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

*Anne-Marie R. DePape[1], Aoju Chen[2,3], Geoffrey B. C. Hall[1] and Laurel J. Trainor[1,4] **

[1] *Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, ON, Canada*
[2] *Utrecht Institute of Linguistics, Utrecht University, Utrecht, The Netherlands*
[3] *Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands*
[4] *Rotman Research Institute, Baycrest Hospital, Toronto, ON, Canada*

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

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

INTRODUCTION

Autism spectrum disorder (ASD) involves impaired social interactions, repetitive and restrictive behaviors, and problems with communication (American Psychiatric Association, 1994). One striking feature of the speech of those with ASD is abnormal prosody (e.g., Baltaxe et al., 1984; Shriberg et al., 2001; McCann and Peppé, 2003; Paul et al., 2005, 2009; Diehl et al., 2009; Green and Tobin, 2009; Sharda et al., 2010; Bonneh et al., 2011; Nadig and Shaw, 2011). Prosody (or intonation) refers to suprasegmental features of speech, including pitch, duration, and intensity. According to Roach (2000) prosody serves important communicative functions at the grammatical, pragmatic, and affective levels. For example, prosody is used to distinguish speech acts such as questions, statements, and imperatives; to convey what is old and new information, and other sorts of pragmatic cues; and, at the affective level, to convey information about a speaker's feeling state (e.g., Halliday, 1967; Nespor and Vogel, 1986; Chun, 1988; Ladd, 1996; Cruttenden, 1997; Gussenhoven, 2004). In the present paper, we examine prosody at the pragmatic level. Abnormal prosody was included in the early descriptions of ASD (Kanner, 1943; Asperger, 1944), but has not been considered a defining feature of ASD, likely because the abnormalities appear to manifest differently across individuals (Baltaxe et al., 1984; Schreibman et al., 1986; Van Lancker et al., 1989; Diehl et al., 2009; Green and

Tobin, 2009; Bonneh et al., 2011). The prosody of ASD speech has been variously described as sounding "robotic," "wooden," "stilted," "monotone," "bizarre," "over precise," and even "singsong" (Fay and Schuler, 1980; Baltaxe and Simmons, 1985; Frith, 1991; Baron-Cohen and Staunton, 1994). Abnormalities appear to include both decreased and increased use of prosodic expression in ASD (Schreibman et al., 1986; Van Lancker et al., 1989), and there is suggestive evidence of "prosodic disorganization" in that prosody is not necessarily used to highlight the intended meaning (e.g., see Green and Tobin, 2009).

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

variability is consistent with recent genetic studies that indicate that although ASD is strongly heritable, it is etiologically heterogenetic, with many loci that each contribute a small amount to genetic susceptibility (e.g., Geschwind, 2009).

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

Most studies of prosody in ASD have examined children rather than adults or even adolescents (e.g., Paccia and Curcio, 1982; Baltaxe et al., 1984; Fosnot and Jun, 1999; Hubbard and Trauner, 2007; Paul et al., 2008; Diehl et al., 2009; Green and Tobin, 2009; Grossman et al., 2010; Sharda et al., 2010; Bonneh et al., 2011; Nadig and Shaw, 2011). Despite descriptions of monotone speech, studies employing acoustic analyses have generally found increased pitch variability in children with ASD, whether the corpus analyzed consisted of isolated words (Bonneh et al., 2011), conversations (Green and Tobin, 2009; Sharda et al., 2010; Nadig and Shaw, 2011), narratives (Diehl et al., 2009), or reading aloud (Green and Tobin, 2009). However, there appear to be individual differences. Baltaxe et al. (1984) found that children with ASD had either very narrow or very wide pitch ranges, suggesting heterogeneity among children. Similarly, Green and Tobin (2009) found that although children with ASD as a group showed larger pitch ranges and larger pitch variability compared to typically developing children, those with ASD could be divided into three distinct groups, consisting of those with narrow, typical, or wide pitch ranges. Similar variance across individuals might also exist for prosodic use of duration, although there is less research on this question. Nadig and Shaw (2011) reported no difference in overall speech rate between children with and without ASD. In other studies, adults with ASD were found to produce less lengthening than controls on stressed syllables in imitative speech (Paul et al., 2008), but children with ASD were found to produce more lengthening than controls on stressed

syllables in spontaneous speech (Grossman et al., 2010). Clearly, more research is needed in order to understand the prosodic use of duration in ASD.

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

Critical to an understanding of prosodic abnormalities in ASD is the question of whether prosody is used to enhance communication. The present paper examines the use of prosody to mark *information structure* in individuals with ASD. In normal conversation, prosody is used to convey what is important in an utterance with respect to the talker's beliefs about the listener's knowledge state (Chafe, 1976; Clark and Haviland, 1977; Prince, 1986). Two of the most widely discussed information structural categories are (1) topic, which refers to what a sentence is about and typically represents given information, and (2) focus, which typically represents new information about the topic (Lambrecht, 1994; Vallduví and Engdahl, 1996). For example, "boy" is the topic and "apple" is the focus of the sentence "The boy is eating an apple" when uttered in response to the question "*What* is the boy eating?". However, "apple" is the topic and "boy" is the focus of the same sentence when uttered in response to "*Who* is eating the apple?". Among typical speakers, focus words are produced with a larger pitch range and longer duration than topic words, all other acoustic features being equal (Chen, 2009). Making focal information more prominent can facilitate language comprehension whereas making the topical information more prominent can delay comprehension (e.g., Nooteboom and Terken, 1982; Birch and Clifton, 1995; Chen, 2010). Inappropriate marking of information can lead to problems in achieving desired communicative intents and produce, among other things, confusion between conversational partners (Fine et al., 1991).

