

Adult-Child Speech Interaction: Speech Database and Psychophysiological Experimental Data



Elena Lyakso¹, Olga Frolova²

¹St. Petersburg University, Russia, lyakso@gmail.com

²St. Petersburg University, Russia, olchel@yandex.ru

ABSTRACT

The paper describes the database “Adult-Child Speech Interaction”: the speech material, recording conditions, equipment for recordings, and additional information about informants. At the present time the database “Adult-Child Speech Interaction” includes: 412 Gb of audio and video records collected from 80 dyads “mother - 4-7-year old children” with typically developing children (n = 40 dyads), children with autism spectrum disorders (n = 20 dyads), children with Down syndrome (n = 20 dyads); 87 Gb of audio and video records of interaction between the experimenter and 10 orphans with light intellectual disabilities and 10 orphans with developmental disorders. The psychophysiological study is aimed at identifying the features of speech interaction in dyads depending on the neurological and psychophysiological state of the child. The practical application of the database that can be used for creating automatic speech and language training systems and socialization skills developing for children with developmental disabilities is discussed.

Key words: adult-child interaction, autism spectrum disorders, Down syndrome, developmental disorders, intellectual disabilities, orphanage, expert analysis, speech database.

1. INTRODUCTION

Modern technologies in the field of speech recognition and synthesis have stepped far forward over the past decades. Robotics and systems aimed at human interaction with automatic devices, in particular gadgets and interfaces, are actively developing. The transition to a new level of development of such systems requires taking into account the social aspect - depending on the features and functionality of the user. There are known computer game and learning systems used in the practice in teaching children at school [1], systems used to develop visual, motor, speech, cognitive skills of children with atypical development [2, 3, 4, 5]. The design of such systems poses a number of additional requirements for implementations, based not only on the age of the child but also on his psychophysiological and neurological state.

The speech databases are the first required condition for creating of any learning speech system. There are a few spontaneous or emotional child speech databases available for child speech research. These include emotional and spontaneous corpora for Mexican Spanish (7-13 year olds) [6], British English (4-14 year olds) [7], German (10-13 year olds) [7, 8], and Russian (3-7 year olds) [9, 10]. Corpora with emotional speech samples are created for tasks of automatic interpretation of emotional behavior. The child speech databases “INFANT.RU” and “Child.Ru” are created on the material of the Russian language [9]. These databases contain spontaneous speech and vocalizations of children from birth to 7 year and find wide application in the study of speech ontogenesis. The speech database “AD-Child.Ru” is currently being created that contains speech materials of 4-16-year old children with atypical development [11, 12]. The CHILDES database including speech material of children growing up in different linguistic environments is used to identify the influence of maternal speech (MR) on the formation of a child’s speech. The authors [13] analyzed the New England longitudinal data from the CHILDES database, comprising transcripts of mother-child conversations at 14, 20, and 32 months. Rates of maternal overlap, but not child overlap, at earlier time-points predicted child language outcomes at later time-points, after controlling for the earlier child mean length of utterance. For early identify children at 3 years of age who could generally be characterized as exhibiting specific language impairment, the National Institute of Child Health and Human Development Study of Early Child Care database of 1364 children was used. It was shown that when classifying a child into the category of speech development disorders, the degree of maternal sensitivity of maternal depression, the ratio of income and family needs, and the quality of the home environment along with indicators of child development were taken into account [14].

Databases of speech interaction between an adult and a child, in particular a mother and a child, an experimenter and a child brought up in an orphanage, are not available for the Russian language. The types of child’s pathology affect the mother’s emotional reactions and her speech [15]. Parents of children with intellectual disabilities adapt to the level of intellectual and linguistic development of their children

facilitating the development of communication skills and attention [16]. Parents of children with Down syndrome (DS) use more emotional speech [17], more directive utterances [18] and ask fewer questions than parents of typically developing (TD) peers. Authors [16] supposed that parents of DS children use speech appropriated to TD children of lower mental age. Mothers of children with autism spectrum disorders (ASD), like mothers of typically developing children, correct their speech depending on the linguistic skills of children. Mothers of ASD children with a high level of verbal skills use more questions and support the speech behavior of the child. Mothers of ASD children with worse verbal skills use more commands and support the motor activity of children [19]. Mothers of ASD children refer to child by name and talk about themselves more often vs. mothers of TD children. Authors considered this finding as the result of ASD children's deficit in engaging in social interactions [20]. The studies of adults' (parents, experimenter) interaction with children of preschool age are few [21]. The child development at maternal deprivation has some features in cognitive, emotional and speech aspects, caused by interaction with the limited number of adults (caregivers and the staff of the institution) [22, 23]. The severity of cognitive impairment in children is associated with the duration of deprivation [24, 25]. The described features of the interaction of parents with children with atypical development and developmental disorders caused the requirement of taking into account the specificity of child development when creating automatic learning systems.

In this regards, the goal of our psychophysiological study is to reveal the correlations of speech strategy of mother's interaction with typically and atypically developing 4-7-year old children. The choice of dyads with ASD and DS children is due to the presence of cognitive and speech disorders, manifested in different ways. ASD are associated with severe impairments in social functioning and language. Multiple symptomatology of disorders, combined into an "autistic triad", includes a violation of social behavior and speech, limited forms of behavior and stereotypes [26]. Children with DS have specificity of 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 [27]. These anatomical peculiarities cause less speech intelligibility and articulation clarity [28, 29] than in TD peers. Social skills of children with DS differ from children with ASD.

