TD speakers who sounded “unnatural” (n = 2; 12%); TD speakers who sounded “natural” (n = 13; 76%). Overall, these findings indicate that a higher proportion of speakers in the ASD group were rated as “unnatural” while a higher proportion of speakers in the TD group were rated as sounding “natural”. The results also suggest that within both groups, some speakers are rated as sounding “natural” or “unnatural”, likely reflecting variations in the human voice

Among the agreed-upon speakers who were rated as sounding “unnatural” by both listeners (i.e., 6 ASD, 35%; 2 TD 12%), mode Likert severity ratings revealed the following: moderate to severe differences in rate, pitch, prosody, and resonance in the ASD group; mild/negligible differences in rate and prosody in the TD group; and moderate differences in pitch and resonance in the TD group. In contrast, the mode Likert severity ratings for speakers who were rated as sounding “natural” by both listeners (i.e., 4 ASD, 24%; 13 TD, 76%) were as follows: mild/negligible differences in pitch and resonance among the speakers with ASD and mild/negligible differences in resonance in speakers who were TD. Therefore, the speakers who were rated as sounding “unnatural” by both listeners exhibited moderate to severe mode Likert severity ratings in at least one descriptor category. Furthermore, the speakers who were rated as sounding “unnatural” by both listeners had perceived differences in at least three descriptor categories. This suggests that it may not be a single descriptor item that contributes to the listeners’ perception of speech that sounds “unnatural”, or different, but rather the pervasiveness and severity of the acoustic features that were perceived by the listeners in the connected speech samples of participants with ASD and TD controls. For example, at least 3 descriptor categories (i.e., differences in pitch, loudness, vocal quality, longer durations within or between syllables, rate of speech, etc.) and/or the severity rating ranging from moderate to severe needed to be present for the listeners’ to rate the connected speech samples as sounding “unnatural” or different, regardless of group membership.

Research Questions

1. Do children with ASD between the ages of 7;10-19;0 years perform with significantly less accuracy on receptive and expressive prosody tasks

according to operational metrics using perceptual and acoustic measures as compared to a TD group matched for age, gender, and language?

- a. H0: Children with ASD between the ages of 7;10-19;0 years do not perform with significantly less accuracy on receptive and expressive prosody tasks according to perceptual and acoustic measures as compared to a TD group matched for age, gender, and language.
- b. H1: Children with ASD between the ages of 7;10-19;0 years do perform with significantly less accuracy on receptive and expressive prosody tasks according to perceptual and acoustic measures as compared to a TD group matched for age, gender, and language.

The investigator rejects the null hypothesis and support the alternative hypothesis. The outcomes of this study support the presence of acoustic and perceptual differences in prosody among those with ASD during elicited prosody tasks as compared to TD peers.

- 2. Is there a significant linear relationship between expressive prosodic abilities and language, motor speech, and pitch processing scores?
 - c. H0: There is not a significant linear relationship between expressive prosodic abilities and language, motor speech, and pitch processing scores.
 - d. H1: There is a significant linear relationship between expressive prosodic abilities and language, motor speech, and pitch processing scores.

The author partially rejects the null hypothesis and accepts the alternative hypothesis. PPVT and group membership account for 63% of the variability in prosody scores (PEPS-C output tasks). Motor speech (AMRs) and EVT-2 scores did not account for a significant amount of the variability in prosody scores above and beyond PPVT-4 scores and group membership. However, a multiple linear regression model that included the variables expressive language (EVT-2) and group membership (ASD vs. TD) in isolation revealed a weak positive relationship between EVT-2 and expressive prosody (PEPS-C output tasks) in the ASD group. A descriptive analysis was conducted to better understand the relationship between expressive prosody and SMRs and pitch processing (FPT). A greater number of participants with ASD exhibited sub-optimal performance on speech motor control (SMR) and pitch processing tasks (59% and 29%, respectively in the ASD group compared to 18% and 6% in the TD group, respectively). Although a statistically significant relationship was not found between prosody and speech motor control and pitch processing in the current study, there appears to be a meaningful relationship between these variables and performance on expressive prosody tasks in the ASD group.

3. Is there an association between group membership (ASD vs. TD) and ratings of naturalness (“natural” vs. “unnatural”) based on two trained listeners’ perceptual ratings of speech, voice, and prosody based on 20-seconds connected speech samples?

- a. H0: There is no association between group membership (ASD vs. TD) and ratings of naturalness (“natural” vs. “unnatural”) based on two trained listeners’ perceptual ratings of speech, voice, and prosody based on 20-seconds connected speech samples.
- b. H1: There is an association between group membership (ASD vs. TD) and ratings of naturalness (“natural” vs. “unnatural”) based on two trained listeners’ perceptual ratings of speech, voice, and prosody based on 20-seconds connected speech samples.

The author rejects the null hypothesis and supports the alternative hypothesis. There was a significant association between group membership and ratings of “naturalness” in the current study.

CHAPTER 5

CONCLUSIONS

5.1 Research Questions, Hypotheses, and Participant Demographics

This investigation addressed three research questions. First, do children with ASD between the ages of 7;10-19;0 years perform with significantly less accuracy on receptive and expressive prosody tasks according to operational metrics using perceptual and acoustic measures as compared to a TD group matched for age, gender, and language? It was predicted that the outcomes of this study would support the presence of acoustic and perceptual differences in prosody among those with ASD during elicited prosody tasks as compared to TD peers. Second, is there a significant linear relationship between expressive prosodic abilities and language, motor speech, and pitch processing scores? It was hypothesized that there would be a significant positive linear relationship between expressive prosody scores and language, speech motor control, and/or pitch processing. Third, is there an association between group membership (ASD vs. TD) and ratings of “naturalness” (“natural” vs. “unnatural”) based on two trained listeners’ perceptual ratings of speech, voice, and prosody based on 20-second connected speech samples? It was hypothesized that there would be an association between group membership and ratings of “naturalness”, as the speech, voice, and prosodic differences among speakers with ASD has been described by human listeners as sounding “different”, “odd”, “exaggerated”, “monotone” or “sing-song” in comparison to their TD peers (e.g., Andrianopoulos et al., 2015; Kanner, 1971; Filipe et al., 2014; Grossman et al., 2013).

A total of 34 participants who met a set of predetermined inclusionary and exclusionary criteria were included in this study. The study consisted of two distinct

groups: seventeen (17) individuals with a formal diagnosis of ASD (Group 1) between the ages of 7;10 and 19;0 years and seventeen (17) typically developing (TD) individuals matched for age, gender, and language (Group 2). The participants with ASD were enrolled in a regular curriculum in school, modified under an Individualized Education Program (IEP). Two first-year graduate students in the SLP master's degree program at the University of Massachusetts Amherst were trained and served as human listeners or judges (Listener 1 and Listener 2).

There were no significant differences between Group 1 (ASD) and Group 2 (TD) in receptive or expressive vocabulary according to standardized scores on the PPVT-4 and the EVT-2. As mentioned previously, two participants with ASD scored “extremely low” on the PPVT-4 and the EVT-2 and one participant with ASD scored “moderately low” on the PPVT-4. The investigator did not eliminate these participants, as it did not have an effect on language ability at the group level. Individuals with autism who have low language abilities are often excluded from research studies examining prosody in autism. This investigator included these participants to better represent the larger, heterogeneous population of individuals with autism with respect to language ability. Furthermore, including individuals with ASD who have below average language allowed the PI to identify if there is a positive relationship between expressive prosody and language.

It is important to mention that some individuals in both groups had comorbid diagnoses. In the ASD group, the following overlapping conditions were reported on the demographic survey during the recruitment process: ADD/ADHD (n = 2); anxiety (n = 4); sensory processing disorder (n = 3); post-traumatic stress disorder (n = 2); mood

disorder NOS (n = 2); oppositional defiant disorder (n = 1); and dysgraphia (n = 1). The co-morbid characteristics among the ASD group are not atypical, as other physical conditions, psychopathologies, emotional behavioral problems, ADD/ADHD, other challenging behaviors, and intellectual disabilities frequently overlap with an autism spectrum disorder. For example, the prevalence rate of ASD and ADHD occurring together is over 50%. Anxiety and depression are two other co-occurring conditions that are reported at high rates in individuals with ASD (Matson & Goldin, 2013). The TD participants were matched to the ASD group on comorbid diagnoses to the extent possible as follows: ADD/ADHD (n = 3) and a mild articulation disorder (n = 1). The two groups were balanced otherwise in terms of age, gender, and sample size.

To address the first research question, seven receptive and expressive prosody tasks with a receptive and expressive component to each were administered using the computerized version of the PEPS-C. The receptive prosody tasks were scored by the PEPS-C software and the expressive prosody tasks were scored perceptually by the PI or Examiner 1 using the PEPS-C scoring procedures. The expressive utterances were recorded for acoustic analysis of duration (seconds) and fundamental frequency/pitch (Hz). The receptive and expressive prosody results will be discussed with respect to percent accuracy on the behavioral tasks (i.e., the PEPS-C), followed by the acoustic characteristics of the audio -recorded expressive Lexical Stress utterances.

5.2 Prosody

Although atypical prosody is not a universal characteristic of ASD, research supports that 33% (Kargas et al., 2016) to 60% (Nadig & Shaw, 2015) of research participants with ASD have differences in receptive or expressive prosodic abilities as

compared to control groups. It has been suggested that individuals with autism have impaired pragmatic and affective prosody while grammatical prosodic abilities may be a relative strength (McCann & Peppé, 2003; Shriberg et al., 2001). Other research has demonstrated that individuals with autism do have impaired grammatical prosody (e.g., Gargan & Andrianopoulos, in review; Paul et al., 2005). We hypothesized that a subgroup of individuals with autism would have impaired pragmatic, affective and/or grammatical prosody, with significant differences being revealed at the group level.

5.2.1 PEPS-C Findings

In support of the first hypothesis, the group with ASD had greater difficulty on receptive and expressive prosody tasks in comparison to TD controls. Overall, the group with ASD scored with significantly lower percent accuracy on the PEPS-C receptive and expressive prosody composite scores. The range of receptive and expressive prosody composite scores in the ASD group was 52-97% and 40-90% accuracy, respectively, while the range of receptive and expressive prosody composite scores in the TD group was 76-99% and 83-98% accuracy, respectively. The wider range of scores, and the proportion of individuals in the group with ASD who scored below competence level, likely contributed to the between group differences.

The group with ASD's mean scores were significantly lower than the TD group's mean scores for 8 out of 14 (57%) of the prosody tasks as follows: Imitation, Turn End Understanding, Lexical Stress Expression, Phrase Stress Understanding, Phrase Stress Expression, Boundary Understanding, Contrastive Stress Understanding, and Contrastive Stress Expression. These findings are consistent with previous literature supporting difficulty with receptive and expressive grammatical and pragmatic prosody among

individuals with autism. For example, published empirical studies have used the same test stimuli from the PEPS-C assessment battery and demonstrated that individuals on the autism spectrum performed with significantly less accuracy when asked to imitate expressive prosodic utterances (e.g., Diehl & Paul, 2012; McCann et al., 2007; Peppe et al., 2007; Peppe et al., 2011; Gargan & Andrianopoulos, in review), receptively distinguish between questions versus statements (e.g., Gargan & Andrianopoulos, in review; Diehl & Paul, 2013), expressively and receptively differentiate nouns versus verbs and compound nouns versus noun phrases (Gargan & Andrianopoulos, in review), understand contrastive stress (Diehl & Paul, 2013; Gargan & Andrianopoulos, in review) and produce contrastive stress (Gargan & Andrianopoulos, in review; McCann et al., 2007; Peppe et al., 2007; Peppe et al., 2011).

