

Research Article

Adult Recognition of the Emotional State and Intonation in Speech of Children with Autism Spectrum Disorders: A Pilot Study

Elena Lyakso*, Olga Frolova

Department of Higher Nervous Activity and Psychophysiology, Saint Petersburg State University, Saint-Petersburg, Russia

*Corresponding author: Elena Lyakso, Department of Higher Nervous Activity and Psychophysiology, Saint Petersburg State University, Saint-Petersburg, Russia. Tel: +79219962492; Email: lyakso@gmail.com

Citation: Lyakso E, Frolova O (2018) Adult Recognition of the Emotional State and Intonation in Speech of Children with Autism Spectrum Disorders: A Pilot Study. Int J Autism & Relat Disabil: IJARD-116. DOI: 10.29011/IJARD-116. 000016

Received Date: 28 August, 2018; **Accepted Date:** 01 October, 2018; **Published Date:** 08 October, 2018

Abstract

The aim of the study is to reveal the ability of adults to recognize the emotional state and intonation contour of speech of children with autism spectrum disorders. 30 children with autism spectrum disorders aged 5-14 years, 60 typically developing coevals (control), and 440 adults (listeners) were participants in the study. Two analysis methods were used: perception study and spectrographic analysis of child speech. Overall, it was found that the adults are able to recognize the child's emotional state by the characteristics of his voice. The adults recognize the states of comfort and discomfort with equal probability when listening to the speech of typically developing children. In the speech of children with autism spectrum disorders adults better recognize the state of discomfort. Correspondence of the intonation contour of the typically developing child's repeated word to the adult's sample were determined better than intonation contour of the words repeated by children with autism spectrum disorders. These data can be used to teaching caregivers interacting with individuals with autism spectrum disorders.

Keywords: Autism Spectrum Disorders; Child Emotional Speech; Intonation Contour; Perception Study; Spectrographic Analysis of Speech

Introduction

Autism Spectrum Disorders (ASD) are associated with severe impairments in social functioning. One of the characteristics of individuals with ASD is a specific expression of emotions and prosody. Prosody plays a leading role in the process of verbal communication and in social contacts between people. Prosodic patterns can be expressed in a single word or across an entire utterance or conversation. Violation of prosody or specific prosodic pattern in individuals with ASD is noted by many researchers [for example:1-4], beginning with Kanner's pioneering works [5], and apparently, can be considered as a central characteristic of autism [6-9]. To a certain extent, the prosodic pattern reflects emotions that play an important role in human life, being a factor in the organization and initiation of behavior, determining the usefulness of information for the organism, and the realization of communication [10]. The ability to recognize the emotional state of adolescents with highly functional ASD for dynamic mimic expression and vocal expressions has been demonstrated [11]. To describe the prosodic features of speech, it is informative to

determine the shape of the Intonation Contour (IC) as a variation of pitch (the fundamental frequency) over time. In the estimation of IC, visual approximation and instrumental analysis of pitch and its intensity are traditionally used. Along with the instrumental estimation, the auditory - perceptual method applies [12,13]. Dynamic characteristics of the pitch - the difference between the maximum and minimum values of pitch - are significant predictor of the assessment of listeners in a perceptual experiment [9]. Listeners in perceptual analysis evaluate independently intensity of contrast in "prosodic minimum pairs" of records containing the same phonemic material that differ in a particular prosodic contrast [9]. Words' repetition in ASD children is based on the specific property of these children - echolalia [5] that could be applied in communication with ASD children [14].

Analysis of speech features of children with severe and moderate severity of autistic disorders is practically absent [13,15]. The literature data and our preliminary results conclude specific features of ASD child speech manifested in unclear articulation and unusual prosodic. These features caused difficulties in interaction of ASD children with other children and unfamiliar adults. Therefore, the aim of the work is to study adult's ability to recognize the emotional state and words' intonation contour in Typically Developing (TD) and ASD child's speech and to reveal

specific acoustic features of ASD child's speech. The hypothesis was verified that emotional state and words' intonation contour of ASD children are recognized by adults worse than TD children. The importance of the research lies in the concretization of data and the addition of auditory perception data by objective data of instrumental speech analysis. This approach can be used in speech therapy practice to correct the pronunciation and for creating automatic training systems for children with atypical development. It is important to find acoustic features of ASD child speech leading to difficulties in emotions recognition by listeners.

