

ЖИВОЙ СТИЛИСТИЧЕСКИЙ СЛОВАРЬ РУССКОГО ЯЗЫКА

Справочный лингвистический веб-сервис

Создан в 2017-2019 гг. в рамках научно-исследовательского проекта
«Лингвистическая подготовка и создание электронного
"Живого стилистического словаря русского языка"»
Linguistic Development and Creation of the Electronic Live Stylistic Dictionary of Russian
при поддержке РФФИ (грант № 17-04-00421)

<http://livedict.syllabica.com/>

Содержание и основные результаты проекта

Руководитель проекта – доктор филологических наук, профессор А.А. Кретов
Автор концепции – доктор филологических наук Г.В. Векшин.
Участники проекта, разработчики: В.В. Леденева, Е.Ю. Кукушкина, К.М. Шилихина, Е.С. Максимов, М.Н. Герцев, Е.В. Андреева, С.Е. Виноградова.

«Стратегии научно-технического развития Российской Федерации»

Направление, в рамках которого будут использованы результаты научного проекта:

Исследования в области понимания процессов, происходящих в обществе и природе, развития природоподобных технологий, человеко-машинных систем, управления климатом и экосистемами, а также исследования, связанные с этическими аспектами технологического развития, изменениями социальных, политических и экономических отношений.

Область знания: 04

Код классификатора: 04-120

Код ГРНТИ: 16.21.65

Почему создан и зачем нужен «Живой стилистический словарь»

Владеть языком и осваивать язык – значит уметь проверять свое понимание слова. Для этого и существуют справочные лингвистические издания, словари. В нашу эпоху лавинообразного роста количества информации необходимость быстро получать самые актуальные сведения о слове, его значении и стилистическом использовании становится особенно острой. Для удовлетворения этих потребностей и создан электронный справочник нового поколения – «Живой стилистический словарь русского языка».

Предлагаемый словарь представляет собой веб-сервис, открытый для любого пользователя электронных устройств. Он дает возможность мгновенного получения разнообразных сведений о стилистическом значении слова – таком, которое возникает в силу его устойчивого употребления в коммуникативной среде разного рода и для воплощения различных ролей и намерений говорящего. До сих пор эти значения, хотя и очень фрагментарно, передавались в составе толкования слова в словарях в виде стилистических помет (*разг., книжн, научн., поэтич., высок., фамиллярн., бранн. и т.п.*). Однако пользоваться этими сведениями было трудно, потому что за многими из определений стояли комбинации разных более простых значений, однако они не охватывали всех видов стилистической окраски, и представление об использовании слова в итоге формировалось на уровне образного мышления пользователя, и очень приблизительно.

Еще одна проблема традиционных печатных словарей в том, что они опираются не на объективную картину коммуникативной практики общества, а на индивидуальное видение этой картины самими составителями словаря. Наконец, стилистические пометы в печатных словарях рискуют быстро устареть и исказить характер реального употребления слова еще до того, как словарь выйдет в свет: в результате печатные словари часто навязывают слову его уже не актуальную стилистическую характеристику, таким образом лишая стилистическую помету своей главной функции – служить гидом социокультурного и эмоционально-оценочного словоупотребления. Печатные словари не успевают следить за реальными изменениями стилистических свойств языковых единиц, не способны оперативно представлять тяготение слова к той или иной среде или его отторжение этих контекстов, следить за изменениями и колебаниями стилистических значений, их динамикой.

Это положение меняет современный электронный словарь – «Живой стилистический словарь русского языка», на основе корпусных данных представляющий окраску слова на основе его фиксируемых применений в характерных текстах, контекстах и сочетаниях.

Современные корпусные методы лингвистики и технологии машинного обучения, реализованные в проекте «Живого стилистического словаря» и его приложения «Стилевой подсказчик», дают возможность мгновенного получения разносторонней информации о контекстах бытования слова, словосочетания, а также, с помощью «Стилевого подсказчика», выявления стилистической ориентации текста, его стилиевой доминанты, – как разговорной, идеологической, религиозной и т.д.

