ABSTRACT

Title of Thesis:     RECEPTIVE PROSODY SKILLS IN INDIVIDUALS WITH
                     HIGH FUNCTIONING AUTISM SPECTRUM DISORDERS

                     Megan Janssen, Master of Arts, 2012


Thesis directed by:  Professor Rochelle Newman
                     Department of Hearing and Speech Sciences

Prosodic differences have been noted in the speech production of individuals with Autism Spectrum Disorders (ASD); however, little is known regarding their ability to perceive and understand features of prosody. It has been determined that children with typical development (TD) can recognize and utilize the prosodic cue of contrastive stress to facilitate interpretation of spoken instructions (Arnold, 2008). We examined this skill in 12 children and adolescents with ASD, and 12 with TD through the analysis of eye fixations to objects during instructions with varying discourse statuses (given or new) and stress patterns (accented or unaccented). Results indicated that both the participants with TD and with ASD were able to perceive and interpret the prosodic cue of contrastive stress within the contextual communication task. No relationships between language, cognitive, or expressive prosody skills and receptive prosody skills were found. Possible explanations, and clinical implications are discussed.

RECEPTIVE PROSODY SKILLS IN INDIVIDUALS WITH HIGH FUNCTIONING
AUTISM SPECTRUM DISORDERS

by

Megan Janssen

Thesis submitted to the faculty of the graduate school of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of Arts
2012

Advisory Committee:

       Professor Rochelle Newman, Chair
       Professor Yi Ting Huang
       Professor Nan Bernstein Ratner

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

## LIST OF TABLES

vi

LIST OF FIGURES

1.      Image of the Podium                                                          25

2.      TD Fixation Proportions to Target during Sentence Onset Region               40

3.      TD Fixation Proportions to Target during Target Disambiguation Region  41

4.      ASD Fixation Proportions to Target during Sentence Onset Region              42

5.      TD Mean Baseline Switch Difference Scores                                    45

6.      ASD Mean Baseline Switch Difference Scores                                   46

7.      Correlation between Expressive and Receptive Prosody Performance             49

Receptive prosody skills in individuals with autism spectrum disorders

When we speak, we convey a significant amount of our meaning not just through our words, but through the way we say them. For example, one could imagine a teenager might say, "I got a new haircut" in a variety of ways, indicating emotions ranging from excitement to disappointment. We utilize pauses within sentences for phrase demarcation, and emphasize words or parts of words. Typical adults are skilled at perceiving and using variations in pitch, duration, amplitude, and rhythm to modify a signal, and understanding the meaning behind those signal modifications. These acoustic modifications, among many others, make up the suprasegmental features of speech that are known as prosody (Fox, 2000). Prosody is not just a separate "add-on" to speech, but is often deeply imbedded in and serves many functions relating to the interpretation of spoken language (Baumann, 2006).

Prosody can be thought of as having two main levels: form and function. The form level consists of the acoustic features of the prosody, such as pitch, duration, or amplitude. The function level is the purpose that these features serve, or what message the speaker is conveying through the production of these features (McCann & Peppé, 2003). The listener must both perceive and understand the prosodic features to recognize the nuances of a message.

**Functions of Prosody**

Prosody serves grammatical, affective, and pragmatic functions in communication (McCann & Peppé, 2003; Korpilahti et al., 2007; Paul, Augustyn, Klin & Volkmar, 2005a). Grammatical prosody consists of modifications to the features of the speech signal that impact the syntax of the sentence, and may include information such as stress

used within words to signal speech class, or pauses to indicate phrase boundaries (Paul et al., 2005a; Peppé, McCann, Gibbon, O'Hare & Rutherford, 2007). Peppé and colleagues (2007) provide the example of "Ellen, the dentist, is here." as opposed to "Ellen, the dentist is here." The placement of the phrase break within the sentence indicates whether the speaker is introducing Ellen, who is a dentist, or if the speaker is notifying Ellen that the dentist is present.

The affective component of prosody provides cues allowing for the interpretation of the speaker's emotions (Peppé et al., 2007). The speaker may modify his or her rate, pitch, or amplitude to demonstrate his or her feelings (Peppé et al., 2007; Paul et al., 2005a). For example, the statement, "The bus is coming." could be said with an increased rate or amplitude to express the feeling of urgency.

Finally, pragmatic prosody is used to emphasize a portion of a sentence, and can be used to draw the listener's attention to information that may be novel or important (Paul et al., 2005a; Peppé et al., 2007). For example, one could say, "I wanted the red SHIRT." or "I wanted the RED shirt." The emphasis placed on the words in the first sentence indicates that the speaker may have been given a red article of clothing, but had wanted the shirt instead. The speaker in the second sentence may have been given a shirt, but had wanted one of a different color. The placement of stress in the sentences signals the needs or intentions of the speaker to the listener.

**Forms of Prosody**

The functions of prosody are manifested through acoustic changes in the signal that define the prosodic form. Although there is some debate over the primary forms of prosody, some main forms include duration, stress/accent, and phrase boundaries (Fox,

2000; Gerken & McGregor, 1998). A range of relatively simple to surprisingly complex acoustic modifications to the signal at both the segmental and suprasegmental levels contribute to the recognition of these forms (Fox, 2000).

Length, or duration, may appear relatively basic, but is actually a complex interaction between segmental as well as suprasegmental structures. That is, both the actual phoneme can be lengthened, and/or a syllable or group of syllables can be lengthened (Fox, 2000). Length can also be influenced by other factors such as the speaker's overall rate of speech, the context of the segment or syllable, or idiosyncratic characteristics of the speaker (Turk, Nakai & Sugahara, 2006). When utilized as a prosodic form, it can serve as a stand-alone structure, or may occur concurrently with and contribute to other forms of prosody, such as accent or stress. Length and duration can be measured acoustically from the onset of the speech segment of interest until the conclusion of the segment (Fox, 2000).

Accent and stress are prosodic forms consisting of acoustic changes in a number of components of the signal. Acoustically, stressed words typically have increased duration, higher fundamental frequency, and increased amplitude (Fox, 2000). Perceptually, this leads to a stressed word produced as longer, higher pitched, and louder. However, the three acoustic cues that create the perception of stress are not equally important. In determining the presence of stress, fundamental frequency holds the most importance, followed by duration, then loudness (Cruttenden, 1997). Stress may be present at the syllable or word level, such as in the differentiation between, "REcord" (as in, "He holds the world record."), and "reCORD" (as in, "Don't forget to record my favorite show."), or at the sentence level.

In determining the stress patterns within sentences, stress at the syllable level provides information towards the sentence level stress pattern (Cruttenden, 1997). Cruttenden (1997) proposed that hierarchies are present, including "rhythm groups" which are based on the pattern of stress at the syllable level, "accent phrases" in which pitch accent over a larger segment is analyzed, and "intonational groups", which encompass the lower levels to create the overall accent pattern of the message (pp. 23-24). These groups consist of primary, secondary, tertiary and unstressed components, based on the location and strength of the pitch prominence (Cruttenden, 1997).

The presence and location of pitch prominence may be determined by the "focus" of a sentence or phrase. This focus can be "narrow", meaning attention is drawn to a single word, or "broad", meaning attention is drawn to a group of words. According to Ladd (2008), within broad focus, two options arise: "normal stress" and "contrastive stress" (p. 216). Ladd (2008) distinguishes these terms by stating, "Normal stress has no meaning or function; it is simply the result of the operation of phonological rules on surface syntactic structures. Any deviation from normal stress, on the other hand, involves 'contrastive stress', which signals some sort of contrast or emphasis on the stressed word." (p. 216). The identification of contrastive and normal stress, however, is dependent on the interpretation of the acoustic cues (i.e., duration, peak and overall fundamental frequencies, maximum and overall amplitude) in conjunction with the semantic and syntactic structures within the sentence (Ladd, 2008). For example, contrastive stress is utilized when stating, "I wanted the BLUE car" (with an emphasis placed on the color) versus "I wanted the blue CAR" (with an emphasis placed on the

object). These cues together determine the prosodic form of stress or accent at the word and sentence levels.

A phrase boundary is another example of a prosodic form. A phrase boundary may be identified by the presence of pausing, movements of pitch, or increases in vowel duration within a word. These acoustic modifications can be present within a phrase to identify clause boundaries, or to mark the end of a phrase (Gerken & McGregor, 1998). In addition, low and high boundary tones can be utilized to demarcate both the termination of a phrase, as well as the presence of a relationship between two phrases (Pierrehumbert & Hirschberg, 1990). For example, a low phrase accent followed by a high boundary tone might indicate that the following phrase is somewhat related to its previous phrase (Pierrehumbert & Hirschberg, 1990).

These and other acoustic cues, such as pitch range and rhythm, are examples of various forms of prosody. Interpretation of the forms allows for meaning to be gained regarding the functions that the prosody is serving.

**Overview of the Visual World Paradigm: A Psycholinguistic Approach**

One method for examining the perception of prosody is the visual world paradigm. When individuals listen to spoken language, they process it in an incremental fashion. The visual world paradigm involves researchers tracking the eye movements of individuals while spoken language is presented to analyze how language is being processed (Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995). For example, when listening to the sentence, "Put the five of hearts that is below the eight of clubs above the three of diamonds," participants shifted gaze quickly from image to image, corresponding to the words being read (Tanenhaus et al., 1995). This is an example of how listeners

parse meaning from sentences in an incremental fashion. Participants also processed

language at such a rate that they shifted their gaze to the correct object even before the

end of the word was uttered when other objects present did not have names that were

similar in onset (Tanenhaus et al., 1995).  The online tracking of language processing in

the visual world paradigm allows researchers to gain a better understanding of how

participants perceive speech, and has been used to research speech perception in topics

ranging from disfluency, to accent, to processing of prosody (Arnold, Tanenhaus,

Altmann, & Fagnano, 2004; Dahan, Tanenhaus & Chambers, 2002; Snedeker & Yuan,

2008).

**Use of Psycholinguistic Approaches in Receptive Prosody Research**

Through the tracking of eye movements, the processing of stress and pitch has

been studied. A confirmation that adult listeners perceive words with a high stress as

words that are "new" to a discourse was provided using this paradigm. The distinction

between "given" and "new" is defined by Dahan et al. (2002) as:

> Given information is typically equated with entities that have been previously
> introduced into the discourse and which share the same linguistic form and
> grammatical role as their antecedents…An expression is described as containing
> new information when it introduces an entity that has not been previously
> mentioned (p. 293).

Dahan et al. (2002) provided participants with pairs of sentences with varying stress

patterns instructing the participants to manipulate objects. Each sentence pair contained

two target words with similar onsets (e.g., "Put the candle/candy above the square"

followed by either "Now put the CANDLE above the square" or "Now put the candle

ABOVE THE SQUARE"). Four conditions were provided. These included combinations

of either given (i.e., previously-mentioned) or new (i.e., previously unmentioned)

discourse contexts, and accented or unaccented target words. Even during the onset of the target word in the second sentence (candle), typical adult participants looked at the previously-unmentioned object (i.e., the "new" object, not stated in the previous sentence) when the word was accented. This indicated that participants tended to consider accented nouns as new or novel information, while unaccented nouns were considered to be objects already mentioned (Dahan et al., 2002). Adults are able to make this judgment because they are able to perceive the prosodic cue, and gain meaning from it. This means that they are able to follow the discourse and identify that the speaker utilized stress to emphasize and draw listener attention to the word, indicating that it is likely new or important.

Arnold utilized the same paradigm with 4- and 5-year olds to investigate developmental aspects of this skill. Findings were similar in that participants were able to utilize prosodic cues within a discourse; however, 4- and 5-year olds only demonstrated the effect of accent during the "given" conditions. That is, when the object mentioned in the first sentence within the instruction set was the same as the object mentioned in the second sentence, participants fixated faster and longer to the target when it was unaccented than when it was accented (Arnold, 2008). It was speculated that it might be the case that an unaccented word is reserved for accessible referents, whereas an accented word could be utilized in a variety of discourse contexts (Arnold, 2008).

To examine typically developing children's (4-6 year-olds) and adults' understanding of the prosodic cue of phrase breaks to disambiguate sentences, Snedeker and Yuan (2008) used a modified methodology of the visual world paradigm. The researchers presented four objects on a podium, and instructed the participant to look at

the center. The objects provided were a "(1) Target Instrument, a full-scale object that could be used to carry out the action; (2) a Target Animal, a stuffed animal carrying a small replica of the Target Instrument (the frog holding a little feather); (3) a Distractor Instrument; a second full-scale object (the candle); and (4) a Distractor Animal, a stuffed animal of a different kind carrying a replica of the Distractor Instrument (the leopard carrying a candle)." (Snedeker & Yuan, 2008, p. 581). Once the participant was fixated, he or she was heard the sentence "You can feel the frog with the feather," with varying pauses/phrase breaks in the sentence. The "modifier" version of the sentence had a pause between the words "feel" and "the frog" (i.e., "[You can feel] [the frog with the feather.]"), which would be more likely to induce an interpretation of touching the frog that was holding the feather. The "instrument" sentence had a pause between the words "frog" and "with." (i.e., "[You can feel the frog] [with the feather.]"), which would be more likely to induce an interpretation of using the feather to touch the frog (Snedeker & Yuan, 2008). The participant then performed the action. The participants' eyes were video-recorded and coded for looking rate and eye movement. Based on the incremental looking patterns of the participants, it was determined that participants were able to interpret the prosodic breaks either before the prepositional phrase or before the noun phrase to comprehend ambiguous sentences (Snedeker & Yuan, 2008). That is, when the break was provided before the prepositional phrase, adults consistently identified that they should touch the frog carrying the small feather, and when the break was before the noun phrase, adults used the feather to touch the frog. Interestingly, the 4-6 year old children only showed this pattern consistently during the first block. After the first block, children appeared to perseverate and no longer showed effects of prosody. In a second

experiment, sentences were also provided in which there were potential prosodic and lexical biases present. Conditions included sentences in which the modifier was a more common completer of the sentence, the instrument was a more common completer of the sentence, and a condition in which there was no significant difference in the frequency of the completer. Adults and children used both prosodic and discourse biases together to interpret an ambiguous instruction, at least initially (Snedeker & Yuan, 2008). These studies utilizing psycholinguistic paradigms confirmed that typically developing adults and children utilize prosody in online processing of sentences (Arnold, 2008; Snedeker & Yuan, 2008; Dahan et al., 2002).

**Overview of Autism and Prosody**

So far, research using this paradigm has been primarily limited to typically developing children and adults. However, prosody has been implicated as being a source of difficulty in several clinical groups, including those identified with Autism Spectrum Disorders (ASD). Autism is a pervasive developmental disorder characterized by a triad of impairments, in which "absence or impairment of social interaction, especially with peers, absence or impairment of development of verbal and nonverbal language, and repetitive, stereotyped activities of any kind" are seen (Wing & Gould, 1979, p. 13). Many people with ASD have been reported to utilize abnormal prosody in their speech (Paul et al., 2005a; Baltaxe & Simmons, 1985) but limited research has investigated how well individuals with ASD comprehend prosodic cues.

In this study, we focus on the receptive prosody of individuals with ASD who are considered "high functioning," such that they are able to speak and communicate. This ensures that intellectual impairments do not serve as a confounding variable in the task.

Although in older children and adolescents, language skills may be present and relatively intact, individuals with High Functioning Autism (HFA) and Asperger's Syndrome (AS) typically show difficulty with social language use (American Psychiatric Association [DSM-IV-TR], 2000). This social language impairment may impact use of prosody. The speech of individuals with HFA has been reported to be monotone in nature, while individuals with AS have been reported to utilize intonation patterns inconsistent with their message. For example, a joke may be told in the form of a statement, or an inappropriately fast rate or loud voice may be used (Rubin & Lennon, 2004). In both cases, individuals with ASD are utilizing inappropriate prosody. Relationships have been found between the mastery of the use of prosody and the social competence ratings of people with ASD, highlighting the importance of appropriate use of prosody during social interactions (Paul, Shriberg, McSweeny, Cicchetti, Klin & Volkmar, 2005b). Because of the importance of appropriate prosody, this is often an area of clinical intervention for individuals with generally intact language function (Bellon-Harn, Harn & Watson, 2007). Understanding the perception of prosody in this group is therefore important in facilitating more efficient prosody intervention.

**Past Research on Autism and Prosody**

A review of the research available on the topic of autism and prosody was performed by McCann and Peppé in 2003. Between 1980 and 2002, only 16 studies were conducted investigating the topic. Most studies focused on the expressive use of prosody, with little emphasis on people with autism's understanding of prosody. This indicates a need for further research in this topic area, seeing as the cause of atypical prosody expressively may lie in deficits in receptive prosody (Peppé et al., 2007; McCann &

Peppé, 2003). Within the studies reviewed, conflicting results were found, such as discrepancies regarding people's ability to use default stress appropriately (Baltaxe & Guthrie, 1987; Shriberg et al., 2001). Such conflicting results were speculated to be the result of variable definitions of terms in the realm of prosody, a lack of standardization, small sample sizes, and limited use of control groups (McCann & Peppé, 2003).

