

## Pragmatic Markers of Russian Everyday Speech: Quantitative Data

**Bogdanova-Beglarian N. V.**

Saint Petersburg State University  
Saint Petersburg, Russia  
n.bogdanova@spbu.ru

**Blinova O. V.**

Saint Petersburg State University,  
HSE University, Saint Petersburg  
Saint Petersburg, Russia  
o.blinova@spbu.ru,  
ovblinova@hse.ru

**Sherstinova T. Ju.**

HSE University, Saint Petersburg  
Saint Petersburg State University,  
Saint Petersburg, Russia  
t.sherstinova@spbu.ru,  
tsherstinova@hse.ru

**Troshchenkova E. V.**

Saint Petersburg State University  
Saint Petersburg, Russia  
e.troshchenkova@spbu.ru

**Gorbunova D. A.**

Saint Petersburg State University  
Saint Petersburg, Russia  
dgorbunova2@gmail.com

**Zajdes K. D.**

Saint Petersburg State University  
Saint Petersburg, Russia  
kristina.zaides@student.spbu.ru

**Popova T. I.**

Saint Petersburg State University  
Saint Petersburg, Russia  
tipopova13@gmail.com

**Sulimova T. S.**

Saint Petersburg State University  
Saint Petersburg, Russia  
spb.sulimova@gmail.com

### Abstract

The article summarizes the results of a large research project dedicated to investigation of pragmatic markers (PM) in Russian everyday speech. Pragmatic markers are essential in spontaneous spoken discourse; thus, the quantitative data on their usage are necessary for solving both theoretical and practical issues related to the study of spoken communication. New results were obtained on the data of two speech corpora: “One Day of Speech” (ORD; mostly dialogues; the annotated subcorpus contains 321 504 tokens) and “Balanced Annotated Text Library” (SAT; monologues; the annotated subcorpus includes 50 128 tokens). Statistical data were calculated for PM in dialogic and monologic speech, pragmatic markers common in both types of speech (e. g., hesitant markers like *vot*, *tam*, *tak*) are identified, as well as PM that are the most typical for monologues (e. g., boundary markers like *znachit*, *nu*, *vot*, *vs'o*) or dialogue (e. g., ‘*xeno*’-markers such as *takoi*, *grit* and metacommunicative markers like *vidish'*, (*ja*) *ne znaju*). Special attention is given to the pragmatic markers usage in different communicative situations.

**Keywords:** Russian everyday speech; speech corpus; pragmatic marker; corpus annotation; monologue; dialogue

**DOI:** 10.28995/2075-7182-2021-20-119-126

## Прагматические маркеры русской повседневной речи: количественные данные

**Богданова-Бегларян Н. В.**

СПбГУ  
Санкт-Петербург, Россия  
n.bogdanova@spbu.ru

**Шерстинова Т. Ю.**

НИУ ВШЭ СПб, СПбГУ,  
Санкт-Петербург, Россия  
tsherstinova@hse.ru  
t.sherstinova@spbu.ru

**Горбунова Д. А.**

СПбГУ  
Санкт-Петербург, Россия  
dgorbunova2@gmail.com

**Попова Т. И.**

СПбГУ  
Санкт-Петербург, Россия  
tipopova13@gmail.com

**Блинова О. В.**

СПбГУ, НИУ ВШЭ СПб  
Санкт-Петербург, Россия  
o.blinova@spbu.ru,  
ovblinova@hse.ru

