

УДК 81-114.2

Семина Т. А.

Московский государственный областной университет

141014, Московская обл., г. Мытищи, ул. Веры Волошиной, д. 24, Российская Федерация

АНАЛИЗ ВЫРАЖЕНИЯ МНЕНИЯ В ПИСЬМЕННОМ ПОЛИТИЧЕСКОМ МЕДИАДИСКУРСЕ

АННОТАЦИЯ

Целью работы является описание особенностей выражения мнения в аналитических политических статьях для выделения тональности.

Процедура и методы. Проведены анализ языкового материала и статистический анализ взаимного расположения основных компонентов высказывания, содержащего мнение, рассмотрено взаимодействие тональности и контрфактивного значения, характеристики отсылки к чужому мнению, а также описан подход к поиску мнения автора по редким для коллекции документов словам.

Результаты. На основе анализа выражения мнения удалось выделить лексические и грамматические особенности выражения тональности в политическом медиадискурсе.

Практическая значимость. В статье сформулированы предложения по работе с указанными особенностями при выделении мнений в аналитических статьях.

КЛЮЧЕВЫЕ СЛОВА

анализ тональности, мнение, модальность, прямая речь, извлечение мнений

СТРУКТУРА

[Введение](#)

[Материал исследования и методы](#)

[Обсуждение результатов](#)

[Выводы](#)

T. Semina

Moscow Region State University

24 Very Voloshinoy ul., Mytishchi 141014, Moscow Region, Russian Federation

ANALYSIS OF EXPRESSION OF OPINION IN WRITTEN POLITICAL MEDIA DISCOURSE

ABSTRACT

Aim is to describe the features of the expression of opinion in analytical articles on politics.

Methodology. The language and the arrangement of the text main components expressing opinion, were analyzed, the interaction of sentiment and counterfactual values, characteristics of referring to someone else's opinion have been considered, as well as the way to find the author's opinion using the words that are rare for a collection of documents.

Results. Identification of lexical and grammatical features used to express opinion in political media discourse.

Research implications. Suggestions for using these characteristics in identifying opinions.

KEYWORDS

sentiment analysis, opinion, modality, direct speech, opinion mining

ВВЕДЕНИЕ

Анализ тональности – область изучения мнений, оценок и эмоций людей по отношению к персонам, объектам или их атрибутам, иногда называется сентимент-анализом или извлечением мнений. Большинство исследователей выделяют во мнении несколько ключевых компонентов: объект или источник, т. е. тот, кому принадлежит мнение; объект или цель, к этому компоненту относится предмет или персона, мнение о котором высказывается; тональность. Под тональностью понимается полярность мнения. Она может быть бинарной, когда мнение признаётся только положительным или отрицательным, тернарной, если добавляется нейтральная оценка, или градуированной, когда вместо набора тэгов для определения тональности задаётся шкала оценивания.

Как правило, материалом для анализа тональности являются тексты из интернета, содержащие явную оценку автором текста некоторого объекта; к текстам такого рода относят рецензии, отзывы, блоги и микроблоги. В последнее время исследовательский интерес сместился в сторону новостных текстов и аналитических статей, так как они содержат не только мнение автора текста, но и мнения сущностей по отношению к другим именованным сущностям. Именованная сущность в современной лингвистике – слово или

словосочетание, обозначающее предмет или явление определённой категории. Примерами именованных сущностей являются персоны, локации, государства и т. д. Универсального перечня именованных сущностей, принятого среди исследователей, нет [3, с. 2].

Анализ тональности текстов на русском языке является широко известной проблематикой, тем не менее нельзя говорить даже о сравнении масштабов исследований в области русского и английского языков. Это связано с дефицитом доступных размеченных корпусов текстов для машинного обучения, что является необходимым фактором для создания прикладной системы [4, с. 55].

Анализ тональности имеет разнообразный спектр применения: он может проводиться на материале отзывов для получения информации о степени удовлетворённости товаром покупателями [7], может быть методом оценки репутации бренда или человека, материалом могут быть тексты твитов, блогов или СМИ [5], наряду с количественным контент-анализом может применяться для лингвистической оценки речей политиков [1, с. 455] или для предсказания вероятности успеха бизнес-стартапа [9].

