ISSN 1063-7710, Acoustical Physics, 2024, Vol. 70, No. 3, pp. 586–591. © Pleiades Publishing, Ltd., 2024.

ACOUSTICS OF LIVING SYSTEMS. BIOLOGICAL ACOUSTICS

Recognition of the Psychoneurological State of Children with Autism Spectrum Disorders Based on Speech Signals: Acoustic and Perceptual Characteristics

A. S. Nikolaev*

St. Petersburg State University, St. Petersburg, 199034 Russia

**e-mail: al.nikolajew@gmail.com*

Received March 20, 2023; revised December 5, 2023; accepted April 24, 2024

Abstract—Recognition by adults of the psychoneurological state of children with autism spectrum disorders (ASD), $n = 35$, and typically developing (TD) children, $n = 47$, aged 5–14 years, was studied. A perceptual analysis was carried out, in which adult native Russian speakers (auditors) took part, $n = 206$. For perceptual analysis, test sequences (audio tests) were created containing words and phrases of children with ASD and TD children, selected from spontaneous speech recordings. The auditors were faced with the problem of determining the psychoneurological state of a child based on auditory perception: typical–atypical development. A spectrographic analysis of the speech material of children with ASD and TD children was carried out. The phrases of children with ASD are characterized by a lower speech rate compared to the phrases of TD children, as well as fewer words, longer duration of stressed and unstressed vowels in words, higher values of the frequency of the fundamental tone in the phrase, word, and stressed and unstressed vowels.

Keywords: child speech, autism spectrum disorder, perceptual analysis, speech recognition, acoustic analysis

DOI: 10.1134/S1063771024602085

INTRODUCTION

One of the areas of speech research is the study of the possibility of identifying, from voice and speech characteristics, and the individual characteristics of the speaker, such as the speaker's age, gender [1, 2], height and weight [3], and race and ethnicity [4]. Numerous studies are aimed at identifying pathological conditions based on speech characteristics. These studies attempt to identify specific acoustic characteristics that could be used as additional diagnostic features and study speech perception by listeners. The speech characteristics of people with various diagnoses are analyzed, such as dysarthria, Parkinson's disease, Down syndrome, and mental retardation [5–9]. A number of studies using material from different languages are devoted to describing perceptual [10–12], prosodic [13], and acoustic [14–16] characteristics of the speech of children with autism spectrum disorders (ASD).

ASD combines a set of common symptoms called the autistic triad and includes disturbance or atypicality of social behavior (especially interpersonal communications), limited forms of behavior, and a tendency towards stereotypical actions, impaired language, and speech [17].

One of the main features of ASD is the underdevelopment or absence of spoken language skills in children compared to typically developing (TD) peers [18, 19]. Depending on the severity of the disorder, speech impairments can manifest themselves at different levels of its organization (articulatory, grammatical, pragmatic) and vary from fairly well-formed speech in children with high-functioning autism [20] to its complete absence (mutism). Common pathological features of the speech of children with ASD include echolalia, poor vocabulary, and impaired grammatical structure of phrases.

The speech of children with ASD is represented mainly by individual words and short phrases [18, 21], and the vocabulary is dominated by nouns [22, 23]. There is peculiar use of words in children with ASD, in particular, with the wrong meaning. Children with ASD lag behind TD children in mastering a number of grammatical categories: prepositions, pronouns, plurals, and auxiliary and modal verbs [24–26].

Most children with ASD exhibit articulation disorders: incorrect or atypical pronunciation of phonemes, unformed affricates, incorrect pronunciation of consonant clusters, etc. [16, 24, 27].

In Kanner's classic work [28], the speech of children with autism is described as monotonous; however, modern studies of the speech of children with ASD based on material from different languages indicate high values of the fundamental frequency (FF)—one of the main characteristics of sounded speech

[29]—and its variability [14, 30–32]. There are also works that show the absence of significant differences between the values of FF of children with ASD and children with TD [33], which, apparently, is due to speech recording situations. Children with ASD have abnormal prosody, and atypical word and phrase stress [13, 14, 34], and a lower speech rate [35]. The described acoustic features of the speech of children with ASD are universal and appear in children regardless of age [36].

