

На правах рукописи



Семина Татьяна Алексеевна

Тональность текста: синтаксические паттерны выражения отношений между
сущностями

Специальность 10.02.19 –
Теория языка

Автореферат диссертации на соискание ученой степени
кандидата филологических наук

Мытищи – 2020

Работа выполнена в Государственном образовательном учреждении высшего образования Московской области Московском государственном областном университете на кафедре теоретической и прикладной лингвистики лингвистического факультета Института лингвистики и межкультурной коммуникации

Научный руководитель: доктор филологических наук, профессор
Максименко Ольга Ивановна

Официальные оппоненты:

Хитина Марина Викторовна, доктор филологических наук, доцент,
Московский государственный лингвистический университет, кафедра
прикладной и экспериментальной лингвистики, профессор

Габриелова Елена Валерьевна, кандидат филологических наук, Высшая
школа экономики, Департамент иностранных языков, кафедра английского
языка для экономических и математических дисциплин, старший
преподаватель

Ведущая организация: Федеральное государственное автономное
образовательное учреждение высшего образования «Российский университет
дружбы народов»

Защита состоится «22» мая 2020г. в 11 ч. 30 м.
на заседании диссертационного совета Д 212.155.04 по филологическим
наукам на базе Государственного образовательного учреждения высшего
образования Московской области Московского государственного областного
университета по адресу: 105082, г. Москва, Переведеновский переулок, д. 5.

С диссертацией можно ознакомиться в библиотеке и на сайте
Государственного образовательного учреждения высшего образования
Московской области Московского государственного областного
университета [http:// mgou.ru](http://mgou.ru).

Автореферат разослан « ___ » _____ 2020 г.

Ученый секретарь
диссертационного совета Д 212.155.04
кандидат филологических наук, доцент

Ю.А. Евграфова

Общая характеристика работы

Реферируемая работа посвящена изучению тональных отношений между сущностями. При анализе текстов с множеством упоминаемых сущностей возникает ряд трудностей при поиске компонентов мнения и установления верных тональных отношений. Под именованной сущностью понимается слово или словосочетание, обозначающее предмет или явление определенной категории. Примерами сущностей являются имена людей, локаций, геополитических организаций, то есть именованная сущность имеет четко определенного референта. В работе также рассматриваются особенности дискурса аналитических статей, влияние модальности на тональность и приводятся характеристики лексического состава аналитических статей с точки зрения тональности.

Современное развитие сети интернет позволяет пользователям каждый день создавать и выкладывать в общий доступ тексты, содержащие оценочный или тональный компонент. Ряд текстов изначально создается с целью выразить оценку, например, тексты рецензий или отзывов о товарах или услугах. Кроме подобных текстов тональность может быть выражена в блогах, микроблогах, комментариях пользователей, а также в публикациях, подготовленных журналистами. Если тексты с явно выраженной оценкой давно являются объектом исследования анализа тональности – раздела лингвистики, изучающего мнения, оценки и эмоции людей по отношению к таким объектам, как товары, услуги, события или именованные сущности, то оценка между объектами внутри текста стала интересовать исследователей только в последнее время. Целью проведения автоматического анализа тональности является выявление субъекта тональности или источника мнения, объекта тональности или цели, а также самой полярности мнения, по-другому называемой тональностью.

Актуальность исследования обусловлена необходимостью изучения выражения тональности между сущностями в текстах, что является достаточно сложной задачей. Существующие на данный момент теории описания оценки не позволяют применять их в полной мере для описания отношений между сущностями, что связано с несовпадением понятий тональности и оценки.

Научная новизна настоящей работы состоит в том, что для русского языка впервые проведен анализ средств выражения тональных отношений между сущностями в тексте. Тексты с преобладающей фактической тональностью редко становятся объектом анализа тональности, проведенное исследование показало несостоятельность некоторых существующих теорий для описания подобной тональности. В ходе работы впервые были предложены синтаксические паттерны, описывающие тональные отношения между сущностями.

Степень разработанности темы диссертации. Описанием средств выражения оценки занимались многие исследователи. Для русского языка широко известен подход при помощи анализа оценочных предикатов [Арутюнова, 1988; Сердобольская, Толдова, 2005; Сердобольская, Толдова, 2014]. Для нескольких языков, в том числе русского, созданы отдельные лингвистические ресурсы, тональные лексиконы, содержащие списки тональных единиц, которые могут не всегда коррелировать с категорией оценочности [Лукашевич, Левчик, 2016; Vaccianella, Esuli, Sebastiani, 2010; Chen, Skiena, 2014; Kochergina, 2015]. Для выявления имплицитных оценок и отношений между компонентами текстов создают тональные паттерны (лексико-семантические и синтаксические) [Entity Based sentiment analysis using syntax patterns and convolutional neural network; Turney, 2002; Almatarneh, Gamallo, 2017] или коннотационные фреймы [Rashkin, Singh, Choi, 2016; Klenner, Amsler, 2016; Лукашевич, Карнаухова, Русначенко, 2018].

