


**РАЗДЕЛ 4. ЛИНГВИСТИЧЕСКАЯ ЭКСПЕРТИЗА: ЯЗЫК И ПРАВО**

УДК 81'42  
ББК Ш105.51  
DOI 10.26170/pl20-01-13


ГСНТИ 16.31.61

Код ВАК 10.02.19

**К. В. Злоказов**

Санкт-Петербургский университет МВД России, Санкт-Петербург, Россия  
ORCID ID: 0000-0002-0664-8444 

**А. А. Рожков**

Санкт-Петербургский университет МВД России, Санкт-Петербург, Россия  
ORCID ID: 0000-0003-1890-0368 

 **E-mail:** zkirvit@yandex.ru; anrozkov@mail.ru.

**Основные экстремистские концепты и их лексические признаки: результаты эмпирического исследования**

**АННОТАЦИЯ.** Исследование проводилось на материале перечня из 3986 наименований текстовых, аудио- и видеофрагментов, включенных в Федеральный список экстремистских материалов. Цель статьи — анализ Федерального списка экстремистских материалов методами когнитивной лингвистики для классификации экстремистских идей и описания типичных лексических признаков их воплощения. Исследование является частью работы, направленной на описание конструкции экстремистского призыва, что позволит впоследствии точнее и быстрее выявлять его новые формы, дополнить существующую стратегию отнесения информации к экстремистской классификационно-критериальными основаниями. Эти задачи становятся особенно актуальными в свете того, что форма представления информации в среде Интернет все чаще от монокодовой переходит к поликодовой, сочетающей в себе текстовые, образные, аудиальные фрагменты. Разнородность текстов требует изменений в методиках, применяемых судебными лингвистами и психологами для выявления признаков экстремистских идей. Методом исследования стал контент-анализ, а методикой — «text mining», разновидность количественно-качественного анализа морфологии и лексики текстов. В ходе анализа корпуса экстремистских текстов статистически определены концепты, представляющие националистский, исламистский, славянский экстремистские дискурсы, а также три их радикальных подвида с выраженными призывами к насилию. Описаны ключевые лексические признаки (стемы) каждого выявленного концепта. Показано, что применение методики «text mining» позволяет дополнить метод экспертной оценки принадлежности текста к экстремистскому установлением частотной лексики и использовать эти возможности при решении задач, связанных с интеллектуальным анализом текстов в интересах противодействия экстремизму.

**КЛЮЧЕВЫЕ СЛОВА:** экстремизм; экстремистские тексты; экстремистские призывы; контент-анализ; корпусная лингвистика; экстремистский дискурс.

**ИНФОРМАЦИЯ ОБ АВТОРЕ:** Злоказов Кирилл Витальевич, кандидат психологических наук, доцент, доцент кафедры педагогики и психологии, Санкт-Петербургский университет МВД России; 198206, Россия, г. Санкт-Петербург, ул. Л. Пилутова, д. 1, кор. 9, каб. 412; e-mail: zkirvit@yandex.ru.

**ИНФОРМАЦИЯ ОБ АВТОРЕ:** Рожков Антон Алексеевич, кандидат педагогических наук, доцент, доцент кафедры юридической психологии, Санкт-Петербургский университет МВД России; 198206, Россия, г. Санкт-Петербург, пос. Стрельна, Санкт-Петербургское шоссе, д. 17, каб. 304; e-mail: anrozkov@mail.ru.

**ДЛЯ ЦИТИРОВАНИЯ:** Злоказов, К. В. Основные экстремистские концепты и их лексические признаки: результаты эмпирического исследования / К. В. Злоказов, А. А. Рожков // Политическая лингвистика. — 2020. — № 1 (79). — С. 97-103. — DOI 10.26170/pl20-01-13.

**БЛАГОДАРНОСТИ.** Исследование выполнено при поддержке РФФИ, проект № 20-012-00415 «Когнитивные интернет-технологии как фактор формирования экстремистского поведения молодежи: механизмы воздействия и профилактика».