There were no significant differences in mean percent accuracy scores between groups on the following prosody tasks in the current study: expressive Turn End; receptive and expressive Affect; receptive and expressive Boundary/Chunking; receptive Lexical Stress; and Auditory Discrimination. These findings are also partially consistent with previous literature. For example, five out of six (83%) previous studies failed to find significant differences between ASD participants and controls on the expressive Turn End task; four out of five (80%) previous studies failed to find significant differences between ASD participants and controls on the expressive Boundary/Chunking task; and three out of five (60%) failed to find significant differences between ASDs and controls on the receptive Boundary/Chunking task.

Some of the current findings are contradictory to previous findings. For example, in contrast to our results, four out of five (80%) previous studies found significant

differences between groups on the expressive Affect task; five out of five (100%) found significant differences between groups on the receptive Affect task; and four out of four (100%) previous studies found significant differences between groups on the Auditory Discrimination task. A comparison of current and previous prosody findings among participants with autism in comparison to control groups without autism using the PEPS-C stimuli can be found in **Table 11**. Overall, the outcomes of this study are not remarkably different with respect to prosodic abilities between participants with ASD and TD controls. It is possible that language ability and other variables such as socioeconomic status, verbal mental age, or chronological age, contributed to some of the differences across studies.

Of note is that the Lexical Stress and Phrase Stress subtests were added to the newer editions of PEPS-C (2015). To date, there are no published studies reporting outcomes on these new subtests to assess lexical stress abilities at the word and phrase level. Therefore, the current findings will be discussed in the context of two previous studies that used comparable experimental stimuli to examine lexical stress in autism during structured experimental tasks.

The group with ASD in the current study had greater difficulty on tasks requiring them to use lexical stress to expressively disambiguate two-syllable noun versus verb word pairs at the word level (i.e. Lexical Stress task). The participants with ASD also had difficulty receptively and expressively differentiating compound nouns versus noun phrases at the sentence level (i.e. Phrase Stress task). The results of this study support differences in the perception and/or production of lexical stress among some individuals on the autism spectrum as compared to controls.

The current finding that some individuals with ASD had significantly more difficulty than controls on the expressive Lexical Stress task, which required them to disambiguate noun and verb word pairs, is consistent with findings reported by Paul et al. (2005). Furthermore, the findings from this study revealed performance that approached statistical significance between groups on the receptive Lexical Stress task ($p = .18$). Similarly, Paul et al., (2005) reported a p-value of .12 for the comparison of receptive lexical stress performance among a group with ASD compared to TD controls in their study. The participants in the current study had a similar mean age as the participants in Paul et al. (2005), which may have contributed to consistent findings across the two studies. In addition, the stimuli for the Lexical Stress task in the current study are comparable to the stimuli used in Paul et al. (2005) with the following exception: the participants in Paul et al. (2005) were provided with sentence level context to support target word meaning and target word production. For this investigation, the stimuli were presented at the single word level with an abstract visual of the stress pattern. Nonetheless, Paul et al. (2005) and the current study found that individuals with ASD had greater difficulty than controls using lexical stress to expressively disambiguate two-syllable noun versus verb word pairs. They also exhibited differences that approached significance on the receptive Lexical Stress task. Taken together, the findings indicate that some individuals with ASD may have difficulty with the perception and production of lexical stress patterns on elicited prosody tasks.

Our findings are contradictory to those reported by Grossman et al. (2010) with respect to significantly less accurate performance among the group with ASD in comparison to controls on the expressive and receptive Phrase Stress tasks. Grossman et

al. (2010) reported that the individuals with HFA in their study performed comparably to the TD controls on an experimental task requiring them to disambiguate compound nouns from noun phrases. Grossman et al. (2010) stated that some TD children in their study might have performed with lower accuracy than expected because they were below the age of competency (i.e., 12 years) for the experimental tasks. According to the authors, this may have lowered the TD group's mean performance and therefore reduced the difference in performance between experimental groups (Grossman et al., 2010). While this may be true, some TD individuals who were below the age of 12;0 in the current study passed the Lexical Stress and Phrase Stress subtests. Therefore, although chronological age may be one variable that contributes to differences in performance within and across groups, it may not fully account for the discrepancies in performance between groups. A stronger explanation for the contradictory findings could be due to differences in the language abilities in the participants with autism in each study. For example, Grossman et al. (2010) reported that the average to above average language abilities in their sample of individuals with HFA may have contributed to accurate performance on the experimental tasks. As mentioned previously, the current study included some individuals with below average language abilities on standardized tests, and some participants who scored in the average range approached 1 SD below the mean (e.g., standard scores of 86, 88 and 89), which may have impacted their performance on the prosody tasks. Another factor that may have contributed to conflicting results across the two studies is that the stimuli in the Phrase Stress task in the current study were different than Grossman et al.'s (2010) stimuli and procedures. Grossman et al. (2010) used picture stimuli. The participants were required to fill in a missing word or phrase

during a sentence completion task while looking at an illustration of the corresponding target utterance, which may have resulted in higher accuracy scores among all participants. In contrast, the participants in the current study did not have a picture stimulus to support comprehension of the target words. This may have contributed to lower performance among some participants with ASD, as they may rely on meaningful visuals to support their success.

To explain between-group differences in performance on the PEPS-C tasks in the current study, the PI examined the mean scores, standard deviation of scores per task, and the range of scores per task. At the group level, the TD group's mean scores were above competence ($> 75\%$ accurate) on all tasks, indicating the tasks were not too difficult for individuals between the ages of 7;11 and 19;0 years of age. The ASD group's mean scores were also at or above competence level on 13 out of 14 (93%) tasks. However, the ASD group had larger standard deviations around the mean on all tasks and a wider range of scores on 13 out of 14 tasks in comparison to the TD group. The within-group variability with respect to percent accuracy on the PEPS-C tasks among the participants with ASD likely contributed to the significant differences between groups. Differences in language ability within the group with ASD may also explain between group differences in performance. The relationship between language and prosody will be discussed in more detail when addressing the second research question.

In summary, the findings across studies that investigated receptive and expressive prosody using structured prosody tasks suggest that at least some individuals with autism have atypical expressive and receptive grammatical, pragmatic, and/or affective prosody. The variables that may contribute to the inconsistent findings across studies include the

sample size, age-range of participants, language ability of the research participants, severity of autism, intellectual ability of participants, socioeconomic status, and/or the presence or absence of comorbid diagnoses.

5.2.2 Prosody: Acoustic Measurements of Pitch and Duration

A prominent finding in the literature is that some individuals with autism are categorically accurate in their expressive prosody, yet they have subtle differences in the acoustic features of their speech. For example, Diehl & Paul (2013) reported that the speakers with ASD in their study accurately differentiated questions versus statements and expressed like versus dislike that was perceived accurately by listeners, but they exhibited longer duration of utterances in comparison to controls. Similarly, Filipe et al. (2014) reported that the speakers with ASD in their study accurately differentiated questions versus statements, but produced utterances that were significantly longer in duration and had greater pitch variability in comparison to the control group. Grossman et al. (2010) did not conduct acoustic analysis of utterances, but reported that the speakers with autism sounded “labored” or “exaggerated” and had long pauses between syllables. Gargan & Andrianopoulos (in review) also reported longer durations of utterances among speakers with ASD in comparison to TD controls on a Lexical Stress and Phrase Stress prosody task. In the current study, the group with ASD exhibited significantly longer durations of utterances in comparison to TD controls. This finding supports previous literature reporting longer durations of utterances among individuals with autism during elicited prosody tasks, despite being categorically accurate in their productions (e.g. Grossman et al., 2010; Filipe et al., 2014; Diehl & Paul, 2013).

The longer durations of utterances among the group with ASD in the current study could be explained by differences in the duration between syllables and/or the duration of each syllable within a word. Some speakers exhibited silent gaps between syllables on some words (e.g., A8, 'DIScount' and 'disCARD'; A4, 'INcrease' and 'disCARD') or lengthened syllables on some words (e.g., A6 and A10, 'DIScount'). The silent gaps between syllables and lengthened syllables are evident when visually inspecting the waveform. While these characteristics may also exist in some speakers in the TD group, it is likely that they exist to a greater extent among speakers in the ASD group, which may have contributed to statistically significant differences in duration of utterances between groups.

No significant differences were found between the ASD and TD groups with respect to mean pitch, minimum pitch, or maximum pitch of Lexical Stress expressive utterances. However, at the group level, the mean pitch was higher among ASD speakers compared to TD controls and the maximum pitch was higher in the ASD group in comparison to TD peers. Although the differences did not reach significance, these patterns are consistent with previous literature demonstrating higher mean pitch and a wider pitch range among speakers with ASD compared to TD controls (Filipe et. al., 2014). A larger sample size and/or a different elicited prosody task (e.g., the Turn End task versus the Lexical Stress task) may have revealed significant differences between groups in mean pitch, minimum pitch, or maximum pitch of expressive utterances.

5.3 Predictors of Expressive Prosody

To address the second research question, the following information was obtained: expressive PEPS-C prosody composite scores; standardized vocabulary scores from the

EVT-2 and PPVT-4; average syllables per second on motor speech AMR tasks; categorical data from the SMR tasks; and percent accuracy on the FPT task.

As previously discussed, six primary hypotheses are reported in the literature to account for atypical prosodic abilities among individuals with autism. Some authors hypothesize that speech differences in ASD could be due to: (1) a lack of one's ability to 'tune up' speech behaviors (Diehl and Paul, 2013; Diehl & Paul, 2012; Shriberg et al., 2011); (2) language deficits (DePape et al., 2012; Peppe et al., 2007; Peppe et al., 2011); (3) social reciprocal interaction impairment (Nakai et al., 2014); (4) the nature of the research tasks (Diehl and Paul, 2013; Peppe et al., 2007); (5) impaired auditory memory (Peppe et al., 2006; Peppe et al., 2007); and (6) motor deficits (Peppe et al., 2007; Diehl & Paul, 2012; Velleman, Andrianopoulos et al., 2009; Andrianopoulos et al., 2015).

Although it has been hypothesized that atypical prosody may be impaired due to level of language ability, speech motor control, enhanced/decreased auditory processing abilities, or a combination of the three, these hypotheses have yet to be directly investigated within the same investigation. Due to the heterogeneity of symptoms, severity, and comorbid conditions in autism, it is reasonable to hypothesize that there are phenotypes or subgroups of individuals with autism who exhibit atypical prosody and comorbid language difficulties, speech motor control differences, and/or auditory processing differences. Statistical and descriptive analysis were administered in this study to identify the predictors of expressive prosodic ability, as well as describe meaningful differences in performance between groups across all variables.

In support of the second hypothesis, receptive vocabulary (PPVT-4) and group membership have a strong, positive relationship with expressive prosody scores in the

group with autism. Receptive vocabulary scores and group membership explained 64% of the variability in prosody scores. It is important to note that while PPVT-4 and group membership were the strongest predictor variables, expressive vocabulary scores on the EVT-2 also revealed a weak positive relationship to prosody in the ASD group.

With respect to speech motor control and auditory processing abilities, patterns in the data were identified through descriptive analysis and are addressed in the next four paragraphs. Although speech motor control based on AMR task performance was not a significant predictor of prosody in the multiple linear regression models, speech motor control differences on AMR and SMR tasks were evident between groups on a quantitative and qualitative level. As such, there appears to be a meaningful difference between groups in speech motor control abilities, which may have been shown to significantly contribute to expressive prosodic abilities using a larger sample size. Regarding auditory processing abilities, it is important to note that the number of participants who completed the FPT task was not large enough to include the data in multiple linear regression analysis. For this reason, the FPT results will only be discussed descriptively.