В основе этих лингвистической методик лежит представление о характере не только контекстуальных притяжений данной единицы, но и характере ее контекстуальных отторжений. Чтобы зафиксировать эти отторжения и притяжения, должны быть выявлены и типологизированы специфичные тексты, легко обнаруживаемые по ключевым словам или другим лингвистическим, в частности жанровым, признакам. Так, невозможность непереносного и/или нецитатного употребления слова печалька в книжных текстах, тяготеющих к стилистической замкнутости (деловой, религиозной, научной литературе) и его высокая частотность в разговорно-бытовых контекстах и жанрах позволяют Словарю квалифицировать его как разговорное; выявление контекстов, характерных для жанров молодежного общения, позволяет представить его возрастную ориентацию; невстречаемость слова в текстах ранее 2000 года укажет на его статус как неологизма.

Составленные и применяемые для выявления этих характеристик корпуса насчитывают более 40 тысяч текстов с преобладающей социокультурной и тональной спецификой.

Живой стилистический словарь обеспечивает базу для целостного автоматического стилистического портретирования русской лексики и фразеологии в синхронии и в динамике ее исторического развития.

Размер словаря составляет более 8 млн. словоупотреблений.

Разработанный коллективом проекта сервис «Стилевой подсказчик» формирует обобщенный стилевой портрет текста на основе базовых стилистических измерений, в их положительном и отрицательном проявлении. Проведенное экспертное тестирование сервиса показывает его более чем 80-процентную точность при идентификации стилевой принадлежности текста.

ЖСС стремится к учету всего комплекса факторов, влияющих на формирование стилистической семантики языковой единицы. Важное отличие «Живого стилистического словаря» от всех существующих проектов, частично решающих сходные задачи, и созданных на их основе сервисов – Национального корпуса русского языка, Генерального интернет-корпуса русского языка, проектов автоматического определения авторства (Houvardas, Stamatas 2006; Романов 2009; Maitra, Ghosh 2015 и др.), идентификации когнитивного стиля (Frias-Martinez, Chen 2007 и др.), выявления пропаганды в СМИ (Hamborg, Donnay 2018; Barrón-Cedeño 2019), автоматического выявления жанров (Santini 2009; Bulygin, Sharoff 2018) и некоторых других – состоит в ориентации составителей его корпусов не на формальные признаки текста, а на его контекстуально-ролевую коммуникативную природу. Так, далеко не всякий газетный текст войдет в идеологический корпус ЖСС, а только такие, где наиболее активно проявляется доминантная для этой сферы социокультурная функция — идеологическая, реализуемая в стремлении автора выступить в качестве идеологически ангажированного лица, заинтересованного в распределении и перераспределении власти, выражающегося в интенсивном представлении соответствующей типической коммуникативной роли и актуализации идеологического контекста. Подобно этому, научный корпус Словаря составят не всякие тексты, формально относимые к научным, и даже не в первую очередь такие, которые успешно решают собственно научные задачи, а скорее такие, где автор стремится активно позиционировать себя в качестве ученого и демонстрировать отнесенность своих занятий к сфере науки. Именно для решения таких, имиджевых задач автору потребуется максимальная концентрация терминов (не всегда оправданных функционально), усложнение синтаксиса (иногда и вопреки логике как главному инструменту научного познания) и т. п. Иными словами, для обнаружения контекстов слова, наиболее активно формирующих его научную стилистическую окраску, составителю потребуются в первую очередь наукообразно написанные тексты, поскольку собственно они и помогут автору демонстрировать типическую коммуникативную роль «я — ученый» и сигнализировать о принадлежности его разысканий к сфере науки: «здесь — наука». Этот базовый для Словаря принцип контекстуально-ролевой ориентации текста позволяет видеть в художественной речи не те средства, которые принципиально помогают автору решить эстетические задачи (это первую очередь область художественных приемов, операций, тактик текстообразования, для решения задач которого могут быть использованы абсолютно различные по своей стилистической семантике единицы), а такие, которые характеризуются устойчивым употреблением там, где автор стремится «говорить, как поэт». В этих случаях и будут наиболее активно проявляться стандартные, отработанные поэтической практикой слова и выражения — лексические и фразеологические поэтизмы, т. е. такие единицы, которые на уровне языкового сознания, словаря, до всякого текста с его уникальным заданием, фигурируют как носители поэтической семантики — в силу их способности реализовать имидж поэта и намекать на принадлежность текста к поэзии. Таким образом, теоретическая база ЖСС позволяет объединять в единый корпус тексты, наиболее определенно свидетельствующие об употребляемых в нем единицах языка как носителях стилистической окраски. Иными