2. STATEMENT OF THE PROBLEM

The computer plays an important role in the life of a modern human, both adult and child. Human – computer interaction always has a number of limitations, especially if there are

certain limitations from the user's side. It is necessary to take into account the social factor of interaction when creating educational systems for children with disabilities and atypical development. This factor would contribute to the successful speech and social skills learning of the child. Thus, to create interfaces for interaction "child with atypical development-computer", it is necessary 1) to create a speech database of interaction between an adult and a child and 2) to analyze the speech interaction in the pair "adult-child" to identify the features of interaction depending on the age and psychophysiological and neurological state of the child.

3. SPEECH DATABASE

3.1 Data collection

The database "Adult-Child Speech Interaction" contains speech material of adults (mothers and experimenters) and children with typical development (TD) and atypical development: children with ASD, children with DS, orphans with intellectual disabilities (ID) and developmental disorders (DD) aged 4-7 years.

All children and their parents were born and living in St Petersburg. For all children, the data on development have been obtained from the moment of birth, for children with ASD additionally - scores on questionnaires, Child Autism Rating Scale (CARS) score [30], diagnoses; for all mothers - information about age, education, presence / absence of diseases, family. Places of audio and video recording were at home, in the laboratory, and in the orphanage. The situation of spontaneous interaction between children and mother, children and the researcher (for orphans), were used: dialogue with the researcher or parents (the child answered to the set of questions), play with the set of toys, picture description, and retelling of the story (monologue).

The total duration of recording was 20-40 minutes each, the duration of model situation was 5-15 min. The recording for every dyad was made from one to ten times. The records were made by the digital recorder "Marantz PMD660" with external microphone "SENNHEIZER e835S" and video camera "SONY HDR-CX560E". Speech files were stored in Windows PCM format WAV, 44100 Hz, 16 bits per sample; video files were in AVI format. Every record is accompanied by a detailed protocol and video recording of the mother's and child's behavior in parallel.

At the present time the database "Adult-Child Speech Interaction" includes: 412 Gb of audio and video records collected from 80 dyads "mother – 4-7-year old children" with TD children (40 dyads), with ASD children (20 dyads), children with DS (20 dyads); 87 Gb of audio and video records of interaction between the experimenter and 10 orphans with light ID and 10 orphans with DD. These

diagnoses (ID and DD) are most common in the orphanage among children without genetic syndromes and severe neurological disorders.

3.2 Database Structure

The “Adult-Child Speech Interaction” database is an organizational tool that provides streamlined access to storing and searching the mother-child’s audio recordings and the documentation of the audio recording. The database consists of aggregated binary files managed by Microsoft SQL Server. Data can be added, modified, and deleted by the user only through the application interface.

Database Management Access Application: The interface of the application contains two blocks. The interface of block-1 includes audio records of mother-child speech in the situation of interaction in accordance with the leading diagnosis of the child (ASD, DS) or TD. Subdirectories of the main catalog are data for each child, which include age (4, 5, 6, 7 years), record situation. For each dyad, there is information about the child’s name, gender, date of birth, birth number in the family, place of birth, diagnosis, and information about parents: families, mother’s age, education. For each record, there is information about the place of recording and the equipment used. The interface of block-2 includes video recordings of mother-child interaction.

The application contains search fields and filters that allow the user to extract mother-child speech interaction data based on the type of the child disease, age of the child, and the recording situation. Additional filters make it possible to select data based on all available metadata: the child’s family number, information about family/orphanage, equipment used during recordings, etc. The application is written in C # language.

All procedures were approved by the Health and Human Research Ethics Committee and written informed consent was obtained from parents (official curator for orphans) of the child participant.

4. EXPERIMENTAL RESULTS

4.1 Experiment 1: Speech Interaction in “Mother-Child Dyads” with TD children, Children with ASD and DS

The design of the study included: expert analysis of audio fragments of mothers’ and children’s speech during the interaction; analysis of nonverbal behavior of the mother and the child in the process of interaction; perceptual, acoustic, statistical analysis of MS directed to children; perceptual analysis of child’s words.

The participants were “mother-child” dyads with 4-7-year old TD children (40), children with ASD (n = 20), DS (n = 10). Expert analysis (n = 10 adults, aged 31.3 ± 9.66 years) of

audio fragments of speech interaction in “mother-child” dyads was conducted. The duration of the fragments (from 4 to 30 fragments for each dyad) was individual for each dyad. The listening time was 27 hours. The criterion for choosing fragments was verbal and/or voice attraction of attention by the mother or the child of one's partner. One expert made a selection of fragments from an audio recording. Each fragment was listened by 10 experts. 21 elements of maternal speech behavior and 19 indicators of child speech level were estimated.

Expert analysis of nonverbal behavior in “mother-child” dyads on the base of video records: Video tests (n = 31) containing fragments of the interaction in the dyad were created by two experts. The duration of each fragment was one min. Expert analysis was performed by 5 experts (aged 33.2 ± 6.8 years) with professional experience of working with children. The experts watched video fragments without audio.