**Трошchenkova Е. В.**

СПбГУ  
Санкт-Петербург, Россия  
e.troshchenkova@spbu.ru

**Зайдес К. Д.**

СПбГУ  
Санкт-Петербург, Россия  
kristina.zaides@student.spbu.ru

**Сулимова Т. С.**

СПбГУ  
Санкт-Петербург, Россия  
spb.sulimova@gmail.com

### Аннотация

В статье подводятся итоги большого исследовательского проекта, посвященного изучению прагматических маркеров (ПМ) русской повседневной речи. Прагматические маркеры являются неотъемлемой частью спонтанного устного дискурса, поэтому количественные данные об их реализации необходимы для решения как теоретических, так и практических задач, связанных с изучением речевой коммуникации. Новые результаты получены на материале двух речевых корпусов: «Один речевой день» (ОРД; преимущественно диалоги; аннотированный подкорпус включает 321 504 токена) и «Сбалансированная аннотированная текстотека» (САТ; монологи; в аннотированном подкорпусе 50 128 токенов). Статистические данные посчитаны для ПМ диалогической и монологической речи, выявлены маркеры, общие для обоих типов речи (хезитативы типа *вот, там, так*), а также те, что свойственны в большей степени монологу (разграничители типа *значит, ну вот, всё*) или диалогу (ксенопоказатели типа *такой, грит* и метакоммуникативы типа *видишь, (я) не знаю*). Особое внимание в работе уделяется употреблению прагматических маркеров в разных условиях коммуникации.

**Ключевые слова:** русская повседневная речь; речевой корпус; прагматический маркер; корпусная разметка; монолог; диалог

## 1 Введение

Под прагматическими маркерами (ПМ) в работе понимаются те единицы устной речи, которые прошли процесс прагматикализации, в результате практически утратили свое исходное лексическое и/или грамматическое значение и приобрели прагматическое, выполняя в дискурсе лишь определенные функции: маркируют границы реплик (*значит, ну вот, всё*) или ввод чужой речи (*такой/ая/ие, типа (того что), грит/грю/грят* и под.), вербализуют хезитацию говорящего (*это самое, как его (её, их), там*), его рефлексия по поводу сказанного либо готовящегося к произнесению (*или как там? или как сказать? или что?*) или самокоррекцию (*это, это самое*), выражают метакоммуникацию (*знаешь, понимаешь, да, (я) не знаю, что ещё?*) и т. п.

Отличия прагматических маркеров от дискурсивных слов, или дискурсивных маркеров (ДМ), под которыми и в зарубежной, и в отечественной лингвистике понимается очень широкий класс функциональных единиц (см., например: [1; 15; 5; 6; 9]), изложены в работе: [12]. Анализ ПМ в настоящем исследовании предшествовало тщательное ручное аннотирование корпус-

ного материала с разграничением ПМ и омонимичных им значимых единиц языка (см., например: [4]). Пилотная разметка корпусных данных выполнялась четырьмя независимыми экспертами, что позволило оценить состоятельность и объективность такой разметки.

Принимая во внимание вариативность ряда форм ПМ, для описания системы ПМ были введены понятия *базового варианта* и *структурных вариантов* (реальных употреблений) ПМ. Объектом количественного анализа в настоящей статье стали именно реальные употребления ПМ, с качественной стороны они описаны в специальном словаре прагматических маркеров<sup>1</sup>.

## 2 Материал и методика исследования

В основу словника прагматических маркеров, включающего 60 единиц, была положена типология ПМ русской устной речи, описанная в [4; 13]. Типология разрабатывалась на основе эмпирических данных двух речевых корпусов: корпуса монологической речи «Сбалансированная аннотированная текстотека» (САТ) и корпуса русской повседневной речи «Один речевой день» (ОРД) (см.: о них: [8; 2; 3; 10; 11]). Были выделены и описаны следующие типы ПМ: хезитативы (Х), рефлексивы (Ф), метакоммуникативы (М), разграничители (Г), ксенопоказатели (К), аппроксиматоры (А), дейктические (Д) и ритмообразующие (Р) маркеры, маркеры самокоррекции (С) и заместители (З).

На двух аннотированных выборках (ОРД, 321 504 токена; и САТ, 50 128 токенов) получены количественные характеристики реальных употреблений ПМ и установлены корреляции между появлением ПМ в повседневной речи и различными факторами: тип речи (диалог – монолог), место и тип коммуникации, социальная роль говорящего в конкретном коммуникативном макроэпизоде, а также его социальные и психологические характеристики. В Таблице 1 представлены данные об объеме проанализированных подвыборок в разных социолектах.

