

VOICE AND SPEECH FEATURES AS A DIAGNOSTIC SYMPTOM

Elena Lyakso, Olga Frolova, & Aleksandr Nikolaev

Department of Higher Nervous Activity and Psychophysiology, Saint Petersburg University (Russia)

Abstract

The study of the peculiarities of speech of children with atypical development is necessary for the development of educational programs, children's socialization and adaptation in society. The aim of this study is to determine the acoustic features of voice and speech of children with autism spectrum disorders (ASD) as a possible additional diagnostic criterion. The multiplicity of symptomatology, different age of its manifestation, and the presence of a leading symptom complex individually for each child make it difficult to diagnose ASD. To determine the specificity of speech features of ASD, we analyzed the speech of children with developmental disabilities in which speech disorders accompany the disease - Down syndrome (DS), intellectual disabilities (ID), mixed specific developmental disorders (MDD). The features that reflect the main physiological processes occurring in the speech tract during voice and speech production are selected for analysis. The speech of 300 children aged 4-16 years was analyzed. Speech files are selected from the speech database "AD_Child.Ru" (Lyakso et al., 2019). Acoustic features of voice and speech, which are specific for different developmental disorders, were determined. The speech of ASD children is characterized by: high pitch values (high voice); pitch variability; high values for the third formant (emotional) and its intensity causing "atypical" spectrogram of the speech signal; high values of vowel articulation index (VAI). The speech of children with DS is characterized by the maximal duration of vowels in words; low pitch values (low voice); a wide range of values of the VAI depending on the difficulty of speech material; low values of the third formant; unformed most of consonant phonemes. The characteristics of speech of children with ID are: high values of vowel's duration in words, the pitch, and the third formant, low values of the VAI; of MDD - low pitch values and high values of the VAI. Based on the identified peculiarities specific to each disease, the set of acoustic features specific to ASD can be considered as a biomarker of autism and used as an additional diagnostic criterion. This will allow a timely diagnose, appoint treatment and develop individual programs for children. Speech characteristics of children with ID, DS, and MDD can be considered to a greater extent in the training and socialization of children and used in the development of training programs taking into account individual peculiarities of children.

Keywords: *Child speech, acoustic features, atypical development, biomarker of autism.*

1. Introduction

Speech disorders accompany many diseases and atypical development of children. The most striking manifestations of speech development disorders are noted in children with autism spectrum disorders (ASD), Down syndrome (DS), and intellectual disabilities (ID). ASD is characterized by impairments in language and social-emotional cognition. Multiple symptomatology of disorders, combined into an "autistic triad", includes a violation of social behavior and speech, limited forms of behavior, and a tendency to stereotypes. The degree of its expression, the age of manifestation of all characteristic symptoms, the presence of a leading symptom complex are individual for ASD children (Wing, 1993). In the pioneering work of Kanner (Kanner, 1943), the speech of patients with ASD is described as monotonous, mechanical, with even, flat intonation. Monotony of speech in ASD is described in other more recent studies: for the speech material of 4-9 year old Japanese children with ASD (Nakai, Takashima, Takiguchi, & Takada, 2010), highly functional autists aged 7-17 years (Grossman, Bemis, Skwerer, & Tager-Flusberg, 2013) and children with Asperger's syndrome (Scharfstein, Beidel, Sims, & Rendon Finnell, 2011) vs TD peers. Other researchers described high values of pitch and its variability in speech of 8-9 year old Portuguese children (Filipe, Frota, Castro, & Vicente, 2014), 4-10 year old bilinguals (Hindi – English) (Sharda et al., 2010), English school students with HFA (Diehl, Watson, Bennetto, McDonough, & Gunlogson, 2009), 4-6.5 year old Israeli children (Bonneh, Levanov, Dean-Pardo, Lossos, & Adini, 2011), 5-14 year old Russian children and adolescents (Lyakso,