Perception and acoustic features of MS addressed to children: The aim of the perceptual study is the review of listeners’ recognition of the function of MS and mother emotional state. The test sequences included: MS directed to children (n = 3 tests, for 10 samples of ASD test, 10 samples of DS test, and 10 samples of TD test). Perceptual analysis of the child speech: 3 test sequences were created. The test sequences included the words of TD, ASD, DS children (for 30 words of ASD test, 30 words of DS test, 30 words of TD test). The listeners had to recognize the meaning of the child’s words via speech. The listeners were Russian-speaking adults (n = 283, 26.7 ± 16.6 y). The factor of the adult’s experience of interaction with children (at the household level) was not significant, so all the data were presented together.

Spectrographic analysis of the speech was carried out in the “Cool Edit Pro” sound editor. Pitch and duration of words and vowels were automatically calculated, based on the algorithms implemented in “Cool Edit Pro” sound editor. The waveform view was used to calculate duration of phrase, words, stressed and unstressed vowels in the word, and the spectral view was used to measure the pitch and formants. We analyzed and compared pitch values, max and min values of pitch, pitch range, formant frequencies (F1, F2), duration of vowels and the stationary part of vowels from mother’s words addressed to TD, ASD, and DS children. Vowel formant triangle areas [31] and vowel articulation index [32] were calculated.

Speech interaction in “mother-child” dyads: expert analysis of the audio fragments of speech samples: It is shown that mothers of TD children and children with ASD and DS use different strategies of speech interaction with the child. Progress in speech development of TD children correlates with MS features: the mother speaks clearly $F(19,270) =$

234.42, $p < 0.0001$, $R^2 = 0.943$ - the child speaks clearly (Beta = 0.150) and uses phrases (Beta = 0.3297); the mother encourages the child $F(19.270) = 16.667$, $p < 0.0001$, $R^2 = 0.540$ and asks questions $F(19.270) = 208.41$, $p < 0.0001$, $R^2 = 0.936$ - the child answers the mother's questions (Beta = 0.584), asks questions (Beta = 0.160) and uses phrases (Beta = 0.154); the mother makes pauses between phrases $F(19.270) = 40.477$, $p < 0.0001$, $R^2 = 0.740$ - the child uses phrases (Beta = 0.276) and asks questions (Beta = 0.100) - Multiple Regression analysis.

The repetition of the words spoken by TD children by the mothers $F(19.270) = 35.568$, $p < 0.0001$, $R^2 = 0.715$; the highlighting of certain words $F(19.270) = 26.909$, $p < 0.001$, $R^2 = 0.654$, the stretching of sounds in words $F(19.270) = 5.625$, $p < 0.001$, $R^2 = 0.233$ are due to the low level of the speech of TD children - speech is indistinct (Beta = 0.356), the replies "yes-no" (Beta = 0.127), the child does not answer the mother's questions and does not ask questions.

The initiative manifested by the mother in interacting with the ASD child correlates with the following characteristics of her speech: emotional (0.56 - Spearman correlation $p < 0.05$), speaks loudly (0.78), clearly (0.94), addresses to the child (0.91), refers by name (0.82), asks questions (0.71), encourages the child (0.55), contains instructions (0.76), repeats the question or the same words (0.87), repeats the words for the child (0.62), specifies what the child said (0.60), grammatically simplifies (0.91), singles out individual words (0.70), stretches sounds in words (0.58), pauses between phrases (0.60). These features of MS correlate $F(33.96) = 34.131$, $p < 0.0001$, $R^2 = 0.921$ with the characteristics of the ASD child's speech: speech is emotional (Beta = -0.157, $p < 0.009$), the children answer the mother's questions (Beta = 0.390, $p < 0.003$), respond to the mother's reply (Beta = 0.390, $p < 0.003$), use a replica of the word (Beta = 0.220, $p < 0.03$), a "yes-no" replica (Beta = 0.209). MS addressed to ASD children is characterized by the repetition of children's words, clarification of child's utterances, simplification of speech, emphasis of words by the voice, stretching of sounds in words, presence of the long pauses between phrases. However, these features of MS do not always lead to progress in the development of the ASD child's speech.

Direct correlations between the characteristics of MS and the features of DS child speech behavior are shown: if the mother speaks loudly, the child uses a loud voice: $F(1.58) = 10.101$, $p < 0.002$ ($R^2 = 0.148$, Beta = 0.385); mother's emotional speech correlates with emotional child speech $F(1.58) = 0.425$, $p < 0.02$ ($R^2 = 0.086$, Beta = 0.292); if the mother demonstrates joy, the child is happy $F(1.58) = 34.511$, $p < 0.0001$ ($R^2 = 0.373$, Beta = 0.610) - Regression analysis. If the mother is angry, the child is also angry $F(3.53) = 518$, $p < 0.0001$ ($R^2 = 0.593$, Beta = 0.660). However, a clear

articulation of the MS words is correlated both with $F(3.56) = 45.607$, $p < 0.001$, $R^2 = 0.710$ clear pronunciation of the words by the child (Beta = 0.653) and with fuzzy articulation of child (Beta = 0.769).

Elements of nonverbal behavior of mothers and children in the process of interaction: expert analysis: The mothers of TD children, children with ASD, and children with DS often looked at the child (94% of the answers of experts for the mothers of TD children, 99% - for mothers of ASD and DS children), rarely showed discontent (19% - DS, 7% - TD, and 3% - ASD) (Figure 1 A).