Since that time, there has been an increase in research in the area of prosody and autism, yet many questions remain. Differences, compared to typically developing individuals, in perception of affective prosody have been confirmed in boys between 9 and 12 years old with AS through the use of auditory event-related potentials (ERPs). When listening to presentations of a single word with two varying fundamental frequency patterns, children with AS demonstrated increased ERP latencies (particularly in their right hemisphere), indicating atypical processing of the acoustic signal and therefore of the prosodic features within that signal (Korpilahti et al., 2007). Although children with AS tend to have less severe deficits with prosody than children with HFA (Peppé, Cleland, Gibbon, O'Hare & Martinez Castilla, 2011), this confirms that even at a neurological level there are differences in the perception of prosody by this population when compared to typically developing individuals (Korpilahti et al., 2007). According to the *Diagnostic and Statistical Manual IV,* individuals with AS do not demonstrate disordered or delayed language prior to age three, whereas individuals with other forms of ASD (such as Autistic Disorder) do show early deficits in language (American Psychiatric Association, 2000). It is possible that language skills and prosodic skills are related. However, children with HFA perform more poorly than verbal-age matched

participants on tasks of prosody, indicating that prosodic difficulties cannot solely be attributed to below-average language skills (Peppé et al., 2011).

Although there is variability in the skills of individuals with ASD, impairments have been demonstrated in a variety of receptive prosody tasks. For example, Peppé and colleagues found that children with ASD demonstrated impairments in prosodic imitation tasks, which do not even require sentence formulation, and may reflect the lack of prosodic forms in the child's skill set (Peppé et al., 2011). Additionally, individuals with ASD have demonstrated difficulties perceiving the form (i.e., identification of prosodic features) of prosody over longer speech streams, which could mean they are at a disadvantage in determining the function (i.e., gaining meaning from the prosodic features) of the message (Jarvinen-Pasley, Peppé, King-Smith & Heaton, 2008). Children with ASD also have been found to show a bias to perceive sentences as declarative as opposed to interrogative (Paul et al., 2005a; Jarvinen-Pasley et al., 2008), and to perceive two signals as "different" even when they are the same (Peppé et al., 2007).

Paul and colleagues (2005a) sought to determine the receptive and expressive prosody skills of individuals ages 14-21 with HFA and AS in the areas of grammatical and affective/pragmatic functions of prosodic features. Tasks measured grammatical and pragmatic perception and production of stress, intonation, and phrasing. Although it was anticipated by the researchers that participants with ASD would have more difficulty in tasks of prosody perception and production with pragmatic, rather than grammatical, functions (due to noted pragmatic deficits in this population), no significant difference between the ASD group and the control group was found on tasks of grammatical versus pragmatic prosody. That being said, there was near perfect performance on 3/6 perception

tasks (i.e., grammatical intonation, pragmatic intonation, pragmatic phrasing), and 2/6

production tasks (i.e., grammatical intonation, pragmatic phrasing) for both groups,

indicating that ceiling effects may have masked significant differences. Participants with

ASD did show difficulties in grammatical and pragmatic functions in the prosodic

component of stress, both receptively and expressively (consistent with Shriberg et al.,

2001); however, many of the tasks required the participants to read their responses. For

example, the grammatical stress perception task required the participants to listen to a

word being said with contrasting stress, and identify the corresponding word in a written

sentence (i.e., choose between "I can't recall his name." versus "They had a recall on this

model car." after hearing "re call"), while the grammatical production task required the

participants to read aloud a word within a sentence with the appropriate stress being

placed on the correct syllable of the word (i.e., participants would read only the

underlined word aloud from the sentence, "I can't recall his name." with the expected

response being the word "re call.") (Paul et al., 2005a). It has been speculated that

prosodic features produced through reading a sentence may be different from prosodic

features used at the conversation level, which may have influenced results (Peppé et al.,

2007). Further investigation to confirm impaired perception and use of stress using more

validated methodology is needed.

In 2007, Peppé and colleagues performed a follow-up investigation to Paul and

colleagues' 2005a study to determine if slightly younger children (6-13 years old) with

HFA have more difficulty in producing and perceiving prosodic features than children

with typical development (TD), and also sought to determine if there is a correlation

between expressive and receptive skills for each of the components of prosody

investigated. All of the participants with HFA were found to have difficulty in at least one component of prosody, but there was variability in which component of prosody was most troublesome. Overall, the children with TD performed better on tasks of prosody perception and expression than the children with ASD. The task of affect identification, which required the listener to determine the feelings (like or dislike) of the speaker, showed the greatest disparity in skill level between the children with TD and the children with HFA (Peppé et al., 2007).

The children's skills were measured using the Profiling Elements of Prosodic Systems in Children (PEPS-C; Peppé & McCann, 2003), which is frequently used to assess prosody in this population (Peppé et al., 2007; McCann & Peppé, 2003). Each subtest of the PEPS-C contains an expressive and receptive component. Significant correlations were noted for the children with HFA between receptive and expressive scores on three subtests (*Overall* – combined scores, *Turnend* – utilizing/interpreting intonation to identify sentence type, and *Chunking* - "segmenting of utterances into prosodic phrases"), but not on the other four subtests (*Affect* – understanding or use of prosody for express emotions, *Focus*- use or understanding of accent or stress for emphasis, *Short-Item Discrimination/Imitation* – identifying whether two non-speech tones are the same or different and imitating words with different intonation patterns, and *Long-Item Discrimination/Imitation* - identifying whether two longer non-speech tones are the same or different and imitating phrases with different intonation patterns) (Peppé et al., 2007). A flaw in the design was that a rater judgment of response determined the child's score on the expressive component of the test. Additionally, a binary (correct or incorrect) system was used to measure perceptual judgment scores for expressive prosody

tasks (with the exception of the imitation of single words and short phrases subtest which allowed a score of "good," "fair," or "poor," corresponding to numerical scores of 1, 0.5, and 0, respectively). Although this scoring method will identify whether a skill is present or absent at the level tested, it would be beneficial to understand the level of mastery the child may or may not have, and have qualitative information regarding production. Understanding the error patterns exhibited may give insight into what contributes to the prosody of people with ASD as being considered "atypical." As the authors noted as a flaw, no acoustical analysis was taken to determine in what ways the deviant responses varied from the typical responses (Peppé et al., 2007).

**Use of Psycholinguistic Paradigms in the Study of Receptive Prosody in Children and Adults with ASD.** Thus far, limited research has been performed utilizing psycholinguistic paradigms with children or adults with ASD to determine the influence of prosody on the interpretation of verbal input. There are many advantages in using the visual world paradigm with this population, including that this paradigm does not require reading by the participants, nor does it require metalinguistic analysis by the participants (Trueswell, Sekerina, Hill & Logrip, 1999). This paradigm also avoids the influences of participants identifying or demonstrating the emotions and feelings of others, such as in other paradigms that require participants to point to corresponding pictures of "happy" or "sad". Finally, the visual world paradigm does not require participants to produce verbal responses, which may complicate the task for children with ASD.

One of the few studies utilizing psycholinguistic paradigms for this purpose was performed by Diehl, Bennetto, Watson, Gunlogson, and McDonough in 2008, using a similar paradigm as Trueswell and colleagues' 1999 study which investigated children's

processing of temporary syntactic ambiguities. Individuals (11-19 year olds) with HFA

and individuals with typical development matched on age, full scale IQ, and receptive

language skills participated. Participants were seated facing a scene in which the

quadrants contained a potential target (e.g., a dog), a potential target placement (e.g., a

basket atop a star), a placement location (e.g., a basket), and an unrelated distractor object

(e.g., a cat). A small picture of a dog and a cat were also located on the vertical midline

of the board, as to mitigate the effects of the lexical bias related to the word "put"

(Snedeker & Trueswell, 2004). Sentences were played in which prosodic cues (i.e., "[Put

the dog in the basket] [on the star]" or "[Put the dog] [in the basket on the star]"), lexical

cues (i.e., "[Put the dog that's in the basket on the star]" or "[Put the dog in the basket

that's on the star]"), or both (i.e., "[Put the dog that's in the basket] [on the star]" or "[Put

the dog] [in the basket that's on the star]") were provided to disambiguate the meanings

(Diehl et al., 2008). When provided with only prosodic cues to differentiate the sentence

meaning, participants with HFA performed significantly more poorly than participants

with typical development overall. When analyzed more closely, both participants with

typical development and participants with HFA performed near ceiling on sentences in

which the sentence structure utilized a verb-phrase attachment (e.g., "[Put the dog] [in the

basket on the star]"), while participants with TD had significantly higher performance

than participants with HFA when the sentence was structured with a noun-phrase

attachment (e.g., "[Put the dog in the basket] [on the star]". This is the same pattern that

five-year olds with TD demonstrated in Trueswell and colleagues' study (1999). The

authors speculated that this pattern of performance might be because of a "lexical bias

toward a VP-interpretation" related to the word "put" (Diehl et al., 2008, p. 149;

Trueswell et al., 1999). This speculation is because of the higher likelihood that a prepositional phrase attachment completes the sentence when the verb "put" is utilized, as in the sentences used in these studies. That is, it is more likely for the prepositional phrase following the noun phrase to be interpreted as a location for the noun to be moved, rather than a modifier of the noun. Questions arose whether participants with ASD had significantly lower performance (as determined by accurate completion of the instruction) than participants with TD because they were unable to perceive and utilize the prosodic cues, or because they were able to perceive the prosodic cues, but were unable to modify their original expectation of the VP completion to the sentence because of the word "put" (Diehl et al., 2008). Performance between the HFA and control groups was similar when given only syntactic cues, or the combination of syntactic and prosodic cues to disambiguate the instruction (Diehl et al., 2008).

A follow-up study was performed to eliminate the confound of lexical bias by using the stimuli from Snedeker and Yuan's 2008 study, and to gain more information about the processing of prosodic input by analyzing the eye movements of the participants during the presentation of the stimuli (Diehl, Snedeker, Tang, Paul, under review). Participants were children and adolescents with HFA ranging from 7-17 years of age, with chronological age, receptive language, full scale IQ and verbal IQ matched TD controls. They were divided into "younger" groups (7-12 year olds) and "older" groups (12-17 year olds). Participants were presented with four objects on a display – a frog holding a small feather, a distractor animal holding a small feather, a candle, and a feather, and heard either "[You can feel the frog] [with the feather]" or "[You can feel] [the frog with the feather]" or "[You can feel the frog with the feather]." The "younger"

participants with HFA demonstrated an effect of prosody on the first block; however, effects of prosody were no longer significant on following blocks, similar to the performance of Snedeker and Yuan's typically developing 4-6 year old participants. There was no significant difference in the performance of the TD and HFA "older" groups, and the TD "younger" group, in which all groups demonstrated an effect of lexical bias and an effect of prosody (Diehl et al., under review). Although participants with HFA performed the task successfully initially, the concept of perseveration or difficulty set shifting in attempts to utilize multiple prosodic cues in succession (as one would experience in the real world) warrants further investigation (Diehl et al., under review).

Given the prior research, the ability of individuals with HFA to utilize a wide-range of prosodic cues serving a variety of functions (i.e., pragmatic, grammatical, and affective) remains unclear. Studies have investigated the ability of individuals with HFA to resolve syntactic ambiguities; however, syntactic ambiguities fall within the category of grammatical prosody, which may be less challenging in this population than tasks of prosody serving the pragmatic function. The question arises if individuals with ASD are only successful in tasks involving grammatical prosody, or if they are successful with all functions of receptive prosody. When given a receptive pragmatic prosody task that children and adults with TD have demonstrated success in, are individuals with ASD also able to perform the task successfully? Both Dahan et al. (2002), and Arnold (2008) have demonstrated that children and adults with TD are able to use the prosodic cue of contrastive stress to provide information regarding the discourse status of a word. We are interested if individuals with ASD are also able to utilize that prosodic cue to aid in the

interpretation of the message. Questions also remain surrounding the ability for an individual with ASD to perceive and revise prosodic interpretations over multiple trials.

**Current Study**

This study utilized the visual world paradigm to investigate the processing of stress/accent by participants with ASD, compared to the processing of stress/accent by participants with TD. Participants were seated facing a podium (similar to that used by Snedeker & Yuan, 2008 and Arnold, 2008), and presented with four objects. Two of the objects were phonologically similar in word onset. Once the child was fixated to the podium (where a camera was positioned to capture eye movement), a set of two sentences was played. Stimuli were modeled after those used in Dahan et al.'s 2002 study. The first sentence instructed the participant to manipulate an object on the podium (e.g., "Put the candy on the square."). Once the participant had manipulated the object, the second instruction played, containing either an unaccented target, or an accented target. The target words in the second sentence contained the same phonological onset as the word manipulated in the first sentence, leading to a period of temporary ambiguity during the onset of the target word in the second sentence (e.g., "Now, put the CANDLE/CANDY on the triangle."). Eye movements were tracked while the stimuli were presented to the participants.

**Hypothesis**

If participants with ASD are sensitive to the discourse function of prosodic cues, then they will demonstrate differential fixation patterns based on the interactions between accent and discourse status. That is, a higher proportion of fixations will be in response to the accented targets during the new condition, and the unaccented targets during the

given condition. This would suggest that individuals with ASD are able to perceive and utilize the prosodic cue of contrastive stress during the interpretation of a contextual communication task. If they are unable to do so, then no interaction between discourse status and presence or absence of accent is expected. This would be consistent with the fact that neurological and behavioral differences have been found in other areas of receptive prosody between individuals with TD and individuals with ASD (Korpilahti et al., 2007; Peppé et al., 2011; Jarvinen-Pasley et al., 2008). This is also predicted since pragmatic skills are typically noted as areas of weakness (e.g., Baron-Cohen, 1988; Wetherby, 1986; Eales, 1993; Volden & Phillips, 2010), and a task requiring the perception of the prosodic feature of accent, and the subsequent interpretation as a pragmatic function of prosody (interpreting speaker intent and emphasis) is expected to highlight deficits in this population. This lack of interaction between discourse status and accent would suggest that individuals with ASD are able to perceive some prosodic cues (perhaps those related to grammatical interpretations (Diehl et al., 2008)), but not prosodic cues related to pragmatic interpretations.

Additionally, relationships between expressive and receptive prosody skills, as well as relationships between receptive prosody skills and set-shifting abilities, receptive language skills, and cognitive skills, will be explored. Implications for intervention will be discussed.

**Methods**

**Participants**

Thirty-one individuals between the ages of 7 and 21 participated in the study (18 individuals with ASD, 13 individuals with TD). Participants were administered the

Receptive Language Index (RLI) subtests of the Clinical Evaluation of Language

Fundamentals Fourth Edition (CELF-4; Wiig, Secord & Semel, 2008) to assess receptive

language skills, and the Kaufman Brief Intelligence Test Second Edition (KBIT-2;

Kaufman & Kaufman, 2004) to assess verbal and nonverbal intelligence. Participants

were also administered a modified version of the Wisconsin Card Sorting Test (WCST;

Berg, 1948) as a measure of cognitive flexibility and set-shifting ability. Participants

were required to score above 1.5 SD below the mean (standard scores of 78 and above)

on the IQ composite score on the KBIT-2, and the RLI of the CELF-4 to be considered

eligible to participate. Performance within this range indicates that intellectual

impairment or significantly below-average receptive language skills will not confound

performance on the task. Four participants with ASD were excluded because of scores

below 1.5 SD below the mean on the KBIT-2. One participant with ASD was excluded

because he was bilingual. One participant with ASD was excluded because of a recording

error, resulting in her TD match being excluded as well. Two individuals with ASD

participated in the experimental session twice because of a recording error during the first

attempt. Prior to the second attempt (approximately 2-3 weeks after the initial attempt),

the participants had not received an explanation of the study, nor were they aware that

eye fixations were being recorded.

The final sample consisted of 12 children and adolescents with ASD ranging from

8 years, 11 months to 17 years, 8 months of age (mean age 12 years, 8 months, SD 2

years, 7 months), and 12 children and adolescents with TD ranging from 7 years, 4

months to 20 years, 9 months of age (mean age 10 years 10 months, SD 4 years, 11

months). This age was chosen since individuals with TD of at least age 9 years and older

have demonstrated utilization of prosodic cues in spoken language interpretation (Snedeker & Yuan, 2008; Diehl et al., under review; Arnold, 2008). Children were also chosen to be above age 7, as they were required to sit and sustain attention to a task for approximately 25-30 minutes at a time. Furthermore, the CELF-4 is standardized for individuals up to age 21, resulting in the receptive language skills of all of the participants measured on the same assessment.

Participants were matched on gender and verbal IQ scores as measured by the Verbal IQ index of the Kaufman Brief Intelligence Test - Second Edition (KBIT-2; Kaufman & Kaufman, 2004). Participants in the ASD group scored an average raw score of 69.5 (SD 15.68) on the KBIT-2 Verbal IQ index, and participants in the TD group scored an average raw score of 68.6 (SD 13.68) on the same measure, with no significant difference in scores between groups (t(11)-.505, p>.05). Average performance on the KBIT-2, CELF-4, and WCST is displayed in Table 1. Ten males and two females participated in each group. All participants were native English speakers, hearing at least 90% English throughout childhood. Participants were recruited through community programs for children with and without special needs, the University of Maryland Hearing and Speech Clinic, fliers, and through word-of-mouth.

Table 1.

*Pre-Test Measure Performance: Standard Scores*

| | RLI: CELF-4 | *<br>Verbal IQ Index:<br>KBIT-2 | Non-Verbal IQ<br>Index:  KBIT-2 | Composite Score:<br>KBIT-2 |
|---|---|---|---|---|
| **ASD** | 101.17 (SD 15.71) | 103 (SD 18.93) | 105.08 (SD 15.63) | 104.92 (SD 17.70) |
| **TD** | 109.83 (SD 13.48) | 117.92 (SD 11.96) | 105.5 (SD 16.64) | 113.83 (SD 11.80) |
| **Overall** | 105.5 (SD 14.98) | 110.46 (SD 17.26) | 105.29 (SD 15.79) | 109.38 (SD 15.40) |

*\* = Significant differences: KBIT Verbal (t(22)=-2.31, p<.05)*
Note: Standard scores reported for CELF-4 and KBIT-2 measures (all mean=100, SD=15).

