ISSN 0022-0930, Journal of Evolutionary Biochemistry and Physiology, 2024, Vol. 60, No. 2, pp. 744–757. © Pleiades Publishing, Ltd., 2024.

EXPERIMENTAL PAPERS

Peculiarities of the Emotions Manifestation in Speech and Facial Expressions by 7–8 Years Old Children

E. A. Kleshnev[a],*, O. V. Frolova[a], and E. E. Lyakso[a],**

[a] *St. Petersburg State University, St. Petersburg, Russia*
**e-mail: kleshnev.egor@gmail.com*
***e-mail: lyakso@gmail.com*

Received July 21, 2023; revised January 18, 2024; accepted February 5, 2024

Abstract—The purpose of the research was to study the peculiarities of the reflection of emotional states "neutral (calm state)–sadness–joy–anger–fear" in the speech and facial expressions of 7–8 years old children and to identify differences in the expression of emotions by boys and girls of this age. The work was carried out as part of the study of the emotional sphere of children with typical and atypical development using the CEDM method for assessing the emotional development of children. 14 children (7 boys, 7 girls)—primary school students—took part in the study. The ability of children to express emotions in the characteristics of voice and facial expressions and to recognize emotions was analyzed. For this purpose, speech recording and video recording of children's facial expressions and behavior, instrumental and perceptual analysis of emotional speech, and automatic analysis of facial expression were carried out. The children's heart rate was recorded, and the leading hemisphere by speech was determined using a dichotic listening. The features of the reflection of the emotional states of boys and girls in the characteristics of speech are revealed. The specificity of reflecting the emotional states of children in facial expression is described. The connections between the psychophysiological indicators of children and their reflection of emotional states in their voice and facial expressions were determined.

DOI: 10.1134/S0022093024020236

Keywords: emotions, emotional speech, acoustic features, reflection of emotions, school age, children's speech, auditory perception

INTRODUCTION

Emotions are one of the factors of adaptive behaviour organisation and manifest the functional state of the organism. Emotional state is manifested in voice characteristics, speech features, non-verbal behaviour (facial expressions, pantomimicry, etc.) and in autonomic components of behavioural reactions [1]. The emotional sphere of a person depends on his or her age, upbringing and cultural environment [2, 3]. Children of primary school age are charactised by a transitional variant of emotion development—between spontaneous manifestation, specific for early age, and the formation of controlled behaviour due to inclusion in a new environment—school. These changes are associated, on the one hand, with social factors, on the other hand, with progressive changes in the morpho-functional organisation of projective and associative areas of the cerebral cortex [4, 5]. School education implies the formation of new social relations between the child and the teacher, the child and peers. New experience influences the expression of emotional states, the amount of social conflicts increases that cause negative emotions increase [6]. Throughout elementary school, children learn more about socially accept-

able rules for expressing emotions, learning which emotions to display (and which to suppress) in specific social situations [7–10].

The prefrontal cortex, which exercises top-down control of emotional reactions originating in the limbic system, does not reach the necessary level of development by the beginning of a child's schooling [11, 12]. Studies using modern neuroimaging methods [13] show that age-related changes in the amygdala, hippocampus, frontal, parietal cortex, and temporal lobe of the right and left hemispheres are non-linear and depend on the child's gender.

A large number of works are devoted to the study of children's perception of emotions with visual (facial expression) and verbal stimuli in the norm [7, 14–20] and in developmental disorders of various etiologies [21–26]. Data on the peculiarities of emotion manifestation in children were obtained mainly on the basis of questionnaires and scales for observing child behaviour [27] and assessing facial expression [28, 29]. These studies did not reveal a strict correlation between the child's gender and the success of recognising the emotions of other people. Analysis of data from 166 works [30] showed intersex differences in the display of emotions. It is shown that girls demonstrate more positive emotions than boys, boys demonstrate anger more often than girls.

Fewer works are devoted to the study of children's emotional speech and recognition of children's emotional states from their speech by adults [31]. Few works use instrumental analysis of children's emotional speech [1, 31, 32]. It was shown that recognition of children's emotional states by their speech by adults depends on the age of children and their neuropsychological state [33]. Atypical development of children may be accompanied by disorders of the emotional sphere—reflection of emotions in voice and facial expressions [25, 34–37].

To analyse the emotional sphere of children with typical and atypical development, in frame of an international study, the method—Child Emotion Development Method—CEDM was developed [38]. Methodological approach includes test tasks for reflecting and recognising emotional states. The test tasks (short version—CEDM-sh) that children with typical and atypical development are able to perform have been identified [38]. This work is the part of an overall study of children's emotional sphere and was carried out using the CEDM-sh test items to further compare the results with data obtained for children with atypical development.

The purpose of this study is to determine the peculiarities of emotion reflection in speech and facial expressions of 7–8 year old children brought up in a Russian-speaking environment, and identify differences in the expression of emotions by boys and girls of this age.

MATERIALS AND METHODS

The study involved 14 children (7 boys, 7 girls) aged 7–8 years (7.3 ± 0.6 years), students of a comprehensive school in St. Petersburg, whose native language is Russian. According to the paediatrician's report, all children developed in accordance with age norms. The criteria for inclusion of children in the study were:

1. parents signed informed consent for the child's participation in the study, approved by the Ethical Committee of St. Petersburg State University;

2. children's age—7–8 years old, attending a primary school;

3. absence of visual and hearing impairments.

The study was conducted in a comprehensive school in St. Petersburg using the Child Emotion Development Method (CEDM-sh) [38].

Study procedure

In the present study, children were offered to perform tasks on:

1) reflection of emotional states (the conversation between the experimenter and the child, "acting" speech—tasks for the child to demonstrate emotions in the voice);

2) recognition of emotional states by children using a video test.

The study included audio recording of children's speech and video recording of children's behaviour in model situations; determination of children's psychophysiological characteristics: heart rate (HR) and leading hemisphere by hearing and speech using the dichotic listening method; analysis of children's speech using perceptual research and instrumental analysis; analysis of children's facial expression in dialogue situations.

We used a standardised recording protocol [38], which included:

1. conversation between the experimenter and the

child, during which the experimenter asked a standard set of questions about school, favorite activities, hobbies, and friends. The questions asked to the child were aimed at evoking different emotional states [39];

2. the task of reflecting emotional states in the child's voice—"acting" speech. This approach was tested earlier for children of different age groups [39].

"Acting" speech is a task in which a child was asked to express five emotional states (neutral–sadness–joy–anger–fear) in the voice.

Children read and then pronounced words and phrases, representing each of the given emotions with their voices. The task included the utterance of words and phrases with a lexically specified emotional colouring and meaningless texts [38]. Meaningless texts are semantically meaningless sentences similar to real sentences, which are widely used in studies aimed at reflecting emotional states in voice characteristics without relying on linguistic context [39–41].

The speech material used was:

1. Words reflecting each of the five proposed emotional states, for example:

crush, hit, push—reflecting the state of anger;

great, beautiful, cool, super—reflecting the state of joy;

normal, well, and, three—reflecting a neutral state;

dumpish, dreary, sad—reflecting the state of sadness;

scary, terrible, horrible, horror, creepy—reflecting the state of fear.

