Rethinking Prosody Production in Autism: Nuanced Insights from Individual Differences and Network Analysis Approaches

Talia Liu[1], Kelsey E. Davison[1], Ayelet M. Kershenbaum[2], Ethan Weed[3], John D. E. Gabrieli[4], Helen Tager-Flusberg[5], Jennifer Zuk[1]

[1]Department of Speech, Language, and Hearing Sciences, Boston University, Boston, MA

[2]Program in Speech and Hearing Bioscience and Technology, Harvard University, Cambridge, MA

[3]Department of Linguistics, Cognitive Science, and Semiotics, Aarhus University, Aarhus, Denmark

[4]McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA

[5]Department of Psychological & Brain Sciences, Boston University, MA

Key words: autism, prosody, language, individual differences, network analysis, community detection

Address correspondence to:

Dr. Jennifer Zuk

jzuk@bu.edu

Abstract

Purpose: Prosodic differences between autistic and non-autistic individuals are recognized, but there is a lack of consensus on the specific prosodic features that characterize the "autistic voice" due to widespread heterogeneity and mixed findings. This study seeks to build further understanding of the nuances of prosody in autism through individual differences and network analyses.

Method: Acoustic analyses were conducted from 66 school-age autistic and non-autistic children and adolescents' narrative generation. Between-group analyses of pitch and timing-related prosodic features were conducted, followed by within-group analyses to investigate associations between prosodic features and individual differences in overall language skills. Thereafter, established network analysis methods (Weed et al., 2023) were adopted to detect communities of participants based on similar prosodic features.

Results: Initial between-group analyses revealed that autistic participants demonstrated greater pitch range and variation, as well as slower speech and articulation rates, compared to non-autistic peers. Subsequent analyses revealed that speech and articulation rates were associated with overall language skills. In line with previous findings, the community detection algorithm identified three clusters of participants: one with more autistic participants, one with more non-autistic participants, and one with largely equal representation of autistic and non-autistic participants.

Conclusions: Although between-group differences consistent with similar previous literature have been indicated, community detection analyses further support the notion that prosody in autism may be "different in different ways" (Weed et al., 2023). This work highlights the importance of moving beyond group-difference approaches in uncovering nuances to individual differences in prosody via within-group and data-driven analysis approaches.

Introduction

Prosody is a central aspect of spoken communication which conveys meaning, emotion, and linguistic meaning and structure via changes in intonation, intensity (loudness), and timing and is used on a day-to-day basis to facilitate effective social communication (Cole, 2015; Edelson & Diehl, 2013). Differences in speech prosody, on the other hand, can become a barrier to communication and lead to communication breakdowns if one's intent is misinterpreted or misperceived. Consequently, these potential breakdowns in communication can adversely impact interpersonal communicative interactions (Wiklund, 2016). As autism is characterized largely by differences in social communication, prosody in autism has been widely studied and of interest to researchers for decades (Baltaxe & Simmons, 1985; McCann & Peppé, 2003). In fact, prosody has been shown to contribute to negative first impressions of autistic individuals based on non-autistic peers' first-impression judgments (Sasson et al., 2017). Unfortunately, these negative perceptions may ultimately relate to and impact autistic individuals' smaller social networks, vocational outcomes, and, importantly, reduced quality of life (van Heijst & Geurts, 2014). Given the importance of prosody within daily functioning, research has particularly sought to address whether autistic individuals present with distinctive prosodic profiles. Although prosodic differences may be perceptually salient in autism (de Marchena & Miller, 2017), existing quantitative findings are incredibly mixed (Asghari et al., 2021; Fusaroli et al., 2017, 2022; McCann & Peppé, 2003). Here, we attempt to contextualize existing inconsistent findings and better understand the nuances of prosody in autism by accounting for individual differences and employing data-driven analytic approaches to characterize participants based on similar prosodic features.

Despite perceptual recognition of prosodic differences in autism, there is limited consensus on whether and how prosody production quantitatively differs between autistic and non-autistic individuals (see Asghari et al., 2021; Fusaroli et al., 2017; Ma et al., 2023 for reviews). Human raters, including both naïve listeners and clinicians, have been shown to reliably distinguish autistic individuals from non-autistic individuals on the basis of audio recordings (Cho et al., 2019; Hubbard & Trauner, 2007; Nadig & Shaw, 2012; Redford et al., 2018; Weed et al., 2023). Perceptual characterizations of prosody in autism range from "monotone" to "singsong" pitch variation when speaking and markedly "fast" or "slow" speech rate (Diehl et al., 2009; Weed et al., 2023; Wiklund, 2023). As a result, it is unsurprising that some studies have demonstrated quantitative "differences" in prosodic features of prosody between autistic and non-autistic individuals (e.g., Diehl et al., 2009; Nadig & Shaw, 2012; Patel et al., 2020), while others have not (e.g., Dahlgren et al., 2018; Nadig & Shaw, 2012). Despite evidence of salient perceptual differences in prosody in autism, inconsistent quantitative findings may be attributed to a number of factors, including variation in the context sampled, a lack of accounting for the well-known widespread heterogeneity of language and social communication skills in autism (Kjelgaard & Tager-Flusberg, 2001; Lord et al., 2020), and differences in methodological approaches employed.

Variability in how prosody production is elicited (i.e., different sampling contexts) and measured (Asghari et al., 2021; Fusaroli et al., 2017) is one likely source of inconclusive findings in much of the existing literature on prosody in autism. A large body of existing research has measured prosody production through controlled and structured laboratory experiments (Asghari et al., 2021). While structured tasks enable the sampling of more controlled and consistent linguistic contexts across participants, these tasks limit the generalizability of findings to natural language contexts where social communication is highly dynamic and context-dependent.

Capturing speech prosody during more naturalistic language sampling offers an important opportunity to quantify prosodic features within a more ecologically valid and generalizable context (e.g., Dahlgren et al., 2018). In particular, narrative contexts elicit spontaneous speech within a semi-structured, more naturalistic manner and have been used in a smaller body of research to assess prosody in autistic children and adolescents (Dahlgren et al., 2018; Diehl et al., 2009; Patel et al., 2020; Thurber & Tager-Flusberg, 1993). In studies where prosody production was specifically measured via narrative language sampling, autistic individuals' speech has been characterized by greater pitch variation (Chan & To, 2016; Diehl et al., 2009) and slower speech rate (Patel et al., 2020) compared to non-autistic peers. Quantifying prosody through tasks like narrative generation provides a valuable opportunity to explore prosodic features in more naturalistic contexts, enhancing the generalizability of findings.

Considering the widespread heterogeneity in language skills in autism (Kjelgaard & Tager-Flusberg, 2001), there is a striking lack of research addressing the potential impact of language skills on prosody in autism, which could likely be another source of mixed existing findings. In particular, existing literature predominantly overlooks the potential role of individual differences in language skills and specifically how prosody and language skills may be intricately linked. Initial evidence has demonstrated that performance on a structured receptive and expressive prosody task is positively associated with a crude indicator of receptive language (i.e., vocabulary) skills (McCann et al., 2007). Although one option for discerning prosodic differences independent of potential effects of language could be to match autistic and non-autistic groups on language skills (e.g., Diehl et al., 2009), this approach can be problematic when it leads to the exclusion of participants based on lower language skills and results in limited representation of heterogeneous language skills. Therefore, it is important to shift attention toward the heterogeneity within autism and capture language variability within study samples (Fusaroli et al., 2022). Acknowledging the highly heterogeneous language skills in autism, further research is needed that directly investigates to what extent prosody may be associated with individual differences in language skills.

Inconsistent findings among existing research on prosody in autism are also likely in part due to the use of different methodological and analytic approaches, resulting in difficulty comparing and generalizing conclusions across studies (Fusaroli et al., 2017; Grice et al., 2023; McCann & Peppé, 2003). Whereas more traditional analysis approaches (i.e., between-group analyses) seek to detect group differences between autistic and non-autistic individuals, a data-driven, network-based approach is naïve to individual participants' diagnosis and enables analysis of multiple prosodic features simultaneously. Recent exploratory findings suggest that an unsupervised, network-based approach – specifically, a community detection algorithm (Reichardt & Bornholdt, 2006) – provides a more nuanced view on prosody in autism (Weed et al., 2023). In this proof-of-concept study, the sample included 26 total autistic and non-autistic adolescents speaking the same, controlled sentences (Weed et al., 2023). Findings revealed that the community detection algorithm identified three clusters of participants based on speech rate, articulation rate, pitch variation, and jitter (features commonly investigated in prosody in autism research): One community contained relatively more autistic participants, one community contained relatively more non-autistic participants, and a third community included a largely equal representation of autistic and non-autistic participants (Weed et al., 2023). While human raters were able to reliably identify autistic participants based on perceptual ratings, none of the communities identified by the algorithm were exclusively distinguished by autism diagnosis. These findings reinforce the idea that not one single prosodic feature, or even a specified set of prosodic features quantified, is sufficient to pinpoint an "autistic voice". Therefore, such prosodic features alone cannot consistently differentiate autistic individuals from non-autistic individuals (Fusaroli et al., 2022; Weed et al., 2023). Thus, Weed and colleagues (2023, p. 14) proposed the notion that prosody in autism, while perhaps perceptually distinct, may be “different in different ways.” In modeling prosody production from autistic and non-autistic participants as a network of nodes defined by prosodic features, this initial study provides an important foundation for moving beyond a between-group differences approach that assesses individual prosodic features in isolation.