The aim of this work was to study adult recognition of the psychoneurological state of children with ASD aged 5–14 years when listening to their speech material.

1. METHOD

The study involved 82 children aged 5–14 years: 35 children with ASD (30 boys, 5 girls) and 47 TD children (37 boys, 10 girls). The children's speech material was obtained during testing using the CEDM method [37] and was additionally selected from the AD_Child.Ru database [38] and included words and phrases cut from recordings of spontaneous speech. The AD_Child.ru database, from which the speech material was selected, contains all the necessary medical information about children.

Children with ASD who participated in the study had a child psychiatrist-confirmed diagnosis and CARS scores [39] of 31–43, which corresponds to mild to moderate severity of autistic disorder. The sample of children with ASD is represented predominantly by boys, which corresponds to the frequency of manifestation of autistic disorders in the population [40], so the group of TD children was formed in a similar way.

For the perceptual experiment, 10 test sequences (audio tests) were created, each of which contained 30 fragments of speech material; each signal in the tests was presented once, and the interval between signals was 5 s. Five test sequences contained phrases from children: four tests included speech material for each age group (5–7 years old, 8–9 years old, 10–11 years old, and 12–14 years old), one test included speech material from children 5–14 years old. Test sequences containing children's words were similarly organized.

A perceptual experiment was conducted in which adult native Russian speakers (auditors; $n = 206$; 25 ± 4.5 years; men 99, women 107; with experience interacting with children, 139 adults; without experience, 67). The auditors were faced with the problem of determining the psychoneurological state of children based on auditory perception: typical–atypical development.

Instrumental analysis of children's speech signals, correctly classified by auditors, was carried out in the Cool Edit Pro 2.0 program. The analysis included determination of the duration of a phrase, length of pauses, and speech rate (number of syllables per second). In the phrase, the word on which the semantic stress falls in the phrase was highlighted. For the selected word, the length, the duration of stressed and unstressed vowels, and the FF value for the word and vowels were determined.

Statistical data processing was carried out using the "STATISTICA 10" program.

2. RESULTS

In the perceptual experiment, auditors listened to tests containing children's words and tests containing phrases. In tests containing words from children, auditors classified 46.8% of signals from children with ASD as "atypical development" and 81% of signals from children with TD as "typical development." In tests containing phrases from children, auditors classified 65.3% of the signals from children with ASD as "atypical development"; 92.8% of signals from TD children were attributed to the "typical development" category. Male auditors were better at recognizing the condition of children with ASD than female auditors ($p < 0.01$); no significant differences were found in recognizing the state of TD children. There were no significant differences in the classification of children's condition between auditors with experience interacting with children and auditors without it.

Auditors are better at recognizing the neuropsychiatric state of children in tests containing phrases than in tests containing children's words: the average recognition completeness (UAR) for tests containing children's words is 0.64; the average recall for tests containing children's phrases is 0.79. Auditors best recognize the condition of children 5–7 and 12–14 years old. The values of average completeness were maximum in tests for determining the state of children aged 5–7 (0.67, words; 0.82, phrases) and 12–14 years (0.65, words; 0.83, phrases). The minimum values of average completeness were in tests to determine the condition of children 8–9 years old: 0.62, words; 0.72, phrases (Fig. 1).

The duration of phrases, words, and pauses between words does not differ between children with ASD and TD children. Phrases of children with ASD are characterized by a smaller number of words (5–7 years, $p < 0.001$; 8–9, 10–11 years—$p < 0.01$; 12–14 years, $p < 0.05$—Mann–Whitney test) and lower speech rate (5–7 years, $p < 0.01$; 8–9 years, $p < 0.05$; 10–11 years, $p < 0.01$; 12–14 years, $p < 0.001$) compared to the phrases of TD children (Fig. 2).