2. Phrases such as:

It's raining outside, I'm sad, it's a sad time—demonstrates a state of sadness;

I watched a scary cartoon, it's dark outside, I'm scared, I want to hide under the blanket from horror—demonstrates the state of fear;

I like to beat and break everything, when everyone annoys me, everything annoys me—demonstrates the state of anger;

I'm fine—demonstrates a neutral state;

I love when everything is beautiful!—demonstrates a state of joy.

Pronunciation of emotionally coloured words and phrases was the easiest task for children of this age, as children could rely on the semantic component of the given text [39]. The task was complicated by the possibility of using meaningless texts—a prose text and a poem text with rhythmic organisation.

3. Meaningless texts:

3.1. An artificial phrase by L.V. Shcherba (1930), created on the basis of the Russian language, in which all root morphemes are replaced by meaningless combinations of sounds [42]:

Glokaya kuzdra shteko budlanula bokra i kurdyachit bokryonka.

3.2. The first quatrain of L. Carroll's poem "Jabberwocky" translated by S. Marshak (1967) in which the author uses non-existent words:

Twas brillig, and the slithy toves
Did gyre and gimble in the wabe:
All mimsy were the borogoves,
And the mome raths outgrabe.

Each child was asked to say each type of speech material twice, trying to express emotional states as expressively as possible.

The study was conducted in a school room, the duration was 40–60 min. Speech was recorded on a tape recorder "Marantz PMD660" with an external microphone "SENNHEIZER e835S", facial expression was recorded on a video camera "SONY HDR-CX560". The distance from the child's face to the microphone was from 3 to 50 cm.

Methods

In order to determine the reflection of children's emotional states in the characteristics of voice and speech, an instrumental analysis of children's speech and a perceptual study were conducted.

1. Spectrographic analysis of children's "acting" and spontaneous speech was carried out in the sound editor "Cool Edit Pro 2.0".

We counted: the values of the pitch values (F0)—average, maximum (F0max), minimum (F0min), intensity values E0 (dB), determined the pitch range values as the difference between its maximum and minimum values (F0max – F0min) for phrases; relative intensity value (E0max / E0min). In the works on the study of acoustic characteristics of emotional speech it was shown [32] that pitch values and intensity are the most informative characteristics.

We determined the acoustic characteristics of speech signals correctly attributed by the listeners to specific emotional states (0.75–1.0 probability of recognising an emotional state by the listeners).

2. Perceptual study. The study consisted of adults determining children's emotional states while listen-

ing to test sequences containing their emotional speech. The children's speech material (spontaneous emotional speech and "acting" speech) was initially annotated by three experts for five emotional states based on the research protocol and viewing of the video recordings. For the perceptual study, only those children's utterances were selected that were unambiguously assigned to a particular emotional state by the three experts (agreement between experts—Cohen's Kappa coefficient 1.0). Children's spontaneous emotional speech represented the child's responses in a dialogue with an adult: "I go to music school", "My friend Kolya", "We have a cat at home"—children's phrases with words reflecting emotional states were excluded from the test sequences. "Acting" speech—nonsense texts "Glokaya kuzdra", "Jabberwocky". "Acting" speech, representing words and phrases reflecting emotional states, was not included in the perceptual experiment so that adult participants of the perceptual experiment could not rely on the linguistic component of children's speech while listening.

Three test sequences with 70 speech samples were created. The interval between speech samples was 5 s. The test sequences were listened to by adult native speakers of Russian (listeners)—53 people (age 23.9 ± 2.1 years, 37 women, 16 men). All listeners had no hearing problems—0–25 dB (hearing thresholds according to tonal audiometry method).

The results of the perceptual analysis were presented in the form of confusion matrices—tables, the rows of which correspond to the specified classes, the columns—to the actual values (real classes). The following was calculated: completeness—the share of samples actually belonging to a given class, relative to all samples of this class in the test sample; precision—the share of samples actually belonging to this class, relative to all samples that listeners attributed to this class; F-measure—a value that allows estimating the ratio of completeness and precision; unweighted average recall (UAR) for all emotional states.

3. In order to control the emotional state of the child in the dialogue with the experimenter, the analysis of children's spontaneous facial expressions was carried out. For this purpose, three experts selected video fragments from the experimenter's dialogue with the child lasting one minute. The video clips were analysed in the program "FaceReader 8v." (Noldus, the Netherlands) (Fig. 1). FaceReader software runs on the Microsoft Azure cloud platform.

"FaceReader 8v." programme (Noldus) automatically determines the time of manifestation of basic emotions in a child's facial expressions, average values of activation and valence of children's emotional expressions based on the algorithms.

4. In order to assess the child's ability to identify emotional states, children were presented with a video test containing 16 graphic images showing sadness, joy, anger, and neutral states (4 images for each state) on a laptop. Images reflecting the state of fear were not included in the video test, based on the recommendations of the St. Petersburg University Ethical Committee on conducting research with children. The child's responses were recorded in the questionnaire, and then confusion matrices were constructed.

To control the degree of activation of the autonomic nervous system (arousal of the child), HR was measured using a ChoiceMMed pulse oximeter before the start of the recording and after its completion.

The leading hemisphere in speech was determined using the dichotic listening method to identify the strategy used by the child in recognising and manifesting emotional states [43]. The coefficient of lateral preference (CLP) was calculated:

$$LPC = (R - L) \times 100 / (R + L)\ (\%),$$

where R is the number of "right choices", i.e. words from the right ear; L is the number of "left choices", i.e. words from the left ear.

Statistical analysis of data was performed in Statistica 10 software using the Mann–Whitney test, Spearman rank correlation (at a significance level of $p < 0.05$), and regression analysis. Agreement between experts in data annotation was assessed on the basis of the Cohen's Kappa coefficient [44, 45].

RESULTS

Characteristics of children's emotional speech. On the basis of instrumental analysis of speech, it is shown that the pitch values of children's speech are the highest in the state of joy—310 ± 61 Hz (mean ± standard deviation) and fear—299 ± 54 Hz. Pitch values for the neutral (calm) state are 254 ± 34 Hz,