Developmentally, the tendency to use a falling pitch contour across a sentence may sometimes override children's ability to mark intended meanings, for example, not using a rising contour when appropriate to ask a question (Wells et al., 2004). One study of Dutch-speaking children found that when answering a question, 7- to 8-year-old, but not 4- to 5-year-old, children emphasized focus words appropriately (Chen, 2011). In particular, the 4- to 5-year-olds accented focus words with several types of accents (e.g., rise, fall, downstepped fall – a fall with a lower peak than the preceding accent) and showed no adult-like preference for falling accents in the sentence-final (object) position, a problem that the author attributed to the children's need to check and seek confirmation (hence the final rise) and a lack of knowledge of the

typical functions of downstepped fall. On the other hand, earlier work on English children and a study of German children suggested that when the focal information is contrastive, even 3- to 4-year-olds showed evidence of using prosody appropriately (Hornby and Hass, 1970; Müller et al., 2006).

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

The small amount of research on prosody in adolescents and adults with ASD suggests that the abnormalities documented in children persist through late development and are resistant to change (Shriberg et al., 2001; Paul et al., 2005; Diehl et al., 2009). Not surprisingly, atypical prosody in adults with ASD can have real-life consequences, such as affecting their ability to make friends and achieve meaningful employment (Van Bourgondien and Woods, 1992; Paul et al., 2005). Thus, a full understanding of the nature of the prosodic deficits is important.

We collected semi-spontaneous speech samples in adults in a controlled but interactive paradigm that enabled us to directly measure pitch and duration features of the same words in focus and topic conditions in sentence-initial and sentence-final

positions. We had three main goals: (1) To compare the general use of prosodic pitch and duration in adults with and without ASD; (2) to examine the use of pitch and duration to convey information structure in adults with and without ASD in short, controlled conversations; and (3) to examine whether individual differences in use of prosody are related to level of language functioning.

MATERIALS AND METHODS
PARTICIPANTS
We tested 12 adult male participants ($M = 25.4$ years; $range = 17$–34 years) with a diagnosis of ASD (**Table 1**). Of these six had receptive vocabulary standard scores of 100 or greater and six had scores below 100 as measured by the standardized Peabody picture vocabulary test-III (PPVT; Dunn and Dunn, 1997; see **Table 1**). ASD participants had been seen at clinic (Offord Centre), assessed using standard instrument batteries (ADOS and ADI; Lord et al., 1989, 1994), and all carried formal psychiatric diagnoses of either AS or HFA. Participants completed the PPVT and a questionnaire on languages spoken and family history of ASD. Previous research has found that scores on the PPVT are correlated with scores on the clinical evaluation of language fundamentals (CELF) test, which includes assessments of morphology, syntax, semantics, and working memory for language (Kjelgaard and Tager-Flusberg, 2001). Thus, the PPVT can be used as a measure of general language functioning. The categorization by current language ability (PPVT) followed their diagnoses, such that all six with scores of 100 or greater (Autism High Language Function, A-highL, group) carried a diagnosis of AS and the others (Autism Moderate Language Functioning, A-moderateL, group) a diagnosis of HFA. In addition, all six in the A-moderateL group experienced early language delay whereas none in the A-highL group experienced early language delay. Six subjects showing typical development (normal controls, NC, group) were also tested ($M = 26.3$ years; $range = 23$–34 years) to provide a standard for comparison purposes, as such detailed comparative acoustic analyses of topic and focus do not exist for English. None of the participants in the NC group had a family member diagnosed with ASD. All participants were monolingual English-speakers and the groups were matched in age ($F < 1$). The A-moderateL group performed significantly worse on receptive vocabulary than the NC ($p = 0.003$)

Table 1 | Demographic and background information by group.

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

PPVT, Peabody picture vocabulary test (standard scores).

and A-highL ($p = 0.006$) groups. NC and A-highL groups did not differ ($p = 0.95$) by *post hoc* Tukey's HSD tests (**Table 1**).

MATERIALS AND PROCEDURE

The research was approved by the McMaster University Research Ethics Board and conformed to the principles set out in the Canadian Tri-Council Ethics Policy. All participants gave informed consent. Testing lasted approximately 1 h and took place in an acoustically treated room. Participants received a debriefing statement after completing the study.

Participants were tested individually playing the "Under the Shape" game (Chen, 2011), in which they were asked questions about pictures presented on a computer. Their verbal responses were recorded for offline acoustic analysis. This task measured how participants vary prosody according to two variables, *information structure* (topic/focus), and *sentence position* (initial/final), adapted from Chen (2011) for use with children and adults. This task was administered on an Acer Notebook using Microsoft Office PowerPoint. Responses were recorded in Sound Studio 3 (Felt Tip Incorporated, 2009) and saved as.wav files at a 44.1-kHz sampling rate with 16 bit resolution using a Mac iBook G4. A microphone (D770 Emotion AKG) was connected to the iBook using a US-122 USB Audio/MIDI Interface. Participants were seated about 2″ away from the microphone.