Stressed vowels of children with ASD aged 8–9, 10–11, and 12–14 years have a longer duration ($p < 0.05$, $p < 0.001$, and $p < 0.001$, respectively) compared to the vowels of TD children (Fig. 3a). Unstressed vowels of children with ASD aged 10–11 and 12–14 years have a longer duration ($p < 0.05$) compared to vowels of TD children (Fig. 3b).

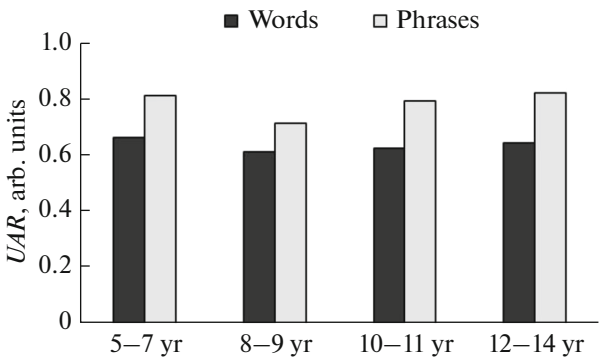

Fig. 1. Average completeness of recognition of psychoneurological state of children by auditors.

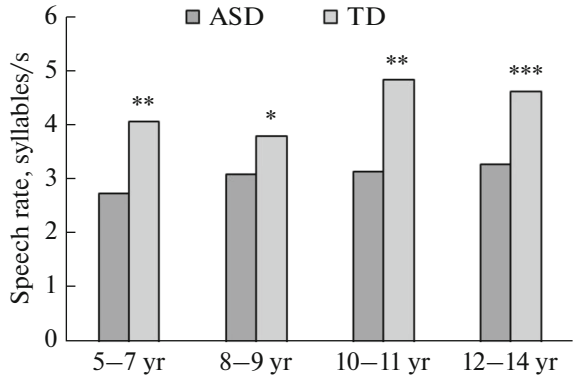

Fig. 2. Speech rate in phrases of children with ASD and TD children; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$, Mann–Whitney test.

In children with ASD aged 10–11 and 12–14 years, the FF values for the phrase and word are higher ($p < 0.01$) than in TD children (Fig. 4).

The FF values of stressed vowels in children with ASD aged 8–9, 10–11, and 12–14 years are higher ($p < 0.05$, $p < 0.05$, $p < 0.001$, respectively) than in TD children (Fig. 5a). The FF values of unstressed vowels in children with ASD aged 8–9 and 10–11 years are higher ($p < 0.05$, $p < 0.01$, respectively) than in TD children (Fig. 5b).

Speech material of children with ASD classified by auditors as belonging to TD children is characterized by higher FF values for words and phrases ($p < 0.05$) and higher values of the duration of the stressed vowel ($p < 0.01$) compared to speech material of TD children, as well as lower durations of stressed vowels compared to the speech material of children with ASD correctly classified by auditors ($p < 0.01$).

Based on correlation analysis (after Spearman, $p < 0.05$) shows the relationship between classifying children's speech signals as "atypical development" and:

(1) number of words in a phrase ($r = -0.47$);

(2) speech rate (−0.5);

(3) duration of phrase (−0.35);

(4) duration of stressed (0.47) and unstressed (0.24) vowel and pauses (0.18);

(5) FF values for a phrase (0.39), word (0.36), and stressed vowel (0.3);

(6) maximum FF values for a phrase (0.46), word (0.45), stressed vowel (0.3), and FF range for a phrase (0.42).