Противодействие экстремизму в Российской Федерации является одной из ключевых задач обеспечения национальной безопасности. В соответствии с федеральным законом от 25.07.2002 № 114-ФЗ «О противодействии экстремистской деятельности», одной из ключевых стратегий противодействия экстремистской деятельности выступает выявление и последующее устранение причин и условий, способствующих ее осуществ-

лению [О противодействии экстремистской деятельности [http](http://www.fedresurs.ru)]. Реализация данной задачи обеспечивается деятельностью правоохранительных органов, общественных институтов, образовательных и культурных организаций.

В Российской Федерации законодательно разработан специальный механизм, регулирующий деятельность по противодействию экстремизму. Одним из его элементов являет-

ся Федеральный список экстремистских материалов — обновляемый перечень текстов, аудио-, видео- и графических материалов, посредством которых осуществлялось распространение экстремистских или радикальных идей. В настоящее время данный список содержит около 4900 записей. Его пополнение осуществляется по решению суда. Обращение к этому перечню позволяет правоохранительным органам определить, является ли информационный материал экстремистским.

Вместе с функцией контроля, данный перечень имеет и иное значение. По существу, содержание включенных в него материалов достаточно точно репрезентирует сущность экстремистских идей, распространявшихся на территории Российской Федерации. Обращение к перечню является безусловно ценным для изучения дискурса экстремистских идей, способствующих радикализации общественных настроений. Наконец, понимание конструкции экстремистского призыва позволяет точнее и быстрее выявлять его новые формы, поскольку дополняет существующую стратегию отнесения информации к экстремистской классификационно-критериальными основаниями [Экстремистский текст и деструктивная личность 2014]. Последнее особенно важно, поскольку форма представления информации в интернет-среде все чаще от монокодовой переходит к поликодовой, сочетающей в себе текстовые, образные, аудиальные фрагменты. Разнородность текстов требует изменений в методиках, применяемых судебными лингвистами и психологами для выявления признаков экстремистских идей.

Достижение этой цели требует разработки специальной методической процедуры. В данной статье нами проводится решение ряда задач, позволяющих приблизиться к ней. Цель представленного в статье сегмента работы — анализ Федерального списка экстремистских материалов методами когнитивной лингвистики для классификации экстремистских идей и описания типичных лексических признаков их воплощения.

Для обработки текста применяется метод контент-анализа, реализующий процедуру количественной (частотной) лингвистики. Выводы делаются на основании статистической обработки названий информационных материалов, включенных в Федеральный список экстремистских материалов. Вследствие ограничений, налагаемых федеральным законодательством в части противодействия экстремизму, в статье не приводятся заголовки данных материалов.

## ТЕОРЕТИЧЕСКИЕ ОСНОВАНИЯ ИЗУЧЕНИЯ ТЕКСТОВ

В последние десятилетия в когнитивной лингвистике активно развивается направление интеллектуального анализа текста (*text mining*) [Feldman, Dagan 1995]. С его помощью решаются задачи обработки содержания текстов и их классификации. Возможности практического применения этого метода достаточно широки. Анализ текстов, разработка правил категоризации используются при работе с информацией в различных областях социальных отношений — от бизнес-процессов (изучение спроса, мониторинг качества продукции и услуг по отзывам клиентов) до изучения общественного мнения в стране. Интеллектуальный анализ текстов позволяет предсказывать потребности и интересы людей, оценивать степень их удовлетворенности оказанными услугами, изучать причины социальных конфликтов и протестных настроений. Материалом для исследований выступает речевая деятельность, представленная в формах сообщений, обращений и жалоб, комментариев. Предметом исследования — семантические (смысловые) значения высказываний.

