The research team:
1. **Bogdanova-Beglaryan Natalia Viktorovna**, Doctor of Philology, professor of the department of the Russian Language (St. Petersburg State University);
2. **Martynenko Grigoriy Yakovlevich**, Doctor of Philology, professor of the department of Computational Linguistics (St. Petersburg State University);
3. **Baeva Ekaterina Mikhailovna**, Candidate of Philology, associate professor of the department of the English Language (St. Petersburg State University);
4. **Blinova Olga Vladimirovna**, Candidate of Philology, associate professor of the department of General Linguistics (St. Petersburg State University);
5. **Sherstinova Tatiana Yurievna**, Candidate of Philology. Sci., associate Professor of department of Computational Linguistics (St. Petersburg State University).

Set of scientific articles: «Everyday Russian Language (corpus study)»

Annotation
Corpus approach to speech studies means analyzing the language in its most natural form, and it is highly important to do spontaneous speech research drawn on corpora of large scopes. What's more, at all stages of corpus research (e.g. collecting, transcribing, annotating and analyzing the data on multi levels) linguists are faced with many difficulties of native, technical, ethical and methodological kind.

The source of most studies of the research group nominated for the award is a speech corpus of everyday communication in Russian, One Speaker's Day, also known as ORD (from *Odin Rechevoj Den*, the Russian title of the project). It is a large corpus of audio data collected by non-stop recording of oral speech of many people. The work on that project started at the faculty of philology of St Petersburg State University back in 2007 and has been done ever since, owing much to the very researchers of the abovementioned team.

The papers present results of linguistic data analysis on phonetic, morphological, lexical, syntactic and discursive levels. *Phonetic* studies describe the main features of spontaneous speech, such as tempo and phonetic realization of words, collocations and word forms; they also focus on prosody in close connection with intonation and the issue of utterance borders. On the morphological level on annotation, there are characteristic changes of nouns and verbs. Lexical level contains vocabulary descriptions and the dynamics of most frequent words. The syntactic level is the most complex and controversial because it is rather impossible to apply descriptive tools of written language while analyzing the spoken one. In oral speech it is difficult to find margins of collocations, not in the least because when building a sentence a speaker often uses various functional speech units which broaden communicative patterns and have several pragmatic goals. The very units are analyzed on discursive level, and the results have enabled to trace their usage and to make a glossary of the most frequent pragmatic marks which are generally even more common than autosemantic words.

The multi-level annotation level allows for studying speech as a communication act. Recorded speech is categorized into micro and macro episodes, enabling minute analysis of communicative roles and various speech acts.

Another promising research field is speech behavior. In spoken speech analysis, for instance, it is possible to study singing behavior in everyday communication.

Moreover, every person's speech reflects certain features of their social background, thus, to some extent, being characteristic not only on individual, but on social level. Any individual's speech should be studied in the context of their gender and age group. The annotation provided for finding differences in the speech of various social groups on all linguistic levels, including discursive. All parameters crucial for further sociolinguistic studies were specified in the especially designed sociolinguistic survey for the participants of the speech recording. The categories mentioned there are gender and age, education level, profession and social status, native language and even the nationality of parents. All of these factors had been conducive to

drawing up a classification of sociolects (which includes about 12 sections) that were later defined using special linguistic features. Each social group was defined not only regarding their vocabulary and grammar but concerning their core values and priorities. This information is vital for sociological studies of contemporary Russian society, also aiming at government social policy optimization.

Most research findings of the team work have been published as a book: «Everyday Russian Language in Different Social Groups» (Sankt Petersburg: LAIKA Publ., 2016. – 244p.).