

Prosodic traits in speech produced by children with autism spectrum disorders – Perceptual and acoustic measurements

SvenOlof Dahlgren

Department of Psychology, University of Gothenburg, Sweden

Annika D Sandberg 

Department of Psychology, University of Gothenburg, Sweden

Sofia Strömbergsson

Department of Clinical Science, Intervention and Technology, Karolinska Institutet, Sweden

Lena Wenhov

Department of Clinical Sciences Lund, Child and Adolescent Psychiatry, Lund University, Sweden

Maria Råstam

Department of Clinical Sciences Lund, Child and Adolescent Psychiatry, Lund University, Sweden

Ulrika Nettelbladt

Department of Clinical Sciences Lund, Logopedics, Phoniatrics and Audiology, Lund University, Sweden

Autism & Developmental Language Impairments

Volume 3: 1–10

© The Author(s) 2018

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/2396941518764527

journals.sagepub.com/home/dli

 SAGE

Abstract

Background: Autism spectrum disorder has been associated with atypical voice characteristics and prosody. In the scientific literature, four different aspects of atypical speech production in autism spectrum disorder have been highlighted; voice quality together with the prosodic aspects pitch, duration and intensity. Studies of prosody in autism spectrum disorder have almost exclusively used perceptual methods. Recently, some studies have used acoustic analyses. In these studies, it has been pointed out that the acoustic differences found are not necessarily perceived as atypical by listeners, which is why it is important to let listeners evaluate perceptual correlates to acoustic findings. The aims of this study were to use both perceptual and acoustic analyses to study prosodic production in children with autism spectrum disorder and to examine if voice and speech characteristics could be used as clinical markers for autism spectrum disorder.

Method: Eleven children within normal range of intelligence diagnosed with autism spectrum disorder and 11 children with typical development participated. Every child was recorded telling a story elicited with the expression, reception and recall of narrative instrument. Excerpts of one minute were extracted from the audio files creating the material underlying the perceptual ratings and in the acoustic analysis. An evaluation procedure, partly based on a standardized voice evaluation procedure developed for clinical practice in Sweden, was designed for the perceptual judgments and analysis. To capture critical prosodic variables, aspects of prosody based on characteristic features of Swedish prosody, prosodic features known to cause Swedish children with language impairment particular problems and current research of prosodic impairments in children with autism, were used as rating variables. The acoustic analysis was based on the four variables fundamental frequency (f_0) average, f_0 range, f_0 variation and speech rate, together with the language production-related variable number of words per utterance.

Corresponding author:

Annika D Sandberg, Department of Psychology, University of Gothenburg, Box 500, SE-40530 Göteborg, Sweden.

Email: annika.dahlgren.sandberg@psy.gu.se



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<http://www.creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Results: In the acoustic analysis, no differences were found with regards to f_0 -related variables or speech rate. However, the children in the autism spectrum disorder-group produced significantly more words per utterance than the typically developing children. The perceptual analysis showed no differences between the groups. Only three children with autism spectrum disorder were correctly identified as such. The narrative ability of these children, according to scores on the narrative assessment profile, was poorer than that of the other eight children. They were also more atypical in fluency and in speech rate. Given the small sample, the results should be interpreted with caution.

Conclusions and implications: The only difference in prosodic production discovered in the acoustic analysis, namely that children with autism spectrum disorder used more words per utterance than the children in the comparison group, was not detected in the perceptual assessment. This implies that it was not perceived as atypical by expert listeners. The results indicate difficulties in using voice and speech characteristics as markers of autism spectrum disorder in clinical settings. The correct identification of some of the children as having autism spectrum disorder or not also indicates that some children with autism spectrum disorder have a prosodic production sufficiently 'atypical' in combination with a limited ability to tell stories to be perceived.

Keywords

Acoustic measurement, autism spectrum disorder, narrative ability, prosody, perceptual measurement

Introduction

Autism spectrum disorder (ASD) has been associated with atypical voice characteristics and prosody ever since Kanner (1943) and Asperger (1944) published the first systematic studies of ASD. These descriptions were followed by others during the next decades (e.g. Goldfarb, Braunstein, & Lorge, 1956; Pronovost, Wakstein, & Wakstein, 1966; Simmons & Baltaxe, 1975).

Despite the consistent descriptions in the first studies from Kanner and Asperger onwards, the findings from more recent research on prosody and autism are contradictory and difficult to interpret. It is unclear whether this stems from methodological problems or from the heterogeneity among individuals with ASD (McCann & Peppé, 2003). Peppé, McCann, Gibbon, O'Hare and Rutherford (2007) found, for example, that expressive prosody varied greatly within a group with ASD. At the same time, all individuals exhibited at least one atypical prosodic trait when compared to typically developing children (TDC), while Grossman, Bemis, Skwerer, and Tager-Flusberg (2010) suggested that individuals with ASD did not seem to have specific difficulties in productive prosody. Different methods as well as differences in cognitive and linguistic level among the participants make it difficult to compare the opposing results. There are indications that the sometimes conflicting results might partly be explained by the heterogeneous symptoms that characterize autism.