Методы интеллектуального анализа текста относятся к сфере контент-анализа, что предполагает использование определенных правил обработки текста [Семёнова, Корсунская 2010]. По характеру процедур выделяют следующие виды методов анализа:

– количественные методы анализа, оценивающие частоту встречаемости конкретных слов, выражений или фраз. Количественные методы обращаются к лексико-грамматической структуре языка, используя для этого процедуры статистической оценки. Правила для оценки текста основываются на закономерностях его структуры, сходстве с заранее определенными словоформами;

– качественные методы анализа, оценивающие содержание текста (эмоциональность, смысл, значение). Данная группа может далее разделяться с учетом источников, на основе которых проводится изучение текста — предписанные (словари, перечни слов) или рецептивные (на основе значений, присвоенных ранее специально отобранной группой экспертов) и пр. Правила для оценки содержания текста опираются на значение и смысл высказываний, а также на их морфологическое и функциональное строение.

При работе с новыми или неизвестными языковыми областями применяются смешанные, качественно-количественные стратегии контент-анализа текста. Исследователи используют качественные методы для разметки текстовых массивов, подготовки

словарей, а затем количественные — для оценки и классификации текстовой информации [Митина, Евдокименко 2010].

Исследование предполагает обращение к репрезентативной выборке слов, фраз и выражений, используемых в экстремистских текстах. В нашем случае источником текста выступает Федеральный список экстремистских материалов.

При исследовании названий материалов, включенных в Федеральный список экстремистских материалов, нами используется стратегия, совмещающая количественный и качественный методы. При этом в рамках данной статьи описывается первичный этап обработки текстового массива — вычленение основных словарных единиц, применяемых при обозначении экстремистских материалов. В дальнейшем это позволит проводить более глубокий семантический анализ словарных единиц экстремистского дискурса.

Работа по исследованию экстремистских материалов строилась по алгоритму, используемому при разработке корпуса новых слов в уже известном языковом пространстве. Схожий алгоритм применяется для создания словарей социальных, профессиональных и этнических групп [Kondratoff 1999].

На первом этапе работы выполнялась очистка массива от некорректных в лексическом смысле записей, затем проводилась токенизация — выделение из массива слов, используемых в заголовках экстремистских материалов. Необходимость проведения токенизации заключается в отсеивании знаков препинания, союзов, предлогов и других элементов, затрудняющих частотный анализ.

На втором этапе проводился стемминг и лемматизация. Стемминг — это морфологическая операция по определению основы слова. Например, слова «русская», «русского», «русским» благодаря стеммингу будут рассматриваться схожими по своему значению, поскольку их основой является слово «русский», а различия возникают из-за окончания. Стемминг позволяет считать данные слова обладающими одинаковым значением.

Лемматизация — более сложная технология нахождения основы слова. Она заключается в приведении различных словоформ к их основному словарному виду. Например, слова «русское», «русских» и «русским» будут приведены к морфологической форме слова, представленной в именительном па-

деже и единственном числе. Лемматизация упрощает обработку массива информации, сокращая объем анализируемых значений.

Стемминг и лемматизация проводятся при обработке корпусов текстов для нахождения концепта — минимального набора слов, используемых для обозначения идеи, характеристики понятия и пр. В рамках нашего исследования изучаемыми концептами являются экстремистские идеи.

Первый и второй этапы осуществлялись посредством утилиты *Text mining* статического пакета *Statistica*, имеющего размеченный словарь русского языка. В результате был получен перечень часто используемых (ключевых) словоформ.

Далее нами выполнялась количественная обработка корпуса текстов. Благодаря этому были дифференцированы совокупности слов, используемых для обозначения различных видов экстремистских материалов — националистических, религиозных, деструктивных. Статистическая оценка выявленных концептов и интерпретация их смысловых значений завершила исследование.

## РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

**Описательная характеристика исследуемых материалов.** Всего Федеральный список экстремистских материалов на момент проведения исследования включал 4699 записей, однако не все они пригодны для анализа, поскольку некоторые не содержат орфографически корректно представленные слова. Поэтому при обработке нами исключены записи, не содержащие лексической информации (например, «45.mp3», «vk.id214942.html» и схожие), а также заголовки, указывающие на источник экстремистской информации («листовка», «публикация в газете»). В результате отбраковки текстовая база исследования составила 4016 записей. С учетом корректно представленного года было определено 3986 записей (см. табл. 1).

