Journal of Autism and Developmental Disorders (2020) 50:3032–3045
https://doi.org/10.1007/s10803-020-04392-9

ORIGINAL PAPER

An Acoustic Characterization of Prosodic Differences in Autism Spectrum Disorder and First-Degree Relatives

Shivani P. Patel[1] · Kritika Nayar[1] · Gary E. Martin[2] · Kathryn Franich[3,5] · Stephanie Crawford[1] · Joshua J. Diehl[4] · Molly Losh[1]

Published online: 13 February 2020

Abstract

This study examined prosody through characterization of acoustic properties of the speech of individuals with ASD and their parents, during narration. A subset of utterances were low-pass filtered and rated for differences in intonation, speech rate, and rhythm. Listener ratings were minimally related to acoustic measures, underscoring the complexity of atypical prosody in ASD. Acoustic analyses revealed greater utterance-final fundamental frequency excursion size and slower speech rate in the ASD group. Slower speech rate was also evident in the ASD parent group, particularly parents with the broad autism phenotype. Overlapping prosodic differences in ASD and ASD Parent groups suggest that prosodic differences may constitute an important phenotype contributing to ASD features and index genetic liability to ASD among first-degree relatives.

Keywords Autism spectrum disorder · Broad autism phenotype · Prosody · Acoustic

Abbreviations

ASD	Autism spectrum disorder
BAP	Broad autism phenotype
F0	Fundamental frequency
nPVI	Normalized pairwise variability index

Introduction

Autism spectrum disorder (ASD) is a genetically-based neurodevelopmental disorder characterized by core deficits in social communication and restricted interests or repetitive behaviors (American Psychiatric Association 2013). Within the domain of social communication, prosody, which includes intonation modulation (changes in how 'high' or 'low' the voice sounds, based on rate of vocal fold vibration), rhythm (how evenly-timed syllables are in speech), and rate (how rapidly syllables are uttered in speech), has been noted as a key area of impairment in ASD (Peppé et al. 2006, 2007; Wells and Peppé 2003; Wells et al. 2004). It is important to note that prosody serves a variety of functions, all of which can impact communication. For example, stress and intonation can be used to encode grammatical information (e.g., the stress difference in the verb conTRAST versus the noun CONtrast), as well as pragmatic information (e.g., differentiating elements of discourse which are under discussion in the sentence "MARY saw the dog," an appropriate answer to the question "Who saw the dog?" vs. "Mary saw the DOG," an appropriate answer to the question "What did Mary see?"). Indeed, inappropriate use of stress in response to either question may disrupt communicative interactions. Furthermore, prosody conveys important information about speaker intent (e.g., sarcasm, persuasion, demand) and affect (e.g., joy, dislike) (Bachorowski 1999; Bachorowski and Owren 1995). Subtle abnormalities in intonation and rhythm have been found to adversely impact listeners' ability to perceive and process speech in the general population (Bent et al. 2016).

In adolescence and adulthood, prosodic differences reported in individuals with ASD include atypical intonation

✉ Molly Losh
m-losh@northwestern.edu

1 Roxelyn and Richard Pepper Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Drive, Frances Searle Building, #2-366, Evanston, IL 60208, USA

2 Department of Communication Sciences and Disorders, St. John's University, Staten Island, NY, USA

3 Present Address: Department of Linguistics and Cognitive Science, University of Delaware, Newark, DE, USA

4 LOGAN Community Resources, Inc, South Bend, IN, USA

5 Department of Linguistics, University of Chicago, Chicago, USA

and stress patterns, aberrant speech rate, lack of affective quality, and poor loudness control (Baltaxe and Simmons 1985; Baltaxe et al. 1984; Baron-Cohen and Staunton 1994; Fay 1969; Pronovost et al. 1966; Shriberg et al. 2001). While these atypical prosodic features don't typically rise to the level of making speech unintelligible to a listener, they are among the first identifiable characteristics that create an impression of "oddness" among peers of individuals with ASD (Mesibov 1992; Van Bourgondien and Woods 1992). While prosodic deficits have been reported since the earliest delineations of ASD by Kanner and Asperger (Asperger and Frith 1991; Kanner 1943), McCann and Peppé's (2003) review of sixteen studies of prosody in individuals with ASD revealed many inconclusive or contradictory findings across studies. For instance, whereas Fosnot and Jun (1999) reported atypical intonation patterns in individuals with ASD, another study (Baltaxe et al. 1984) found that intonation patterns did not differ from those of controls with typical development. Similarly, studies have reported both increased and decreased speech rate in individuals with ASD (Baron-Cohen and Staunton 1994; Shriberg et al. 2001). McCann and Peppé (2003) cite insufficient sample sizes, limited control data, and lack of standardized measures and sufficiently detailed methodology as explanations for the variability in findings. Furthermore, of those early studies reviewed, only two (Baltaxe et al. 1984; Fosnot and Jun 1999) made acoustic measurements of prosody, while others used subjective measures (i.e., perceptual judgments), which while clinically valid, offer only global characterization of differences.

More recent work applying acoustic analyses across a variety of communicative contexts has contributed to a more comprehensive understanding of specific prosodic atypicalities in individuals with ASD. For instance, examining speech of high-functioning individuals with ASD (HFA) and controls during a task involving retelling an emotional story, Edelson et al. (2007) found that the HFA group had significantly higher pitch overall and a wider pitch range compared to the control group. Diehl et al. (2009) also identified greater fundamental frequency (F0) variation during narrative production for both adolescent and child HFA groups compared to their respective control groups. Similarly, Nadig and Shaw (2012) found increased F0 range in speakers with HFA compared to an age- and language level-matched control group. In a study by Paul et al. (2008) involving the production of nonsense syllables, individuals with ASD demonstrated significantly less of a distinction in duration between stressed and unstressed syllables, as well as a pattern of increased F0 range for both stressed and unstressed syllables compared to controls, though this difference was not statistically significant. Using the Profiling Elements of Speech-Prosody in Communication (PEPS-C; Peppé and McCann 2003), Diehl and Paul (2013) found that individuals with ASD had a significantly greater F0 range than controls when using focus to highlight new information. However, the ASD group's standard deviation (SD) and range of F0 did not differ from the control group in subtests involving expressing affect, producing appropriate utterance-final intonation for statements and questions, and signaling appropriate phrasal boundaries. In sum, although prior findings indicate variable differences in prosody across studies employing acoustic analysis versus perceptual judgments of prosody, together, prior work clearly demonstrates that prosody is impacted in ASD. It is likely that atypicalities along these acoustic and temporal dimensions contribute to impressionistic evaluations of speech in individuals with ASD as being overly 'sing-songy' (perhaps a reflection of greater F0 range in individuals with ASD) or machine-like (perhaps a reflection of lesser differentiation between stressed and unstressed syllables).

Differences in prosody have also been observed among clinically unaffected first degree relatives of individuals with ASD and may constitute a linguistic feature of the broad autism phenotype (BAP). The BAP refers to a constellation of subclinical language and personality features documented at elevated rates among relatives that parallel the defining features of ASD in quality (Losh et al. 2008; Losh et al. 2012a; Piven et al. 1997; Ruser et al. 2007). Such phenotypes are thought to reflect underlying genetic liability, and might afford better understanding of the range of phenotypic expression of ASD liability, and links to underlying biology. To this end, prior studies have shown a specific neuropsychological profile associated with the BAP in parents of individuals with ASD, where differences in social cognition and language processing abilities were evident in the subgroup of parents showing BAP features, whereas those without these subclinical phenotypes performed more similarly to controls (e.g., Losh et al. 2009; Nayar et al. 2018).

Subtle differences based on clinical-behavioral ratings of prosodic (also referred to as suprasegmental) language features (e.g., intonation and rhythm) have been reported in individuals with the BAP (Landa et al. 1992; Losh et al. 2008, 2012b; Piven et al. 1997). Further, a study exploring affective prosody perception in individuals with ASD and their siblings demonstrated that siblings exhibited some difficulty accurately perceiving emotion from prosody, though differences from controls were more subtle than those observed in the ASD group (Oerlemans et al. 2014). These findings suggest that prosodic differences could constitute an important feature of the BAP in unaffected relatives that indexes genetic liability to ASD. However, studies to date have only applied perceptual ratings to examine the acoustic properties of prosody in relatives of individuals with ASD, leaving unclear how these perceptually-based differences in prosody might bear out in objective acoustic measurements.