Аналитические политические статьи являются непростым материалом для анализа тональности по ряду причин: во-первых, в соответствии с жанром информационной журналистики автор текста не имеет права открыто, эксплицитно выражать собственное мнение об обсуждаемых в статье персонах или событиях; во-вторых, большое количество именованных сущностей позволяет устанавливать тональные отношения между ними; в-третьих, тональность выражается не оценочной лексикой, а фактической [8]. В последнем случае можно говорить об изменении понимания термина «мнение», которое в современном анализе тональности включает в себя и так называемые *good for/ bad for events* – действия одного объекта по отношению к другому, которые положительно или отрицательно влияют на объект тональности [6].

МАТЕРИАЛ ИССЛЕДОВАНИЯ И МЕТОДЫ

Материалом исследования послужили статьи с сайта *inosmi.ru* – веб-ресурса, на котором публикуются переводы статей иноязычной прессы. Стоит отметить, что, несомненно, при переводе может искажаться смысл и даже меняться тональность текста или его фрагмента, тем не менее при рассмотрении формальных составляющих выражения мнения, таких как грамматические и лексические особенности, соотнесение текста с действительностью и текстом на языке оригинала не обязательно. Кроме того, грамматические и лексические особенности, выделенные в переводах, будут применимы и к русскоязычным статьям. В ходе исследования удалось вывести несколько характерных черт выражения мнения, которые следует учитывать при про-

ведении автоматического анализа тональности. Данные характерные черты можно разделить на лексические и синтаксические, хотя деление и является весьма условным.

К методам, использованным в данном исследовании, относятся *описание, статистический анализ, классификация*.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

К лексическим особенностям относится эксплицитное употребление тональных высказываний [10, р. 575], однако в аналитических политических статьях фактическая лексика по количеству будет превалировать над оценочной. Найти лексику такого рода несложно, возможно применение тонального лексикона, в котором она будет иметь маркировку полярности тонального отношения. Имея подобный лексикон, можно создать систему анализа тональности, которая будет считать количество положительных и отрицательных слов по заданному списку тональных высказываний и выдавать предположение о тональности. Создание такого рода системы не является трудной задачей. Более интересным случаем является поиск взаимной тональности, которую выражает некоторая лексика.

1) *В то время как Соединенные Штаты и Россия немного приблизились к расширению военного сотрудничества на земле в Сирии...¹*

В примере 1 слово *сотрудничество*, несомненно, является положительно окрашенным тональным высказыванием, тем не менее оно показывает сразу два отношения, потому что в этом предложении говорится об улучшении как отношения России к США, так и наоборот. Учёт лексики, накладывающей на пару сущностей взаимную тональность, поможет в выявлении имплицитной тональности, т. е. тональности, не имеющей явного тонального высказывания. Подобной лексики немного, в настоящее время мы занимаемся сбором лексем-показателей взаимной тональности.

Другим важным лексическим фактором, связанным с выражением мнений, является нейтральная лексика, влияющая на тональность. Существует ряд лексем, которые невозможно отнести к ряду тональных, потому что они не выражают ни оценки, ни фактической информации, влияющей на некоторый объект. Однако данная лексика может менять степень тональности высказывания (усиливать или ослаблять) или менять полярность высказывания на противоположную.

2) *Как Россия будет использовать Брексит для отмены санкций²*.

¹ Рогин Д. Керри похвалил позицию России в отношении группировок в Сирии // ИноСМИ.Ру. [2016]. URL: <https://inosmi.ru/politic/20160713/237169965.html> (дата обращения: 13.11.2019).

² Таллиз Д. Как Россия будет использовать Брексит для отмены санкций // ИноСМИ.Ру. [2016]. URL: <https://inosmi.ru/politic/20160706/237087099.html> (дата обращения: 13.11.2019).