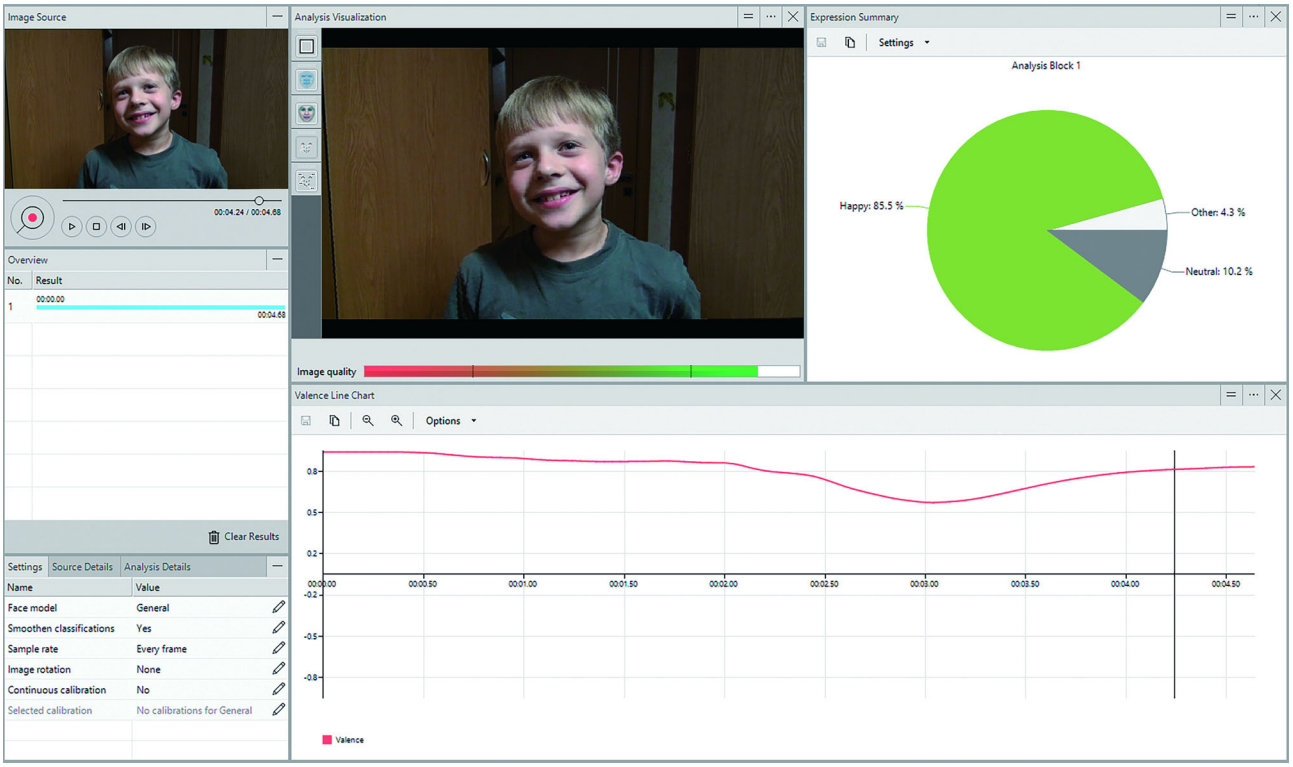

Fig. 1. Program FaceReader 8v.—facial expression of state of joy (boy 8 years old).

Table 1. Comparison of spontaneous and "acting" speech of boys and girls

Gender	Boys				Girls		
Emotions	N	S	J	A	N	S	J
F0			A > Sp		A > Sp		A > Sp
p			0.002		0.002		0.012
F0max – F0min	A > Sp		A > Sp	A > Sp	A > Sp		A > Sp
p	0.043		< 0.001	0.004	0.001		< 0.001
E0max / E0min	A > Sp	A > Sp	A > Sp	A > Sp	A > Sp	A > Sp	A > Sp
p	< 0.001	< 0.001	0.002	< 0.001	< 0.001	0.012	< 0.001

Designations: A—"acting" speech, Sp—spontaneous speech, N—neutral, S—sadness, J—joy, A—anger, *p*—*p*-value (Mann–Whitney test).

sadness—271 ± 45 Hz, anger—292 ± 59 Hz. The characteristics of spontaneous and "acting" speech of boys and girls were compared (Table 1). Significant differences were found in intensity values: in "acting" speech the intensity values were higher in the neutral state, sadness state, and joy state for boys and girls. Pitch range values are higher in "acting" speech in the state of joy and neutral state for boys and girls. In the anger state, pitch range values and intensity values are higher in "acting" speech in boys.

For spontaneous speech, differences were found in the pitch values of boys and girls in the sad state—boys had higher vs. girls ($p < 0.05$—Mann–Whitney test). In boys' spontaneous speech, significant differences were found between the sad and neutral states—F0max in the neutral state is lower than in the sad state ($p < 0.05$); sad and angry states—pitch

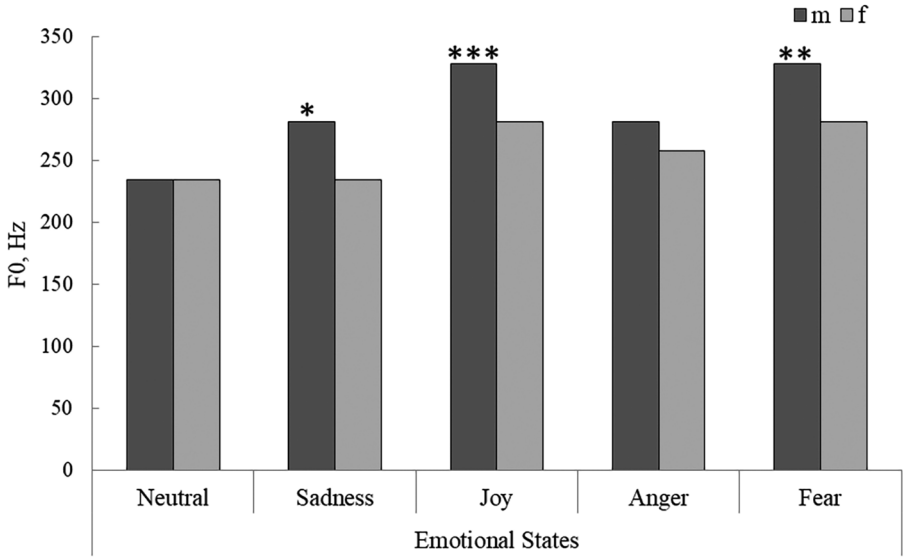

Fig. 2. The pitch values (F0) of boys and girls speech in emotional states—neutral, sadness, joy, anger, fear. Emotional states are marked on the horizontal axis, pitch values, Hz are marked on the vertical axis. Columns m are pitch values for boys, columns f are pitch values for girls. ***—$p < 0.001$—Mann–Whitney test, differences between boys and girls.

Table 2. Comparison of the acoustic characteristics of "acting" speech in different emotional states in boys and girls

Gender	Acoustic characteristics	Comparison of children's emotional states							
Boys	F0	N < S	N < J	N < A	N < F	S < J	S < F	J > A	F > A
	p	0.019	< 0.001	0.016	< 0.001	< 0.001	0.025	< 0.001	0.012
	F0max – F0min	N < S	N < J	N < C	J > A	J > F			
	p	0.048	< 0.001	0.037	< 0.001	< 0.001			
	E0max / E0min	N < S	N < J	N < A	N < F				
	p	0.001	0.001	0.004	0.002				
Girls	F0	N < J	N < A	N < F	S < J				
	p	< 0.001	0.001	< 0.001	0.008				
	F0max – F0min	N < J	N < A	N < C	S < J	J > A	J > F		
	p	< 0.001	0.033	0.022	0.004	0.011	0.016		
	E0max / E0min	N < J	S < J						
	p	0.021	0.039						

Designations: N-neutral, S—sadness, J—joy, A—anger, F—fear; *p*—*p*-value (Mann–Whitney test).

range values and F0max in the sad state is higher than in the angry state ($p < 0.05$). In girls' spontaneous speech—F0min is higher in the joy state than in the neutral state ($p < 0.05$).