В центре настоящего исследования находятся синтаксические средства, выражающие тональные отношения между сущностями в тексте. Отдельное внимание уделено выражению мнения автора, что связано с особенностями аналитического жанра публицистических текстов, которые были материалом исследования.

Теоретическая значимость настоящего исследования состоит в изучении синтаксических особенностей выражения мнения между сущностями в аналитическом тексте, то есть в тексте, не предполагающем наличие оценки в традиционном понимании. В работе проведен анализ представленности оценочных предикатов, который показал практически полное их отсутствие. При этом при проведении количественного анализа было показано, что в текстах содержится большое количество тональных лексем. Эти два факта могут указывать на преимущественное выражение тональности при помощи неоценочной лексики, что не рассматривается в современных теориях оценки. Кроме того, при анализе отношений между сущностями удалось выявить семантический класс лексем, останавливающий взаимное (или двунаправленное) тональное отношение.

Практическая значимость заключается в возможности повышения качества действующих систем автоматического извлечения мнений. Материалы диссертационного исследования включают: корпус аналитических политических текстов, отдельные списки лексики, оказывающей влияние на тональность, а также описание грамматических структур выражения мнения между сущностями. Эти данные могут применяться в системах анализа тональности, агрегирующих и анализирующих данные в режиме реального времени и отслеживающих изменения в тональности. Мониторинг тональных отношений между сущностями может быть полезным и для выявления «фейковых новостей», текстов, умышленно искажающих действительность, поскольку резкое отличие в тональности пар сущностей в тексте от среднего показателя в

определенный промежуток времени может свидетельствовать о представлении в тексте фактов, несоответствующих действительности.

Цель представленного исследования состояла в анализе корпуса статей с множественным упоминанием сущностей и выявлении паттернов выражения тональности, которые будут включать в себя компоненты мнений, тональную лексику, а также грамматические характеристики этих объектов.

В соответствии с поставленной целью исследования выдвигаются следующие **задачи**:

1. Создать корпус статей с множеством упоминаемых сущностей.
2. Провести количественный анализ расположения основных компонентов мнения относительно друг друга.
3. Проанализировать влияние контрфактивных значений на тональность.
4. Извлечь примеры конъюнктива и модальности с целью выявления влияния данных категорий на тональность;
5. Провести анализ нейтральной лексики, влияющей на тональные высказывания.
6. На основе результатов составить паттерны выражения отношений между сущностями.

Методологическую основу исследования определяют общенаучные методы анализа и синтеза, сравнения, обобщения и интерпретации научных данных. Кроме того, в исследовании применялись такие методы как количественный и качественный контент-анализ, классификация, интент-анализ, дискурс-анализ, методы лингвостатистики.

Для извлечения примеров, поиска тональных слов или предложений с заданными характеристиками не использовались специализированные программы, вместо этого был написан код на языке программирования Python 3. Он позволил проводить поиск по лемматизированным формам (приведение к нормальной форме осуществлялось при помощи пакета

PyMorphy2), составить частотный словарь для корпуса, предварительно исключив стоп-слова, выявлять тональные слова в текстах (в качестве тонального словаря использовался PyСентиЛекс), извлекать предложения с несколькими заданными характеристиками.

Объектом исследования является выражение мнений при множестве упоминаемых сущностей в текстах на русском языке.

Предметом исследования являются синтаксические и лексические особенности, оказывающие влияние на выражение мнения.

Материалом исследования послужили аналитические политические статьи сайта inosmi.ru, общий объем корпуса равен 120 статьям (103665 словоформ). Отбирались статьи, в которых присутствовало минимум пять разных именованных сущностей. Тексты являются переводами на русский язык с других языков, авторы текстов – журналисты. Для анализа были выбраны тексты с этого портала из-за большого количества доступной информации, подходящей под заданные требования. Факт перевода текстов не влияет на ход исследования, поскольку целью является описание формальных характеристик: лексики и структуры текста. Единственное ограничение, которое накладывает перевод текстов, состоит в неотделимости автора от переводчика, сопоставление оригиналов и перевода не проводилось из-за неактуальности этого анализа для настоящего исследования. Тем не менее, в связи с этим было сделано допущение, в рамках которого опускается личность переводчика и возможность его влияния на тональность в тексте. Несомненно, тональность текстов на языке оригинала и языке перевода может отличаться, но для описания структуры тональных отношений это роли не играет.

На защиту выносятся следующие положения:

1. Выражение мнений между сущностями является отдельным случаем оценки;
2. Ряд характеристик сущностей-компонентов мнения имеет закономерности в выражении;

3. Существует отдельный класс лексем, указывающий на взаимное тональное отношение между сущностями;
4. Ряд проблем анализа тональности может быть решен только при помощи создания отдельных лингвистических ресурсов.