This study aimed to build on prior work by comparing acoustic profiles of prosody among individuals with ASD and their parents, with respective control groups. Overall, we tested the hypothesis that prosody is a linguistic marker of genetic liability to ASD by examining objective acoustic measures of prosody in individuals with ASD and their parents. Based on the literature reviewed above, we predicted that individuals with ASD and their parents would demonstrate overlapping areas of prosodic differences and that such differences would co-segregate with ASD severity in individuals with ASD and features of the BAP in parents. Additionally, we predicted that acoustic measurements from each group would relate to broader pragmatic (i.e., social) language atypicalities during conversational interactions in individuals with ASD and the BAP. Sex differences were also explored. Finally, listener-based perceptual ratings of prosody were collected in the ASD and ASD control groups to investigate how acoustic patterns might map to listeners' perceptions of prosodic differences in individuals with ASD.

Based on prior literature characterizing prosody in individuals with ASD, we predicted that acoustic analyses would reveal differences in measurements of F0 variability (SD and range), rhythm, and rate in the ASD group. Given reports of prosodic differences in a subset of parents of individuals with ASD who exhibit features of the BAP, we predicted similar differences in acoustic measurements of F0 variability, rhythm, and rate in the ASD Parent group, and that these differences would be driven by parents with the BAP. Furthermore, we predicted that differences in acoustic measurements would be related to increased clinical-behavioral atypicalities, including increased pragmatic language violations in the ASD and ASD Parent groups, as well as ASD symptom severity in the ASD group. Finally, we predicted that the confluence of variables in which acoustic differences emerge would significantly contribute to listeners' perceptual ratings of prosody in the ASD and ASD control groups.

Method

Participants

Study participants included 55 individuals with ASD (ASD group), 39 typically developing individuals with no family history of ASD (ASD control group), 96 parents of individuals with ASD (ASD parent group), and 48 parent controls without any family history of ASD (parent control group). Both the ASD and ASD control groups included children and adults. Parent groups included both parents when possible (ASD parent group: n = 28 couples; parent control group: n = 5 couples). Participants were recruited through a broader family-genetic study of ASD, which included individuals with ASD, their parents, and respective controls. Additional inclusionary criteria for all participants included having no history of brain injury, major psychiatric disorder, or known genetic syndrome or neurodevelopmental disorder (other than ASD), and being a native and fluent speaker of English. Furthermore, participants in either control group were excluded if they had first- or second-degree relatives with ASD, related genetic disorders (e.g., fragile X syndrome) or dyslexia. Individuals in the ASD parent group were excluded if they had a diagnosis of ASD.

Diagnosis of ASD was confirmed for all individuals in the ASD group with research reliable administration and scoring of the Autism Diagnostic Observation Schedule-2nd Edition (ADOS-2; Lord et al. 2000), as well as the Autism Diagnostic Interview-Revised (ADI-R; Lord et al. 1994) when time permitted. The ADOS-2 was also administered to individuals in the ASD control group to rule out ASD. ADOS-2 Overall, Social Affect, and Restricted and Repetitive Behaviors (RRB) calibrated severity scores were used to determine ASD severity (Hus et al. 2014; Lord et al. 2012).

The Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler 1999), Wechsler Adult Intelligence Scale-Third or Fourth Editions (WAIS; Wechsler 1997, 2008), or the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV; Wechsler 2003) were used to assess IQ for all participants. Table 1 summarizes the overall age, full-scale IQ, verbal IQ, and performance IQ for each group. The ASD group had a significantly lower full-scale IQ, VIQ, and PIQ than the ASD control group ($ts > -2.63$, $ps < .01$). The ASD parent group had a significantly lower performance IQ and higher mean chronological age than parent controls ($ts > -2.18$, $ps < .03$).

Narrative Elicitation

The 24-page wordless picture book, *Frog, Where Are You?* (Mayer 1969), was used to elicit spontaneous narratives from all participants across groups. This story has been used extensively in previous studies of narrative discourse in typically developing populations, as well as those with developmental disorders including ASD (Capps et al. 2000; Diehl et al. 2006; Losh and Capps 2003; Tager-Flusberg and Sullivan 1995). The story is about the adventures of a boy and his dog as they search for the boy's missing pet frog. Participants were asked to narrate the story as each page was presented to them on a computer screen. All narrations were recorded using either a Blue Snowball USB microphone or a Logitech USB Desktop Microphone (980186-0403). The microphone was positioned approximately 8 inches from the participant's mouth during the narration.

Table 1 Group characteristics

	ASD group M (SD) Range	ASD control group M (SD) Range	ASD parent group M (SD) Range	Parent control group M (SD) Range
Males:females	45:10[a]	19:20	38:58	20:28
Chronological age	16.57 (6.62)	18.99 (5.21)	46.67 (7.67)*	43.00 (9.87)
	6.45–35.10	12.38–32.03	28.38–65.76	25.95–63.89
Full-scale IQ	104.22 (12.03)*	115.45 (12.03)	111.09 (11.41)	114.81 (11.36)
	83.00–131.00	89.00–142.00	85.00–136.00	85.67–136.00
Verbal IQ	105.10 (13.79)*	117.50 (11.36)	109.41 (12.23)*	111.20 (12.85)
	82.00–146.00	93.00–142.00	80.00–132.00	82.00–138.00
Performance IQ	102.88 (14.58)*	111.08 (14.19)	110.40 (11.10)	115.08 (12.93)
	68.00–131.00	79.00–143.00	83.00–133.00	86.00–148.00

Significant differences ($p < .05$) between the ASD Group or ASD Parent Group and their respective control groups are indicated with *

ASD Autism Spectrum Disorder, *M* Mean, *SD* standard deviation

[a]Sex distribution in the ASD Group largely reflects sex bias noted in the prevalence of ASD (Werling and Geschwind 2013; Fombonne 2009; Idring et al. 2015)

Utterance Selection

In order to obtain relatively comparable speech samples from participants' narratives, the first utterance from each page of the narrative was used in analysis, provided that the first utterance of a page did not meet exclusionary criteria consistent with criteria used in prior studies (i.e., the utterance contained character speech, a question, unfinished words, fewer than two words, an interruption by the examiner, or was unintelligible, directed towards someone else in the room and not related to the narrative, or abandoned; e.g., Shriberg et al. 2001). Importantly, the use of these exclusionary criteria promoted consistency in utterance length and type (e.g., no questions) across participants. Complete utterances with greater than two words were necessary for analysis of utterance-final F0 excursion size (described below). Additionally, these criteria ensured that included utterances were part of the narration rather than included as part of an interrupting conversation or other interaction. In cases where exclusionary criteria were met, the next utterance from the same page was chosen. If another utterance from the same page was unavailable or all utterances within the page met exclusionary criteria, an utterance from another page within the same section of the book (i.e., beginning, middle, or end) was selected. This resulted in a maximum of 24 utterances from each narration, each with a structure that typically corresponded to a maximal intonational phrase (i.e., the highest unit of structure within linguistic models of the prosodic hierarchy; Selkirk 2009, 2011). A minimum of 20 qualifying utterances were included from each participant. All audio recordings were analyzed using Praat (Boersma and Weenink 2017; http://www.fon.hum.uva.nl/praat/; version 6.0.29), a program for acoustic analysis of speech signals.

Automatic Alignment of Speech Samples

The onset and offset of selected utterances, which were assessed by finding the points in the acoustic signal where a participant began and finished an utterance, were manually marked using TextGrids in Praat and subsequently automatically aligned at the level of words and phones (i.e., speech sounds) using FAVE (Forced Alignment and Vowel Extraction; Rosenfelder 2013).[1] Each file was manually checked for inconsistencies and hand-corrected where necessary by the authors (all trained in phonetic analysis). Due to poor recording quality which prevented proper alignment, one adolescent subject and one parent subject were excluded from analysis, which is reflected in the sample sizes reported above. A Praat script with pitch tracking ranges dependent on speaker age and sex (see Table 2) was implemented in order to minimize pitch tracking errors in Praat. The script extracted F0 (in Hz) within each force-aligned utterance with a timestep of 0.01 s. Logarithmic transformation was applied to F0 values to approximate the scale on which pitch is perceived. After applying the logarithmic transformation, the mean, SD, and range of F0 for each utterance was calculated. Range of F0 was calculated by subtracting the maximum and minimum F0 values obtained from each utterance. Mean F0

[1] FAVE uses modified acoustic models for English from the Penn Phonetics Lab Forced Aligner (p2fa; Yuan and Liberman 2008) to predict where words and phones begin and end within the acoustic signal. More information about forced alignment with FAVE can be found at https://github.com/JoFrhwld/FAVE/wiki/Using-FAVE-align.