Based on regression analysis data, there is a relationship between classifying speech signals as "atypical development" and:

(1) speech rate F (6, 113) = 8.3045, $p < 0.0000$ ($R^2 = 0.269$, $\beta = -0.552$);

(2) maximum FF values for a phrase F (7, 112) = 7.7005, $p < 0.0000$ ($R^2 = 0.283$, $\beta = 0.548$);

3) minimum values of FF for the phrase F (7, 112) = 7.7005, $p < 0.0000$ ($R^2 = 0.83$, $\beta = -0.472$);

(4) FF range of stressed vowel F (6, 79) = 8.7813, $p < 0.0000$ ($R^2 = 0.355$, $\beta = 0.477$);

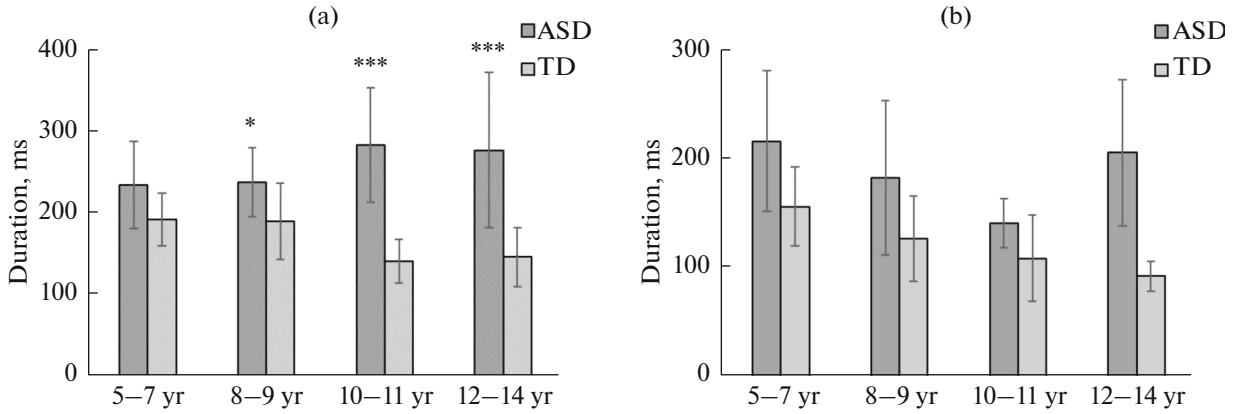

Fig. 3. Duration of stressed (a) and unstressed (b) vowels from phrases of children with ASD and TD children; * $p < 0.05$; *** $p < 0.001$, Mann–Whitney test.

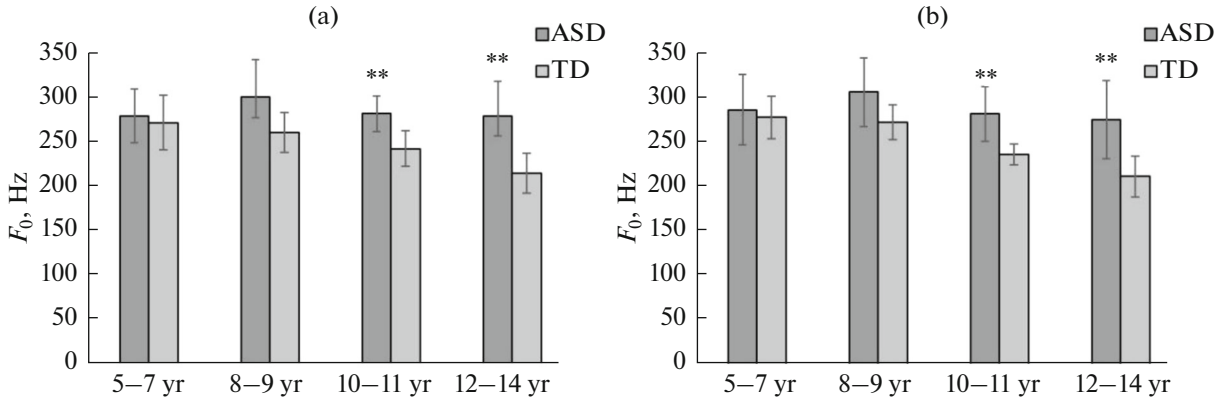

Fig. 4. FF values for phrase (a) and word (b) in children with ASD and TD children; ** $p < 0.01$, Mann–Whitney test.