The mothers of all groups of children attract the attention of the child with eyes (44% - TD, 43% - ASD, 40% - DS), use gestures (56%, 60%, 56%). Mothers of children with ASD and with DS are more likely to touch children in the process of interaction (93% - ASD, 80% - DS) compared to mothers of TD children (30%), attract the child's attention more through gestures (66% - ASD, 57% - DS) than mothers of TD children (41%). Mothers of DS children are more often dissatisfied with their children (19%) than mothers of ASD children (3%) and TD children (7%); they smile less (36%) than mothers of ASD children (50%) and TD children (51%); less often look around (14%) vs. mothers of ASD children (51%) and TD children (48%).

Nonverbal behavior of ASD children and children with DS differs from the behavior of TD children by a large number of behavioral elements: TD children smile more often (56%) than ASD children (19%) and DS children (33%), others show discontent (1% - TD children, 19% - ASD children, 47% - children with DS), more often look at the mother (71%), use contact "eye-to-eye" (56%), look at the object (85%). Children with ASD are more likely to look around (91%) than children with DS (76%) and TD children (62%); touch mothers more often (51%) than children with DS (44%) and TD children (16%) (Figure 1 B).

Perception and acoustic features of MS addressed to the child: MS directed to TD children and ASD children (38% and 37.2%) is aimed at attracting the attention of the child more than MS addressed to children with DS.

Mothers of DS children more comment child's utterances (20.9%) than mothers of ASD children (12.7%) and TD children (16%). MS addressed to ASD and DS more stimulates children to verbal response (43.8% and 37.6% ASD and DS, respectively) than MS addressed to TD children (32%) with already well-formed speech. MS addressed to TD children and children with ASD contains equally few encouragements for the child (5% and 4.3%), mothers of DS children more often encourage their children (8.3%).

Determining the emotional state of the mothers, the listeners noted a calm state for mothers of TD, ASD, and DS children, but the number of listener’s answers was minimal for mothers of DS children (32.2% for DS, 64% - TD, 52.2% - ASD). According to the responses of listeners, the mothers of ASD children manifested the state of joy more frequently (19.7%) vs. mothers of DS children (17.9%) and TD children (15%). The listeners noted the state of angry (23.5% - DS, 10.5% -

ASD, 6% - TD), sadness (10.2% - DS, 7.8% - ASD, 7% - TD), and aggressive for mothers of DS children (10.1%) more often vs. mothers of ASD (7.2%) and TD children (2%). The mother’s utterances addressed to ASD children are longer ($p < 0.001$ - the Mann-Whitney test) than the utterances directed to TD children and children with DS, contain shorter phrases ($p < 0.005$).

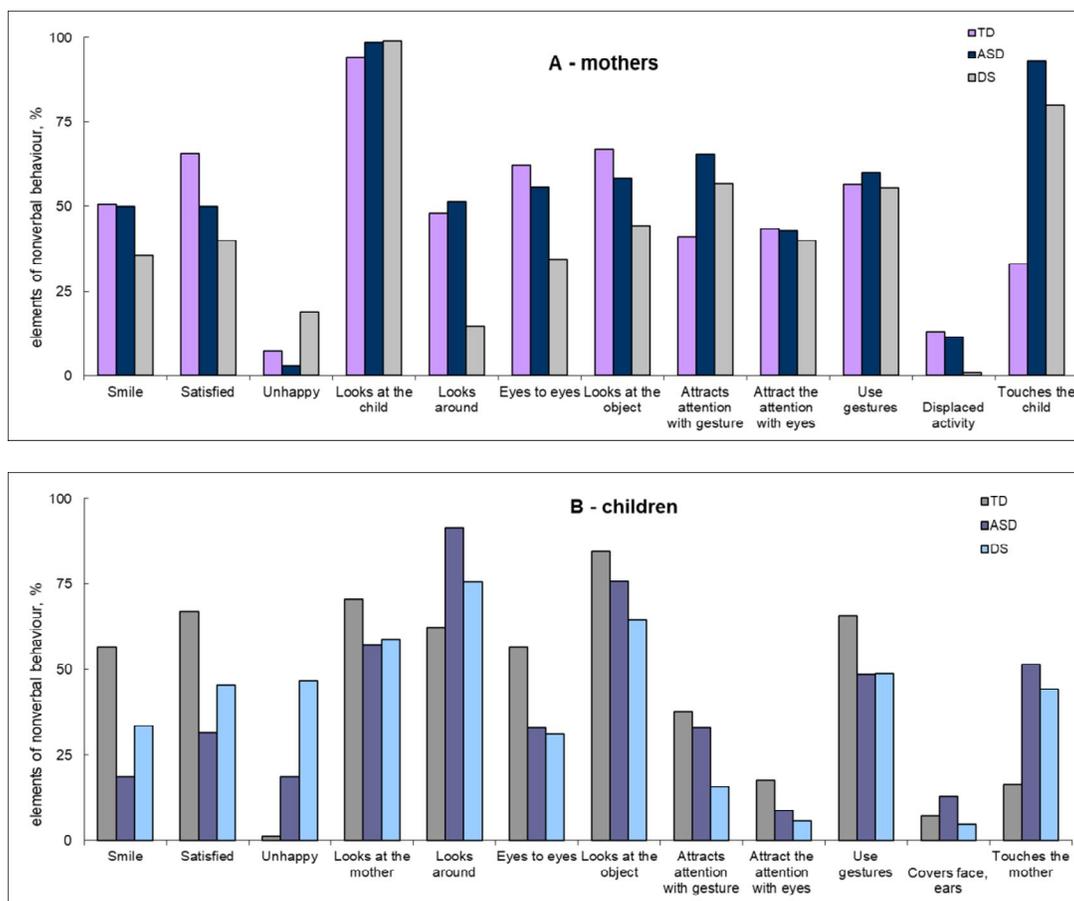


Figure 1: Elements of non-verbal behavior in “mother-child” dyads: expert analysis. A – data for mothers’ non-verbal behavior; B – similarly for children. Vertical axis – elements of non-verbal behavior, %

In MS addressed to children with DS, the duration of pauses between phrases in utterances is longer vs. the MS directed to TD children ($p < 0.005$). The pauses between phrases ($p < 0.005$) and between words ($p < 0.005$) are longer in MS addressed to ASD vs. TD children.

