

**E.E.Lyakso, M.A.Bogorad, U.S.Gaikova, A.D. Gromova,  
A.V. Kurazhova, A.V. Ostrouhov, O.V. Frolova  
“CHILDRU”: Speech database of 4-6 years old children**

Institute of Physiology of A.A. Uchtomsky, Saint-Petersburg State University  
Russia, Saint-Petersburg, Universitetskaya emb., 7/9  
Phone (812) 321-3361; Fax (812) 323-2454; E-mail: [Lyakso@mail.ru](mailto:Lyakso@mail.ru)

*The pilot database “CHILDRU” contains the records of children speech at the age from 4 to 6 years. It continues database “INFANTRU” that contains longitudinal recordings of children of first three years of life. Recording of children’s speech material was conducted in situations of interaction with adult: spontaneous speech, answers to questions, reading, poetry or retelling a tale, count and alphabet, play. The information about children, about mothers, recording conditions and equipment is in the database. Speech material is presented by the following way: There are original files by duration up to 5 minute, that demonstrate the situation: spontaneous speech, answers to questions, reading, poetry or retelling a tale, count and alphabet, play. Phrases; dialogues of child with adult; questions; sounds and syllables, pronounced separately; words, that contain one, two, three, four, five and more syllables; reading the words and phrases; the mistakes that caused by: different variants of changes, gaps and transpositions of phoneme /r/ in word; changes, gaps and transpositions of other phonemes and syllables in the word; mistakes in the phrase were chosen from original file. The phonetic transcription in symbols of IPA and child meaning of the words that contains mistakes is presented. Speech files format is Windows PCM, 22050 Hz, 16 bit. VDB.EXE program was created for the work with the base. The program allows choosing speech material of all children or of every child depending on age and the situation.*

Currently, there are limited amount of child speech databases. It is caused by small amount of investigation of child’s speech signal acoustics. Consequently there are some problems when using the existing technologies for the recognition of child speech. So generally accepted approaches to format of early child (from the first year of life) vocalizations databases and to the materials contained in these databases is absent. In particular, the corpus of child speech (children of the school age) that contains the reading material of long-lasting phrases and separate words, but without the spontaneous speech was the part of speech database SPEECON, made by Russian Acoustic Company “ODITEK”

Nowadays speech databases are integral part of speech investigations. Their structure and technical characteristics depend on the investigation tasks. Database size (the amount of speakers), speech material: reading speech – pre-defined word-commands, sentences for phonemic representatively; induced, spontaneous speech; communication channel – stationary phone communication, mobile communication, broadband channel etc – are defined by the particular aim of use [1]. The modern speech databases are created mainly for the purposes of automatic speech recognition (for instance: SpeechDat-I, SpeechDat-II, SpeechDat-E, SpeechDat-Car, Speecon) and speaker verification [2]. They contains of adult speakers recording. Existing generally accepted technologies of speech recognition (Hidden Markov Models) define the databases format. The great amount of different speakers (several thousand), the presenting of all phonemes (monophones) of the language, valid presenting of diphones and threephones is necessary. Such characteristics as utterance prosodic organization, dialogues structure are not considered, as a rule.

The database “INFANTRU” of speech and sound signals of children from 3 to 36 months of age was created by the group of employees [3]. Analogues of similar bases do not exist, as we know. Directivity of base «INFANTRU» is accompaniment and using of its materials when conducting scientific investigation in the field of speech ontogenesis and the practice use for making the educating programs. Speech bases, that contain speech material of children (early pre-school age) from 4 to 6 years old are absent for Russian language.

Presented version of the pilot database «CHILDRU» contains the speech material of 21 children. Common amount of incorporated records is 6651. Final variant of database would contain speech material of 50 children for every age: 4year, 4 year 6 months, 5 years, 5 years 6 months, 6years, 6 years 6 months. The recording of speech material was conducted in 2006 year. Database contains speech material of that children, which sound and speech material was in database «INFANTRU», and other “new” children that were recorded from 4 years of age. Longitude recording of children speech is conducted with the period of 6 months. Every record is accompanied by detailed protocol and parallel videotaping. Tape recording of children’s speech materual was conducted in

situations of interaction with adult: spontaneous speech, answers to questions, reading, poetry or retelling a tale, count and alphabet, play.

The information about children is in the database: child name (abbreviated indication of first and second name); sex; date of birth; place of residence of child; child number in the family; presence or absence of brothers and sisters; the presence or absence of prenatal and chronic diseases; if child visits a kindergarten - common or logopedic; marks by scale RCDI: the correspondence of normative development of child or not; number of child in database «INFANTRU»; the place of recording. The information about family of child: the period of mother's residences in St-Petersburg to the time of child's birth, mother's education level, complete or incomplete family, who is the main caregiver: mother, father, grandmother, nanny.

Presentation of speech material in database "CHILDRU": There are original files by duration up to 5 mine, that demonstrate the situation: spontaneous speech, answers to questions, reading, poetry or retelling a tale, count and alphabet, play. Original files could contain as child speech, mother's speech, speech of investigator, other children, as different noises in play situations. Phrases; dialogues of child with adult; questions; sounds and syllables, pronounced separately; words, that contain one, two, three, four, five and more syllables; the mistakes that caused by different variants of changes, gaps and transpositions of phoneme /r/ in word; changes, gaps and transpositions of other phonemes and syllables in the word; mistakes in the phrases - wrong construction of phrase; reading words and phrases were chosen from original file. As description of words, as the phonetic transcription in symbols of IPA and the child's meaning of the words that contains mistakes are presented. The format of speech files is Windows PCM, 22050 Hz, 16 bit. The structure of database «CHILDRU» repeats the structure of precede database.