In particular, there are four different aspects of atypical speech production in ASD reported in the scientific literature; voice quality together with the prosodic aspects pitch, duration and intensity (Fusaroli, Lambrechts, Bang, Bowler, & Gaigg, 2017; Titze, 1994). One of the most common descriptions is that

the speech of individuals with ASD is monotonous including atypical pitch and pitch variation (Baltaxe, Simmons, & Zee, 1984; Fay & Schuler, 1980; Goldfarb, Goldfarb, Braunstein, & Scholl, 1972; Kaland, Swerts, & Krahmer, 2013; Paccia & Curcio, 1982; Pronovost et al., 1966). There are also descriptions indicating atypical voice characteristics, where children with ASD are described as having a hoarse or harsh voice (Baltaxe, 1981; Pronovost et al., 1966; Sheinkopf, Mundy, Oller, & Steffens, 2000), with a hyper nasal resonance (Shriberg, Paul, McSweeney, Klin, Cohen, & Volkmar, 2001). Goldfarb et al. (1972), Simmons and Baltaxe (1975) and Baltaxe (1981) described the speech as being too slow or too quick. Others have reported the speech as being too loud or too quiet, and sometimes shifting between these two extremes (Goldfarb et al., 1972; Pronovost et al., 1966; Shriberg et al., 2001; Shriberg, Paul, Black, & van Santen, 2011).

Different standardized diagnostic instruments also include atypical prosody as part of the diagnosis e.g. the Autism Diagnostic Interview (ADI; Lord, Rutter, & Le Couteur, 1994; Rutter, Le Couteur, & Lord, 2003) and the Autism Diagnostic Observation Schedule (ADOS) (Lord, Rutter, DiLavore & Risi, 2000; Nadig & Shaw, 2012) with a focus on the person's use of prosody to express a certain content rather than on prosodic deviation in, e.g. pitch.

Studies of prosody in ASD have almost exclusively used perceptual methods, but in recent years, some studies have used acoustic analyses (e.g. Diehl, Watson, Bennetto, McDonough, & Gunlogson, 2009; Nadig & Shaw, 2012; Nakai, Takashima, Takiguchi, & Takada, 2014). Acoustic analysis suggests that individuals with ASD do not seem to have specific difficulties in productive prosody (Grossman et al., 2010;

Kaland et al., 2013). However, based on the children's longer expressions, Grossman et al. (2010) suggested that children with high functioning autism (HFA) have an atypical prosody production in natural settings. Kaland et al. (2013) found that adults with ASD had a narrower pitch range and were more monotonous in their speech, and Nakai et al. (2014) found more monotonous speech in their school-aged children with ASD, with the degree of monotonous speech being related to the degree of social interaction. It has also been pointed out that the acoustic differences found are not necessarily perceived as atypical by listeners, which is why it is important to let listeners evaluate perceptual correlates to acoustic findings (Diehl & Paul, 2013).

Referring to the research above the purpose of this study was, first, to examine prosodic characteristics in an objective acoustic analysis in school-aged children diagnosed with ASD within normal range of intelligence compared to TDC. As mentioned earlier, prosody is used as a characteristic of the disorder in clinical settings through the diagnostic instruments ADI and ADOS. Since prosody is considered important in such work, a second purpose was to explore whether a perceptual assessment made by experienced speech and language pathologists (SLPs) could capture any differences in productive prosody, rated as deviances from productive prosody in the typical population. Third, as another way of studying the usability of voice and speech characteristics as clinical markers for ASD, we examined to what extent group membership, i.e. an ASD-diagnosis or not, could be predicted from the ratings of productive prosody.

Method

Participants

Twenty-two children took part in the study; 11 children within normal range of intelligence diagnosed with ASD, 10 boys and 1 girl, and 11 children with typical development, six boys and five girls. The children with ASD met the DSM-IV-TR (American Psychiatric

Association, 2000) criteria for autistic disorder or Asperger syndrome. The ADOS—Generic (Lord et al., 2000) and the ADI-Revised (ADI-R; Lord et al., 1994) were used in the diagnostic procedure. Four of the children were also diagnosed with attention deficit/hyperactivity disorder (ADHD), one with attention deficit disorder (ADD) and one with tics. The children's cognitive level was assessed by means of Wechsler Intelligence Scale for Children (WISC-4, Wechsler, 2003). The participants with ASD were recruited from the Child and Adolescent Psychiatric (CAP) clinic, Lund University Hospital. The assessments were performed by a multidisciplinary team specialised in neurodevelopmental disorders, and consisted of a child psychiatric examination, including neurological status by a child psychiatrist, a clinical interview with the parents and teachers, an interview with the child, and a neuropsychological assessment of the child. The two groups of children were matched on chronological age on a group level. The TDC in the comparison group were recruited from schools in the same region. They were all judged by teachers and parents, independently, to be of normal intelligence and according to information from parents and teachers they had no history of contact with a SLP or psychologist. All children were tested for linguistic competence using scores from Test for Reception of Grammar (TROG2) to define linguistic age. See Table 1 for participants' characteristics.