Table 2 Fundamental frequency detection ranges

	Minimum	Maximum
Males 11 years and younger	130	400
Males 12–18 years of age	70	400
Males 19+ years	70	250
Females (all ages)	130	400

is a measure of the rate of vocal fold (also known as "vocal cord") vibration and is the physical correlate of pitch, or how "high" or "low" an individual's voice is. The SD and range of F0 measure the extent to which an individual's pitch varies during speech. Additionally, the F0 range (in Hz) of the final word of each utterance was calculated via Praat script as a measure of 'excursion size', or overall change in F0 (whether rising or falling) in utterance-final position (subsequently referred to as utterance-final F0 excursion size). Z-score transformation was applied to normalize data within the ASD, ASD control, ASD parent, and parent control groups.

Speech rate for each utterance was automatically calculated using a script in *R* statistical software. Using input from the force aligned TextGrid files, the script computed the number of vocalic intervals (representative of the number of syllables) per utterance and divided these totals by the duration of each utterance, resulting in a measure of syllables per second (including within-utterance pauses).

Speech rhythm was calculated based on Normalized Pairwise Variability Index (nPVI; Grabe and Low 2002), a measure of durational variability between pairs of syllables in an utterance. The measure is calculated based on the following equation:

$$nPVI = 100 \times \left[\sum_{k-1}^{m-1} \left| \frac{d_k - d_{k+1}}{d_k + d_{k+1}/2} \right| \Big/ (m-1) \right]$$

The equation calculates the absolute value of the difference in duration *d* of adjacent syllables divided by the mean duration of the two syllables. These quotients are then averaged for the utterance[2] and multiplied by 100. This method is used for evaluating relative changes in duration across syllables while controlling for effects of speech rate. Following the procedures of (Low and Grabe 1995, 2000), nPVI was calculated for individual utterances, as opposed to across all utterances. A lower nPVI indicates less variability in duration across syllables (i.e., a more uniform or staccato rhythm). We calculated nPVI for each utterance in the analysis, which typically corresponded with an intonational phrase.[3] To ensure that any differences in nPVI across groups were not driven solely by differences in the use of utterance-initial fillers, which tend to be longer in duration than typical words and are also known to be utilized less by individuals with ASD (Gorman et al. 2016), utterance-initial fillers (i.e., "um" and "uh") and utterance-initial syllables with a duration of greater than two standard deviations above the mean (300 ms) were deleted, as these elongated syllables can function in a similar way to fillers in delaying the onset of more regular, rhythmic speech (Clark and Fox Tree 2002). Henceforth, utterance-initial "um" and "uh" and elongated syllables will be collectively described as "fillers," unless otherwise specified. Finally, utterances for which speech rate or nPVI were greater or less than two standard deviations from the mean were treated as outliers and removed. This resulted in a trimming of less than 5% of the total data.

Clinical-Behavioral Measures and Correlates

Assessment of the BAP Personality Features

The Modified Personality Assessment Schedule (MPAS; Tyrer 1988) was used to assess the presence of personality traits of the BAP in the ASD parent group. Participants were asked a series of questions regarding personality traits that constitute the BAP, including social reticence, rigidity, and untactfulness. All interviews were rated by two coders, who were blind to participant group. Inter-rater reliability was 81%. Scores range from 0 (trait absent) to 2 (trait definitely present). Individuals were characterized as BAP(+) if they scored a 2 on the Social, Rigid, or Untactful traits of the MPAS. BAP(−) status was assigned when coders did not endorse traits on any of those subscales (scores < 2). These personality features are thought to mirror in quality the core domains of impairment in individuals with ASD, and have been shown to reliably distinguish ASD relatives from controls (Losh et al. 2008; Piven et al. 1997).

Pragmatic Language Ability

Pragmatic language skill was assessed using the Pragmatic Rating Scale-School Age (PRS-SA; Landa 2011) for individuals in the ASD and ASD control groups and using the Pragmatic Rating Scale (PRS; Landa et al. 1992) for individuals in the parent groups. The PRS-SA was rated from semi-structured interactions from the ADOS-2. Coders blind to group rated these interactions for pragmatic language

[2] nPVI is calculated on a by-utterance basis since speech rate is known to vary from utterance to utterance even in the speech of a single individual (Miller et al. 1984). Calculating the value for each utterance therefore gives us a more fine-grained measure of rhythmic variability for each participant.

[3] In theoretical models of prosody (Nespor and Vogel 1986; Pierrehumbert and Beckman 1988; Selkirk 1986), the intonational phrase is typically posited as the highest level within a hierarchy of prosodic constituents.

features on a three-point scale as absent (0), unknown (1), or present (2) with an average item-level reliability of 76%. The PRS was coded based on a semi-structured conversational interview (Life History Interview) in which an examiner asks the participant a series of questions about their social relationships, family, academic achievements, and occupation. As in the PRS-SA, two independent coders blind to group status rated videos of these interactions on a three-point scale as absent (0), mildly present (1), or clearly present (2) for specific pragmatic language features, which include subjective measures of prosody such as intonation patterns and speech rate. Average item-level reliability for the PRS was 86%. Both the PRS-SA and PRS include several items that comprise sub-domains of pragmatic skill, including a suprasegmental domain that includes ratings of atypical intonation, speech rate, and volume. Additional domains contributing to overall pragmatic language skill assessed on the PRS-SA include: (1) presupposition/theory of mind, which involves understanding the perspective of the listener to decide what and how information should be shared; (2) discourse management, which includes topic initiation, maintenance, and conversational reciprocity; and (3) non-verbal communicative behaviors, such as gestures, affect, and proxemics. Additional domains on the PRS include management of listener expectations, such as providing clarification as needed and reciprocating during conversation, and dominating conversation, which includes overly detailed language use and topic preoccupation.

Listener-Based Perceptual Ratings

Listener-based perceptual ratings were conducted on a subset of utterances produced during narrations from the ASD and ASD control groups. The third, eighth, and fourteenth utterance from each participant's narration was selected to represent utterances from the beginning, middle, and end of the picture book. Audio recordings of each of the selected utterances were low-pass filtered from 0 to 400 Hz to preserve prosodic information from the signal while removing higher frequencies important for speech intelligibility. Each utterance was rated on an 11-point Likert scale from −5 to 5 on variables of intonation (i.e., "flat" to "overly variable/sing-songy," respectively), rate (i.e., "too slow" to "too fast," respectively), and rhythm (i.e., "too staccato/choppy sound pattern" to "overly variable sound pattern," respectively), with ratings of 0 indicating "typical." To determine if ratings of the ASD group's intonation, rate, and rhythm differed from ratings of the ASD control group, all ratings were transformed and analyzed as deviant from a rating of "typical." After rating the intonation, rate, and rhythm, of each utterance, raters provided a judgement about how likely (definitely not, probably not, probably yes, definitely yes) it was that the speaker had ASD ("ASD likelihood"). For the purpose of planned analyses, ASD likelihood was recoded to a binary variable. Raters included 14 individuals ages 21–30 years with varying levels of experience with individuals with ASD (e.g., no direct experience, direct or indirect experience within an educational environment, interventionist, involvement with ASD research, family member with ASD). Raters with four or more years of experience working with individuals with ASD were classified as "expert raters" (n = 6) for subsequent analyses of perceptual ratings of prosody to determine if prior experience with individuals with ASD influenced ratings.

Analysis Plan

A series of mixed effects linear regression models were conducted in the ASD and parent groups separately, investigating mean F0, SD of F0, range of F0, utterance-final F0 excursion size, speech rate, and nPVI were fitted using the *lme4* package (Bates et al. 2014) for *R* statistical software. All models including the ASD and ASD control groups included fixed effects of group, utterance length (number of syllables), and chronological age. Though the groups differed in IQ, the inclusion of full-scale IQ did not strengthen the model, so it was excluded. Models including the parent groups included fixed effects for group and utterance length (number of syllables). Age-related differences in F0 would not be expected in the parent groups, so age was not included in the model. The model for nPVI also included a fixed effect of speech rate and a two-way interaction between family diagnosis and speech rate, as speech rate and nPVI are known to covary. By-subject random intercepts were also included in all models, as well as random slopes corresponding to all fixed effects. To investigate potential sex-related differences in prosody, additional models separately comparing the males and the females in the child and parent groups were conducted using the same parameters described above. In addition to assessing overall group differences in the ASD parent and parent control groups, models were run to assess differences based on BAP status using Tukey multiple comparison tests. It is important to note that BAP status was not available for all participants due to missing data. The BAP(+) group included 44 individuals (males: 22; females: 22) and the BAP(−) group included 41 individuals (males: 12; females: 29). Finally, binary mixed effects logistic regression models were run to determine how well acoustic variables of mean F0, SD of F0, range of F0, speech rate, nPVI, and utterance-final F0 excursion size predicted diagnosis. By-subject random slopes corresponding to each of the three fixed effect variables were included in the model. For all models, continuous predictors were mean-centered to reduce collinearity. Kappa values for the models were all < 6, indicating that collinearity between predictors was low and unlikely to affect model results. All *p* values were estimated

based on Kenward–Rogers degrees of freedom estimation using the package *lmerTest* for R. Finally, correlations between participant acoustic characteristics, clinical ratings of pragmatic language and ASD severity were examined.