Средний объем знаков, составляющих одну запись, равен 42 знакам, минимальное значение — 2 знака, максимальное — 1133 знака. Общий объем информации составил 168 983 знака. Токенизация массива записей позволила выделить 5923 слова, сведенных посредством стемматизации к 32 стемам (см. табл. 2).

**Таблица 1.** Количество записей в Федеральном списке экстремистских материалов  
(Table 1. The number of entries in the Federal List of Extremist Materials)

Год	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Орфографически корректные записи	80	179	148	255	290	518	465	380	665	532	294	161	19

*Примечание:* отметка «год» — время внесения материала в Федеральный список экстремистских материалов; приведено количество записей после лексико-орфографической обработки

**Таблица 2.** Стематизация заголовков записей Федерального списка экстремистских материалов  
(Table 2. Stemming of the titles of entries in the Federal List of Extremist Materials)

№ п/п	Стема	Частота использования	Количество записей	№ п/п	Стема	Частота использования	Количество записей
1	русск	473	415	17	аб	64	59
2	росс	247	230	18	алл	62	57
3	обращен	119	112	19	войн	60	57
4	эт	93	85	20	прот	57	55
5	хач	89	75	21	чурк	52	50
6	скинхед	79	75	22	исламск	49	47
7	джихад	78	69	23	убит	49	45
8	наш	77	72	24	убива	49	42
9	прав	74	69	25	ислам	48	47
10	бел	72	66	26	мир	48	48
11	Кавказ	72	68	27	кавказц	47	45
12	жид	70	62	28	чурок	47	45
13	народ	70	65	29	пут	46	42
14	скин	68	60	30	слав	46	43
15	амир	67	59	31	исл	45	44
16	смерт	65	55	32	слов	45	45

Рассмотрение полученного перечня стем позволяет заключить, что они представляют разные направления экстремистского дискурса. Видно, что некоторые понятия относятся к двум областям экстремистских идей: националистической (леммы «русский», «Россия», «скинхед»), радикальной исламской (леммы «амир», «джихад», «ислам»). Кроме того, часть стем отражает радикализацию экстремизма (леммы «смерть», «убит», «убивать»). Однако часть стем представлена наречиями и предлогами, что, по сути, отражает грамматическую структуру экстремистского текста.

Соответственно, необходимо определить отношения между выявленными стемами, дифференцировав на их основании близкие, нейтральные и противоположные группировки — концепты. Определение концептов осуществляется посредством статистической оценки частоты соотношения слов в изучаемом нами массиве записей методом преобразования матрицы слов и показателей их встречаемости в текстах. Данная операция базируется на предположении о том, что применение слов с определенной частотой в предложении говорит о схожести их значений. Кроме того, показателем важности «веса» этих слов выступает доля их представленности в конкретном тексте. Чем больше слов, описывающих концепт по от-

ношению к другим словам в тексте, тем большую значимость они имеют для характеристики концепта. К примеру, экстремистский текст, включающий слова агрессивно-насильственной направленности, как правило, короче, чем текст, не имеющий их. При соотношении частот употребления насильственная семантика будет иметь большую значимость в концепте, поскольку эти слова встречаются чаще и составляют большую часть текста.

Определение концептов проводилось посредством сингулярного разложения матрицы текстовых записей. Сущностно данный метод близок к методу главных компонент, применяемому в эксплораторном факторном анализе. Цель сингулярного разложения — сокращение количества факторов (концептов), описывающих группировки слов.

По критерию отношения сингулярных значений к доле объясняемых с их помощью взаимоотношений в массиве нами было выделено 6 концептов. В совокупности они объясняют 96 % применений 5923 словоформ, образующих 32 стемы. Меньшее количество концептов приводило к повышению доли необъяснимых комбинаций слов, а большее — к ухудшению их объяснительного потенциала, появлению схожих по смыслу концептов. Выявленные концепты представлены в таблице 3.