словами — автоматически формировать стилистическую характеристику слова и перечень рекомендуемых для него стилистических помет.

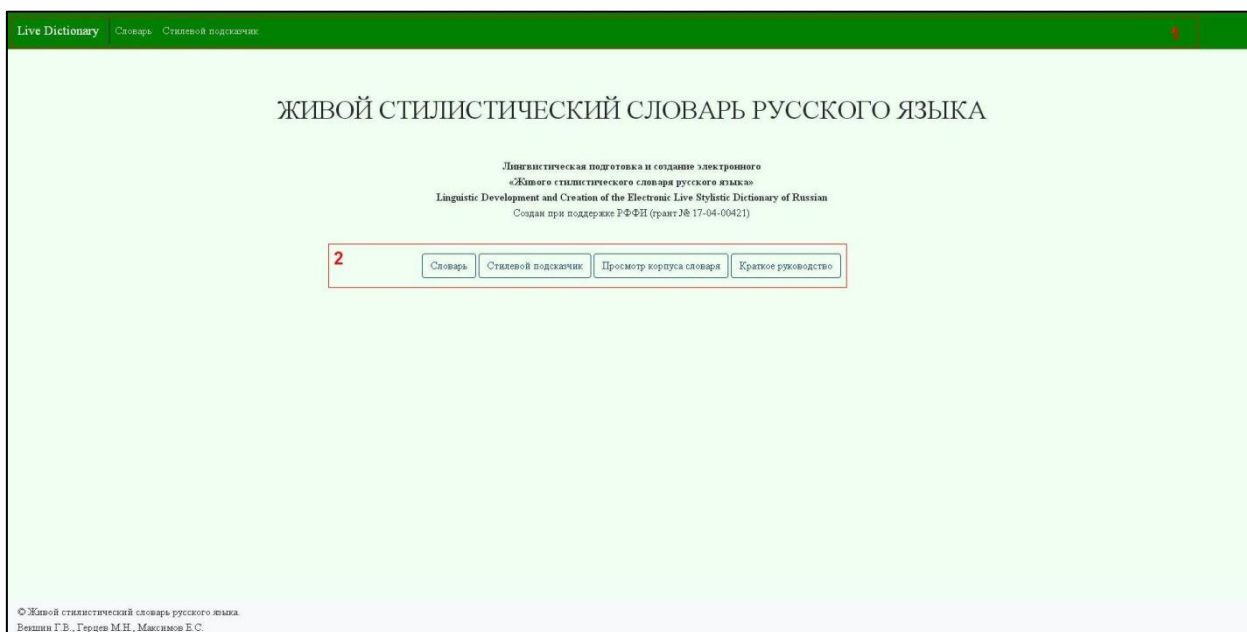
Электронных словарей-справочников, которые давали бы многомерную стилистическую характеристику слову и выражению, во всех их основных социокультурных и тональных измерениях, до сих пор не существовало не только в России, но и за рубежом.