В примере 2 *отмена* – нейтральная лексема, тем не менее она инвертирует негативную тональность лексемы *санкции*. Если случаи с усилением или ослаблением тональности могут не играть роли при бинарной (положительная и отрицательная) классификации, то в случаях градуальной оценки эти лексемы могли бы давать дополнительную информацию для анализа. В наших текстах только со словом *отмена* было 14 предложений, 13 из которых относятся к описываемой нами ситуации. В дальнейшем будет расширен список слов, инвертирующих, усиливающих или ослабляющих тональность.

В число лексических особенностей можно включить и выбор лексики для выражения своего мнения. Жанр информационной журналистики не позволяет автору эксплицитно высказывать свою позицию, кроме того, мнение автора в таких текстах высказывается в меньшем количестве случаев по сравнению с мнениями сущностей друг о друге. Тем не менее журналист может выразить своё мнение, выбирая специфическую лексику для описания некоторых событий или явлений.

3) *Между тем резервы (1.2759) эти и так с гулькин (2.1790) нос (1.7019), а перспектива (0.5258) резкой (0.6349) девальвации (1.5769) рубля (1.0998) после деноминации (2.1790), призванной (0.9237) укрепить (0.3936) к нему доверие (0.7988), – это для белорусского (1.7019) руководства (0.5355) просто ужас (1.7019), летящий (2.1790) на крыльях (2.1790) ночи (1.4800). Так что для спасения (1.4800) от обвала (1.4008) очередной (0.8568) конфликт (0.3728) с Москвой надо бы уладить (2.1790) как можно скорее (0.6349). ... Капитуляция (1.7019) в этом вопросе (0.1456) может лишь раззадорить (2.1790) Кремль, станут (0.0998) дожимать (2.1790) и в другом (0.0998)*³.

В примере 3 выделена лексика, выражающая мнение автора по отношению к именованным сущностям. В приведённом примере (в корпусе встречались похожие данные) мнение выражено употреблением лексики, редкой для подобного жанра статей, и найти такую лексику, редкую в корпусе, можно при помощи метрики обратной частоты документа (*idf*). В примере после слов указаны показатели *idf* для всех слов, за исключением стоп-слов. Можно заметить, что у редкой лексики, выражающей мнение автора, этот показатель выше, чем у остальной. Максимальная *idf* = 2.1790 в наших расчётах (слово *гулькин*) встретилась в нашем корпусе только в одной статье. Подобный поиск подойдёт только для монотематичного корпуса, например, максимальную *idf* в нашем корпусе получило слово *деноминация*, это обусловлено тематическим отклонением в той статье. Если бы мы учитывали и частоту использования термина в документе, этого можно было бы избежать [2].

За рамки оценки лексики выходят случаи, влияющие на поиск тональности и на сам факт необходимости установления тонального отношения. Для начала

³ Класковский А. Москва будет выбивать у Лукашенко газовый долг // ИноСМИ.Ру. [2016]. URL: <https://inosmi.ru/economic/20160714/237186452.html> (дата обращения: 13.11.2019).

стоит отметить вариации расположения субъекта и объекта тональности по отношению друг к другу. Чаще субъект предшествует объекту, при этом субъект и объект находятся в пределах одного предложения. В таблице 1 представлены результаты ручной разметки 10 статей: были найдены примеры, в которых эксплицитно выражено мнение одной сущности о другой, при этом, кроме порядка *субъект – объект* или *объект – субъект* учитывалось и расположение компонентов мнения в пределах одного предложения.

Таблица 1 / Table 1
Распределение мнений / Opinions distribution

	Субъект – объект	Объект – субъект
Пересекает границу предложения	12,99%	6,49%
Не пересекает границу предложения	57,14%	23,38%

Несмотря на превалирование порядка *субъект – объект*, обратное расположение компонентов встретилось примерно в 30% случаев, что является достаточно серьёзным количеством для анализа грамматических характеристик, объясняющих подобное расположение. До настоящего времени не был проведён глубокий анализ подобных примеров, тем не менее можно указать, что к ним часто относятся случаи прямой речи, так же как и использование предиката в страдательном залоге.