**Таблица 3.** Смысловые группировки (концепты) лексики текстов, представленной в Федеральном списке экстремистских материалов  
(Table 3. Semantic groupings (concepts) of vocabulary of texts presented in the Federal List of Extremist Materials)

Концепт 1 «Националистский»	Концепт 2 «Радикально-националистский»	Концепт 3 «Смешанный»	Концепт 4 «Славянский»	Концепт 5 «Радикально-исламский»	Концепт 6 «Исламский»
Русск* (102) Росс* (70) Бел* (51) Народ* (42) Убива* (41)	Хач* (67) Скин* (47) Убива* (44) Смерт* (40)	Хач* (104) Смерт* (74) Скин* (55) Джихад* (50) Наш* (49) Убива* (44)	Росс* (190) Слав* (52) Жид* (55) Народ* (45)	Эт* (120) Наш* (61) Джихад* (56) Кавказ* (45)	Эт* (100) Обращен* (91) Амир* (44) «Алла*» (41)
29 % сочетаний словоформ в массиве	17 % сочетаний словоформ в массиве	14 % сочетаний словоформ в массиве	13 % сочетаний словоформ в массиве	12 % сочетаний словоформ в массиве	11 % сочетаний словоформ в массиве

*Примечание:* в таблице представлены группировки стем (концепты). Знак перед коэффициентом в матрице сингулярного разложения указывает на частоту встречаемости в текстах. Коэффициенты округлены до целых чисел.

Названия концептов даны в соответствии с семантикой включенных в них стем. Первый концепт, объясняющий 29 % группировки словоформ, основывается на использовании стем «русск», а также «бел», «народ» и «убива». Такой набор слов типично применяется для обоснования националистских идей и может быть представлен как в агрессивной, так и ксенофобической формах.

Второй концепт объясняет 17 % словоформ, также относится к националистическому варианту экстремистских текстов, однако содержит большее количество лексики насилия. В таблице она представлена стемами «убива» и «смерт». Концепт присутствует в материалах, содержащих прямые призывы к убийству или конфликту с лицами неславянской национальности.

Третий концепт, объясняющий 14 % словоформ, включает понятия, свойственные не только националистской, но и исламистской экстремистской идеологии. В нашей интерпретации он получил название «Смешанный». Причина группировки объясняется смешением лексики насилия с понятиями националистской и исламистской радикальной риторики. Об этом говорит сочетание стем «смерт», «джихад», «убива» с обозначениями «хач» и «скин». Такое смешение лексики возможно из-за схожести призывов, применяемых идеологами радикального националистского и исламистского направлений.

Четвертый концепт — «Славянский» (13 % словоформ) — обобщает две семантические группы: славянскую и антисемитскую. Концепт построен на возвышении славянской культуры и ценностей в противопоставлении иудейской культуре. Цель — посредством псевдоисторических исследований, искажения фактов, предоставления ложных сведений сформировать ксенофо-

бическую установку по национальному и религиозному признаку. В ряде информационных материалов подобного содержания представлена только славянская лексика, в ряде она смешивается с антисемитской.

Пятый концепт — «Радикально-исламский» (12 % словоформ). Данный концепт обобщает лексику, используемую при формулировании призыва к джихаду (священной войне с неверными) и реализуется в стемах «джихад», «Кавказ», а также грамматических конструкциях «это», «наш». Конструкция данного концепта типична и включает описанные стемы или их словоформы.

Шестой концепт — «Исламский» — объясняет 12 % словоформ и включает применяемый в радикальной исламской пропагандистской риторике набор понятий: «амир», «алла\*», а также стратегию «обращение». Эти понятия раскрывают «мягкую» технологию воздействия на представления верующих. Она заключается в радикальной трактовке положений ислама отдельными лицами, позиционирующими себя в качестве общественных деятелей — «амиров». Жанр «обращения» предполагает религиозное обоснование конкретной модели поведения и указание на необходимость ее соблюдения верующими. Как правило, содержание обращений охватывает различные стороны жизни отдельных мусульман и религиозных сообществ. Однако экстремистские обращения типичны, поскольку подобные выступления нацелены на формирование вражды по признаку принадлежности к религии, пропагандируют несоблюдение норм «светского» поведения.