Гендер				
мужчины		женщины		
171 497		158 390		
Возрастные группы				
младшая		средняя		старшая
143 805		67 089		118 993
Образование				
неоконч. высшее / среднее		среднее специальное		высшее
47 634		36 313		215 540
УРК (уровень речевой компетенции)				
низкий		средний		высокий
19 983		245 586		54 825
Профессиональные группы				
ГУМ	ЕСТ	ИНЖ	ИТ	ОБР
30 493	11 988	30 897	29 105	29 738
ОФ	РАБ	СИЛ	СО	ТВОР
8 989	13 290	1 587	27 041	17 581

Таблица 1: Объем подвыборок с учетом различных социальных параметров (в токенах)

Рассмотрим основные полученные результаты.

## 3 Общие частоты ПМ (реальные словоупотребления)

Анализ корпусного материала позволил получить статистику конкретных словоупотреблений ПМ (370 реализаций 60-ти базовых ПМ) (о базовых ПМ и их вариантах см.: [13]). На аннотиро-

<sup>1</sup> Мы отдаем себе отчет в том, что с момента записи наших корпусов прошло уже довольно много времени и полученные данные нельзя считать отражением *сегодняшней* устной коммуникации, но когда речь идет о корпусных массивах данных, которые требуют больших трудозатрат по сбору, систематизации, расшифровке, разметке и разнообразной обработке, иначе, по-видимому, и быть не может. Это близко к ситуации создания любого словаря, который к моменту выхода отчасти уже утрачивает адекватность реальному состоянию нашего языка и тем более нашей повседневной речи.

ванных подвыборках было выделено 370 и 133 ПМ соответственно. В среднем ПМ показали частоту 29 611 ipm (2,96 %) в диалогической речи и 19 300 ipm (1,93 %) в монологах. Видно, что ПМ в диалоге используются на треть чаще.

Верхние десять позиций общего частотного словника (топ-10) заняли следующие единицы: ВОТ, ТАМ, ДА, ТАК, КАК БЫ, ГОВОРИТ (ГРИТ), ЗНАЧИТ, ЗНАЕШЬ, НУ ВОТ, СЛУШАЙ. Абсолютным «лидером» во всех социолектах является маркер ВОТ – его частотность составляет около четверти от общего количества ПМ в речи всех говорящих (1827/20,7 %) (здесь и везде далее в скобках через слеш показаны абсолютные и относительные величины). Достаточно частотен также маркер ТАМ, имеющий ранг 2 и в общем словнике (976/11,1 %), и в большинстве социолектов. Частотность остальных маркеров верхней зоны частотного списка существенно ниже: от 3 до 5 % (от 243 до 462 употреблений).

Интересно, что это общее распределение употреблений ПМ полностью (хотя и не всегда в одинаковом порядке) повторилось в речи женщин, людей среднего возраста и носителей языка с высоким уровнем речевой компетенции (УРК). Верхние зоны частотных списков речи мужчин, а также носителей языка среднего возраста отличаются отсутствием в них контактоустанавливающих маркеров ЗНАЕШЬ и СЛУШАЙ, которые в общем списке имеют ранги, соответственно, 8 и 10 и входят практически во все другие частотные списки. Только в речи старших говорящих место маркера СЛУШАЙ занял контактный глагол ПОНИМАЕШЬ. В речи молодежи заметен высокий ранг маркеров КОРОЧЕ (5) и ТИПА (9).

Прагматический маркер КОРОЧЕ, именно в таком, редуцированном, варианте, а не в полном (базовом) КОРОЧЕ (ГОВОРЯ), в целом маркирует речь мужчин (ранг 8), младших говорящих (ранг 5), в том числе не имеющих высшего образования и высокого УРК, а также речь детей (ранг 10). Из профессиональных групп говорящих этот ПМ имеет высокий ранг 3 в речи рабочих и представителей силовых структур.