Pitch values are significantly higher ($p < 0.001$ - Mann-Whitney test) in MS addressed to ASD vs. MS addressed to TD children and children with DS in utterances, phrases, words, and stressed vowels. The pitch values of voice of ASD child’s mothers are higher if the child is older $F(1.74) = 4.531, p < 0.03 (R^2 = 0.057, Beta = 0.240)$ – Regression analysis. Mother’s voice addressed to ASD boys is higher than directed to ASD girls $F(1.74) = 4.118, p < 0.04 (R^2 = 0.054, Beta = 0.231)$. The lower voice of ASD child’s mothers

(speech is less emotional) is, the higher the child’s CARS score is $F(1.74) = 37.397, p < 0.0001 (R^2 = 0.336, Beta = -0.579)$. Discriminant analysis showed the difference in pitch values of MS as a function of the TD child’s gender $F(6.47) = 2.879, p < 0.01$ Wilks' Lambda - 0.731 Wilks' = 0.800. Mothers’ communication with TD boys is more emotional (pitch values are higher) than with the girls $F(6.47) = 2.872, p < 0.01 (R^2 = 0.268, Beta = -1.815)$. For MS addressed to DS children, the correlation $F(1.21) = 5.790, p < 0.002 (R^2 = 0.216, Beta = -0.465)$ – Regression analysis) between the child’s age and pitch values of mother’s phrases was revealed. The mother’s voice is higher if the child is younger. The DS children’s age and gender of are not correlated with speech features of their mothers.

Maximal values of stressed vowel formant triangle areas (341498 conv. units) and minimal values of unstressed vowel formant triangle areas (135067.8 conv. units) were revealed in MS addressed to children with DS. The values of stressed vowel formant triangle areas in MS addressed to children with ASD (180670 conv. units) are higher vs. values in MS addressed to TD children (177800.5 conv. units). The values of VAI for stressed vowels in words from MS directed to children with DS (1.16) are higher than values of VAI for stressed vowels for MS directed to children with ASD (0.95) and TD (0.93). Data on the values of the vowel formant triangle areas and VAI indicate more clear articulation in MS addressed to children with speech disorders – ASD and DS vs. TD children. The articulation of mothers of DS children is clearer vs. mothers of TD and ASD children.

Perception data of child speech: Listeners recognized correctly 65.8% words of TD children, 44.8% words of ASD children and 17.1 % words of DS children in tests. The number of words unrecognized by listeners is significantly more ($p < 0.001$) for DS children vs. ASD and TD children ($p < 0.005$), ASD children vs. TD children.

4.2 Experiment 2: Speech Interaction in Pairs “Experimenter-Orphan” with DD and ID

The design of the study (experiment 2) was similar to experiment 1: expert analysis of audio fragments of experimenter’s speech and children’s speech during the interaction; analysis of nonverbal behavior of the child in the process of interaction; perceptual analysis of the child’s words. The behavior of the experimenter was standardized during interaction with all orphans: the set of the same questions was used in the dialogue, and the similar toys and pictures were provided.

Expert analysis ($n = 5$ adults, aged 27 ± 5.43 years) of audio fragments of speech interaction in “experimenter-child” pairs was conducted (100 fragments). The criterion for choosing fragments was a verbal attraction of attention of the child by the experimenter. One expert made a selection of fragments from an audio recording (similar to study 1). The same experts made expert analysis of the orphan’s nonverbal behavior in “experimenter-child” pairs on the base of video records. Video tests ($n = 20$) containing fragments of the interaction in the pairs were created by two experts. Perceptual analysis of the child’s speech: the test sequences included the words of DD and ID orphans (30 words). The listeners had to recognize the meaning of the child’s words via speech. The listeners were Russian-speaking adults ($n = 182$, age 19 ± 2.3 y).

Speech interaction in “experimenter-child” pairs: The speech of the experimenter was standardized, but there are some correlations between adult’s and child’s speech

features: the adult asks questions – the child answers the questions $F(1.48) = 22.972$, $p < 0.0001$ ($R^2 = 0.324$, Beta = 0.569) – Regression analysis for DD children; $F(1.48) = 47.284$, $p < 0.0001$ ($R^2 = 0.496$, Beta = 0.704) – for ID children. The number of DD children’s answers increased if the adult highlighted the certain words $F(1.48) = 6.654$, $p < 0.01$ ($R^2 = 0.122$, Beta = 0.349), stretched sounds in the words $F(1.48) = 9.84$, $p < 0.01$ ($R^2 = 0.17$, Beta = 0.12); the number of ID children’s answers increased if the adult used the emotional speech $F(1.48) = 6.49$, $p < 0.01$ ($R^2 = 0.119$, Beta = 0.345). The number of replies presented by phrases in DD children increased if the speech of the experimenter contained the instructions $F(1.48) = 13.141$, $p < 0.001$ ($R^2 = 0.215$, Beta = 0.464). For children with ID, the correlation between grammatical complexity and (the replies are presented by words or phrases) and the adult’s speech features was revealed. The number of yes/no replies increased in children with ID and DD when the number of adult’s questions increased $F(1.48) = 4.785$, $p < 0.03$ ($R^2 = 0.091$, Beta = 0.301) – Regression analysis for DD children, $F(1.48) = 10.06$, $p < 0.01$ ($R^2 = 0.173$, Beta = 0.416) – for ID.