Предоставляемая словарем информация может быть успешно использована при решении множества практических задач, связанных с выбором речевых средств и тактик общения: при **обучении языку** (как родному, так и русскому как иностранному) и при самостоятельном освоении русского языка, в преподавании русского языка и стилистики в школе и вузе, его данные могут быть использованы в **сферах идеологической и коммерческой деятельности, перевода, копирайтинга, редактирования и рерайтинга, как дополнительный инструмент для установления авторства текста др.**

Данные Словаря могут быть полезны при реализации программ **машинного перевода** (в частности, обеспечивая сопровождение перечня синонимов стилистическими пометами), в системах **автоматического реферирования/аннотирования** (automatic summarization), требующих стилистической однородности аннотации. На основе программного приложения «Стилевой подсказчик» могут быть усовершенствованы существующие и созданы новые системы **проверки стилистической правильности высказывания.**

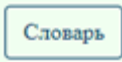
"Живой стилистический словарь русского языка" будет помощником всякому, кто стремится проверить и усовершенствовать свое знание русского языка, понять тенденции употребления слов, проверить стилистическую уместность средств речи, повысить ее эффективность.

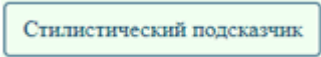
Как пользоваться Словарем

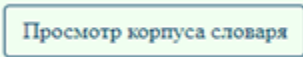


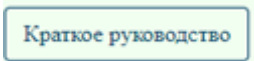
1: Главное меню

2: Кнопки перехода на основные страницы сайта 3: Различные материалы сайта

Кнопка  переводит пользователя на страницу просмотра стилистического словаря

Кнопка  переводит пользователя на страницу стилистического подсказчика

Кнопка  позволяет пользователю изучить обучающую базу данных

Кнопка  открывает данный файл

Стилистическая картина, запрашиваемого пользователем, слова или выражения строится непосредственно нейронной сетью в "live" режиме.

Словарь

Live Dictionary | Словарь | Стилистический подсказчик

Статистика по словам

1 Введите текст

Отобразить

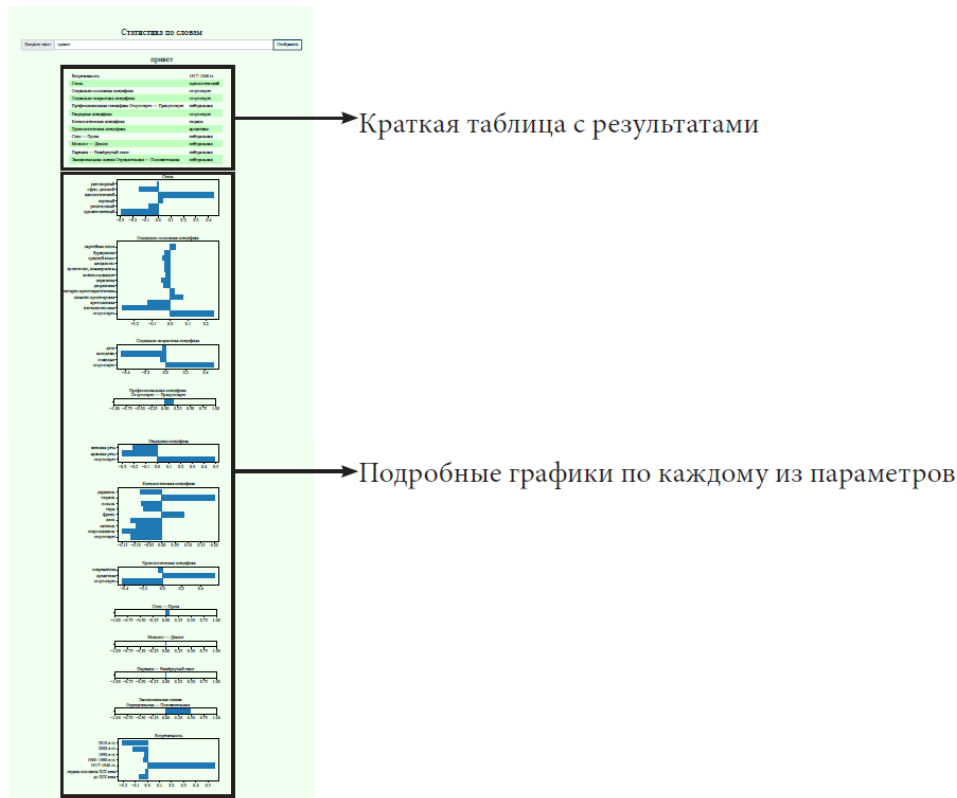
Требуется ввести слово!