Дети, которые попали в число говорящих только как коммуниканты в корпусе ОРД, имеют свой собственный частотный список ПМ (топ-10), не совпадающий с данными по взрослой речи: ТАМ, КАК БЫ, ВОТ, НЕ ЗНАЮ, ДА, ТАК, ТИПА, ВОТ ТАК ВОТ, ЗНАЕТЕ, КОРОЧЕ.

Маркер ТИПА можно признать специфическим показателем речи младших говорящих (ранг 9), в том числе детей (ранг 7), а также речи представителей силовых структур (ранг 6) и творческой интеллигенции (ранг 10).

При рассмотрении условий коммуникации наибольшие отличия демонстрирует общение в кафе и ресторанах. В целом по корпусу (преимущественно это общение *дома, в офисе, на улице, в казарме, в медицинских учреждениях*) ранг 1 имеет маркер ВОТ, ранг 2, за исключением казармы, – маркер ТАМ. В «казарменном» общении на втором месте в частотном списке оказался маркер КОРОЧЕ, совсем немного уступающий по употребительности маркеру ВОТ (31/19,4 vs 32/20 %). После ТАМ в этом списке следуют маркеры ТИПА и КАК БЫ, также весьма типичные в разговорах между курсантами. *Кафе и рестораны* дают совсем другую картину: на первом месте в соответствующих словниках – с одинаковой частотой – находятся маркеры ТАМ и ДА (по 18/17,4 %), затем идет ПМ СЛУШАЙ (13/12,4 %) и лишь затем ВОТ (9/8,6 %).

«Лидирующее» положение маркера ВОТ (ранг 1) сохраняется во всех **социальных ролях** говорящих. Особенно высока доля этого ПМ в речи «родителей» (132/40 %). Ранг 2 почти во всех социальных ролях имеет маркер ТАМ. Единственным исключением стала роль «однокурсника», в которой ТАМ отошло на четвертую позицию, уступив место маркерам ДА и КОРОЧЕ. Очень распространенный в нашей повседневной речи ПМ КОРОЧЕ встретился еще в верхней зоне частотного списка социальной роли «друга» (ранг 4), а позицию 7 в большинстве ролей («друг», «коллега», «муж»/«жена», «однокурсник» и «родители») на удивление устойчиво занимает маркер КАК БЫ. Самый высокий его ранг 3 зафиксирован в роли «ребенка» («сын»/«дочь»), в социальной роли «подруги» этот маркер имеет ранг 4, а в роли «клиент – сервис» он вообще не попал в верхнюю зону частотного списка прагматических маркеров.

*Метакоммуникативы* вошли в верхнюю зону (топ-10) частотных списков ПМ для всех социальных ролей, особенно их много в роли «родителей» (ПОНИМАЕШЬ, ЗНАЕШЬ, ПРЕДСТАВЛЯЕШЬ, СЛУШАЙ) и «подруги» (ЗНАЕШЬ, ПОНИМАЕШЬ, СЛУШАЙ). Видимо, именно в этих ролях говорящий в наибольшей степени стремится установить, а затем и удержать контакт с собеседником.

Интересно было также проанализировать распределение в разных социальных ролях маркеров-ксенопоказателей ГОВОРИТ (ГРИТ) и ГОВОРЮ (ГРЮ), свидетельствующих о том, что говорящие в этих ролях часто пересказывают чужие (или свои) слова и мнения. Оба эти ПМ обнаружались в социальной роли «коллеги» (ранги 3 и 5 соответственно), в ролях «мужа»/«жены» и «клиент – сервис» в топ-10 вошел только маркер ГОВОРИТ (ранги 8 и 3 соответственно). В других ролях эти ПМ в верхних зонах частотных списков не отмечены.

Распределение употреблений ПМ с учетом *психотипа* говорящего (экстраверты – интроверты) показало, что первые две позиции в частотных списках маркеров в речи обоих психотипов занимают единицы ВОТ и ТАМ, причем первых в 2-3 раза больше, чем вторых: 342/21,5 vs 166/10,5 % у экстравертов, 210/26,9 vs 70/8,95 % у интровертов. Речь экстравертов отличает также наличие в зоне топ-5 ксенопоказателя ГОВОРИТ (ранг 3), а речь интровертов – наличие в той же зоне метакоммуникативов ДА и ПОНИМАЕШЬ (ранги, соответственно, 3 и 4).