Frolova, & Grigorev, 2016, 2017). The increased pitch range is supposed to be an indicator of the delay in the speech development of ASD children (Sharda, et al., 2010). Abnormal prosody of speech in patients with ASD has been identified as a main feature (Bonneh et al., 2011; Paul, Augustyn, Klin, & Volkmar, 2005; Hubbard & Trauner, 2007). A number of studies have noted the inability of patients to control pitch, atypical word and phrasal stress (Grossman et al., 2010; Diehl & Paul, 2012). Along with high pitch values and pitch range, an atypical spectrum of speech signals is described (Bonneh et al., 2011). These speech characteristics can be used in the diagnosis of autism. Down syndrome is one of chromosomal abnormalities. It is shown that in children with DS non-verbal skills develop according to psychological age and the deficit of verbal skills is maintained and strengthened (Dodd & Thompson, 2001). Children with DS have specifics in the vocal tract structure - a smaller volume of the oral and nasal cavities, lowering of the lower jaw, a narrow palate, and shorter length of the vocal tract vs. TD children. The specificity of children and adults with DS is a large folded tongue and muscular hypotonia (Kanamori, Witter, Brown, & Williams-Smith, 2000). These anatomical peculiarities cause less speech intelligibility and articulation clarity (Kent & Vorperian, 2013), the less difference between values of formant frequencies of the cardinal vowels /i/ and /u/ (Moura et al., 2008; Bunton & Leddy 2011) than in TD peers. Mixed specific developmental disorders (MDD) is a residual category for disorders in which there is some admixture of specific developmental disorders of speech and language, of scholastic skills, and motor function, but in which none predominates sufficiently to constitute the prime diagnosis. The disorders are usually associated with some degree of general impairment of cognitive functions. Mild and moderate mental retardation (ID) caused a lower level of language development. Children with ID are characterized by a delay in speech development, which affects different areas: syntax, vocabulary, speech intelligibility (Kaiser, Hester, & McDuffie, 2001; Facon, Facon-Bollengier, & Grubar, 2002). The descriptions of acoustic features of speech of children with ID are few (Lyakso, Frolova, & Karpov, 2019).

The aim of this study is to determine the acoustic features of voice and speech of children with ASD as a possible additional diagnostic criterion.

2. Methods

2.1. Data collection

Speech material of children was taken from the speech database "AD_Child.Ru" (Lyakso et al., 2019). The participants in the study were children aged 4-16 years: with ASD (n = 95 children), DS (n = 25), ID (n = 24), MDD (n = 39), and typically developing (TD) children (coevals, n=150). To assess the severity of autistic disorders, we used the Child Autism Rating Scale (CARS) score (Schopler, Reichler, DeVellis, & Daly, 1980), which was filled out by parents of children, and the child psychiatrist's report. For this study, the ASD sample was divided into two groups according to developmental features: the presence of development reversals at the age 1.5-3.0 years (first group - ASD-1) and developmental risk diagnosed at the infant birth (second group - ASD-2). For these children, the ASD is a symptom of neurological diseases associated with brain disturbed. Recording of children's speech and behavior was carried out in the most similar situations: a dialogue with the experimenter, during which the children were asked a standardized set of questions, viewing pictures and, if possible, a story about them or answers to questions. Recording was carried out in laboratory conditions. The recordings were made by the "Marantz PMD 660" recorder with a "SENNHEIZER e835S" external microphone. Speech files are stored in Windows PCM format, 44100 Hz, 16 bits per sample.