Similar to analyses of acoustic measurements, a series of mixed effects linear regression models investigating listener-based perceptual ratings of intonation, rate, and rhythm were fitted using the *lme4* package (Bates et al. 2014) for *R* statistical software and included a fixed effect of group. Binary mixed effects logistic regressions were conducted to ascertain the effects of listener-based perceptual ratings of intonation, rate, and rhythm on raters' estimates of actual group membership. For logistic regressions, values for predictor variables were centered around the ASD control group mean ratings. Analyses were conducted including responses from all raters and subsequently repeated including responses from expert raters only. By-rater and by-utterance-number random slopes corresponding to all three fixed effects were included in the model. Pearson correlations were conducted to assess associations between listener-based perceptual ratings of prosody and acoustic measurements. Statistically, in addition to p values $< .05$, p values $< .10$ were additionally reported in order to address the exploratory nature of some of our aims, and to increase transparency of reporting to inform further research with larger sample sizes. Importantly, findings with p values falling within the .05–.10 range are interpreted with caution.

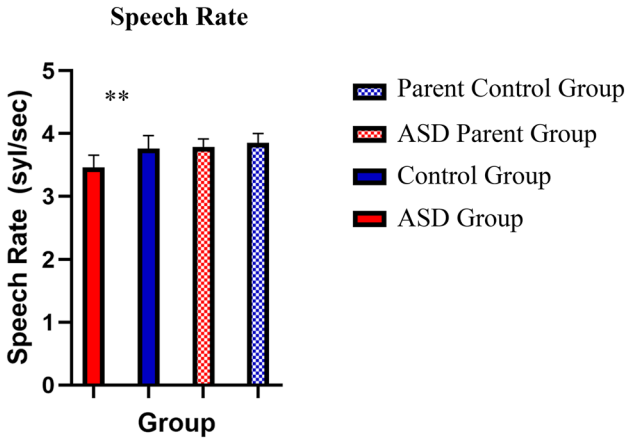

Fig. 1 Speech rate across groups. Individuals with ASD exhibited significantly slower speech rate compared to controls. Error bars denote the standard error. ** denotes $p < .01$

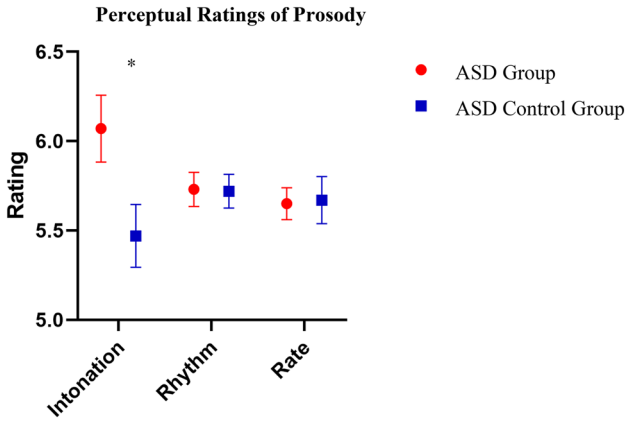

Fig. 2 Perceptual ratings of prosody in ASD and ASD Control groups. Listeners identified significantly atypical intonation, but not rhythm or rate in the ASD group. Error bars denote the standard error. * denotes $p < .05$

Results

ASD and ASD Control Groups

Group Differences in Acoustic Measurements

No significant group differences were observed for mean F0 ($\beta = .22$, $p = .20$), SD of F0 ($\beta = -.001$, $p = .96$), or range of F0 ($\beta = -.06$, $p = .67$). A marginally significant group difference was detected for utterance-final F0 excursion size, with the ASD group exhibiting a larger change in F0 in the utterance-final position ($\beta = -.12$, $p = .05$). Additionally, a main effect of group for speech rate ($\beta - .59$, $p < .01$) revealed a slower speech rate among individuals with ASD (see Fig. 1). There was no effect of group for nPVI (i.e., rhythm; $\beta = 1.74$, $p = .48$), indicating no group difference in paired syllable durations. When examining acoustic prosodic differences in males only, males with ASD demonstrated a larger utterance-final F0 excursion size than controls ($\beta = -.20$, $p < .05$), but no differences in mean F0 ($\beta = .11$, $p = .56$), SD of F0 (β .01, $p = .95$), range of F0 ($\beta = -.08$, $p = .70$), or nPVI ($\beta = 2.27$, $p = .48$). Males with ASD demonstrated slower speech rate than controls ($\beta = -.56$, $p < .05$). In analyses of females only, there was a marginally significant main effect of group on mean F0 ($\beta = .46$, $p = .09$) such that females with ASD exhibited a lower mean F0 compared to female controls. Additionally, females with ASD demonstrated a marginally larger utterance-final F0 excursion size than female controls ($\beta = -.17$, $p = .07$). There were no significant group differences in SD ($\beta = .24$, $p = .39$), range of F0 ($\beta = .06$, $p = .85$), speech rate ($\beta = -.25$, $p = .88$), or nPVI ($\beta = 2.69$, $p = .64$). Mixed effects binary logistic regression indicated no combination of acoustic variables were significant predictors of ASD group membership.

Group Differences in Listener-Based Perceptual Ratings

A significant group difference was observed for ratings of intonation ($\beta = -.60$, $p = .03$; see Fig. 2). For ratings by expert raters only, a main effect of group for intonation ($\beta = -.69$, $p = .03$) was detected, along with a marginal effect of group for ASD likelihood ($\beta = -.09$. $p = .05$) such that expert raters accurately identified individuals with a diagnosis of ASD as being more likely to have ASD based on the excerpted acoustic sample. Non-expert raters showed no such distinction ($\beta = -.04$. $p = .37$). Raters (expert and non-expert) showed no differences in ratings of rate or rhythm (all raters, rate: $\beta = .04$. $p = .72$, rhythm: $\beta = -.05$. $p = .72$; expert raters only, rate: $\beta = -.01$, $p = .98$, rhythm: $\beta = -.12$, $p = .44$).

A binary logistic regression was conducted to determine the types of perceptual ratings based on the short excerpted acoustic samples that best predicted group membership. Findings indicated that raters' perceptions of atypical speech rate and the likelihood of ASD diagnosis both significantly predicted ASD group membership ($\beta = -.306$, $z = -3.372$, $p < .001$; $\beta = .474$, $z = 2.935$, $p < .01$, respectively). Ratings of intonation were not a significant predictor of ASD group membership ($\beta = -0.005$, $z = -.10$, $p = .92$), nor were ratings of rhythm ($\beta = .06$, $z = 1.11$, $p = .26$). The model correctly predicted 91% of utterances from individuals in the ASD group and 14% of the control group's samples, giving an overall percentage correct prediction rate of 59%. Thus, while the model showed quite high sensitivity, specificity was low.

Relationships Between Acoustic Measurements and Listener-Based Perceptual Ratings

Pearson correlations revealed significant positive associations between F0 variability (SD of F0 and utterance-final F0 excursion) and listener-based perceptual ratings of intonation ($rs > .42$, $ps < .01$), rhythm ($rs > .33$, $ps < .05$), and rate ($rs > .28$; $ps < .05$). Additionally, range of F0 was positively correlated with listener-based perceptual ratings of intonation ($r = .32$, $p < .05$). Finally, speech rate was positively associated with perceptual ratings of rate and rhythm ($rs > .33$; $ps < .05$).

ASD Parent and Parent Control Groups

Significant parent group differences were observed for mean F0 ($\beta = -.46$, $p < .001$) and range of F0 ($\beta = .28$, $p = .01$), with the ASD parent group exhibiting a lower mean F0 and higher range of F0 compared to the parent control group. The ASD parent and parent control groups did not differ in SD of F0 ($\beta = .05$, $p = .68$) or utterance-final F0 excursion size ($\beta = .07$, $p = .15$). While there was no group difference in speech rate ($\beta = -.05$, $p = .44$), there was an interaction between group and utterance length ($\beta = -.06$, $p < .05$), such that parents of individuals with ASD showed greater slowing of speech rate as utterance length increased (see Fig. 3). There were no significant group differences on nPVI ($\beta = .08$, $p = .88$).