For "acting" speech, differences in pitch values of boys and girls are shown (Fig. 2). The pitch values of boys in the states of sadness ($p < 0.001$), joy ($p < 0.001$) and fear ($p < 0.001$) are significantly higher than the pitch values of girls in these states.

Acoustic characteristics of "acting" speech of boys and girls in different emotional states are presented in Table 2. Neutral state in boys is characterised by

Table 3. Confusion matrix for determining emotions based on spontaneous speech of boys, % of listeners' responses

State	Neutral	Sadness	Joy	Anger
Neutral	**65**	18	8	9
Sadness	37	**44**	7	12
Joy	38	5	**52**	5
Anger	8	22	6	**64**
Recall	0.65	0.44	0.52	0.64
Precision	0.44	0.49	0.71	0.71
F1-score	0.52	0.46	0.60	0.67
UAR (Unweighted Average Recall)—0.56				

The % of listeners' correct responses in recognising emotional states is highlighted in bold.

Table 4. Confusion matrix for determining emotions based on spontaneous speech of girls, % of listeners' responses

State	Neutral	Sadness	Joy	Anger
Neutral	**53**	26	18	3
Sadness	35	**56**	3	6
Joy	13	9	**76**	2
Anger	58	30	8	**4**
Recall	0.53	0.56	0.76	0.04
Precision	0.33	0.51	0.74	0.07
F1-score	0.41	0.51	0.74	0.07
UAR (Unweighted Average Recall)—0.47				

The % of listeners' correct responses in recognising emotional states is highlighted in bold.

the lowest pitch values and intensity, in girls—by lower pitch values, pitch range values compared to the states of joy, anger and fear. Boys in the sadness state have lower pitch values than in the joy and fear states, and girls have lower pitch values than in the joy state. The state of joy is characterised by the highest pitch range values in girls, and in boys—by higher pitch range values compared to the neutral state, the states of anger and fear. In the state of anger in boys, pitch values are lower than in the states of joy and fear.

The relationship between speech material and E0max/E0min was determined $F_{(1, 339)} = 65.737$ $p < 0.0001$ ($R^2 = 0.162$, $\beta = -0.403$)—regression analysis. Children's spontaneous speech had lower intensity values than "acting" speech. A correlation between the children's gender and the acoustic characteristics of their speech was found: F0 $F_{(1, 339)} = 34.541$ $p < 0.0001$ ($R^2 = 0.111$, $\beta = -0.333$); F0max – F0min—$F_{(1, 339)} = 9.379$ $p < 0.01$ ($R^2 = 0.033$, $\beta = -0.181$). The pitch values and pitch range values were higher in boys than in girls. The relationship between children's emotional state and acoustic characteristics of "acting" speechwas revealed: F0 $F_{(1, 289)} = 36.577$ $p < 0.001$ ($R^2 = 0.312$, $\beta = 0.312$); F0max – F0min $F_{(1, 289)} = 12.600$ $p < 0.05$ ($R^2 = 0.036$, $\beta = 0.190$); E0max/E0min $F_{(1, 289)} = 18.699$ $p < 0.001$ ($R^2 = 0.052$, $\beta = 0.229$)—based on these characteristics, differences between different emotional states are shown.

Recognition of children's emotional states from their speech by adults (perceptual study data). No significant differences were found between men and women in recognition of children's emotional states, so the data are presented for the total sample of listeners. When listening to children's spontaneous speech, listeners are better at recognising states of joy (64%—correct answers) and neutral states (57%), worse at recognising states of sadness (50%) and anger (48%). The UAR = 0.52. In the spontaneous speech of boys, listeners recognise neutral state (65%) and anger state (64%) better, and sadness state (44%) worse (Table 3). From the spontaneous speech of girls, listeners recognise the state of joy better (76%). The state of anger in girls is worse recognised by listeners (4%) (Table 4).

When recognising children's emotional state from the text of the meaningless text "Jabberwocky", listeners are better at recognising the neutral state (63%) and the state of joy (60%), worse at recognising the states of fear (52%), sadness (54%) and anger (47%). The UAR = 0.55. From the speech of boys, listeners recognise fear better (85%) and sadness worse (54%) (Table 5). From the speech of girls the listeners recognise neutral state better (64%), anger state worse (35%) (Table 6).

When recognising children's emotional state via the meaningless text—"Glokaya kuzdra" listeners are better at recognising the state of sadness (68%), neutral state (69%) and the state of joy (60%), worse

Table 5. Confusion matrix for determining emotions by the "acting" speech of boys—"Jabberwocky", %

State	Neutral	Sadness	Joy	Anger	Fear
Neutral	**61.5**	14	13	1.5	10
Sadness	28	**54**	2	12	4
Joy	25	4	**60**	4	7
Anger	23	8	4	**65**	0
Fear	15	0	0	0	**85**
Recall	0.62	0.54	0.60	0.65	0.85
Precision	0.40	0.68	0.76	0.79	0.80
F1-score	0.49	0.60	0.67	0.71	0.83
UAR—0.65					

The % of listeners' correct responses in recognising emotional states is highlighted in bold.

Table 6. Confusion matrix for determining emotions by the "acting" speech of girls—"Jabberwocky", %

State	Neutral	Sadness	Joy	Anger	Fear
Neutral	**64**	19	0	2	15
Sadness	32	**54**	0	0	14
Joy	31	0	**61**	1	7
Anger	27	3	8	**35**	27
Fear	31	4	19	2	**44**
Recall	0.64	0.54	0.61	0.35	0.44
Precision	0.35	0.68	0.69	0.88	0.41
F1-score	0.45	0.60	0.65	0.50	0.43
UAR—0.52					

The % of listeners' correct responses in recognising emotional states is highlighted in bold.

Table 7. Confusion matrix for determining emotions by the "acting" speech of boys—"Glockaya kuzdra", %

State	Neutral	Sadness	Joy	Anger	Fear
Neutral	**54**	40	2	0	4
Sadness	9	**73**	3	1	14
Joy	15	4	**67**	6	8
Anger	21	4	4	**58**	13
Fear	14	29	10	7	**40**
Recall	0.54	0.73	0.67	0.58	0.40
Precision	0.48	0.49	0.78	0.81	0.51
F1-score	0.51	0.58	0.72	0.67	0.45
UAR—0.58					

Table 8. Confusion matrix for determining emotions by the "acting" speech of girls—"Glockaya kuzdra", %

State	Neutral	Sadness	Joy	Anger	Fear
Neutral	**84**	7	6	1	2
Sadness	26	**64**	0	0	10
Joy	45	1	**52**	0	2
Anger	36	12	18	**32**	2
Fear	25	14	19	5	**37**
Recall	0.84	0.64	0.52	0.32	0.37
Precision	0.39	0.65	0.55	0.84	0.70
F1-score	0.53	0.65	0.53	0.46	0.48
UAR—0.54					

at recognising the state of anger (45%) and fear (40%). The UAR = 0.56. When identifying the emotional states of boys, the listeners recognise the state of sadness better (73%), and the state of fear worse (40%) (Table 7). When recognising the emotional states of girls, the listeners are better at recognising the neutral state (84%), worse at recognising the anger state (32%) (Table 8).