Materials

As measures of language comprehension, the TROG-2 (Bishop, 2003), and the vocabulary subtest from WISC IV (Wechsler, 2003) were used. Scores from TROG2 were converted to linguistic age in years. Based on voice recorded speech elicited by means of the Expression, Reception and Recall of Narrative instrument (ERRNI, Bishop, 2004), language production and narrative ability were assessed using the narrative assessment profile (NAP, Bliss, McCabe, & Miranda, 1998). In the ERRNI, the child is asked to look closely at a

Table 1. Descriptive characteristics of the children in the two groups.

Variable	ASD (<i>n</i> = 11, M/F 10/1)			Comparison group (<i>n</i> = 11, M/F 6/5)			<i>p</i> values
	M	SD	Range	M	SD	Range	
Chronological age (years)	11.1	1.10	9.3–12.9	11.1	0.47	10.5–12.1	.889
Linguistic age (years)	10.4	3.86	4.8–16.4	14.7	2.37	9.7–16.4	.005
WISC vocabulary, scaled scores	4.8	2.79	1–10	9.6	4.13	4–16	.007
Narrative assessment profile (NAP) (max 18)	14.9	1.45	12–17	16.1	1.22	14–18	.052

ASD: autism spectrum disorder; M/F: quotient male/female; SD: standard deviation.

picture sequence consisting of 15 pictures, and, as a second step, produce a spoken story from the sequence, i.e. using visual cues. There are two different stories and for this study, the so-called Beach story was used. No other prompts than encouraging sounds to make the child continue are allowed. The test includes a 'warming up'-picture to introduce the child to the task. All narratives were then transcribed using the Codes for the Human Analysis of Transcriptions (CHAT) format (<https://talkbank.org/manuals/CHAT.pdf>). The recordings were made with a video camera, Japan Victor Company (JVC), Everio GZ-MG335, hard disc camcorder, in quiet rooms at the CAP clinic. In the case of the children in the comparison group, the recordings were made in their schools, in quiet rooms.

Data analyses

Perceptual analysis. The authors of the original master thesis in speech and language pathology (Dotevall & Lendt, 2014) developed an evaluation procedure specifically designed for the perceptual judgments and analysis. Parts of the procedure were based on a standardized voice evaluation procedure developed for clinical practice in Sweden (the Stockholm Voice Evaluation Approach (SVEA), Hammarberg, 2000). Swedish is often referred to as a pitch accent language regarded as having a relatively complex prosody compared to English (Cruttenden, 1997). For example, pitch accent may constitute a discriminative feature between two segmentally identical words, like 'stegen' /'ste:gen/ (Eng. *the steps*) and 'stegen' /'ste:gen/ (Eng. *The ladder*). To capture critical prosodic variables, we included aspects of prosody partly based on characteristic features of Swedish prosody (Bruce, 2012) as well as prosodic features known to cause Swedish children with language impairment particular problems (Samuelsson & Nettelbladt, 2004). The selection of variables was checked against current research of prosodic impairments in children with autism (Fusaroli et al., 2017; McCann & Peppé, 2003; Titze, 1994), resulting in the variables: pitch, intonation, speech rate, intensity, length of utterance, timbre, nasality and fluency. The ratings varied from 1 to 2, 3 or 4 depending on question with a higher number indicating a greater deviation from TDC. The SLPs filled in a protocol with the eight rating scales, marking the intended figure. Each scale was introduced with a written instruction. Only the end points were assigned a name. For timbre and fluency, the scale was: No deviation, small, medium and large, i.e. from 1 to 4. Nasality was rated as existing (2) or not (1). We also asked the SLPs to rate the 'overall impression', where we asked the raters to try to disregard of details in the other variables. Overall impression of prosody was

rated as adequate, atypical or very atypical. As a last question, the SLPs were asked to state if they thought that the child had an ASD-diagnosis or not.