Подключение к анализу материала данных о *темпераменте и уровне невротизма (нейротизма)* говорящего не поколебало высокого первого ранга маркера ВОТ во всех случаях, хотя доля его оказалась весьма низкой (по сравнению с другими группами) в речи говорящих с низким уровнем невротизма (65/12,8 %) и в смешанной группе холериков-флегматиков (14/12,4 %). И эти же две группы говорящих «любят» маркер КОРОЧЕ: его ранг 3 в обоих случаях. Высокий ранг 2 маркера ТАМ сохраняется в речи говорящих с низким и средним уровнем невротизма, а также в речи сангвиников, флегматиков и смешанной группы холериков-флегматиков. В других группах говорящих на эту позицию выдвинулся маркер ДА (говорящие с высоким уровнем невротизма, меланхолики и холерики).

#### 4 Словарь прагматических маркеров

Наличие различных функций ПМ в устном тексте, часто совмещенных в одной единице, – при размытом, ослабленном или вовсе отсутствующем семантическом их наполнении – вынудило поставить задачу создания специального словаря таких единиц (далее – Словарь ПМ), который должен отражать богатство и разнообразие как самого списка ПМ, так и функций, выполняемых ими в устной речи. Такой ресурс может быть исключительно полезен специалистам самых разных направлений.

К настоящему времени словарь ПМ создан, он включает в себя перечень наиболее частотных, регулярно используемых в русской устной речи, прагматических маркеров, с указанием их типа, функции и примеров употребления. Словарь подготовлен в двух версиях – бумажном варианте [7] и электронной версии с аудиопримерами. Словарные статьи в Словаре ПМ построены как лексикографические эссе, что обусловливается спецификой самого материала (ср.: «Путеводитель по дискурсивным словам русского языка» [1]). В таких эссе дается описание функций (прагматических значений) и особенностей функционирования всех выявленных ПМ. При этом разные функции ПМ в спонтанном тексте часто совмещаются в одной единице – при размытом, повторим, ослабленном или вовсе отсутствующем семантическом ее наполнении, т. е. ПМ оказались почти всегда полифункциональны и очень зависимы в своем статусе от контекста, что еще более укрепило нас в осознании необходимости такого словаря.

Словарь ПМ может быть полезен специалистам самого разного толка: собственно лингвистам, исследователям повседневной русской речи (коллоквиалистика, когнитивистика, социо- и психолингвистика); создателям грамматики речи (или модели языка, основанной на употреблении – *usage-based theory*, см.: [14]), которая, вне всякого сомнения, отличается от грамматики языка; специалистам по корпусной лингвистике, разрабатывающим системы автоматического аннотирования и анализа корпусного материала на разных уровнях; переводчикам спонтанных текстов на другие языки, хотя бы в рамках художественного произведения, при передаче речи персонажей; преподавателям русского языка иностранцам, которые вынуждены учиться воспринимать и правильно понимать нашу спонтанную речь как устно, так и письменно, при чтении русскоязычных текстов. В ряду других словарей, построенных на материале устных корпусов, Словарь ПМ отражает лексическое и дискурсивное своеобразие повседневной русской речи.

Электронная версия словаря доступна онлайн на сайте <https://www.ord-multimedia-dict.com/>.

## 5 Заключение

В статье представлены количественные данные о частоте использования прагматических маркеров устной речи, полученные на основе двух проаннотированных речевых корпусов (ОРД и САТ).

Основой для анализа стали ранжированные частотные списки ПМ с информацией об относительной частоте их употребления в диалоге (выборка ОРД) и монологе (выборка САТ); а также частотные списки функциональных типов прагматических маркеров с привязкой к типу речи (диалог vs монолог). Кроме того, получены частотные списки ПМ для разных условий коммуникации (т. е. с учетом типа разговора, социальной роли говорящего и др.) и для разных групп говорящих (социолектов). Все эти данные суммированы в словаре ПМ, имеющем, кроме бумажной, электронную версию, дающую возможность прослушать аудиопримеры на каждый тип ПМ.