*Pre-Test Measure Performance: Raw Scores*

| | +<br>Verbal IQ Raw:<br>KBIT-2 | Non-Verbal IQ<br>Raw:  KBIT-2 | *<br>Modified WCST | *<br>SCQ Scores |
|---|---|---|---|---|
| **ASD** | 69.50 (SD 15.68) | 34 (SD 5.70) | 6.75 (SD 2.00) | 22.17 (6.77) |
| **TD** | 68.58 (SD 13.68) | 29.75 (SD 8.29) | 4.5 (SD 2.94) | 2.9 (SD 1.66) |
| **Overall** | 69.04 (SD 14.39) | 31.88 (SD 7.29) | 5.63 (SD 2.72) | 13.41 (SD 11.03) |

*\* = Significant difference: SCQ (t(20)=8.75, p<.0001) Modified WCST (t(22)=2.19, p<.05)*
*+ =Participant groups were matched on this measure (Raw Verbal IQ: KBIT-2)*
Note: Average raw scores for the CELF-4 were not reported since the Receptive Language Index is composed of varying subtests depending upon participant age; average total number of completed sets reported on the Modified WCST.

Parent report of diagnoses by certified professionals indicated that within the ASD group, four participants were diagnosed with "PDD-NOS", two participants were diagnosed with "PDD-NOS/ASD", three participants were diagnosed with "Asperger's Syndrome", and three were diagnosed with "Asperger's Syndrome/ASD." Parents of all participants under the age of 18 were given the Social Communication Questionnaire (SCQ; Rutter, Bailey & Lord, 2003) as an independent confirmation of the diagnosis of

the participant. The SCQ is a brief screening measure in which parents circle "yes" or "no" to a series of questions regarding social and communication development. It is recommended that any individual who scores 15 or more points "Fails" the screener and should be referred for a full evaluation for an ASD (Rutter, Bailey & Lord, 2003). Participants in the TD group scored from 0-6 points (mean 2.9 points, SD 1.66). Two participants in the TD group were not given the SCQ, as they were over the age of 18 and parents were not available to complete the questionnaire. Participants in the ASD group scored from 11-32 points (mean 22.17 points, SD 6.77). Participants in the ASD group scored significantly higher (i.e., showed more symptoms of ASD) than participants with TD on this measure (t(20)=8.75, p<.0001). Two participants with Asperger's Syndrome scored slightly below the cut-off score (11 points and 14 points). This is not surprising, however, as individuals with Asperger's Syndrome, by nature of their disorder, have no developmental clinical impairments in language and cognitive skills (American Psychiatric Association [DSM-IV-TR], 2000) while the SCQ includes Communication items. Additionally, on a measure of SCQ validity, when participants were divided by their diagnosis, individuals with ASD ("including Asperger's Syndrome and PDD-NOS, but not Rett syndrome or childhood disintegrative disorder") received a mean score of 13.7 (SD 7.0) (p. 21, Rutter, Bailey & Lord, 2003). This indicates that individuals with ASD who are high functioning (such as individuals with Asperger's Disorder and PDD-NOS) may score slightly below the prescribed 15-point cut-off, and still be accurately diagnosed.

**Stimuli**

**Apparatus**

The participants were tested in a quiet space in their home, or in a testing room at the University of Maryland. Each participant was seated facing an apparatus similar to that utilized by Arnold (2008) and Snedeker and Yuan (2008). The apparatus consisted of a podium with four spaces in each corner for the objects. Four shapes (star, circle, square, triangle) were positioned in the upper and lower corners of the podium, with round magnets in the centers to serve as locations for the objects to be placed. The shapes and center magnets were painted white, and the shapes were outlined in black. A hole in the center contained a video camera. The camera was positioned to capture the participants' eyes during the experiment. The podium was constructed of wood and covered with white shelf liner. The base measurements were approximately 2 feet by 3 feet. The face of the podium intersected with the base at an approximately 45-degree angle. See Figure 1 for an image of the podium.
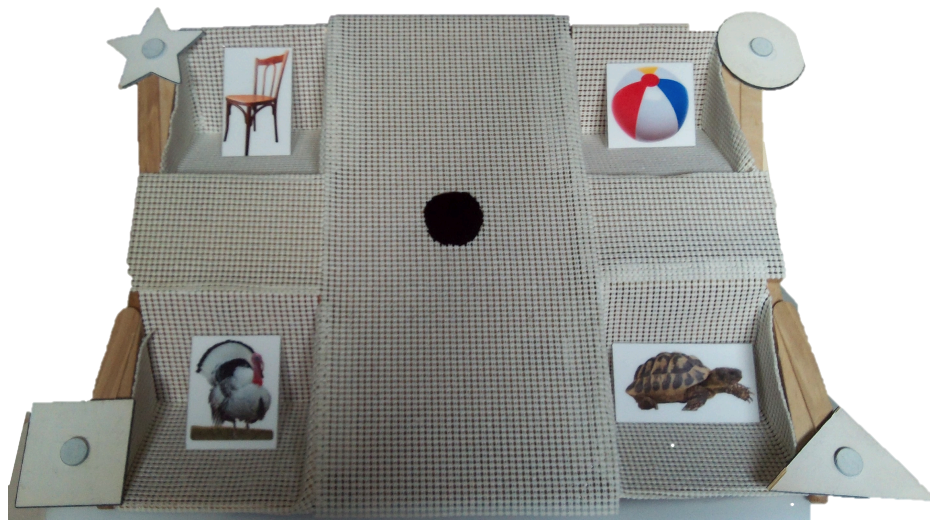


*Figure 1.* Image of the podium holding stimulus objects in each quadrant, with magnetic shapes in each corner.

An additional video camera was positioned at an angle behind the participant, to capture

the movements of the participant interacting with the objects on the podium. This allowed

for any experimenter errors (such as object placement errors) or participant errors (such

as object movement errors) to be recorded and excluded from the analysis.

**Stimulus Items**

Pictures of eight objects familiar to children and adolescents were utilized as

critical items (i.e., "targets" mentioned in the instruction or "competitor" objects with

similar onsets) in the study. These items were a bell, bed, sandwich, sandal, turkey, turtle,

candle, and candy. Items were selected to be word pairs with similar onsets, and were

originally utilized in Dahan et al.'s study (2002). Eight additional objects were used as

"distractors." These items included a bowl, apple, fork, tree, cup, shoe, chair and ball.

Each set of target items (i.e., candle/candy, sandwich/sandal, bell/bed, turtle/turkey), was

matched with a set of distractor items (see Table 2 for the list of set items). This was

designed to limit predictions of the participants, since the two target items were already

required to always appear together by nature of the task. Having the sets of four items

always appearing together led to a greater consistency of trials, and aided in smoother

transitions between trials.

*Table 2. Stimulus Item Sets*

|  | Set 1 | Set 2 | Set 3 | Set 4 |
|---|---|---|---|---|
| Target Items | *Candy* *Candle* | *Turkey* *Turtle* | *Bed* *Bell* | *Sandal* *Sandwich* |
| Distractor Items | *Bowl* *Apple* | *Ball* *Chair* | *Cup* *Shoe* | *Fork* *Tree* |

The images of objects ranged from 2-5 inches in height, and were shown in color.

The object images were laminated using 10-millimeter lamination pouches, creating firm

object images for the participants to manipulate. During a pre-trial phase, each participant

demonstrated familiarity with the objects by pointing to each when the items were said

aloud. Two participants preferred to label the objects themselves, rather than point to the

objects. If an error was made, the correct object label was told to the participant.

Following the corrections, the participants were asked to point again to erred items, and

all participants pointed with 100% accuracy. This pre-trial phase ensured that participants

were knowledgeable of the vocabulary, and provided a preview of all of the items to

increase familiarity with each image prior to the test-phase.

      A 2x2 design manipulating the presence of contrastive stress and reference to the

new or given referent was utilized in this study. Pairs of sentences were constructed using

a similar design as Dahan et al. (2002). The instructions consisted of a simple two-step

command involving moving the objects in the quadrants of the podium to the shape

locations on the upper and lower corners of the podium (e.g. "Put the sandwich on the

star. Now, put the SANDWICH on the square."). In each instruction set, the first

sentence was prepared with no specific accent pattern; however, a high boundary tone

was utilized at the end of the first instruction sentence, indicating that the second

sentence was in some way related to the first (Pierrehumbert & Hirschberg, 1990). The

second instruction was produced in one of four ways (stressed/accented words are

indicated by all capital letters, "target words" are underlined): given-unaccented target

(e.g., "Put the candy on the star. Now put the candy on the circle."), given-accented target

(e.g., "Put the candy on the triangle. Now put the CANDY on the square."), new-

unaccented target (with similar onset) (e.g., "Put the candle on the circle. Now put the

candy on the triangle."), or new-accented target (with similar onset) (e.g., "Put the candle

on the square. Now put the CANDY on the star." (Dahan et al., 2002). In addition to these test conditions, non-test conditions (e.g., Put the candy on the square. Now put the apple on the circle.") were included to, as Dahan et al. (2002) proposed, prevent against listener expectations of focusing on objects that started the same way. For a complete list of stimuli sentences and object placement locations, see Appendix A.

The spoken sentences were recorded by a female speaker of American English, using a Shure SM81 microphone and a Mackie 1202 VLZ mixer/amplifier, recorded at a 44.1 kHz sampling rate in a sound-treated booth. As in the Dahan et al. (2002) study, the second instruction in the set in the accented condition contained a pitch accent on the target word, with no accent on the preposition (e.g., "Now put the CANDY on the square"). In the unaccented condition, the pitch accent was placed on the prepositional phrase, with no pitch accent placed on the target word (e.g., "Now put the candy ON THE TRIANGLE"). The presence of accent on the prepositional phrase was not of particular interest, as looking times were expected to be influenced primarily by the accent or lack of accent on the target word – particularly during the phonological ambiguous onset portion of the target word, which occurs prior to the stressed prepositional phrase.

Length, loudness, and pitch of the target words were analyzed to ensure that the target words were appropriately accented/unaccented through the use of PRAAT (Boersma & Weenink, 2011). Target words in the stressed condition were significantly longer (stressed M=.62 sec, SD=.13; unstressed M=.46 sec, SD=.12; t(30)=3.40, p<.01), had greater average word amplitude (stressed M=-13.79dB, SD=1.87; unstressed M=-20.24dB, SD=3.4; t(30)=6.72, p<.0001), higher average word pitch (stressed

M=301.55hz, SD=18.11; unstressed M=247.28hz, SD=10.45; t(30)=10.38, p<.0001), and higher maximum word pitch (stressed M=379.85hz, SD=.31.76; unstressed M=291.28hz, SD=27.97; t(30)=8.37, p<.0001).

**Blocks/Orders**

**Balancing/randomization – blocks.** Stimuli sentences were divided into eight blocks. Each block consisted of six trials, one in each of the four conditions of interest, with two additional "filler" trials. The eight blocks were randomized into 12 orders. Each order was completed by one individual with ASD and one TD individual.

**Filler trials.** Filler trials instructed the participants to move the distractor items, and were included to reduce participant predictions that the objects to be mentioned will always be the objects with similar onsets. For each target pair (e.g., candle/candy or sandwich/sandal), four filler sentences were utilized. Within the sets of four, at least three out of four possible conditions were presented with distractor words included rather than target words. Two of the four instructions within the sets included a target object mentioned in the first sentence, followed by a distracter object mentioned in the second sentence. The other two filler sentences within each set included distractor objects in the first sentences, with either target objects, or distractor objects in the second sentences.

**Balancing/randomization – objects.** Blocks were balanced such that each of the conditions appeared equally as often in the first half of all of the blocks as in the second half. Within the visual array, the object in the first sentence and the target in the second sentence occurred equally often in each position (i.e., upper right corner, upper left corner, lower right corner, lower left corner). Target items moved in the first instruction sentences were never placed on the shape in the same quadrant as the eventual "target" or

"competitor" (i.e., the object in the display with a similar onset as the target). Instructions were designed in this manner because if a target and competitor were both located in the same quadrant during the second instruction sentence, looks to that quadrant would be unable to be differentiated between target and competitor looks during coding.

**Balancing/randomization – locations.** The locations were designed such that within an object set (e.g., sandwich/sandal), within the first instruction, the same target was not placed on the same location more than two times. Nor, within the second instruction, was the same target placed on the same location more than two times. This was designed to limit the ability of participants to predict the object stated in the second sentence based on factors within the first sentence (i.e., predict based on a pattern of location placements if the word will be a distractor or a target word).

**Pre-Test Measures**

**KBIT-2 and CELF-4 administration.** During the first visit, participants were administered the KBIT-2, followed by the RLI of the CELF-4, a modified version of the WCST, then a picture description task. All participants were tested by the same examiner, a speech-language pathology graduate student. The examiner followed standard administration guidelines for the KBIT-2 and CELF-4. Breaks were given as needed.

**Modified WCST administration.** During the WCST, participants were given an instruction sheet (see Appendix B), and given a brief demonstration of the task. Participants were required to sort 64 cards with different shape options (triangle, circle, cross, star), different color options (red, blue, yellow, green), and a different number of shapes (one, two, three, four). The participants were directed to sort the cards based on color, shape, or number, that the examiner would say "yes" or "no" to indicate if the card

was sorted into the correct category, and that the category may change as the game progressed. Categories were pre-determined in a random order, and once the participant sorted five consecutive cards into the correct category, the category was changed. Scores were determined by the total number of categories that the participant completed. Participant scores ranged from zero and nine sets completed (M= 5.79, SD= 2.73).

**Picture description task.** A picture description task was completed by the participants to acquire a rating of the participants' expressive prosody. Participants were given a copy of the "Cookie Theft" picture from the Boston Diagnostic Aphasia Examination (BDAE; Goodglass & Kaplan, 1983). The participants were instructed to take a few minutes and tell the examiner everything they saw happening in the picture. They were also instructed that follow-up questions might be asked. The follow-up questions included topics such as, "What time of year do you think it is?" and "What do you think is going to happen next?" These questions served to elicit a longer speech sample from the participants, and obtain a more conversational sample of their speech. Participants were also asked, "Is the girl stealing the cookies?" in an attempt to elicit expressive contrastive stress (i.e., a response of "No, the BOY is stealing the cookies.").

Following the picture description task, six judges blind to the factor of "group" each rated the naturalness of all of the participants' expressive prosody samples on a scale of 1-10 (with 1 being the least natural and 10 being the most natural). Raters were instructed to, "Please rate the naturalness of the participants' prosody (consider pitch, loudness, duration, phrase breaks, etc.)." The raters had either recently received a Master's of Arts degree in Speech Language Pathology, or were in their final semester of their graduate program in speech-language pathology. These judges were specifically

chosen, as it is more likely that individuals with clinical experience and knowledge of typical and atypical speech and language development would have the ability to parse prosodic differences from concomitant articulation disorders and voice disorders present in the sample. Participants in the ASD group scored an average rating of 5.49 (SD 1.71), while participants in the TD group scored an average rating of 6.65 (SD 1.03). A one-tailed Mann-Whitney U test was performed on the data, since ordinal ratings were analyzed, and since it was expected that the average prosody of the participants with ASD was poorer than the prosody of TD individuals. Participants in the TD group were determined to have significantly more natural prosody than the participants with ASD (U(22)=40.5, p<.05), although these differences were relatively slight. It is anticipated that with longer, more conversational speech samples, even larger differences in expressive prosody would be apparent.

**Experimental Session**

**Design.** The pre-test sessions and experimental sessions were held on separate days, or with a break (of at least an hour) between sessions to limit fatigue and distractibility. First, the participants engaged in a preliminary phase to become acquainted with the images used in the study and to confirm that the participants were comfortable with the object labels. The participants were first presented with all sixteen items (eight target items and eight distractor items), and instructed to point to each item given verbal instructions, (i.e., "Point to the ____."). Three participants preferred to independently label the objects, rather than point. Any errors in producing or pointing during the first attempt were corrected, and all participants performed the task with 100% accuracy on their first or second attempt.

Following this preliminary step, participants were seated facing the test apparatus, with the center of the podium (the camera location) at eye-level. Both the control group (TD participants) and the experimental group (participants with ASD) underwent the same procedures. The sound files were organized and played through PsyScope (Cohen, MacWhinney, Flatt & Provost, 1993). The program advanced to play the next stimuli when a key on a MacBook Pro laptop was pressed. The participants were instructed that they were going to follow instructions presented from a laptop computer and move pictures of objects onto different shapes on the podium. Participants were provided with a spoken sample instruction set (e.g., the experimenter said aloud "Put the ball on the circle. Now, put the chair on the star."), with a demonstration of the object movements. Sample sentences included at least one distractor item.

Then, the first four objects were placed on the podium by the experimenter. The experimenter referred to a list of items and positions to appropriately place each item in the correct position throughout the experiment.

Stimuli were played to the participant either aloud, or through Sennheiser HD201 headphones. The stimuli were played through headphones to sixteen participants, while eight heard the stimuli through computer speakers. This variation in presentation method is because of technical difficulties (7) or participant preference (1). When headphones were used, a splitter was attached to the computer headphone jack to provide the output through both the participant headphones, and Logitech Z130 external speakers. This allowed the sound to be played aloud so the cameras could capture the audio for later coding and analysis.

