Pragmatic markers (PMs) are indispensable elements of spoken discourse in any language. They are speech elements, having major influence on a pragmatic aspect of spoken discourse and being practically devoid of their own referential meaning. In spite of PMs wide circulation, they are very poorly studied. The current research demonstrates an interdisciplinary approach to study of PMs based on two representative speech corpora – ORD corpus of Russian Everyday Speech known as “One Day of Speech”–corpus and the “Balanced Annotated Collection of Texts” (SAT corpus). The research involves methodologies of different linguistics branches (phonetics, discourse analysis, sociolinguistics, psycholinguistics, corpus linguistics, etc.), making it possible to build formal statistical schemes which may be used both for theoretical linguistic studies and the improvement of NLP tasks.

The research is supported by the Russian Science Foundation, project # 18-18-00242 “Pragmatic Markers in Russian Everyday Speech”

Applications

The results of the project will find their practical application: 1) in the field of the applied linguistics, informational and speech technologies – to support the systems of automatic speech monitoring, voice search, speech synthesis and recognition systems, artificial intelligence, voice dialog systems when communicating with a computer or robot, 2) for teaching Russian as a foreign language, and 3) for conducting linguistic and forensic expertise based on audio records of speech communication.

Acknowledgements

The research is supported by the Russian Science Foundation, project # 18-18-00242 “Pragmatic Markers in Russian Everyday Speech”