Three tokens of AMRs for the syllables [p[^]], [t[^]] and [k[^]] and SMRs for the syllable sequence [p[^]t[^]k[^]] were administered to each participant using the instructions outlined by Fletcher (1972). Among the speakers with ASD, five out of 14 (36%) were qualitatively described as having irregular rhythm or rate during AMR productions according to the PI. In contrast, only three TDs (17%) were described as having irregular rhythm, but to a lesser extent than the ASD group. For example, one participant in the TD group was described as having “mild” irregularities in rhythm during some AMR

productions, whereas a participant in the ASD group had “moderate” irregularities in rhythm, such as fluctuating pitch that was consistent throughout all AMR productions. In addition, six participants (35%) in the ASD group were described as having “slurred” AMR productions, whereas zero participants in the TD group sounded “slurred” during the AMR tasks.

Qualitative differences were also noted during the SMR tasks with respect to the type of errors made when producing the [p^] [t^] [k^] syllable sequences across groups. For example, the following errors were documented in the ASD group: (1) switching the sequence of syllables, such as starting with [p^t^k^] and switching to [p^k^t^] or [k^p^t^] (n=4); (2) sound additions, such as adding an extra [k^] or [p^] between trials a few times (n=3); (3) lengthened syllables (n=2); (4) incorrect stress on syllables, such as starting with [p^t^k^] and switching to [p^Tlk^] (n=2); (5) vowel distortions, such as [poo-ti-k^] or [p^tee-k^] (n=3); (6) sounding “effortful” (n=3); (7) “slurred” (n=1); (8) “slow” (n=1); and (9) “choppy”, such as [p^..t^...k^].....[p^..t^..k^].... (n=1). Contrary to these findings, the TD participants had fewer documented errors. Moreover, the PI described the errors in the TD group as “mild” or “typical” errors that were self-corrected. Some of the participants in the TD group were amused by these tasks and their production of them, and self-corrected, demonstrating that TD participants may have been more aware of their errors. The errors in the TD group were described as follows: (1) very minor errors, but self-corrected (n=3); (2) mild incorrect stress (n=1); (3) changed the vowel (n=1); and; (4) minor error once (n=1). The results from the SMR task can be interpreted to suggest that the ASD group exhibited characteristics consistent with a motor speech disorder, whereas errors in the TD group appeared to be mild and “typical” errors,

possibly due to fatigue. Interestingly, among the 7 participants with ASD who failed the expressive PEPS-C composite score, three participants (43%) performed with less than 75% accuracy on average on the SMR tasks. This pattern may have been more prominent in a larger sample of participants.

With respect to auditory processing of pitch, five participants with ASD (50%) did not pass the FPT task in at least one ear. In contrast, only one TD participant failed the FPT in their left ear. The remaining participants in both groups passed the test in both ears based on their ages. Although a higher number of individuals with ASD failed the FPT test in each ear, the results did not reach significance ($p > .05$). Qualitative differences were observed regarding the mode of response on the FPT. For example, five individuals with ASD and one TD participant responded by humming, rather than verbally stating the pitch pattern that they heard. According to Musiek (2002), humming the response accurately, with difficulty verbally stating the pattern, may indicate a problem in the left hemisphere and/or transferring information from the right to left hemisphere. Another interesting observation relates to the high number of reversal responses produced among some participants with ASD (e.g., saying “high high low” (HHL) when the stimulus item was low-low-high (LLH)). Poor performance on the FPT task, such as a high number of reversals, may be indicative of a central pathology (Musiek, 2002). Interestingly, one participant with ASD imitated the correct pitch while saying the wrong words. For example, if he heard the tone pattern HLH, he responded by saying “low-high-low” paired with a high-low-high pitch pattern when he spoke. This participant did not appear to be aware of his errors. For example, he responded at a consistent rate, he did not demonstrate an attempt to self-correct his responses, and he did

not exhibit any other obvious behavioral signs (e.g., frustration) to show that he was having difficulty with the task. The participant's responses appeared to be reversals in his verbal productions, with the accurate ability to hum the response. According to Musiek (2002) left hemisphere lesions or pathologies may contribute to responses of this nature. One can speculate that humming the tones accurately while using incorrect verbal labels reflects difficulty transferring auditory and linguistic information from the right to left hemisphere, central pathology, and/or a combination of the two.

The participants' age and maturity may have contributed to their performance on the FPT in that as age increased in the current study, participants' FPT scores or mode of response also improved. For example, the participants in the ASD group who hummed their responses ranged in age from 7;10 to 12;0 years, and the TD participant who hummed their responses was 8;0 years old. In addition, the one TD participant who failed the FPT was 8;0 years old. These results could suggest that there is a developmental trajectory in performance on the FPT task. Kelly (2002) stated that by approximately age 12;0, however, children should be more accurate in their ability to discriminate frequencies. Therefore, age may not be the only variable contributing to inaccurate performance, as four TD participants and two ASD participants who were below the age of 12 years old passed the FPT in both ears in the current study. In general, the FPT results support the hypothesis stated in the literature that although enhanced pitch-processing abilities may be present in some individuals with autism, it is not a universal characteristic among those with ASD (Mayer et al., 2016). Among the individuals with ASD who failed the expressive PEPS-C composite score, only one participant failed the

FPT. This suggests that there may not be a positive relationship between this variable and prosody.

In summary, the current findings provide support for the hypothesis that differences in prosodic performance are related to level of language ability (e.g., DePape et al., 2012). While the current study sheds insight into the speech motor control and auditory processing differences between groups, a study with a larger sample size may be needed to reveal a significant positive relationship between speech motor control and auditory processing of pitch as predictors of expressive prosody. The current study did not incorporate SMR data in the multiple linear regression models and therefore it is currently unknown if SMR performance predicts expressive prosody. Nonetheless, there does appear to be a meaningful relationship between SMR performance and expressive prosody in this sample of participants.

5.4 Speech, Voice, and Prosodic Naturalness

To address the third research question, individual trained listener responses from Part 1 of the Naturalness Perceptual Rating Scale (Gargan & Andrianopoulos, in review) are discussed descriptively. In addition, Chi-Square analysis was completed for Part 2 of the rating scale responses, using the agreed-upon naturalness ratings between the two listeners to determine if group membership and naturalness ratings are associated.

Some individuals with ASD have been described in the literature as sounding “monotonic”, “machine-like”, “sing-song”, “awkward”, “odd”, “labored”, and “different” (Andrianopoulos et al., 2015; Filipe et al., 2014; Grossman, 2015; Grossman et al., 2013; Kanner, 1971; Shriberg et al., 2001). Human listeners can perceive speech, voice and prosodic differences under controlled conditions that distinguish children with ASD from

typically developing (TD) peers during the production of oral narratives (Andrianopoulos et al., 2015; Dahlgren, 2018; Redford, Kapatsinski, & Cornell-Fabiano, 2018). Similarly, during elicited grammatical prosody tasks, Filipe et al. (2014) reported that human listeners rated the speech of children with Asperger syndrome as sounding “odd” when producing statements versus questions, yet the speakers were rated as categorically accurate in their productions. Grossman et al. (2010) reported that the individuals with HFA sounded “slow” and “labored”, despite high accuracy on lexical stress tasks. Shriberg et al. (2001) reported that more than 50% of the speakers with autism in their study demonstrated inappropriate production of stress and 40% of the individuals with autism were described as sounding hypernasal. Based on these findings, it was hypothesized that the speakers with ASD would be described by trained listeners as having moderate to severe differences in their speech, voice, and prosody in at least one category (speech sound errors, rate, pitch, prosody, resonance, or overall “naturalness”) based on a 20-second audio-recorded connected speech sample. In addition, it was predicted that a higher proportion of speakers with autism would be described as having differences in their prosody, pitch, rate of speech, and resonance in comparison to controls. Lastly, it was also hypothesized that there would be an association between overall naturalness ratings and group membership.

Among the agreed-upon speakers who were rated as sounding “unnatural” by both listeners, mode Likert severity ratings reveal the following: moderate to severe differences in rate, pitch, prosody, and resonance in the ASD group; mild/negligible differences in rate and prosody in the TD group; and moderate differences in pitch and resonance in the TD group. In contrast, the mode Likert severity ratings for speakers who

were rated as sounding “natural” by both listeners were as follows: mild/negligible differences in pitch and resonance among the speakers with ASD and mild/negligible differences in resonance in speakers who were TD. Therefore, the speakers who were rated as sounding “unnatural” by both listeners exhibited moderate to severe mode Likert severity ratings in at least one descriptor category. Furthermore, the speakers who were rated as sounding “unnatural” by both listeners had perceived differences in at least three descriptor categories. This suggests that it may not be a single descriptor item that contributes to sounding “unnatural”, but rather moderate to severe differences in at least 3 descriptor categories, regardless of group membership.

To determine the proportion of speakers who were characterized as having differences in their prosody, pitch, rate of speech, and resonance in comparison to controls, the PI took a frequency count for how many speakers received a mild to severe rating by both listeners per category (i.e., the agreed-upon ratings per category) and divided the total count by the number of speakers in each group (i.e., $n = 17$). More individuals with ASD were perceptually described as having differences in their rate of speech, pitch, prosody, and resonance by both raters in comparison to controls. The two listeners agreed that 59% of speakers with ASD had “odd” or “atypical” prosody; 47% had differences in their pitch; 35% had differences in their rate of speech; and 24% had differences in resonance. In contrast, only 12-24% of speakers in the TD group were rated by both listeners as having differences in their pitch, rate of speech, and/or prosody (Please see **Table 12**). The results indicate that some speakers were rated as sounding “natural” or “unnatural” within both groups, likely reflecting some typical variations in

the human voice. The results also revealed a significant association between group membership and ratings of overall naturalness.

5.5 Summary

In summary, the current findings support that at least some individuals with autism perform with significantly less accuracy on structured expressive and/or receptive grammatical and pragmatics tasks in comparison to TD controls matched for age, gender, and language. In support of previous findings, the participants with ASD in the current study had significantly longer utterance durations.. Although they also exhibited wider average pitch ranges than the TD comparison group, the results did not reach significance.

Consistent with the hypothesis that language ability plays a role in prosodic functioning (e.g., DePape et al., 2012), standardized receptive vocabulary scores (PPVT-4) and group membership were significant predictors of expressive prosody in the current study. There also appear to be meaningful differences within and across groups with respect to performance across expressive vocabulary, speech motor control, and auditory processing variables. For example, the following sub-groups were identified: two individuals with ASD performed sub-optimally on 4+ variables; 5 individuals with ASD performed sub-optimally on 3+ variables; 3 individuals with ASD performed sub-optimally on 2 variables; 4 individuals with ASD performed sub-optimally on 1 variable; and only 1 participant with ASD performed well on all variables. It is worth mentioning that the two individuals with ASD who performed sub-optimally on 4+ variables had below average PPVT-4 and EVT-2 scores. Overall, 59% of participants with ASD had difficulty on SMR tasks and were rated as sounding “unnatural” by both listeners; 41%

failed the expressive PEPS-C composite score; 29% failed the FPT; 24% failed the receptive PEPS-C composite scores, and 12-18% had expressive and receptive vocabulary scores that were more than 1 SD below the mean. This demonstrates a spectrum of abilities on the variables examined in the current sample of participants with ASD. In contrast, the following was observed in the TD comparison group: 5 TD participants performed sub-optimally on 1 variable; 1 TD participant performed sub-optimally on 2 variables; and 10 TD participants performed well on all variables. Overall, only 18% of TD participants had difficulty on SMR tasks, 12% were rated as sounding “unnatural” by both listeners; and 6% had expressive vocabulary scores more than 1 SD below the mean or failed the FPT in at least 1 ear.

Lastly, there was a significant association between group membership and ratings of “naturalness”. It is hypothesized that there may not be one single descriptor item that contributes to sounding “unnatural”. Rather, moderate to severe differences in at least 3 descriptor categories could result in sounding “unnatural” to trained listeners, regardless of group membership.