The first sentence in the instruction set was performed by the participant, then the next instruction was played and the participant performed that action. Between each instruction set, the experimenter removed the items and replaced them with items for the next instruction. Once the object pictures were set, the experimenter presented the next instruction by pushing a key to advance the PsyScope program. Each trial lasted between 45 and 75 seconds, with this portion of the study taking approximately 20-35 minutes. If requested, a break was provided. Participants were debriefed by asking them if they noticed anything about the way the instructions were said. They were told that some words were said with more emphasis, and that we were interested in how that affected which object they looked at when they heard the instruction. Parents were provided a full explanation of the study. Participants under age 18 were provided with a small prize for participating in the study, while participants over 18 were provided with monetary compensation ($8/hour) for participating.

**Coding and Analysis**

Three sets of video/audio recordings were obtained for each participant – one set from the pre-test measures and two from the experimental session. Any unclear participant responses during the pre-test measures were marked during the test session and later addressed by referring to the recordings of the pre-test session. A research assistant re-scored all test protocols for accuracy in totaling items and calculating standard scores. From the experimental session, both "eye-gaze videos" (i.e., videos captured by the camera set inside the podium which captured the participants' eye movements), and "backlog videos" (i.e., videos captured by the camera set behind the

participant which captured the participants' manipulations of the objects on the podium) were analyzed.

**Backlog coding.** First, backlog videos were uploaded onto a Macbook Pro laptop computer. A master worksheet of participant orders and object positions was utilized to ensure that for each trial, the objects were positioned accurately and the sound files were played correctly. Participant errors in placement or object utilized, as well as experimenter sound play or placement errors were recorded.

**Eye-gaze audio coding.** The audio from the eye-gaze videos from the experimental sessions was uploaded into Audacity (Mazzoni & Dannenberg, 2000), and the time of onset of the initiation of the second sentence (i.e., the onset of the word, "Now") was determined and recorded for each target trial for each participant. Times were recoded in an hours-minutes-seconds-frames format.

**Eye-gaze video coding.** The eye-gaze videos were then uploaded to a computer and coded frame-by-frame using *VCode* (Hagedorn, Hailpern & Karahalios, 2008). The videos were coded with the sound off, to prevent bias in the coding. Coding began at the time of the onset of each target trial (as determined by the Audacity audio coding). The eye gazes were coded as being fixated on the center, away from the display (i.e., the participant's eyes were fixated to an object away from the podium), off screen (i.e., the participant's eyes were out of the scope of the camera and therefore unable to be coded) or on one of the quadrants of the podium – upper right, upper left, lower right, or lower left. This resulted in text files containing series of eye-gaze codes in milliseconds for each participant.

The experimenter coded all of the eye-gaze videos in the study. A second individual re-coded all critical trials for four of the twenty-four participants, with the sound off as well. The data was divided into frame-by-frame fixations for each participant. Looks away, off-screen, and center were collapsed into a single "away" category, since these fixations are all away from the objects (or unable to be determined in the case of "off-screen" marks). Each frame was determined to be a match between the two coders, or not a match. It was determined that the two coders matched on 87.8% of the frames. The codes from the primary coder were utilized in the analysis.

**Post-coding processing.** Eye coding files were converted from milliseconds to frames, and divided into separate trials. Audio trial onset times for each participant were combined with that participant's eye fixation trials, to remove eye codes marked prior to the onset of each trial. R syntax (R Development Core Team, 2006) was utilized to separate the eye codes into separate areas of fixation per frame, based on the reference list of object locations per trial. As a result of the R syntax, participant fixations for each trial were then in reference to target, competitor, and distractors for the various conditions, rather than simply gazes to locations on the podium.

**Time Regions.** Time-course information is necessary in the interpretation of the eye-gaze fixation data. For this reason, the stimuli sentences were analyzed and divided into two time regions of interest: the onset (e.g., "Now, put the"), and the disambiguation period (e.g., "SANDWICH on the square."). Traditionally, a delay of approximately 200ms after the onset of the spoken word is sufficient to allow for eye saccade movements (Allopena, Magnuson & Tanenhaus, 1998). For that reason, 200ms were added to the onsets of each region of interest, so the labeled regions reflected the

participant eye movements from those regions. However, in addition, the critical region was extended to the entire duration of the target word (e.g. SANDWICH) rather than just the phonologically ambiguous onset portion of the target word (e.g., SAND-) in this analysis to account for slower fixations demonstrated by children and individuals with ASD. Arnold (2008), for example, analyzed a critical region of 500-1200ms post onset of the target word to account for the slower fixations of the child participants compared to the adult participants. The fixation proportions over the entire region were expected to reflect the differences in prosodic processing.

Following Arnold's findings, if participants are able to interpret and utilize the prosodic cue of contrastive stress within a contextual communication task, it is predicted that during the "given" conditions, they will look to the target longer and faster upon hearing the unaccented targets than the accented targets (2008). This fixation difference is expected to occur during the target word disambiguation period, beginning at the onset of the target word as that is the first instance the participants have access to the prosodic cue of contrastive stress to influence their fixations.

The time and location of fixations for each time region were analyzed, to determine if participants looked at objects with the same phonological onset differentially based solely on the interaction between the presence or absence of contrastive stress, and discourse status (mentioned previously in the contextual communication task, or new to the task). A set of 2x2 ANOVAs with the main effects of stress/accent and discourse status were performed for each group. First, analysis of the eye fixations of participants in the TD group will be discussed, to verify replication of prior studies, followed by analysis of the data from the participants in the ASD group.

**Results**

**Action Results**

Results of the placements and manipulation of the objects on the podium were determined. A total of fourteen trials were excluded because of experimenter placement errors and four trials were excluded because of experimenter sound-play errors. Two additional trials were excluded because of environmental interferences, such as a telephone ringing or talking. Therefore, 20 critical trials were excluded because of experimenter errors or environmental interferences.

Twenty-two additional trials were excluded because of participant errors, including errors in "shape" (i.e., the participant placed the correct object onto the incorrect shape), and errors in "object" (i.e., the participant moved the incorrect object). Fifteen of the twenty-two errors occurred on the first sentence in the instruction set, before any influences of prosody can be implicated. Of the seven errors that occurred during the second sentence in the instruction set, six errors were related to "location", while one error was related to "object". No errors occurred in the given-unaccented condition, two errors occurred in the given-accented condition, two errors occurred in the new-unaccented condition, and three errors occurred in the new-accented condition. The one error related to "object" was in the new-unaccented condition, and was an erred substitution of the competitor object for the target object (the felicitous expectation). Five of the 22 participant errors (23%) were performed by individuals in the TD group, while 17 of the errors (77%) were performed by individuals in the ASD group.

Of the 768 critical trials presented to participants, 2.6% were excluded because of experimenter errors and environmental interferences, and 2.9% were excluded because of

participant errors. Overall, 94.5% of the total critical trials were included in the final analysis. Since all of the critical trials analyzed involved the appropriate movement of objects, "performance" in all instances following refers to proportions of eye movements to the target and competitor objects during sentence processing, and not to participant manipulations of the objects.

      **Sentence onset region: TD group.** Eye fixations were coded as either towards the target, competitor, or distractor items, per frame during each condition. For each participant, fixations to the objects were averaged over the entire time region for each trial, and then condensed over each condition for those time regions. Finally, the fixations were condensed by group to create an average looking proportion (as defined by looks to the target divided by total looks) by region and condition for each group. This proportion of target looks served as the performance data for each condition. A 2x2 ANOVA with the main effects of accent and discourse status was performed on the proportion of eye gazes to the target and competitor words by the TD participants during the sentence onset region (prior to the onset of the target word). Although a slight preference to look towards the previously-mentioned object (the object that had been moved in the first sentence of the instruction set) was demonstrated, no significant main effect of accent ($F(1,11) = 1.285$, $p>.05$), or discourse status ($F(1,11) = .797$, $p>.05$), and no significant interaction ($F(1,11) = .282$, $p>.05$) was found; indicating that TD participants did not demonstrate a significant baseline preference during the sentence onset region (See Figure 2).
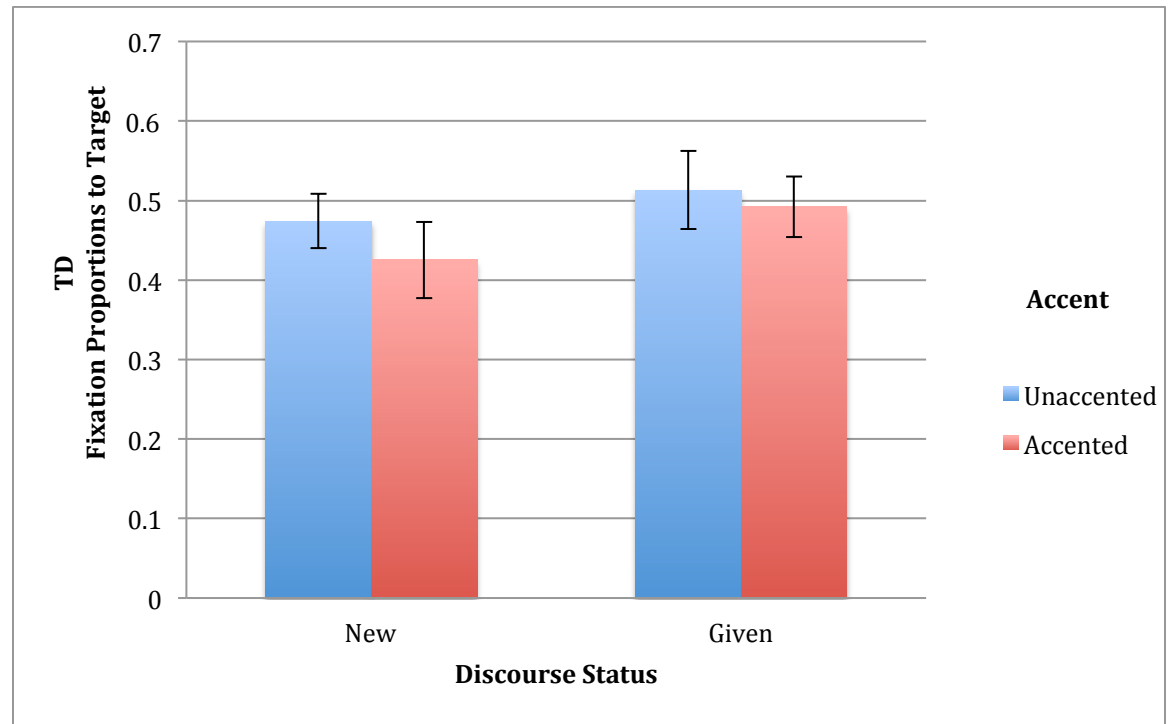
*Figure 2.* TD fixation proportions to the target during the onset of the instruction sentence does not demonstrate a significant preference to look towards objects that are given (i.e., were mentioned in the first sentence of the instruction set).

**Target disambiguation region: TD group.** A 2x2 ANOVA performed on the eye-fixation data from the TD group with the main effects of accent and discourse status indicated a significant main effect of accent ($F(1,11)=21.167$, $p<.01$), and a significant interaction between accent and discourse status ($F(1,11)=5.859$, $p<.05$). The significant main effect of accent was driven by a participant preference to look in higher proportions towards the target upon hearing an unaccented condition than when hearing an accented condition. Follow-up t-tests performed on the critical finding (the interaction between accent and discourse status) indicated that participants looked longer to the unaccented target when it was given, rather than new ($t(11)=-3.260$, $p<.01$), and longer to the target in the given condition when it was unaccented, rather than accented ($t(11)=5.175$, $p<.001$) (see Figure 3). Thus, the TD participants were able to perceive the prosodic cue

of contrastive stress, and utilize that cue within the contextual communication task to predict that an unaccented object often refers to an object that was mentioned previously within the contextual communication task. This finding replicates the previous finding by Arnold (2008) in children who are typically developing, and indicates that the methods used in this study were appropriate to determine the receptive prosody abilities of individuals with ASD as well.
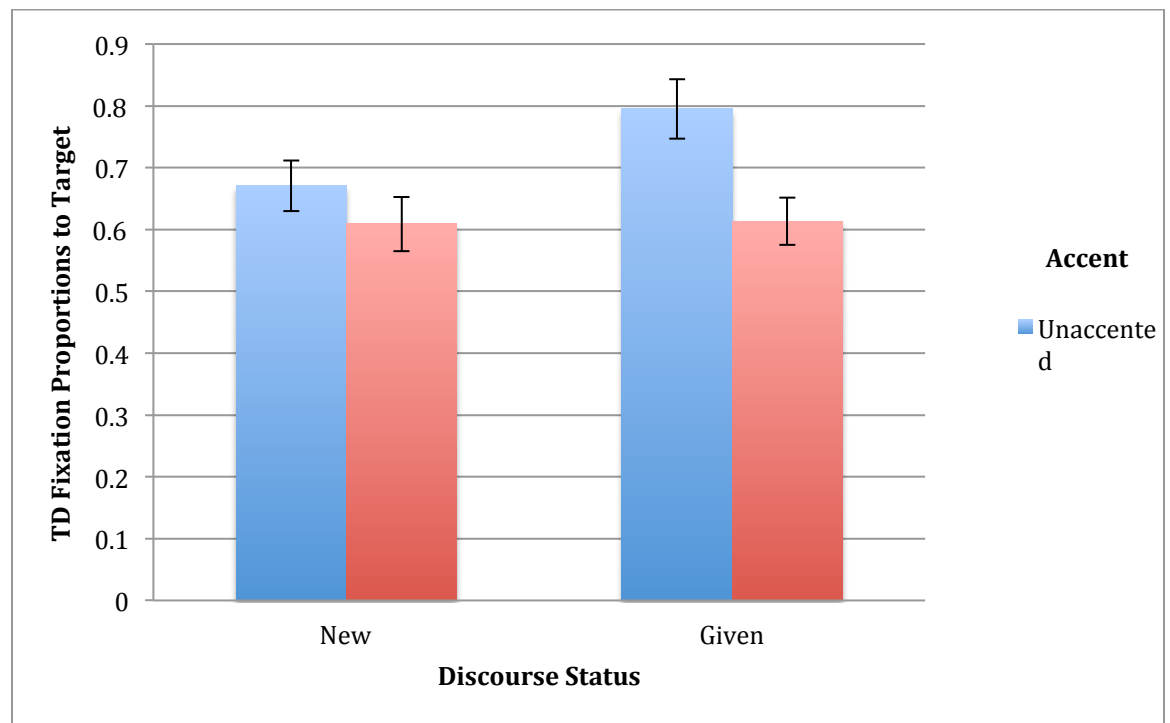


*Figure 3.* Fixation proportions to the target of participants who are TD during the disambiguation portion of the instruction sentence demonstrates a significant interaction between discourse status and accent, such that TD participants fixated to the target in higher proportions in response to a target word in the given condition when it was unaccented.

**Sentence onset region: ASD group.** In the ASD group, during the onset region (before the target word had been heard) a significant main effect of discourse status was found ($F(1,11) = 8.433$, $p<.05$), with no significant main effect of accent ($F(1,11) = .010$, $p>.05$) and no significant interaction ($F(1,11) = 1.826$, $p>.05$). The mean proportion of

looks towards the target object during the given conditions was significantly greater than the mean proportion of looks to the target during the new conditions (.61 and .57 for given conditions, compared to .42 and .45 for new conditions). In other words, after having moved an object in the first sentence, children with ASD showed a bias to continue looking at the same object. Participants with ASD demonstrated a statistically significant preference to look towards given objects (objects mentioned in the first sentence of the instruction), rather than new objects, when comparing looks between the new-accented and given-accented conditions (t(11)=-2.849, p<.05), as well as between the new-unaccented and given-unaccented conditions (t(11)=-2.584, p<.05) (See Figure 4).
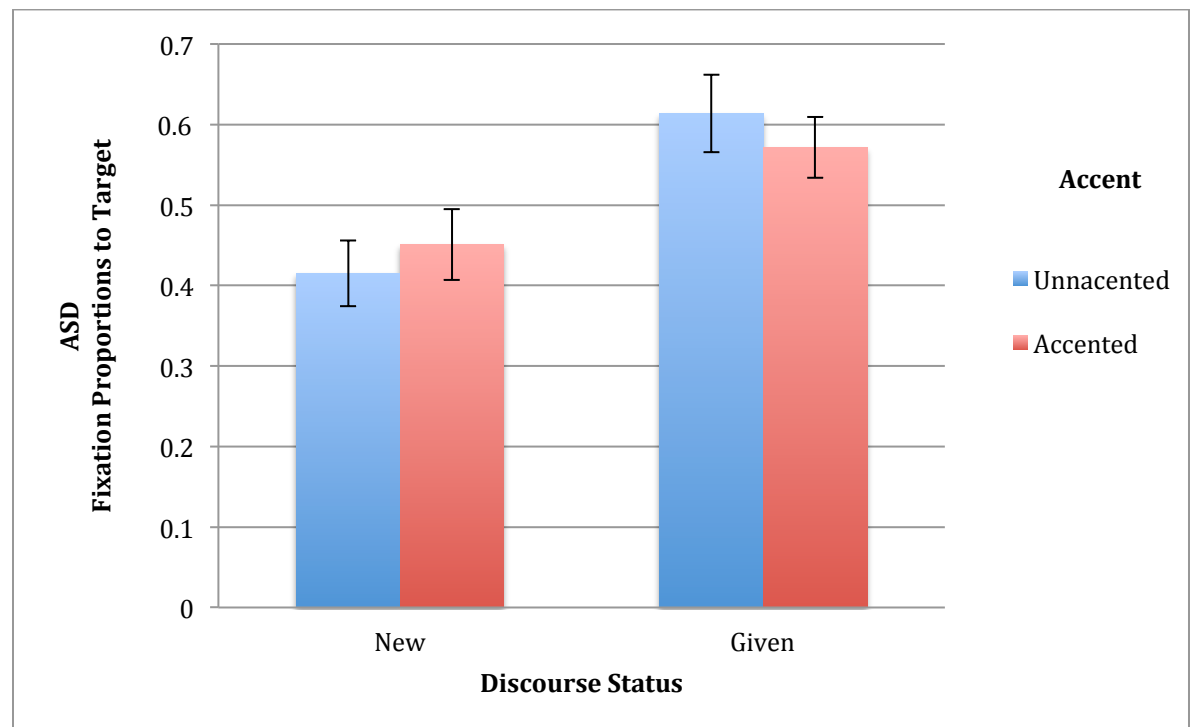


*Figure 4.* Fixation proportions to the target of participants with ASD during the onset portion of the instruction sentence demonstrate a significant preference to look towards objects that are given (i.e., were mentioned in the first sentence of the instruction set).