Acoustic characteristics of speech signals correctly attributed to emotional states

It is shown that the most significant characteristics in recognising emotional states by listeners are minimum pitch values, pitch range values and intensity. A correlation between correct recognition of emotional state and acoustic characteristics of speech was revealed: F0min—$F_{(1, 208)}$ = 171. 45, $p < 0.01$ ($R^2 = 0.449$, $\beta = 0.672$); E0max / E0min—$F_{(1, 208)}$ = 60.163 $p < 0.001$ ($R^2 = 0.225$, $\beta = 0.474$); F0max – F0min—$F_{(1, 208)}$ = 8.634, $p < 0.01$ ($R^2 = 0.035$, β = 0.199).

Neutral state—F0min—214 ± 23.5 Hz; F0max – F0min—114 ±5 3.9 Hz; E0max / E0min—3.5 ± 3.4; sadness state—F0min—214 ± 32 Hz; F0max – F0min—127 ± 73 Hz; E0max / E0min—6 ± 4.4; joy state—F0min—220 ± 32. 6 Hz; F0max – F0min—

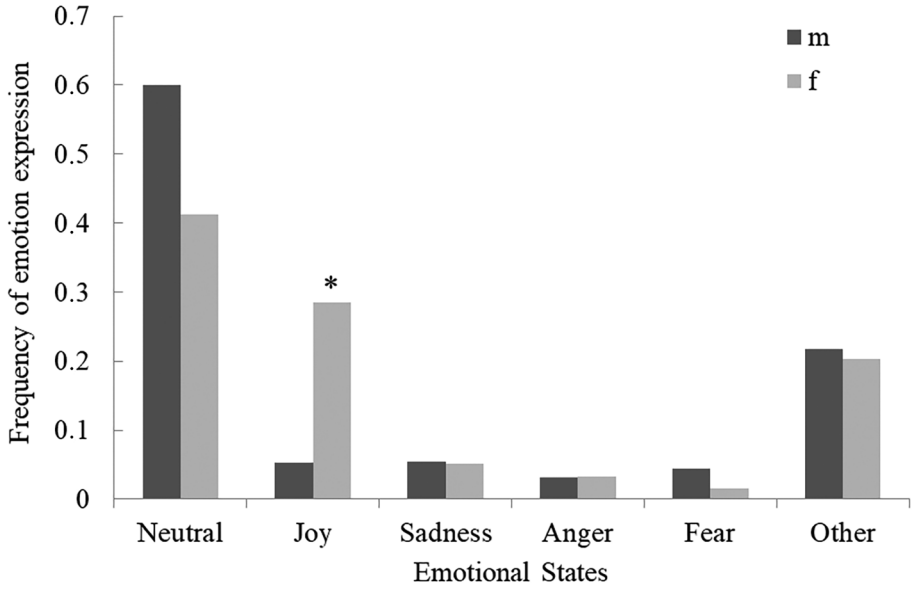

Fig. 3. The frequency of emotions manifestation in children's facial expressions. Emotional states are indicated on the horizontal axis, and the frequency of emotion expression is indicated on the vertical axis. Columns m—frequency of emotions in boys, columns f—frequency of emotions in girls. *—$p < 0.05$—Mann–Whitney test, differences between boys and girls.

143.5 ± 65 Hz; E0max / E0min—6.5 ± 6.8; anger state—F0min—211 ± 25.2 Hz; F0max – F0min—77. 8 ± 38 Hz; E0max / E0min—8.5 ± 3.3; fear state—F0min—249.7 ± 27.1 Hz; F0max – F0min—78.3 ± 54.3 Hz; E0max / E0min—6.7 ± 9.

The neutral state is characterised by lower pitch range values compared to the joy and fear states ($p < 0.05$); the anger state by higher intensity values compared to the neutral state ($p < 0.05$); the joy state by higher pitch range values compared to the anger state ($p < 0.05$).

Analysis of children's facial expression

It is shown that in girls and boys the predominant emotional state in spontaneous facial expressions is the neutral state (0.6 and 0.41—frequency of manifestation in boys and girls, respectively). Girls show the state of joy (0.29) more often than boys (0.05) in facial expression. Fear in facial expressions is expressed more often in boys (0.04) than in girls (0.01) (Fig. 3).

In facial expression of boys, emotions with negative valence prevail (–0.113—mean value of valence), in girls—with positive valence (0.204).

A correlation was found between child gender and: frequency of facial expressions of emotions with positive valence $F_{(1, 339)} = 284.80$, $p < 0.0001$ ($R^2 = 0.456$, $\beta = 0.675$); negative valence $F_{(1, 339)} = 223.55$, $p < 0.0001$ ($R^2 = 0.397$, $\beta = -0.630$)—regression analysis. Girls more often than boys show emotions with positive valence.

Children's recognition of emotional states by video test

Children are better at recognising the emotional states of joy (94%—number of correct answers) and sadness (92%), while they are worse at recognising the neutral state (74.5%) and the state of anger (62.5%). The UAR = 0.81.

Boys (82.2%) recognise emotional states better than girls (79.2%) on the video test ($p < 0.05$) (Table 9). Girls are better at recognising the sadness state (100%) than boys (83.3%) ($p < 0.001$) (Table 10). Boys are better at recognising the neutral state (87%) than girls (62.5%) ($p < 0.001$).

Relationships between the expression of emotional states, their recognition and psychophysiological characteristics of children

On the basis of regression analysis the links between the leading hemisphere on the child's speech and:

listeners' recognition of his emotional states $F_{(1, 339)} = 28.081$, $p < 0.0001$ ($R^2 = 0.081$, $\beta = 0.290$); positive valence in facial expressions—$F_{(1, 339)} = 6.078$, $p < 0.05$ ($R^2 = 0.176$, $\beta = 0.133$); child's correct recognition of images attributed to a neutral state—

Table 9. Confusion matrix for determining emotions by video test for boys, %

State	Neutral	Sadness	Joy	Anger
Neutral	**87**	0	8.7	4.3
Sadness	4.2	**83.3**	4.2	8.3
Joy	8.3	0	**91.7**	0
Anger	12.5	16.7	4.2	**66.7**
Recall	0.87	0.833	0.92	0.67
Precision	0.78	0.83	0.84	0.84
F1-score	0.82	0.83	0.88	0.74
UAR—0.82				