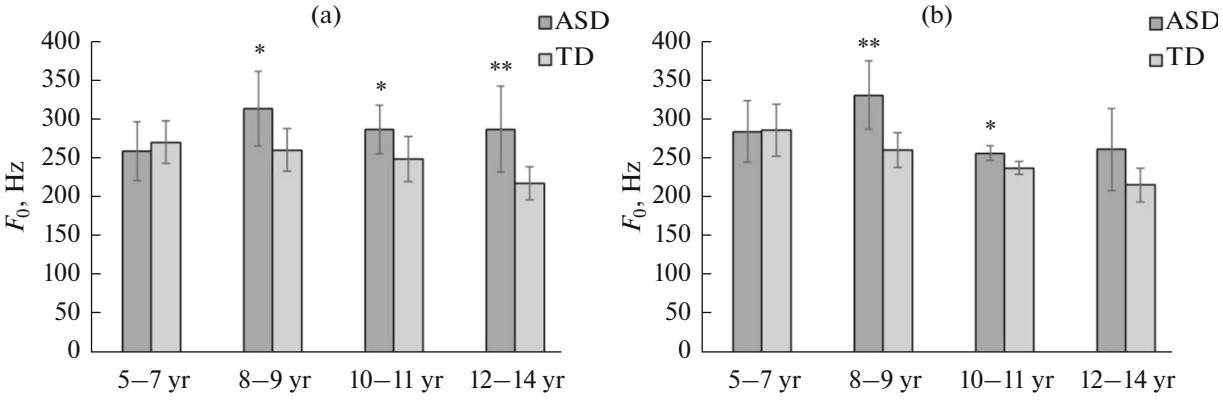

Fig. 5. FF values of stressed (a) and unstressed (b) vowels from phrases of children with ASD and TD children. * $p < 0.05$; ** $p < 0.01$—Mann–Whitney test.

(5) duration of unstressed vowel $F(6, 79) = 8.7813$, $p < 0.0000$ ($R^2 = 0.355$, $\beta = 0.267$).

3. DISCUSSION

The study showed adults' ability to correctly classify the psychoneurological state of children based on speech signals. Adults determine the condition of TD children better than the condition of children with ASD. The age of children with ASD does not affect how adults recognize their condition, but the organization of the test material does: auditors recognize children's condition better from phrases than from individual words. The data obtained agree with the results of studies conducted on speech material of children with ASD aged 11–12 years [10].

The speech material of children with ASD, classified by auditors as belonging to children with atypical development, is characterized by high FF values and its variability in phrases, words, stressed vowels, and low speech rate. These acoustic characteristics are distinctive features of the speech of children with ASD [11, 14, 31, 35]. The speech material of children with ASD, classified as belonging to children with typical development, also differs from both the speech signals of TD children and from the signals of children with ASD, correctly recognized by auditors. In particular, these are differences in the duration of stressed vowels and the FF values in words and phrases.

Auditors also classified speech signals of children with ASD with a small number of words in a phrase as "atypical development." This may indicate that auditors, when recognizing a child's psychoneurological state, rely not only on voice characteristics, but also on the grammatical structure of the utterance.

CONCLUSIONS

The study obtained data on recognition by adults, based on auditory perception, of the psychoneurological state of children with ASD and the acoustic characteristics that influence auditors' classification of children's speech signals into the category "atypical development."

The likelihood of auditors recognizing the psychoneurological state of children is higher when listening

to test material containing phrases of children compared to test material containing individual words. The speech signals of children with ASD, correctly classified by the auditors, are characterized by a lower speech rate compared to the speech signals of children with disabilities, fewer words in phrases, longer duration of stressed and unstressed vowels, higher FF values for a phrase, word, stressed and unstressed vowel, and a wider FF range for a stressed vowel.