**Baseline Switch Analysis**

An important effect of the preference of participants with ASD to fixate towards the given object prior to the onset of the target word is its impact on the interpretation of fixations during the target disambiguation region (i.e., the critical region). The preference makes it difficult to determine if significant effects during the critical region are carry-over fixations related to this initial preference, or if they are a true representation of the receptive prosody skills of the participants with ASD. A baseline switch analysis was performed on the data from both the TD and ASD groups, to account for the difference in object preferences between groups even prior to the onset of the target word and the introduction of prosodic factors. This measurement sorted the data into looks "to the target" (i.e., when the participant had been looking at an object other than the target in the previous frame), and looks "off of the target" (i.e., when the participant had been looking at the target prior in the previous frame). The proportion of switches to the target when the participant had been looking away from the target, and the proportion of switches away from the target when the participant had been looking at the target were calculated. The proportion of switches to the target minus the proportion of switches away from the target was calculated to create a "difference score". This indicated the likelihood that the participant would look to or away from the target based on the prosodic cue, rather than a baseline preference. This difference score included information from both when the participants were showing an initial preference to look away from the target, and towards the target. Since the baseline switch difference analysis includes the most representative information from the data acquired, results will be reported below for both the TD and ASD groups. The baseline switch difference was taken in 100ms sections (to capture time

course information) from slightly after the onset of the target word (about 200ms post-onset), until 1200ms post-onset, corresponding with the end of the stimuli sentence.

**TD baseline switch difference analysis.** A 2x2 ANOVA with the factors of accent and discourse status (new or given) was performed on the proportion of switches to the target minus the proportion of switches away from the target for the TD participants. This difference provides the likelihood that given the interaction of the variables of discourse status and accent, the participants will switch to the target. A significant main effect of accent ($F(1,11)=10.576$, $p<.05$) indicated that participants switched to the target more often during unaccented conditions than accented conditions, and a minor, nonsignificant interaction between accent and discourse status was found ($F(1,11)=62.319$, $p=.179$) (see Figure 5). Follow-up T-tests replicated the initial analysis of the TD data, and indicated that participants switched to the target significantly more often during the given condition when the target was unaccented than accented ($t(8)=3.464$, $p<.01$), and switched to the target more often during the new condition when the target word was accented ($t(8)=2.852$, $p<.05$) rather than unaccented. Since this baseline switch difference analysis replicated findings from the initial analysis while correcting for baseline preferences, this analysis was utilized on the data from the participants with ASD as well.
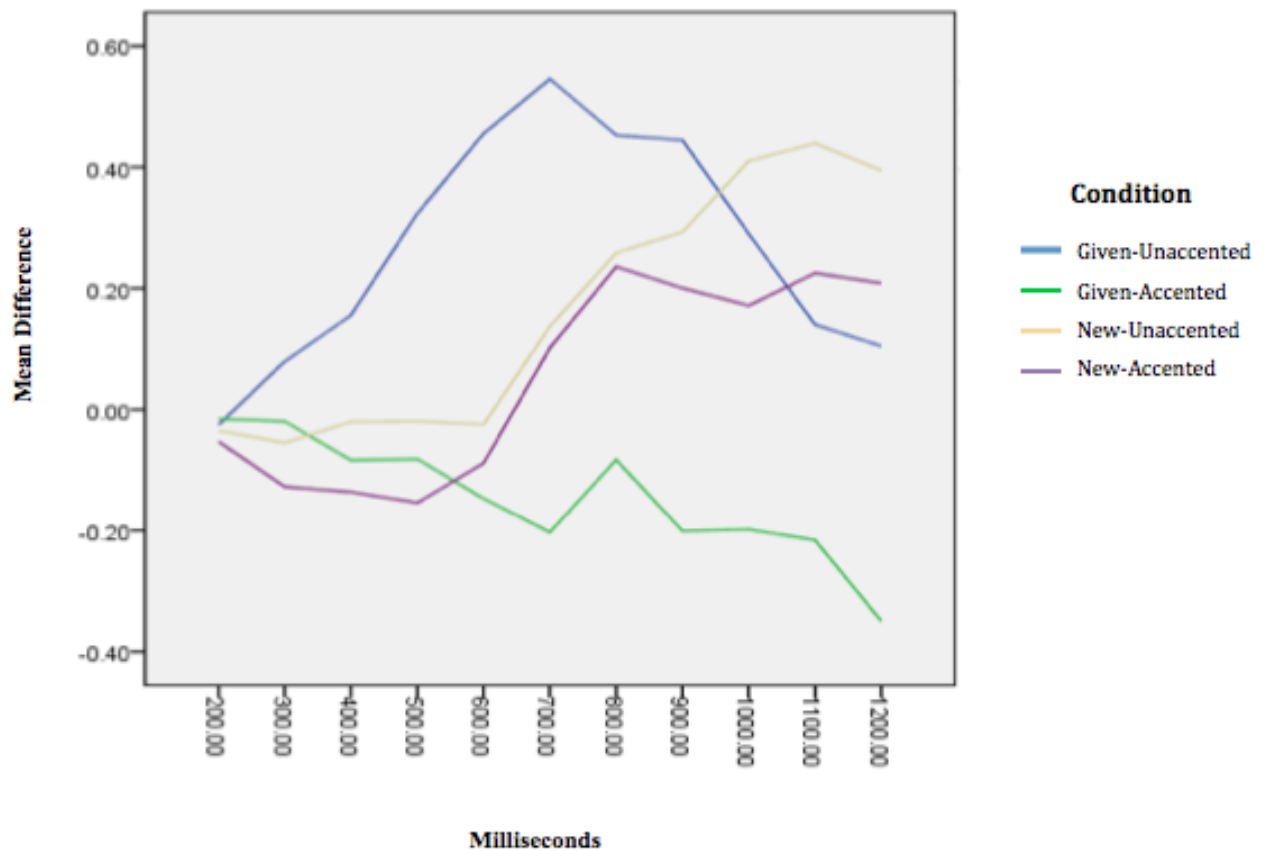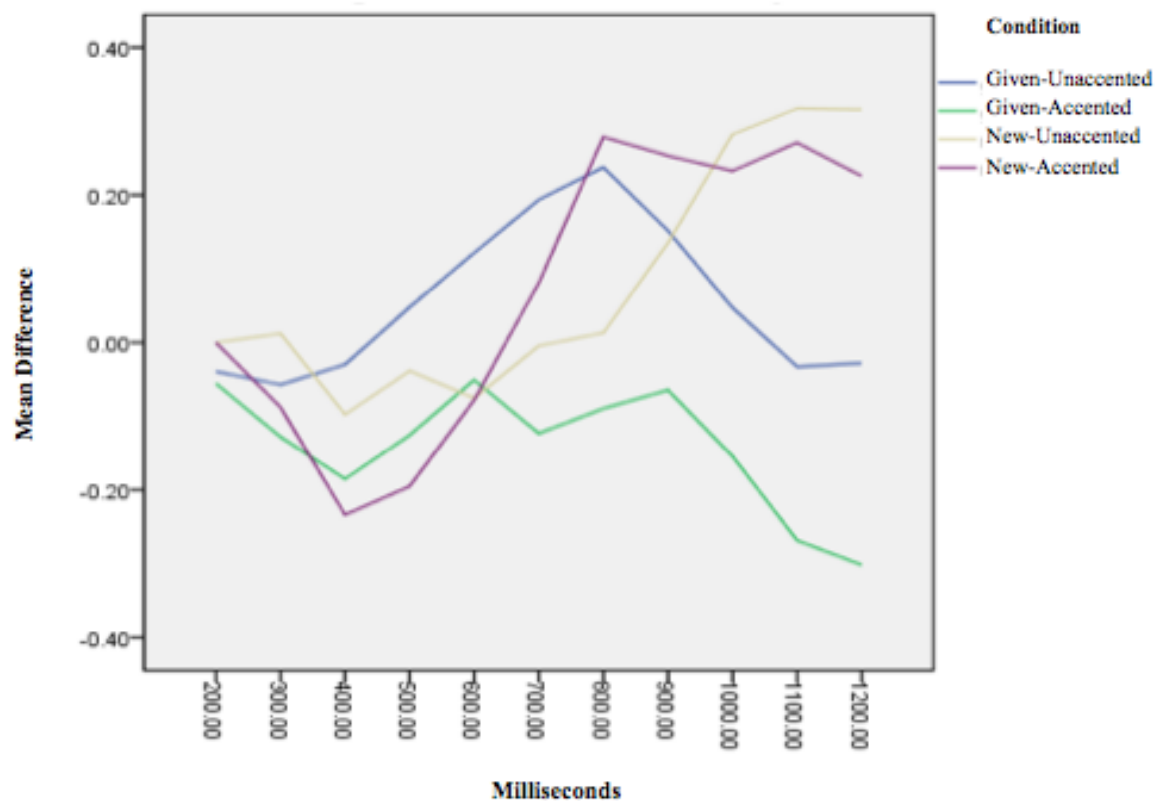
*Figure 5.* Mean difference scores of switches to the target and away from the target demonstrate a higher likelihood of switches to the target during the given-unaccented condition (blue line) than during the given-accented condition (green line) for TD participants.

**ASD baseline switch difference analysis.** A 2x2 ANOVA with the factors of accent and discourse status (new or given) was performed on the proportion of switches to the target minus the proportion of switches away from the target for the ASD participants as well. No significant main effects or interactions were found through this analysis; however, since the critical comparison of accented versus unaccented is anticipated only during the given condition, follow-up t-tests were performed to gain specific information concerning that condition. It was found that like the TD children,

participants with ASD switched to the target significantly more often during the given

condition when the target was unaccented than accented (t(11)=2.406,

p<.05). A slight difference was also found, such that participants with ASD switched to

the target more often during the new condition when the target word was accented than

unaccented (t(8)=1.552, p=.159) (see Figure 6 for a graphic representation of these

findings).



*Figure 6.* Mean difference scores of switches to the target and away from the target
demonstrate a higher likelihood of switches to the target during the given-unaccented
condition (blue line) than during the given-accented condition (green line), and a
marginally higher likelihood was found to participants with ASD to switch to the target
during the new-accented condition (purple line) than the new-unaccented condition (tan
line).

**Comparison of ASD and TD difference of baseline switches.** A 2x2 ANOVA with the factors of group and accent was performed using the difference scores from the baseline switch correction to compare the looking patterns of participants in the TD and ASD groups during the given condition. A significant main effect of accent was found, such that all participants had a higher proportion of switches towards the given target when the object was unaccented rather than accented ($F(1,22)=15.803$, $p<.01$). This indicates that both groups were sensitive to the discourse function of prosody. No significant interaction between accent and group was found ($F(1,22)=.016$, $p>.90$), indicating that during the given condition, ASD and TD participants were both using the prosodic cue of contrastive stress to a similar extent to guide their looks to the correct object.

A 2x2 ANOVA with the factors of group and accent performed on the difference scores during the new condition yielded no significant main effects or interactions. Thus, participants in neither group were able to utilize the prosodic cue of contrastive stress to make predictions within a contextual communication task when a new word was presented.

**Regression**

A multiple regression equation was performed to determine if any of the pre-test measures related to eye-movement performance on the task. Performance was defined as the baseline switch difference score of switches toward the target during the unaccented – given condition, divided by the baseline switch difference score of switches to the target during the accented – given condition. For example, a performance score of .40 could correspond to a switch to the target percentage of 90, minus a switch away from the

target percentage of 10 during the unaccented – given condition, divided by a switch to the target percentage of 30 minus a switch away from the target percentage of 10 during the accented – given condition $(((90-10)/(30-10))*100 = .40)$. This would indicate that the participant switched to the target 40% more during the unaccented – given condition (the felicitous condition) than the accented – given condition. Higher performance scores indicate that the participant showed a stronger ability to use the prosodic cue to guide his or her on-line sentence interpretation.

No significant correlations were found between receptive prosody eye-movement performance and scores on the KBIT-2 Verbal IQ, Non-verbal IQ, Composite IQ Score, RLI of the CELF-4, Modified WCST, or SCQ. Table 3 displays the correlation coefficient values, as well as significance levels. This indicates that none of the factors, including receptive vocabulary scores, cognitive abilities, or set-shifting abilities, were related to performance on the receptive prosody measure. It is important to highlight, however, that the participants in this study were required to adhere to standards of performance on the cognitive and language measures in order to qualify to participate in this study. It is possible that with a broader range of performance abilities on these measures, a relationship would be evidenced between language or cognitive abilities and receptive prosody abilities.

*Table 3. Correlations between Pre-Test Measures and Receptive Prosody Eye-Movement Performance Scores*

| Pre-Test Measure | | Receptive Prosody Eye-Movement Performance |
|---|---|---|
| KBIT-2 Verbal IQ Standard Score | Pearson Correlation<br>Sig. (2-tailed)<br>N | .298<br>.157<br>24 |
| KBIT-2 Non-verbal IQ Standard Score | Pearson Correlation<br>Sig. (2-tailed)<br>N | -.133<br>.536<br>24 |
| KBIT-2 Composite IQ Standard Score | Pearson Correlation<br>Sig. (2-tailed)<br>N | .111<br>.607<br>24 |
| RLI of the CELF-4 Standard Score | Pearson Correlation<br>Sig. (2-tailed)<br>N | .079<br>.714<br>24 |
| Modified WCST Total Categories | Pearson Correlation<br>Sig. (2-tailed)<br>N | -.271<br>.201<br>24 |
| SCQ Total Score | Pearson Correlation<br>Sig. (2-tailed)<br>N | -.255<br>.229<br>24 |

**Relationship Between Expressive Prosody Scores and Receptive Prosody Scores**

A correlation was also performed between the expressive prosody "naturalness" scores of the participants in the ASD group and the receptive prosody performance scores (looks to the target divided by the sum of looks to the target and looks to the competitor). As the expressive prosody scores were ratings, a Spearman's Rank Order Correlation was utilized. The correlation indicated no significant relationship between expressive prosody naturalness ratings and receptive prosody performance (as determined by the difference score between participant switches to and away from the target in the given condition during accented compared to unaccented targets) ($r_S(11) = -.319$, $p = .312$). In fact, a weak, non-significant relationship appeared such that individuals with ASD who had lower expressive prosody ratings actually performed somewhat more strongly on the

receptive task than individuals with higher expressive prosody ratings (see Figure 7).

This finding should be interpreted with caution; however, since the expressive prosody

ratings were measured perceptually, over a very short sample (1-3 minutes), and in a

somewhat unnatural language context (describing a picture).
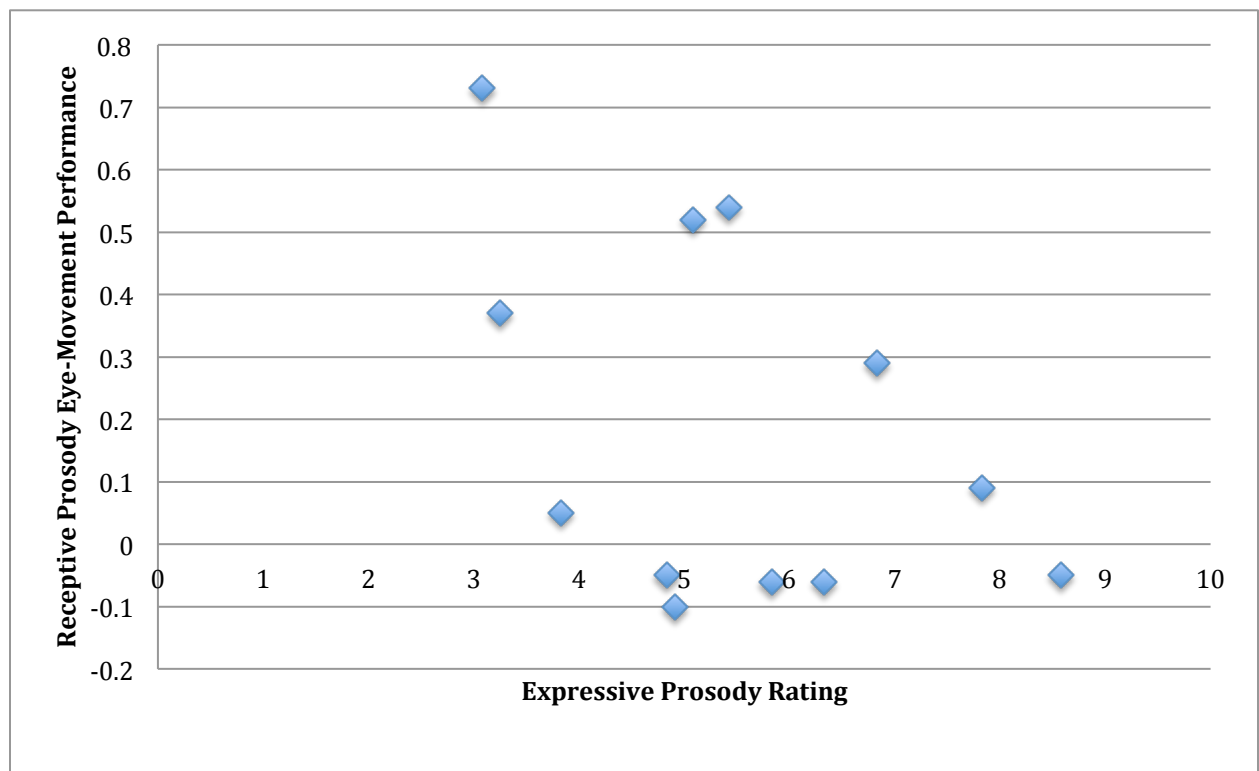


Figure 7. No significant correlation was found between performance on tasks measuring expressive and receptive prosody skills.