The rating scale was pre-tested by two experienced SLPs, both specialists in voice. These SLPs did not take further part in the project. Excerpts of one minute each were extracted from the audio files, creating the material per child to be rated. The mean length of the produced stories was 139 seconds and the extracted sequences were identified as 30 seconds before and 30 seconds after the midpoint of the recordings, with some adjustments to avoid cutting the recording in the middle of an utterance. The ratings were made by three expert listeners, all three SLP specialists in voice. These experts were selected because of their experience of making perceptual analyses of voices, of children's speech, of being used to making consensus judgements and of their experience with the actual type of rating scales. They were informed verbally and with a written instruction of the purpose of the study but had no knowledge of the children. First, the recorded voice of a child, the same one-minute excerpt that was used for the acoustic analysis, was assessed individually by each rater, using their experience of children's speech in their ratings. The raters were allowed to repeat each recording as many times as they wanted but were not allowed to go back to an earlier recording. Thereafter, the ratings were discussed by the three raters until a consensus judgment was reached, in this phase without repetitions of the recordings. The speech excerpts were presented randomly, based on group belonging and gender.

Acoustic analysis. The audio signals were extracted from the video recordings by means of VideoLan Creator (VLC) Media Player (VideoLan Organization, n.d.). Sequences relevant for the current study were identified and converted from stereo to mono format in Audacity®, version 2.0.5. Transcription and annotation were then performed in Wavesurfer (Sjölander & Beskow, 2000), by the two aforementioned SLP students as part of their master's thesis. The children's utterances were identified with visual support from both waveforms and spectrograms, and orthographically transcribed. Hence, an utterance here is operationally defined as speech produced by the child and delimited by audibly and visually confirmed silence. Based on this manual annotation of utterances, relevant speech material was automatically extracted. The same one-minute sequences that were used for the perceptual analyses were then selected for further acoustic analyses.

The selection of variables included for acoustic analysis was motivated by previous reports indicating these features as being atypical in individuals with ASD. Whenever possible, acoustic variables correlating with the selected perceptual components were included.

Hence, f_o AVERAGE, f_o RANGE and f_o VARIATION were chosen as representing acoustic correlates of perceptual features previously identified as characteristic of ASD speech, i.e. pitch and pitch variation (Baltaxe, et al., 1984; Fay & Schuler, 1980; Goldfarb et al., 1972; Kaland et al., 2013; Paccia & Curcio, 1982; Pronovost et al., 1966). Similarly, SPEECH RATE was included as an acoustic variable since previous research has reported atypical speech rate in individuals with ASD (Baltaxe, 1981; Goldfarb et al., 1972; Simmons and Baltaxe, 1975). For other prosodic features, however, acoustic analysis was deemed unsuitable. This pertained, for example, to the analysis of loudness/intensity, where the lack of control of recording environment and microphone distance would make acoustic analysis unreliable. For similar reasons, no acoustic analysis of voice quality was performed.

Fundamental frequency (f_o) was extracted by means of Praat (Boersma, 2001), using default settings. Apart from the f_o AVERAGE of the utterance, f_o RANGE (expressed in semitones) was also registered. For each utterance, f_o VARIATION was estimated in accordance with Edlund and Heldner (2007), as the standard deviation of the f_o for the utterance, expressed in semitones.

The calculation of SPEECH RATE was based on the orthographic transcriptions and estimated by dividing the number of consonant to vowel transitions by the duration of the utterance, in seconds. Hence, this is an estimation of the number of syllables per second.

Following the recommendation by Hubbard and Trauner (2007), UTTERANCE LENGTH (represented as the number of words per utterance) was added to complement the acoustic variables, as a measure reflecting language production. Although this is not an acoustic measure in the strict sense but derived from the annotation of utterances in the recorded speech, we will report it as an acoustic variable for reasons of readability.

Table 2 presents an overview of the analysed acoustic and perceptual components. For some of the components, there are exact correspondences between the two methods. For some there are not, since not all prosodic components could be captured with both methods.

Statistical analyses

For the statistical analyses, International Business Machines Statistical Package for Social Sciences (IBM SPSS) Statistics (ver. 23) was used. To minimize the risk for Type I errors, comparisons were made with one-way analysis of variance, $p \leq .05$ instead of t tests (Hinkle, Wiersma, & Jurs, 2003). For non-parametric analyses, Mann-Whitney U test was used.

Because of small sample size, effect sizes, where applicable, were computed by means of Cohen's d .

Table 2. Acoustically and perceptually analysed components.

Acoustic components	Perceptual components
f_o average (Hz)	Pitch
f_o range (semitones)	Intonation
f_o variation (semitones)	Intonation
Speech rate (c2v ^a /second)	Speech rate
Words per utterance	Length of utterance
	Intensity
	Timbre
	Nasality
	Fluency
	Overall impression

^aConsonant-to-vowel transitions per second.