The Present Study

In an effort to clarify mixed findings and further our understanding of prosody production in autism, the present study investigated prosody during narrative generation in a sample of autistic and non-autistic children and adolescents while accounting for language abilities using three analytic approaches: a between-group differences approach, an individual differences approach, and a data-driven approach. As an initial step, we first took a more traditional analysis methodology to determine whether prosodic features differ between autistic and non-autistic participants in the current sample of children and adolescents specifically completing a narrative task. In line with existing studies on prosody in autism within a narrative context, we hypothesize that autistic participants’ prosody would be characterized by greater pitch range and variation (Chan & To, 2016; Diehl et al., 2009) and slower speech rate (Patel et al., 2020) and compared to non-autistic peers. In light of calls to account for individual differences (Fusaroli et al., 2022) and the unmet need to account for language variability in autism, the second analysis approach moved beyond between-group (autistic vs. non-autistic) comparisons by investigating individual differences in prosody in relation to overall language skills separately within groups. Here, we hypothesized that prosodic features would be associated with overall language skills (as indicated by a composite score). The third analysis method utilized Weed and colleagues’ (2023) novel data-driven approach to identify clusters of participants based on similar acoustic profiles, making no *a priori* assumptions about participants' diagnostic status. In the present study, we sought to replicate and extend their approach by eliciting prosody production from a more naturalistic context (i.e., narrative generation versus reciting standard, memorized sentences in Weed et al. (2023)) in a larger sample that includes more variable language skills among autistic participants. In line with this closely related work from Weed and colleagues, a similar pattern of classification is hypothesized, such that classification may reinforce the notion that there may not be a singular "autistic voice"; rather, autistic voices may be "different in different ways" (Weed et al., 2023, p. 14). Taken together, the present study aimed to illuminate the importance of moving beyond group-difference approaches to uncover nuances to individual differences in prosody via within-group and data-driven analysis approaches.

Methods

Participants

Sixty-six participants ranged from 5-17 years old and included 30 autistic (26 male, M = 10.4 years, SD = 2.85) and 36 non-autistic participants (24 male, M = 10.2 years, SD = 2.44) (Table 1). Participants were recruited from the greater Boston area as part of a larger longitudinal neuroimaging study (Lu et al., 2016; O'Brien et al., 2022). All participants were English speakers, right-handed, and born after 32 weeks gestation. Additional inclusion criteria included caregiver-reported normal hearing, no history of head injury, and no co-occurring psychiatric, neurologic, or genetic conditions. All participants in the sample demonstrated nonverbal cognitive skills with standard scores greater than 75, as measured by the Kaufman Brief Intelligence Test (KBIT; Kaufman & Kaufman, 2004) Matrices subtest. Caregivers completed written informed consent, and all participants provided written informed assent. All procedures were approved by the

Committee on the Use of Humans as Experimental Subjects (COUHES) at MIT and conducted in accordance with the Declaration of Helsinki.

Diagnostic Group Classification

Autistic participants were first recruited based on caregiver-reported existing community-based autism diagnoses and subsequently verified through administration of the Autism Diagnostic Observation Schedule (ADOS-2) as part of study participation. Depending on participants' age, trained research staff administered the ADOS-2 Module 3 or 4 (Lord et al., 2012) to all study participants (both autistic and non-autistic). Autism diagnostic status was confirmed based on ADOS threshold scores for diagnostic classification. Moreover, non-autistic participants had no first-degree autistic relatives, per caregiver report.

Table 1. Participant Demographics

	Autistic (n = 30)	Non-autistic (n = 36)	Test Statistic (χ^2 or t)	Significance (p)
Sex (assigned at birth)			2.56	.110
M	26 (86.7%)	24 (66.7%)		
F	4 (13.3%)	12 (33.3%)		
Age in years [Mean (SD)]	10.4 (2.85)	10.2 (2.44)	0.31	.755
Household income			0.58	.900
Less than $30,000	1 (3.3%)	0 (0.0%)		
$30-60,000	0 (0.0%)	3 (8.3%)		
$60-100,000	8 (26.7%)	5 (13.9%)		
$100,000 or more	14 (46.7%)	24 (66.7%)		
Not reported	7 (23.3%)	4 (11.1%)		
Race			1.27	.531
White	27 (90.0%)	33 (91.7%)		
Black or African American	0 (0.0%)	1 (2.8%)		
More than one race	3 (10.0%)	2 (5.6%)		
Ethnicity			0.02	.890
Hispanic/Latino	3 (10.0%)	4 (11.1%)		
Non-Hispanic/Latino	26 (86.7%)	31 (86.1%)		
Not reported	1 (3.3%)	1 (2.8%)		
CELF-4 Core Language Score [Mean (SD)]	93.3 (20.4)	115 (10.8)	-5.18	<.001***
KBIT Matrices Standard Score [Mean (SD)]	106 (17.3)	113 (12.6)	-1.91	.062
ADOS-2 Comparison Score [Mean (SD)]	6.77 (1.96)	1.28 (0.659)	12.33	<.001***
Narrative Generation				
Total number of utterances in narrative	20.9 (8.56)	28.3 (10.6)	-3.14	.003**
Total word count in narrative	161 (97.8)	235 (94.1)	-3.10	.003**
Total duration of narrative (s)	75.0 (47.9)	96.9 (40.1)	-1.98	.052
Mean length of utterance (MLU) in morphemes	8.13 (2.13)	9.25 (1.21)	-2.57	.014*

Note: **p* < .05, ***p* < .01, ****p* < .001.

Standardized Language Measures

To assess participants' overall language skills, all participants were administered the Clinical Evaluation of Language Fundamentals–4th Edition (CELF-4) (Semel et al., 2003). The CELF-4 Core Language Score (population mean = 100, SD = 15) was derived as a composite language score from subtests depending on the child's age. For 5-8-year-olds, subtests included Concepts and Following Directions, Word Structure, Recalling Sentences, and Formulated Sentences. For 9-12-year-olds, subtests included Concepts and Following Directions, Recalling Sentences, Formulated Sentences, and Word Classes. For participants aged 13 or older, subtests included Recalling Sentences, Formulated Sentences, Word Classes, and Word Definitions. Notably, relative to non-autistic peers, the autistic group demonstrated significantly lower [$t(42.11) = -5.18$, $p < .001$] and more variable [$F(29,35) = 3.60$, $p < .001$] Core Language Scores.

Prosody Production Elicitation and Manual Coding

Language samples were elicited from the ADOS-2's "Telling a Story from a Book" task using standard ADOS procedures. All participants completed this task using the wordless picture book, *Tuesday* (Wiesner, 1991). Two trained research staff (the first and second authors) were masked to participants' diagnostic status and manually transcribed and segmented utterances in Praat (Boersma & Weenink, 2024) according to Systematic Analysis of Language Transcripts (SALT) (Miller & Iglesias, 2012) conventions. Sixteen participant files (19.5%) were double coded by both research staff for reliability. Participants' utterances were segmented based on communication units (C-units), defined by "an independent clause with its modifiers" (https://www.saltsoftware.com/media/wysiwyg/tranaids/CunitSummary.pdf). In accordance with previous literature, utterances that were not directly related to narrative generation (i.e., extratextual talk not related to the task, questions directed at the examiner, spontaneous reactions to the book), single word utterances, character speech, and animal sounds were not segmented or transcribed for analysis (Patel et al., 2020; Shriberg et al., 2001). Within segmented utterances, disfluencies and fillers (e.g., "fro-frogs", "um") were transcribed, while non-linguistic fillers such as lip smacking or heavy sighs were not transcribed). In accordance with existing literature on participants' pausing during storytelling, pauses were defined as greater than 250 ms (Redford, 2013). Pauses were coded in a separate Praat tier.