Applications are currently being developed to support people with atypical development, their socialization, and education, e.g., [41]. The creation of such applications requires determination, first of all, of the psychoneurological state of people in order to further recognize their emotional state, taking into account the diagnosis and status. The data obtained in the study on the possibility of adults recognizing the psychoneurological state of children with ASD, the selected acoustic characteristics of children's speech, on which auditors rely, can be taken into account when developing automatic speech recognition systems.

FUNDING

The study was supported by the Russian Science Foundation (grant no. 22-45-02007).

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

All studies were conducted in accordance with the 1964 Declaration of Helsinki and its subsequent amendments. The study was approved by the Ethics Committee of St. Petersburg State University (no. 115-02-3 of April 19, 2023). Each participant (for children, their parents or legal or legal guardians) in the study gave voluntary written informed consent after being explained the potential risks and benefits and the nature of the forthcoming research.

CONFLICT OF INTEREST

The author of this work declares that he has no conflicts of interest.

REFERENCES

1. V. N. Sorokin and I. S. Makarov, Acoust. Phys. **54** (4), 571 (2008).
2. H. Goy, M. K. Pichora-Fuller, and P. van Lieshout, J. Acoust. Soc. Am. **139** (4), 1648 (2016).
3. L. Bruckert, J. S. Lienard, A. Lacroix, M. Kreutzer, and G. Leboucher, Proc. R. Soc. B: Biol. Sci. **273**, 83 (2006).
4. T. K. Perrachione, J. Y. Chiao, and P. C. M. Wong, Cognition **114**, 42 (2010).
5. H. N. Jones, K. D. Crisp, M. Kuchibhatla, L. Mahler, Th. Risoli, Jr., C. W. Jones, and P. Kishnani, Am. J. Intellect. Dev. Disabil. **124** (4), 324 (2019).
6. R. D. Kent, *Clinical Management of Sensorimotor Speech Disorders* (Thieme, 2009).
7. D. O'Leary, A. Lee, C. O'Toole, and F. Gibbon, Clin. Linguist. Phon. **34** (1-2), 72 (2020).
8. V. Verkhodanova, D. Trckova, M. Coler, and W. Lowie, in *Proc. Int. Conf. on Speech and Computer (SPECOM 2020)* (St. Petersburg, 2020).
9. E. E. Lyakso, O. V. Frolova, A. S. Nikolaev, and A. S. Grigor'ev, Ross. Fiziol. Zh. im. I. M. Sechenova **11** (107), 1395 (2021).
10. O. Frolova, V. Gorodnyi, A. Nikolaev, A. Grigorev, S. Grechanyi, and E. Lyakso, in *Proc. Int. Conf. on Speech and Computer (SPECOM 2019)* (Istanbul, 2019).
11. M. A. Redford, V. Kapatsinski, and J. Cornell-Fabiano, Lang. Speech **61** (2), 277 (2018).
12. E. E. Lyakso, O. V. Frolova, S. V. Grechanyi, Yu. N. Matveev, O. V. Verkholyak, and A. A. Karpov, in *Voice Portrait of a Child with Typical and Atypical Development,* Ed. by E. E. Lyakso and O. V. Frolova (Izd.-poligraf. Assotsiatsiya vysshykh uchebnykh zavedenii, 2020) [in Russian].
13. J. Diehl and R. Paul, Appl. Psycholinguist. **34** (1), 135 (2013).
14. Y. S. Bonneh, Y. Levanov, O. Dean-Pardo, L. Lossos, and Y. Adini, Front. Hum. Neurosci. **4** (237), 1 (2011).
15. E. E. Lyakso and O. V. Frolova, in *Proc. Int. Conf. on Multimodal Interaction ICMI'20* (Utrecht, 2020), p. 514.
16. L. Wolk and C. Brennan, Speech, Lang. Hear **19** (2), 121 (2016).
17. L. Wing, Eur. Child Adolesc. Psychiatry **2** (1), 61 (1993).
18. E. E. Lyakso, O. V. Frolova, A. S. Grigor'ev, V. D. Sokolova, and K. A. Yarotskaya, Ross. Fiziol. Zh. im. I. M. Sechenova **6** (102), 729 (2016).
19. L. R. Watson, E. Patten, G. T. Baranek, M. Poe, B. A. Boyd, A. Freuler, and J. Lorenzi, J. Speech Lang. Hear. Res. **54** (6), 1562 (2011).
20. R. B. Grossman, L. R. Edelson, and H. Tager-Flusberg, J. Speech Lang. Hear. Res. **56** (3), 1035 (2013).
21. I. Rapin, M. A. Dunn, D. A. Allen, M. C. Stevens, and D. Fein, Dev. Neuropsychol. **34** (1), 66 (2009).
22. S. Tek, L. Mesite, D. Fein, and L. Naigles, J. Autism Dev. Disord. **44** (1), 75 (2014).
23. J. Boorse, M. Cola, S. Plate, L. Yankowitz, J. Pandey, R. T. Schultz, and J. Parish-Morris, Mol. Autism **10**, 14 (2019).
24. J. Boucher, J. Child Psychol. Psychiatry **53** (3), 219 (2012).
25. G. Mazzaggio and A. Shield, J. Autism Dev. Disord. **50**, 1425 (2020).
26. A. Terzi, T. Marinis, A. Zafeiri, and K. Francis, Front. Psychol. **10**, 1301 (2019).
27. J. Cleland, F. E. Gibbon, S. J. Peppé, A. O'Hare, and M. Rutherford, Int. J. Speech-Lang. Pathol. **1** (1), 69 (2010).
28. L. Kanner, Nervous Child **2**, 217 (1943).
29. D. A. Vol'f and R. V. Meshcheryakov, Acoust. Phys. **62** (2), 215 (2016).