Прямая речь сама по себе является интересным явлением. В статьях нередко приводят цитаты официальных представителей или руководителей государств. Прямая речь также может содержать мнение именованной сущности о другой именованной сущности. В подобных ситуациях возникает вопрос о целесообразности установления тональных отношений между сущностями, ведь они могут не отражать реального расположения дел. В работе мы не делаем различий между реальными и нереальными событиями, однако при создании рабочей автоматической системы это может стать серьёзной проблемой, ведь автор цитаты может лишь высказывать своё мнение о тональных отношениях между двумя сущностями, но не объективно передавать его. Важной для анализа является ситуация, в которой субъектом мнения будет автор прямой речи, в качестве объекта будет представлена другая именованная сущность, подобные примеры попали в таблицу 1 в столбец «объект – субъект».

Кроме прямой речи, может быть и отсылка на мнение или слова другого человека, обычно выраженные при помощи косвенной речи. Можно сказать, это частный случай прямой речи, так как поиск, анализ и последующая правильная автоматическая обработка данных примеров позволят раскрыть те же аспекты мнений, что и прямая речь.

Модальность и условные предложения являются отдельными феноменами для изучения, так как в них выражается и мнение автора, и неуверенность в реальности отношения между сущностями. Сослагательное наклонение является полисемичным, в русском языке оно выражает события, не относящиеся к реальным, которые выражают предположения автора или именованной сущности об отношениях между сущностями.

4) *А если Турция активизирует борьбу с «Исламским государством»⁴ в Сирии, Вашингтон может снизить свою поддержку курдских группировок⁵.*

Контрфактивное высказывание в примере 4 выражает несколько мнений именованных сущностей (табл. 2). Мнения выписаны в виде кортежа <субъект тональности; объект тональности; тональность или изменения тональности; эксплицитное тональное высказывание, если оно выделяется в примере>. Вместо «курдских группировок» в таблице указана «Турция» – это связано с тем, что *курдские группировки* также должны являться возможным объектом тональности, поэтому для иллюстрации они заменены на эквивалентную именованную сущность.

Таблица 2 / Table 2
Тональная структура предложения / Sentiment structure of the sentence

Субъект	Объект	Тональность	Тональное высказывание
Турция	Исламское государство	Отрицательная тональность	Активизирует борьбу
Вашингтон	Турция	Снижение тональности по шкале	Может снизить свою поддержку

Первое отношение выражает негативное значение, второе отношение снижает положительную тональность в отношении пары *Вашингтон – Турция*. Снижение тональности было бы особо полезным при небинарной оценке тональности, тем не менее даже в случае классификации мнений на положительные и отрицательные эта информация имеет большое значение. Отдельно стоит отметить, что некоторые лексемы тональных высказываний в таблице 2 являются оценочными, они внесены в тональные лексиконы. Встаёт вопрос о необходимости выделения этих тональных отношений: они

⁴ Организация запрещена на территории Российской Федерации.

⁵ Бэндоу Д. Турция помирилась с Россией, Америке тоже надо прекратить новую холодную войну с Москвой // ИноСМИ.Ру. [2016]. URL: <https://inosmi.ru/politic/20160706/237092665.html> (дата обращения: 13.11.2019).

не связаны с реальным миром, компоненты высказывания не прошли бы проверки критериями истинности, следовательно, этих отношений нет. Модальность похожим образом влияет на тональность, однако эти случаи сложнее из-за большей многозначности модальности в языке. С одной стороны, модальность является категорией, выражающей отношение говорящего к содержанию высказывания, тем самым показывая тональность. С другой стороны, модальность может быть показателем отношения содержания высказывания к действительности, делая нежелательным выделение тональности в ситуациях ложного содержания.

ВЫВОДЫ

Особенности выражения мнений при проведении анализа тональности зависят от ряда факторов, в первую очередь от материала. Аналитические политические статьи обладают рядом особенностей, которые требуют особого внимания при обработке. Если некоторые особенности лишь затрудняют установление мнений, то другие меняют полярность оценки или в принципе делают выделение отдельного мнения нежелательным, как в случаях с несоответствием содержания действительности.

К лексическим особенностям можно отнести высказывания с взаимной тональностью и нейтральную лексику, влияющую на тональные высказывания. При предварительном анализе этой лексики можно расширить число тональных отношений, частично покрыв имплицитную тональность, и избежать части ошибок классификации тональных отношений.