Compared to autistic participants, non-autistic participants produced a greater number of total narrative utterances ($t(63.94) = -3.14$, $p = .003$), more words in total ($t(60.94) = -3.10$, $p = .003$), and utterances with greater morphosyntactic complexity (as indicated by mean length of utterance (MLU) in morphemes ($t(44.08) = -2.57$, $p = .014$, see Table 1 for an overview). Autistic participants' narratives tended to be shorter in total duration compared to non-autistic participants' narratives, but this difference did not reach statistical significance ($t(56.76) = -1.98$, $p = .052$).

Extraction of Prosodic Features

Prosodic features extracted for the present study were chosen in accordance with existing literature suggesting differences between autistic and non-autistic individuals (as described in recent meta-analyses: Asghari et al., 2021; Fusaroli et al., 2017), as well as features previously derived from spontaneous language sampling during narrative contexts (e.g., Dahlgren et al., 2018; Patel et al., 2020). Therefore, the following prosodic features were extracted: speech rate, articulation rate, f_0 (fundamental frequency, which is perceived as pitch) mean, f_0 range, and f_0 standard deviation, as well as jitter (variability or perturbation of fundamental frequency). The f_0 variables are hereafter referred to as "pitch mean", "pitch range", and "pitch variation", respectively. Scripts used to extract speech timing and vocal acoustics can be found on Open Science Framework (OSF; https://osf.io/syctq/).

Speech and Articulation Rates

Speech rate was defined as syllables per second and was averaged across all segmented and transcribed utterances. Articulation rate was defined as syllables per second after the duration of pauses were subtracted out of the total duration of the utterance. Speech and articulation rates were calculated using a custom Python script. The Python library *textgrid* (https://github.com/kylebgorman/textgrid) was used to parse Praat TextGrid files. Syllable counting was automated based on each utterance's transcription using the Python library *syllapy* (https://github.com/mholtzscher/syllapy).

Vocal Acoustics

Vocal acoustic measures extracted from Praat included pitch mean, pitch range, pitch variation, and jitter. Due to the variability in sex and age within the sample and in order to ensure proper pitch tracking in Praat, a custom protocol was established for this study to ensure the most accurate reflection of pitch data. Initially, for each participant, pitch settings were manually adjusted in accordance with Patel et al. (2020) such that the pitch floor and pitch ceiling were 70 Hz and 400 Hz for males older than 12 years old and 130 Hz and 400 Hz for all other participants (i.e., males younger than 12 years old and all females). Next, a custom Praat script extracted individual participants' 10th and 90th percentile pitch values. For each participant, the pitch floor and pitch ceiling were then manually adjusted to reflect the 10th and 90th percentile values, respectively, in order to exclude outliers in Praat's pitch tracking. Finally, a research assistant trained on the protocol completed audio-visual inspection of the pitch tracking for each segmented utterance and annotated, in a new Praat tier, intervals to include in the vocal acoustic extraction. These included intervals had proper pitch tracking with alignment of the glottal pulses marked by Praat. Segments of utterances that were excluded from feature extraction included doubling/halving errors in pitch tracking, vocal fry, and pitch tracking visible during clearly unvoiced sounds. After manually completing pitch tracking protocol for each individual participant, pitch metrics (mean, range, and standard deviation) and jitter were extracted for each segmented interval using a Praat script (using each individual's pitch floor and pitch ceiling settings), and then averaged across intervals.

Analysis Plan

Welch's two sample *t*-tests were used to address the first aim of examining group differences in the following prosody metrics: speech rate, articulation rate, pitch mean, pitch range, pitch variation, and jitter. Follow-up analyses of covariance (ANCOVAs) were employed to assess how overall language skills may additionally account for variation in the prosody metrics. For the second aim, partial Spearman's correlations controlling for age, sex, and nonverbal cognitive skills were employed to assess relationships between prosody metrics and participants' overall language skills, as measured by the CELF-4 Core Language Score. Correlation analyses were first conducted across the whole group, followed by within-group analyses (within the autistic and non-autistic groups). All statistical analyses were conducted in R Statistical Software (R Core Team, 2022).

Community Detection

To address the third aim, community detection analyses were conducted based on methods described in Weed et al. (2023). With the goal of replication and extension of this previous work leveraging a larger and more heterogeneous sample in the present study, each participant was represented by the same four prosodic features used in Weed et al. (2023): speech rate, articulation rate, pitch variation, and jitter. Consistent with Weed et al. (2023), pitch was converted from Hz to a semitone scale using the *hmisq* R package (Quene, 2022), with the default 50 Hz as the reference frequency for this analysis. The *qgraph* R package (Epskamp et al., 2012) was used to create a correlation matrix of participants, which was then plotted as a network graph. A spinglass community detection algorithm (Reichardt & Bornholdt, 2006) was implemented in the *igraph* R package (Csárdi & Nepusz, 2006) in order to maximize positive edges (correlations) and minimize negative edges within communities, while maximizing negative edges and minimizing positive edges outside of communities. To account for the non-deterministic nature of the community detection algorithm which may result in different numbers of communities across multiple runs, we performed the analysis 1000 times with different random seeds and assessed consistency of the community detection results across runs (Djelantik et al., 2020; Weed et al., 2023). The median number of communities was identified after the 1000 runs, then the final network was chosen randomly from the resulting networks that reflected this median. The *pca* function from the *AMR* R package (Berends et al., 2022) was used to determine which prosodic features are weighted more heavily on the principal components and therefore hold the greatest predictive power in distinguishing communities.

Results

Group Differences in Prosodic Features

Welch's two sample *t*-tests revealed that autistic participants had slower speech rate ($t(58.16) = -2.06$, $p = .044$) and articulation rate ($t(60.01) = -2.09$, $p = .041$) compared to non-autistic participants (Figure 1). Regarding the pitch metrics, autistic participants had wider pitch range ($t(51.24) = 2.26$, $p = .028$) and greater pitch variation ($t(51.03) = 2.13$, $p = .038$) compared to their non-autistic peers (Figure 2). The differences in pitch mean and jitter between autistic and non-autistic participants were not statistically significant. For a full summary of group comparisons, see Table 2.

Table 2. Prosodic features by group

Prosodic feature	Autistic (n = 30)	Non-autistic (n = 36)	t	p
Speech rate (syllables/second)	3.07 (0.69)	3.41 (0.61)	-2.06	.044*
Articulation rate (syllables/second without pauses)	3.44 (0.64)	3.76 (0.6)	-2.09	.041*
Pitch mean (Hz)	223.3 (56.69)	224.82 (40.34)	-0.12	.903
Pitch range (Hz)	72.46 (29.45)	57.95 (21.04)	2.26	.028*
Pitch variation (Hz)	18.29 (7.74)	14.70 (5.5)	2.13	.038*
Jitter	0.02 (0.01)	0.02 (0.01)	1.00	.320

Note: $*p < .05$.

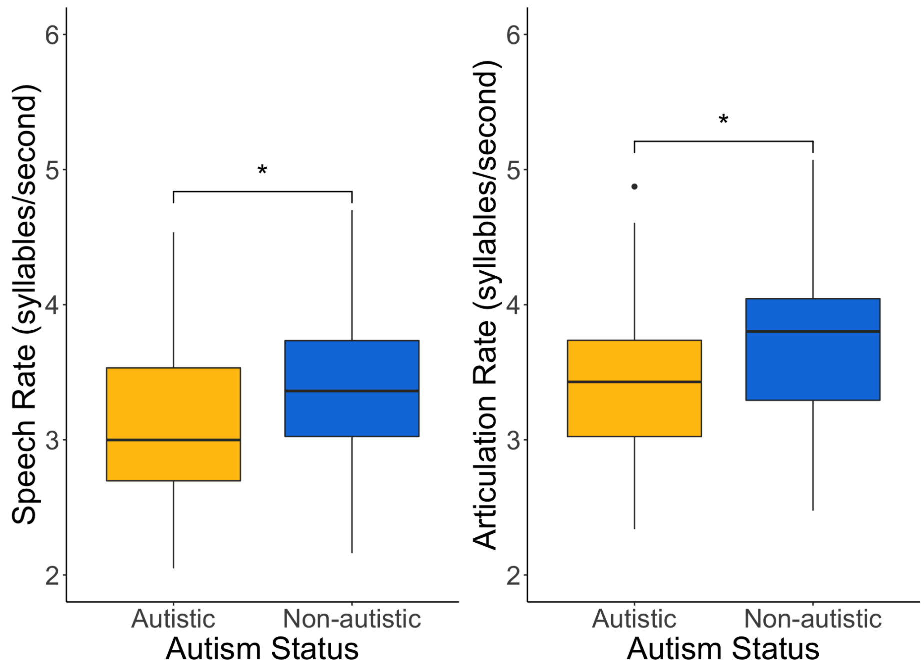

Figure 1. Box plots depicting significant group differences in speech timing aspects of prosody. Autistic participants' speech rate (left) and articulation rate (right) were slower than non-autistic participants'.