Results

Data from the acoustic analysis, aimed at answering research question one: examining prosodic characteristics in school-aged children diagnosed with ASD within normal range of intelligence compared to TDC, as presented in Table 3, showed that the children with ASD used more words per utterance than the children in the comparison group. This was the only statistically significant difference. The effect sizes in the acoustic analysis strengthen results.

Information pertaining to the second purpose of this study and with relevance for clinical settings, i.e. to explore whether a perceptual assessment could capture differences in productive prosody in children diagnosed with ASD within normal range of intelligence compared to TDC, is presented in Table 4. There were no differences between the ratings of the productive prosody in the two groups.

The third question regarded the ability to judge, with the perceptual ratings of productive prosody as a basis, if each child had an ASD-diagnosis or not. No one in the comparison group was judged to have an ASD-diagnosis. A correct judgment was made only for three of the children with an ASD-diagnosis, i.e. only three of the 11 children with an ASD-diagnosis were identified as such.

To examine if there was any difference among the children in the ASD-group that could explain correct identification or not, nonparametric statistical tests were performed. The analyses showed that the three children identified by raters as having ASD performed worse on NAP ($Md=13$ versus $Md=15$, $U=3.00$, $p=.033$) and were more atypical in fluency ($Md=4$ versus $Md=2$, $U=.500$, $p=.015$) and speech rate ($Md=3$ versus $Md=2$, $U=.000$, $p=.009$) than the eight children who were not correctly identified as having ASD.

Table 3. Acoustic analysis.

Variable	ASD		Comparison group		<i>F</i>	<i>p</i> value	Cohen's <i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
<i>f</i> ₀ average (Hz)	202.25	29.98	203.98	23.02	0.023	.881	−0.06
<i>f</i> ₀ range (semitones)	11.44	3.41	11.74	2.83	0.051	.824	−0.10
<i>f</i> ₀ variation (semitones)	23.73	2.58	23.98	2.15	0.061	.807	−0.11
Speech rate (syllables per second)	3.79	0.86	3.87	0.42	0.082	.778	−0.19
Words per utterance	6.31	1.57	4.81	1.45	5.421	.031	0.99

ASD: autism spectrum disorder; SD: standard deviation.

Table 4. Perceptual analysis.

Variable	ASD		Comparison group		Mann-Whitney <i>U</i>	<i>p</i> value
	<i>Md</i>	<i>Min–max</i>	<i>Md</i>	<i>Min–max</i>		
Pitch (1–4)	1.00	1–3	1.00	1–2	56.00	0.797
Intonation (1–4)	2.50	1–3	2.00	1–2	37.00	0.223
Speech rate (1–4)	2.00	1–4	2.00	1–3	44.00	0.300
Length of utterance (1–4)	2.00	1–4	2.00	1–3	60.00	1.000
Intensity (1–4)	2.00	1–3	1.00	1–2	35.50	0.173
Timbre (1–4)	1.00	1–3	2.00	1–3	58.50	0.898
Nasality (1–2)	1.00	1–2	1.00	1–2	44.00	0.300
Fluency (1–4)	2.00	1–4	2.00	1–4	58.00	0.898
Overall impression (1–3)	1.00	1–2	1.00	1–2	43.50	0.426

ASD: autism spectrum disorder.

Note: Numbers within parenthesis show the end points of the rating scale, higher number higher deviation.

Discussion

In this study, both perceptual and acoustic analyses were used to study prosodic production in children with ASD and to examine if voice and speech characteristics could be used as clinical markers for ASD. Although there were significant differences between the study groups on measures of language, the only statistically significant difference between the groups was found in the acoustic analysis, consisting of children with ASD using more words per utterance than the children in the comparison group. Interestingly, this variable may be considered the variable most closely reflecting language production. The other acoustic variables were not sensitive to any group difference. The children with ASD produced longer utterances, in terms of the number of words, compared to the children in the comparison group. While this is in line with similar findings reported by Hubbard and Trauner (2007), it is contradictory to other reports (e.g. Park, Yelland, Taffe, & Gray, 2012; Thurber & Tager-Flusberg, 1993), where children with ASD have been found to produce shorter utterances compared to

typically developing peers. There are only few studies using more natural settings like eliciting narratives in connection with acoustic analysis and results based on narrative tasks, such as in the present study. Results based on conversational speech in more interactive settings (as in Park et al., 2012), or in more controlled story completion tasks (as in Hubbard & Trauner, 2007) are not necessarily comparable. Moreover, the divergent results may also be explained by different definitions of what constitutes an utterance. In the present study, an utterance was defined as speech delimited by silence, whereas in most previous studies, the definition of utterance is more like a sentence, i.e. a unit defined by its syntactic, semantic and/or pragmatic features. However, if this difference can explain the divergent results, it also implies that the only statistically significant difference in the present study would disappear if changing our definition of utterance. An additional factor that may explain the seemingly contradictory results are differences in participant characteristics between the children in the present study and those participating in the other studies, e.g. intellectual level or gender.