Материал и результаты диссертационного исследования прошли **апробацию** и были представлены в докладах и выступлениях на международных практических и научных конференциях в 2017–2019 гг., в 12 публикациях автора, в том числе 4 публикациях в научных изданиях, включенных в реестр ВАК РФ.

Структура и объем диссертации соответствуют поставленным задачам. Работа состоит из введения, двух глав, заключения, списка условных обозначений, словаря терминов, списков библиографии и приложений. Каждая глава завершается выводами, в заключении приводятся основные положения и выводы исследования.

Основное содержание работы

Во **Введении** представлены основные характеристики диссертационного исследования: определены цели и задачи исследования, обоснована актуальность и научная новизна исследования, приведена научная гипотеза, инициировавшая исследование, раскрыта теоретическая и практическая значимость работы, представлен материал исследования и приведены примеры апробации полученных результатов.

В **первой главе** диссертации раскрываются ключевые теоретические понятия, такие как оценка, тональность, компоненты мнения, приведены классификации мнений и оценок. Описан анализ тональности как направление лингвистических исследований и его взаимодействие с различными уровнями языка. Характеризуются аналитические и новостные статьи как материал для исследования тональности.

Еще со времен античности изучалась языковая реализация оценки, но важным является само понимание оценки как явления. Е.М. Вольф разделяет

понятие «оценка» и «квалификация», «оценка в узком смысле слова – это противопоставление по признаку хорошо / плохо. В понятие квалификации входит собственно оценка и некоторые другие свойства, например, параметрические» [Вольф, 1985, с. 280].

По мнению Е.М. Вольф, оценка представляет собой структурно-организованную модельную рамку, «которая накладывается на высказывания и не совпадает ни с его логико-семантическим построением, ни с синтаксическим» [Вольф, 1985, с. 24]. В модельную рамку входит субъект и объект оценки, а также связывающий их предикат. Кроме того, могут присутствовать факультативные элементы модельной рамки, например, аспект, оценочный классификатор, мотивировка, средства выражения интенсификации и деинтенсификации и др. Это является показателем сложности оценочной структуры в языке [Минемуллина, 2015, с. 25–26].

Оценка может считаться наиболее ярким представителем прагматического компонента, то есть значения, приобретаемого в ситуации речи. Оценка в наибольшей степени связана с говорящим субъектом, она может отражать как его личное мнение, так и социальное влияние [Арутюнова, 1988, с.6].

Оценочные предикаты относятся к предикатам состояния. Они могут иметь заполненные валентности для субъекта и объекта оценки, или субъект может отсутствовать. Это не единственно возможные валентности для оценочных предикатов, но именно они имеют ключевое значение при выражении оценки [Сердобольская, Толдова, 2005, с. 436]. Для синтаксической теории оценочные предикаты интересны неканоническим дативным подлежащим [Сердобольская, Толдова, 2014, с. 443]. Кроме дативного подлежащего, субъект оценки может кодироваться при помощи конструкции из предлога «для» и прономинала или именной группы [Там же. С. 445]. Субъектом оценки в работах Н.В. Сердобольской и С.Ю. Толдовой [Сердобольская, Толдова, 2005; Сердобольская, Толдова, 2014], а также в работе А.В. Циммерлинга [Циммерлинг, 1999] обозначается человек,

производящий выраженную оценку. Объектом оценки считается оцениваемый объект или ситуация. Кроме этих компонентов оценки может быть представлен ориентир оценки, относительно которого описана оценочная ситуация [Сердобольская, Толдова, 2014, с. 445].

Тональность (не в трактовке дискурс-анализа), оценка, эмоция и позиция имеют важную общую характеристику: они могут быть представлены в виде субъектно-объектных отношений. Отличается только характер этих отношений.

Оценка предполагает наличие субъекта и объекта оценки, и характер выражения субъектом мнения представлен в привычном для носителей языка понимании термина *мнение*. Объектом, в таком случае, может являться предмет, ситуация, другой человек и т.д. Оценка имеет вполне устойчивые средства вербализации, например, оценочные предикаты.

Эмоция также предполагает наличие субъекта и объекта, но характер отношений между ними представлен иначе. Можно сказать, что эмоция направлена от объекта к субъекту, поскольку именно объект является причиной изменения психологического состояния субъекта или выражения им эмоции. Можно выделить средства вербализации эмоций, но явно не в том же виде, что для оценки. В устной речи эмоции представлены скорее на уровне просодии, в письменном модусе устойчивых средств вербализации меньше.

Позиция является необычным способом представления предпочтений субъекта. Она не эквивалентна тональности, но репрезентирует факт поддержки объекта субъектом.

Понятие тональности включает в себя все эти компоненты, объединяя в себе и оценку, и эмоции, так как эмоциональная и оценочная лексика фиксируется в тональных лексиконах.