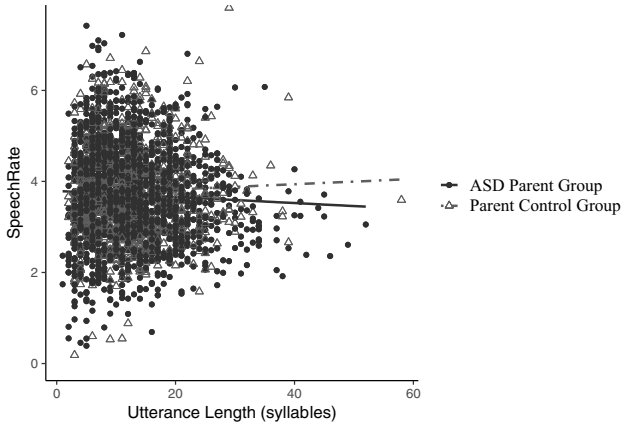

Fig. 3 Interaction between speech rate and utterance length in the ASD parent and parent control groups. The ASD parent group showed significantly slower speech rate at higher utterance lengths compared to the parent control group ($p < .05$)

When examining acoustic prosodic differences in fathers only, no significant group differences were detected for mean F0 ($\beta = -.31$, $p = .62$), SD of F0 ($\beta = .0003$, $p = 1.00$), range of F0 ($\beta = -.02$, $p = .96$), speech rate ($\beta = -.17$, $p = .42$), or nPVI ($\beta = 2.56$, $p = .34$). However, fathers of individuals with ASD demonstrated a marginally smaller utterance-final F0 excursion size than control fathers ($\beta = .13$, $p = .08$). Analyses of mothers revealed that mothers of individuals with ASD exhibited a significantly lower mean F0 and greater range of F0 compared to control mothers ($\beta = -.47$, $p < .001$; $\beta = .31$, $p < .01$, respectively). No significant group differences in SD of F0 ($\beta = .04$, $p = .79$) or utterance-final F0 excursion size ($\beta = .05$, $p = .33$), speech rate ($\beta = -.15$, $p = .41$), or nPVI ($\beta = 1.07$, $p = .65$) were detected in mothers.

BAP Group Differences

Differences in parent groups for mean F0 were driven by the BAP(+) group ($\beta = -.58$, $p < .01$). In particular the BAP group differences in mean F0 were driven by mothers with the BAP ($\beta = -.58$, $p = .02$; fathers: $\beta s < .34$, $ps > .10$). There was a marginal effect of BAP status on range of F0 ($\beta = .28$, $p < .09$) such that the BAP(+) group had a marginally higher range of F0 compared to parent controls. This marginal finding did not appear to be driven by sex (mothers: $\beta s < .29$, $ps > .13$; fathers: $\beta s < .09$, $ps > .96$). BAP status did not significantly influence SD of F0 or utterance-final F0 excursion size ($\beta s < .08$, $ps > .18$). Furthermore, sex differences by BAP group were not identified for SD of F0 (mothers: $\beta s < .07$, $p > .91$; fathers:

ßs < .14, *p* > .92) or utterance-final F0 excursion size (mothers: ßs < 3.0, ps > .71; fathers: *ßs* < 10.5, *ps* > .30). Analysis of the effect of BAP status on speech rate revealed a significant interaction between speech rate and utterance length, demonstrating greater slowing of speech rate with increased utterance length in the BAP(+) group compared to parent controls ($\beta = .08$, $p < .05$). Differences between the BAP(+) and BAP(−) groups did not reach significance ($\beta = .04$, $p = 0.31$). The interaction did not reach significance for either sex (mothers: $\beta = -.09$, $p = .11$; fathers: $\beta = -.02$, $p = .63$), though the relationship was stronger for mothers than for fathers. There were no significant differences in nPVI by BAP status overall (*ßs* < .36, *ps* > .50) or by sex (mothers: $\beta = .52$, $p = .53$; fathers: $\beta = .24$, $p = .75$).

Associations Between Acoustic Measurements of Prosody, ASD Severity, and Pragmatic Language

ASD and ASD Control Groups

In the ASD group, decreased speech rate was correlated with increased overall and social affect symptom severity based on the ADOS-2 ($r = -.33$, $p = .04$ and $r = -.39$, $p = .01$, respectively). Similarly, decreased speech rate was correlated with increased overall score on the ADOS-2 in the ASD control group ($r = -.79$, $p = .03$). Increased F0 range and decreased speech rate in the ASD group, were correlated with greater impairment in pragmatic language features related to theory of mind (e.g., not providing sufficient background information) based on the PRS-SA (r = .295, p = .049; r = .36, p = .02, respectively). Increased range of F0 in the ASD group was positively correlated with greater impairments in the suprasegmental domain of the PRS-SA ($r = .32$, $p = .03$). F0 excursion in the ASD group was positively correlated with ASD symptom severity in restricted and repetitive behaviors, driven primarily by males with ASD, ($r = .48$, $p < .01$), as well as increased pragmatic language impairment based on the PRS-SA ($r = .41$, $p < .01$).

ASD Parent and Parent Control Groups

Increased nPVI in the ASD parent group was associated with increased pragmatic violations overall, based on the PRS ($r = .32$, $p = .02$). This association was driven by the dominating conversation factor (e.g., overly detailed, tangential; $r = .34$, $p = .01$). Decreased range of F0 in the ASD parent group, particularly in mothers, was associated with increased violations in pragmatic language features associated with the managing listener expectations factor (e.g., fails to reciprocate, vague) during the PRS ($r = -.34$, $p < .01$). There were no significant associations between acoustic variables of prosody and clinical-behavioral measures in controls.

Discussion

The present study examined acoustic profiles of prosody in individuals with ASD and their parents during a narrative task, and examined relationships with clinical-behavioral features in individuals with ASD and among parents. Analyses also compared acoustic measures with listener-based perceptual ratings of prosody for individuals with ASD. Findings from analysis of acoustic measures in the ASD group were perhaps most revealing in the relatively few group differences that were evident, particularly in light of more robust differences that emerged in perceptual ratings for the ASD group, which were in line with prior literature using listener-based ratings (Baltaxe and Simmons 1985; Fine et al. 1991; Fosnot and Jun 1999; Landa et al. 1992; Losh et al. 2008, 2012b; Piven et al. 1997). For instance, no differences were observed between the ASD group and controls in key acoustic features such as F0 variability or rhythm (measured using nPVI). Objective acoustic measures were also minimally related to listener-based judgments, underscoring the complexity of the prosodic differences associated with ASD. That is, the clear distinctions in prosody observed by listeners (and repeatedly documented in prior studies) do not appear to be reflected in any basic, or readily measured set of acoustic properties. Acoustic features that differed in the ASD group were most notable in cases where differences were also detected among clinically unaffected parents, and driven by the subgroup of parents who demonstrated features of the BAP. These findings together may point towards a core set of acoustic features influenced by genetic liability to ASD.

Specifically, our study found speech rate to be an important variable distinguishing prosodic profiles in individuals with ASD and the BAP, supporting prior reports from perceptual ratings of speech in individuals with ASD and the BAP (Landa et al. 1992; Losh et al. 2012a). Findings that speech rate was overall slower for individuals with ASD, and also slower with increased utterance lengths in BAP(+) parents, could implicate cognitive processing differences in individuals with ASD and the BAP. Importantly, prior research in typically developing populations has shown associations between increased cognitive load and reduced speech rate (Griffin and Williams 1987; Huttunen et al. 2011). Further research is necessary to understand the various cognitive and social-pragmatic factors that may influence prosodic features, such as speech rate, in individuals with ASD and how these may interact with one another.

In addition to reduced speech rate, individuals with ASD exhibited increased utterance-final F0 excursion. While this differs from overall differences in intonation variability (e.g., SD of F0, range of F0) previously identified in individuals with ASD (Diehl et al. 2009; Edelson et al. 2007; Nadig and

Shaw 2012), use of more dynamic utterance-final F0 excursion may be related to prior descriptions of individuals with ASD inappropriately or ineffectively using questioning and statement intonation patterns (Paul et al. 2005; Peppé et al. 2007). Furthermore, despite a lack of overall acoustic differences in F0 variability, listener-based perceptual ratings revealed differences in ratings of intonation, suggesting that listeners are perceiving intonation differences in individuals with ASD that may not be due exclusively to differences in the F0 measurements examined here. One possibility is that additional acoustic measures of F0 need to be examined, such as the relative timing of pitch rises and falls, in addition to overall measures of pitch variability. Another possibility, consistent with prior work (Van Santen et al. 2010), is that listener perceptions of disordered prosody do not map directly to acoustic measures. It is well known, for example, that pitch and timing interact such that changes in timing can be perceived as changes in pitch, and vice versa (Lake et al. 2014; Yu 2010). Furthermore, it may be that differences in perceptual ratings reflect the confluence of speech properties impacting a listener's ratings, rather than any individual component of the speech signal. This is supported by associations identified between increased F0 variability (i.e., SD and range of F0) and listener-based perceptual ratings of increased atypicalities in intonation, speech rate, and rhythm. Ratings of ASD likelihood based on the individual's intonation, rate, and rhythm differed based on a rater's prior experience with individuals with ASD even though ratings of the aforementioned prosodic components did not differ between expert and non-expert raters. This further suggests that there may be additional speech characteristics, not necessarily contained within the prosodic components highlighted in this study (e.g., vocal intensity/volume, timing of pitch rises or falls), influencing ratings of ASD likelihood and that raters with more experience with individuals with ASD may be more attuned to these features. This conclusion is also supported by recent work by Redford et al. (2018) in which naïve listeners presented with short speech segments from individuals with ASD and typically developing controls identified the speech of individuals with ASD as atypical based on speech patterns related to speech motor control and voice quality, rather than prosodic components in particular. More specifically, listeners stated that they based their judgements on what they perceived to be slurred speech and differences in speech rate, leading Redford et al. (2018) to posit that intelligibility was largely responsible for discriminating the speech of an individual with ASD. Similarly, in the present study, ratings of speech rate were the strongest predictors of ASD group membership, suggesting that even when presented with filtered speech in which speech intelligibility is unpreserved, speech rate remains important in distinguishing atypical speech patterns.