As discussed by Grossman et al. (2010) although expressive speech of individuals with ASD may not be perceived as being inaccurate, it may be quantifiably different from the expressive speech of other individuals, for example, when analysed acoustically. These differences may contribute to the verbal expressions being perceived as atypical. The perceptual analyses in the present study, revealed only slight differences between the children with ASD and the children in the comparison group, corroborating the previously demonstrated finding (see e.g. McCann & Peppé, 2003) that perceived distinctive atypical prosodic and vocal characteristics of individuals with ASD are difficult to quantify operationally. In spite of the difference in the acoustic analysis in this study, the SLPs did not perceive the prosodic characteristics in the ASD group as atypical. This could mean that the differences were not sufficiently different to affect the children's prosodic production in a natural setting, a result similar to that of Diehl and Paul (2013). In addition, the excerpts to be assessed by the SLPs might have been too limited to get a sufficient picture of a possible prosodic deviance.

The third research question regarded the ability to predict the ASD diagnosis with the perceptual ratings of produced prosody as a basis. The expert listeners did not judge any child in the comparison group as having ASD. However, only three children with ASD were correctly judged as children with ASD while the other children with an ASD diagnosis ($n = 8$) were judged as TDC. This suggests that not all children with ASD have a general atypical produced prosody different enough to be discovered in a natural setting (McCann & Peppé, 2003). But, it also suggests that there are children with ASD whose prosodic and vocal expression is distinctive or atypical enough to be registered in a perceptual assessment. It is also worth noting that none of the children who were judged to have ASD had a comorbid diagnosis.

In the statistical analyses of the differences between the children who were correctly judged as having ASD and those who were wrongly judged to be typically developing, three statistically significant differences were found. The three children who were correctly identified as having ASD performed worse on the NAP and were more atypical in fluency and speech rate than the other eight children in the ASD-group. The latter two differences confirm the results of Grossmann et al. (2010) and suggest that the expert listeners noticed both some aspects of atypical prosodic production together with a limited ability to retell a story. The difference in NAP score highlights one important question, namely whether children with autism can use prosody to tell stories in a manner that increases the possibility for the listener to correctly understand the narrative or not. There seems to be something

in the narratives that correctly attracts the attention of the raters when making the judgement. This suggestion is in line with the results in a study on children with language impairment where a subgroup of children with pragmatic problems showed prosodic deviance at the discourse level, assessed in brief conversational and narrative tasks (Samuelsson & Nettelbladt, 2004). It is possible that such features, for example, reflected as atypical placement or expression of word stress, would have been captured with acoustic measures not included in the present work. Diehl et al. (2009) found a larger prosodic within-participant variation in narratives of children with ASD than in those of TDC. They also found a relation to clinical judgment of impaired communication. However, like in our study, there were no statistically significant group differences in average f_0 , indicating that something else than pitch variation is needed to explain the within-participant variation. To further study prosodic production in children with ASD within normal range of intelligence, it would be interesting e.g. to have the child express different feelings by different intonation only, i.e. in line with the symptomatology of ASD according to DSM-5 (American Psychiatric Association, 2013).

Limitations

Although all recordings took place in quiet rooms, the quality of the recordings was not optimized for acoustic analysis. However, these conditions reflect common clinical conditions and contribute to ecological validity. Moreover, as the same audio files were used both for the perceptual and for the acoustic analysis, conditions were kept constant to ensure a mirroring analysis. The methodological choice to restrict the analysis to one-minute stretches of speech, although longer samples may have given listeners more cues in their rating decisions, are also assumed to reflect realistic clinical conditions. The conclusions should be interpreted cautiously since there were only 11 children in each group and only three raters. However, the effect sizes for the acoustic variables corroborate the conclusions.

The gender imbalance between the ASD group and the TD group – with higher boys: girls' ratio in the ASD group – warrants careful interpretation of f_0 averages. The risk cannot be eliminated that a lower f_0 -level in boys compared to girls in these ages (Baken & Orlikoff, 2000) may have obscured a potential difference between the groups. However, Fusaroli et al. (2017) found in a recently published review that only two (Filipe, Frota, Castro, & Vicente, 2014; Sharda et al., 2010) out of 16 studies investigating pitch mean, reported a significant group difference with a higher pitch mean in the ASD groups. In seven of the 14 studies with nonsignificant results, the groups were

matched for gender while there was no gender match in the only two reporting higher pitch mean in the ASD group. Based on these findings (Fusaroli et al., 2017), we conclude that the lack of difference in pitch average between groups in our study cannot be explained solely by gender imbalance.