Термин «анализ тональности» стал применяться начиная с 2003 г. Одной из первых работ, в которой появился термин, была научная статья Т. Насукавы и Дж. Йи, где исследовалось извлечение мнений по отношению к

заданному предмету [Nasukawa, Yi, 2003]. В этом же году вышла работа К. Дейва, С. Лоуренса и Д. Пеннока, в которой авторы говорили об «извлечении мнений» из отзывов о различных продуктах [Dave, Lawrence, Pennock, 2003].

Можно выделить четыре основных подхода к анализу тональности: анализ со словарем; по правилам; с применением машинного обучения и гибридный подход.

Анализ тональности может проводиться на уровне документа, предложения, фразы, сущностей, событий или аспектов.

Сущность (или именованная сущность) – это слово или словосочетание, предназначенное для конкретного предмета или явления, выделяющее этот предмет или явление из ряда однотипных предметов или явлений. Именованная сущность обычно пишется с заглавной буквы и обязательно имеет референт. Извлечение именованных сущностей является важным этапом обработки естественного языка для ряда задач [Можарова, Лукашевич, 2016, с. 81].

Целью в системах анализа тональности на уровне фрагментов является фрагмент текста, а не сама вершина фразы, относящейся к цели. Цель в таких системах оценивается подсчетом коэффициента сходства извлеченного фрагмента по отношению к эталонной фразе [Yang, Cardie, 2013], в то время как при анализе на уровне сущностей/событий цель оценивается по отношению к эталонному слову (т.е. вершине именной или глагольной группы), что является более точной оценкой.

И аналитические, и новостные статьи относятся к неперсонализированным текстам, но степень свободы автора в выражении собственного мнения в них различается. Если в новостных статьях это запрещено, то в аналитических автор, исходя из целей подобных текстов, может давать оценку значимости события.

Мнение между сущностями в тексте будет выражено скорее фактической лексикой, или лексикой, описывающей событие с положительным или отрицательным влиянием на сущность.

Во **второй главе** исследования описывается собранный корпус текстов с точки зрения тональной и нетональной лексики. Рассматривается расположение сущностей-субъектов и сущностей-объектов мнения относительно друг друга, приводятся паттерны, выражающие разные виды контекстов. Характеризуется класс лексики, накладывающей на сущности отношения взаимной тональности, представлены грамматические характеристики компонентов трех паттернов в этих отношениях. Приведены нетональные лексемы, инвертирующие, усиливающие или ослабляющие тональность. Также описаны особенности выражения мнения автора текста и автора цитаты. Кроме того, проанализировано влияние субъективной, объективной и иллюкативной модальности на тональность.

Характеристика лексического состава текстов

Количество тональных лексем в текстах может быть показателем наличия мнения. В собранном корпусе из 8932 уникальных слов без учета стоп-слов было выделено 1038 единиц с отрицательной тональностью и 508 единиц с положительной тональностью. Выходит, что лексический состав уникальных единиц на 17,3% состоит из тональных лексем, что является свидетельством наличия оценки в аналитических политических статьях. Подсчет лексем проводился при помощи написанного кода на языке Python, пакета Rymorphy2 [Электронный ресурс] для лемматизации, модуля регулярных выражений для предварительного удаления знаков препинания для более корректной лемматизации, списка стоп-слов для русского языка из библиотеки Natural Language Tool Kit [Электронный ресурс] и тонального лексикона RuСентиЛекс [Лукашевич, Левчик, 2016], из которого были сгенерированы списки лексем с положительной и отрицательной тональностью. Если подсчитать количество словоупотреблений положительной и отрицательной лексики, получится 5880 и 9564 соответственно (при общем числе словоупотреблений 103655).

Нейтральная лексика, как правило, не несет в себе информацию, необходимую для установления тональных отношений. Тем не менее, это не

всегда верно. Нейтральная лексика, а точнее, определенные виды лексики, могут оказывать влияние на тональную лексему, меняя степень положительного или отрицательного мнения, или же меняя его на противоположное. Под инвертированием тональности понимается изменение полярности оценки на противоположную под влиянием каких-то вербальных или структурных компонентов [Семина, 2018б].

Для анализа случаев инвертирования тональности из корпуса были автоматически извлечены предложения, содержащие хотя бы одно слово-триггер инвертирования и хотя бы одно тональное слово. В качестве набора триггеров были взяты лексемы «отменять», «отмена», «отмениться», «прекращение», «прекращать», «прекратиться», «аннулирование», «аннулировать», «аннуляция». Поиск проводился по лемматизированным при помощи пакета Rmorphy2 [Электронный ресурс] текстам, похожие списки были составлены для других классов нейтральной лексики, влияющих на тональную.

Случаи усиления и ослабления тональности имеют схожую с инвертированием структуру, которая представлена в Таблице 1.