5.6 Limitations and Future Directions

There are several limitations to this study, which can be addressed in future research. First, duration (seconds) and pitch (Hz) were only examined for one PEPS-C subtest (i.e., Lexical Stress). It is recommended that future studies examine these same acoustic variables across more than one subtest in the same study, to determine if there are differences in duration and pitch of utterances regardless of prosodic function. In addition, conducting acoustic analysis for other subtests, such as the Turn End subtest, will allow for comparison of results to previous studies. For example, the current study

and Filipe et al. (2014) both reported that there were no significant differences between groups with respect to percent accuracy on the expressive Turn End subtests. However, Filipe et al. (2014) reported that there were significant differences between groups with respect to duration and pitch of utterances, paired with perceptual ratings of sounding “odd” to human listeners. Therefore, it is possible that on the PEPS-C subtests that failed to find statistical differences in performance between groups in terms of percent accuracy, there may be differences on an acoustic level that contribute to sounding “odd” to listeners. Second, linear regression analysis including a larger sample of participants, SMR data, and FPT scores may reveal a positive relationship between these variables and expressive prosody. Third, the current study only examined the relationship between language, speech motor control, pitch processing, and expressive prosody. We did not examine the relationship between these variables and receptive prosodic abilities. Fourth, the Naturalness Perceptual Rating tool did not include a category to assess vocal quality on its own. Instead, vocal quality was merged with the resonance category. Fifth, future studies should include a larger number of listeners to allow for more advanced statistics and examine if trained listeners are more likely to identify speech, voice, and prosody as sounding “unnatural” among speakers with ASD in comparison to the speech, voice, and prosody of TD peers. Sixth, although the audio-recorded speech samples were randomized in the current study in terms of group membership, it is recommended that more than one randomized CD is used with a larger number of listeners to control for the potential order effect of the speech samples. Lastly, future studies could compare listener ratings of speech, voice, and prosody using the Naturalness Perceptual Rating Scale across elicited prosody tasks, spontaneous speech, narrative tasks, and picture description

tasks among participants with autism, those who are TD, and other clinical groups (e.g., Childhood Apraxia of Speech).

Table 1. Demographic Characteristics of the Participants

Characteristic	ASD	TD
Number of Participants	17	17
Gender		
Male	12 (71%)	11 (65%)
Female	5 (29%)	6 (35%)
Age Range (years; months)	7;10-19;0	8;1-17;11
Mean Age (Months)	160.70	160.70
Mean (SD) EVT-2	103(22)	108(14)
Mean (SD) PPVT-4	103(20)	114(11)
Co-Morbid Diagnosis		
ADD/ADHD	2	3
Anxiety	4	0
Sensory Processing Disorder	3	0
Post-Traumatic Stress Disorder	2	0
Mood Disorder – NOS	2	0
Operational Defiant Disorder	1	0
Dysgraphia	1	0
Mild Articulation Disorder	0	1

Note: ASD= Autism Spectrum Disorder; TD = typically developing; ADHD = Attention Deficit/Hyperactivity Disorder; ADD = Attention Deficit Disorder; EVT-2 = Expressive Vocabulary Test, Second Edition (Williams, 2007); PPVT-4 = Peabody Picture Vocabulary Test, Fourth Edition (Dunn & Dunn, 2007)

Table 2. PEPS-C 2015 Results

Group		ASD Group (N=17)		TD Group (N=17)	
PEPS-C Tasks	t	M (SD)	Range %	M (SD)	Range %
Discrimination	-1.6921	92.05 (16.10)	50 - 100	90.82 (14.02)	56 - 100
Imitation	-2.5285*	87.23 (17.39)	38 - 100	98.05 (2.98)	91 - 100
Turn End	-2.9601**	84.70 (19.68)	50 - 100	98.94 (2.35)	94 - 100
Understanding					
Turn End Expression	-2.0039	84.64 (20.35)	38 - 100	95 (6.28)	75 - 100
Affect Understanding	-.92255	90.58 (12.02)	63 - 100	93.88 (8.48)	69 - 100
Affect Expression	-1.5044	75.29 (18.13)	36 - 100	84.47 (17.42)	38 - 100
Lexical Stress	-1.3567	69.17 (18.18)	38 - 94	76.94 (15.03)	50 - 100
Understanding					
Lexical Stress	-2.9572**	73 (19.54)	38 - 94	89.11 (11.08)	68 - 100
Expression					
Phrase Stress	-2.6565*	75.52 (15.97)	50 - 100	89.52 (14.72)	44 - 100
Understanding					
Phrase Stress	-4.1579***	75.72 (13.17)	50 - 100	90.64 (6.72)	75 - 100
Expression					
Boundary	-2.7028*	84.37 (18.90)	43 - 100	97.52 (4.77)	88 - 100
Understanding					
Boundary Expression	-1.0117	84.25 (17.95)	44 - 100	89.41 (9.99)	68 - 100
Contrastive Stress	-2.3786*	79.35 (25.70)	16 - 100	94.64 (6.50)	75 - 100
Understanding					
Contrastive Stress	-2.9227**	75.38 (29.36)	0 - 100	96.47 (4.77)	88 - 100
Expression					

*p <.05; ** p <.01; *** p <.001

Table 3. Inter-Judge Reliability for the PEPS-C 2015 Expressive Prosody Tasks

Reliability Measures	K	Interpretation (Landis & Koch, 1977)
Kappa between Examiners 1 & PI		
T2_Lexical Stress	.718	<i>Substantial</i>
T2_Phrase Stress	.812	<i>Substantial</i>
T2_Affective	1	<i>Almost Perfect</i>
T2_Contrastive	1	<i>Almost Perfect</i>
A9_Lexical Stress	.625	<i>Substantial</i>
A9_Phrase Stress	.906	<i>Almost Perfect</i>
A9_Affective	.906	<i>Almost Perfect</i>
A9_Contrastive	1	<i>Almost Perfect</i>
A3_Lexical Stress	1	<i>Almost Perfect</i>
A3_Phrase Stress	1	<i>Almost Perfect</i>
A3_Affective	.718	<i>Substantial</i>
A3_Contrastive	1	<i>Almost Perfect</i>
A6_Lexical Stress	1	<i>Almost Perfect</i>
A6_Phrase Stress	.43	<i>Moderate</i>
A6_Affective	1	<i>Almost Perfect</i>
A6_Contrastive	1	<i>Almost Perfect</i>

Table 4. Intra-Judge Reliability for the PEPS-C 2015 Expressive Prosody Tasks

Reliability Measures	K	Interpretation (Landis & Koch, 1977)
Kappa between PI original score and PI 4 months later		
T2_Lexical Stress	1	<i>Almost Perfect</i>
T2_Phrase Stress	1	<i>Almost Perfect</i>
T2_Affective	1	<i>Almost Perfect</i>
T2_Contrastive	.71	<i>Substantial</i>
A1_Lexical Stress	.93	<i>Almost Perfect</i>
A1_Phrase Stress	1	<i>Almost Perfect</i>
A1_Affective	.71	<i>Substantial</i>
A1_Contrastive	1	<i>Almost Perfect</i>
A2_Lexical Stress	.71	<i>Substantial</i>
A2_Phrase Stress	1	<i>Almost Perfect</i>
A2_Affective	.43	<i>Moderate</i>
A2_Contrastive	.71	<i>Substantial</i>
T13_Lexical Stress	.9	<i>Almost Perfect</i>
T13_Phrase Stress	1	<i>Almost Perfect</i>

Table 5. Acoustic Measurements of Pitch and Duration

Parameters	M	SD	Range
Duration (seconds)*			
ASD	1.01	.19	.76 – 1.31
TD	.87	.19	.61 – 1.49
Mean Pitch			
ASD	181.95	43.55	94.67 – 236.38
TD	174.79	35.28	109.12 – 221.43
Minimum Pitch			
ASD	139.15	32.36	74.99 – 185.04
TD	131.18	26.96	84.96 – 184.05
Maximum Pitch			
ASD	244.60	59.11	135.85 – 306.42
TD	242.08	54.88	148.05 – 321.4

*p < 0.05

Table 6. Frequency Pattern Test Right Ear Results, separated by group and rating

Group	“Pass”	“Fail”
ASD	n = 6	n = 4
TD	n = 10	n = 0

p = .08

Table 7. Frequency Pattern Test Left Ear Results, separated by group and rating

Group	“Pass”	“Fail”
ASD	n = 5	n = 5
TD	n = 9	n = 1

p = .1409

Table 8. Frequency Pattern Test Results

	Right		Left		
	%Accuracy	Pass/Fail	% Accuracy	Pass/Fail	Notes
A1	57%	Fail	30%	Fail	Correct pitch; incorrect
A2	53%	Fail	69%	Fail	Reversals
A3	100%	Pass	100%	Pass	Gestures with speech
A4	33%	Fail	40%	Fail	Hummed
A5	87%	Pass	53%	Fail	Seemed to lose interest
A6	73%	Fail	30%	Fail	Hummed
A7	100%	Pass	100%	Pass	Hummed
A8	100%	Pass	100%	Pass	Hummed
A9	100%	Pass	100%	Pass	
A10	93%	Pass	100%	Pass	Hummed
T1	93%	Pass	93%	Pass	
T2	70%	Pass	70%	Pass	
T3	100%	Pass	93%	Pass	
T4	77%	Pass	93%	Pass	
T6	100%	Pass	100%	Pass	Hummed
T7	93%	Pass	93%	Pass	
T11	100%	Pass	100%	Pass	
T12	93%	Pass	93%	Pass	
T13	100%	Pass	100%	Pass	
T15	83%	Pass	30%	Fail	

Table 9. Inter-Judge Reliability for the Naturalness Perceptual Rating Scale

Reliability Measures	K	Interpretation (Landis & Koch, 1977)
Kappa between Listeners 1 & 2		
ASD Speakers	.16	Slight
TD Speakers	.76	Substantial
ASD & TD Speakers	.47	Moderate

Table 10. Chi-Square Analysis Results for the Naturalness Perceptual Rating Scale

	Unnatural	Natural
ASD	6	4
TD	2	13

Table 11. Systematic Literature Review of PEPS-C Results

	Diehl, J. & Paul, R. (2013)		Diehl, J.J., & Paul, R. (2012)		Peppe et al., 2011		Peppe et al., 2007		McCann et al., 2007		Filipe et al., 2012		Gargan & Andrianopoulos (in review)		Gargan & Andrianopoulos (in progress)	
PEPS-C Tasks	R	E	R	E	R	E	R	E	R	E	R	E	R	E	R	E
Turn end (R vs. E)		Long durations. Categorically accurate					X	X	X	X	X	Categorically accurate, long durations.		X		X
Affect (R vs. E)		Long durations. Categorically accurate.													X	X
Boundary/Chunking (R vs. E)	X	X					X	X	X	X				X		X
Focus / Contrastive Stress (R vs. E)		higher f0, higher variance of f0, & higher intensity					X		X							
Lexical Stress (R vs. E)														+ long durations	X	+ long durations
Phrase Stress (R vs. E)																
Aud. Discrimination (R)															X	
Imitation (E)																

Red = significant difference between groups;

Light red = acoustic differences only, but categorically accurate productions;

Pink with white text = significantly different on PEPS-C AND specified acoustic feature

Black X = examined but not significant

Table 12. Descriptive Analysis for the Naturalness Perceptual Rating Scale

	ASD	TD
Rate (too fast/too slow)	35%	12%
Pitch (monotone)	35%	24%
Pitch (too high/too low)	12%	0%
Pitch overall	47%	24%
Prosody (odd/atypical, sing-song)	59%	24%
Resonance (hyper/hypo-nasal)	24%	0%