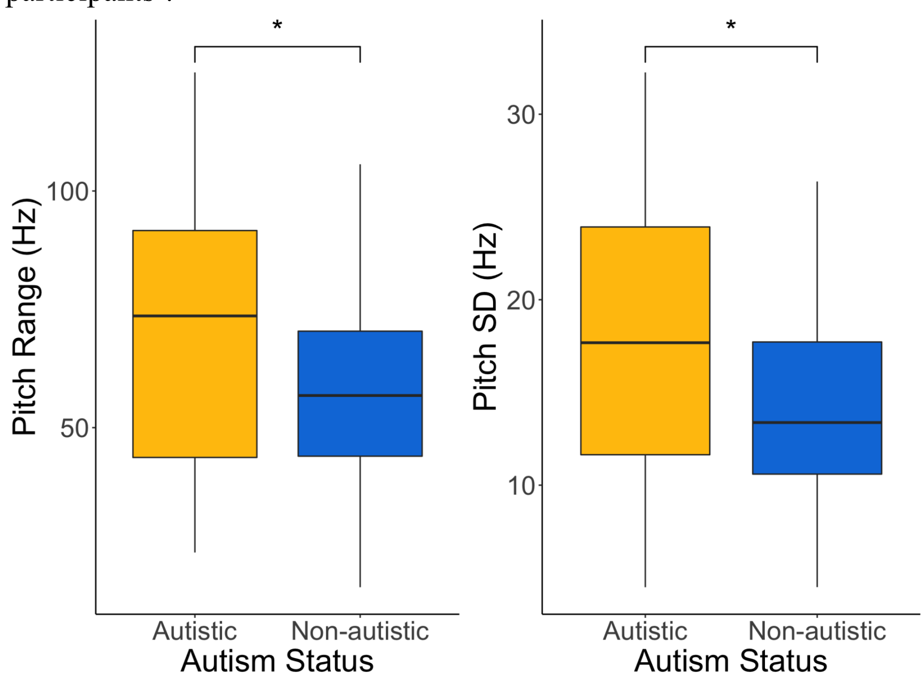

Figure 2. Box plots depicting significant group differences in pitch aspects of prosody. Autistic participants demonstrated wider pitch range (left) and greater pitch variation (right) compared to non-autistic participants.

Since group differences in overall language skills (as indicated by CELF-4 Core Language Scores) were indicated, subsequent between-group ANCOVAs were employed to determine whether observed effects remained significant while controlling for overall language skills. Accordingly, ANCOVAs revealed that significant effects of diagnostic status on pitch range [$F(1,63) = 5.67$, $p = .020$] and pitch variation [$F(1,63) = 6.32$, $p = .015$] remained when controlling for language. However, when accounting for overall language skills, significant main effect of diagnostic status were no longer significant for the timing aspects of prosody (speech rate: $F(1,63) = 2.03$, $p = .160$; articulation rate: $F(1,63) = 1.63$, $p = .206$). For full ANCOVA results, see Supplementary Table 1.

Associations Between Prosodic Features and Individual Differences in Overall Language Skills

Whole-Group Correlations

Partial Spearman's correlations controlling for age, sex, and nonverbal cognitive skills revealed that, in the whole sample, both speech rate and articulation rate were significantly positively associated with participants' overall language skills (Figure 3). No significant correlations were observed for any pitch metrics or jitter in relation to participants' overall language skills, p's > .05. For an overview of all correlations, see Table 3.

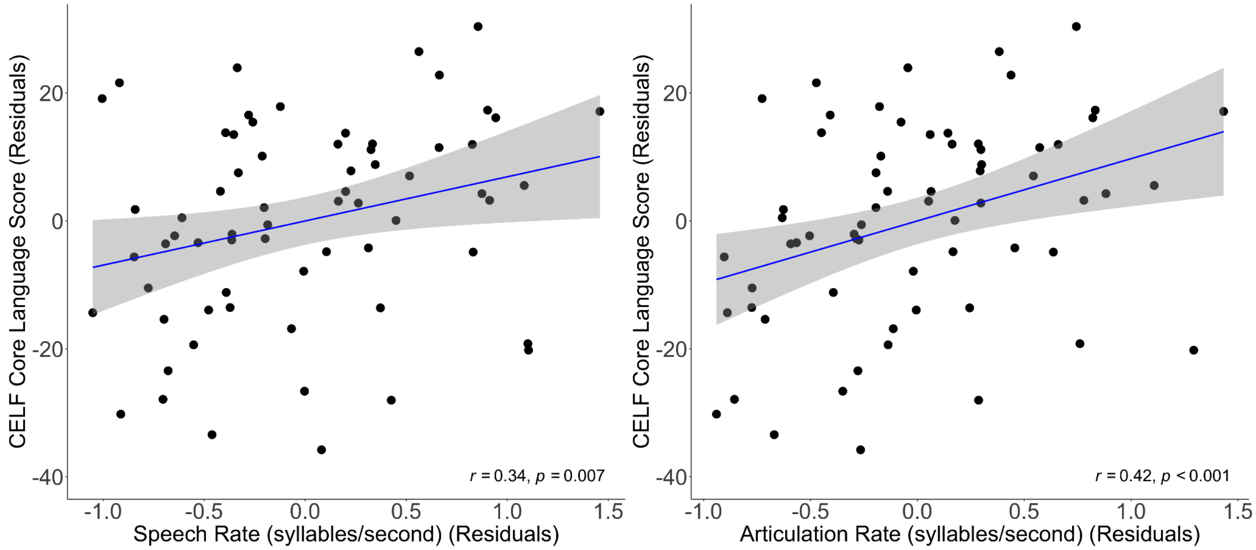

Figure 3. Residual scatterplot depicting a significant correlation between speech rate (left), articulation rate (right) and overall language skills in the whole sample when controlling for age, sex, and nonverbal cognitive skills.

Table 3. Partial correlations between prosodic features and overall language skills

	Whole group		Autistic		Non-autistic	
Prosodic feature	r_s	*p*	r_s	*p*	r_s	*p*
Speech rate (syllables/second)	0.335	0.007**	0.113	0.575	0.435	0.011*
Articulation rate (syllables/second without pauses)	0.416	<0.001***	0.253	0.204	0.508	0.003**
Pitch mean (Hz)	-0.008	0.952	-0.165	0.41	0.05	0.781
Pitch range (Hz)	-0.079	0.538	-0.141	0.484	0.217	0.225
Pitch variation (Hz)	-0.023	0.857	-0.013	0.949	0.186	0.299
Jitter	0.15	0.241	0.177	0.377	0.319	0.071

Note: All correlations are partial Spearman correlations controlling for age, sex, and nonverbal cognitive skills. Overall language skills are measured by CELF-4 Core Language Score standard scores. *$p < .05$, **$p < .01$, ***$p < .001$.

Within-Group Correlations

Further within-group correlation analyses revealed that the correlations between speech timing metrics and overall language skills were only significant in the non-autistic group (Table 3). No other within-group correlations were significant, *p*'s > .05.

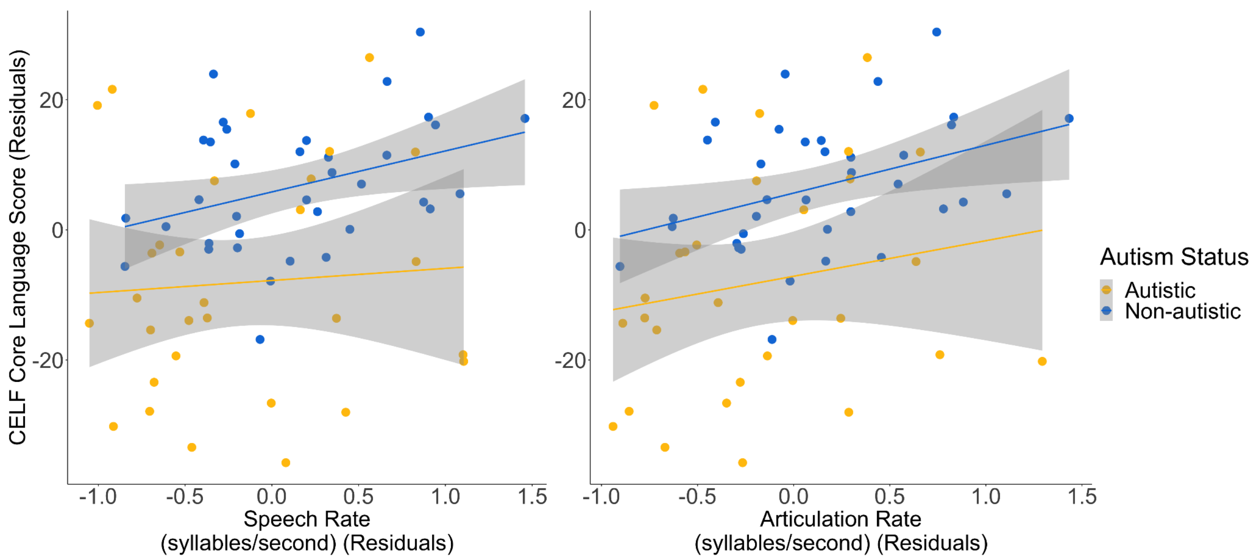

Figure 4. Residual scatterplots depicting that speech rate (left) and articulation rate (right) and overall language skills are significantly correlated in the non-autistic group, but not in the autistic group. Correlations control for age, sex, and nonverbal cognitive skills.