Таблица 1 – Характеристика тональной лексики при разных типах триггера

Тип триггера	Тональное слово
существительное	именная группа в генитиве
	модификатор именной группы в генитиве
глагол	именная группа с существительным в аккузативе
	модификатор именной группы с существительным в аккузативе

Взаимное расположение компонентов мнения

К основным компонентам мнения, если говорить об анализе тональности на уровне именованных сущностей, можно отнести субъект, объект, тональное высказывание и тональность.

Анализ тональности на уровне сущностей представляет собой два этапа: поиск двух (и более) сущностей, состоящих в тональных отношениях, и определение полярности этого отношения. Для реализации первого этапа важным является взаимное расположение субъекта и объекта. После ручного анализа статей общим объемом 424 предложения, были выделены примеры мнений с разным расположением субъекта и объекта относительно друг друга. Кроме того, для каждой группы отдельно выделялись случаи пересечения границы предложения, то есть случаи, когда субъект и объект находились в разных предложениях. Полученные распределения представлены в Таблице 2.

Таблица 2 – Статистика взаимного расположения компонентов мнения

Тип	Субъект-объект	Объект-субъект
пересекает границу предложения	12,99 %	6,49 %
не пересекает границу предложения	57,14 %	23,38%

Варианты с пересечением границы предложения не так просты для анализа, в таких случаях сложно говорить о наличии закономерностей в выражении мнений. При этом в предложениях с сущностями могут присутствовать кореферентные сущностям языковые единицы. Для порядка объект-субъект характерна одна из следующих ситуаций:

1. Пассивный предикат;
2. Субъект оценки в дативе с предлогом «для»;
3. Тональное высказывание со значением «испортить отношение субъекта тональности к себе»;
4. Прямая или косвенная речь, где субъектом тональности является автор цитаты.

При субъекте, предшествующем объекту, представлено большее разнообразие в организации структуры мнения. Могут быть представлены и

структуры с отсутствующим субъектом или объектом, но подобные случаи не рассматриваются в рамках данного диссертационного исследования. В случаях употребления в пределах предложения и субъекта, и объекта, можно выделить закономерные характеристики компонентов. В Таблице 3 представлены падежи именованных сущностей, занимающих слоты в структуре мнения.

Таблица 3 – Характеристика сущностей при расположении субъект-объект

ИС – субъект	ИС – объект
является подлежащим (номинатив)	датель
генитив (с предлогом «у»)	генитив
номинатив + ИС в генитиве	датель + ИС в генитиве
датель (с предлогом «для»)	номинатив в придаточной клаузе
	аккузатив

Взаимная тональность между сущностями

Взаимная тональность – ситуация, в которой между субъектом и объектом мнения есть взаимное тональное отношение с одной полярностью оценки. Выделение взаимной тональности в отдельную группу случаев выражения мнения позволит найти в тексте часть имплицитной информации. Встает вопрос о способах выражения такой тональности. Соответственно, на некотором отрезке текста есть два компонента мнения, при этом их статус может быть разным; в ряде случаев можно говорить о равноправии элементов, в других – один элемент скорее является субъектом, тем не менее, объект также имеет оценочное направленное отношение по отношению к субъекту.

Для поиска примеров взаимной тональности в корпусе был составлен ядерный набор лексем-показателей и проведен поиск по корпусу, после чего список был расширен. В ядерный набор вошли такие слова как «отношения», «связь», «сотрудничество», «разногласие», «союзник», при этом, все слова, кроме «отношения» искались по документу, в котором слова были приведены в нормальную форму при помощи лемматизации через пакет

PyMorphy2 для Python с предварительной очисткой пунктуации при помощи регулярных выражений. Ядерный набор был составлен в ходе просмотра частотного словаря для корпуса текстов.

Выделено три паттерна взаимной тональности в зависимости от расположения компонентов мнения (две именованные сущности и лексема-показатель взаимной тональности). Наиболее частотным оказался паттерн, в котором показатель предшествует сущностям. В Таблице 4 представлена структура паттерна и требуемый падеж именованных сущностей.

Таблица 4 – Структура паттерна SOOS – Named Entity – Named Entity

Окружение	ИС {падеж}	Коннектор	ИС {падеж}
конфликт, взаимодействие, союз, договоренность, отношения, разногласия, напряжение	-		NE {gent} NE {gent}
	между	и	NE {abl}
сотрудничество, связь, отношения, конфликт, взаимодействие	NE {gent}	с	NE {abl}

Следующим по встречаемости оказался паттерн, в котором показатель взаимной тональности расположен между именованными сущностями, в нем лексемой-показателем может быть и глагол (Таблица 5).

Таблица 5 – Структура паттерна Named Entity-SOOS-Named Entity

Именованная сущность {падеж}	Окружение	Именованная сущность {падеж}
NE {nomn}	сотрудничать, дружить с	NE {abl}
NE {nomn}	Verb партнерство, взаимодействие, отношения, связи, сотрудничество, дружба с	NE {abl}
NE {nomn}	союзник, партнер, враг, друг	NE {gent}

Наиболее редким случаем расположения лексемы взаимной тональности относительно именованных сущностей, участвующих в тональном отношении, является предшествование обеих сущностей показателю (Таблица 6).