Method

Design of the Experimental Study Included

- Participants' selection on the base of diagnoses and questionnaires scores and ASD children distribution to groups according developmental features;
- Interview of informant's parents and completion of questionnaires about family and child development;
- Child speech recordings and video recordings of child's behavior;
- Perceptive analysis of child speech by groups of adult listeners, native speakers;
- Acoustic analysis of child speech;
- Statistical analysis of the data.

Participants in the study were children with ASD (F84 according to ICD-10), biologically aged 5-14 years ($n = 30$ children) and TD coevals ($n = 60$). For this study the ASD sample was divided into two groups according developmental features and medical conclusion: presence of development reversals at the age 1.5-3.0 years (first group - ASD-1, $n = 15$) and developmental risk diagnosed at the infant birth (second group - ASD-2, $n = 15$) - for these children the ASD was a symptom of neurological diseases associated with brain disturbed. Mean Child Autism Rating Scale [16] total scores were calculated for each group. In order to assess whether differences in autism severity varied across groups, a one-way ANOVA was conducted for two groups. The ASD child's groups don't differ significantly.

Two types of experimental methods of speech analysis were performed: perceptual (by listeners) and spectrographic.

Two Experimental Studies were Carried Out

Study 1 - Perception and spectrographic analysis of the child emotional speech.

For the first study recording conditions in the model experiment included playing with a standard set of toys; repetition of words from a toy-parrot in the game store situation; watching the

cartoon and the retelling the story. Situations of speech recording for TD children and children with ASD were maximally standardized. A model experiment was performed with heart rate control (HR) using a pulse oximeter "Choicemmed MD300C318".

Study 2 - Perception analysis of the child's repeated word to the sample (words spoken by the experimenter) with the purpose of determining the correspondence of the word repeated by the child to the sample by intonation.

Participants of this study were children aged 5-12 years. We used the model of word repeating by children [13], which allows to evaluate the formation of the articulation; the development of verbal memory - saving the words image; development of auditory memory - segmentation and emphasizing of phonemes from the speech stream; formation of the child's attention.

The recording of speech was carried out in laboratory. The recording time was 20-40 minutes. The recordings were made by the "Marantz PMD660" recorder with a "SENNHEIZER e835S" external microphone.

The child's emotional state was revealed based on recording situation and video fragment analysis by 5 speech experts. The test sequences were presented to 240 adults (native Russian speakers, age 21.3 ± 5.1 years) for perceptual analysis of emotional speech. 5 test sequences contained 30 words each were created. Adults were given the task of listening to tests to determine the three emotional states of the child "discomfort - neutral - comfort". The number of signals reflecting child's different emotional states in the tests was the same, while processing the personal data was taken as 100%. Participants of the second perception study were 200 adults (age 22.4 ± 6.2 year). The aim of the second perceptual study is the review of listeners' recognition of the correspondence of the word repeated by the child to the sample by the meaning and intonation contour. The test sequences included words from repetition words ($n=12$ tests, "adult sample - child response" for 35 samples each). The tests sequences included "adult sample - child response" for TD children aged 5-7 years (4 tests), for TD children aged 8-12 years (4 tests), and 5-12 years old children with ASD (4 tests). The tests 1 and tests 3 for each group of children contained the words requiring complex difficult articulation.

Spectrographic analysis of speech was carried out in the Cool Edit (Syntrillium Soft. Corp. USA) sound editor. Pitch values, max and min values of pitch, pitch range (F_0 max- F_0 min), and values of third formant frequency (F_3 , emotional formant) were analyzed and compared. The analysis of the words intonation contour was carried out in the Praat v. 6.0. 36 (<http://www.fon.hum.uva.nl/praat/>) (figure 1). Statistical analysis was made in "STATISTICA" program using Mann-Whitney test, Spearman correlation, one-way ANOVA.