Figure 1. Lexical Stress and Duration Boxplot

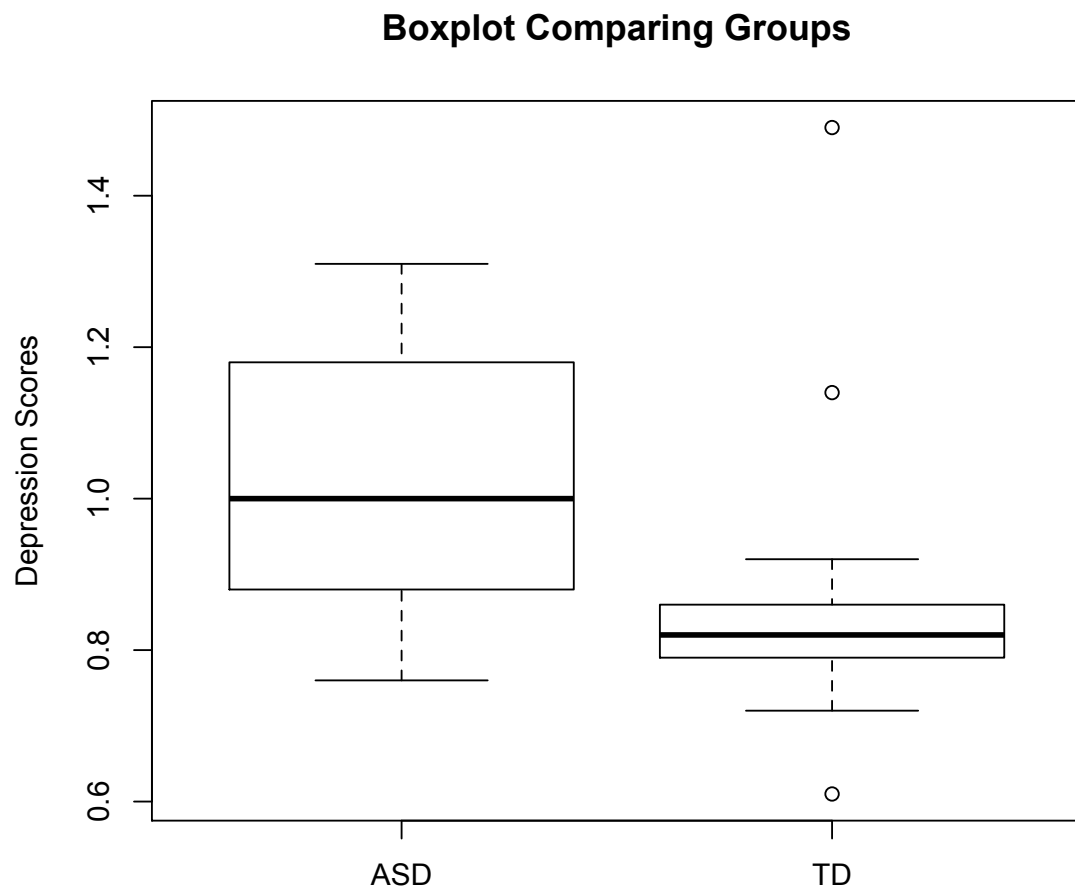
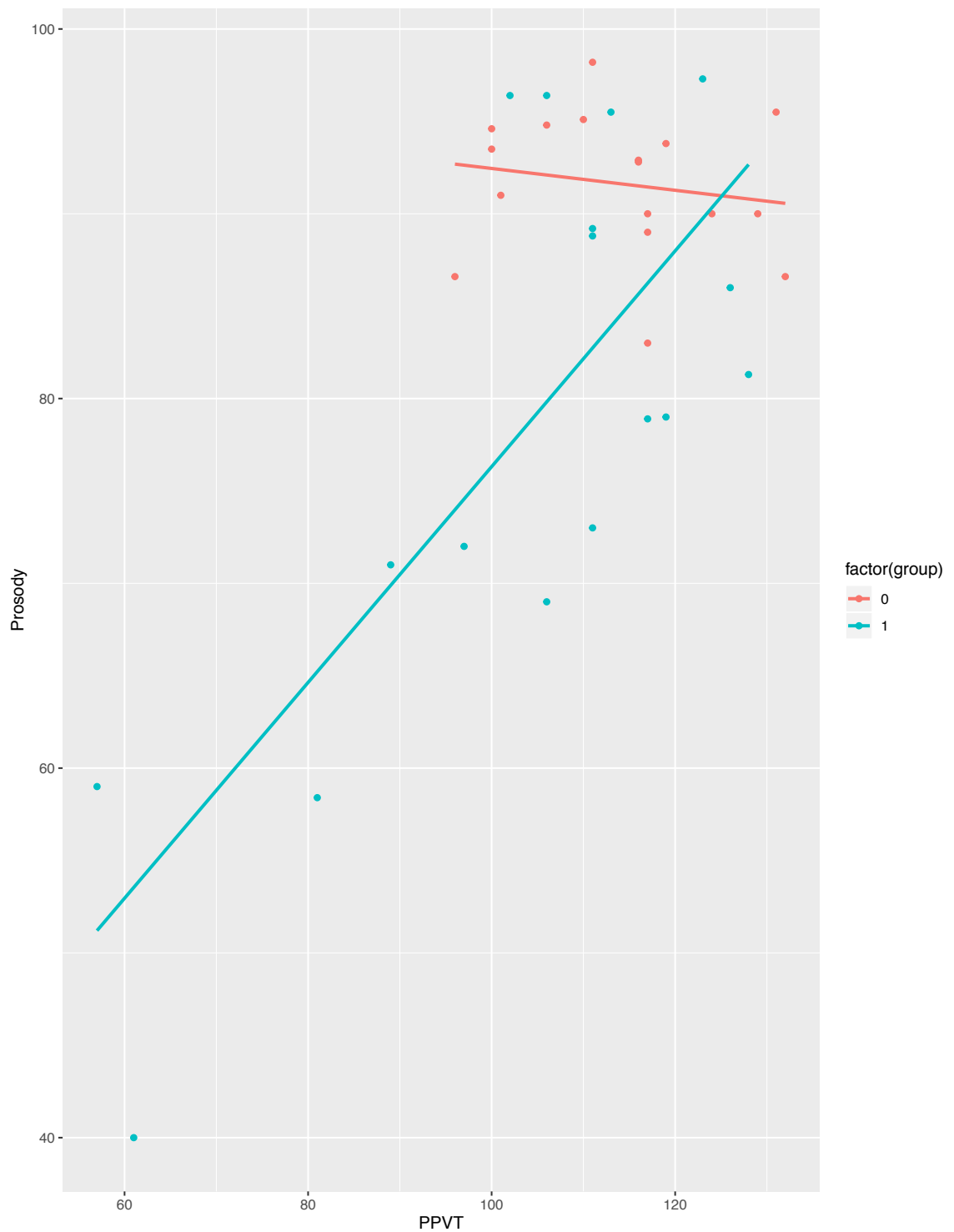


Figure 2. Interaction of Group and PPVT Scores on Expressive Prosody



Note: Individuals with Autism Spectrum Disorder = 0; Individuals who were Typically Developing = 1

APPENDIX A
DEMOGRAPHIC SCREENING QUESTIONS

The parent/guardian will be asked these screening questions when they contact investigators about the study only AFTER they provide the investigators verbal consent (over the phone) to answer the screening questions.

Investigator will say:

"Before you officially enroll in this research study, I will be asking you to a few questions. It should take you no more than 10 minutes to complete. This questionnaire is a screening tool that will ask you to determine your child's eligibility for participation in the study. If your child is determined ineligible to participate, the completed questionnaire will be destroyed. If your child is determined eligible to participate, the completed questionnaire will become part of the study materials, and we will protect the information as confidential and safeguard it from unauthorized disclosure. Only research personnel will have access to this information. Do you have any questions?" YES or NO.

Consent Obtained? YES or NO

I, _____, hereby give consent for Dr. Andrianopoulos and/or Colleen Gargan (your name, printed) or their research staff to ask me these screening questions about my child's potential participation in this study.

Name:

Date of Birth:

Your town/city, state, zipcode:

Your Phone Number:

School:

To participate in this study the individual must meet the following criteria:

DIAGNOSTIC HISTORY:

- a) Formally diagnosed with an Autism Spectrum Disorder by a specialist. **YES or NO**
 - a. If yes, please **bold** reported specialist: pediatrician, autism specialist, psychologist, neurologist)
 - b. If yes, would you be willing to show documentation of a diagnosis on the day your child participates in the study? The researcher will not keep this document. The researcher will just look at the document to confirm that a formal diagnosis was given. This will be informative, since the researcher is not administering standardized tests to confirm a diagnosis of autism. If you choose not to share the formal document or you do not have access to it, that is okay. **YES or NO**
- b) At what age was your child diagnosed with ASD?
- c) What type of interventions have they had?

GENERAL CHARACTERISTICS:

- d) Between the ages of 8;0-13;11 years? **YES or NO**
- e) What gender do they identify as?
- f) Are they a monolingual speaker of English? **YES or NO**
- g) Do they live in a monolingual, English speaking, household? **YES or NO**
- h) Have no history within the past 6 months of self-injury or injury to others or damage to property? **YES or NO**
- i) Vocalize (communicate at least 10 words orally and at least 50 words using some other form of communication, such as PECS, ASL, AT/AAC, etc.)? **YES or NO**
- j) Hearing or visual impairments? **YES or NO**
- k) Cranial-facial malformations (e.g., cleft palate or cleft lip)? **YES or NO**
- l) Other co-morbid neurological, medical, or behavioral problems? **YES or NO**
 - a. If yes, please list co-morbid diagnoses:
 - i. _____
 - ii. _____

ACADEMICS:

- m) Are they enrolled in the general curriculum at school? (even if modified under an IEP): **YES or NO**
- n) Do they read at current grade level? **YES or NO**
 - a. If no, please indicate what grade level they read at:
 - i. _____
- o) What is their IQ (if applicable)?

- a. Would you be willing to bring documentation of IQ testing and results on the day of the experiment to let the researcher look at and document an IQ score?

If you choose not to share this information, that is okay. **YES or NO**

Dr. Mary Andrianopoulos

Department of Communication Disorders

358 North Pleasant Street

University of Massachusetts, Amherst, MA 01003

Email: mva@comdis.umass.edu

Voice: 413-545-0551

APPENDIX B

SPEECH MOTOR CONTROL TASK

The following instructions were obtained directly from Fletcher (1972). Diadochokinetic rates will be obtained for each participant on the following sounds:

/puh/, /tuh/, /kuh/, /puh-tuh-kuh/.

Each subject will be seated and the following instructions will be given:

I want you to say some sounds for me. They aren't words, just sounds. I'll show you how to make it first, then you can say it with me. Then you try it yourself as fast as you can. The first sound is ... [for example, /puh, puh, puh . . . puh/].
NOW try it with me-[first practice trial of approximately three seconds] OK,
that's the way.

Now do it by yourself, as fast as you can . . . [second practice trial of approximately three seconds]. Good.

Now I want you to do it once more. This time it has to be a long one. I'll tell you when to start. Don't stop until I tell you. Ready. Start. [Repetitions counted in this third trial.]

The next sound is _____.

A similar pattern of instructions will be followed for each sound (Fletcher, 1972).

APPENDIX C

SPONTANEOUS SPEECH SAMPLE

Participants will engage in two, 3-minute spontaneous speech samples with the examiner.

The recording with sufficient spontaneous language (i.e., multiple segments of at least 3 second utterances / connected speech per sample) will be selected for analysis.