Community Detection

After 1000 runs, the community detection algorithm settled on three communities 986 times (98.6%) and four communities 14 times (1.4%). Overall, community membership was very consistent: Out of the 1000 runs, only four participants were occasionally placed in a fourth community; however, they were placed in this fourth community less than 2% of the time (see Table 4). Each of these participants was placed in their "preferred community" (i.e., the community they were placed in >98% of the time) in the final network (Figure 5). Therefore, a three-community solution was used for the final network.

Table 4. Participants that were inconsistently assigned to communities across 1000 runs and the percentages of runs in which they were assigned to each community.

ID	Diagnostic Status	Community 1	Community 2	Community 3	Community 4
1	Non-autistic	0	99.3007	0	0.599401
2	Autistic	0	99.2008	0.0999	0.599401
3	Non-autistic	99.1009	0.0999	0	0.699301
4	Autistic	0	98.002	0	1.898102

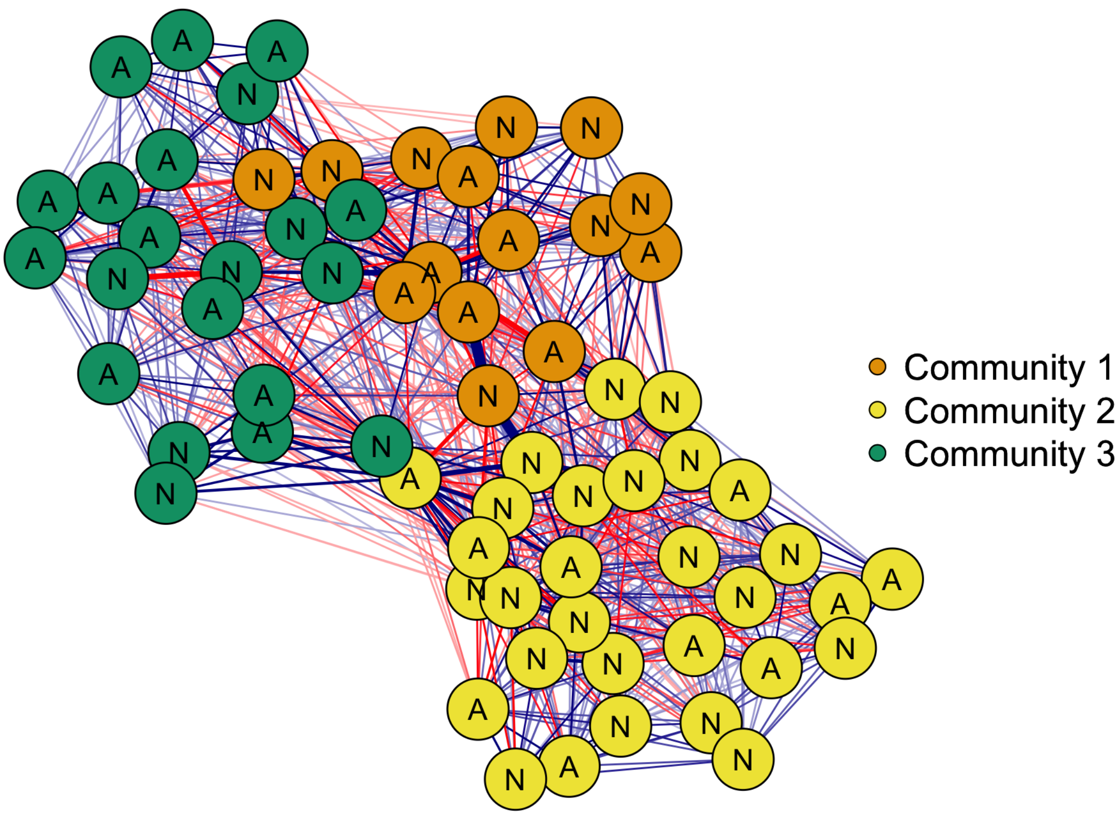

Figure 5. Network of communities as defined by the spinglass community detection algorithm, based on speech rate, articulation rate, pitch variation, and jitter. Each node represents a single participant (A = autistic, N = non-autistic), and each color denotes community membership (orange = Community 1, yellow = Community 2, green = Community 3). Blue lines between participants indicate positive correlations, while red lines between participants indicate negative correlations. Line thickness represents the correlation's absolute strength.

Network modularity

To assess the strength of the communities detected, a modularity score was calculated, that is, a measure of the degree of "clumpiness" in the network that ranges from -1 to 1 in which modularity values closer to 1 represent stronger partitioning of communities within the network (Newman & Girvan, 2004). The modularity of the network was determined to be 0.55, suggesting a network with strong community structure (Newman & Girvan, 2004).

Clustering of Diagnosis and Prosodic Features

In the final network, autistic and non-autistic participants were distributed within all three communities as follows: In Community 1, autistic participants made up roughly half (47%) of all participants within the community (Table 5). In Community 2, autistic participants made up one-third of participants, while Community 3 consisted of more autistic participants (62%) than non-autistic participants (38%).

Table 5. Distribution of autistic and non-autistic participants within each community

Community	Autistic		Non-autistic	
	n	%	*n*	%
1	7	46.7%	8	53.3%
2	10	33.3%	20	66.6%
3	13	61.9%	8	38.1%

To better illustrate the patterns of prosodic features that characterized each of the three communities, the key prosodic features included (speech rate, articulation rate, pitch variation, and jitter) were submitted to a principal component analysis. The first component (PC1), is largely represented by variation in the speech timing variables (speech and articulation rates), accounted for 50.4% of the variance. The second component (PC2), influenced by pitch variation and jitter (vocal acoustic measures), accounted for 31.1% of the variance. Visualization of individual participants and communities along the two dimensions of the PCA (Figure 6) revealed the following: all three communities occupy a similar area in the center of the PCA space, but participants in Community 1 (orange, largely equal distribution of autistic and non-autistic participants) and Community 3 (green, more autistic participants) cover more broad and distinct ranges influenced by pitch variation and jitter (PC2), whereas Community 2 (yellow, more non-autistic participants) was more heavily influenced by the speech timing variables (speech and articulation rates).

Finally, ANOVAs were employed to determine whether variation in participants' overall language skills predicted their community membership. Across all participants, overall language skills did not differ by community [$F(1,64) = 1.06$, $p = .307$]. ANOVAs to assess differences across community membership based on overall language skills were also conducted within the autistic and non-autistic groups separately. Overall language skills did not significantly differ by community membership for participants in either diagnostic group (autistic group: $F(1,28) = 0.023$, $p = .879$; non-autistic group: $F(1,34) = 0.581$, $p = .451$).

Similarly, none of the key demographic characteristics were found to differ by community membership among the whole sample or within each diagnostic group, including age (autistic group: $F(1,28) = 4.17$, $p = .051$; non-autistic group: $F(1,34) = 1.67$, $p = .204$; whole sample: $F(1,64) = 0.472$, $p = .495$), sex (autistic group: $\chi^2 = 0.748$, $p = .804$; non-autistic group: $\chi^2 = 2.03$, $p = .408$; whole sample: $\chi^2 = 2.63$, $p = .322$), nor nonverbal cognitive skills (autistic group: $F(1,28) = 0.243$, $p = .626$; non-autistic group: $F(1,34) = 0.937$, $p = .340$; whole group: $F(1,64) = 0.149$, $p = .701$).