Elements of nonverbal behavior of mothers and children in the process of interaction: expert analysis:

Orphans with DD and ID smile when interacted with the experimenter (60% of DD children and 68% of ID children), were satisfied (64% and 60% - DD and ID), rarely demonstrated discomfort. Children with ID looked at the adult more often than children with DD (64% and 75% - DD and ID), but they rarely looked at the object (87% - ID and 63% - DD). Children with ID more often compared with DD children closed the face and ears by hands – demonstrated the displacement activity (4% - DD and 10% - ID), touched the adult (2% - DD and 12,5% - ID) (Figure 2). In general, the nonverbal behavior of orphans (especially children with DD) during interaction with the experimenter is more similar to the behavior of typically developing children when interacting with their mothers than children with developmental disorders (ASD and DS) interacting with mothers.

Perception data of child speech: Listeners recognized correctly 39% of words of ID children and 43% of words of DD children from the orphanage.

4.3 Conclusion and Discussion

Differences in speech behavior strategies of mothers during interaction with TD children, children with ASD, and children with DS are revealed. The strategies of the mother’s speech behavior leading to the progress of the speech development of TD children are defined. It is shown that mothers of children with ASD adapt their speech to the level of the child’s speech development and are guided by the degree of autistic disorders of the child, but this does not always lead to progress in the child’s speech development. The speech of mothers of DS children is characterized by a

clearer articulation of stressed vowels in words and longer duration of pauses between phrases in the utterance than the speech of mothers of TD children. Our data, in general, correspond to the results obtained on Italian mother-child dyads with the child's developmental age of two years and the different chronological age [20]. Comparing "mother-child" dyads with older children, we revealed the differences between speech strategies between ASD and DS child's mothers. Speech behavior strategies of mothers of children with DS are more connected with the emotional manifestation in the child's behavior. This study highlights the strengths and weaknesses of parental communication with children

with Down syndrome and helps to identify areas of potential improvement through intervention [17]. However, it was shown that properly organized MS is not enough to correct the defects that cause the specificity of child development.

On the base of perceptual analysis, the differences in recognition of children's words by listeners were revealed – the words of children with DS are recognized worst, the words of TD children are recognized with maximal probability. The words of children with ASD and orphans are recognized with medium probability.

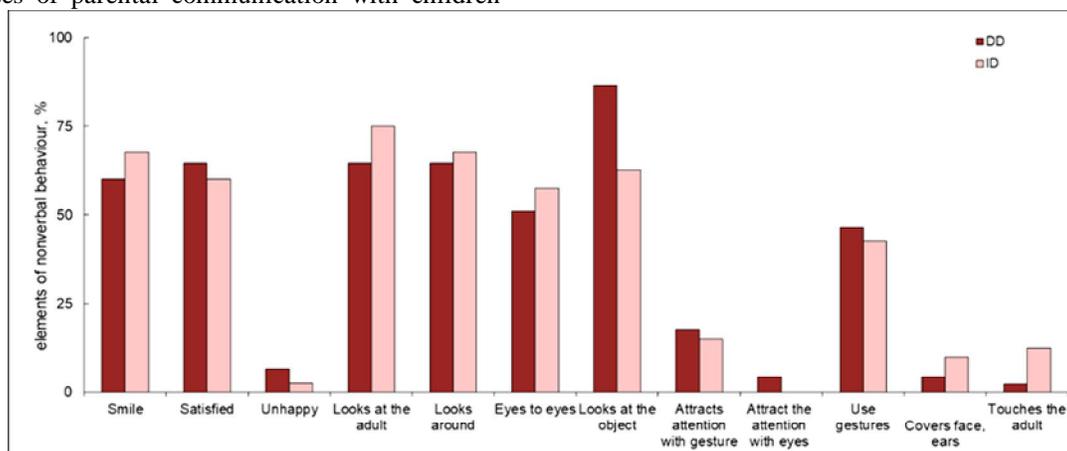


Figure 2: Elements of non-verbal child's behavior in "experimenter-orphan" pairs: expert analysis. Vertical axis – elements of non-verbal behavior, %

Our results show a lack of input leading to violations in speech acquisition in orphans that corresponds to the data of other studies [23]. The strategies of mothers' speech behavior during interaction with ASD and DS children described in the work are useful for working with parents of children with atypical development. The features of mothers' behavior associated with a high level of speech development of the child can be used for training the staff working with children of preschool age with atypical development.

5. CONCLUSION

Creating the database "Adult-Child Speech Interaction" is the main result of our work. The database structure is traditional for speech bases [11, 12]. However, the specificity of this work is the psychophysiological research, which allowed identifying the features of interaction between mother and child in dyads with typically and atypically developing children of preschool age. The presented examples of using the database material to determine the specificity of mother-child speech interaction make it possible to widely use the created base in psychophysiological research, at the research of automated child speech and emotions recognition systems. At the present time the machine emotions classification systems for different languages are developed for sound speech [10] and texts [34], [35]. Database "Adult-Child Speech Interaction" could be considered as the

basement for such applied studies and multidisciplinary fundamental researchers.