1. The examiner will be instructed to engage the participant in a conversation about school, hobbies, or a favorite game, book, or activity. Sample questions and prompt items will be generated for the examiner to use. The prompt questions will ensure that the spontaneous speech samples are controlled for content. A list of possible questions or prompts is provided below:

- Tell me about your favorite game.
- What is the goal of the game?
- Who is your favorite character? Why? Can you describe them?
- What is your favorite book? What was it about?
- What is your favorite thing to do when you're not at school?

2. ADOS-2 Cartoon Task & Picture Description Task (Lord, Luyster, Gotham, & Guthrie, 2012).

APPENDIX D

NATURALNESS PERCEPTUAL RATING SCALE

Participant #: _____ Age: _____ Gender: _____

Instructions: After listening to a 20-second connected speech sample per participant, please consider your opinion of the speaker's speech, voice and prosody. You may listen to the recording 2x each. Please complete all ratings at the same time.

For each descriptor, please indicate the degree of severity:

Descriptor	Severity of Descriptor Items			
	0 = Not present	1 = Mild "Negligible"	2 = Moderate	3 = Severe
SPEECH SOUND ERRORS				
/r/ /s/ Other:				
RATE OF SPEECH				
Too Fast				
Too Slow/Long Durations				
PITCH				
Too High				
Too Low				
Monotone				
PROSODY				
"Odd"				
"Atypical"				
"Sing-Song"				
RESONANCE				
Hypo/Hyper nasal				
Hoarse				
Strain				

Please indicate your overall opinion of this speaker's speech, voice, and prosody:

Category	0	1
Rating of "Naturalness"	<i>The speaker sounds natural – appropriate and expected expression with differences in the voice that are negligible</i>	<i>The speaker sounds unnatural – consistently sounds "different" or "socially awkward"</i>

(Gargan & Andrianopoulos; adapted from Andrianopoulos et al., 2015)

APPENDIX E

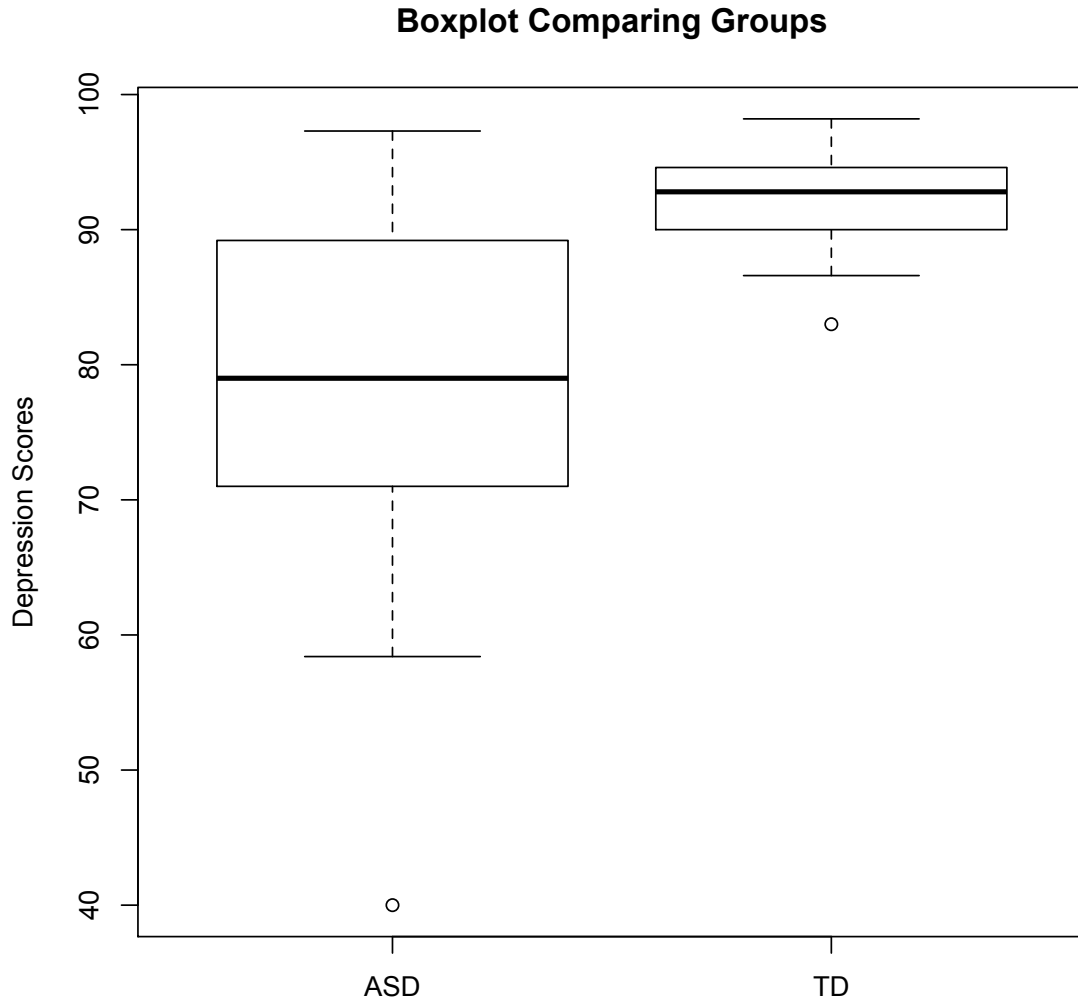
SUMMARY OF RESULTS WITHIN AND BETWEEN GROUPS

Participant	EVT	PPVT	R. PEPS-C	E. PEPS-C	FTP	SMRs	"Unnatural"
A1					X	X	X
A2					X		X
A3						X	X
A4	X	X	X	X	X		X
A5					X		
A6					X	X	X
A7						X	
A8		X	X	X			
A9						X	
A10						X	
ASD 1	X	X	X	X			X
ASD 5						X	X
ASD 7				X		X	X
ASD 12							
ASD 13				X			
ASD 14				X		X	X
ASD 16			X	X		X	
T1						X	
T2							
T3							
T4							
T6						X	
T7							X
T11						X	
T12							
T13							
T15					X		X
TD 2							
TD 3							
TD 4							
TD 6	X						
TD 9							
TD 10							
TD 11							

Note: An 'X' was assigned if: EVT & PPVT >1 standard deviation below the mean; <75% on the PEPS-C composite score; failed the Frequency Pattern Test in at least one ear; <75% accurate on Sequential Motor Rates; sounded "unnatural" to both listeners.

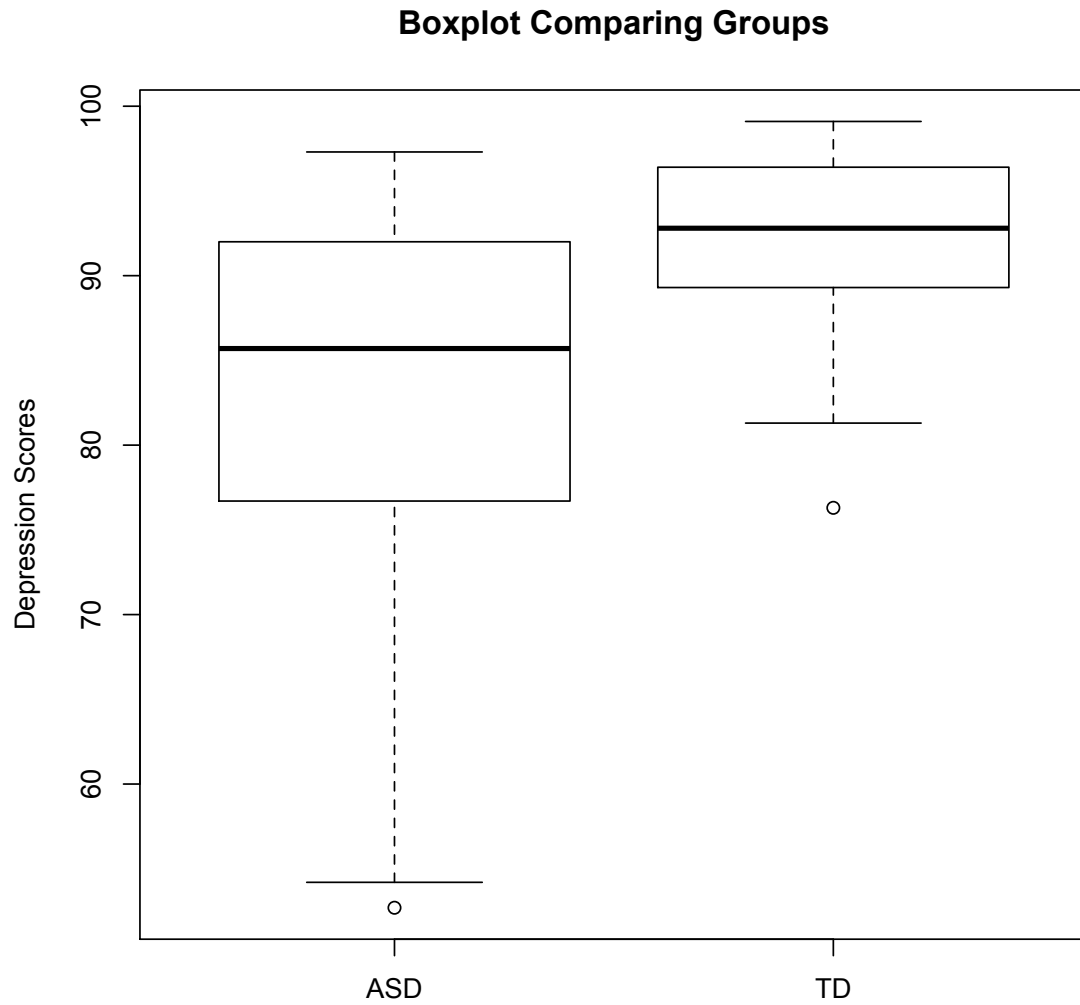
APPENDIX F

BOXPLOT COMPARING PEPS-C EXPRESSIVE PROSODY COMPOSITE SCORES BETWEEN GROUPS



APPENDIX G

BOXPLOT COMPARING PEPS-C RECEPTIVE PROSODY COMPOSITE SCORES BETWEEN GROUPS



BIBLIOGRAPHY

- Adams, L. (1998). Oral-motor and motor-speech characteristics of children with autism. *Focus on Autism and Other Developmental Disabilities, 13*(2), 108-112.
- Andrianopoulos, M.V., Zaretsky, E., McGuigan, C., Warshaw, R. (2015). Perceptual characteristics of spontaneous speech in autism: are we discriminating listeners? Proceedings of the 39th Boston University Conference on Language Development. Cascadilla Press, Boston, MA.
- American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author.
- American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders: DSM-5. Washington, D.C: American Psychiatric Association.
- Baio J, Wiggins L, Chistensen DL, Maenner, M.J., Daniels, J., Warren, Z., et al. Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years – Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014. *MMWR Surveillance Summary* 2018;67(No. SS-6):1-23. DOI <http://dx.doi.org/10.15585/mmwr.ss6706a1>
- Baron-Cohen, S., Leslie, A., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition, 21*(1), 37-46.
- Bellis, T.J. (2003). *Assessment and management of central auditory processing disorders in the educational setting: From science to practice*. Clifton Park, NY: Thomson/Delmar Learning.
- Bonneh, Y., 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*(237), 1-7.
- Bonnel, A., Mottron, L., Peretz, I., Trudel, M., Gallun, E., Bonnel, A.M. (2003). Enhanced pitch sensitivity in individuals with autism: a signal detection analysis. *Journal of Cognitive Neuroscience, 15*(2), 226-235.
- Boucher, M.J. (2013). “Evaluation of Motor Speech and Intervention Planning for Children with Autism”. *Open Access Dissertations. 729*.
https://scholarworks.umass.edu/open_access_dissertations/729
- Campisi, P., Tewkif, T., Manoukian, J., Schloss, M., Pelland-Blais, E., Sadeghi, N. (2002). Computer-assisted voice analysis: establishing a pediatric database. *Archives of Otolaryngology Head and Neck Surgery, 128*(2), 156-160.