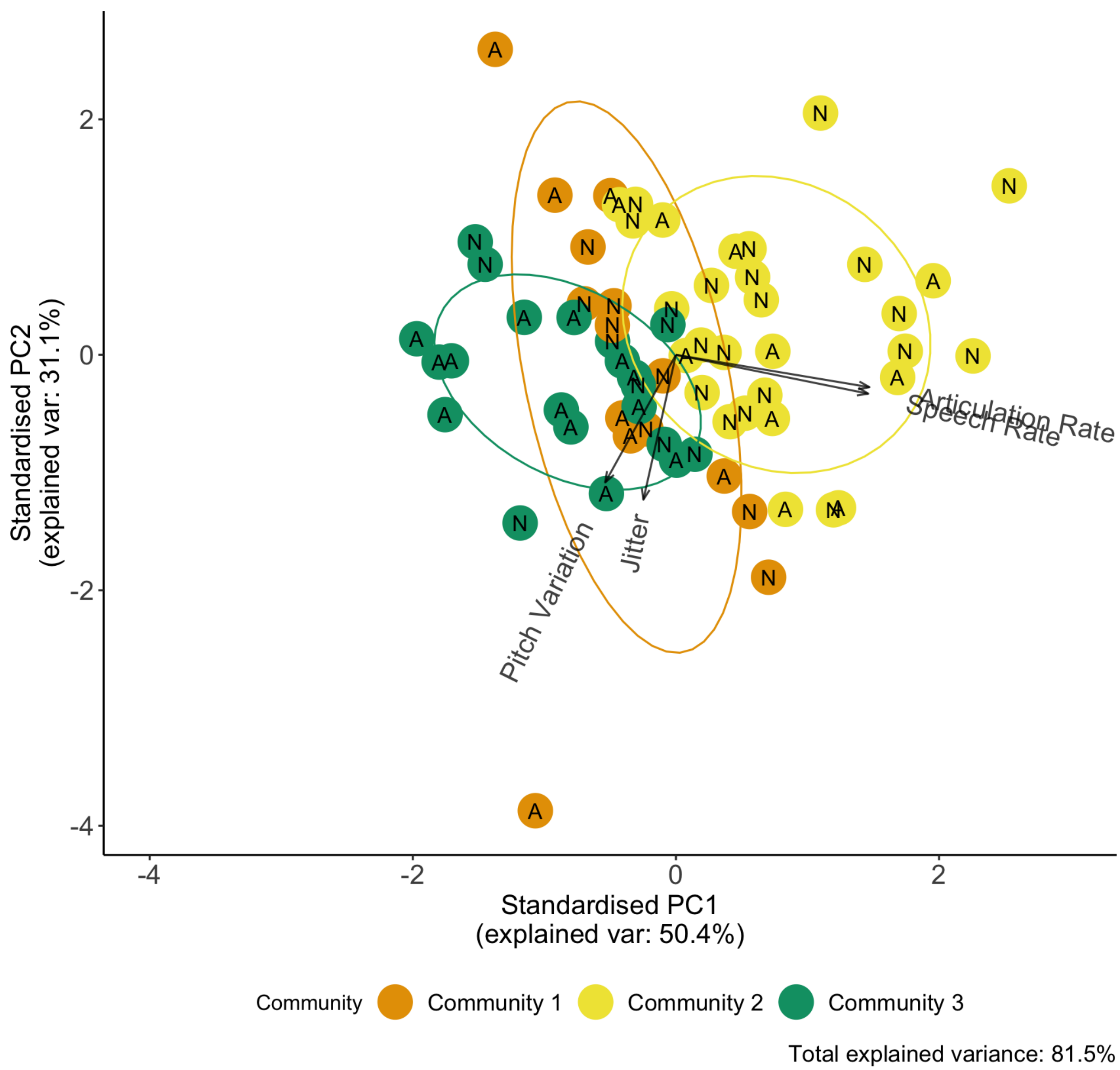

Figure 6. Participants (A = autistic, N = non-autistic) and communities (orange = Community 1, yellow = Community 2, green = Community 3) visualized along the first two dimensions of a principal components decomposition. Arrows indicate the influence of each prosodic feature on the components.

Discussion

The present study sought to clarify characteristics of speech prosody production associated with autism within a naturalistic, narrative context through multifaceted analytic approaches. In line with previous literature examining prosody production during narrative elicitation (Chan &

To, 2016; Diehl et al., 2009; Patel et al., 2020), initial between-group analyses indicated greater pitch range and variation as well as slower speech and articulation rates in autistic compared to non-autistic participants. However, group differences in speech and articulation rates were no longer significant when accounting for overall language skills. In line with these findings, individual differences (correlational) analyses revealed that speech and articulation rates were associated with individual differences in overall language skills. Taken together, these between-group and individual differences analyses illuminate the importance of accounting for individual differences in overall language skills when investigating prosody in autistic and non-autistic individuals. Lastly, the data-driven, network analysis approach using a community detection algorithm to identify clusters of participants characterized by similar acoustic-prosodic profiles resulted in three strong communities: one with more autistic participants, one with more non-autistic participants, and one with largely equal representation of autistic and non-autistic participants. Taken together, the present findings further reinforce the nuanced perspective that prosody in autism may be "different in different ways" (as originally coined by Weed et al., 2023, p. 14).

Our initial, traditional analysis approach demonstrated significant between-group differences, such that autistic participants demonstrated greater pitch range and pitch variation compared to their non-autistic peers. This finding supports a substantial body of evidence pointing toward differences in pitch variation between autistic and non-autistic children and adolescents. The significant difference in pitch *range* (i.e., the difference between the maximum and minimum fundamental frequencies) aligns with repeatedly demonstrated findings indicated in existing meta-analyses and reviews (Asghari et al., 2021; Fusaroli et al., 2017) and has even been shown to hold across contexts (e.g., during both conversation and structured communication tasks; Nadig &

Shaw, 2012). Differences in pitch range in particular has been suggested to be one of the few consistent findings in existing prosody in autism literature (Fusaroli et al., 2017). While existing meta-analyses suggest that pitch *variation*, when solely defined by the standard deviation of fundamental frequencies, does not show consistent differences among autistic and non-autistic individuals, autistic individuals show increased variability when pitch variation is considered in conjunction with pitch range (Asghari et al., 2021). Importantly, this finding is consistent with other studies specifically utilizing a more naturalistic narrative context for language sampling as in the present study (Asghari et al., 2021; Diehl et al., 2009). Furthermore, in the present study, group differences in pitch variation were stable after controlling for participants' overall language skills. Thus, the present findings add to a substantial body of literature pointing toward differences in pitch variability between autistic and non-autistic individuals.

Accounting for overall language skills in between-group analyses of covariance and individual differences analyses revealed that timing aspects of prosody – specifically, speech rate and articulation rate – seem to be linked with overall language skills. Although initial group comparisons indicated group differences in speech and articulation rate, these effects were no longer significant when controlling for overall language skills. While group differences in speech rate have been indicated in some but not all previous literature, the present findings suggest these mixed findings may be attributed to individual differences in language. Moreover, individual difference analyses revealed whole-group associations between timing aspects of prosody (i.e., speech rate and articulation rate) and overall language skills. Within-group correlations revealed significant correlations between these timing aspects of prosody and overall language skills among non-autistic participants but not among autistic participants, suggesting observed effects may be driven by non-autistic participants. The non-significant correlation between autistic participants' speech and articulation rates and overall language skills may be at least partly explained by the increased variability in overall language skills in our autistic sample compared the non-autistic sample, reflective of the heterogeneity of language skills in autism in general (Kjelgaard & Tager-Flusberg, 2001). Thus, the present findings suggest that individual differences in language may at least partly explain mixed findings in the prosody in autism literature, particularly with respect to timing aspects of prosody.

Finally, moving beyond the traditional between-group differences approach, the present study employed a data-driven, network analysis approach in an attempt to identify clusters of participants characterized by similar acoustic profiles based on a set of prosodic features. Using this approach, we sought to replicate and extend highly relevant recent findings, which clustered participants based on speech rate, articulation rate, pitch variation, and jitter (Weed et al., 2023). The present study both complements and extends findings from Weed and colleagues (2023) by employing the same algorithm based on the same set of prosodic features extracted from more naturalistic language sampling (rather than reciting memorized sentences) from a larger sample with a more heterogeneous range of language skills in the autistic sample. Comparable to the findings indicated by Weed and colleagues, the community-detection algorithm built a strong network of three communities: one with more autistic participants, one with more non-autistic participants, and one with largely equal representation of autistic and non-autistic participants. Therefore, similar to the final network observed by Weed and colleagues, none of the resulting three communities of participants were exclusively distinguished by autism diagnosis. Further, post hoc analyses confirmed that these communities did not seem to be classified by demographic characteristics or overall language skills.