Rather than being a limitation, it is also worth noting that the study concerns children with ASD within normal range of intelligence, a growing percentage of the total group of the population with an ASD diagnosis (Idring et al., 2015).

Implications

The acoustic analysis showed that children with ASD differed in speech production, using more words per utterance, compared to typically developing peers matched for chronological age. However, this difference was not discovered in the perceptual assessment by three experienced SLPs specialised in voice, and as such was not perceived as atypical by listeners. The listeners' difficulty to detect differences between the children in the ASD and TDC groups is further found in their judgment of diagnosis where only three of the children with ASD were correctly judged as children with ASD. However, that three children with ASD were correctly identified also suggests that some children with ASD have an atypical prosodic production which can be perceived, but only at an individual level and, in the case of the children in this study, in combination with a limited ability to tell stories. The results indicate that there are difficulties even for specialized SLPs to detect and use voice and speech characteristics as clinical markers of ASD in clinical settings. This finding should be studied more thoroughly in relation to productive prosody, narrative ability and theory of mind (McCann, Peppé, Gibbon, O'Hare, & Rutherford, 2007).

Acknowledgements

The authors express their gratitude to the children who participated in the study and their parents. We also thank the students of speech and language pathology Nadia Chouaiby, MarieHelene Dotevall, Louise Lendt, Helena Moreau, Louise Mucchiano and Rebecca Rindhagen who collected the data.

Ethical statement

The study was approved by the Regional Ethical Review Board in Lund, Sweden and informed consent was collected in written form from parents and from all participating children.


Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Annika D Sandberg  <http://orcid.org/0000-0001-7960-0487>.

References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders*. 4th ed. Washington DC: Author.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. 5th ed. Washington DC: Author.
- Asperger, H. (1944). Die 'Autistischen Psychopathen' im Kindesalter. *Archiv für Psychiatrie und Nervenkrankheiten*, 117(1), 76–136.
- Audacity Team. (2014). Audacity® Version 2.0.5. Retrieved from <http://audacityteam.org>
- Baken, R. J., & Orlikoff, R. F. (2000). *Clinical measurement of speech and voice*. 2nd ed. San Diego, CA: Singular.
- Baltaxe, C. A. M., Simmons, J. Q., & Zee, E. (1984). Intonation patterns in normal, autistic and aphasic children. In A. Cohen & M. van de Broecke (Eds.), *Proceedings of the 10th International Congress of Phonetic Sciences* (pp. 713–718). Dordrecht: Foris.
- Baltaxe, C. (1981). Acoustic characteristics of prosody in autism. In P. Mittler (Ed.), *Frontier of knowledge in mental retardation* (pp. 223–233). Baltimore: University Park Press.
- Bishop, D. V. M. (2003). *Test for reception of grammar: TROG-2 version 2*. Pearson Assessment.
- Bishop, D. V. M. (2004). *Expression, reception, and recall of narrative instrument* [Measurement instrument]. London, UK: Harcourt Assessment.
- Bliss, L. S., McCabe, A., & Miranda, A. E. (1998). Narrative assessment profile: Discourse analysis for school-age children. *Journal of Communication Disorders*, 31, 347–363.
- Boersma, P. (2001). Praat: A system for doing phonetics by computer. *Glott International*, 5, 341–345.
- Bruce, G. (2012). *Allmän och svensk prosodi (General and Swedish Prosody)*. Lund: Studentlitteratur.
- Cruttenden, A. (1997). *Intonation*. New York: New York University Press.
- 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(01), 135–161.
- 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), 385–404.
- Dotevall, M. H. & Lendt, L. (2014). *Prosodi hos skolbarn med autismspektrumtillstånd*. (Prosody in school aged children with Autism Spectrum Disorder). Unpublished Master's thesis. Lund: Avd. för logopedi, foniatri och audiologi; Lunds universitet.
- Edlund, J., & Heldner, M. (2007). Underpinning /nailon/: Automatic estimation of pitch range and speaker relative pitch. In C. Müller (Ed.), *Speaker classification i:*