Table 10. Confusion matrix for determining emotions by video test for girls, %

State	Neutral	Sadness	Joy	Anger
Neutral	**62.5**	16.7	16.7	4.2
Sadness	0	**100**	0	0
Joy	4	0	**96**	0
Anger	16.7	16.7	8.3	**58.3**
Recall	0.63	1	0.96	0.58
Precision	0.75	0.75	0.79	0.93
F1-score	0.68	0.86	0.87	0.72
UAR—0.79				

$F_{(1, 339)} = 74.874$, $p < 0.001$ ($R^2 = 0.192$, $\beta = -0.438$). Emotional states of children with a leading right hemisphere are better recognised by listeners from speech. Children whose leading hemisphere by speech is left hemisphere show fewer emotions with positive valence in facial expressions and are better at recognising neutral states on the video test. Correlations were found between child HR values and acoustic characteristics of speech: pitch range values—F0max – F0min $F_{(1, 339)} = 10.792$, $p < 0.01$ ($R^2 = 0.03$, $\beta = 0.176$); E0max / E0min—$F_{(1, 339)} = 4.215$, $p < 0.05$ ($R^2 = 0.012$, $\beta = 0.110$)—the higher the HR values, the higher the pitch range values and voice intensity. A correlation was shown between the child's HR after recording "acting" speech and the level of emotion activation in facial expressions $F_{(1, 339)} = 5.6290$, $p < 0.05$ ($R^2 = 0.163$, $\beta = 0.128$)—the more excited the child is, the more vividly they show emotions.

It is shown that children who show more emotions with positive valence are worse at recognising the emotional states of anger and neutral state on the video test. Correlations were found between the frequency of emotions with positive valence in the child's facial expressions and the number of images correctly recognised by the child demonstrating the neutral state $F_{(1, 339)} = 14.380$, $p < 0.05$ ($R^2 = 0.044$, $\beta = -0.201$) and the anger state $F_{(1, 339)} = 24.826$, $p < 0.0001$ ($R^2 = 0.073$, $\beta = -0.271$).

The frequency of children's expressions of emotions with negative valence in facial expression was related to the number of images correctly recognised by the child reflecting anger $F_{(1, 339)} = 56.763$, $p < 0.0001$ ($R^2 = 0.153$, $\beta = 0.391$). Children who show more emotions with negative valence are better at recognising anger in images.

DISCUSSION

In this work, based on the standardised CEDM-sh methodology, data on the features of emotion reflection in speech and facial expressions of 7–8 year old children have been obtained.

Differences in acoustic characteristics of children's speech reflecting different states "neutral–sadness–joy–anger–fear" are shown. Our data agree with the results obtained in studies of children's and adult speech on the material of different languages [32, 38, 39, 46–49]: the states of anger and fear are characterised by high pitch values, pitch range values and intensity, the state of sadness—by average (in comparison with other emotional states) pitch values, pitch range values, for the neutral state—by low pitch values, its range and intensity [46]. The connection between the manifestation of emotional states and acoustic characteristics of children's emotional speech, revealed in the work for children 7–8 years old, agrees with the data obtained on the material of Russian and Tamil languages for children 8–12 years old [39].

Researchers have linked the increase in pitch values, pitch range values and intensity to the degree of emotional arousal [31]. For example, for preschool children, there is a connection between emotional arousal in the state of joy and an increase in pitch values, pitch range values and intensity [50]. Physio-

logical changes due to activation of the autonomic nervous system are reflected in the acoustic characteristics of emotional speech [49, 51]. All emotional states except neutral (calm) are characterised by an increase in HR [52], which agrees with the results of our study, in which it was shown that changes in HR are related to the acoustic characteristics of children's speech, which listeners rely on when correctly recognising emotional states.

The perceptual study showed that adult native speakers are more successful in identifying basic emotions from children's "acting" speech than from spontaneous speech. These data are consistent with the results of a cross-cultural study on the material of Russian and Tamil languages on the recognition of emotions in spontaneous and "acting" speech of children 8–12 years old [39]. When listening to the spontaneous speech of boys, the listeners better recognise the state of anger and neutral state, worse at recognising the state of sadness; in the spontaneous speech of girls, they were better at recognising the state of joy, worse at recognising the state of anger. From girls' "acting" speech, adults also identified anger states worse than other emotional states. In general, adults identified neutral states well from children's "acting" speech, which corresponds to the data on recognition of adult "acting" speech obtained on the material of the Russian language [53]. In Morozov's research it is noted that the state of fear is recognised better than joy by the "acting" speech of adults [53]. In our work, this pattern is shown only for the "acting" speech of boys (the text "Jabberwocky"), which may be related to the age of informants and recording conditions.

Children of 7–8 years old when performing tasks in the CEDM method (dialogue with the experimenter) more often demonstrate facial expressions corresponding to a neutral state. Girls more often than boys show emotions with positive valence—joy. In studies for children growing up in different cultural environments, it is shown that boys of this age more often show anger, and girls—joy [30]. Significant differences between boys and girls in the manifestation of positive emotions are absent in infancy and preschool age, in the manifestation of anger differences between boys and girls are not expressed in infancy, appear in preschool age, and in adolescence girls demonstrate more negative emotions than boys. Men more often demonstrate neutral state and anger, women—the state of joy [32].

From the images included in the video test, children were better at identifying the states of joy and sadness than the neutral state and anger. The connection between the child's ability to show negative emotions in facial expressions and to recognise the state of anger from images was established. In foreign works [14, 16] it is shown that children of 7–8 years old recognise the state of joy better than the state of anger and sadness by facial expression. Russian children 8–9 years old recognise joy more successfully than sadness, and sadness more successfully than anger [20]. The same pattern was found for German preschool children [15].

The links between psychophysiological features of children and their ability to manifest emotions in speech have been revealed—listeners better recognise emotional states from the speech of children with a leading right hemisphere in speech, which agrees with the ideas about the role of the right hemisphere in reflecting emotional states in human speech [54]. The right hemisphere is associated with the manifestation and regulation of emotions in the voice—changes in pitch range value and intensity [55, 56].

Thus, the normative data on the peculiarities of reflection and perception of emotional states of typically developing children of 7–8 years old growing up in Russian-speaking environment have been obtained. Differences in the manifestation of emotions by boys and girls of this age have been revealed. The acoustic characteristics of children's emotional speech are described, the most informative characteristics that listeners rely on to determine the emotional state from the child's voice—F0max – F0min, E0max / E0min—are indicated: the correlation between these characteristics and HR values is revealed.

CONCLUSIONS

1. Differences in acoustic characteristics of boys' and girls' speech reflecting emotional states "neutral–sadness–joy–anger–fear" were revealed. It is shown that the pitch values of boys are higher than that characteristic of girls in the states of sadness, joy and fear.

2. The peculiarities of recognition by adults of children's emotional states by speech were deter-

mined. The accuracy of recognising emotional states from boys' speech is higher than from girls' speech, listeners recognise emotional states better from "acting" speech.

3. The features of reflecting children's emotional states in facial expressions are described: in the situation of dialogue with the experimenter, girls more often show the state of joy in facial expressions than boys.

4. The links between the leading hemisphere of the child's speech and its reflection of emotional states in the characteristics of the voice, between the values of HR and acoustic characteristics of the voice were shown.