In an effort to better understand how the selected prosodic features characterize each community, subsequent implementation of principal component analysis revealed, in accordance with that of Weed and colleagues, that one principal component was dominated by vocal acoustic measures (i.e., pitch variation and jitter), while the other was dominated by timing aspects of prosody (i.e., speech and articulation rates). Visual inspection via plotting these two principal components revealed apparent differences in the representation of these principal components among communities, which suggests that the community detection algorithm did form qualitatively different clusters of communities based on the prosodic features selected. Specifically, variation in Communities 1 and 3 appeared to be accounted for by differences in pitch variation and jitter, whereas variation in Community 2 appeared to be accounted for primarily by differences in the timing aspects of prosody. Regarding the distribution of autistic and non-autistic participants in each community: Community 1, which is roughly evenly comprised of autistic and non-autistic participants, heavily overlaps with Communities 2 and 3 at the center of the PCA plot. Communities 2 and 3, which show less overlap, are the communities represented by more non-autistic individuals and more autistic individuals, respectively. Taken together, our findings with a larger and more heterogeneous sample than in Weed et al. (2023) support the complexity and nuance of prosody in autism and provide further evidence that perhaps there is no singular definition of an "autistic voice" (Fusaroli et al., 2022; Weed et al., 2023), but rather, different clusters of participants characterized by different prosodic characteristics.

Limitations and Future Directions

The present study offers new perspectives on prosody in autism by illuminating the potential role of overall language skills in impacting between-group comparisons. In addition, by using data-driven methods to replicate and extend existing findings with a larger and more heterogeneous sample, the present study reinforces the importance of moving beyond traditional, between-group analysis approaches. At the same time, the present findings are to be interpreted in light of several considerations and limitations.

As evidenced by the community detection findings, the specific set of prosodic features examined in the present study is not sufficient to reliably distinguish participants by diagnostic status, but it is important to acknowledge that prosody is far more complex and multifaceted than the quantitative approaches presently investigated. It is possible that incorporating additional or alternate prosodic features could improve diagnostic classification in the network model, which emphasizes the importance of feature selection (Fusaroli et al., 2017). For instance, the present work reflects global measures of prosodic features, derived by averaging metrics across all of a given participant's narrative utterances. Future work involving the characterization of more local features and fluctuations in prosody during production, such as final sentence declination of pitch or the use of mid-sentence pauses (Benjamin & Schwanenflugel, 2010), may be one avenue to furthering our understanding of the most salient prosodic features. Although data-driven classification has yet to be achieved with the prosodic features presently investigated, it is important to note that both naïve and expert human raters are able to distinguish audio recordings by diagnostic status based on perceptual judgment (Cho et al., 2019; Nadig & Shaw, 2012; Redford et al., 2018; Weed et al., 2023). Therefore, investigating perceptual ratings of audio samples in conjunction with acoustic and network analyses, as modeled by Weed and colleagues (2023), is a powerful multifaceted approach to assess whether perceptual, acoustic, and data-driven methods are capable of alignment, or rather reveal distinct classification patterns. Overall, future work is needed to address discrepancies between perceptual ratings and acoustic and data-driven quantitative approaches to clarify our understanding of prosody in autism.

The present findings are also to be interpreted in light of a few design and sampling considerations. Although no group-level differences in age were observed, it is important to acknowledge that participants' ages ranged from school age to adolescence, spanning a wide range of developmental stages, including puberty. Regarding language heterogeneity, while the present study included all variation in overall language skills within the autistic group rather than group matching based on overall language skills (which would have excluded participants with lower language skills), it is important to acknowledge that all participants in the study sample were speaking communicators. Although the present focus on speaking communicators was pertinent to establish associations between prosody and overall oral language skills, we wish to acknowledge that this study is one among the majority of prosody production research in autism that has not included minimally speaking autistic individuals. Preliminary work focused on minimally speaking autistic children has suggested trends for greater pitch variation among minimally speaking autistic children than in speaking autistic and non-autistic children (Thorson et al., 2016). Thus, future work is warranted to clarify these findings and to build an overall characterization of prosody in autistic individuals with a wider range of language skills represented (DePape et al., 2012). In addition, considering known prosodic differences based on sex and gender, race and ethnicity, and geographic region (Armstrong et al., 2022), future work is needed that includes more representation across socioeconomic, demographic, and sociolinguistic aspects of variation in the participants sampled.

Lastly, we acknowledge that the present study did not involve direct consultation or inclusion of autistic perspectives in its design and implementation. Recognizing calls for increased participatory research in the field, future prosody in autism research could model an initial step to include autistic voices by aligning with autistic participants' priorities regarding research on prosody and autism (O'Brien et al., 2024). While much of the existing prosody in autism literature focuses on prosodic differences in autistic individuals at an individual level, new research directions should further consider how prosody may be impacted by contextual factors (e.g., task differences (Wynn et al., 2024). Although a strength of the present study involved measuring prosody production from a more naturalistic language sampling context, future research is warranted that extends beyond this by examining prosody during even more naturalistic and ecologically valid contexts, especially conversational interactions. Moreover, there is an unmet need to address the potential impact of additional contextual factors, such as familiarity with communication partners, the sensory environment, and individuals' use of various strategies to control their prosody (e.g., prosodic masking; O'Brien et al., 2024).

Conclusions

Overall, the present findings suggest that comprehensive investigation utilizing multifaceted analysis approaches reveals nuanced individual differences in prosody among both autistic and non-autistic participants. In doing so, we provide further evidence that while some group differences are indicated, these group differences may be attributed to individual differences in overall language skills. Furthermore, network analyses reveal that pitch variation, jitter, and speech and articulation rate are insufficient to reliably distinguish autistic from non-autistic participants. Moving away from traditional between-group comparisons that are often interpreted from a "deficit-based" perspective, employing an unsupervised data-driven approach offers a new perspective on prosody in autism that challenges existing literature that often proposes characteristic "prosodic differences" in autism. This new perspective lays a foundation for future work to gain more insights into the complexities of prosody in autism, and the prevalence of individual differences within this heterogeneous population. In consultation with the autistic community, these insights can be used to rethink and inform how prosody may be used as a tool for effective communication. With this better understanding of prosody in autism, all individuals can ultimately become more aware and inclusive of different communication strategies.

Acknowledgments

This work was funded by National Institute on Deafness and Other Communication Disorders R01DC011339 (to JDEG and HTF), the Hartwell Foundation (to JZ), and National Institute on Deafness and Other Communication Disorders T32DC013017 (to TL and KED). We thank all participating families for their contribution to our study. We are grateful to the MIT and BU research teams that contributed to data collection, especially Lisa Wisman Weil, as well as students who assisted with prosody analyses, especially Ione Madsen-Hardy, Therese Ancheta, and Antonia Nikolaidou. We also thank Stefanie Shattuck-Hufnagel for providing formative insights in consultation with coauthors in the early stages of establishing acoustic analysis approaches.

Data Availability Statement

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

Armstrong, M., Breen, M., Gooden, S., Levon, E., & Yu, K. M. (2022). Sociolectal and Dialectal Variation in Prosody. *Language and Speech*, *65*(4), 783–790. https://doi.org/10.1177/00238309221122105

Asghari, S. Z., Farashi, S., Bashirian, S., & Jenabi, E. (2021). Distinctive prosodic features of people with autism spectrum disorder: A systematic review and meta-analysis study. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-02487-6

Baltaxe, C. A. M., & Simmons, J. Q. (1985). Prosodic Development in Normal and Autistic Children. In E. Schopler & G. B. Mesibov (Eds.), *Communication Problems in Autism* (pp. 95–125). Springer US. https://doi.org/10.1007/978-1-4757-4806-2_7

Benjamin, R. G., & Schwanenflugel, P. J. (2010). Text Complexity and Oral Reading Prosody in Young Readers. *Reading Research Quarterly*, *45*(4), 388–404. https://doi.org/10.1598/RRQ.45.4.2

Boersma, P., & Weenink, D. (2024). Praat: doing phonetics by computer [Computer program]. Version 6.4.13, retrieved from http://www.praat.org/

Chan, K. K. L., & To, C. K. S. (2016). Do Individuals with High-Functioning Autism Who Speak a Tone Language Show Intonation Deficits? *Journal of Autism and Developmental Disorders*, *46*(5), 1784–1792. https://doi.org/10.1007/s10803-016-2709-5

Cho, S., Liberman, M., Ryant, N., Cola, M., Schultz, R. T., & Parish-Morris, J. (2019). Automatic Detection of Autism Spectrum Disorder in Children Using Acoustic and Text Features from Brief Natural Conversations. *Interspeech 2019*, 2513–2517. https://doi.org/10.21437/Interspeech.2019-1452

Cole, J. (2015). Prosody in context: A review. *Language, Cognition and Neuroscience*, *30*(1–2), 1–31. https://doi.org/10.1080/23273798.2014.963130

Csárdi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal*, *Complex Systems*, *1695*(5), 1-9. https://igraph.org.