- Chenausky, K., Brignell, A., Morgan, A., & Tager-Flusberg, H. (2019). Motor speech impairment predicts expressive language in minimally verbal, but not low verbal, individuals with autism spectrum disorder. *Autism and Developmental Language Impairments*, 4, 1-12. <https://doi.org/10.1177/2396941519856333>
- Chenausky, K., Nelson, C. & Tager-Flusberg, H. (2017). Vocalization rate and consonant production in toddlers at high and low risk for autism. *Journal of Speech, Language and Hearing Research*, 60, 866-876.
- Chevallier, C., Noveck, I., Happe, F., & Wilson, D. (2011). What's in a voice? Prosody as a test case for the theory of mind account of autism. *Neuropsychologia*, 49, 507-517. doi:10.1016/j.neuropsychologia.2010.11.042
- Chowdhury, R., Sharda, M., Foster, N., Germain, E., Tryfon, A., Doyle-Thomas, K. et al. (2017). Auditory pitch perception in autism spectrum disorder is associated with nonverbal abilities. *Perception*, 46(11), 1298-1320.
- Chiew, J. & Kjelgaard, M. (2017). The perception of affective prosody in children with autism spectrum disorders and typical peers. *Clinical Archives of Communication Disorders*, 2(2), 128-141.
- Cohen J. (1988). Statistical Power Analysis for the Behavioral Sciences. New York, NY: Routledge Academic
- Colich, N. L., Wang, A., Rudie, J. D., Hernandez, L. M., Bookheimer, S. Y., & Dapretto, M. (2012). Atypical neural processing of ironic and sincere remarks in children and Adolescents with Autism Spectrum Disorders. *Metaphor and Symbol*, 27(1), 70-92.
- Courchesne, E., Redcay, E., & Kennedy, D. (2004). The autistic brain: birth through adulthood. *Current Opinion in Neurology*, 17(4), 489-496.
- Delacre, M., Lakens, D., & Leys, C. (2017). Why Psychologists Should by Default Use Welch's *t*-test Instead of Student's *t*-test. *International Review of Social Psychology*, 30(1), 92–101. DOI: <http://doi.org/10.5334/irsp.82>
- DePape, A., Chen, A., Hall, G., Trainor, L. (2012). Use of prosody and information structure in high functioning adults with Autism in relation to language ability. *Frontiers in Psychology*, 3(72), 1-13.
- Devadiga, D., & Bhat, J. (2012). Oral diadokochinetic rate – An insight into speech motor control. *International Journal of Advanced Speech and Hearing Research*, 1(1), 10-14.

- Diehl, J. & Paul, R. (2012). Acoustic differences in the imitation of prosodic patterns in children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 6, 123-134.
- Diehl, J., & Paul, R. (2013). Acoustic and perceptual measurements of prosody production on the profiling elements of prosodic systems in children by children with autism spectrum disorders. *Applied Psycholinguistics*, 34(1), 135-161. DOI: <https://doi.org/10.1017/S0142716411000646>
- 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.
- Dunn, L., & Dunn, D. (2007). *PPVT-4: Peabody picture vocabulary test*. Minneapolis, MN. Pearson Assessments.
- Dziuk, M., Larson, J., Apostu, A., Mahone, E., Denckla, M., Mostofsky, S. (2007). Dyspraxia in autism: association with motor, social, and communicative deficits. *Developmental Medicine and Child Neurology*, 49(1), 734-739.
- Eigsti, I.M. & Fein, D. (2013). More is less: pitch discrimination and language delays in children with optimal outcomes from autism. *Autism Research*, 6, 605-613.
- Eigsti, I., Schuh, J., Mencl, E., Schultz, R. T., & Paul, R. (2012). The neural underpinnings of prosody in autism. *Child Neuropsychology*, 18(6), 600-617.
- Esposito, G., Rostagno, M., Venuti, P., Haltigan, J., & Messinger, D. (2014). Brief Report: Atypical expression of distress during the separation phase of the strange situation procedure in infant siblings at high risk of ASD. *Journal for Autism and Developmental Disorders*, 44, 975-980.
- Filipe, M.G., Frota, S., Castro, S.L., Vicente, S.G. (2014). Atypical prosody in Asperger syndrome: perceptual and acoustic measurements. *Journal of Autism and Developmental Disorders* 44, 1972- 1981.
- Fletcher, S.G. (1972). Time-by-count measurement of diadochokinetic syllable rate. *Journal of Speech and Hearing Research*, 15, 763-770.
- Gebauer, L., Skewes, J., Hørlyck, L., & Vuust, P. (2014). Atypical perception of affective prosody in Autism Spectrum Disorder. *NeuroImage: Clinical*, 6, 370-378.
- Fletcher, S. (1972). Time by count measurement of diadochokinetic syllable rate. *Journal of Speech and Hearing Research*, 15, 763-770.

- Friberg, J., McNamara, T. (2010). Evaluating the reliability and validity of (central) auditory processing tests: A preliminary investigation. *Journal of Educational Audiology*, 16, 4-17.
- Fulceri, F., Grossi, E., Contaldo, A., Narzisi, A., Apicella, F., Parrini, I., et al. (2019). Motor skills as moderators of core symptoms in Autism Spectrum Disorders: Preliminary data from an exploratory analysis with artificial neural networks. *Frontiers in Psychology*, 9(2683), 1-12.
- Gargan & Andrianopoulos (in review). Receptive and expressive lexical stress in adolescents with autism.
- Gargan & Andrianopoulos (2018). Evidence of motor speech impairments and abnormal prosody in Autism Spectrum Disorder. Paper submitted to the Organization for Autism Research, 1-37.
- Gernsbacher, M., Goldsmith, H., O'Reilly, M., Sauer, E., DeRuyter, J., & Blanc, M. (2002). *Infant motor dyspraxia as a predictor of speech in childhood autism*. Proceedings from a Poster presented at the International Conference on Infant Studies, Toronto.
- Gernsbacher, M.A., Geye, H., Hill Goldsmith, H., Sauer, E., & Schweigert, E. (2008). Infant and toddler oral- and manual-motor skills predict later speech fluency in autism. *Journal of Child Psychology and Psychiatry*, 49, 43-50.
- Golan, O., Baron-Cohen, S., Hill, J., & Rutherford, M. (2007). The 'Reading the Mind in the Voice' test-revised: a study of complex emotion recognition in adults with and without autism spectrum conditions. *Journal of Autism and Developmental Disorders*, 37(6), 1096-1106. DOI: [10.1007/s10803-006-0252-5](https://doi.org/10.1007/s10803-006-0252-5)
- Grossman, R. (2015). Judgments of social awkwardness from brief exposure to children with and without high-functioning autism. *Autism*, 19(5), 580-587.
- Grossman, R., Edelson, L., & Tager-Flusberg, H. (2013). Emotional facial and vocal expressions during story retelling by children and adolescents with high functioning autism. *Journal of Speech, Language, and Hearing Research*, 56(3), 1035-1044.
- Grossman, R., & Tager-Flusberg, H. (2012). "Who said that?" Matching of low- and high-intensity emotional prosody to facial expressions by adolescents with ASD. *Journal of Autism and Developmental Disorders*, 42, 2546-2557.
- Grossman, R., Bemis, R., Skwerer, D. & Tager-Flusberg, H. (2010). Lexical and affective prosody in children with high-functioning autism. *Journal of Speech, Language, and Hearing Research*, 53(3), 778-793.

- Hall, N. (2013). "An investigation of the efficacy of direct and indirect AAC service provision via telepractice". Doctoral Dissertations Available from Proquest. AAI3589032. <https://scholarworks.umass.edu/dissertations/AAI3589032>
- Happe, F., & Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 36(1), 5-25.
- Harris, S. (2017). Early motor delays as diagnostic clues in autism spectrum disorder. *European Journal of pediatrics*, 176(9), 1259-1262. doi: 10.1007/s00431-017-2951-7
- Heaton, P. (2003). Pitch memory, labeling and disembedding in autism. *Journal of Child Psychology and Psychiatry*, 44(4), 543-551.
- Heaton, P., Davis, R., & Happe, F. (2008a). Research note: Exceptional absolute pitch perception for spoken words in an able adult with autism. *Neuropsychologia*, 46, 2095-2098.
- Heaton, P., Hudry, K., Ludlow, A., & Hill, E. (2008b). Superior discrimination of speech pitch and its relationship to verbal ability in autism spectrum disorders. *Cognitive Neuropsychology*, 25(6), 771-782.
- Heaton, P., Williams, K., Cummins, O., Happe, F. (2008c). Autism and pitch processing splinter skills. *Autism*, 12(2), 203-219.
- Hesling, I., Dilharreguy, B., Galera, C., Etchegoyen, K., Bouvard, M., & Allard, M. (2009). The integration of prosodic speech in High Functioning Autism: An fMRI study. *NeuroImage*, 47.
- Humes, L. E., Coughlin, M., & Talley, L. (1996). Evaluation of the use of a new compact disc for auditory perceptual assessment in the elderly. *Journal of the American Academy of Audiology*, 7(6), 419-427.
- Jarvinen-Pasley, A. & Heaton, P. (2007). Evidence for reduced domain-specificity in auditory processing in autism. *Developmental Science*, 10(6), 786-793.
- Jarvinen-Pasley, A., Wallace, G., Ramus, F., Happe, F., & Heaton, P. (2008). Enhanced perceptual processing of speech in autism. *Developmental Science*, 11(1), 109-121.

- Jarvinen-Pasley, A., Wallace, G., Ramus, F., Happe, F., & Heaton, O. (2008b). Enhanced perceptual processing of speech in autism. *Developmental Science*, 11(1), 109-121. doi:10.1111/j.1467-7687.2007.00644.x.
- Jolliffe, T. & Baron-Cohen, S. (1999). A test of central coherence theory: linguistic processing in high-functioning adults with autism or Asperger syndrome: is local coherence impaired? *Cognition*, 149-185.
- Jones, C., Happe, F., Baird, G., Simonoff, E., Marsden, A., Tregay, J. et al. (2009). Auditory discrimination and auditory sensory behaviors in autism. *Neuropsychologia*, 47, 2850-2858.
- Kaland, C., Krahmer, E., & Swerts, M. (2012). Contrastive intonation in autism: The effect of speaker-and-listener-perspective. In Proceedings of the 13th Annual Conference of the International Speech Communication Association (Interspeech). (Vol. 1, pp. 1046-1049). Red Hook, U.S.A.: Curran Associates, Inc.
- Kaland, C., Swerts, M., & Krahmer, E. (2013). Accounting for the listener: Comparing the production of contrastive intonation in typically-developing speakers and speakers with autism. *The Journal of the Acoustical Society of America*, 134(3), 2182-2196.
- Kanner, L. (1971). Follow-up study of eleven autistic children, originally reported in 1943. *Journal of Autism and Childhood Schizophrenia*, 1, 119-145.
- Kargas, N., Lopez, B., Morris, P., & Reddy, V. (2016). Relations among detection of syllable stress, speech abnormalities, and communicative ability in adults with autism spectrum disorders. *Journal of Speech, Language, and Hearing Research*, 59, 206-215.
- Karhson, D. & Golob, E. (2016). Atypical sensory reactivity influences auditory attentional control in adults with autism spectrum disorders. *Autism Research*, 9(10), 1079-1092. DOI: 10.1002/aur.1593
- Kay Pentax, 1993. *Multi-Speech*. [Computer program] (Pine Brook NJ: Author).
- Kay Pentax, 1993. *Multi-Dimensional Voice Program (MDVP)*. [Computer program] (Pine Brook NJ: Author).
- Kay Pentax, 1993. *Multi-Dimensional Voice Program (MDVP) Manual*.
- Kazdin, A. E. (2011), 2nd edition. *Single-Case Research Designs: Methods for Clinical and Applied Settings*. New York: Oxford University Press.