#### Database structure

Speech database structure is as follows:

```
VDB.EXE
BORLNDMM.DLL
CC3260MT.DLL
RTL60.BPL
VCL60.BPL
VCLX60.BPL
\DATA
  \CHILD.DB
  \RECORD.DB
  \<NAME><DD><SC>.WAV
\DOC
  \DESIGN.DOC
VDB.EXE      Database program shell
BORLNDMM.DLL Backup file
CC3260MT.DLL Backup file
RTL60.BPL    Backup file
VCL60.BPL    Backup file
VCLX60.BPL   Backup file
\DATA        Data catalogue (speech files and summary tables)
CHILD.DB     Child information table
RECORD.DB    Recording session information table
\DOC         Documentation catalogue
DESIGN.DOC   DB Description
```

#### Speech file format

Speech files are Windows PCM format, 22050 Hz, 16 bit. File name structure is as follows:

Files have following structure of name.

```
<NAME><DD><SC><SPc>.WAV
```

Indication	Description	Examples
NAME	Brief child's name and first letters of the second name	Skar
DD	Age: year (text), month (text)	4y 4y6m
SC	Situation code	S, T, A
SPc	Subcode of situation	W(n)F,D,Q,M, MW,MF, S, SY, R

<DD>- main file name by duration before 5 mines

#### Codes of situations

Code	Description
S	Spontaneous speech (child speaks about subjects, chosen by himself without adult's help)
R	Read
A	Answers to questions.
T	Poetry and tale
C	Count and alphabet
P	Play (interaction in play situation)

#### Subcodes of situations

Code	Description
W(π)	Word (amount of syllables)
F	Phrase
D	Dialogue
Q	Question
M	Mistakes
mw	Mistake in word
mf	Mistake in phrase
s	Sounds
sy	Syllables – read
r	Words, phrases – read

#### CHILD.DB file format

CHILD.DB is a table in text format. Record separator is line feed, record fields are in quotes.

Field number	Description	Format
1	Child number	number
2	Child number in database INFANTRU: disc, point, serial number	D.nnn
3	Date of birth	dd.mm.yyyy
4	Name	текст
5	Sex	m/f
6	Place of birth	Text
7	Number of child in family	first/not first
8	Visits kindergarten	Yes/No
9	Kindergarten: common, logopedic	C/L
10	Prenatal disorders	Text
11	Chronic disorders	Text
12	Mother's period of stay in St-Petersburg at the child birth (years)	number
13	Mother's education level at child records	average / average special / unfinished high / high

**RECORD.DB file format**

RECORD.DB is a table in text format.

Record separator is line feed, record fields are in quotes.

Field number	Description	Format
1	Child number	Numeral
2	Place of record: home, kindergarten	Text
3	Situation	S- Spontaneous speech (child speaks about subjects, chosen by himself without adult's help), R- Read, A- Answer questions; T- Poetry and tale, C- Count and alphabet, P- Play (interaction in play situation)
4	Sub situation	W(n) Word (syllables number), F- phrases, D- dialogue, Q- questions, M – mistakes, Mw- mistakes in word, mf – mistakes in phrase, s- sounds, syllables, r- reading of words and phrases
5	Child age (year, point.; month )	Numeral
6	Marks by RCDI-scale	Yes/no
7	Family type	complete / incomplete
8	Single child in the family	Yes/no
9	Caregiver (who brings up the child)	text (mother / father / grandmother / nanny)
10	Visits child a kindergarten	common logopedic no
11	Recording equipment	Tape recorder Marantz
12	Microphone type	SENNHEIZER e835S

The VDB.EXE program was created for effective work with database. This program allows choosing speech material in accordance with following parameters: Original record; Phrases; Dialogue; Questions; Mistakes in word in phoneme «R»; other mistakes in word – different variants of gap, change, transposition of phonemes and syllables; Mistakes in phrase; Separately pronounced sounds; Syllables; Reading of words and phrases; Words with one syllable; Words with two syllables; Words with three syllables; Words with four syllables; Words with five or more syllables. Choice of speech material could be made as for all children, as for every child.

From our point of view, children speech material from the database could be the basis for scientific projects about Russian language mastering by Russian children.

The work is financially supported by the Russian Fund for Humanities (projects № 06-06-00519a, 06-06-12623b).

**REFERENCES**

1. Galunov V.I., Kochanina Y.L., Solovyev A.N., et. al. Wideband Speech Database for Russian // International Workshop. Speech and computer. - SPB.- 2002. – P.113-115.
2. Victorov A.B., Victorova K.O., Voroncova A.V et all. Speech databases for problems of automatic recognition of speech and verification of speaker // Modern speech technologies. Proceedings of IX Session RAS.- M.:GEOS, 1999. – P.142-145.
3. Lyakso E.E., Gromova A.D., Bogorad M.A., Razumichin D.V., Frolova O.V., Kurazhova A.V., Romanova O.A., Novikova I.V., Galunov V.I. Sounds and speech database of children of first three years of life. // Proceedings XVI session RAS «Speech acoustics. Medical and biological acoustics. Architectural and building acoustics. Noises and vibrations », M. «GEOS», V.3., M. 2005, P.65-68.