2

[Посмотреть alpha версию словаря на основании технологии fasttext](#)

- 1: Поле ввода слова (или выражения)
- 2: Ссылка для перехода на alpha-версию словаря

Пример результата работы



Стилистический подсказчик

Live Dictionary | Словарь | Стилистический подсказчик

1 Стилистический подсказчик

Введите текст

Отобразить

в отличие от традиционных энциклопедий таких как энциклопедия Брокгауза и Ефрона, ни одна статья в википедии не проходит формального процесса экспертной оценки: любая статья википедии может редактироваться как с учётной записью участника так даже и без регистрации на проекте за исключением некоторых страниц защищённых частому вандализму которые доступны для изменения только определённым категориям участников или в крайних случаях только администраторам википедии и при этом все внесённые в статью изменения немедленно становятся доступными для просмотра любым пользователем поэтому википедия не гарантирует истинность своего содержимого: ведь в любом случае между моментом когда в статью будет внесена какая-то недостоверная информация и моментом когда эта информация будет удалена из статьи другим участником википедии более компетентным в данной области знания пройдет определенное время естественно для того чтобы обнаружить и удалить из статьи явный вандализм нужно намного меньше времени чем для того чтобы освободить статью от недостоверной информации когда подобная недостоверность не является особо очевидной

2

стилистика текста

Категория	Значение
разговорный	-0.8
офиц.-деловой	0.8
идеологический	0.8
научный	0.8
религиозный	0.8
художественный	0.2

1: Поле для ввода текста, и кнопка «Отобразить» для проведения его анализа стилистическим подсказчиком;

2: График общей стилистической картины текста.