Dahlgren, S., Sandberg, A. D., Strömbergsson, S., Wenhov, L., Råstam, M., & Nettelbladt, U. (2018). Prosodic traits in speech produced by children with autism spectrum disorders – Perceptual and acoustic measurements. *Autism & Developmental Language Impairments*, *3*, 2396941518764527. https://doi.org/10.1177/2396941518764527

de Marchena, A., & Miller, J. (2017). "Frank" presentations as a novel research construct and element of diagnostic decision-making in autism spectrum disorder. *Autism Research*, *10*(4), 653–662. https://doi.org/10.1002/aur.1706

DePape, A.-M. R., Chen, A., Hall, G. B. C., & Trainor, L. J. (2012). Use of Prosody and Information Structure in High Functioning Adults with Autism in Relation to Language Ability. *Frontiers in Psychology*, *3*, 72. https://doi.org/10.3389/fpsyg.2012.00072

Diehl, J. J., Watson, D., Bennetto, L., Mcdonough, J., & Gunlogson, C. (2009). An acoustic analysis of prosody in high-functioning autism. *Applied Psycholinguistics*, *30*(3), 385–404. https://doi.org/10.1017/S0142716409090201

Djelantik, A. A. A. M. J., Robinaugh, D. J., Kleber, R. J., Smid, G. E., & Boelen, P. A. (2020). Symptomatology following loss and trauma: Latent class and network analyses of prolonged grief disorder, posttraumatic stress disorder, and depression in a treatment-seeking trauma-exposed sample. *Depression and Anxiety*, *37*(1), 26–34. https://doi.org/10.1002/da.22880

Edelson, L., & Diehl, J. (2013). Prosody. In F. R. Volkmar (Ed.), *Encyclopedia of Autism Spectrum Disorders* (pp. 2413–2417). Springer. https://doi.org/10.1007/978-1-4419-1698-3_366

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). Qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software,* 48(4), 1–18. Retrieved from http://www.jstatsoft.org/v48/i04/

Fusaroli, R., Grossman, R., Bilenberg, N., Cantio, C., Jepsen, J. R. M., & Weed, E. (2022). Toward a cumulative science of vocal markers of autism: A cross-linguistic meta-analysis-based investigation of acoustic markers in American and Danish autistic children. *Autism Research*, *15*(4), 653–664. https://doi.org/10.1002/aur.2661

Fusaroli, R., Lambrechts, A., Bang, D., Bowler, D. M., & Gaigg, S. B. (2017). "Is voice a marker for Autism spectrum disorder? A systematic review and meta-analysis." *Autism Research*, *10*(3), 384–407. https://doi.org/10.1002/aur.1678

Grice, M., Wehrle, S., Krüger, M., Spaniol, M., Cangemi, F., & Vogeley, K. (2023). Linguistic prosody in autism spectrum disorder—An overview. *Language and Linguistics Compass*, *17*(5), e12498. https://doi.org/10.1111/lnc3.12498

Hubbard, K., & Trauner, D. A. (2007). Intonation and Emotion in Autistic Spectrum Disorders. *Journal of Psycholinguistic Research*, *36*(2), 159–173. https://doi.org/10.1007/s10936-006-9037-4

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman brief intelligence test (KBIT-2)*. Pearson.

Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An Investigation of Language Impairment in Autism: Implications for Genetic Subgroups. *Language and Cognitive Processes*, *16*(2–3), 287–308. https://doi.org/10.1080/01690960042000058

Lord, C., Michael Rutter, DiLavore, P. C., Risi, S., & Gotham, K. (2012). *Autism diagnostic observation schedule (ADOS-2)* (2nd ed.). Western Psychological Services.

Lord, C., Brugha, T. S., Charman, T., Cusack, J., Dumas, G., Frazier, T., Jones, E. J. H., Jones, R. M., Pickles, A., State, M. W., Taylor, J. L., & Veenstra-VanderWeele, J. (2020). Autism spectrum disorder. *Nature Reviews Disease Primers*, *6*(1), 5. https://doi.org/10.1038/s41572-019-0138-4

Ma, W., Xu, L., Zhang, H., & Zhang, S. (2024). Can Natural Speech Prosody Distinguish Autism Spectrum Disorders? A Meta-Analysis. *Behavioral Sciences*, *14*(2), Article 2. https://doi.org/10.3390/bs14020090

McCann, J., & Peppé, S. (2003). Prosody in autism spectrum disorders: A critical review. *International Journal of Language & Communication Disorders*, *38*(4), 325–350. https://doi.org/10.1080/1368282031000154204

McCann, J., Peppé, S., Gibbon, F. E., O'Hare, A., & Rutherford, M. (2007). Prosody and its relationship to language in school-aged children with high-functioning autism. *International Journal of Language & Communication Disorders*, *42*(6), 682–702. https://doi.org/10.1080/13682820601170102

Miller, J. & Iglesias, A. (2012). Systematic Analysis of Language Transcripts (SALT), Research Version 2012 [Computer Software]. Middleton, WI: SALT Software, LLC.

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. https://doi.org/10.1007/s10803-011-1264-3

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, *69*(2), 026113. https://doi.org/10.1103/PhysRevE.69.026113

O'Brien, A. M., Liu, T., Cardinaux, A., Gabrieli, J. D. E., & Zuk, J. (2024). *Perceiving and producing prosody: Perspectives and insights from autistic adults*. OSF Preprints.

https://doi.org/10.31219/osf.io/f7mqn

Patel, S. P., Nayar, K., Martin, G. E., Franich, K., Crawford, S., Diehl, J. J., & Losh, M. (2020). An Acoustic Characterization of Prosodic Differences in Autism Spectrum Disorder and First-Degree Relatives. *Journal of Autism and Developmental Disorders*, *50*(8), 3032–3045. https://doi.org/10.1007/s10803-020-04392-9

Peppé, S., Cleland, J., Gibbon, F., O'Hare, A., & Castilla, P. M. (2011). Expressive prosody in children with autism spectrum conditions. *Journal of Neurolinguistics*, *24*(1), 41–53. https://doi.org/10.1016/j.jneuroling.2010.07.005

Quene, H. (2022). hqmisc: Miscellaneous convenience functions and dataset. R package version 0.2-1, https://CRAN.R-project.org/package=hqmisc

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Redford, M. A. (2013). A comparative analysis of pausing in child and adult storytelling. *Applied Psycholinguistics*, *34*(3), 569–589. https://doi.org/10.1017/S0142716411000877

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

Reichardt, J., & Bornholdt, S. (2006). Statistical mechanics of community detection. *Physical Review E*, *74*(1), 016110. https://doi.org/10.1103/PhysRevE.74.016110

Sasson, N. J., Faso, D. J., Nugent, J., Lovell, S., Kennedy, D. P., & Grossman, R. B. (2017). Neurotypical Peers are Less Willing to Interact with Those with Autism based on Thin Slice Judgments. *Scientific Reports*, *7*, 40700. https://doi.org/10.1038/srep40700

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical evaluation of language fundamentals–*

Fourth edition (CELF-4) (4th ed.). Pearson.

Shriberg, L. D., Paul, R., McSweeny, J. L., Klin, A., Cohen, D. J., & 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. https://doi.org/10.1044/1092-4388(2001/087)

Thorson, J. C., Usher, N., Patel, R., & Tager-Flusberg, H. (2016). *Acoustic Analysis of Prosody in Spontaneous Productions of Minimally Verbal Children and Adolescents with Autism*. 40th Annual Boston University Conference on Language Development. https://www.bu.edu/bucld/files/2016/09/BUCLD40_Thorson1.pdf

Thurber, C., & Tager-Flusberg, H. (1993). Pauses in the narratives produced by autistic, mentally retarded, and normal children as an index of cognitive demand. *Journal of Autism and Developmental Disorders*, *23*(2), 309–322. https://doi.org/10.1007/BF01046222

Weed, E., Fusaroli, R., Simmons, E., & Eigsti, I.-M. (2023). Different in Different Ways: A Network-Analysis Approach to Voice and Prosody in Autism Spectrum Disorder. *Language Learning and Development*, *0*(0), 1–18. https://doi.org/10.1080/15475441.2023.2196528

Wiesner, D. (1991). *Tuesday*. Clarion Books.

Wiklund, M. (2016). Interactional challenges in conversations with autistic preadolescents: The role of prosody and non-verbal communication in other-initiated repairs. *Journal of Pragmatics*, *94*, 76-97. https://doi.org/10.1016/j.pragma.2016.01.008

Wiklund, M. (2023). Prosody. In M. Wiklund (Ed.), *Speech and Interaction of Preadolescents with Autism Spectrum Disorder: Focus on Prosody, Disfluencies and Comprehension Problems* (pp. 17–52). Springer Nature. https://doi.org/10.1007/978-981-19-8117-3_2

Wynn, C. J., Barrett, T. S., & Borrie, S. A. (2024). Conversational Speech Behaviors Are Context Dependent. *Journal of Speech, Language, and Hearing Research*, *67*(5), 1360–1369. https://doi.org/10.1044/2024_JSLHR-23-00622