Furthermore, in addition to a more slowed speech rate with increased utterance length, results revealed a reduced mean and increased range of F0 in parents of individuals with ASD, specifically those with the BAP. Interestingly, the pattern of reduced mean and increased range of F0 was driven by females with the BAP. While evidence of clear sex-specific differences in the expression of the BAP is limited (Rubenstein and Chawla 2018; Sasson et al. 2014; Seidman et al. 2012), recent work has demonstrated stronger associations between mothers with the BAP and their children on measures of language fluency compared to fathers with the BAP (Nayar et al. 2018). Results from the present study add to such evidence, perhaps suggesting a stronger maternal effect in the language-related phenotypes reflecting genetic liability to ASD. More broadly, lower mean F0 among females in the general population has been shown to index aspects of social identity, such as intellectual status or "nerdiness." Females may use such aspects of their speech to signal related virtues, such as intellectual involvement over social status, while at the same time signaling a rejection of mainstream ideals of femininity (Bucholtz 2012). In the context of this broader literature, female-specific speech patterns observed in the present study may similarly be socially-driven. Indeed, females with ASD have been shown to have a greater ability to communicate and socialize than males (Hiller et al. 2014; Sedgewick et al. 2016), and ASD has also been associated with academic talent among individuals without intellectual impairment (Baron-Cohen et al. 2007). It has also been cautiously proposed that there is a genetic link between ASD and certain types of academic talent (e.g., mathematical aptitude; Warrier and Baron-Cohen 2016), suggesting that such links may also be evident in individuals with the BAP.

Associations between acoustic measures of prosody and clinical-behavioral features provided further insight into the role of prosody within the overall presentation of ASD and the BAP. In the ASD and ASD control groups, decreased speech rate was associated with greater overall ASD severity, particularly in the social affect domain for individuals with ASD. Furthermore, correlations between acoustic measures during narrative and clinical-behavioral ratings of overall pragmatic language ability during semi-structured interactions suggest that cross-contextual atypicalities present differently in individuals with ASD and parents of individuals with ASD. In the ASD group, speech rate and range of F0 were related to broader pragmatic language atypicalities, including intonation and theory of mind-related pragmatic language skills. Relatedly, parents' speech rhythm and range of F0 were related to overall atypicalities in pragmatic language, suggesting that these prosodic elements may co-occur with or contribute to pragmatic language difficulties. Of note, however, is that group differences revealed increased

range of F0 in the ASD Parent group while decreased range of F0 was associated with increased pragmatic language atypicalities in this group. This discrepancy may be due, at least in part, to contextual differences in the collection of these measurements as prosody was assessed during narrative and pragmatic language atypicalities were measured during a conversational task. Taken together, findings suggest that prosody plays a key role in the larger social communication atypicalities that are observed in these groups, and may be contributing to what is qualitatively perceived as an "odd" or somewhat "socially awkward" interaction.

Clinical Implications

In light of a lack of specificity between perceptual ratings of prosody and acoustic differences in individuals with ASD, it may be beneficial to focus on multiple features of prosody (e.g., intonation modulation, rate, rhythm) rather than restricting focus to any singular feature. Results also support the use of a multi-method approach that includes both acoustic measurements and listener-based perceptual ratings of prosody. Overall, findings highlight the importance of targeting prosody in speech-language interventions and that such intervention may benefit overall pragmatic language abilities. Yet, limited prosody-focused interventions for individuals with ASD exist (Diehl and Paul 2009; Paul 2001), underscoring the need for collaboration between clinicians and researchers in this area. Indeed, acoustic differences in speech rate and utterance-final F0 excursion identified in the present study may serve as important objective benchmarks for speech-language interventions.

Limitations and Future Directions

The present study assessed prosody in individuals with ASD and their parents within a specific context. To better understand patterns of strengths and challenges, as well as genetically meaningful phenotypes associated with prosody, it will be important for future work to systematically assess prosody in more and less structured tasks in individuals with ASD and their family members given that varying social and cognitive demands across contexts may impact prosody. While the present study made several contributions to the literature in terms of assessesing sex differences in prosody in individuals with ASD, future work should seek to include a larger sample of females to expand on the present findings. Importantly, though the present study identified slower speech rate in the ASD group overall, this finding appeared to be driven by males with ASD, suggesting that speech rate may be an area of relative strength for females with ASD. Additional evidence is necessary to support this conclusion and would have broader implications for speech-language interventions tailored for females with ASD, who are known to exhibit a different clinical presentation from males (Head et al. 2014; Mandy et al. 2012; Sedgewick et al. 2016; Van Wijngaarden-Cremers et al. 2014). Furthermore, the present study focused solely on prosodic *production*. However, further insight into the prosodic profile in individuals with ASD may be gained through investigation of *receptive* prosody skills.

Conclusion

In sum, this study revealed some key prosodic differences in individuals with ASD and the BAP, using a comprehensive suite of acoustic analysis in relation to clinical-behavioral characteristics, as well as perceptual ratings of prosody in the ASD group. In particular, prosodic differences in speech rate appear impacted in both individuals with ASD and clinically unaffected relatives with the BAP, in addition to differences in mean and range of F0 in individuals with the BAP, particularly among females with the BAP. Complementary analyses on acoustic measures and listener-based perceptual ratings demonstrate complexity of the prosodic differences in individuals with ASD such that listeners' access to the combination of features available in the speech signal distinguish individuals with ASD even when objective measurements of discrete prosodic components may not.

Acknowledgements The authors are thankful to all of the individuals who participated in this research. The authors would also like to thank Chun Liang Chan for his assistance with Praat scripting and computer programming. This study was funded by Grants from the National Institutes of Health (R01DC010191, R01MH091131), the National Science Foundation (BCS-0820394, DGE-1324585), and P30 HD03110.

Author Contributions ML conceived the study and oversaw all data collection, analyses, and manuscript preparation. SP, KN, KF, and SC assisted in data acquisition, processing, and analysis, and contributed to manuscript preparation. GEM and JJD assisted with data interpretation and manuscript preparation. All authors read and approved the final manuscript.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in this study involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Asperger, H., & Frith, U. (1991). "Autistic psychopathy" in childhood. In *Autism and Asperger syndrome* (pp. 37–92). Cambridge: Cambridge University Press.

Bachorowski, J.-A. (1999). Vocal expression and perception of emotion. *Current Directions in Psychological Science, 8*(2), 53–57. https://doi.org/10.1111/1467-8721.00013.

Bachorowski, J.-A., & Owren, M. J. (1995). Vocal expression of emotion: Acoustic properties of speech are associated with emotional intensity and context. *Psychological Science, 6*(4), 219–224. https://doi.org/10.1111/j.1467-9280.1995.tb00596.x.

Baltaxe, C., Simmons, J. Q., & Zee, E. (1984). Intonation patterns in normal, autistic and aphasic children. In *Proceedings if the tenth internation congress of phonetic sciences* (pp. 713–718).

Baltaxe, C. A. M., & Simmons, J. Q. (1985). Prosodic development in normal and autistic children. In *Communication problems in Autism* (pp. 95–125). New York: Plenum Press https://doi.org/10.1007/978-1-4757-4806-2_7.

Baron-Cohen, S., & Staunton, R. (1994). Do children with autism acquire the phonology of their peers? An examination of group identification through the window of bilingualism. *First Language, 14*(42–43), 241–248. https://doi.org/10.1177/014272379401404216.