Таблица 6 – Структура паттерна Named Entity - Named Entity – SOOS

ИС {падеж}	Коннектор	ИС {падеж}	Окружение	
NE {nomn}	и	NE {nomn}	Verb	Связи Сотрудничество
			Сотрудничать	
NE {datv}	и	NE {datv}	Modal verb, Verb	Связи Сотрудничество Партнерство Взаимодействие

Особенности выражения мнения автора

При анализе аналитических статей можно увидеть, что ряд тональных отношений имеет в качестве субъекта не сущность внутри текста, а автора текста, что не подразумевается жанром информационной журналистики. Тем не менее, автор текста не выражает мнение напрямую, используя эксплицитные оценочные конструкции, вместо этого применяется ирония, сарказм и стилистически сниженная лексика [Семина, 2019б]. Примером подобной лексики будет выражение «ужас, летящий на крыльях ночи» и «с гулькин нос» в примере:

Между тем резервы эти и так с гулькин нос, а перспектива резкой девальвации рубля после деноминации, призванной укрепить к нему доверие, — это для белорусского руководства просто ужас, летящий на крыльях ночи [Класковский, 2016, Электронный ресурс].

Нужно отметить, что автор текста также может считаться именованной сущностью, потому что для текста публикации известен автор статьи или издание, опубликовавшее текст. Автор статьи имеет явно определенного

референта в реальности, что касается и названия газеты или интернет-ресурса, следовательно, они являются именованными сущностями и на этом уровне анализа тональности могут быть объектом исследования. Отличие от остальных сущностей заключается в том, что мнение автора может быть высказано в статье в любом месте, при этом использование эксплицитных оценочных предикатов не представляется возможным.

Так как для части статей осуществлялась ручная разметка тональных отношений, был проведен анализ частотности выражения мнения автором.

На рисунке 1 представлено отношение статей, содержащих мнение автора, ко всем статьям, для которых была проведена разметка. Можно увидеть, что в 25% статей был минимум один случай выражения мнения автором.

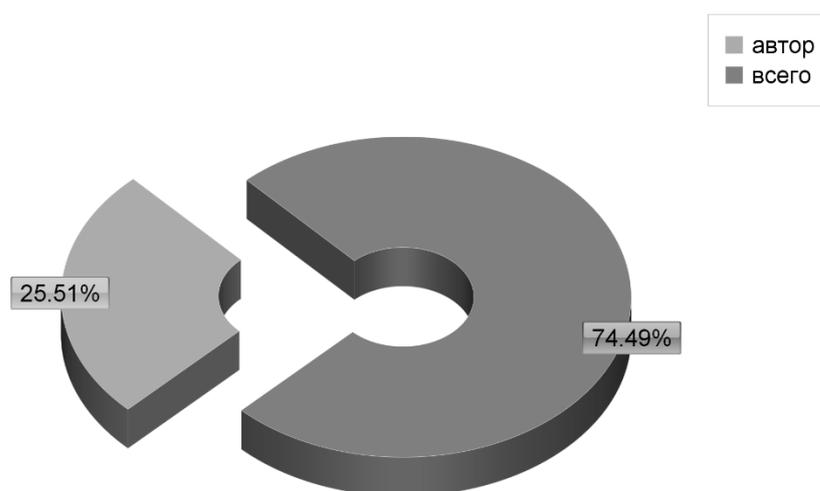


Рисунок 1

При выражении мнения автор обычно использует нетипичную для стиля аналитических статей лексику. Метрика idf позволит найти слова, которые редко встретились не в языке, а в отобранных текстах [Семина, 2019б]. Для этой задачи можно использовать и простой поиск редких слов в коллекции на основе частотного словаря корпуса. Нужно отметить, что между единичным упоминанием в корпусе и упоминанием только в одном документе есть различия. Преимуществом обратной частоты документа

является возможность привлечения показателя частоты термина в документе (tf), что могло бы быть полезным в поиске текстов, тематически отличающихся от основного массива.

Прямая речь

Стоит выделить возможные ситуации выражения прямой речи, которые могут оказать влияние на тональность:

1. автор прямой или косвенной речи может высказывать собственное мнение о сущностях,
2. автор мнения может передавать факты, показывающие отношения сущностей друг к другу,
3. автор может высказывать отношения о сущностях, не являющиеся истинными [Семина, 2019д].

В первом случае автор прямой речи может высказывать свое мнение о некоторой именованной сущности, тогда целесообразно устанавливать его субъектом мнения во всех тональных отношениях, как в примере:

«Многие члены НАТО не платят по счетам, не делают то, что должны», – заявил миллиардер [= Дональд Трамп] [«Агент Путина»: рассуждения Трампа напугали американские СМИ, 2016, Электронный ресурс].