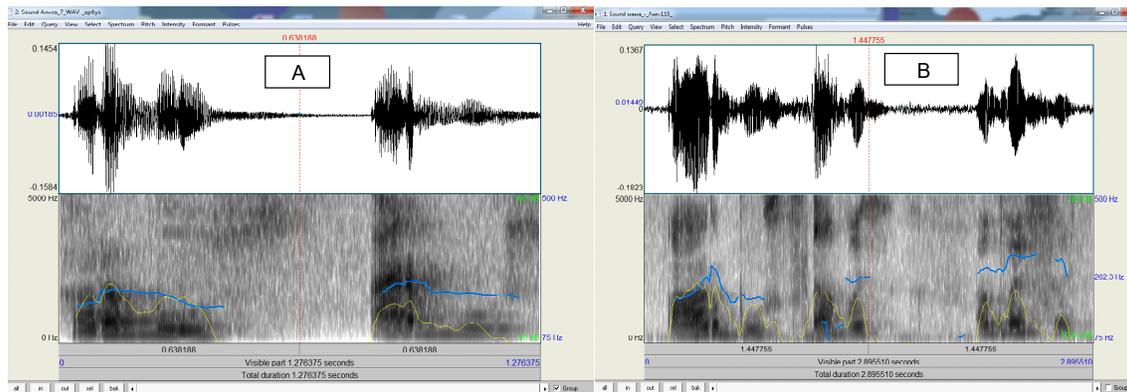


Figure 1: The Praat program window: analog form and IC of the words /watermelon - arbuz/ in the sample and in the repetition child word for girls: TD - Alisa (A) and Anna with ASD (B); Praat v. 6.0 36 (<http://www.fon.hum.uva.nl/praat/>).

Ethical Consideration

All procedures were approved by the Health and Human Research Ethics Committee (HHS, IRB 00003875, St. Petersburg State University) and written informed consent was obtained from parents of the child participant.

Results

Study 1 - Perception and Acoustic Features of Child Emotional Speech

The excitation emotional state of the TD children led to an increase in heart rate: 82 - 90 beats / min - in a calm state, 85 - 110 beats / min - in an emotional state. The values of heart rate in children with ASD were 80-90 beats / min in a calm state, 85-115 beats per minute in emotional state (125 beats per minute in one child), and did not differ significantly from the corresponding HR values in TD children.

Both discomfort and comfort conditions in the speech of TD children were recognized by adults with the perception rate of 0.75-1.0 better compared to the neutral condition. Positive correlation between TD age and recognition of discomfort state $r = 0.9747$ ($p < 0.05$ Spearman) was revealed. Adults recognized discomfort state in the speech of ASD children better according to Mann-Whitney test ($p < 0.01$) than comfort and neutral state. The listeners rated more speech signals of ASD-1 group children as reflection discomfort state (50% speech signals), than neutral (33%) and comfort (33%) states. According adult's conclusion more ($p < 0.01$) of speech signals of ASD-2 group children reflected the discomfort state (100%) than the speech signals of the ASD-1 group children (50%). The perception analysis data didn't reveal speech signals of ASD-2 group children reflecting comfort state.

Spectrographic analysis revealed that speech interpreted by listeners as discomfort, neutral and comfort is characterized by a set

of acoustic features. Discomfort TD children's speech samples are characterized by highest maximum pitch values ($p < 0.01$), average pitch values ($p < 0.05$) and pitch range values (F0max-F0min) ($p < 0.05$) vs. neutral speech samples. Discomfort state doesn't differ significantly from comfort state on the base of average pitch values of stress vowels from words. Correctly recognized by adults discomfort and comfort speech do not differ in pitch variation values. Changes of comfort and neutral state recognition with a child's age are bonded together: positive correlations between recognition of comfort and neutral test samples were revealed ($r = 0.9$). Discomfort state is mostly characterized by falling pitch contour type (intonation contour - IC), comfort state - by rising, and neutral - by flat pitch contour. Discomfort ASD children's speech samples are characterized by vowels' highest average pitch values, pitch range, and third formant frequency of words ($p < 0.001$) than comfort and neutral speech samples. Pitch variation values (F0max-F0min) in ASD-1 child's discomfort, neutral and comfort speech are significantly higher ($p < 0.001$) than in ASD-2 child's speech. The F3 values in discomfort speech of ASD-1 children are significantly higher than in corresponding voice features in ASD-2 children ($p < 0.01$) and TD peers ($p < 0.01$). IC type does not change depending on the emotional state of ASD children.

To clarify the data on the variance between the IC of ASD child's speech and the words IC normative for the Russian language, reflecting the neutral state, the state of comfort and discomfort, the experiment was conducted with the child repeating words for the experimenter.