30. M. G. Filipe, S. Frota, S. L. Castro, and S. G. Vicente, J. Autism Dev. Disord. **44** (8), 1972 (2014).
31. E. Lyakso, O. Frolova, and A. Grigorev, n *Proc. Int. Conf. on Speech and Computer (SPECOM 2017)* (Hatfield, 2017).
32. M. Sharda, T. P. Subhadra, S. Sahaya, Ch. Nagaraja, L. Singh, R. Mishra, A. Sen, N. Singhal, D. Erickson, and N. Singh, Neurosci. Lett. **478** (1), 42 (2010).
33. L. S. Shriberg, Rh. Paul, L. M. Black, and J. P. van Santenc, J. Autism Dev. Disord. **41** (4), 405 (2011).
34. R. Paul, A. Augustyn, A. Klin, and F. Volkmar, J. Autism Dev. Disord. **35**, 205 (2005).
35. S. P. Patel, K. Nayar, G. E. Martin, K. Franich, S. Crawford, J. J. Diehl, and M. Losh, J. Autism Dev. Disord. **50** (8), 3032 (2020).
36. E. Lyakso, O. Frolova, and A. Nikolaev, *Psychological Applications and Trends* (Science Press, Lisboa, 2021), p. 259.
37. E. Lyakso, O. Frolova, E. Kleshnev, N. Ruban, M. Mekala, and K. V. Arulalan, in *Proc. Int. Conf. on Multimodal Interaction ICMI'22* (Bengaluru, 2022).
38. E. Lyakso, O. Frolova, and A. Karpov, in *Proc. Int. Conf. on Sensor Networks and Signal Processing. SNSP 2018* (Xi'an, 2018), p. 175.
39. E. Schopler, R. J. Reichler, R. F. DeVellis, and K. Daly, J. Autism Dev. Disord. **10** (1), 91 (1980).
40. J. S. Nicholas, J. M. Charles, L. A. Carpenter, L. B. King, W. Jenner, and E. G. Spratt, Ann. Epidemiol. **18** (2), 130 (2008).
41. C. Frauenberger, J. Good, and N. Pares, in *Proc. CHI Conf. on Human Factors in Computing Systems* (San Jose, CA, 2016), p. 3373.