The obtained data on the peculiarities of the reflection of emotional states by children of primary school age will be supplemented by a larger sample of children living in a Russian-speaking environment and can be used for comparative analysis in the study of the emotional sphere of children with developmental disorders.

AUTHORS' CONTRIBUTION

Idea and design of the experiment (E.E.L.), conducting the experiment and collecting data (E.A.K., O.V.F.), data processing (E.A.K.), writing and editing the text (E.A.K., O.V.F., E.E.L.).

FUNDING

This work was funded from the budget of the Russian Science Foundation (Project no. 22-45-02007). No additional grants were received to conduct or supervise this particular study.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

All studies were conducted in accordance with the principles of biomedical ethics set out in the 1964 Declaration of Helsinki and its subsequent amendments. They were also approved by the Ethical Committee of St. Petersburg State University, Minutes no. 115-02-3 of April 19, 2023.

Each participant (for children, from their parents) 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 authors of this work declare that they have no conflicts of interest.

REFERENCES

1. Lyakso EE, Frolova OV (eds) (2020) Voice portrait of a child with typical and atypical development. Publ house—polygraph association of higher educat institut. (In Russ).
2. Ekman P (1982) Methods for measuring facial action. Handbook of methods in nonverbal behavior research. Cambridge Univer Press, New York.
3. Izard CE (1977) Theories of Emotion and Emotion-Behavior Relationships. In: Human Emotions, 19–42. https://doi.org/10.1007/978-1-4899-2209-0_2
4. Hudspeth WJ, Pribram KH (1990) Stages of Brain and Cognitive Maturation. J Educ Psychol 82: 881–884. https://doi.org/10.1037/0022-0663.82.4.881
5. Hudspeth WJ, Pribram KH (1992) Psychophysiological indices of cerebral maturation. Int J Psychophysiol 12: 19–29. https://doi.org/10.1016/0167-8760(92)90039-E
6. Hernández MM, Eisenberg N, Valiente C, Spinrad TL, Berger RH, VanSchyndel SK, Silva KM, Diaz A, Thompson MS, Gal DE, Southworth J (2018) Bidirectional associations between emotions and school adjustment. J Pers 86: 853–867. https://doi.org/10.1111/jopy.12361
7. de Bordes PF, Hasselman F, Cox RFA (2021) Children's perception of facial expressions. Dev Psychol 57: 506–518. https://doi.org/10.1037/DEV0000979
8. Thrasher C, Grossmann T (2021) Children's emotion perception in context: The role of caregiver touch and relationship quality. Emotion 21: 273–282. https://doi.org/10.1037/EMO0000704
9. Jones DC, Abbey BB, Cumberland A (1998) The Development of Display Rule Knowledge: Linkages with Family Expressiveness and Social Competence. Child Dev 69: 1209–1222. https://doi.org/10.1111/J.1467-8624.1998.TB06168.X
10. Sauter DA, Panattoni C, Happé F (2013) Children's recognition of emotions from vocal cues. Br J Dev Psychol 31: 97–113. https://doi.org/10.1111/J.2044-835X.2012.02081.X
11. Casey BJ, Heller AS, Gee DG, Cohen AO (2019) Development of the emotional brain. Neurosci Lett 693: 29–34. https://doi.org/10.1016/J.NEULET.2017.11.055
12. Plessen KJ, Kabicheva G (2010) Hjernen og følelser—fra barn til voksen. Tidsskr den Nor

Laegeforening 130: 932–935. https://doi.org/10.4045/TIDSSKR.09.0255
13. Giedd JN, Blumenthal J, Jeffries NO, Castellanos FX, Liu H, Zijdenbos A, Paus T, Evans AC, Rapoport JL (1999) Brain development during childhood and adolescence: a longitudinal MRI study. Nat Neurosci 210 (2): 861–863. https://doi.org/10.1038/13158
14. Widen SC (2013) Children's interpretation of facial expressions: The long path from valence-based to specific discrete categories. Emot Rev 5: 72–77. https://doi.org/10.1177/1754073912451492
15. Covic A, Von Steinbüchel N, Kiese-Himmel C (2020) Emotion Recognition in Kindergarten Children. Folia Phoniatr Logop 72 (4): 273–281. https://doi.org/10.1159/000500589
16. Brechet C (2017) Children's Recognition of Emotional Facial Expressions Through Photographs and Drawings. J Genet Psychol 178: 139–146. https://doi.org/10.1080/00221325.2017.1286630
17. Kawahara M, Sauter DA, Tanaka A (2021) Culture shapes emotion perception from faces and voices: changes over development. Cogn Emot 35: 1175–1186. https://doi.org/10.1080/02699931.2021.1922361
18. Dmitrieva ES, Gel'man VYA (2018) Perception of emotional auditory and visual information and the success of learning in younger schoolchildren. Psiholog nauka i obrazovan 23: 29–39. (In Russ).
19. Dmitrieva ES, Gel'man VYa, Zajceva KA, Orlova AM (2006) Study of the perception of the emotional component of a speech signal at different stimulus durations. Fiziol chelov 32: 36–41. (In Russ).
20. Dmitrieva ES, Anderson MN, Gel'man VYA (2016) A comparative study of visual and auditory perception of emotions in children of primary school age. Eksper psihol 9: 38–52. (In Russ).
21. Paine AL, van Goozen SHM, Burley DT, Anthony R, Shelton KH (2023) Facial emotion recognition in adopted children. Eur Child Adolesc Psychiatry 32: 87–99. https://doi.org/10.1007/S00787-021-01829-Z/TABLES/3
22. Cannon SA, Chatterjee M (2019) Voice Emotion Recognition by Children with Mild-to-Moderate Hearing Loss. Ear Hear 40: 477–492. https://doi.org/10.1097/AUD.0000000000000637
23. Maire J, Galera C, Roux S, Bioulac S, Bouvard M, Michel G (2019) Facial emotion recognition in children with or without Attention Deficit/Hyperactivity Disorder: Impact of comorbidity. Encephale 45: 114–120. https://doi.org/10.1016/J.ENCEP.2018.01.006
24. Liu Z, Liu J, Zhang Z, Yu H, Hu F (2020) Facial Emotion Recognition and Polymorphisms of Dopaminergic Pathway Genes in Children with ASD. Behav Neurol 4: 6376842. https://doi.org/10.1155/2020/6376842
25. Portnova GV, Skorokhodov IV, Mayorova LA (2023) The Levels of Auditory Processing during Emotional Perception in Children with Autism. J Integr Neurosci 22: 112. https://doi.org/10.31083/J.JIN2205112/2279F1232C14CCC07A661F8C52328386
26. Löytömäki J, Laakso ML, Huttunen K (2023) Social-Emotional and Behavioural Difficulties in Children with Neurodevelopmental Disorders: Emotion Perception in Daily Life and in a Formal Assessment Context. J Autism Dev Disord 53: 4744–4758. https://doi.org/10.1007/S10803-022-05768-9/FIGURES/3
27. Li HCW, Lopez V (2005) Children's Emotional Manifestation Scale: development and testing. J Clin Nurs 14: 223–229. https://doi.org/10.1111/J.1365-2702.2004.01031.X
28. Lyakso E, Frolova O, Matveev Y (2020) Facial expression: Psychophysiological study. Handb Res Deep Learn Image Anal Under Constrained Unconstrained Environ 266–289. https://doi.org/10.4018/978-1-7998-6690-9.CH014
29. Jones AC, Gutierrez R, Ludlow AK (2021) Emotion production of facial expressions: A comparison of deaf and hearing children. J Commun Disord 92: 106113. https://doi.org/10.1016/J.JCOMDIS.2021.106113
30. Chaplin TM, Aldao A (2013) Gender differences in emotion expression in children: A meta-analytic review. Psychol Bull 139: 735–765. https://doi.org/10.1037/A0030737
31. Lyakso E, Frolova O (2015) Emotion state manifestation in voice features: Chimpanzees, human infants, children, adults. Lect Notes Comput Sci (includ Subser Lect Notes Artif Intell Lect Notes Bioinformat) 9319: 201–208. https://doi.org/10.1007/978-3-319-23132-7_25
32. Lyakso E, Frolova O, Ruban N, Mekala AM (2021) Child's Emotional Speech Classification by Human Across Two Languages: Russian and Tamil. Lect Notes Comput Sci (includ Subser Lect Notes Artif Intell Lect Notes Bioinformat) 12997 LNAI: 384–396. https://doi.org/10.1007/978-3-030-87802-3_35
33. Lyakso EE, Frolova OV, Nikolaev AS, Grigorev AS (2022) Perceptual Analysis by Adults of the Speech of Children with Autism Spectrum Disorders, Down's Syndrome, and Intellectual Disabilities. Neurosci Behav Physiol 52: 582–590. https://doi.org/10.1007/S11055-022-01279-3
34. Frolova O, Nikolaev A, Grave P, Lyakso E (2023)