Таким образом, в результате лексической обработки наименований, включенных в Федеральный список экстремистских материалов, были описаны несколько взаимосвязанных группировок слов (концептов),

представляющих различные виды экстремистских идей: «националистскую», «исламистскую», «смешанную» и «славянскую». Кроме того, зафиксированы два концепта, пропагандирующие применение насилия в поддержку националистских и исламистских радикальных идей. Выделение концептов проводилось посредством обобщения частот применяемых слов, выявления основных словоформ и поиска их морфологических оснований. Полученные результаты могут быть использованы в системах автоматической классификации текстов с целью блокировки текстов, содержащих экстремистские идеи, а также призывы к насилию. Полученные результаты показывают перспективы применения методов корпусной лингвистики для выявления типичных признаков экстремистских текстов, оценки степени радикальности экстремистских текстов на основе извлечения находящейся в них лексики насилия.


#### ЛИТЕРАТУРА


1. Араева, Л. А. Судебно-лингвистическая экспертиза по криминальным проявлениям экстремизма / Л. А. Араева, М. А. Осадчий. — Текст : непосредственный // Уголовный процесс. — 2006. — № 4 (16). — С. 45—56.
2. Баранов, А. Н. Лингвистическая экспертиза текста: теория и практика / А. Н. Баранов. — Москва : Флинта : Наука, 2007. — 592 с. — Текст : непосредственный.
3. Бидова, Б. Б. Преступления, квалифицированные наличием экстремистских мотивов / Б. Б. Бидова. — Текст : непосредственный // Молодой ученый. — 2013. — № 11. — С. 515—518.
4. Ворошилова, М. Б. Русское лото 18 из 1488: опыт комплексной психолого-лингвистической экспертизы / М. Б. Ворошилова, К. В. Злоказов. — Текст : непосредственный // Уральский филологический вестник / ФГБОУ ВПО «Урал. гос. пед. ун-т»; гл. ред. Т. А. Гридина. — Екатеринбург, 2013. — Вып. 4. — С. 69—82.
5. Гагина, О. В. Психолого-лингвистическое исследование кратких текстов (в том числе креолизованных) по делам, связанным с противодействием экстремизму (на материале демоматериалов) / О. В. Гагина, В. О. Кузнецов, Ю. Н. Зубкова. — Текст : непосредственный // Теория и практика судебной экспертизы. — 2013. — № 2 (30). — С. 43—44.
6. Гальперин, И. Р. Текст как объект лингвистического исследования / И. Р. Гальперин. — Москва : Наука, 1981. — Текст : непосредственный.
7. Грищенко, А. И. Экспрессивные этнонимы как приметы языка вражды / А. И. Грищенко, Н. А. Николина. — Текст : непосредственный // Язык вражды и язык согласия в социокультурном контексте современности : коллективная моногр. / отв. ред. И. Т. Вепрева, Н. А. Купина, О. А. Михайлова. — Екатеринбург : Издательство Уральского университета, 2006. — С. 175—187. — (Труды / Урал. МИОН ; вып. 20).
8. Злоказов, К. В. Репрезентация идей националистической направленности в креолизованном тексте / К. В. Злоказов. — Текст : непосредственный // Политическая лингвистика. — 2014. — № 3. — С. 236—241.
9. Митина, О. В. Формализованные методы исследования текстов: опыт применения к анализу технической документации / О. В. Митина, А. С. Евдокименко. — Текст : непосредственный // Вестник Томского государственного университета. Филология. — 2010. — № 1. — Т. 9. — С. 60—69.
10. О противодействии экстремистской деятельности : Федеральный закон : от 25 июля 2002 г. № 114-ФЗ : (ред. от 25.12.2012) // Консультант Плюс: справочно-правовая система : сайт. — URL: <http://base.consultant.ru/cons/cgi/online.cgi?req=doc;base=LAW;n=139786> (дата обращения: 01.05.2019). — Текст : электронный.
11. Семёнова, А. В. Контент-анализ СМИ: проблемы и опыт применения / А. В. Семёнова, М. В. Корсунская ; под ред. В. А. Мансурова. — М. : Институт социологии РАН, 2010. — 324 с. — Текст : непосредственный.
12. Экстремистский текст и деструктивная личность : моногр. / Ю. А. Антонова, Л. Е. Веснина, М. Б. Ворошилова, К. В. Злоказов, Ю. П. Тагильцева, А. А. Карапетян ; Урал. гос. пед. ун-т. — Екатеринбург : [б. и.], 2014. — 272 с. — Текст : непосредственный.
13. Feldman, R. KDT — knowledge discovery in texts / R. Feldman, I. Dagan. — Text : unmediated // Procedia of the First International Conference on Knowledge Discovery (KDD). — 1995. — P. 112—117.
14. Kodratoff, Y. Knowledge discovery in texts: A definition and applications / Y. Kodratoff. — Text : unmediated // Lecture Notes in Computer Science. — 1999. — No 1609. — P. 16—29.