Синтаксическими особенностями считаются вопросы прямой речи и отсылки к чужому мнению, а также модальность и условные предложения; последние относятся к контрфактивным высказываниям, ставящим под вопрос возможность выделения мнений с такими значениями.

Исследование особенностей выражения мнения в политическом дискурсе важно как для теоретической лингвистики, так и для автоматической обработки естественного языка.

ЛИТЕРАТУРА

1. Беляков М. В., Максименко О. И. Прагматика метафоры в русском дипломатическом дискурсе // IV Фирсовские чтения. Язык в современных дискурсивных практиках: материалы докладов и сообщений Международной научно-практической конференции, Москва, 22–23 октября 2019 г. / под ред. А. С. Борисовой, А. В. Игнатенко, Т. В. Лариной, О. В. Ломакиной. М.: РУДН, 2019. С. 453–459.
2. Семина Т. А. Извлечение мнения автора через обратную частоту документа // Вестник Московского государственного областного университета

(электронный журнал). 2019. № 2. URL: www.evestnik-mgou.ru (дата обращения: 04.10.2019).

3. Юсупов И. Распознавание именованных сущностей с использованием синтактико-семантических признаков и нейросетей [Электронный ресурс] // Диалог: труды Международной конференции по компьютерной лингвистике и интеллектуальным технологиям, Москва, 1–4 июня 2016 г. Вып. 15 (22). [2016]. URL: <http://www.dialog-21.ru/media/3475/yusupov.pdf> (дата обращения: 04.10.2019).
4. Baymurzina D. R., Kuznetsov D. P., Burtsev M. S. Language model embeddings improve sentiment analysis in Russian // Компьютерная лингвистика и интеллектуальные технологии: по материалам ежегодной международной конференции «Диалог», Москва, 29 мая – 1 июня 2019 г. М.: Издательство РГГУ, 2019. С. 53–62.
5. Chiranjeevi P., Santosh D. T., Vishnuvardhan B. Survey on Sentiment Analysis Methods for Reputation Evaluation // Cognitive Informatics and Soft Computing. Singapore: Springer, 2019. P. 53–66.
6. Choi Y., Wiebe J., Mihalcea R. Coarse-grained +/- Effect Word Sense Disambiguation for Implicit Sentiment Analysis // The IEEE Transactions on Affective Computing. 2017. Vol. 8. № 4. P. 471–479.
7. Li X., Wu C., Mai F. The effect of online reviews on product sales: A joint sentiment-topic analysis // Information & Management. 2019. Vol. 56. № 2. P. 172–184.
8. Rusnachenko N., Loukachevitch N. V. Extracting Sentiment Attitudes from Analytical Texts via Piecewise Convolutional Neural Network // DAMDID/RCDL, October 9th–12th, 2018. Moscow, 2018. P. 186–192.
9. Saura J. R., Palos-Sanchez P., Grilo A. Detecting indicators for startup business success: Sentiment analysis using text data mining // Sustainability. 2019. Vol. 11. № 3. URL: <https://www.mdpi.com/2071-1050/11/3/917> (дата обращения: 04.10.2019).
10. Zhang L., Liu B. Identifying noun product features that imply opinions // Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers. Vol. 2. Shoumen, Bulgaria: Association for Computational Linguistics, 2011. P. 575–580.

REFERENCES

1. Belyakov M. V., Maksimenko O. I. [Pragmatics of metaphor in Russian diplomatic discourse]. In: Borisova A. S., Ignatenko A. V., Larina T. V., Lomakina O. V., eds. *IV Firsovskie chteniya. Yazyk v sovremennykh diskursivnykh praktikakh: materialy dokladov i soobshchenii Mezhdunarodnoi nauchno-prakticheskoi konferentsii, Moskva, 22–23 oktyabrya 2019 g.* [IV Firsov readings. Language in modern discourse practices: proceedings of the International scientific and practical conference, Moscow, October 22nd–23rd, 2019]. Moscow, Peoples' Friendship University of Russia Publ., 2019, pp. 453–459.
2. Semina T. A. [Author's opinion mining using inverse document frequency]. In: *Vestnik Moskovskogo gosudarstvennogo oblastnogo universiteta (elektronnyy*