Speech Features of Children with Mild Intellectual Disabilities. ACM Int Conf Proceed Ser 406–413. https://doi.org/10.1145/3610661.3616236
35. Deschamps PKH, Been M, Matthys W (2014) Empathy and empathy induced prosocial behavior in 6 and 7 year-olds with autism spectrum disorder. J Autism Dev Disord 44: 1749–1758. https://doi.org/10.1007/s10803-014-2048-3
36. Charman T, Ricketts J, Dockrell JE, Lindsay G, Palikara O (2015) Emotional and behavioural problems in children with language impairments and children with autism spectrum disorders. Int J Lang Commun Disord 50: 84–93. https://doi.org/10.1111/1460-6984.12116
37. Baron-Cohen S, Golan O, Ashwin E (2009) Can emotion recognition be taught to children with autism spectrum conditions? Philos Trans R Soc B 364: 3567–3574. https://doi.org/10.1098/rstb.2009.0191
38. Lyakso E, Frolova O, Kleshnev E, Ruban N, Mary Mekala A, Arulalan KV (2022) Approbation of the Child's Emotional Development Method (CEDM). ACM Int Conf Proceed Ser 201–210. https://doi.org/10.1145/3536220.3563371
39. Lyakso E, Ruban N, Frolova O, Mekala MA (2023) The children's emotional speech recognition by adults: Cross-cultural study on Russian and Tamil language. PLoS One 18: e0272837. https://doi.org/10.1371/JOURNAL.PONE.0272837
40. Castro SL, Lima CF (2010) Recognizing emotions in spoken language: a validated set of Portuguese sentences and pseudosentences for research on emotional prosody. Behav Res Methods 42: 74–81. https://doi.org/10.3758/BRM.42.1.74
41. Gilam G, Hendler T (2016) Deconstructing anger in the human brain. Curr Top Behav Neurosci 30: 257–273. https://doi.org/10.1007/7854_2015_408
42. Glokaya kuzdra.?: languagehat.com. https://languagehat.com/glokaya-kuzdra/
43. Vartanyan IA, Galunov VI, Dmitrieva ES, Morozov VP (1988) Speech perception. Issues of functional brain asymmetry. Nauka, SPb. (In Russ).
44. McHugh ML (2012) Interrater reliability: the kappa statistic. Biochem Med 22: 276. https://doi.org/10.11613/bm.2012.031
45. Rashidah N, Juremi MD, Zulkifley MA, Hussain A, Diyana WM, Zaki W (2017) Inter-rater reliability of actual tagged emotion categories validation using Cohen's Kappa coefficient. J Theor Appl Inf Technol 2017: 31.
46. Laukka P, Thingujam NS, Iraki FK, Elfenbein HA, Rockstuhl T, Chui W, Althoff J (2016) The expression and recognition of emotions in the voice across five nations: A lens model analysis based on acoustic features. J Pers Soc Psychol 111: 686–705. https://doi.org/10.1037/PSPI0000066
47. Liu P, Pell MD (2012) Recognizing vocal emotions in Mandarin Chinese: a validated database of Chinese vocal emotional stimuli. Behav Res Methods 44: 1042–1051. https://doi.org/10.3758/S13428-012-0203-3
48. Juslin PN, Laukka P (2003) Communication of emotions in vocal expression and music performance: different channels, same code? Psychol Bull 129: 770–814. https://doi.org/10.1037/0033-2909.129.5.770
49. Scherer KR (1995) Expression of emotion in voice and music. J Voice 9: 235–248. https://doi.org/10.1016/S0892-1997(05)80231-0
50. Nwokah EE, Davies P, Islam A, Hsu H, Fogel A (1993) Vocal affect in three-year-olds: A quantitative acoustic analysis of child laughter. J Acoust Soc Am 94: 3076–3090. https://doi.org/10.1121/1.407242
51. Scherer KR (2003) Vocal communication of emotion: A review of research paradigms. Speech Commun 40: 227–256. https://doi.org/10.1016/S0167-6393(02)00084-5
52. Ekman P, Levenson RW, Friesen WV (1983) Autonomic Nervous System Activity Distinguishes Among Emotions. Science 221: 1208–1210. https://doi.org/10.1126/SCIENCE.6612338
53. Morozov VP (2013) Emotional hearing. Experimental—psychological research. Psikhol zhurn 34: 45–62. (In Russ).
54. Patel S, Oishi K, Wright A, Sutherland-Foggio H, Saxena S, Sheppard SM, Hillis AE (2018) Right Hemisphere Regions Critical for Expression of Emotion Through Prosody. Front Neurol 9: e224. https://doi.org/10.3389/FNEUR.2018.00224
55. Shapiro BE, Danly M (1985) The role of the right hemisphere in the control of speech prosody in propositional and affective contexts. Brain Lang 25: 19–36. https://doi.org/10.1016/0093-934X(85)90118-X
56. Lindell AK (2006) In your right mind: Right hemisphere contributions to language processing and production. Neuropsychol Rev 16: 131–148. https://doi.org/10.1007/S11065-006-9011-9/METRICS

Translated by A. Dyomina