Качественный анализ конкретных употреблений ПМ вынужденно оставлен за рамками настоящей статьи, равно как и количественные данные по распределению базовых вариантов ПМ и конкретных их функциональных типов. Это может стать предметом отдельных статей.

В целом анализ корпусного материала показал, что ПМ действительно регулярно встречаются в повседневной речи говорящих всех социальных групп и во всех без исключения коммуникативных ситуациях. Однако частота использования определенных базовых ПМ и их вариантов меняется в зависимости от характеристик как самих говорящих, так и условий коммуникации. Наиболее типичные тренды использования ПМ описаны в настоящей статье. Следует принимать во внимание, что в речи отдельных говорящих могут наблюдаться определенные «выбросы», отличающие идиостиль от «среднего по группе». Индивидуальная вариативность ПМ и ее мера заслуживает специального рассмотрения.

Полученные данные расширяют теоретические представления об употреблении ПМ в реальном повседневном общении и могут быть использованы в разнообразных практических приложениях – от разработки систем в области речевых технологий до задач лингвистической экспертизы, практики перевода и преподавания русского как иностранного.

## Благодарность

Исследование проведено при финансовой поддержке гранта Санкт-Петербургского государственного университета (проект № 75254082 «Моделирование коммуникативного поведения жителей российского мегаполиса в социально-речевом и прагматическом аспектах с привлечением методов искусственного интеллекта»).

## References

- [1] Baranov A. N., Plungyan V. A., Rakhilina E. V. (1993), Russian Discourse Words Guide [Putevoditel' po diskursivnym slovam russkogo jazyka]. Moscow.
- [2] Bogdanova-Beglarian N.V., Blinova O.V., Sherstinova T.Yu., Martynenko G.Ya. (2019a), Corpus of Russian Everyday Speech “One Day of Speech”: present state and prospects [Korpus russkogo jazyka povsednevnogo obshchenia «Odn rechevoj den'»: tekushchee sostojanie i perspektivy] // Proceedings of the V. V. Vinogradov Russian Language Institute. Vol. 21. Russian National Corpus: Research and Development [Trudy In-ta russkogo jazyka im. V. V. Vinogradova. Vyp. 21. Nacional'nyj korpus russkogo jazyka: issledovania i razrabotki], Moscow, pp. 101–110.
- [3] Bogdanova-Beglarian N.V., Blinova O.V., Zajdes K.D., Sherstinova, T.Yu. (2019b), “Balanced Annotated Text Library” (SAT; monologues): Studying the Specifics of Russian Monological Speech [“Sbalansirovannaya annotirovannaya tekstoteka” (SAT): izuchenie spetsifiki russkoj monologicheskoi rechi] // Proceedings of the V. V. Vinogradov Russian Language Institute. Vol. 21. Russian National Corpus: Research and Development [Trudy In-ta russkogo jazyka im. V. V. Vinogradova. Vyp. 21. Nacional'nyj korpus russkogo jazyka: issledovania i razrabotki]. Moscow, pp. 111–126.
- [4] Bogdanova-Beglarian N.V., Blinova O.V., Martynenko G.Ya., Sherstinova, T.Yu., Zajdes K.D., Popova T.I. (2019c), Annotation of Pragmatic Markers in the Russian Speech Corpus: Problems, Searches, Solutions, Results [Annotirovanie pragmaticheskikh markerov v russkom rechevom korpuse: problemy, poiski, reshenia, rezul'taty] // Computational Linguistics and Intellectual Technologies [Kompjuternaya lingvistika i intellektual'nye tekhnologii]: Proceedings of the International Conference “Dialogue 2019”. Vol. 18 (25). Moscow, pp. 72–85.