Study 2 - Perception Data of the Words Intonation Contour

Determining the correspondence of the intonation contour of the child's repeated word to the sample caused difficulty for the listeners. The listeners define the correspondence of the words' IC to the IC of the sample better for the TD child's aged 8-12 years

- 77.8 % and 5-7 years old TD child's - 71.6 %, worse for the 5-12 years children with ASD - 64.4 %. For TD children, the number of words defined by adults as coinciding with the sample according to IC varies in different tests and does not depend on the complexity of the articulation of words included in the test sequence (figure 2). In three tests containing the "adult sample - ASD child response", the number of words coinciding in the IC with the sample indicated by the adults did not differ (64.8%, 66.5%, 67.5%, respectively, in test 1, 2, 3), in the fourth test - lower (not significantly - 58.5% of the words).

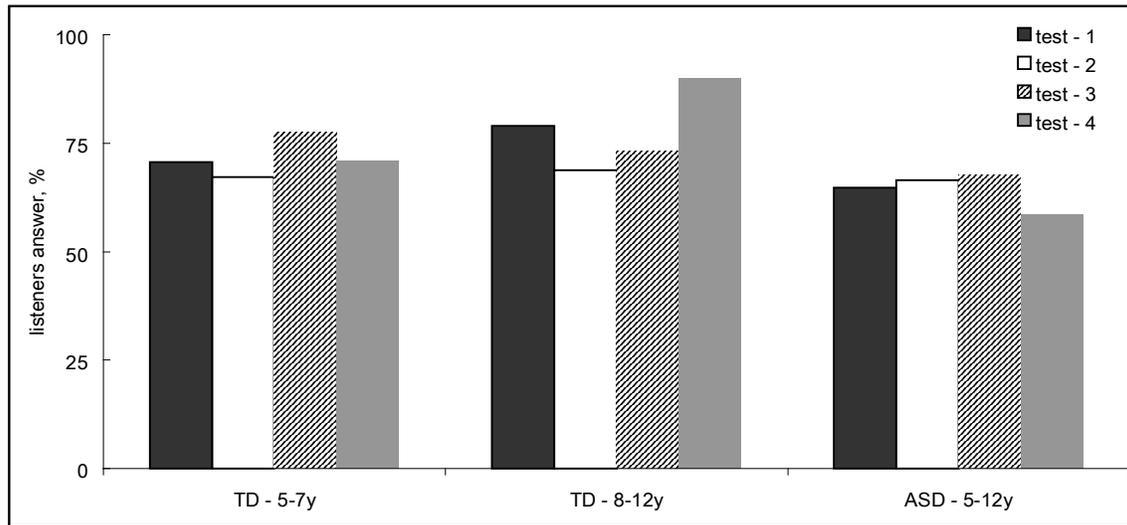


Figure 2: Correspondence of the word repeated by the child to the sample by the intonation contour. Vertical axis - listeners answer, %. Test - 1 data - black, test - 2 - white, test - 3 - oblique shading, test - 4 - gray.

Discussion

The specificity of adult's recognition of emotional states in TD and ASD children is revealed. It was shown that adults recognize with equal probability the states of comfort and discomfort by speech signals of TD children and attribute more number of speech signals of ASD children to discomfort. Listeners recognize the emotional state of children from ASD-2 group worse comparing with children from ASD-1 and TD groups. They have difficulties with comfort state determination in ASD-2 children, while discomfort state of ASD-2 children is recognized better vs. TD children. Skills of TD children in dampening emotional outbursts and in displaying unfelt "emotional fronts" increases between 5 and 12 years of age [17], emotions of children with ASD are more wild, natural and less socially conditioned. The specific of emotional state manifestation in informants with true autism - ASD-1 and with ASD as a symptom of neurological diseases - ASD-2 is shown first time in our study. The studies of ASD participants' emotional state recognition by humans are deficiently, but these data are important for effective communication [15,18]. In the study of adults it was shown that listeners were more accurate at identifying the emotion context from speech produced by ASD participants compared to TD participants but rated ASD emotional speech as sounding less natural [18]. The presented data on the listeners' recognition of the emotional state of children with

ASD does not confirm the data obtained for ASD adults. The data of the instrumental analysis showed that pitch range in ASD-1 child's emotional speech is higher than in ASD-2 child's speech. Apparently, the high pitch range noted by the researchers [4,11,19] and revealed in this paper, which caused the variety of prosodic characteristics and the degree of affect, can be one of the features for speech signals attribution to different emotional states. The third formant values in discomfort speech of ASD-1 children are higher than in ASD-2 and TD children. According listeners' answers ASD child's IC correspond to the adult's sample in fewer words than in TD children. Our study about the possibility of repeating the intonation contour by children with ASD is confirmed by the study of imitation of prosodic models by children with ASD and TD children using a more complex task in the PEPS-C program [8].