Просмотр корпуса словаря

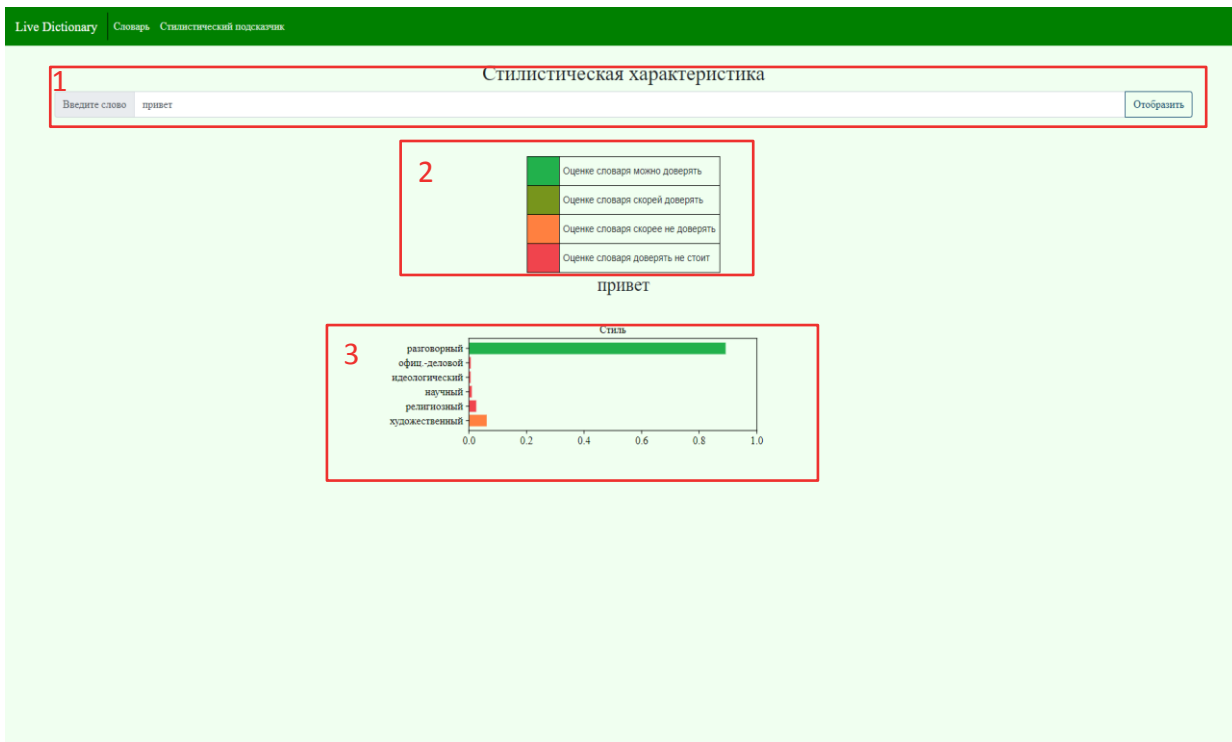
Live Dictionary		Информация							
N	user	Адрес	Автор	Название	Дата	Источник	Пол автора	Текст	
1	8	https://cyberleninka.ru/a	Ланской Г.Н.	О взаимоотношениях	11	1,5	2	8180929	
2	11	https://elitsy.ru/commun	Преподобные	Наставления старцев.	2	5	2	11181002	
3	11	http://mbrsm.ru/2018/09	Без автора	Христос – весь	11	5	2	11181002	
4	11	https://credo.press/1308	Игумен Сергей	«Покайтесь!» Открыто	11	5	2	11181002	
5	11	http://verapraslavnyaya	Святитель Игнатий	О терпении скорбей	5	2,5	2	11181002	
6	11	http://www.sedmitza.ru/	Посиф Игумен	Посифа Игумена, иже	2	2,5	2	11181002	
7	11	HTTP://WWW.PRAVOS	старец Пансий	В КАКИХ СЛУЧАЯХ	1	2,5	2	11181002	
8	11	http://podvorie-	Без автора	Житие Прп. Давида	1	5	2	11181002	
9	11	http://www.sedmitza.ru/	преподобный Посиф	О еже, како подобает	2	2,5	2	11181002	
10	11	https://txt.drevle.com/te	Петр Могила	Символ, рекше	2	5	2	11181002	
11	11	https://xn---	Без автора	Молитва 8-я Господу	1	5	2	11181002	
12	11	http://xn---	Без автора,	Святым	1	5	2	11181002	
13	11	http://mymolitva.ru/polz	Киприан	Полная молитва	2	5	2	11181002	
14	11	https://3rm.info/publicat	Преподобный	Две сети шавола.	11	5	2	11181002	
15	11	https://molitva.info/mol	Без автора,	Молитва во еже	2	5	None	11181002	
16	11	http://www.hram-	преподобный	Наставления св.	1	5	2	11181002	
17	11	http://www.biblioteka3	Амвросий Оптинский	Душеполезные	1	5	2	11181002	
18	11	http://www.biblioteka3	Свт. Игнатий	Аскетическая	4	2,5	2	11181002	
19	13	https://azbyka.ru/otchn	Епископ Мефодий	Святоотеческое	1	2,5	2	13181002	
20	12	https://azbyka.ru/otchn	Епископ Мефодий	Святоотеческое	4	2,5	2	13181002	

- 1: Кнопка для показа общей информации о базе данных (выводится на белом поле, при нажатии на которое информация убирается)
- 2: Таблица с краткой информацией о записях в базе данных
- 3: Кнопки переключения страниц (на странице выдается сто записей корпуса) 4: Номер записи. Нажатие выводит подробную информацию о данной записи.