- [5] Discursive Words of the Russian language: An Experience of Context-Semantic Description [Diskursivnye slova russkogo jazyka: Opyt kontekstno-semanticheskogo opisani]. (1998), Moscow.
- [6] Discursive Words of the Russian language: Variation and Semantic Unity. Collection of articles [Diskursivnye slova russkogo jazyka: varjирование i semanticheskoe jedinstvo. Sb. Statej] (2003), Moscow, 207 p.
- [7] Pragmatic Markers of Russian Everyday Speech: Dictionary-Monograph Pragmaticheskie [Markery russkoj povsednevnoj rechi: slovar'-monografija] (2021), St. Petersburg. In Print.
- [8] Asinovsky A.S., Bogdanova N.V., Rusakova M.V., Ryko A.I., Stepanova S.B., Sherstinova T.Yu. (2009), The ORD Speech Corpus of Russian Everyday Communication "One Speaker's Day": Creation Principles and Annotation // Matoušek, V., Mautner, P. (eds.) TSD 2009. LNAI, vol. 5729. Springer, — Berlin-Heidelberg, 2009, pp. 250–257.
- [9] Beliao Julie, Lacheret Anne. Disfluency and Discursive Markers: when Prosody and Syntax Plan Discourse // DiSS 2013: The 6th Workshop on Disfluency in Spontaneous Speech, — Stockholm, Sweden. № 54 (1), 2013, pp. 5–9.
- [10] Bogdanova N.V., Sherstinova T.Yu., Blinova O.V., Martynenko G.Yu. An Exploratory Study on Sociolinguistic Variation of Spoken Russian // SPECOM 2016. Lecture Notes in Artificial Intelligence, LNAI, vol. 9811. — Springer, Switzerland, 2016. — pp. 100–107.
- [11] Bogdanova N.V., Sherstinova T.Yu., Blinova O.V., Baeva E.M., Martynenko G.Ya., Ryko A.I. Sociolinguistic Extension of the ORD Corpus of Russian Everyday Speech // SPECOM 2016, Lecture Notes in Artificial Intelligence, LNAI, vol. 9811. — Springer, Switzerland, 2016b — pp. 659–666.
- [12] Bogdanova N.V., Filyasova Yu.A. Discourse vs Pragmatic Markers: A Contrastive Terminological Study // 5th International Multidisciplinary Scientific Conference on Social Sciences and Arts, SGEM 2018. — Vienna ART Conference Proceedings. Vol. 5, Iss. 3.1, 2018. — pp. 123–130.
- [13] Bogdanova N.V., Blinova O.V., Sherstinova T.Yu., Troshchenkova E.V., Gorbunova D.A., Zaides K.D. Pragmatic Markers of Russian Everyday Speech: the Revised Typology and Corpus-Based Study // Proceedings of the 25th Conference of Open Innovations Association, FRUCT. — Helsinki, Finland, 2018. — pp. 57–63.
- [14] Boye Kasper, Harder Peter. A Usage-based Theory of Grammatical Status and Grammaticalization // Language, 88 (1), 2012. — pp. 1–44.
- [15] Shiffrin Deborah. Discourse Markers. Cambridge University Press, Cambridge, UK, 1996.