Baron-Cohen, S., Wheelwright, S., Burtenshaw, A., & Hobson, E. (2007). Mathematical talent is linked to autism. *Human Nature, 18*(2), 125–131. https://doi.org/10.1007/s12110-007-9014-0.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *ArXiv Preprint* ArXiv:1406.5823.

Bent, T., Baese-Berk, M., Borrie, S. A., & McKee, M. (2016). Individual differences in the perception of regional, nonnative, and disordered speech varieties. *The Journal of the Acoustical Society of America, 140*(5), 3775–3786. https://doi.org/10.1121/1.4966677.

Boersma, P., & Weenink, D. (2017). *Praat: Doing phonetics by computer*. Retrieved from http://www.praat.org/.

Bucholtz, M. (2012). *White kids, language, race, and styles of youth identity*. New York: Cambridge University Press.

Capps, L., Losh, M., & Thurber, C. (2000). The frog ate a bug and made his mouth sad: Narrative competence in children with autism. *Journal of Abnormal Child Psychology, 28,* 193–204.

Clark, H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition, 84*(1), 73–111. https://doi.org/10.1016/S0010-0277(02)00017-3.

Diehl, J. J., Bennetto, L., & Young, E. C. (2006). Story recall and narrative coherence of high-functioning children with autism spectrum disorders. *Journal of Abnormal Child Psychology, 34*(1), 87–102. https://doi.org/10.1007/s10802-005-9003-x.

Diehl, J. J., & Paul, R. (2009). The assessment and treatment of prosodic disorders and neurological theories of prosody. *International Journal of Speech-Language Pathology*, *11*(4), 287–292. https://doi.org/10.1080/17549500902971887.

Diehl, J. J., & Paul, R. (2013). Acoustic and perceptual measurements of prosody production on the profiling elements of prosodic systems in children by children with autism spectrum disorders. *Applied Psycholinguistics, 34*(1), 135–161. https://doi.org/10.1017/S0142716411000646.

Diehl, J. J., Watson, D., Bennetto, L., Mcdonough, J., & Gunlogson, C. (2009). An acoustic analysis of prosody in high-functioning autism. *Applied Psycholinguistics, 30*(03), 385.

Edelson, L., Grossman, R., & Tager-Flusberg, H. (2007). *Emotional prosody in children and adolescents with autism*. Poster session presented at the annual international meeting for Autism Research, Seattle, WA.

Fay, W. H. (1969). On the basis of autistic echolalia. *Journal of Communication Disorders, 2*(1), 38–47. https://doi.org/10.1016/0021-9924(69)90053-7.

Fine, J., Bartolucci, G., Ginsberg, G., & Szatmari, P. (1991). The use of intonation to communicate in pervasive developmental disorders. *Journal of Child Psychology and Psychiatry, 32*(5), 771–782.

Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. *Pediatric Research*, *65*(6), 591–598. https://doi.org/10.1203/PDR.0b013e31819e7203.

Fosnot, S. M., & Jun, S. (1999). Prosodic characteristics in children with stuttering or autism during reading and imitation. In *14th international congress of phonetic sciences* (pp. 103–115).

Gorman, K., Olson, L., Hill, A. P., Lunsford, R., Heeman, P. A., & van Santen, J. P. H. (2016). Uh and um in children with autism spectrum disorders or language impairment. *Autism Research: Official Journal of the International Society for Autism Research, 9*(8), 854–865. https://doi.org/10.1002/aur.1578.

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Laboratory Phonology, 7,* 515–546.

Griffin, G. R., & Williams, C. E. (1987). The effects of different levels of task complexity on three vocal measures. *Aviation, Space and Environmental Medicine, 58*(12), 1165–1170.

Head, A. M., McGillivray, J. A., & Stokes, M. A. (2014). Gender differences in emotionality and sociability in children with autism spectrum disorders. *Molecular Autism, 5*(1), 19. https://doi.org/10.1186/2040-2392-5-19.

Hiller, R. M., Young, R. L., & Weber, N. (2014). Sex differences in autism spectrum disorder based on DSM-5 criteria: Evidence from clinician and teacher reporting. *Journal of Abnormal Child Psychology, 42*(8), 1381–1393. https://doi.org/10.1007/s10802-014-9881-x.

Hus, V., Gotham, K., & Lord, C. (2014). Standardizing ADOS domain scores: Separating severity of social affect and restricted and repetitive behaviors. *Journal of Autism and Developmental Disorders, 44*(10), 2400–2412. https://doi.org/10.1007/s10803-012-1719-1.

Huttunen, K. H., Keränen, H. I., Pääkkönen, R. J., Päivikki Eskelinen-Rönkä, R., & Leino, T. K. (2011). Effect of cognitive load on articulation rate and formant frequencies during simulator flights. *The Journal of the Acoustical Society of America, 129*(3), 1580–1593. https://doi.org/10.1121/1.3543948.

Idring, S., Lundberg, M., Sturm, H., Dalman, C., Gumpert, C., Rai, D., … Magnusson, C. (2015). Changes in prevalence of autism spectrum disorders in 2001-2011: Findings from the Stockholm youth cohort. *Journal of Autism and Developmental Disorders*, *45*(6), 1766–1773. https://doi.org/10.1007/s10803-014-2336-y.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 35*(2), 217–250. https://doi.org/10.1105/tpc.11.5.949.

Lake, J. I., LaBar, K. S., & Meck, W. H. (2014). Hear it playing low and slow: How pitch level differentially influences time perception. *Acta Psychologica, 149,* 169–177. https://doi.org/10.1016/j.actpsy.2014.03.010.

Landa, R. (2011). *Pragmatic rating scale for school-age children*. Baltimore: Kennedy Krieger Institute.

Landa, R., Piven, J., Wzorek, M. M., Gayle, J. O., Chase, G. A., & Folstein, S. E. (1992). Social language use in parents of autistic individuals. *Psychological Medicine, 22*(1), 245–254.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., …, Rutter, M. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*(3), 205–223.

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule–2nd edition*. Los Angeles, CA: Western Psychological Corporation.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasice developmental disorders. *Journal of Autism and Developmental Disorders, 24*(5), 659–685.

Losh, M., Adolphs, R., Poe, M. D., Couture, S., Penn, D., Baranek, G. T., et al. (2009). Neuropsychological profile of autism and the broad autism phenotype. *Archives of General Psychiatry, 66*(5), 518–526. https://doi.org/10.1001/archgenpsychiatry.2009.34.

Losh, M., & Capps, L. (2003). Narrative ability in high-functioning children with autism or Asperger's syndrome. *Journal of Autism and Developmental Disorders, 33*(3), 239–251.

Losh, M., Childress, D., Lam, K., & Piven, J. (2008). Defining key features of the broad autism phenotype: A comparison across parents of multiple- and single-incidence autism families. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics: The Official Publication of the International Society of Psychiatric Genetics, 147B*(4), 424–433. https://doi.org/10.1002/ajmg.b.30612.

Losh, M., Klusek, J., Martin, G. E., Sideris, J., Parlier, M., & Piven, J. (2012a). Defining genetically meaningful language and personality traits in relatives of individuals with fragile X syndrome and relatives of individuals with autism. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics: The Official Publication of the International Society of Psychiatric Genetics, 195B*(6), 660–668. https://doi.org/10.1002/ajmg.b.32070.

Losh, M., Klusek, J., Martin, G. E., Sideris, J., Parlier, M., & Piven, J. (2012b). Defining genetically meaningful language and personality traits in relatives of individuals with fragile X syndrome and relatives of individuals with autism. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics: The Official Publication of the International Society of Psychiatric Genetics, 159B*(6), 660–668. https://doi.org/10.1002/ajmg.b.32070.

Low, E. L., & Grabe, E. (1995). Prosodic patterns in Singapore English. *Proceedings of the XIIIth International Congress of Phonetic Science, 3,* 636–639.

Low, E. L., & Grabe, E. (2000). Quantitative characterizations of speech rhythm: Syllabletiming in Singapore English. *Language and Speech, 43,* 377–401.

Mandy, W., Chilvers, R., Chowdhury, U., Salter, G., Seigal, A., & Skuse, D. (2012). Sex differences in autism spectrum disorder: Evidence from a large sample of children and adolescents. *Journal of Autism and Developmental Disorders, 42*(7), 1304–1313. https://doi.org/10.1007/s10803-011-1356-0.

Mayer, M. (1969). *Frog, where are you?*. New York: Dial Press.

McCann, J., & Peppé, S. (2003). Prosody in autism spectrum disorders: A critical review. *International Journal of Language & Communication Disorders/Royal College of Speech & Language Therapists, 38,* 325–350. https://doi.org/10.1080/13682820310001542 04.