В примере выше автор прямой речи высказывает мнение об именованной сущности *НАТО*, других тональных отношений в предложении нет.

При упоминании нескольких сущностей внутри прямой речи сложнее выделить тональные отношения. Немаловажным является установление истинности высказывания и отнесенности его к реальному миру. Высказывание может быть, например, описанием желательного события для автора высказывания, пример:

Министр обороны Италии Пинотти: «В борьбе с терроризмом НАТО сблизится с Москвой, Европа станет более интегрированной» [Валентино, 2018, Электронный ресурс].

В примере были указаны фактические события, говорится о факте борьбы с терроризмом, но отношения НАТО и Москвы в примере нельзя считать фактическими, так как лишь говорится о их возможном установлении.

Тональность при прямой и косвенной речи может выражаться по-разному, при упоминании в речевом фрагменте только одной именованной сущности можно установить тональные отношений между автором цитаты и сущностью в цитате. При наличии нескольких сущностей и тональных отношений между ними, тональные отношения устанавливаются по обычным правилам. Если некоторые сущности не находятся с другими сущностями внутри цитаты в тональных отношениях, можно установить отношения между ними и автором речевого высказывания.

Влияние модальности на тональность текста

Проанализировав случаи выражения модальности в текстах, удалось установить влияние каждого типа модальности на тональность.

Субъективная модальность наиболее сильно коррелирует с тональностью, что и ранее отражалось в исследованиях, но проведенный анализ подтвердил наличие модальных единиц со значением пропозициональной установки, соответствия ожиданию, онтологическую, деонтическую и эпистемическую возможность. Модальность может быть как показателем мнения автора текста или цитаты, она может затрагивать и отношения между другими сущностями.

Объективная модальность не прямо связана с тональностью, но является важным аспектом для осмысления необходимости выражения этого мнения и прохождения мнением критериев истинности. Для анализа объективной модальности были рассмотрены примеры с сослагательным наклонением, анализ показал наличие большого количества нейтральной модальности, что может объясняться характеристиками аналитической статьи как жанра. Примеры контрфактивности были представлены мало, но в них тональность могла меняться на противоположную, потому что само

понимание термина «контрфактивное значение» говорит о несоответствии высказывания критериям истинности и фактам реального мира.

Иллокутивная модальность оказывает влияние на тональность не во всех случаях. При ассертивной модальности сложно говорить о ее прямом влиянии на тональность, скорее при этом типе модальности тональность выражается только при помощи лексики и структурного оформления высказывания.

Экскламатив является эксплицитным указанием на тональность или эмоциональность фрагмента текста. В примере ниже показано неравнозначное отношение к ЕС и НАТО – по мнению автора статьи, выход из НАТО является более серьезным шагом, и вызывает более эмоциональную реакцию, что выражено только при помощи экскламатива.

Президент Чехии Милош Земан, который уже давно крестится в сторону России, заявил, что он не согласен с теми, кто агитирует за выход из ЕС, но сделает все, чтобы организовать для них референдум, на котором будет возможность высказаться за выход из ЕС и (!) НАТО [Гордон, 2016, Электронный ресурс].

Императивная модальность была описана про помощи скорректированной теории речевых актов, различные виды актов в разной степени связаны с тональностью.

Вопросительная иллокутивность может быть показателем авторского стиля или следствием формата ведения дискурса, если статья имеет фрагменты интервью, в таком случае сложно говорить о прямой связи с тональностью. Тем не менее, в ряде случаев вопросительная иллокутивная модальность является прагматическим выбором автора, цель которого состоит в дополнительном привлечении внимания читателей к тексту статьи.

У вас есть сосед, с которым вы хотели бы подружиться, но с которым при этом бывает невероятно тяжело? Сосед, который может быть действительно приятным человеком, если удастся с ним поладить, но

ссориться с ним страшно, тем более что он уже настроил против себя весь квартал? [Клеберг, 2016, Электронный ресурс].

Вопрос соотнесения вопросительной иллокуции и тональности является неоднозначным. С одной стороны, употребление вопросов может быть обосновано типом дискурса, с другой – это прагматическое намерение автора текста, который хочет вступить с читателем в своеобразную дискуссию, тем самым привлекая большее внимание к фрагменту текста. Во втором случае можно говорить о корреляции с тональностью, поскольку этот тип иллокутивной модальности сознательно применяется к определенному типу информации, следовательно, может иметь место авторское мнение по отношению к этой информации, к персонам или к описываемой ситуации.

Все виды модальности могут пересекаться в пределах одного предложения и оказывать влияние на одно тональное отношение.

В **Заключении** приводятся основные итоги проведенного исследования и намечаются перспективы для дальнейшей работы, связанные с изучением других аспектов проблемы.