- Fundamentals, features, and methods* (pp. 229–242). Berlin: Springer.
- Fay, W. H., & Schuler, A. L. (1980). Emerging language in autistic children: Language intervention studies. London: Edward Arnold.
- Filipe, M. G., Frota, S., Castro, S. L., & Vicente, S. G. (2014). Atypical prosody in Asperger syndrome: Perceptual and acoustic measurements. *Journal of Autism and Developmental Disorders*, 44, 1972–1981.
- 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, 384–407.
- Goldfarb, W., Braunstein, P., & Lorge, I. (1956). Childhood schizophrenia: Symposium, 1955: 5. A study of speech patterns in a group of schizophrenic children. *American Journal of Orthopsychiatry*, 26(3), 544.
- Goldfarb, W., Goldfarb, N., Braunstein, P., & Scholl, H. (1972). Speech and language faults of schizophrenic children. *Journal of Autism and Childhood Schizophrenia*, 2(3), 219–233.
- Grossman, R. B., Bemis, R. H., Skwerer, D. P., & Tager-Flusberg, H. (2010). Lexical and affective prosody in children with high-functioning autism. *Journal of Speech, Language, and Hearing Research*, 53(3), 778–793.
- Hammarberg, B. (2000). Voice research and clinical needs. *Folia Phoniatrica et Logopaedica*, 52(1–3), 93–102.
- Hinkle, D.E., Wiersma, W., & Jurs, S. G. (2003). *Applied statistics for behavioural sciences*. Boston, MA: Houghton Mifflin.
- Hubbard, K., & Trauner, D. A. (2007). Intonation and emotion in autistic spectrum disorders. *Journal of Psycholinguistic Research*, 36(2), 159–173.
- Ildring, 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, 1766–1773.
- Kaland, C., Swerts, M., & Krahmer, E. (2013). Accounting for the listener: Comparing the production of contrastive intonation in typically-developing speakers and speakers with autism. *The Journal of the Acoustical Society of America*, 134(3), 2182–2196.
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2, 217–250.
- 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 pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24(5), 659–685.
- Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (2000). *Autism diagnostic observation schedule (ADOS)*. Los Angeles, CA: Western Psychological Services.
- McCann, J., & Peppé, S. (2003). Prosody in autism spectrum disorders: A critical review. *International Journal of Language & Communication Disorders*, 38(4), 325–350.
- McCann, J., Peppé, S., Gibbon, F., O'Hare, A., & Rutherford, M. (2007). Prosody and its relationship to language in school-aged children with high-functioning autism. *International Journal of Language & Communication Disorders*, 42, 682–702.
- 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.
- Nakai, Y., Takashima, R., Takiguchi, T., & Takada, S. (2014). Speech intonation in children with autism spectrum disorder. *Brain Development*, 36(6), 516–522. doi: 0.1016/j.braindev.2013.07.006. PubMed PMID: 23973369.
- Paccia, J. M., & Curcio, F. (1982). Language processing and forms of immediate echolalia in autistic children. *Journal of Speech, Language, and Hearing Research*, 25(1), 42–47.
- Park, C. J., Yelland, G. W., Taffe, J. R., & Gray, K. M. (2012). Morphological and syntactic skills in language samples of pre-school aged children with autism: Atypical development? *International Journal of Speech-Language Pathology*, 14(2), 95–108.
- Peppé, S., McCann, J., Gibbon, F., O'Hare, A., & Rutherford, M. (2007). Receptive and expressive prosodic ability in children with high-functioning autism. *Journal of Speech, Language, and Hearing Research*, 50, 1015–1028.
- 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, 19–26.
- Rutter, M., LeCouteur, A., & Lord, C. (2003). *Autism diagnostic interview-revised*. Los Angeles, CA: Western Psychological Services.
- Samuelsson, C., & Nettelbladt, U. (2004). Prosodic problems in Swedish children with language impairment: Towards a classification of subgroups. *International Journal of Language and Communication Disorders*, 39, 325–344.
- Sharda, M., Subhadra, T. P., Sahay, S., Nagaraja, C., Singh, L., Mishra, R., . . . Singh, N. C. (2010). Sounds of melody–pitch patterns of speech in autism. *Neuroscience Letters*, 478, 42–45.
- Sheinkopf, S. J., Mundy, P., Oller, D. K., & Steffens, M. (2000). Vocal atypicalities of preverbal autistic children. *Journal of Autism and Developmental Disorders*, 30(4), 345–354.
- Shriberg, L. D., Paul, R., McSweeney, 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 & Hearing Research*, 44(5), 1097–1115.
- Shriberg, L. D., Paul, R., Black, L., & van Santen, J. (2011). The hypothesis of apraxia of speech in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 41(4), 405–426.
- Simmons, J., & Baltaxe, C. (1975). Language patterns in adolescent autistics. *Journal of Autism and Childhood Schizophrenia*, 5, 333–351.
- Sjölander, K., & Beskow, J. (2000). WaveSurfer – An open source speech tool. In B. Yuan, T. Huang & X. Tang (Eds.), *Proceedings of ICSLP 2000* (pp. 464–467). Beijing, China: Chinese Military Friendship Publisher.
- 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.
- Titze, I. R. (1994). *Principles of voice production*. Englewood Cliffs, NJ: Prentice Hall.
- Wechsler, D. (2003). *Wechsler intelligence scale for children*. 4th Ed. Swedish version 2007.
- VideoLan Organization. (n.d.). VLC media player [Software]. Retrieved from <http://videolan.org/vlc/>