- Kelley, A. (2007). Normative data for behavioral tests of auditory processing for New Zealand school children aged 7 to 12 years. *Australian and New Zealand Journal of Audiology*, 60-64.
- Kent, R. D., Kent, J. & Rosenbek, J. (1987). Maximum performance tests of speech production. *Journal of Speech and Hearing Disorders*, 52, 367-387.
- Kern, J. et al. (2013). Handgrip strength in autism spectrum disorder compared with controls. *Journal of Strength and Conditioning Research*, 27(8), 2277-2281. DOI: 10.1519/JSC.0b013e31827de068.
- Kleinman, J., Ventola, P., Pandey, J., Verbalis, A., Barton, M., Hodgson, S. (2008). *Journal of Autism and Developmental Disorders*, 38, 606-615.
- Korpilahti, P., Jansson-Verkasalo, E., Mattila, M., Kuusikko, S., Suominen, K., Rytty, S., Moilanen, I. (2006). Processing of affective speech prosody is impaired in Asperger Syndrome. *Journal of Autism and Developmental Disorders*, 37(8), 1539-1549.
- Kozlowski, A., Matson, J., Horovitz, M., Worley, J., & Neal, D. (2011). Parents' first concerns of their child's development in toddlers with autism spectrum disorders. *Developmental Neuropsychology*, 14(2), 72-78.
- Lord, C., Luyster, R., Gotham, K., Guthrie, W. (2012). *Autism Diagnostic Observation Schedule, Second Edition (ADOS-2)*. Torrance, CA: Western Psychological Services, 2012.
- Lord, C., Risi, S., DiLavore, P., Shulman, C., Thurm, A., Pickles, A. (2006). Autism from 2 to 9 years of age. *Arch Gen Psychiatry*, 63, 694-701.
- Loveland, K. A. (2008). Judgments of auditory-visual affective congruence in adolescents with and without autism: A pilot study of a new task using fMRI. *Perceptual and Motor Skills PMS*, 107(6), 557.
- Matson, J. L., & Goldin, R. L. (2013). Comorbidity and autism: trends, topics and future directions. *Research in Autism Spectrum Disorders*, 7, 1228-1233.
- Mayer, J., Hamment, I., Heaton, P. (2016). Mapping the developmental trajectory and correlates of enhanced pitch perception on speech processing in adults with ASD. *Journal of Autism and Developmental Disorders*, 46, 1562-1573.
- McCaleb, P. & Prizant, B. (1985). Encoding of new versus old information by autistic children. *Journal of Speech and Hearing Disorders*, 50(3), 230-240.
- McCann, J. & Peppe, S. (2003). Prosody in autism spectrum disorders: a critical review. *International Journal of Language and Communication Disorders*, 38(4), 325-350.

- McCann, J., Peppe, S., Gibbon, F., O'Hare, A., & 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(6), 682-702.
- McDermott, E., Smart, J., Boiano, J., Bragg, L., Colon, T., Hanson, E., Emanuel, D., & Kelly, A. (2016). Assessing auditory processing abilities in typically developing school-aged children. *Journal of the American Academy of Audiology*, 27(2), 72-84. DOI: 10.3766/jaaa.14050.
- Minagawa-Kawai, Y., Naoi, N., Kikuchi, N., Yamamoto, J., Nakamura, K., & Kojima, S. (2009). Cerebral laterality for phonemic and prosodic cue decoding in children with autism. *NeuroReport*, 20(13), 1219-1224.
- Ming, X., Brimacombe, M., & Wagner, G. (2007). Prevalence of motor impairment in autism spectrum disorders. *Brain and development*, 29(9), 565-570.
- Mottron, L., Dawson, M., Soulières, I., Hubert, B., Burack, J. (2006). Enhanced perceptual functioning in autism: an update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders*, 36(1), 27-43.
- Musiek, F. (1994). *Duration Pattern Test*. Storrs, CT: Audiology Illustrated.
- Musiek, F. (2002). The frequency pattern test: A guide. *The Hearing Journal*, 55(6), 58.
- Musiek, F.E. & Pinheiro, M.L. (1987). Frequency Patterns in Cochlear, Brainstem, and Cerebral Lesions. *Audiology*, 26, 79-88.
- Nadig A., & Shaw, H. (2012). Acoustic and perceptual measurement of expressive prosody in high-functioning autism: Increased pitch range and what it means to listeners. *Journal of Autism and Developmental Disorders*, 42, 499-511. <https://doi.org/10.1007/s10803-011-1264-3>
- Nadig, A. & Shaw, H. (2015). Acoustic marking of prominence: how do preadolescent speakers with and without high-functioning autism mark contrast in an interactive task? *Language, Cognition and Neuroscience*, 30, 1-2, 32-47. DOI: 10.1080/01690965.2012.753150
- Nakai, Y., Takashina, R., Takiguchi, T., Takada, S. (2013). Speech intonation in children with autism spectrum disorder. *Brain & Development*, 36(6), 516-522. <https://doi.org/10.1016/j.braindev.2013.07.006>
- O'Connor, K. (2012). Auditory processing in autism spectrum disorder: a review. *Neuroscience and Behavioral Reviews*, 36, 836-854.

- O’Riordan, M. & Passetti, F. (2006). Discrimination in autism with different sensory modalities. *Journal Autism Developmental Disorders*, 36, 665-675.
- Ozonoff, S., Young, G., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., et al. (2011). Recurrence risk for autism spectrum disorders: A baby siblings research consortium study. *Pediatrics*, 128(3), 488-495. doi: 10.1542/peds.2010-2825
- Page, J. & Boucher, J. (1998). Motor impairments in children with autistic disorder. *Child Language Teaching and Therapy*, 14(3), 233-259..
- Parish-Morris, J., Sariyanidi, E., Zampella, C., Bartley, G., Ferguson, E., Pallathra, A., Bateman, L., Plate, S., Cola, M., Pandley, J., Brodtkin, E., Schultz, R., & Tunc, B. (2018). Oral-motor and lexical diversity during naturalistic conversations in adults with autism spectrum disorder. In K. Loveys, K. Niederhoffer, E. Prud’hommeaux, R. Resnik, & P. Resnik (Eds.), *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic* (pp. 147-157). Association for Computational Linguistics <https://doi.org/10.18653/v1/W18-0616>
- Paul, R., Augustyn, A., Kin, A., Volkmar, F. (2005). Perception and production of prosody by speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 35(2), 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., Shriberg, L., McSweeney, J., Cicchetti, D., Klin, A., & Volkmar, F. (2005). Brief report: Relations between prosodic performance and communication and socialization ratings in high functioning speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 35(6), 861-869. DOI: 10.1007/s10803-005-0031-8
- Peppé, S. (2015). Profiling Elements of Prosody in Speech Communication (PEPS-C).
- Peppé, S., Cleland, J., Gibbon, F., O’Hare, A., 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.

- Plumb, A. & Wetherby, A. (2013). Vocalization development in toddlers with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research*, 56, 721-734.
- Pusponegoro, H., Efar, P., Soedjatmiko, Soebadi, A., Firmansyah, A., Chen, H., & Hung, K. (2016). Gross motor profile and its association with socialization skills in children with Autism Spectrum Disorders. *Pediatric Neonatology*, 57(6), 501-507.
- R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Rinehart, N., Bradshaw, J., Moss, S., Brereton, A., Tonge, B. (2001). A deficit in shifting attention present in high-functioning autism but not Asperger's disorder. *Autism*, 5(1), 67-80.
- Rutherford, M., Baron-Cohen, S., Wheelwright, S. (2002). Reading the mind in the voice: A study with normal adults and adults with Asperger syndrome and high functioning autism. *Journal of Autism and Developmental Disorders*, 32(3), 189-194.
- Scheinkopf, S., Mundy, P., Oller, D., Steffens, M. (2000). Vocal atypicalities of preverbal autistic children. *Journal of Autism and Developmental Disorders*, 30(4), 345-354.
- Schoen, E., Paul, R., & Chawarska, K. (2011). Phonology and vocal behavior in toddlers with autism spectrum disorders. *Autism Research*, 4, 177-188.
- Seal, B. & Bonvillian, J. (1997). Sign language and motor functioning in students with autistic disorder. *Journal of Autism and Developmental Disorders*, 27(4), 437-466. DOI: [10.1023/a:1025809506097](https://doi.org/10.1023/a:1025809506097)
- Shetreat-Klein, Shiinar, S., & Rapi, I. (2014). Abnormalities of joint mobility and gait in children with Autism Spectrum Disorders. *Neurology*, 36(2), 91-96.
- Shriberg, L., Kwiatowski, J., Rasmussen, C., Lof, G., & Miller, J. (1992). *The prosody-voice screening profile (PVSP): Psychometric data and reference information for children*. (Tech. Rep. No. 1). Phonology Project, Waisman Center on Mental Retardation and Human Development, University of Wisconsin-Madison.
- Shriberg, L., Paul, R., McSweeney, J., Klin, A., Cohen, D., Volkmar, F. (2001). Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. *Journal Speech Language and Hearing Research*, 44(5), 1097-1115.

- Sparks, B., Friedman, S., Shaw, D., Aylward, E., Echelard, D., Artru, A., et al. (2002). Brain structural abnormalities in young children with autism spectrum disorder. *Neurology*, 59(2), 184-192.
- Stanutz, S., Wapnick, J., & Burack, J. (2013). Pitch discrimination and melodic memory in children with autism spectrum disorders. *Autism*, 18(2), 137-147.
- Velleman, S. (2015). *Speech Sound Disorders* (New Edition ed.). Lippincott Williams & Wilkins.
- Velleman, S. (2002). VMPAC and PEPS: Effective new tool for differential diagnosis? Presented at Childhood Apraxia of Speech Research Symposium, Tucson, AZ.
- Velleman, Shelley, Andrianopoulos, Mary, Boucher, Marcil, Perkins, Jenny, Averbach, Karen, Currier, Alyssa, Marsello, Michael, Van Emmerik, Richard. (2009). Motor speech disorders in children with autism. In Rhea Paul & P. Flipsen (Eds.), *Speech sound disorders in children* (pp. 141-180). San Diego: Plural Publishing.
- Waldie, K. & Saunders, A. (2014). The neural basis of autism: A review. *International Journal of School and Cognitive Psychology*, 1(3), 1-4.
- Wang, A. T. (2006). Neural basis of irony comprehension in children with autism: The role of prosody and context. *Brain*, 129(4), 932-943.
- Wang, A. T., Lee, S. S., Sigman, M., & Dapretto, M. (2007). Reading affect in the face and voice. *Arch Gen Psychiatry Archives of General Psychiatry*, 64(6), 698.
- Warlaumont, A., Richards, J., Gilkerson, J., & Oller, D. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science*, 25(7): 1314–1324.
- Watanabe, T., Yahata, N., Abe, O., Kuwabara, H., Inoue, H., Takano, Y., Yamasue, H. (2012). Diminished medial prefrontal activity behind autistic social judgments of incongruent information. *PLoS ONE*, 7(6), 1-11. DOI: 10.1371/journal.pone.0039561
- Williams, K.T. (2007). *Expressive Vocabulary Test, Second Edition*. Circle Pines, MN: AGS Publishing.
- Wilson, D. (1987). Voice problems of children (pp. 118-119). Baltimore, MD: Williams and Wilkins.