## Список литературы

- [1] Баранов А. Н., Плулгян В. А., Рахилина Е. В. Путеводитель по дискурсивным словам русского языка. — М.: Помовский и партнеры, 1993.
- [2] Богданова-Бегларян Н. В., Блинова О. В., Мартыненко Г. Я., Шерстинова Т. Ю. Корпус русского языка повседневного общения «Один речевой день»: текущее состояние и перспективы // Труды Ин-та русского языка им. В. В. Виноградова. — Вып. 21. Национальный корпус русского языка: исследования и разработки / Гл. ред. А. М. Молдован. Отв. ред. вып. В. А. Плулгян. — М., 2019а, сс. 101–110.
- [3] Богданова-Бегларян Н. В., Блинова О. В., Зайдес К. Д., Шерстинова Т. Ю. Корпус «Сбалансированная аннотированная текстотека» (САТ): изучение специфики русской монологической речи // Труды Ин-та русского языка им. В. В. Виноградова. — Вып. 21. Национальный корпус русского языка: исследования и разработки / Гл. ред. А. М. Молдован. Отв. ред. вып. В. А. Плулгян. — М., 2019б, сс. 111–126.
- [4] Богданова-Бегларян Н. В., Блинова О. В., Мартыненко Г. Я., Шерстинова Т. Ю., Зайдес К. Д., Попова Т. И. Аннотирование прагматических маркеров в русском речевом корпусе: проблемы, поиски, решения, результаты // Компьютерная лингвистика и интеллектуальные технологии: По м-лам ежегодной международной конференции «Диалог» / Гл. ред. В. П. Селегей. — Вып. 18 (25), 2019в, сс. 72–85.
- [5] Дискурсивные слова русского языка: Опыт контекстно-семантического описания / Под ред. К. Л. Киселевой, Д. Пайара. — М.: Метатекст, 1998.
- [6] Дискурсивные слова русского языка: варьирование и семантическое единство. Сб. статей / Под ред. К. Л. Киселевой, Д. Пайара. — М.: Азбуковник, 2003.
- [7] Прагматические маркеры русской повседневной речи: Словарь-монография / Сост., отв. ред. и автор предисловия Н. В. Богданова-Бегларян. — СПб.: Нестор-История, 2021. — В печати.
- [8] Asinovsky, A., Bogdanova, N., Rusakova, M., Ryko, A., Stepanova, S., Sherstinova, T. The ORD Speech Corpus of Russian Everyday Communication «One Speaker's Day»: Creation Principles and Annotation // Matoušek, V., Mautner, P. (eds.) TSD 2009. — LNAI, Vol. 5729. — Springer, Berlin-Heidelberg, 2009. — P. 250–257.

- [9] Beliao Julie, Lacheret Anne. Disfluency and Discursive Markers: when Prosody and Syntax Plan Discourse // DiSS 2013: The 6th Workshop on Disfluency in Spontaneous Speech, Stockholm, Sweden. — № 54 (1), 2013. — P. 5–9.
- [10] Bogdanova-Beglarian, N., Sherstinova, T., Blinova, O., Martynenko, G. An Exploratory Study on Sociolinguistic Variation of Spoken Russian // SPECOM 2016. Lecture Notes in Artificial Intelligence. — LNAI, Vol. 9811. — Springer, Switzerland, 2016a. — P. 100–107.
- [11] Bogdanova-Beglarian, N., Sherstinova, T., Blinova, O., Baeva, E., Martynenko, G., Ryko, A. Sociolinguistic Extension of the ORD Corpus of Russian Everyday Speech // SPECOM 2016, Lecture Notes in Artificial Intelligence. — LNAI, vol. 9811. — Springer, Switzerland, 2016b. — P. 659–666.
- [12] Bogdanova-Beglarian, N. V., Filyasova, Yu. A. Discourse vs Pragmatic Markers: A Contrastive Terminological Study // 5th International Multidisciplinary Scientific Conference on Social Sciences and Arts, SGEM 2018. Vienna ART Conference Proceedings, 19-21 March, 2018. — Vol. 5, Iss. 3.1. — P. 123–130.
- [13] Bogdanova-Beglarian, N. V., Blinova, O. V., Sherstinova, T. Yu., Troshchenkova, E. V., Gorbunova, D. A., Zajdes, K. D. Pragmatic Markers of Russian Everyday Speech: the Revised Typology and Corpus-Based Study // Proceedings of the 25th Conference of Open Innovations Association FRUCT / S. Balandin, V. Niemi, T. Tuytina (eds.). — Helsinki, Finland, 2019. — P. 57–63.
- [14] Boye Kasper, Harder Peter. A Usage-based Theory of Grammatical Status and Grammaticalization // Language. — 88 (1) — 2012. — P. 1–44.
- [15] Shiffrin Deborah. Discourse markers. — Cambridge University Press, Cambridge, UK, 1996.