N	user	Адрес	Автор	Название	Дата	Источник	Пол автора	Текст
1	8	https://cyberleninka.ru/a	Ланской Г.Н.	О взаимоотношениях	11	1,5	2	8180929
2	11	https://elitsy.ru/commun	Преподобные	Наставления старцев.	2	5	2	11181002
3	11	http://mbrsm.ru/2018/09	Без автора	Христос – весь	11	5	2	11181002
4	11	https://credo.press/1308	Игумен Сергей	«Покайтесь!» Открыто	11	5	2	11181002

N	user	Адрес	Автор	Название	Дата	Источник	Пол автора	Текст
1	8	https://cyberleninka.ru/a	Ланской Г.Н.	О взаимоотношениях	11	1,5	2	8180929
2	11	https://elitsy.ru/commun	Преподобные	Наставления старцев.	2	5	2	11181002
3	11	http://mbrsm.ru/2018/09	Без автора	Христос – весь	11	5	2	11181002
4	11	https://credo.press/1308	Игумен Сергей	«Покайтесь!» Открыто	11	5	2	11181002

Alpha-версия словаря



- 1: Поле ввода слова/выражения
- 2: Краткое описание расцветки графика
- 3: График оценки стилистики слова

Здесь производится оценка уровня от нуля до единицы по каждому стилю. Система предлагает оценку достоверности показателей: от достоверной до сомнительной.

Теоретическая база проекта, его методы и результаты отражены в публикациях:

1. Стилистическая семантика русского языка: Коммуникативная роль – контекст – слово : к основаниям «Живого стилистического словаря»: Монография. / Под ред. В.В. Леденевой. – М.: Издательский центр «Азбуковник», 2019. – 288 с.
2. Авамилова Э. А., Векшин Г. В., Кретов, А. А. Максимов Е. С. Алгоритмы составления корпусов «Живого стилистического словаря русского языка» и его программная разработка // Книга в современном мире: место в культурной парадигме общества в условиях цифровой революции: Сборник научных статей / науч. ред. Ж. В. Грачева. – Воронеж : ООО «Принт», 2019. С.4-19.
3. Векшин Г.В. Основы стилистической семантики: Пособие для обучающихся по направлению подготовки кадров высшей квалификации (аспирантура) 05.13.01. Системный анализ, управление и обработка информации / Моск. технологич. ун-т. – М.: Московский технологический университет, 2017. – 185 с.
4. Кретов А.А. Распределение слова по группам текстов и его стилистическая окраска, Вестник ВГУ. Серия Лингвистика и межкультурная коммуникация: Научный журнал, Воронежский государственный университет, Воронеж, 2017, 3, 5 – 15.
5. Кретов А.А. Лексические приметы религиозного стиля русского языка, «Studios Rossica Gedanensia» [Gdansk Russian Studies], Гданьск, 2017, 4, 82 – 93.
6. Векшин Г.В. Социальная среда и стилистическая семантика (к описанию лексики в проекте "Живого стилистического словаря русского языка") // Проблемы изучения живого русского слова на рубеже тысячелетий. – Воронеж, ВГПУ, 2017. – С. 9-16.
7. Шилихина К.М., Векшин Г.В.. Об источниках словника «Живого стилистического словаря русского языка» // Вестник ВГУ. Серия: Лингвистика и межкультурная коммуникация. - 2017. - № 3. - С. 16-20.
8. Векшин Г. В. , Лемешева М.М. Семантика поэтизма и стихотворческий дискурс (из опыта работы над «Живым стилистическим словарем русского языка» // Вестник ВГУ. Серия: Лингвистика и межкультурная коммуникация. 2019, № 4. С.33-43.
9. Векшин Г. В. , Лемешева М.М. Поэт как речевая роль: к семантике и прагматике русского поэтизма // Вестник РУДН. Серия: Теория языка. Семиотика. Семантика. 2019, № 4. – С. 35-44.
10. Шилихина К.М. Маркеры метакоммуникативной рефлексии говорящего в переводе // Алексеева Е.А., Корнева В.В., Полянчук О.Б., Шилихина К.М., Фененко Н.А. Языковая картина мира: особенности вербализации и ревербализации. Воронеж: Издательский дом ВГУ, 2019.
11. Шилихина К.М. Семантическое развитие прилагательных *ядовитый* и *токсичный*: что «отравляет» нашу жизнь? // Вестник ВГУ. Серия: Лингвистика и межкультурная коммуникация. 2019. №2. С. 86-93.
12. Шилихина К.М. Жанры-«двойники»: пародийные жанры в компьютерно-опосредованном дискурсе // Коммуникативные исследования. 2019. Т.6. № 3. С. 810-828.
13. Шилихина К.М. Юмористические фейковые новости как коммуникативная технология и жанр онлайн-коммуникации // Известия Волгоградского государственного педагогического университета. 2019. № 5(138). С. 139-143.
14. Векшин Г.В. Ксенологический компонент в стилистической семантике языковых единиц // Язык и мышление: Психологические и лингвистические аспекты. – М., ИЯз РАН, 2016. – С. 98 – 103.