During the *Familiarization Phase*, participants were told that they would see pictures of people, animals and objects performing different actions. They were asked to report aloud what they saw on the screen (e.g., "rabbit"), when they were shown a picture. This phase included 30 pictures presented in a fixed order and took about 2 min to complete. The purpose was to ensure that participants could identify and use a consistent label for each picture. Participants were asked to remember these labels as they would see the same pictures in the next phase of the game.
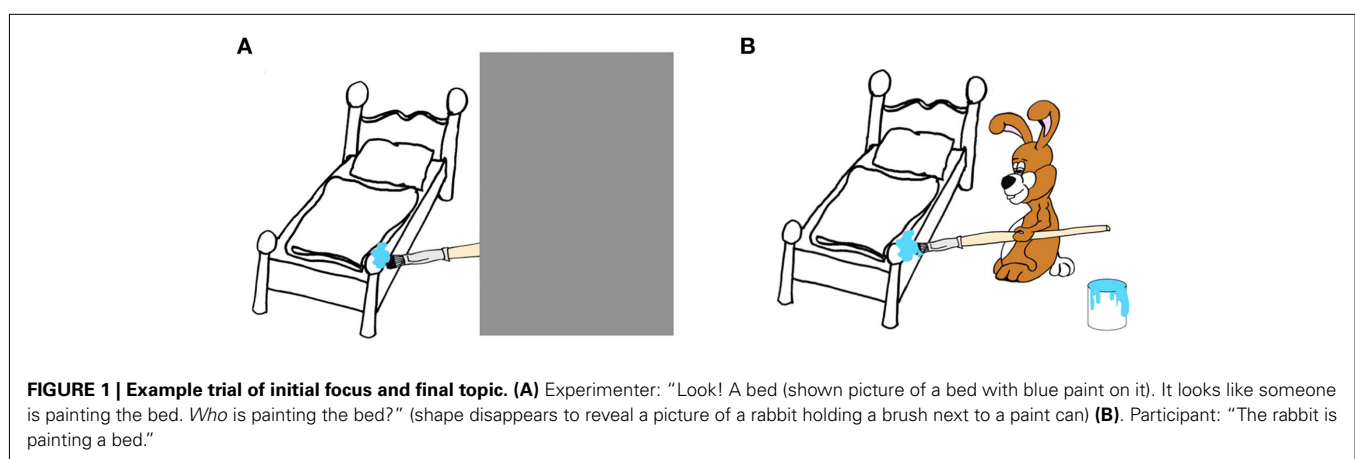
During the *Experimental Phase*, the "Under the Shape" game was played. Two referents, which could be people, animals, or objects, were presented on the screen at the same time but one was covered by an opaque rectangle. The experimenter posed a *who* or a *what* question. When the experimenter pressed a button on the keyboard, the rectangle was removed and the participant was then able to answer the experimenter's question (see **Figure 1**). This procedure measured how participants converse with a live speaker. The experimenter received training so that all questions were asked using the same prosody, with prominence placed on the first word, which was either *who* or *what*.

Responses to *who* and *what* question types differed in terms of whether the new information (focus) occurred in the sentence-initial position (subject) and the given information (topic) in the sentence-final (object) position or vice versa. Note, however, that the subject was always at the beginning of the sentence and the object at the end, regardless of which was the focus in terms of containing new information. For example, when "*WHO* is painting the bed?" (see **Figure 1**) was asked, the new information (focus) occurred in the initial position, "The *RABBIT* is painting the bed." Conversely, when "*WHAT* is the rabbit painting" (see **Figure 2**) was asked, the new information (focus) occurred in the final position, "The rabbit is painting the *BALL*." For each sentence position (initial/final), all nouns were used in topic and focus contexts in order to ensure that the acoustic analyses compared the same words across different contexts. To avoid boredom, every combination of subject and object nouns occurred only once during the experiment. Participants were required to respond to all questions using a full sentence. This response format ensured that each sentence contained a subject in the sentence-initial position and an object in the sentence-final position. Following four practice trials, participants completed 22 trials in the experimental phase, with equal numbers of *who* and *what* questions.

ACOUSTIC ANNOTATION

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

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

given a value that matched the rise (range-fall of zero). We also measured word duration.

The subject and object words were acoustically annotated by examining the waveform using the wide-band spectrum and pitch track in Praat 5.1.0.7 (Boersma and Weenink, 2009) and checked for octave errors by comparing visual displays of pitch tracks with auditory perceptions. The data were coded by the first author after receiving sufficient training from the second author. All data were checked independently by the second author for both accuracy and consistency and corrections were made by the two transcribers together. Three F0-related landmarks were labeled in each word, as illustrated in **Figure 3**:
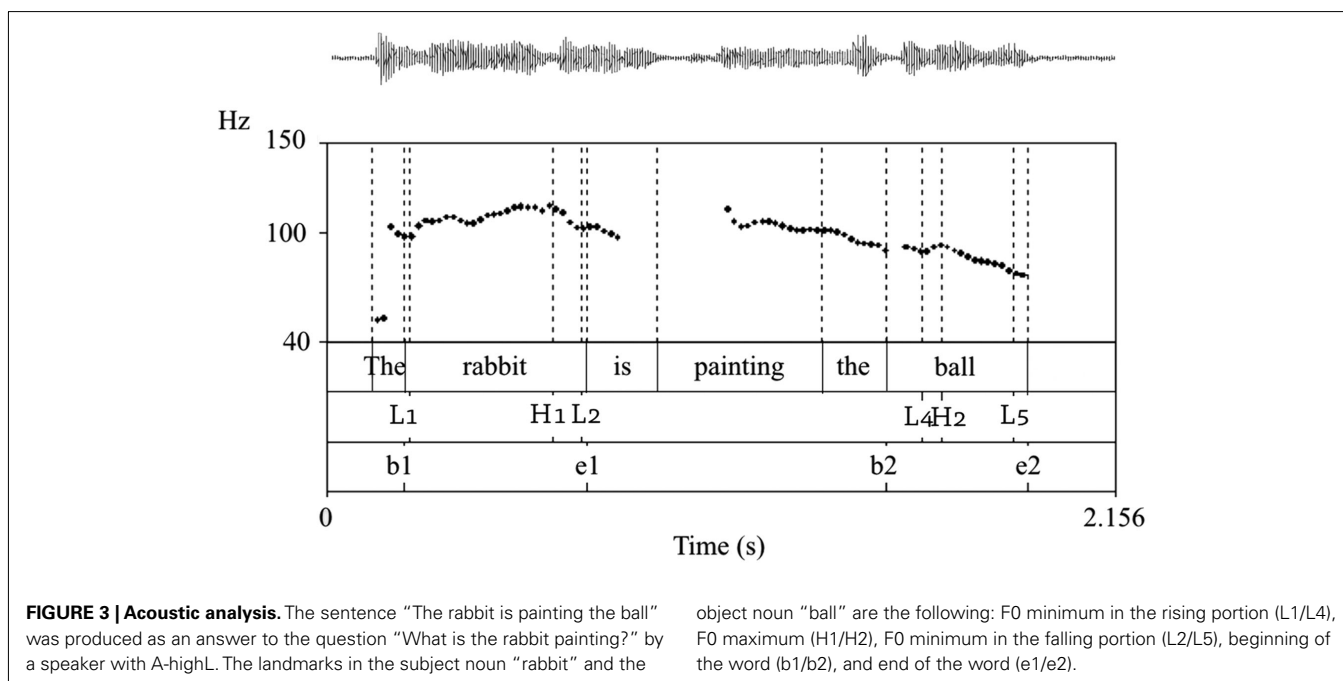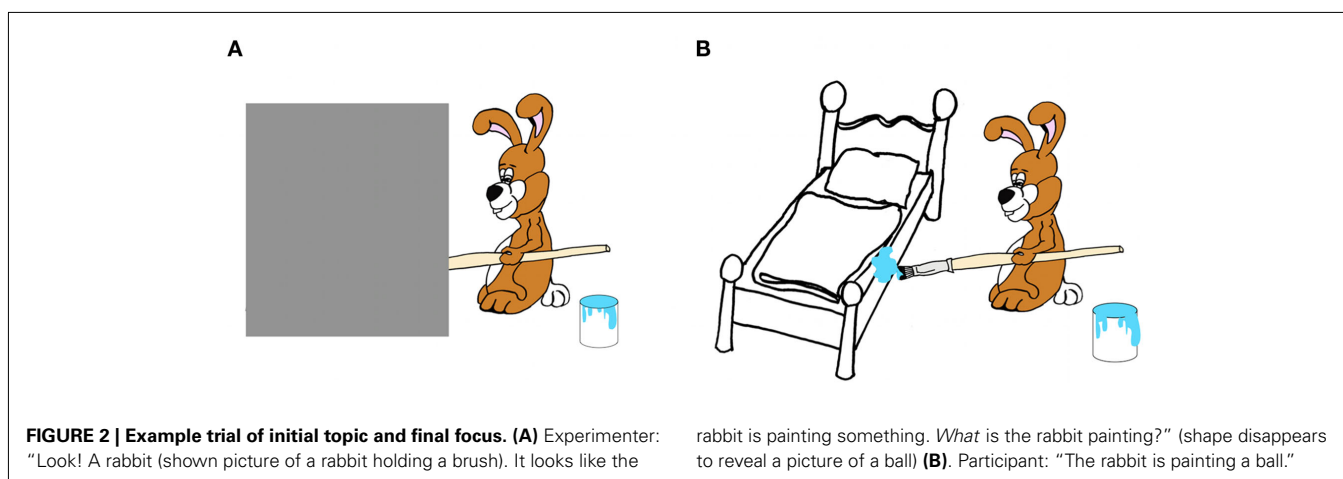
- Beginning F0 minimum: the initial lowest pitch in the subject noun (L1) and in the object noun (L4).
- F0 maximum: the highest pitch in the subject noun (H1) and in the object noun (H2) before the beginning of the pitch fall.