REFERENCES

1. R. D. Hawkins, G. A. R. M. Ferreira, and J. M. Williams. **The development and evaluation of 'Farm Animal Welfare': an educational computer game for children**, *Animals*, Vol. 9, no 3, pii: E91, March 2019. <https://doi.org/10.3390/ani9030091>
2. H. C. Lin, Y. H. Chiu, Y. J. Chen, Y. P. Wuang, C. P. Chen, C. C. Wang, C. L. Huang, T. M. Wu, and W. H. Ho. **Continued use of an interactive computer game-based visual perception learning system in children with developmental delay**, *Int J Med Inform*, Vol. 107, pp. 76-87, Nov. 2017. <https://doi.org/10.1016/j.ijmedinf.2017.09.003>
3. M. Li, X. Li, L. Xie, J. Liu, F. Wang, and Z. Wang. **Assisted therapeutic system based on reinforcement learning for children with autism**, *Comput Assist Surg*, sup. 2, pp. 94-104, Oct. 2019. <https://doi.org/10.1080/24699322.2019.1649072>
4. J. DiPietro, A. Kelemen, Y. Liang, and C. Sik-Lanyi. **Computer- and robot-assisted therapies to aid social and intellectual functioning of children with autism spectrum disorder**, *Medicina*, Vol. 55, no. 8, pii: E440, Aug. 2019. <https://doi.org/10.3390/medicina55080440>

5. A. A. A. Dzulkarnain, S. Rahmat, A. W. Ismail, R. Musa, M. Badzis, and T. Z. H. Tengku Zam Zam. **A novel computer-based simulated learning environment in audiology with learning assistance: preliminary findings**, *Med J Malaysia*, Vol. 74, no. 2, pp. 168-173, Apr. 2019.
6. H. Perez-Espinosa, C. A. Reyes-Garcia, and L. Villasenor Pineda. **"EmoWisconsin: an emotional children speech database in Mexican Spanish,"** *Affective Computing and Intelligent Interaction, LNCS*, Vol. 6975, pp.62–71, Oct. 2011.
https://doi.org/10.1007/978-3-642-24571-8_7
7. A. Batliner, M. Blomberg, S. D'Arcy, D. Elenius, D. Giuliani, M. Gerosa, C. Hacker, M. J. Russell, S. Steidl, and M. Wong. **"The PF_STAR children's speech corpus,"** *Proc. Interspeech*, Lisbon, Sep. 2005, pp. 2761–2764.
8. A. Batliner, S. Steidl, and E. Noth. **"Releasing a thoroughly annotated and processed spontaneous emotional database: the FAU Aibo Emotion Corpus,"** in *Proc. Satellite Workshop of LREC-2008 on Corpora for Research on Emotion and Affect*, Marrakeh, 2008, pp. 28–31.
9. E. Lyakso, O. Frolova, A. Kurazhova, and J. Gaikova. **Russian infants and children's sounds and speech corpora for language acquisition studies**, in *Proc. INTERSPEECH-2010*. Makuhari, 2010, pp. 1981–1888.
10. E. Lyakso, O. Frolova, E. Dmitrieva, A. Grigorev, H. Kaya, A. A. Salah, and A. Karpov. **"EmoChildRu: emotional child Russian speech corpus,"** *LNCS*, Vol. 9319, pp. 144–152, Sep. 2015.
https://doi.org/10.1007/978-3-319-23132-7_18
11. E. Lyakso, O. Frolova, and A. Karpov. **A new method for collection and annotation of speech data of atypically developing children**, in *Proc. of 2018 International IEEE Conference on Sensor Networks and Signal Processing (SNSP 2018)*, Xi'an, 2018, pp. 175-180.
<https://doi.org/10.1109/SNSP.2018.00040>
12. E. Lyakso, O. Frolova, A. Kaliyev, V. Gorodnyi, A. Grigorev, and Yu. Matveev. **AD-Child.Ru: speech corpus for Russian children with atypical development**, *LNAI*, Vol. 11658, pp. 299-308. August 2019.
https://doi.org/10.1007/978-3-030-26061-3_31
13. E. S. Che, P. Brooks, M. Alarcon, F. Yannaco, S. Donnelly. **Assessing the impact of conversational overlap in content on child language growth**, *J Child Lang*, Vol. 45, no. 1, pp. 72-96, Jan. 2018.
<https://doi.org/10.1017/S0305000917000083>
14. K. M. La Paro, L. Justice, L. E. Skibbe, and R. C. Pianta. **Relations among maternal, child, and demographic factors and the persistence of preschool language impairment**, *Am J Speech Lang Pathol*, Vol. 13, no. 4, pp. 291-303, Nov. 2004.
[https://doi.org/10.1044/1058-0360\(2004/030\)](https://doi.org/10.1044/1058-0360(2004/030))
15. C. Cardoso-Martins and C. B. Mervis. **Maternal speech to prelinguistic children with Down syndrome**, *Am J Ment Defic*, Vol. 89, no. 5, pp. 451-458, Mar. 1985.
16. M. Legerstee and T. Fisher. **Coordinated attention, declarative and imperative pointing in infants with and without Down syndrome: sharing experiences with adults and peers**, *First Language*, Vol. 28, no. 3, pp. 281-311, Aug. 2008.
17. S. de Falco, P. Venuti, G. Esposito, and M. H. Bornstein. **Maternal and paternal pragmatic speech directed to young children with Down syndrome and typical development**, *Infant Behav Dev*, Vol. 34, no. 1, pp. 161–169, Feb. 2011.
<https://doi.org/10.1016/j.infbeh.2010.12.002>
18. M. Legerstee, J. Varghese, and Y. van Beek. **Effects of maintaining and redirecting infant attention on the production of referential communication in infants with and without Down syndrome**, *J Child Lang*, 2002, Vol. 29, no. 1, pp. 23-48, Feb. 2002.
19. M. Konstantareas, H. Zajdeman, S. Homatidis, and A. McCabe. **Maternal speech to verbal and higher functioning versus nonverbal and lower functioning autistic children**, *J Autism Dev Disord*, Vol. 18, no. 4, pp. 647-656, Jan. 1989.
<https://doi.org/10.1007/BF02211882>
20. P. Venuti, S. de Falco, G. Esposito, M. Zaninelli, and M. H. Bornstein. **Maternal functional speech to children: a comparison of autism spectrum disorder, Down syndrome, and typical development**, *Res Dev Disabil*, Vol. 33, no. 2, pp. 506-517, Nov. 2012.
21. E. Lyakso and O. Frolova. **Speech interaction in "mother-child" dyads with 4–7 years old typically developing children and children with autism spectrum disorders**, *LNAI*, Vol. 11096, pp. 347-356, Sep. 2018.
22. O. Frolova and E. Lyakso. **Emotional speech of 3-years old children: norm-risk-deprivation**, *LNAI*, Vol. 9811, pp. 262-270, Aug. 2016.
https://doi.org/10.1007/978-3-319-43958-7_31
23. M. J. Bakermans-Kranenburg, H. Steele, Ch. H. Zeanah, R. J. Muhamedrahimov, P. Vorria, N. A. Dobrova-Krol, M. Steele, M. H. I. Jzendoorn, F. Juffer, and M. R. Gunnar. **Attachment and emotional development in institutional care: characteristics and catch up**. *Monographs of the Society for Research in Child Development*, Vol. 76, no. 4, pp. 62-91, Dec. 2011.
24. M. Rutter. **Developmental catch-up, and deficit, following adoption after severe global early privation**, *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, Vol. 39, no. 4, pp. 465-476, May 1998.
<https://doi.org/10.1111/1469-7610.00343>
25. J. Bick, N. A. Fox, C. Zeanah and C. A. Nelson. **Early deprivation, atypical brain development, and internalizing symptoms in late childhood**, *Neuroscience*, Vol. 342, pp. 140-153, Feb. 2017