2.2. Data analysis

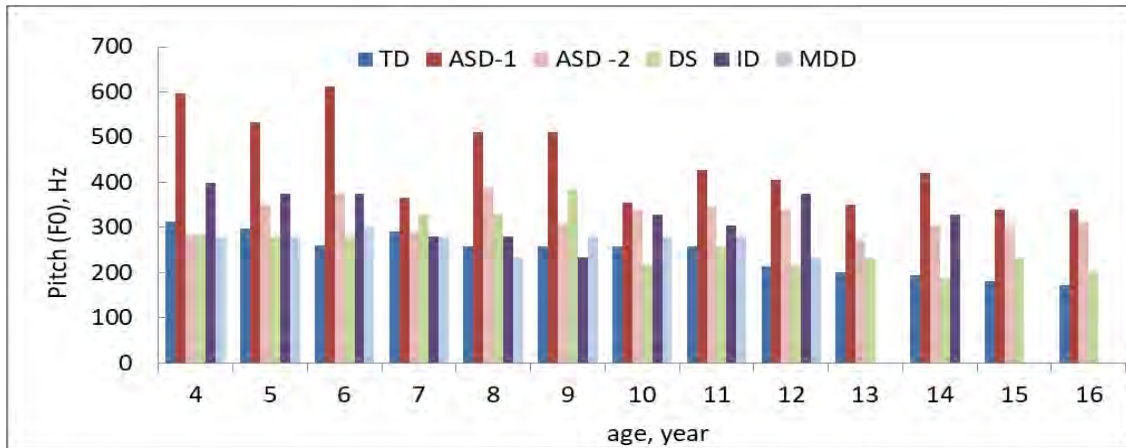
Spectrographic analysis of speech was carried out in the Cool Edit Pro 2.1 sound editor. We analyzed and compared pitch values (F0), pitch range [F0max-F0min], formants (F1 - first formant, F2 - second formant, F3 - third formant), formant intensity (E), and duration for vowels and the stationary part of vowels (the articulation is stable). The accuracy of vowel articulation was determined on the basis of values of vowel articulation index (VAI) (Roy, Nissen, Dromey, & Sapir, 2009). The analyzed acoustic features reflect the basic physiological processes occurring in the vocal tract during voice and speech production. The pitch values indicated the frequency of vocal folds vibration, the values of the first two formants - the articulation process in the oral cavity, the third formant is considered emotional formant. Statistical analysis was performed using "STATISTICA 10". Non-parametric criteria - Mann - Whitney test, Multiple regression analysis were used.

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

3. Results

A comparative analysis of the acoustic features of the speech of children with ASD, DS, ID, and MDD was carried out in order to determine the characteristics specific for disease. The analysis of the spontaneous speech of children showed that in all ages the values of pitch of children with ASD are significantly higher ($p < 0.001$ – Mann -Whitney test) than corresponding features of TD children, in ASD-1 children they are significantly higher ($p < 0.001$) than in ASD-2 children and children with ID, DS, MDD (Figure 1).

Figure 1. The pitch values of vowels in speech samples of children. Vertical axis – pitch values, Hz, horizontal axis – child age, years. Different colors indicated the pitch values of speech samples of children with different diagnoses.



The acoustic features of speech specific for different diseases were revealed:

ASD-1 (autism spectrum disorders – leading diagnosis): minimum duration of stressed vowels in the words (171.4 ± 124.1 ms average \pm std, median (med) - 145 ms); high values of pitch ($F_0 = 345.3 \pm 112.2$ Hz, med 328.5 Hz — significantly higher than corresponding data for speech of TD children, children with ASD-2, DS, MDD, on the tendency – for speech of children with ID); pitch range (61.2 ± 67.6 Hz, median – 40 Hz) higher vs ASD-2 (57.2 ± 75 Hz, med - 31 Hz), but lower than for children with DS and MDD; F_3 values (med 3269 Hz) are higher than in ASD-2 (2242 Hz) and DS; the intensity of F_3 is higher than the intensity of F_2 (E_3 / E_0 is higher vs E_2 / E_0) in comparison with data for DS and ID children; high values of VAI (0.92) vs VAI for ASD-2 children (0.88), indicating the possibility of a clear pronunciation of vowels in words.

ID: longer duration of stressed vowels in words (204.6 ± 100.9 ms, med - 186 ms), but lower vs DS; high pitch values (356.4 ± 107 Hz med - 328 Hz), higher than in children with ASD-2, DS, and MDD. The pitch range (60.8 ± 59.1 Hz, med - 47 Hz) is higher vs data for ASD-2 and DS, but does not differ from MDD in median values (47 Hz); high F_3 values (med - 3656 Hz); low VAI values (0.84).

DS: maximum values of the duration of stressed vowels (284.7 ± 195.2 ms, med - 225 ms) compared with children with ID and MDD; minimum pitch values of stressed vowels (266 ± 79.9 Hz, med 262 Hz) compared to ASD, ID, MDD and minimum F_3 values (med - 2164 Hz); minimum frequency range for the second and third formant [F_3-F_2]; the VAI varies from minimal values (0.76) for vocalizations, to high values for clear pronunciation of simple words (0.97).