- Final F0 minimum: the lowest pitch reached following the F0 maximum in the subject noun (L2) and in the object noun (L5).

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

Further, two segmental landmarks were labeled in each noun:

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

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

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

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

Three measurements were then obtained for each noun:

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

STATISTICAL ANALYSIS AND RESULTS

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

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

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

When measuring how acoustic features are varied across topic and focus, it is important that the same words are compared. This is because the intrinsic pitch of vowels causes some words to have larger pitch ranges than others, and different segmental markup causes some words to be longer in duration than others.

For the "Under the Shape" game (Chen, 2011), some participants occasionally used different labels on different trials for the same object (e.g., "bunny" and "rabbit"), an error that was made on a total of 19 out of 396 word pairs (4.8%). These cells were replaced with the mean for that word for that particular group given that replacing up to 5% of data in this manner has been found to be acceptable (Rubin et al., 2007).

PITCH AND DURATION

The ANOVA conducted on absolute pitch revealed a main effect of sentence position, $F(1, 15) = 42.63$, $p < 0.001$, $\eta^2 = 0.74$, with pitch falling from sentence-initial ($M = 119.34$ Hz, SEM $= 4.55$ Hz) to sentence-final words (declination; $M = 104.09$ Hz, SEM $= 3.74$ Hz), but no main effect of group, $F(2, 15) = 2.87$, $p = 0.09$ ($\eta^2 = 0.28$). There was also no significant interaction between group and sentence position ($F < 1$), suggesting no overall differences in pitch range across the sentences.

The ANOVA conducted on absolute duration revealed a main effect of sentence position, $F(1, 15) = 6.29$, $p = 0.02$, $\eta^2 = 0.30$, with shorter durations for the sentence-initial ($M = 0.33$ s, SEM $= 0.01$ s) than for the sentence-final words ($M = 0.36$ s, SEM $= 0.02$ s), but no main effect of group, $F(2, 15) = 1.87$, $p = 0.20$. There was no significant interaction between group and sentence position ($F < 1$), indicating no overall differences between groups in duration and suggesting similar durational variation across the sentences.

SENTENCE-INITIAL (SUBJECT)

Initial range-rise

In the initial (subject) position, the ANOVA on range-rise revealed no significant effects (**Figure 4A**). Planned Mann–Whitney tests revealed that for topic words, the A-moderateL group used a significantly smaller range-rise than the NC ($U = 2.00$, $p = 0.01$, $r = 0.74$) and A-highL ($U = 2.00$, $p = 0.01$, $r = 0.74$) groups. There were no significant differences across groups for focus words.

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

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

Initial range-fall

The ANOVA on range-fall revealed significant main effects of information structure, $F(1, 15) = 17.31$, $p = 0.001$, $\eta^2 = 0.54$ and of group, $F(2, 15) = 3.56$, $p = 0.05$, $\eta^2 = 0.32$ (see **Figure 4C**). Post hoc tests using Tukey's HSD showed that the main effect for group was due to a significantly smaller range-fall overall in the A-highL compared to the NC group ($p = 0.04$).