Conclusions

Adults are able to recognize the child's emotional state by the characteristics of his voice. The adults recognize the states of comfort and discomfort with equal probability when listening to the speech of TD children. In the speech of ASD children adults better recognize the state of discomfort. On the base of perceptual experiment the recognition of ASD and TD child correspondence of the intonation contour of the words of the child's repeated word to the sample with less recognition of ASD child was revealed.

Acknowledgments

This study is financially supported by the Russian Foundation for Basic Research (projects 16-06-00024a, 17-06-00503a-OGN, 18-013-01133a).

References

1. Shriberg LD, Paul R, McSweeney JL, Klin AM, Cohen DJ, et al. (2001) Speech and prosody characteristics of adolescents and adults with high functioning autism and Asperger syndrome. *Journal of Speech, Language, and Hearing Research* 44: 1097-1115.
2. Shriberg LD, Paul R, Black LM, van Santen JP (2011) The hypothesis of apraxia of speech in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders* 41: 405-426.
3. Paul R, Augustyn A, Klin A, Volkmar FR (2005) Perception and production of prosody by speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders* 35: 205-220.
4. Bonnef YS, Levanov Y, Dean-Pardo O, Lossos L, Adini Y (2011) Abnormal speech spectrum and increased pitch variability in young autistic children. *Frontiers in Human Neuroscience* 4: 237.
5. Kanner L (1943) Autistic disturbances of affective contact. *Nervous Child* 2: 217-250.
6. Tager-Flusberg H (2000) Understanding the language and communicative impairments in autism. *International Review of Research in Mental Retardation* 23: 185-205.
7. Baron-Cohen S (2000) Theory of mind and autism: a review. *International Review of Research in Mental Retardation* 23: 169-184.
8. Diehl JJ, Paul R (2012) Acoustic differences in the imitation of prosodic patterns in children with autism spectrum disorders. *Research in Autism Spectrum Disorders* 6: 123-134.
9. van Santen JP, Prud'hommeaux ET, Black LM, Mitchel M (2010) Computational prosodic markers for autism. *Autism* 14: 215-236.
10. Liebenthal E, Silbersweig DA, Stern E (2016) The language, tone and prosody of emotions: neural substrates and dynamics of spoken-word emotion perception. *Frontiers in Neuroscience* 10: 506.
11. Grossman RB, Edelson LR, Tager-Flusberg H (2013) Emotional facial and vocal expressions during story retelling by children and adolescents with high-functioning autism. *Journal of Speech, Language, and Hearing Research* 56: 1035-1044.
12. van Santen JPH, Prud'hommeaux RT, Black LM (2009) Automated Assessment of Prosody Production. *Speech Communication* 51: 1082-1097.
13. Lyakso E, Frolova O, Grigorev A (2017) Perception and acoustic features of speech of children with autism spectrum disorders. *LNAI* 10458: 602-612.
14. Saad AG, Goldfeld M (2009) Echolalia in the language development of autistic individuals: a bibliographical review. *Pró-Fono Revista de Atualização Científica* 21: 255-260.
15. Lyakso E, Frolova O, Grigorev A (2016) A comparison of acoustic features of speech of typically developing children and children with autism spectrum disorders. *LNAI* 9811: 43-50.
16. Schopler E, Reichler RJ, DeVellis RF, Daly K (1980) Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *Journal of Autism and Developmental Disorders* 10: 91-103.
17. Saarni C, Mumme DL, Campos JJ (1998) Emotional development: action, communication, and understanding. In W. Damon and N. Eisenberg (Ed.), *Handbook of child psychology: Social, emotional, and personality development*. Hoboken, NJ, US: John Wiley and Sons Inc: 237-309.
18. Hubbard DJ, Faso DJ, Assmann PF, Sasson NJ (2017) Production and perception of emotional prosody by adults with autism spectrum disorder. *Autism Research* 10: 1991-2001.
19. Sharda M, Subhadra TP, Sahay S, Nagaraja Ch, Singh L, et al. (2010) Sounds of melody-pitch patterns of speech in autism. *Neuroscience Letters* 478: 42-45.