## Discussion

The purpose of this study was to determine the ability of individuals with ASD to

utilize the prosodic cue of contrastive stress within a contextual communication task

through the use of the visual world paradigm. Eye fixation proportions to objects with

phonologically similar onsets (target and competitor objects), as well as distractor

objects, were recorded for analysis of the influences of prosody on the participants'

online processing of the instructions. During analysis, fixations were initially divided into

regions of interest – sentence onset (e.g., "Now put the"), target disambiguation period (e.g., "CANDLE"), and sentence ending (e.g., "on the square."). Because of an early preference for participants with ASD to look towards the object mentioned in the previous sentence, a switch analysis was performed to assess the likelihood of participants to switch to, or away from the target after hearing the onset of the target word. This analysis provided an indication of the effects of the prosodic cue of contrastive stress in the online sentence interpretation of the participants with ASD and the TD participants. Our hypothesis predicted that if the comprehension of prosody is served by an ability to generate these features in production, then we might expect expressive differences in prosody between TD and ASD children to lead to differences in comprehension across these two groups. If the comprehension of prosody is independent of production processes, then we might expect similar performance across these groups. Minimally, it was anticipated that the TD children and adolescents in this study would utilize the prosodic cue of contrastive stress during the given condition, by demonstrating an increased proportion of looks to the target in the unaccented, rather than the accented condition, replicating the findings of Arnold (2008). Additional areas of interest included the possibilities of a perseverative effect occurring in the participants with ASD, and a possible relationship between receptive language and cognitive skills, and receptive prosody skills.

**Baseline Differences**

In addressing the main hypothesis, an unexpected difference arose in the fixation patterns of the individuals with ASD during the "sentence onset" region of the stimuli sentence. This difference occurred even before the target word was heard, and reflected a

preference for the participants with ASD to fixate towards the target object during the

"given" conditions. This means that during the "sentence onset" region, participants with

ASD fixated in higher proportions towards the object that had been mentioned in the first

instruction. For example, after hearing "Put the candy on the circle." participants showed

a preference to fixate towards the "candy" while hearing, "Now, put the" even prior to

the onset of the target word.

**Perseveration**

We speculate that this difference might be related to a perseverative behavior by

the participants with ASD. Perseverative patterns have been demonstrated by younger TD

children (4-6 year olds), and slightly older children with ASD (7-12 year olds) in prior

research (Snedeker & Yuan, 2008; Diehl et al., under review). There are two ways that

children could perseverate during the study. They could perseverate by continuing to look

at the object they moved in the first sentence of the instruction set, or they could notice

the condition (i.e., discourse status and presence or absence of accent) utilized in the first

trial, and continue to be biased towards that pattern.

As a result of randomness in the counterbalancing of the study, five of the twelve

participants in each group were exposed to the given – unaccented condition in their first

trial. Participants with ASD may have been demonstrating a perseveration on the first

condition they heard, which for almost half of the participants was the given condition. If

that was the case, it is unsurprising that they continued to fixate to the given word with

higher proportions during the onset of the second sentences.

Therefore, it is valuable to assess whether the participants with ASD in this study

demonstrated perseveration while completing the pre-test measure of the WCST (a

measure of cognitive set shifting). Surprisingly, the participants with ASD actually

scored significantly higher on the WCST than the TD participants (t(22)=2.074, p<.05).

This may be because the non-standardized administration of the WCST may not be an

accurate measure of set shifting abilities in this population. Standard administration does

not include an instruction sheet or a demonstration, which were provided in this study.

Current standard administration also indicates that participants continue the test until six

sets have been completed or 128 cards have been sorted, whereas during this study, the

task was discontinued when the entire stack of cards (64 cards) had been sorted (Heaton,

Chelune, Talley, Kay & Curtiss, 1993 as cited in Ozonoff, 1995). Although the

standardized version of the WCST has been shown to be a reliable measure when utilized

with individuals with ASD, perhaps the participants responded well to this task because

of the modified administration techniques used (Ozonoff, 1995). Ozonoff (1995) found

that when the WCST is administered in non-standardized formats to individuals with

ASD, reliability decreases significantly. Additionally, scoring techniques traditionally

include a measure of how many cards have been placed in a set that has been

discontinued, as a more accurate measure of perseveration, along with the total number of

correct sets. This measure was not included in this administration, and the data collection

sheet utilized during the study did not record the specific error made, rather, just if the

object placement was correct. This non-standardized administration technique may

explain the unexpected results on the WCST. Since it has been found that some

individuals with ASD demonstrate perseveration, it is possible that the preference to

continue to look at the object they had just moved in the prior instruction could be

explained by a perseverative looking pattern regardless of WSCT scores.

**Main Hypothesis**

To evaluate the main hypothesis, an ANOVA was first performed on the target proportion data for the TD group. This analysis yielded results replicating the findings of Arnold (2008). During the given condition, TD participants utilized the presence or absence of stress to aid in the online processing of the instructions. Because of the baseline preference demonstrated by the ASD group, a target switch measure was then computed for each group. This allowed for an interpretable analysis of the interaction between accent and discourse status in the ASD population as well. The target switch resulted in the same pattern demonstrated by the TD group, indicating that this computation retains the integrity of the data in its original form. Through the analysis of switch difference scores it was determined that the ASD group, like the TD group, switched to the target item in greater proportions during the given condition when the target word was unaccented, rather than accented. Therefore, the fixation patterns of both the TD and ASD groups are consistent with prior findings for TD individuals (Arnold, 2008). There was no difference in performance found between groups. This indicates that all participants were able to identify the presence of the prosodic cue of contrastive stress, and utilize that cue to help make predictions within the contextual communication task.

It was initially postulated that individuals with ASD might demonstrate increased difficulty compared to TD individuals when asked to perform a task investigating receptive skills in the area of the pragmatic function of prosody. This was anticipated since individuals with ASD often demonstrate difficulties with pragmatic skills, and because individuals with ASD often utilize atypical expressive prosody. That being said, in past research, individuals with ASD have demonstrated the ability to perform receptive

prosody tasks in the area of syntactic prosody (Diehl et al., under review). The current finding indicates that individuals with ASD are able to perform receptive prosody tasks in the area of pragmatic prosody as well. This, coupled with prior research, means that individuals with high functioning ASDs are actually successful in perceiving a range of prosodic forms and interpreting those forms to serve a variety of functions.

This finding indicates a gap between receptive and expressive skills in prosody for individuals with high functioning ASD. A foundation in perceiving and understanding prosodic cues is likely a prerequisite to effectively utilizing those cues. It seems that individuals with ASD have difficulty monitoring or modulating their own production of prosodic cues effectively. Clinically, this is a positive sign. Knowledge that individuals with ASD are able to perceive and interpret prosodic cues will allow clinicians to focus on targeting prosodic output. It also can guide clinical practices to utilize strategies such as negative practice, or recording and replaying client productions to increase the client's awareness of his or her own prosody production. Expressive prosody production can be shaped progressively to more natural sounding prosody. Of course, because of the high variability of skills in individuals with ASD, it is important to investigate if there are any specific forms or functions of prosody that pose particularly challenging, however; knowledge that even atypical expressive use of those prosodic features may not indicate a lack of understanding of those features may facilitate more efficient remediation techniques.

**Relationship between Expressive and Receptive Prosody**

To address the hypothesis that there might be a relationship between receptive and expressive prosody skills, a Spearman's Rank Order Correlation was performed between

the naturalness ratings of the participants' expressive prosody, and a switch difference proportion to quantify participants' receptive prosody skills. No significant correlation was found between the participants' expressive and receptive prosody skills.

One factor that should be considered is the wide range of prosodic cues present in communication. Prosody serves multiple functions in language, and takes many forms. These forms include, but are not limited to, modifications in pitch, duration, amplitude, rhythm, and pauses or phrase breaks. It is reasonable to consider that the prosodic feature of contrastive stress assessed in this study is not an area of weakness in the realm of prosody for a proportion of the participants in this study. Perhaps the participants with ASD presented with less natural prosody than TD participants in this study because of weaknesses with the use of other prosodic cues such as word duration, or pitch movement.  That would explain why participants seem to have an understanding of contrastive stress used by others, but are still demonstrating atypical prosody themselves. However, contrastive stress serves a pragmatic function within a discourse. It was predicted that prosodic cues serving the pragmatic function might be more challenging than prosodic cues used to serve the grammatical function. Individuals with ASD have now demonstrated the ability to understand prosodic cues serving both functions, while still demonstrating atypical prosody expressively (Diehl et al., under review). This contribution strengthens the case that overall, individuals with ASD are able to understand prosodic cues.

Another interpretation of the results is the idea that high-functioning individuals with ASD have the pre-requisite skills of understanding prosodic cues, but have not been able to make the transition to utilizing those cues in their expressive language, and

perhaps there is something special about this population that makes this transition particularly challenging. In many populations, children demonstrate a gap between expressive and receptive skills, such that receptive skills tend to be stronger than expressive skills. This discrepancy is seen beyond semantic development, but can be seen in articulation, phonology, and morphology development, and in both first and second language acquisition (Gibson, Oller, Jarmulowicz & Ethington, 2012; Snowling & Hulme, 1994; Honig, 2007). So, it may be the case that individuals with ASD have made some progress in acquiring an understanding of contrastive stress, but have not been able to transfer that understanding to the level of production. Peppé and colleagues (2007) also noted a lack of relationship between the expressive and receptive prosody of the participants with ASD during the section related to contrastive stress. But, what is it about production of prosodic cues in this population that makes the jump from understanding to expressing so challenging? Perhaps it more of an issue of output realization, and utilizing nuances of prosody. It is possible that individuals with ASD follow the "rules" of when to use certain prosodic structures, but are slightly off in their execution. This theory was suggested by Grossman, Bernis, Skwerer and Tager-Flusberg (2010), when they also investigated relationships between expressive and receptive prosody in this population using a paradigm similar to the PEPS-C, and found results similar to the findings of the current study. Participants with ASD performed similarly to TD participants in the receptive tasks (investigating the affective and lexical functions of prosody), but demonstrated atypical prosody expressively. In completing the expressive task, individuals with ASD were said to have been "*Accurate* in their productions of lexical stress but still quantifiably *different* from their TD peers." (Grossman et al., 2010,

p.788). Additionally, in a summary of acoustic findings of the use of prosody by individuals with ASD, Diehl and Paul advocated the notion that in individuals with HFA, even when use of prosodic cues appropriately serves its function, there are inherent differences in prosodic production (Diehl & Paul, in press, as cited in Diehl & Paul, 2012).

Consider the realm of musical performance. There are those who do not sing well, because they are not able to produce certain notes, but have awareness of their poor singing. Then, there are those who do not sing well because they are tone deaf, and are not able to perceive their own productions of notes. It follows that individuals with ASD either struggle with expressive prosody because they are performing more like the "tone deaf" group, in that they have difficulty perceiving and monitoring their own productions, or the "production" group, in that they are unable to produce the fine motor movements required to modulate their output effectively. This might be the case, as individuals with ASD have demonstrated fine motor impairments relating to dexterity, coordination, and praxis (e.g., Bhat, Landa, & Galloway, 2011). These motor impairments could impact speech production as well.

This finding may actually be very promising for prosody intervention with individuals with ASD. Realizing that individuals with high-functioning ASD might have an understanding of the rules of using prosody to serve a variety of functions can streamline therapy, and allow speech-language pathologists to focus on fine-tuning the use of prosody to sound more natural.

**Limitations and Future Research**

   **Participant experience in speech-language therapy.** In this study, participants

included were at such an age that most had received speech and language therapy, which

have been confounding variable for their strong performance on the task. Of the families

of the participants with ASD who responded to questions regarding their child's

experience in speech-language therapy, 9/9 participants had received therapy, and it was

noted that prosody might have been explicitly addressed in therapy for at least three of

the participants. Explicit instruction in prosody could have made two contributions to the

interpretation of this data. First, it could have influenced the participants' expressive

prosody skills, therefore creating a sample of participants not entirely representative of

the entire population of individuals with ASD, including those who have not received

speech/language therapy. Or, therapy might not have corrected their aberrant expressive

use of prosody, but might have brought an awareness of prosodic cues in the speech of

others to the individuals with ASD, supporting their ability to notice and interpret

prosodic cues, and leading to strong receptive prosody performance.  In the future, more

specific information regarding the participants' experience in speech and language

therapy would be beneficial.

   **Task naturalness.** Another limitation is the lack of naturalness of the task.

Within this study, prosodic cues were presented at the sentence level, which is more

natural than simply word level prosodic cues; however, the repetition within the task, and

lack of social exchange might have benefitted the individuals with ASD. Individuals with

ASD often benefit from routine and repetition, which this task included. The receptive

prosody portion of the study also was devoid of much social interaction, which may have

led to improved performance as well. As Ozonoff (1995) found that individuals with ASD performed better on the WSCT when presented on a computer screen, as opposed to in an interactive style, perhaps individuals with ASD would have performed differently if the sentences had been spoken aloud, rather than played from a computer.

**Factors related to contrastive stress.** Contrastive stress within a contextual communication task was chosen because it is skill developed at a young age (demonstrated in TD children ages 4-5) (Arnold, 2008), and includes many salient cues such as pitch, duration, and amplitude. Since prosody does not exist in a vacuum, the question arises if other cues present in the sentence, such as duration or amplification of words during the onset portion of the sentence, influenced looks prior to the onset of the target word. If this were the case, it might actually be more impressive, since the participants would be picking up more subtle prosodic cues. That being said, it does not appear to be the case, since there was no significant interaction between accent and discourse status during the onset region of the sentences in either group.

**Diagnosis/age.** This study consisted of a small, relatively homogeneous sample, which may have also contributed to the findings and limited ability to draw widespread conclusions about receptive prosody skills in this population. Individuals who are high functioning and have successfully acquired language are more likely to be able to pick up nuances of the message than individuals who are struggling to learn language. That being said, this topic is most relevant to individuals who are high functioning, since prosody should be addressed once intact language abilities are in place (Bellon-Harn et al., 2007). In future studies, including a larger number of participants within each ASD subtype, individuals with PDD-NOS, Asperger's Syndrome, and HFA, might provide more insight

into the specific strengths and weaknesses of individuals with each subtype.
Additionally, multiple comparison groups, including comparison groups functionally
matched and age matched might provide further insight into the factors influencing
performance by individuals with ASD. For example, the stronger performance on by the
ASD group on the WCST, as compared to the TD group, may be explained by the
chronologically younger age of the TD participants. Including multiple comparison
groups may provide beneficial insight in the future.

Since the individuals in the study were high functioning, some participants at the
time of testing did not demonstrate particularly poor expressive prosody. Utilizing an
experimental group with distinctly impaired expressive prosody would facilitate a more
clear relationship between expressive and receptive prosody skills.

**Expressive prosody ratings.** Additionally, large variability was present in the
expressive prosody ratings of both TD individuals and the participants with ASD. Ideally,
a longer language sample with both perceptual and acoustic measures would have
provided a more precise characterization of the participants' expressive prosody skills.
Also, since the specific receptive prosody skill being investigated in this study was
contrastive stress, it would have been beneficial to measure that specific skill in the
participants' expressive prosody. This was attempted during the language sample, but
when the examiner asked the follow-up question of, "Is the girl stealing the cookies?"
(with the intention of the response, "No, the BOY is stealing the cookies.") most
participants (including TD participants) did not utilize contrastive stress in their
responses. Another task eliciting expressive prosody in this population would have been
more appropriate, such as a sequence story. This would provide the participants with an

opportunity to include "new" and "given" distinctions in their productions. To address

the theory that individuals with ASD have difficulty producing nuances of expressive

prosody, very detailed analysis (ideally an acoustic analysis for research use) of the

participants' expressive prosody would be beneficial. Clinically, an assessment such as

*The Prosody-Voice Screening Profile (PVSP*; Shriberg, Kwiatkowski, & Rasmussen,

1990), which provides guidelines to perceptually assess the factors of rate, stress (with

subcategories such as, "Multisyllabic word stress", "Reduced/Equal Stress", or

"Excessive/Equal/Misplaced Stress"), pitch, quality, and resonance could be beneficial in

highlighting client strengths and weaknesses for use in therapy targeting prosody skills.