Planned Mann–Whitney tests revealed that for topic words, the A-moderateL group used a significantly smaller range-fall than the NC ($U = 5.00$, $p = 0.04$, $r = 0.60$) and A-highL ($U = 5.00$, $p = 0.04$, $r = 0.60$) groups. For focus word, the NC group used a significantly larger range-fall than the A-highL group ($U = 4.00$,
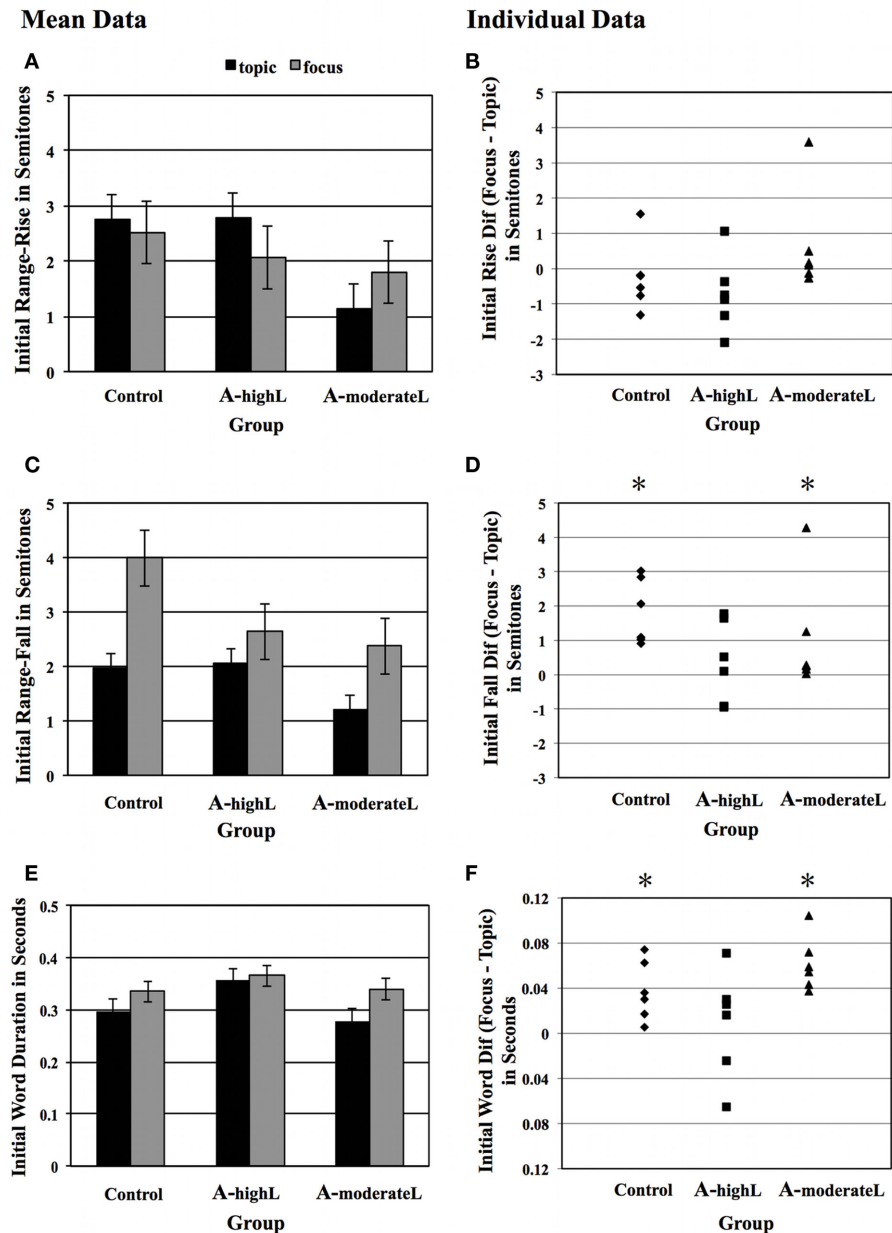
FIGURE 4 | Sentence-initial results. (A) Mean range-rise and SE by group. **(B)** Individual data for range-rise difference (focus–topic) by group. Note that no difference between topic and focus is represented by the zero line. **(C)** Mean range-fall and SE by group. **(D)** Individual data for range-fall difference (focus–topic) by group. **(E)** Mean word duration and SE by group. **(F)** Individual data for duration difference (focus–topic) by group. $*p < 0.05$.

$p = 0.03$, $r = 0.65$) and there was a trend for the NC group to use a larger range-fall than the A-moderateL group ($U = 9.00$, $p = 0.15$, $r = 0.42$). This is consistent with the greatest marking of information structure by the control group.

Planned Wilcoxon signed-rank tests revealed significantly larger range-falls for focus than topic in the NC ($p = 0.03$, $d = 0.96$) and A-moderateL ($p = 0.03$, $d = 0.64$) groups, but not in the A-highL ($p = 0.46$) group (see **Figure 4D**).

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

Initial duration

The ANOVA on word duration revealed a significant main effect of information structure, $F(1, 15) = 20.01$, $p < 0.001$, $\eta^2 = 0.57$, with longer word durations for focus than topic, no main effect of group, $F(2, 15) = 1.97$, $p = 0.17$, and a significant interaction

between information structure and group, $F(2, 15) = 3.57$, $p = 0.05$, $\eta^2 = 0.32$ (see **Figure 4E**).

Planned Mann–Whitney tests revealed that the A-highL group used a longer duration for topic words than the A-moderateL group ($U = 5.00$, $p = 0.04$, $r = 0.60$) but there were no significant effects for focus words.

Planned Wilcoxon signed-rank tests revealed a significant difference between topic and focus for the NC ($p = 0.03$, $d = 0.43$) and A-moderateL ($p = 0.03$, $d = 0.70$), groups, but not for the A-highL ($p = 0.46$), group (see **Figure 4F**).

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

Initial correlations with PPVT

Finally, across the entire sample, there were significant (or approaching significant) Pearson correlations between PPVT and the size of the sentence-initial range-rise (subject), $r = 0.48$, $p = 0.04$, and range-fall, $r = 0.46$, $p = 0.06$, but not between PPVT and duration, $p > 0.23$ (**Table 2**), again suggesting that differences in language ability underlie the different prosodic strategies.

SENTENCE-FINAL (OBJECT)
Final range-rise
In the final (object) position, the ANOVA on range-rise revealed no significant effect of, or interactions involving, group. However, there was a significant main effect of information structure, $F(1, 15) = 14.21$, $p = 0.002$, $\eta^2 = 0.49$, with a larger range-rise for focus than for topic (see **Figure 5A**).