MDD: the duration of stressed vowels in words (177 ± 73.9 ms, med - 167 ms) is shorter than in children with DS and ID, but higher than in ASD; pitch values (295.4 ± 64 Hz, med - 281 Hz) are higher than those of children with DS, but lower than ASD and ID; low intensity of formant frequencies; VAI values (0.96) are high.

The correlation between the acoustic features of speech and the children's age was traced (*table 1*). Age was taken as a dependent variable in order to combine data for all children in one table. The pitch values decrease with age was found for children with different diagnoses. A decrease in the duration and pitch values of stressed vowels in the stationary parts, an increase in the values of the second formant (F_2) with the age of TD children were shown. The duration of stressed vowels and pitch values, pitch range, the third formant values decrease with the age of ASD children, while the second formant values increase. In the speech of children with DS, pitch values and the duration of stressed vowels decrease with an increase in the duration of words. In speech of ID children, the duration of vowel's stationary part decreases and the duration of the vowel increases. No correlation between the acoustic features of speech and the MDD child's age was found.

Table 1. Correlation between age of children and acoustic features of speech: Multiple regression analysis data.

R ²	F	Independent variable	Beta	SE of Beta	B	SE of B	T	p-level
Dependent variable: child's age							t(1552)	
TD 0.492	(5,1552) 303.03	Dur st part vow	-0.451	0.025	0.041	0.002	17.978	0.000
		F0 vow	-0.432	0.021	-0.018	0.001	-20.56	0.000
		F2	0.092	0.018	0.000	0.000	4.984	0.000
ASD 0.128	(2,737) 53.359						t(737)	
		Dur-stress vow	-0.098	0.034	0.002	0.001	-2.839	0.004
		F0 aver-vow	-0.334	0.034	0.01	0.001	-9.652	0.000
0.188	(5,614) 28.433						t(614)	
		F0 range	-0.155	0.039	-0.009	0.002	-3.970	0.000
		F0 st part vow	-0.282	0.059	-0.007	0.001	4.728	0.000
		F2	0.216	0.050	0.0009	0.000	4.304	0.000
DS 0.303	(4,552) 59.902						t(552)	
		F0 aver-vow	-0.294	0.118	-0.012	0.004	-2.472	0.017
		F0 max	-0.280	0.108	-0.001	0.003	-2.585	0.009
0.285	(3,549) 73.113						t(549)	
		Word duration	0.118	0.038	0.000	0.000	3.102	0.002
ID	(15,35) 7.184	Dur-stress vow	0.271	0.061	0.006	0.001	4.376	0.000
		Dur st part	-0.418	0.059	-0.051	0.007	-7.021	0.000

Note: R² - correlation coefficient (R) - square; SE - standard error; B - regression coefficient. Legend: Dur - duration, st part - stationary part, aver - average values, vow - vowel, F0 - pitch values, F2 - values of the second formant, F3 - values of the third formant.

4. Discussion & conclusion

The set of acoustic features of speech specific for children with four different types of psychiatric or neurological diagnosis: autism spectrum disorders, Down syndrome, mixed specific developmental disorders, intellectual disabilities, were revealed. The fact of the stability of the revealed acoustic features in a wide age range - 4-16 years is noteworthy.

Based on the identified peculiarities specific to each disease, the set of acoustic features specific to ASD can be considered as a biomarker of autism. High pitch values; pitch variability, defined as the difference between the maximum and minimum pitch values; well-marked high-frequency in spectrum and intensity of the third formant, causing an atypical spectrogram of the speech signals; high values of the vowel articulation index, indicating the possibility of clear pronunciation of vowels in the words - set of these features can be used as an additional diagnostic criterion of autism. This will allow a timely diagnose, appoint treatment and develop individual programs for children.

The greatest similarity of acoustic features of speech was revealed for children with ASD and ID, which was also shown for linguistic characteristics. They point to the similarity of speech disorders in children with ID and ASD. These data regard the need for a more detailed comparative analysis of the speech characteristics of children with ASD and ID at all levels of the speech organization.

This work was financially supported by the Russian Science Foundation (project 18-18-00063).

References

- Bonneh, Y.S., 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, 1-7.
- Bunton, K., & Leddy, M. (2011). An evaluation of articulatory working space area in vowel production of adults with Down syndrome. *Clinical Linguistics and Phonetics*, 25, 321-334.
- Diehl, J. J., & 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.