Mesibov, G. B. (1992). Treatment issues with high-functioning adolescents and adults with autism. In *High-funtioning individuals with autism* (pp. 143–155). New York: Springer.

Miller, J., Grosjean, F., & Lomanto, C. (1984). Articulation rate and its variability in spontaneous speech: An analysis and some implications. *Phonetica, 41,* 215–225.

Nadig, A., & Shaw, H. (2012). Acoustic and perceptual measurement of expressive prosody in high-functioning autism: Increased pitch range and what it means to listeners. *Journal of Autism and Developmental Disorders, 42*(4), 499–511.

Nayar, K., Gordon, P. C., Martin, G. E., Hogan, A. L., La Valle, C., McKinney, W., …, Losh, M. (2018). Links between looking and speaking in autism and first-degree relatives: Insights into the expression of genetic liability to autism. *Molecular Autism*, *9*, 51. https://doi.org/10.1186/s13229-018-0233-5.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris Publications.

Oerlemans, A. M., van der Meer, J. M. J., van Steijn, D. J., de Ruiter, S. W., de Bruijn, Y. G. E., de Sonneville, L. M. J., …, Rommelse, N. N. J. (2014). Recognition of facial emotion and affective prosody in children with ASD (+ADHD) and their unaffected siblings. *European Child & Adolescent Psychiatry*, *23*(5), 257–271. https://doi.org/10.1007/s00787-013-0446-2.

Paul, R. (2001). *Language disorders from infancy through adolescence: Assessment and intervention* (2nd ed.). St. Louis, MO: Mosby.

Paul, R., Augustyn, A., Klin, A., & Volkmar, F. R. (2005). Perception and production of prosody by speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 35*(2), 205–220.

Paul, R., Bianchi, N., Augustyn, A., Klin, A., & Volkmar, F. R. (2008). Production of syllable stress in speakers with autism spectrum disorders. *Research in Autism Spectrum Disorders, 2*(1), 110–124. https://doi.org/10.1016/j.rasd.2007.04.001.

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 & Phonetics, 17*(4–5), 345–354.

Peppé, S., McCann, J., Gibbon, F., O'Hare, A., & Rutherford, M. (2006). Assessing prosodic and pragmatic ability in children with high-functioning autism. *Journal of Pragmatics, 38,* 1776–1791. https://doi.org/10.1016/j.pragma.2005.07.004.

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*(4), 1015–1028.

Pierrehumbert, J. B., & Beckman, M. (1988). Japanese tone structure. In *Linguistic inquiry monograph 15*. Cambridge: MIT Press.

Piven, J., Palmer, P., Jacobi, D., Childress, D., & Arndt, S. (1997a). Broader autism phenotype: Evidence from a family history study of multiple-incidence autism families. *American Journal of Psychiatry, 154*(2), 185–190.

Piven, J., Palmer, P., Landa, R., Santangelo, S., Jacobi, D., & Childress, D. (1997b). Personality and language characteristics in parents from multiple-incidence autism families. *American Journal of Medical Genetics, 74*(4), 398–411.

Pronovost, W., Wakstein, M. P., & Wakstein, D. J. (1966). A longitudinal study of the speech behavior and language comprehension of fourteen children diagnosed atypical or autistic. *Exceptional Children, 33*(1), 19–26.

Redford, M. A., Kapatsinski, V., & Cornell-Fabiano, J. (2018). Lay listener classification and evaluation of typical and atypical children's speech. *Language and Speech, 61*(2), 277–302. https://doi.org/10.1177/0023830917717758.

Rosenfelder, I. (2013). *Forced Alignment & Vowel Extraction (FAVE): An online suite for automatic vowel analysis*. Philadelphia: University of Pennsylvania Linguistics Lab.

Rubenstein, E., & Chawla, D. (2018). Broader autism phenotype in parents of children with autism: A systematic review of percentage estimates. *Journal of Child and Family Studies, 27*(6), 1705–1720. https://doi.org/10.1007/s10826-018-1026-3.

Ruser, T. F., Arin, D., Dowd, M., Putnam, S., Winklosky, B., Rosen-Sheidley, B., …, Folstein, S. (2007). Communicative competence in parents of children with autism and parents of children with specific language impairment. *Journal of Autism and Developmental Disorders*, *37*(7), 1323–1336. https://doi.org/10.1007/s10803-006-0274-z.

Sasson, N. J., Faso, D. J., Parlier, M., Daniels, J. L., & Piven, J. (2014). When father doesn't know best: Selective disagreement between self-report and informant report of the broad autism phenotype in parents of a child with autism. *Autism Research: Official Journal of the International Society for Autism Research, 7*(6), 731–739. https://doi.org/10.1002/aur.1425.

Sedgewick, F., Hill, V., Yates, R., Pickering, L., & Pellicano, E. (2016). Gender differences in the social motivation and friendship experiences of autistic and non-autistic adolescents. *Journal of Autism and Developmental Disorders, 46*(4), 1297–1306. https://doi.org/10.1007/s10803-015-2669-1.

Seidman, I., Yirmiya, N., Milshtein, S., Ebstein, R. P., & Levi, S. (2012). The Broad Autism Phenotype Questionnaire: Mothers versus fathers of children with an autism spectrum disorder. *Journal of Autism and Developmental Disorders, 42*(5), 837–846. https://doi.org/10.1007/s10803-011-1315-9.

Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology Yearbook, 3,* 371–405.

Selkirk, E. (2009). On clause and intonational phrase in Japanese: The syntactic grounding of prosodic constituent structure. *Gengo Kenkyu*, *136*, 37–73.

Selkirk, E. (2011). The syntax-phonology interface. In J. Goldsmit, J. Riggle, & A. Yu (Eds.), *The handbook of phonological theory* (2nd ed.). Oxford: Blackwell.

Shriberg, L., Paul, R., McSweeny, J., Klin, A., & Volkmar, F. R. (2001). Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. *Journal of Speech, Language, and Hearing Research, 44*(5), 1097–1115.

Tager-Flusberg, H., & Sullivan, K. (1995). Attributing mental states to story characters: A comparison of narratives produced by autistic and mentally retarded individuals. *Applied Psycholinguistics, 16*(03), 241–256. https://doi.org/10.1017/S0142716400007281.

Tyrer, P. (1988). Personality assessment schedule. In P. Tyrer (Ed.), *Personality disorders: Diagnosis, management, and course* (pp. 51–100). London: Butterworth and Company.

Van Bourgondien, M. E., & Woods, A. V. (1992). Vocational possibilities for high-functioning adults with autism. In *High-functioning individuals with autism* (pp. 227–239). New York: Springer.

Van Santen, J. P. H., Prud'hommeaux, E. T., Black, L. M., & Mitchell, M. (2010). Computational prosodic markers for autism. *Autism: The International Journal of Research and Practice, 14*(3), 215–236. https://doi.org/10.1177/1362361310363281.

Van Wijngaarden-Cremers, P. J. M., van Eeten, E., Groen, W. B., Van Deurzen, P. A., Oosterling, I. J., & Van der Gaag, R. J. (2014). Gender and age differences in the core triad of impairments in autism spectrum disorders: A systematic review and meta-analysis. *Journal of Autism and Developmental Disorders, 44*(3), 627–635. https://doi.org/10.1007/s10803-013-1913-9.

Warrier, V., & Baron-Cohen, S. (2016). Genetics of mathematical aptitude. In *eLS* (pp. 1–6). American Cancer Society. https://doi.org/10.1002/9780470015902.a0026848.

Wechsler, D. (1997). *Wechsler adult intelligence scale-third edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence (WASI)*. Cambridge, UK: Pearson As.

Wechsler, D. (2003). *Wechsler intelligence scale for children-fourth edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2008). *Wechsler adult intelligence scale-fourth edition*. San Antonio, TX: The Psychological Corporation.

Wells, B., & Peppé, S. (2003). Intonation abilities of children with speech and language impairments. *Journal of Speech, Language, and Hearing Research : JSLHR, 46*(1), 5–20.

Wells, B., Peppé, S., & Goulandris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language, 31*(4), 749–778. https://doi.org/10.1017/S030500090400652X.

Werling, D. M., & Geschwind, D. H. (2013). Understanding sex bias in autism spectrum disorder. *Proceedings of the National Academy of Sciences*, *110*(13), 4868–4869. https://doi.org/10.1073/pnas.130160211.

Yu, A. C. L. (2010). Tonal effects on perceived vowel duration. *Laboratory Phonology, 10,* 151–168.

Yuan, J., & Liberman, M. (2008). Speaker identification on the SCOTUS corpus. *The Journal of the Acoustical Society of America, 123*(5), 3878. https://doi.org/10.1121/1.2935783.