Planned Mann–Whitney tests revealed no differences between groups. Planned Wilcoxon signed-rank tests revealed significant differences between topic and focus for the NC ($p = 0.03$, $d = 0.48$) and A-moderateL ($p = 0.03$, $d = 0.49$) groups, but not for the A-highL ($p = 0.25$) group (see **Figure 5B**).

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

Final range-fall
The ANOVA on range-fall revealed a significant main effect of information structure, $F(1, 15) = 15.83$, $p = 0.001$, $\eta^2 = 0.51$, with a larger range-fall for focus than for topic. There was also a significant main effect of group, $F(2, 15) = 5.75$, $p = 0.01$, $\eta^2 = 0.43$, and an interaction between information structure and group, $F(2,$

Table 2 | Pearson correlations between receptive vocabulary and prosody ($n = 18$ speakers).

	Receptive vocabulary (PPVT)
Sentence-initial range-rise	0.48*
Sentence-initial range-fall	0.46[†]
Sentence-initial duration	0.30
Sentence-final range-rise	0.01
Sentence-final range-fall	0.48*
Sentence-final duration	0.27

*$p < 0.05$; [†]$p < 0.06$.
PPVT, Peabody picture vocabulary test (standard scores).

$15) = 4.67$, $p = 0.03$, $\eta^2 = 0.38$ (see **Figure 5C**). *Post hoc* tests using Tukey's HSD revealed an overall larger pitch range in the A-highL compared to A-moderateL group ($p = 0.01$).

Planned Mann–Whitney tests revealed that for topic words, the A-highL group showed a significantly larger range-fall compared to the NC ($U = 0.00$, $p = 0.004$, $r = 0.83$) and A-moderateL group ($U = 0.00$, $p = 0.004$, $r = 0.83$), consistent with overall exaggerated pitch excursions in the A-highL group. The A-moderateL group showed a significantly smaller range-fall for focus words compared to the NC ($U = 5.00$, $p = 0.04$, $r = 0.60$) and A-highL ($U = 3.00$, $p = 0.016$, $r = 0.69$) groups, consistent with smaller pitch excursions in the A-moderateL group.

Planned Wilcoxon signed-rank tests revealed a significantly larger range-fall for focus than topic for the NC group ($p = 0.03$, $d = 1.2$), but not for the A-highL ($p = 0.34$) and A-moderateL ($p = 0.34$) groups (see **Figure 5D**).

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

Final duration
For word duration, there was a significant main effect of information structure, $F(1, 15) = 24.17$, $p < 0.001$, $\eta^2 = 0.62$, with longer word durations for focus than for topic (see **Figure 5E**). The main effect of group was not significant, $F(2, 15) = 1.23$, $p = 0.32$, but there was a significant interaction between information structure and group, $F(2, 15) = 8.17$, $p = 0.004$, $\eta^2 = 0.52$.

Planned Mann–Whitney tests revealed no significant differences between groups. Planned Wilcoxon signed-rank tests revealed that the difference between topic and focus was significant for the NC ($p = 0.03$, $d = 0.47$) and A-moderateL ($p = 0.03$, $d = 0.38$) groups, but not for the A-highL group ($p = 0.25$; see **Figure 5F**).

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

Final correlations with PPVT
Finally, across the entire sample ($n = 18$) there was a significant Pearson correlation between PPVT and the size of the sentence-final (object) range-fall, $r = 0.48$, $p = 0.04$, although not between PPVT and the size of the range-rise, $p > 0.97$, or duration, $p > 0.27$ (**Table 2**), again suggesting that differences in language ability underlie the different prosodic strategies.