      **Design and stimuli presentation.** In future research endeavors, minor changes in

the design and presentation of stimuli might strengthen the study. Regarding the design of

the stimuli, sentences either contained stress on the target word (stressed condition; e.g.,

"Put the CANDLE on the square." or on the prepositional phrase (unstressed condition;

e.g., "Put the candle ON THE SQUARE." modeled after Dahan et al. (2002). A more

felicitous location of stress would have been on the destination shape (e.g., "Put the

candle on the SQUARE."), since the placement location is the factor that changed

between sentences within the unstressed condition (Arnold, 2008).  Additionally, instead

of holding the object placement locations constant for each target word (e.g., all of the

sentences with target word "sandwich" would be moved to the triangle in the first

sentence, and the square in the second sentence – regardless of stress), object placements

were randomized with the exception of controlling that the objects with similar onsets

never appeared in the same quadrant together. This increased the amount of variability in

eye movement, since the participants were always required to search for the new location.

Regarding stimuli presentation, some participants heard the stimuli played aloud from laptop computer speakers, and some wore headphones. The task might have been more challenging, and looking times might have been slightly delayed for participants who heard the stimuli played through the computer speakers because of a lower quality signal, as well as a greater chance of background noise interference within the participants' homes. Thirteen participant errors were performed by participants who heard the stimuli through headphones (41% of the errors), and nine participant errors were performed by those who heard the stimuli through speakers (59% of the errors). Regardless of these discrepancies, overall, both participants with ASD and TD participants demonstrated the use of the prosodic cue successfully, indicating that they were able to overcome these additional challenges.

**Future directions.** New areas of research surrounding this topic could extend into additional types of prosodic cues within varying contexts. As discussed earlier, prosody serves a variety of functions including pragmatic, affective, and syntactic functions. Although individuals with ASD demonstrated the ability to utilize the pragmatic prosodic cue of contrastive stress, further investigation into other (perhaps more subtle) types of affective and pragmatic cues might prove to be challenging for this population. Additionally, perhaps individuals with ASD would be less proficient with this skill in a more natural environment. Individuals with ASD have increased difficulty compared to individuals with typical development in separating signals from background noise (Teder-Sälejärvi, Pierce, Courchesne & Hillyard, 2005), so performance on this task in conditions realistic to normal conversation (i.e., signal to noise ratios of +5dB to -7dB) might be impacted (Crandell & Smaldino, 2000). Investigating the performance of

children who are TD and children with ASD with background noise may have implications for performance in the classroom or in groups.

**Conclusion**

In this study, we found that children and adolescents with ASD and TD children and adolescents interpret the prosodic cue of contrastive stress within a contextual communication task in similar ways. Within the given condition (i.e., when the target had been mentioned in the first sentence of the instruction set), participants in both groups used the prosodic cue of contrastive stress to guide online sentence interpretation. One difference found between the groups actually existed prior to the onset of the target word. TD participants did not show a preference to look at any particular object on the podium before the target word was heard, but participants with ASD looked significantly more at the "given" target word, that is, the word that had been mentioned in the first sentence of the instruction set. A single switch analysis was performed to correct for that initial preference so that the effects of prosody could be analyzed. We found that participants with ASD and participants with typical development both utilized the prosodic cue of contrastive stress within the contextual communication task. Both groups fixated to the unaccented, rather than accented, target with greater proportions during the given, rather than new, condition. No correlations between scores on the WCST, Receptive Language Index of the CELF-4, KBIT-2, the SCQ, or to the receptive prosody performance scores were found. Additionally, no correlation was found between the participants' expressive prosody skills (as determined by "naturalness ratings" on brief language samples) and receptive prosody skills.

These findings are a positive sign for the treatment of the expressive prosody of individuals with ASD. Children and adolescents with ASD who have intact cognitive skills and a developed language framework (similar to the profile of the participants in this study) are the appropriate population for targeting prosody (Bellon-Harn et al., 2007). Within this study, individuals with HFA and AS demonstrated success in perceiving and understanding the prosodic cue of contrastive stress within a contextual communication task.

Typically, the ability to understand is a prerequisite to be able to use a skill itself. Knowing that individuals with ASD may already be able to understand prosody will allow clinicians to spend more time targeting output and appropriate use, and particularly the nuances of prosody use to make individuals with ASD sound more natural. Strategies such as negative practice could be beneficial in allowing individuals with ASD recognize differences between their own production and the production used by a clinician. Once a client can recognize the appropriateness of the prosody in his or her speech, treatment can focus on modulating and shaping those productions to create more natural prosody. Although future research is necessary to investigate the ability of individuals with ASD to perceive and produce other prosodic cues serving various functions, and to investigate best practices in the treatment of deficits in expressive and receptive prosody, this study provides additional proof that individuals with impaired expressive prosody may have an intact ability to understand prosodic cues.

*Appendix A. List of audio stimuli and object locations for each critical and filler trial.*
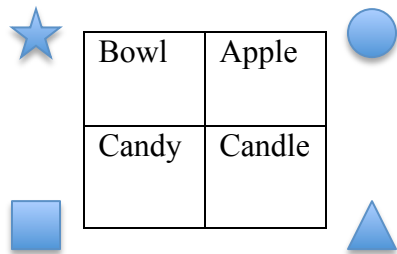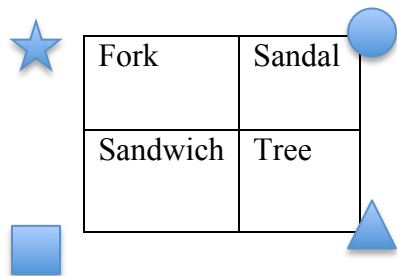


| Bowl | Apple |
|------|-------|
| Candy | Candle |

Put the candy on the star. Now, put the candy on the circle.



| Fork | Sandal |
|------|--------|
| Sandwich | Tree |

Put the sandwich on the star. Now, put the SANDWICH on the square.



| Tree | Sandwich |
|------|----------|
| Sandal | Fork |
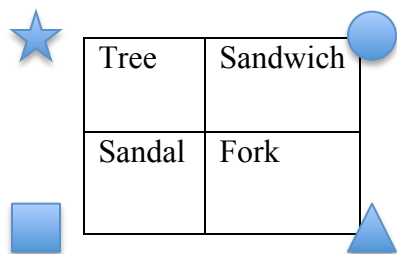
Put the sandal on the circle. Now, put the FORK on the square.

|        |       |
|--------|-------|
| Candle | Bowl  |
| Candy  | Apple |
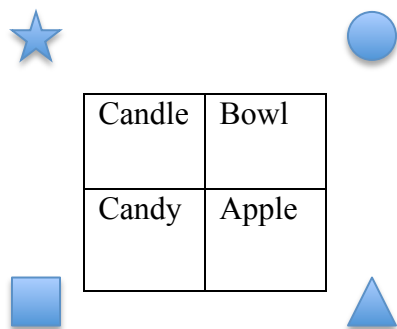
Put the candy on the square. Now, put the candle on the triangle.

|     |      |
|-----|------|
| Cup | Bell |
| Bed | Shoe |
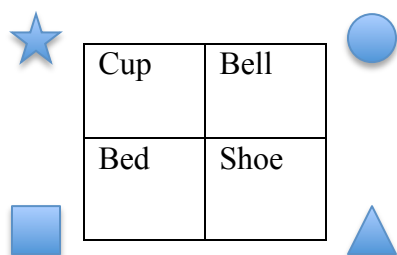
Put the bell on the square. Now, put the cup on the triangle.

|      |      |
|------|------|
| Bed  | Shoe |
| Bell | Cup  |

Put the bell on the circle. Now, put the BED on the triangle.

|          |        |
|----------|--------|
| Sandwich | Sandal |
| Tree     | Fork   |
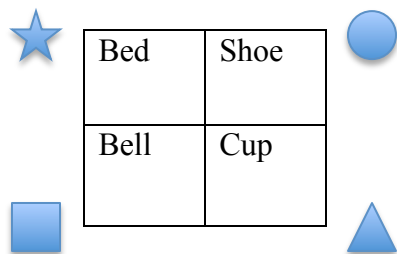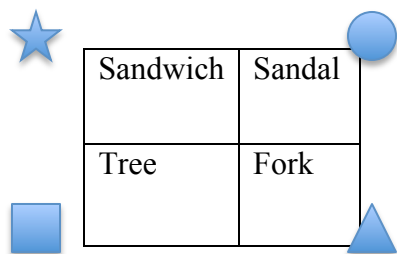
Put the sandwich on the square. Now, put the SANDAL on the star.

| Cup | Shoe |
|-----|------|
| Bed | Bell |
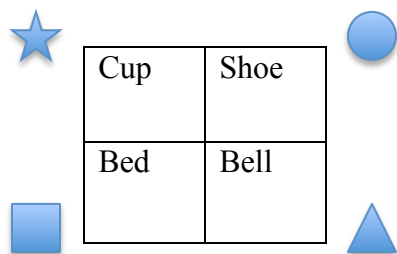
Put the bell on the star. Now, put the bed on the circle.

| Sandal | Fork |
|--------|------|
| Tree | Sandwich |
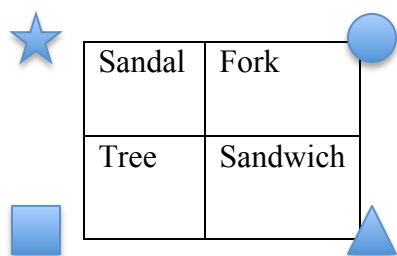
Put the sandwich on the triangle. Now, put the sandwich on the circle.

| Turkey | Turtle |
|--------|--------|
| Chair | Ball |
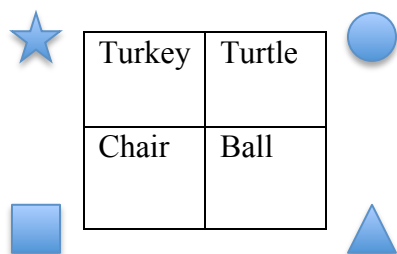
Put the turkey on the triangle. Now, put the TURKEY on the star.

| Bell | Cup |
|------|-----|
| Shoe | Bed |
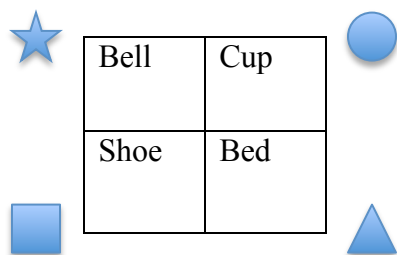
Put the shoe on the star. Now, put the shoe on the square.

| Ball | Turtle |
|------|--------|
| Turkey | Chair |
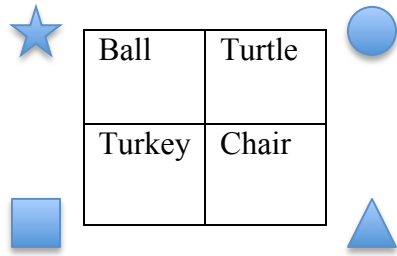
Put the turtle on the square. Now, put the ball on the star.
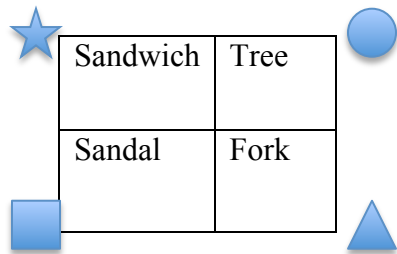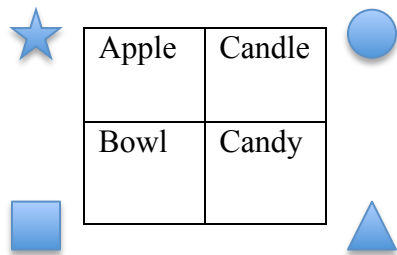
| Sandwich | Tree |
|----------|------|
| Sandal | Fork |

Put the sandal on the circle. Now, put the sandwich on the triangle.

| Apple | Candle |
|-------|--------|
| Bowl | Candy |

Put the candy on the star. Now, put the CANDLE on the triangle.

| Cup | Bed |
|-----|-----|
| Bell | Shoe |

Put the bell on the triangle. Now, put the BELL on the star.

| Bowl | Candy |
|------|-------|
| Candle | Apple |

Put the candy on the square. Now, put the apple on the circle.



| Turtle | Chair |
|--------|-------|
| Ball | Turkey |

Put the turkey on the circle. Now, put the turkey on the square.



| Ball | Chair |
|------|-------|
| Turtle | Turkey |

Put the chair on the circle. Now, put the turtle on the square.



| Bowl | Candle |
|------|--------|
| Apple | Candy |

Put the candle on the square. Now, put the CANDY on the star.

| Bowl | Candle |
|-------|--------|
| Candy | Apple |
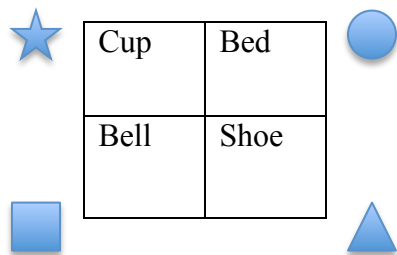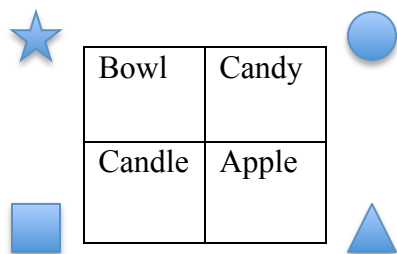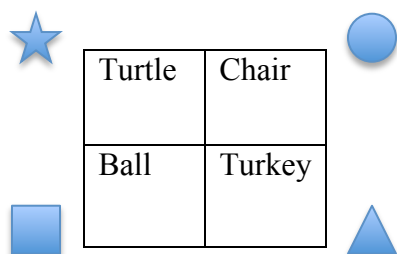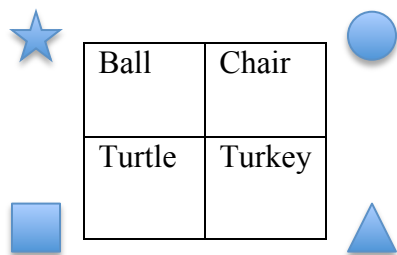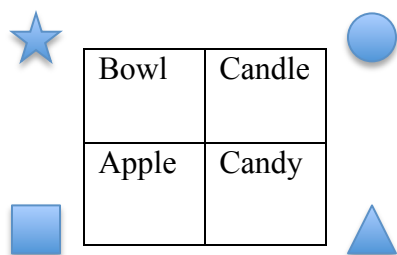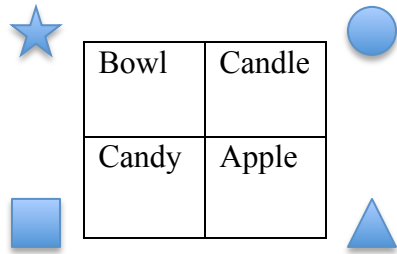
Put the candle on the star. Now, put the BOWL on the square.

| Chair | Turtle |
|--------|--------|
| Turkey | Ball |
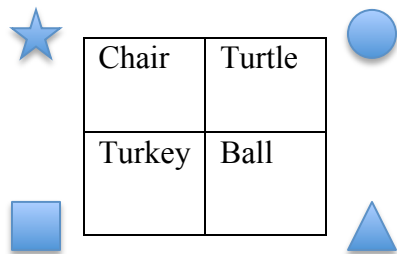
Put the turkey on the star. Now, put the turtle on the triangle.

| Turtle | Turkey |
|--------|--------|
| Chair | Ball |

Put the turkey on the star. Now, put the CHAIR on the circle.

| Fork | Sandwich |
|------|----------|
| Tree | Sandal |
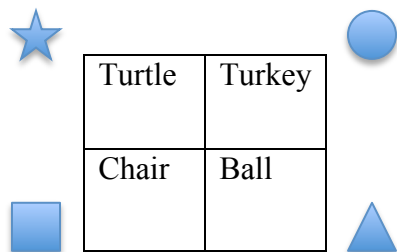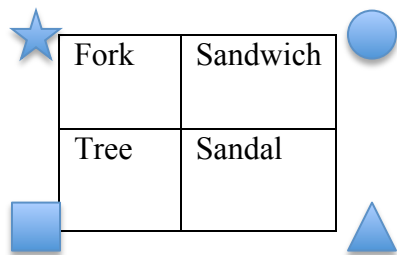
Put the sandal on the triangle. Now, put the SANDAL on the star.

★ 
| Bell | Bed |
|------|-----|
| Cup | Shoe |
◯

■ ▲
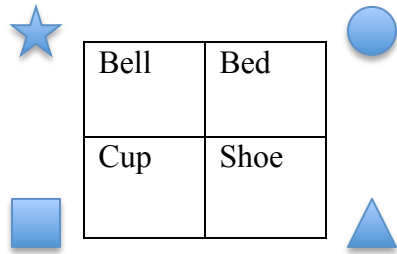
Put the bed on the triangle. Now, put the bed on the square.