Список использованной литературы включает перечень научных трудов, электронных ресурсов и ресурсов иллюстративного материала, использованных в ходе исследования.

В **Приложениях** представлен фрагмент частотного словаря для корпуса, контексты паттернов взаимной тональности и пример представления текста в виде частот лексем.

Основные положения диссертации отражены в следующих публикациях:

Статьи, опубликованные в рецензируемых изданиях, рекомендованных ВАК Минобрнауки РФ:

1. Семина Т.А. Дихотомия субъективность vs. объективность и тональная релевантность в задачах анализа тональности // Вестник

Московского государственного областного университета. Сер.: Лингвистика. М.: ИИУ МГОУ, 2018. № 1. С. 38–45.

2. Максименко О.И., Семина Т.А. Создание корпуса текстов для анализа тональности // Ученые записки Национального общества прикладной лингвистики. М.: НОПриЛ, 2019. С. 106–113.

3. Семина Т.А. Извлечение мнения автора через обратную частоту документа // Вестник Московского государственного областного университета (Электронный журнал), 2019. № 2. С. 171–179.

4. Семина Т.А. Анализ выражения мнения в письменном политическом медиадискурсе // Вестник Московского государственного областного университета (Электронный журнал), 2020. № 3.

Публикации в других научных изданиях:

5. Семина Т.А. Автоматическое выявление отношений между сущностями в мультиобъектных текстах // Человек в информационном пространстве: понимание в коммуникации: сб. научн. тр. / под общ. ред. Н.В. Аниськиной, Л.В. Уховой. В 2 тт. Ярославль: Изд-во ЯГПУ, 2017. Т. 1. С. 304–309.

6. Семина Т.А. Анализ тональности в текстах на политическую тему на примерах Twitter и аналитических статей // Большое евразийское партнёрство: лингвистические, политические и педагогические аспекты: сб. ст. по материалам Международн. научно-практич. Конф. (г. Москва, 15 декабря 2017 г.) [Электронный ресурс] / редкол.: И.Ф. Беляева (отв. ред.), Е.П. Савченко (отв. сек.). Электрон. текстовые дан. (21,2 Мб). М.: ИИУ МГОУ, 2018. С. 749–754.

7. Семина Т.А. Влияние нейтральной лексики на тональные высказывания // Военно-гуманитарный альманах. Сер.: Лингвистика: материалы XII Международн. научн. конф. по актуальным проблемам языка и коммуникации. Язык. Коммуникация. Перевод (г. Москва, Военный

университет, 29 июня 2018 г.) / под общ. ред. Н.В. Иванова. М.: ИД «Международные отношения», 2018. Вып. 3. Том 2. С. 85–90.

8. Семина Т.А. Триггеры инвертирования тональности в средствах массовой коммуникации // Язык. Коммуникация. Культура–2018: сборник научных материалов II Международной научно-практической конференции молодых ученых. М.: Гос. ИРЯ им. А.С. Пушкина, 2018. С. 106–107.

9. Семина Т.А. Гендерный аспект при анализе тональности речи политиков // Профессиональная картина мира: кросс-культурный диалог: материалы Международн. научно-практич. конф. Москва, 22–23 ноября 2019 г. / ред. колл.: Е.Ф. Тарасов (отв. ред.), С.В. Мыскин, В.П. Синячкин. М.: ООО «Агентство социально-гуманитарных технологий», 2019. С. 315–317.

10. Семина Т.А. Корпусные исследования и анализ тональности // Слово. Словарь. Термин. Лексикограф [Электронный ресурс]: сб. ст. по материалам Международн. научно-практич. конф. памяти д. филол. н., проф. Юрия Николаевича Марчука (г. Москва, 1–2 марта 2019) / ред. колл.: отв. ред. И.И. Валуйцева, отв. сек. Е.П. Савченко, ред. англ. текста И.А. Улиткин; под общ. ред. И.И. Валуйцевой. Электрон. текстовые дан. (15,6 Мб). М.: ИИУ МГОУ, 2019. С. 560–566.

11. Семина Т.А. Описание языковых единиц для анализа тональности // IV Фирсовские чтения. Язык в современных дискурсивных практиках: материалы док. и сообщ. Международн. научно-практич. конф. (Москва 22–23 октября 2019 г.) / под ред. А.С. Борисовой, А.В. Игнатенко, Т.В. Лариной, О.В. Ломакиной. М.: РУДН, 2019. С. 295–299.

12. Семина Т.А. Проблема прямой речи при извлечении мнений // Наука на благо человечества–2019: материалы Международн. научн. конф. молодых учёных (МГОУ, г. Москва, 15–26 апреля 2019 г.): ст. преп. и асп. [Электронный ресурс] / отв. ред. и сост. Е.А. Певцова. Электрон. текстовые дан. (9,46 Мб). М.: ИИУ МГОУ, 2019. С. 61–65.