- zhurnal) [Bulletin of Moscow Region State University (e-journal)], 2019, no. 2. Available at: www.evestnik-mgou.ru (accessed: 04.10.2019).
3. Yusupov I. [Named entity recognition using syntactic and semantic features and neural networks]. In: *Dialog: Trudy mezhdunarodnoi konferentsii po kompyuternoi lingvistike i intellektual'nykh tekhnologiyam, Moskva, 1–4 iyunya 2016 g.* [Dialogue: Proceedings of the international conference on computational linguistics and intellectual technologies, Moscow, June 1st–4th, 2016], vol. 15 (22). Available at: <http://www.dialog-21.ru/media/3475/yusupov.pdf> (accessed: 04.10.2019).
4. Baymurzina D. R., Kuznetsov D. P., Burtsev M. S. Language model embeddings improve sentiment analysis in Russian. In: *Komp'yuternaya lingvistika i intellektual'nye tekhnologii: po materialam ezhegodnoi mezhdunarodnoi konferentsii «Dialog», Moskva, 29 maya – 1 iyunya 2019 g.* [Computational linguistics and intellectual technologies: Based on the materials of the annual international conference “Dialogue”, Moscow, May 29th – June 1st, 2019]. Moscow, Russian State Humanitarian University Publishing, 2019, pp. 53–62.
5. Chiranjeevi P., Santosh D. T., Vishnuvardhan B. Survey on Sentiment Analysis Methods for Reputation Evaluation. In: *Cognitive Informatics and Soft Computing*. Singapore, Springer, 2019, pp. 53–66.
6. Choi Y., Wiebe J., Mihalcea R. Coarse-grained +/- Effect Word Sense Disambiguation for Implicit Sentiment Analysis. In: *The IEEE Transactions on Affective Computing*, 2017, vol. 8, no. 4, pp. 471–479.
7. Li X., Wu C., Mai F. The effect of online reviews on product sales: A joint sentiment-topic analysis. In: *Information & Management*, 2019, vol. 56, no. 2, pp. 172–184.
8. Rusnachenko N., Loukachevitch N. V. Extracting Sentiment Attitudes from Analytical Texts via Piecewise Convolutional Neural Network. In: *DAMDID/RCDL, October 9th–12th, 2018*. Moscow, 2018, pp. 186–192.
9. Saura J. R., Palos-Sanchez P., Grilo A. Detecting indicators for startup business success: Sentiment analysis using text data mining. In: *Sustainability*, 2019, vol. 11, no. 3. Available at: <https://www.mdpi.com/2071-1050/11/3/917> (accessed: 04.10.2019).
10. Zhang L., Liu B. Identifying noun product features that imply opinions. In: *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. Vol. 2*. Shoumen, Bulgaria, Association for Computational Linguistics, 2011, pp. 575–580.

ДАТА ПУБЛИКАЦИИ

Статья поступила в редакцию: 19.11.2019

Статья размещена на сайте: 22.07.2020

ИНФОРМАЦИЯ ОБ АВТОРЕ / INFORMATION ABOUT THE AUTHOR

Семина Татьяна Алексеевна – аспирант кафедры теоретической и прикладной лингвистики Московского государственного областного университета; e-mail: tanasemina@gmail.com

Tat'yana A. Semina – Postgraduate Student, Department of Theoretical and Applied Linguistics, Moscow Region State University; e-mail: tanasemina@gmail.com

ПРАВИЛЬНАЯ ССЫЛКА НА СТАТЬЮ / FOR CITATION

Семина Т. А. Анализ выражения мнения в письменном политическом медиадискурсе // Вестник Московского государственного областного университета (электронный журнал). 2020. № 3. URL: www.evestnik-mgou.ru.

Semina T. A. Analysis of expression of opinion in written political media discourse. In: *Bulletin of Moscow Region State University (e-journal)*, 2020, no. 3. Available at: www.evestnik-mgou.ru.