★
| Cup | Bed |
|-----|-----|
| Bell | Shoe |
◯

■ ▲
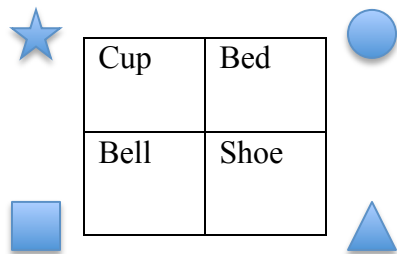
Put the bed on the triangle. Now, put the CUP on the circle.

★
| Candy | Bowl |
|-------|------|
| Apple | Candle |
◯

■ ▲
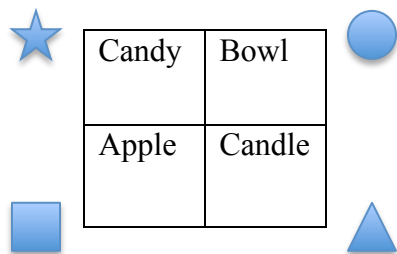
Put the bowl on the circle. Now, put the bowl on the triangle.

★
| Cup | Shoe |
|-----|------|
| Bed | Bell |
◯

■ ▲
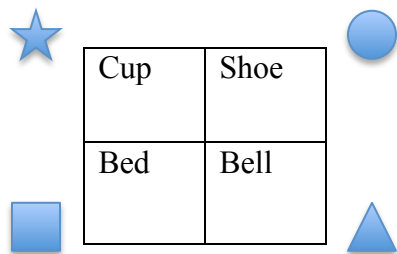
Put the bed on the star. Now, put the BELL on the circle.

|  |  |
|---|---|
| Tree | Fork |
| Sandal | Sandwich |

Put the sandal on the circle. Now, put the sandal on the square.

|  |  |
|---|---|
| Chair | Turkey |
| Turtle | Ball |

Put the turtle on the square. Now, put the turkey on the triangle.

|  |  |
|---|---|
| Bowl | Candy |
| Apple | Candle |

Put the candle on the star. Now, put the CANDLE on the square.

|  |  |
|---|---|
| Ball | Chair |
| Turtle | Turkey |

Put the turtle on the circle. Now, put the turtle on the square.

|        |      |
|--------|------|
| Bed    | Cup  |
| Shoe   | Bell |

Put the cup on the circle. Now, put the shoe on the star.

|        |          |
|--------|----------|
| Sandal | Fork     |
| Tree   | Sandwich |
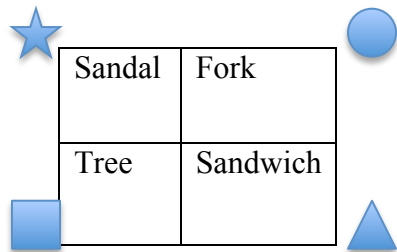
Put the tree on the triangle. Now, put the tree on the star.
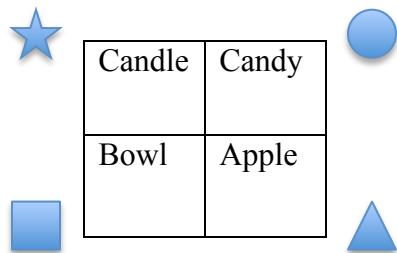
|        |       |
|--------|-------|
| Candle | Candy |
| Bowl   | Apple |

Put the candy on the triangle. Now, put the CANDY on the square.

|      |          |
|------|----------|
| Tree | Sandwich |
| Fork | Sandal   |

Put the sandal on the square. Now, put the SANDWICH on the star.

| Sandal | Fork |
|---|---|
| Sandwich | Tree |

Put the sandwich on the triangle. Now, put the sandal on the circle.

| Candy | Apple |
|---|---|
| Candle | Bowl |

Put the candle on the triangle. Now, put the candle on the circle.

| Turtle | Turkey |
|---|---|
| Ball | Chair |

Put the turtle on the square. Now, put the TURTLE on the star.

| Sandwich | Tree |
|---|---|
| Fork | Sandal |

Put the sandwich on the triangle. Now, put the tree on the star.

|      |       |
|------|-------|
| Bed  | Bell  |
| Cup  | Shoe  |

Put the bed on the square. Now, put the bell on the triangle.

|        |          |
|--------|----------|
| Tree   | Fork     |
| Sandal | Sandwich |
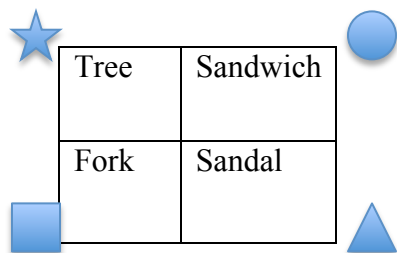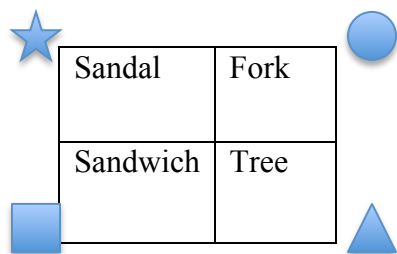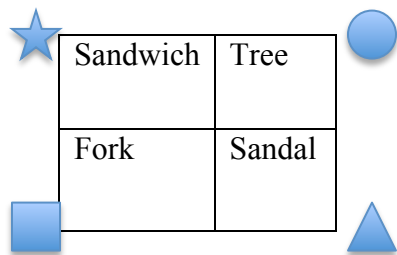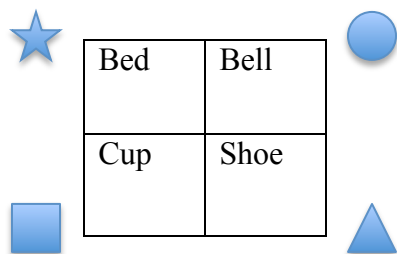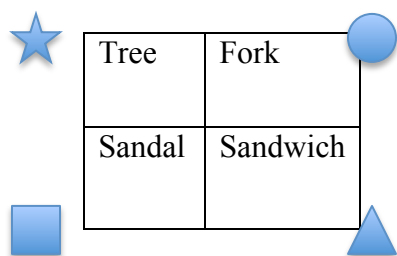
Put the fork on the square. Now, put the FORK on the triangle.
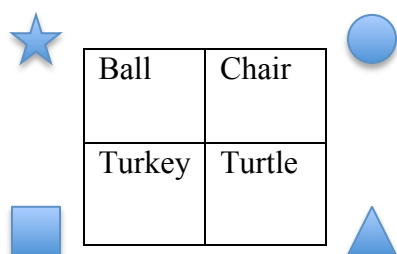
|        |        |
|--------|--------|
| Ball   | Chair  |
| Turkey | Turtle |

Put the turtle on the circle. Now, put the TURKEY on the star.

|      |      |
|------|------|
| Shoe | Bell |
| Cup  | Bed  |
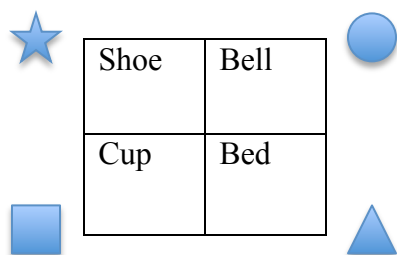
Put the bed on the square. Now, put the BED on the star.
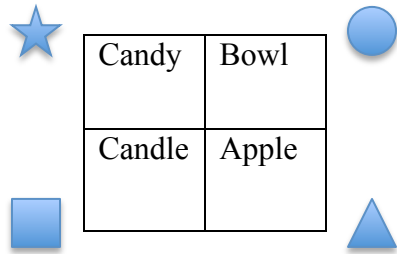
Put the candle on the circle. Now, put the candy on the triangle.
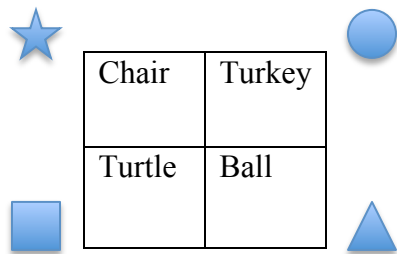


Put the ball on the star. Now, put the chair on the circle.



Put the turkey on the circle. Now, put the TURTLE on the square.



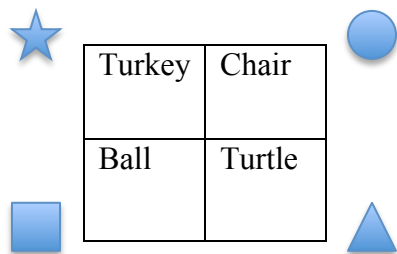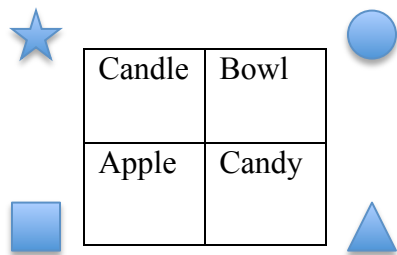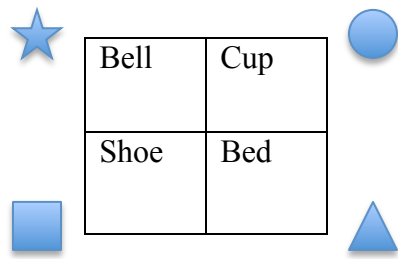Put the apple on the triangle. Now, put the candle on the circle.

| Bell | Cup |
|------|-----|
| Shoe | Bed |

Put the bell on the circle. Now, put the bell on the square.

*Appendix B. Participant Instructions for the Modified WCST*

Sort the cards based on color, shape, or number. I will tell you if you are right or wrong by saying yes or no. The correct way to sort may change during the activity.

References

Allopena, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38,* 419-439.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Arnold, J.E. (2008). *THE BACON* not *the bacon*: How children and adults understand accented and unaccented noun phrases. *Cognition*, *108,* 69-99.

Arnold, J. E., Tanenhaus, M. K., Altmann, R., & Fagano, M. (2004). The old and thee, uh, new: disfluency and reference resolution. *Psychological Science*, *15*(9), 578-582.

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

Baltaxe, C., & Simmons, J. (1985). Prosodic development in normal and autistic children. In Schopler, E., & Mesibov, G*., Communication Problems in Autism* (pp. 95-118). New York: Plenum Press.

Baron-Cohen, S. (1988). Social and pragmatic deficits in autism: Cognitive or affective? *Journal of Autism and Developmental Disorders*, *18*, 379–402.

Baumann, S. (2006). Information structure and prosody: Linguistic categories for spoken language annotation. In S. Sudhoff, D. Lenertova, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter, & J. Schlieber  (Eds.), *Methods in empirical prosody research* (pp. 153-180). Berlin: Walter de Gruyter GmbH & Co.

Bellon-Harn, M.L., Harn, W.E., & Watson, G.D. (2007). Targeting prosody in an eight-year-old child with high-functioning autism during an interactive approach to therapy. *Child Language Teaching and Therapy, 23(*2), 157-179.

Berg, E. A. (1948). A simple objective technique for measuring flexibility in thinking. *Journal of General Psychology*, *39*, 15-22.

Bhat, A.N., Landa, R.J. & Galloway, J.C. (2011). Current perspectives on motor functioning in infants, children, and adults with Autism Spectrum Disorders. *Physical Therapy, 91*(7), 1116-1129.

Boersma, P. & Weenink, D., (2011): Praat: Doing phonetics by computer (Version 5.2.35) [Computer program]. Retrieved August 5, 2011, from http://www.praat.org/

Cohen J.D., MacWhinney B., Flatt M., & Provost J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25*(2), 257-271.

Crandell, C.C., & Smaldino, J.J. (2000). Classroom acoustics for children with normal hearing and with hearing impairment. *Language, Speech, and Hearing Services in Schools, 31*, 362-370.

Cruttenden, A. (1997). *Intonation* (2nd ed.). Cambridge, UK: The Press Syndicate of the University of Cambridge.

Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory & Language, 47*(2), 292.

Diehl, J., Bennetto, L., Watson D., Gunlogson, C., & McDonough, J. (2008). Resolving ambiguity: A psycholinguistic approach to understanding prosody processing in high-functioning autism. *Brain & Language, 106,* 144-152.

Diehl, J. J., & Paul, R. (in press) Acoustic and perceptual measurements of prosody production on the PEPS-C by children with autism spectrum disorders. *Applied Psycholinguistics*.

Diehl, J.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., Snedeker, J., Tang, K., & Paul, R. (under review). *Cognitive and Developmental Factors in Prosody Comprehension in Children with Autism Spectrum Disorders.* Manuscript submitted for publication.

Eales, M. (1993). Pragmatic impairments in adults with childhood diagnoses of autism or developmental receptive language disorder. *Journal of Autism and Developmental Disorders, 23*(4), 593-617.

Fox, A. (2000). *Prosodic features and prosodic structure.* New York: Oxford University Press Incorporated.

Gerken, L., & McGregor, K. (1998). An overview of prosody and its role in normal and disordered child language. *American Journal of Speech-Language Pathology, 7(*2), 38-48.

Gibson, T.A., Oller, D.K., Jarmulowicz, L., & Ethington, C.A. (2012). The receptive-expressive gap in the vocabulary of young second-language learners: Robustness and possible mechanisms. *Bilingualism: Language and Cognition, 15*(1), 102-116.

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

Goodglass, H., & Kaplan, E. (1983). *The Boston Diagnostic Aphasia Examination.* Boston: Lea & Febiger.

Hagedorn, J., Hailpern, J., & Karahalios, K. (2008). VCode and VData: Illustrating a new framework for supporting the video annotation workflow. *Extended Abstracts of AVI.*

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test manual: Revised and expanded*. Odessa, FL: Psychological Assessment Resources.

Honig, A.S. (2007). Oral language development. *Early Child Development and Care, 177* (6-7), 581-613.

Jarvinen-Pasley, A., Peppe, S., King-Smith, G., & Heaton, P. (2008). The relationship between form and function level receptive prosodic abilities in autism. *Journal of Autism & Developmental Disorders*, *38*(7), 1328-1340.

Kaufman, A.S., & Kaufman, L. (2004). Kaufman Brief Intelligence Test, Second Edition. Bloomington MN: Pearson, Inc.

Korpilahti, P., Jansson-Verkasalo, E., Mattila, M., Kuusikko, S., Suominen, K., Rytky, S., Pauls, D. & Moilanen, I. (2007). Processing of affective speech prosody is impaired in Asperger Syndrome. *Journal of Autism & Developmental Disorders, 37*(8), 1539-1549.

Ladd, D.R. (2008). *Intonational Phonology* (2nd ed.). New York: Cambridge University Press.

Mazzoni, D., & Dannenberg, R. (2000). Audacity (Version 1.3.14-beta) [Computer program]. Retrieved December 8, 2011, from http://audacity.sourceforge.net/.

McCann, J., & Peppé, S. (2003). Prosody in autism spectrum disorders: a critical review. *International Journal of Language and Communication Disorders, 38*, 325–350.

Ozonoff, S., (1995). Reliability and validity of the Wisconsin Card Sorting Test in studies of autism. *Neuropsychology, 9*(4), 491-500.

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

Paul, R., Shriberg, L. D., McSweeny, J., Cicchetti, D., Klin, A., & Volkmar, F. (2005b). 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.

Peppé, S., Cleland, J., Gibbon, F., O'Hare, A., & Martinez Castilla, A. (2011). Expressive prosody in children with autism spectrum conditions. *Journal of Neurolinguistics*, *24*(1), 141-153.

Peppé, S., & McCann, J. (2003). Assessing intonation and prosody in children with atypical language development: The PEPS-C test and the revised version. *Clinical Linguistics and Phonetics*, *17*, 345–354.

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.

Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in Communication* (pp. 271–311). Cambridge, MA: MIT Press.

R Development Core Team (2006). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rubin, E., & Lennon, L. (2004). Challenges in social communication in Asperger Syndrome and High-Functioning Autism. *Topics in Language Disorders, 24*(4), 271-285.

Rutter, M., Bailey, A., & Lord, C., (2003). The social communication questionnaire (SCQ). Los Angelos, CA: Western Psychological Services.

Shriberg, L. D., Kwiatkowski, J., & Rasmussen, C. (1990). *The Prosody-Voice Screening Profile*. Tucson, AZ: Communication Skill Builders.

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

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology, 49,* 238-299.

Snedeker, J., & Yuan, S. (2008). Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of Memory & Language, 58*(2), 574-608.

Snowling, M., & Hulme, C. (1994). The development of phonological skills. *Philosophical Transactions: Biological Sciences, 356(*1315), 21-27.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632–1634.

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

Trueswell, J.C., Sekerina, I., Hill, N.M. & Logrip, M.L. (1999). The kindergartenpath effect: studying on-line sentence processing in young children. *Cognition, 73*, 89-134.

Turk, A., Nakai, S. & Sugahara, M. (2006). Acoustic segment durations in prosodic research: A practical guide. In S. Sudhoff, D. Lenertova, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter, & J. Schlieber (Eds.), *Methods in empirical prosody research* (pp. 1-27). Berlin: Walter de Gruyter GmbH & Co.

Volden, J., & Phillips, L. (2010). Measuring pragmatic language in speakers with Autism Spectrum Disorders: Comparing the Children's Communication Checklist-2 and the Test of Pragmatic Language. *American Journal of Speech-Language Pathology, 19*(3), 204-212.

Wetherby, A. M. (1986). Ontogeny of communicative functions in autism. *Journal of Autism and Developmental Disorders,* 16, 295-316.

Wiig, E. H., Secord, W. A., & Semel, E., (2008). Clinical evaluation of language

      fundamentals, fourth edition (CELF-4). Toronto, Canada: The Psychological

      Corporation/A Harcourt Assessment Company.