26. L. Kanner. **Autistic disturbances of affective contact**, *Nervous Child*, Vol. 2, pp. 217–250, 1943.
27. K. Polišenská and S. Kapalková. **Language profiles in children with Down syndrome and children with language impairment: implications for early intervention**, *Res Dev Disabil*, Vol. 35, no. 2, pp. 373-382, Feb. 2014.
<https://doi.org/10.1016/j.ridd.2013.11.022>
28. K. Bunton and M. Leddy. **An evaluation of articulatory working space area in vowel production of adults with Down syndrome**, *Clin Linguist Phon.*, Vol. 25, no. 4, pp. 321–334, Apr. 2011.
29. R. D. Kent and H. K. Vorperian. **Speech impairment in Down syndrome: a review**, *J Speech Lang Hear Res*, Vol. 56, no. 1, pp. 178–210, Feb. 2013.
30. E. Schopler, R. J. Reichler, R. F. DeVellis, K. Daly. **Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS)**, *J Autism Dev Disord*, Vol. 10, no. 1, pp. 91-103, Mar. 1980.
<https://doi.org/10.1007/BF02408436>
31. E. E. Lyakso and A. S. Grigor'ev. **Dynamics of the duration and frequency characteristics of vowels during the first seven years of life in children**, *Neurosci Behav Physiol*, Vol. 45, no. 5, pp. 558-567, June 2015.
32. N. Roy, S. L. Nissen, C. Dromey, and S. Sapir. **Articulatory changes in muscle tension dysphonia: evidence of vowel space expansion following manual circumlaryngeal therapy**, *J Commun Disord*, Vol. 42, no. 2, pp. 124-135, Dec. 2009.
<https://doi.org/10.1016/j.jcomdis.2008.10.001>
33. E. Lyakso, A. Grigorev, A. Kurazova, and E. Ogorodnikova. **"INFANT.MAVS" - multimedia model for infants cognitive and emotional development study**, *LNAI*, Vol. 8773, pp. 284-291, Oct. 2014.
34. P. Thakur, R. Shrivastava. **A Review on Text Based Emotion Recognition System**, *International Journal of Advanced Trends in Computer Science and Engineering*, Vol. 7, no. 5, pp. 67-71, Sept.-Oct. 2018.
<https://doi.org/10.30534/ijatcse/2018/01752018>
35. I. S. Makki, F. Alqurashi. **An Adaptive Model for Knowledge Mining in Databases “EMO_MINE” for Tweets Emotions Classification**, *International Journal of Advanced Trends in Computer Science and Engineering*, Vol. 7, no. 3, pp. 52-60, May-June 2018.
<https://doi.org/10.30534/ijatcse/2018/04732018>