DISCUSSION
Even with only six participants in each of the subgroups, with the detailed acoustic analyses we performed, we found robust and marked differences in performance between those with ASD with stronger language skills (A-highL group) compared to those with weaker language skills (A-moderateL group). Regardless of information structure, compared to controls, we found larger pitch ranges for those with ASD with strong language skills, and smaller pitch ranges for those with moderate language skills. It is worth
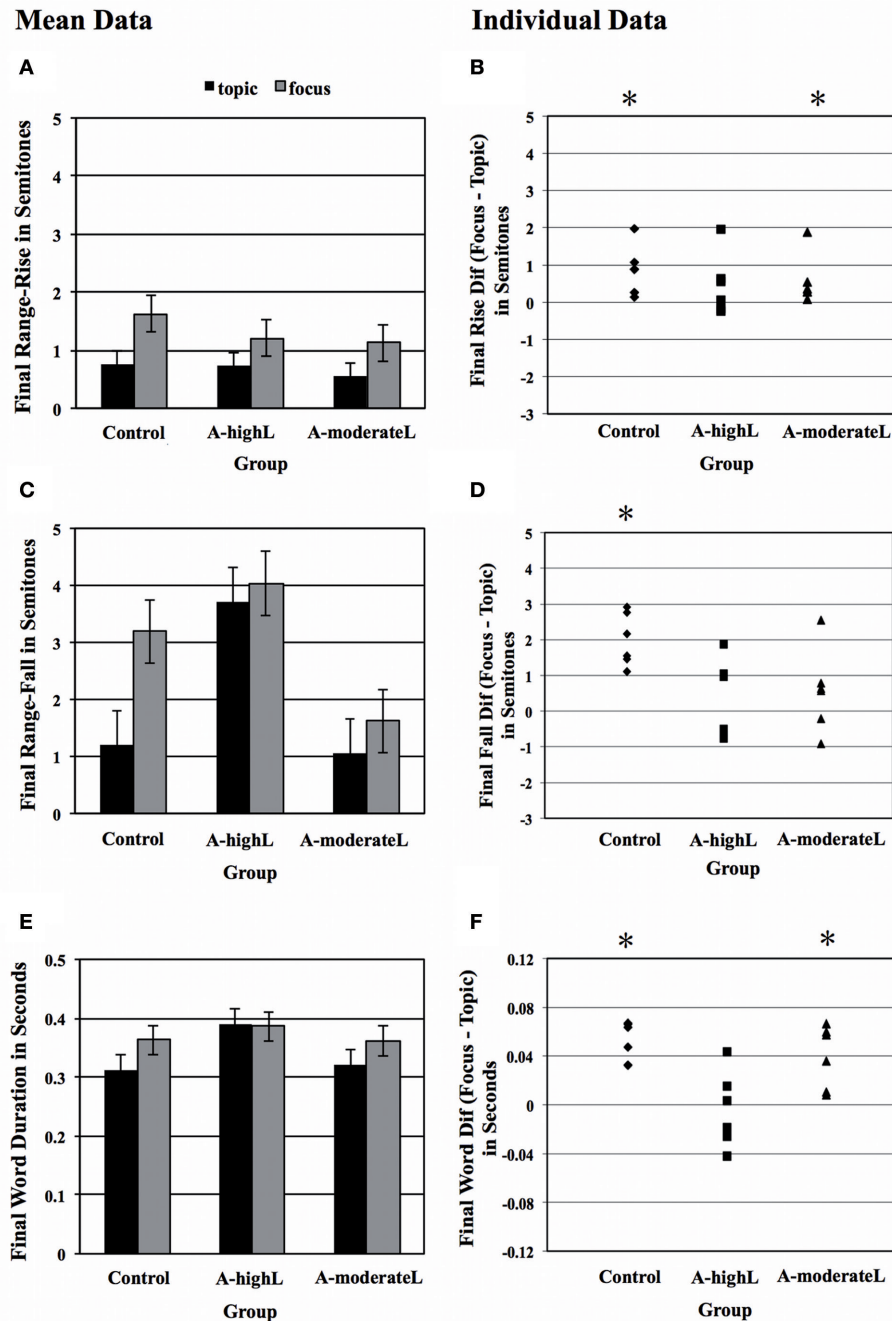
FIGURE 5 | Sentence-final results. (A) Mean range-rise and SE by group. **(B)** Individual data for range-rise difference (focus–topic) by group. Note that no difference between topic and focus is represented by the zero line. **(C)** Mean range-fall and SE by group. **(D)** Individual data for range-fall difference (focus–topic) by group. **(E)** Mean word duration and SE by group. **(F)** Individual data for duration difference (focus–topic) by group. $*p < 0.05$.

noting that these differences cannot be explained by potential differences in overall pitch height as the three groups did not differ significantly in initial absolute (starting) pitch. The small pitch range of those with ASD and moderate language skills is consistent with a monotone quality to their speech, whereas the large pitch range of those with ASD and stronger language skills is consistent with a singsong quality. It would be interesting to test this notion

further in future studies to see whether speech with these different prosodic pitch characteristics is indeed perceived as monotone and singsong, respectively. With respect to duration, we did not find any significant group differences in how this acoustic feature was varied in general when information structure was not considered. Thus, pitch appears to be the primary contributor to general abnormal prosody in ASD, a finding that could help to inform

future remediation programs in speech and language. Our finding that individuals with ASD could be divided into subgroups who use either a smaller or a larger pitch range than normal is consistent with previous reports of heterogeneity in this regard (e.g., Baltaxe et al., 1984; Green and Tobin, 2009). Furthermore, our results extend previous studies by indicating that in ASD, use of a smaller pitch range is associated with moderate language skill, whereas use of a larger pitch range is associated with high language skill.

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

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

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

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

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

Importantly, the present study also contributes to the finding that the prosodic abnormalities identified in children with ASD

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

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

Regardless of the origin of the differences, both the A-highL and A-moderateL groups used abnormal prosody, which would affect their ability to communicate effectively. Although those with moderate language skills used pitch and duration cues to mark information structure, they varied pitch to a lesser extent than controls, and this would likely give the impression that they were uninterested in conversation. Indeed, in real communicative contexts, such use of monotonous speech might override the fact

that those in the A-moderateL group mark information structure appropriately for the most part. On the other hand, those with high language skills used more prosodic variation relative to controls and those in the A-moderateL group (average size of range-fall across sentence positions was approximately 0.5 semitones and 1.5 semitones larger than control and A-moderateL groups, respectively), but the way that they did so with respect to information structure was not useful to listeners. This use of prosody is likely distracting because the indiscriminant use of large pitch excursions does not direct the listener's attention to focus words. It remains for future research to document the precise effects of different prosodic abnormalities related to information structure on typical listeners, but it is evident that abnormal prosody can have serious consequences for social communication (Wells et al., 2004; Peppé et al., 2006, 2007).

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

REFERENCES

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

Asperger, H. (1944). Die autistischen Psychopathen im Kindesalter. *Arch. Psychiatr. Nervenkr.* 117, 76–136.

Baltaxe, C. (1984). Use of contrastive stress in normal, aphasic, and autistic children. *J. Speech Hear. Res.* 24, 97–105.

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

Baltaxe, C. A. M., and Guthrie, D. (1987). The use of primary sentence stress by normal, aphasic and autistic children. *J. Autism Dev. Disord.* 17, 255–271.

Baltaxe, C. A. M., and Simmons, J. Q. (1985). "Prosodic development in normal and autistic children," in *Communication Problems in Autism*,

eds E. Schloper and G. Mesilov (New York: Plenum), 95–125.

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

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

Birch, S., and Clifton, C. (1995). Focus, accent and argument structure: effects on language comprehension. *Lang. Speech* 38, 365–391.

Boersma, P., and Weenink, D. (2009). *Praat: Doing Phonetics by Computer [Computer Program]. Version 5.1.07.* Available at: http://www.praat.org/ [retrieved 16 August 2009].

Bonneh, Y. S., Levanon, Y., Dean-Pardo, O., Lossos, L., and Adini, Y. (2011). Abnormal speech spectrum and increased pitch variability in young autistic children. *Front. Hum. Neurosci.* 4:237. doi:10.3389/fnhum.2010.00237

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

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

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

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

Chen, A. (2011). Tuning information packaging: intonational realisation of topic and focus in child Dutch. *J. Child Lang.* 38, 1055–1083.