- Diehl, J. J., Watson, D. G., Bennetto, L., McDonough, J., & Gunlogson, C. (2009). An acoustic analysis of prosody in high-functioning autism. *Applied Psycholinguistics*, *30*, 385–404.
- Dodd, B. & Thompson, L. (2001). Speech disorder in children with Down's syndrome. *Journal of Intellectual Disability Research*, *45*, 308–316.
- Facon, B., Facon-Bollengier, T., & Grubar, J. (2002). Chronological age, receptive vocabulary and syntax comprehension in children and adolescents with mental retardation. *American Journal of Mental Retardation*, *107*, 91–98.
- 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.
- 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*, 778–793.
- Hubbard, K. & Trauner, D. A. (2007). Intonation and emotion in autistic spectrum disorders. *Journal of Psycholinguist Research*, *36*, 159-173.
- Kaiser, A. P., Hester, P.P., & McDuffie, A.S. (2001). Supporting communication in young children with developmental disabilities. *Mental Retardation and Developmental Disabilities Research Reviews*, *7*, 143–150.
- Kanamori, G., Witter, M., Brown, J., & Williams-Smith, L. (2000) Otolaryngologic manifestations of Down syndrome. *Otolaryngologic Clinics of North America.*, *33*, 1285–1292.
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, *2*, 217–250.
- Kent, R. D., & Vorperian, H. K. (2013). Speech impairment in Down syndrome: a review. *Journal of Speech, Language, and Hearing Researches*, *56*, 178–210.
- Lyakso, E., Frolova, O., & Grigorev, A. (2016). A comparison of acoustic features of speech of typically developing children and children with autism spectrum disorders. *Lecture Notes in Computer Science*, *9811*, 43-50.
- Lyakso, E., Frolova, O., & Grigorev, A. (2017). Perception and acoustic features of speech of children with autism spectrum disorders. *Lecture Notes in Computer Science*, *10458*, 602–612.
- Lyakso, E., Frolova, O., & Karpov, A. (2019). A new method for collection and annotation of speech data of atypically developing children. *Proceedings of 2018 International Conference on Sensor Networks and Signal Processing, SNSP 2018*, 175-180.
- Lyakso, E., Frolova, O., Kaliyev, A., Gorodnyi, V., Grigorev, A., & Matveev, Y. (2019). AD-Child.Ru: Speech corpus for Russian children with atypical development. *Lecture Notes in Computer Science*, *11658*, 299-308.
- Moura, C. P., Cunha, L. M., Vilarinho, H., Cunha, M. J., Freitas, D., Palha, M., & Pais-Clemente, M. (2008). Voice parameters in children with Down syndrome. *Journal of Voice*, *22*, 34–42.
- Nakai, Y., Takashima, R., Takiguchi, T., & Takada, S. (2014). Speech intonation in children with autism spectrum disorder. *Brain and Development*, *36*, 516–522.
- Paul, R., Augustyn, A., Klin, A., & Volkmar, F. (2005). Perception and production of prosody by speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *35*, 205–220.
- Roy, N., Nissen, S. L., Dromey, C., & Sapir, S. (2009). Articulatory changes in muscle tension dysphonia: Evidence of vowel space expansion following manual circumlaryngeal therapy. *Journal of Communication Disorders*, *42*, 124-135.
- Scharfstein, L. A., Beidel, D. C., Sims, V. K., & Rendon Finnell, L.R. (2011). Social skills deficits and vocal characteristics of children with social phobia or Asperger's disorder: A comparative study. *Journal of abnormal child psychology*, *39*, 865–875.
- Schopler, E., Reichler, R. J., DeVellis, R. F., & Daly, K. (1980). Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *Journal of Autism and Developmental Disorders*, *10*, 91-103.
- Sharda, M., Subhadra, T. P., Sahaya, S., Nagaraja, Ch., Singh, L., Mishra, R., Sen, A., Singhal, N., Erickson, D., & Singh, N. (2010). Sounds of melody—Pitch patterns of speech in autism. *Neuroscience Letters*, *478*, 42–45.
- Wing, L. (1993). The definition and prevalence of autism: a Review. *European Child and Adolescent Psychiatry*, *2*, 61–74.