15. Кретов А.А. Именительный единственного, правда и сингуляроиды / // Вестник Московского государственного областного университета. Серия "Русская филология". 2019, № 5.
16. Кретов А.А. Содержательная интерпретация распределения слов по длине // Информатика: проблемы, методология, технологии: сборник материалов XIX международной научно-методической конференции / под редакцией Д. Н. Борисова; Воронеж, Воронежский государственный университет, 14-15 февраля 2019 г. – Воронеж : Издательство «Научно-исследовательские публикации» (ООО «Вэлборн»), 2019. – С.1609-1614.
17. Кретов А.А. Маркемный анализ поэзии М. Кузмина // Универсалии русской литературы. 7 : сб. статей / науч. ред. : А.А. Фаустов, М. Фрайзе. – Воронеж : Издательский дом ВГУ, 2019, С.193-206.
18. Кретов А.А. О «потомках» в морфемно-деривационном контексте / А.А. Кретов // Проблемы изучения русского слова на рубеже тысячелетий: Материалы X Международной научной конференции (г. Воронеж, 26-27 октября 2019) / научный редактор А.Д. Черенкова. – Воронеж: Воронежский государственный педагогический университет, 2019, С. 73-78.
19. Леденёва В.В. Об использовании единиц эмоциональной лексики с семантикой 'недоброе, недружественное отношение'//Рациональное и эмоциональное в русском языке – 2019 [Электронный ресурс] : сборник трудов Международной научной конференции, посвящённой памяти профессора П. А. Леканта (г. Москва, 19 ноября 2019 г.) / ред. колл.: Н. Б. Самсонов (отв. ред.) и др. – Электрон. текстовые дан. (8,70 Мб). – М. : ИИУ МГОУ, 2019.
20. Векшин Г.В. Стилистическая окраска языковых единиц как дейктическое средство (К основаниям «Живого стилистического словаря русского языка») // Язык и мышление: Психологические и лингвистические аспекты.– М.: Ин-т языкознания РАН; Орехово-Зуево: ГОУ ВО МО «Государственный гуманитарно-технологический университет», ФГНУ «Психологический институт» Российской академии образования, 2018. – С. 37-41.
21. Кукушкина Е.Ю. Словотворчество в блогосфере // Труды Института русского языка им. В.В. Виноградова. Вып. 19. Материалы международной научной конференции "Вторые Григорьевские чтения: Неология как проблема лингвистической поэтики" -- М., 2019. С. 222--227.