Chun, D. M. (1988). The neglected role of intonation in communicative competence and proficiency. *Mod. Lang. J.* 72, 295–303.

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

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

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

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

Fay, W., and Schuler, A. L. (1980). *Emerging Language in Autistic Children*. Baltimore: University Park Press.

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

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

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

Foster, B., and King, B. H. (2003). Asperger syndrome: to be or not to be? *Curr. Opin. Pediatr.* 15, 491–494.

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

Geschwind, D. H. (2009). Advances in autism. *Annu. Rev. Med.* 60, 367–380.

Green, H., and Tobin, Y. (2009). Prosodic analysis is difficult…but worth it: a study in high functioning autism. *Int. J. Speech Lang. Pathol.* 11, 308–315.

Grossman, R. B., Bemis, R. H., Skwerer, D. P., and Tager-Flusberg, H. (2010). Lexical and affective prosody in children with high-functioning autism. *J. Speech Lang. Hear. Res.* 53, 778–793.

Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.

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

Hanssen, J., Peters, J., and Gussenhoven, C. (2008). "Prosodic effects of focus in Dutch declaratives," in *Proceedings of the 4th Conference on Speech Prosody*, eds P. A. Barbosa, S. Madureira, and C. Reis (Campinas: Editora RG/CNPq), 609–612.

Hornby, P. A., and Hass, W. A. (1970). Use of contrastive stress by preschool children. *J. Speech Hear. Res.* 13, 359–399.

Howlin, P., and Asgharian, A. (2007). The diagnosis of autism and Asperger syndrome: findings from a survey of 770 families. *Dev. Med. Child Neurol.* 41, 834–839.

Hubbard, K., and Trauner, D. A. (2007). Intonation and emotion in autistic spectrum disorders. *J. Psycholinguist. Res.* 36, 159–173.

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

Kjelgaard, M., and Tager-Flusberg, H. (2001). An investigation of language impairment in autism: implications for genetic subgroups. *Lang. Cogn. Process.* 16, 287–308.

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

Lambrecht, K. (1994). *Information Structure and Sentence Form: Topics, Focus, and the Representations of Discourse Referents*. Cambridge: Cambridge University Press.

Lord, C., and Paul, R. (1997). "Language and communication in autism," in *Handbook of Autism and Pervasive Development Disorders*, eds D. J. Cohen and F. R. Volkmar (New York: John Wiley), 195–225.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., and Schopler, E. (1989). Autism diagnostic observation schedule: a standardized observation of communicative and social behavior. *J. Autism Dev. Disord.* 19, 185–212.

Lord, C., Rutter, M., and Le Couteur, A. (1994). Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J. Autism Dev. Disord.* 24, 659–685.

McCaleb, P., and Prizant, B. M. (1985). Encoding of new versus old information by autistic children. *J. Speech Hear. Disord.* 50, 230–240.

McCann, J., and Peppé, S. (2003). Prosody in autistic spectrum disorders: a critical review. *Int. J. Lang. Commun. Disord.* 38, 325–350.

McCann, J., Peppé, S., Gibbon, F. E., O'Hare, A., and Rutherford, M. (2007). Prosody and its relationship to language in school-aged children with high-functioning autism. *Int. J. Lang. Commun. Disord.* 42, 682–702.

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

Nadig, A., and Shaw, H. (2011). Expressive prosody in high-functioning autism: increased pitch range and what it means to listeners. *J. Autism Dev. Disord.* [Epub ahead of print].

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

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

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

Norbury, C. F. (2005). Barking up the wrong tree? Lexical ambiguity resolution in children with language impairments and autistic spectrum disorders. *J. Exp. Child. Psychol.* 90, 142–171.

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

Paccia, J., and Curcio, F. (1982). Language processing and forms of immediate echolalia in autistic children. *J. Speech Hear. Res.* 25, 42–47.

Paul, R., Augustyn, A., Klin, A., and Volkmar, F. (2005). Perception and production of prosody by speakers with autistic spectrum disorders. *J. Autism Dev. Disord.* 35, 205–220.

Paul, R., Bianchi, N., Augustyn, A., Klin, A., and Volkmar, F. (2008). Production of syllable stress in speakers with autism spectrum disorders. *Res. Autism Spectr. Disord.* 2, 110–124.

Paul, R., Orlovski, S., Marchinko, H., and Volkmar, F. (2009). Conversational behaviors in youth with high-functioning autism and Asperger syndrome. *J. Autism Dev. Disord.* 39, 115–125.

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

Peppé, S., McCann, J., Gibbon, F., O'Hare, A., and Rutherford, M. (2007). Receptive and expressive prosodic ability in children with high-functioning autism. *J. Speech Lang. Hear. Res.* 50, 1015–1028.

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

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

Roach, P. (2000). *English Phonetics and Phonology*. Cambridge: Cambridge University Press.

Rubin, L. H., Witkiewitz, K., St. Andre, J., and Reilly, S. (2007). Methods for handling missing data in the behavioural neurosciences: don't throw the baby rat out with the bath water. *J. Undergrad. Neurosci.* 5, A71–A77.

Schreibman, L., Kohlenberg, B. S., and Britten, K. R. (1986). Differential responding to content and intonation components of a complex auditory stimulus by nonverbal and echolalic autistic children. *Anal. Interv. Dev. Disabil.* 6, 109–125.

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

Shriberg, L., Paul, R., McSweeney, J., Klink, A., Cohen, D., and Volkmar, F. (2001). Speech and prosody characteristics of adolescents and adults

with high functioning autism and Asperger syndrome. *J. Speech Lang. Hear. Res.* 44, 1097–1115.

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

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

Van Lancker, D., Cornelius, C., and Kreiman, J. (1989). Recognition of emotional-prosodic meanings in speech by autistic, schizophrenic,

and normal children. *Dev. Neuropsychol.* 5, 207–226.

Wells, B., Peppé, S., and Goulandris, A. (2004). Intonation development from five to thirteen. *J. Child Lang.* 31, 749–778.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.