**K. V. Zlokazov**

St. Petersburg University of the Ministry of Internal Affairs, St. Petersburg, Russia  
ORCID ID: 0000-0002-0664-8444 

**A. A. Rozhkov**

St. Petersburg University of the Ministry of Internal Affairs, St. Petersburg, Russia  
ORCID ID: 0000-0003-1890-0368 

 **E-mail:** [zkirvit@yandex.ru](mailto:zkirvit@yandex.ru); [anrozkov@mail.ru](mailto:anrozkov@mail.ru).

## Extremist Concepts and Their Lexical Features: Empirical Research

**ABSTRACT.** *The sample consists of 3986 written, audio- and video-fragments included in the Federal list of extremist materials. The aim of the study is to analyze the texts from the Federal list of extremist materials by means of cognitive linguistics for classification of extremist ideas and description of the typical lexical features of their realization. The study is a part of a comprehensive investigation targeted at description of the construction of extremist appeal which would allow us in the future to reveal its new forms in a quicker and more precise manner and to enrich the existing strategy of classification of information as extremist by categorization principles. These tasks become especially urgent due to the fact that the form of presentation of material over the Internet in more and more cases shifts from a monocode to a policode one combining written, visual and audial fragments. The text heterogeneity demands changes in the methods used by forensic linguists and psychologists to reveal the typical features of extremist ideas. Content analysis became the leading research method incorporating the procedure of text mining as a kind of quantitative-qualitative analysis of text morphology and lexicon. In the course of extremist text corpus analysis, the authors have singled out concepts of nationalist, Islamic, Slavic and extremist discourses, as well as three radical subtypes displaying salient appeals for violence. The article describes the key lexical features (stememes) of each concept singled out. It shows that the technology of text mining allows complementing the method of ex-*

pert evaluation of the text as an extremist one through identification of frequently used vocabulary and makes it possible to use this potential to solve problems connected with intellectual text analysis in order to counteract extremism.

**KEYWORDS:** *extremism; extremist texts; extremist appeals; content analysis; corpus linguistics; extremist discourse.*

**AUTHOR'S INFORMATION:** *Zlokazov Kirill Vital'evich, Candidate of Psychology, Associate Professor of Department of Pedagogy and Psychology, St. Petersburg University of the Ministry of Internal Affairs, St. Petersburg, Russia.*

**AUTHOR'S INFORMATION:** *Rozhkov Anton Alekseevich, Candidate of Pedagogy, Associate Professor of Department of Legal Psychology, St. Petersburg University of the Ministry of Internal Affairs of Russia, St. Petersburg, Russia.*

**FOR CITATION:** *Zlokazov, K. V. Extremist Concepts and Their Lexical Features: Empirical Research / K. V. Zlokazov, A. A. Rozhkov // Political Linguistics. — 2020. — No 1 (79). — P. 97-103. — DOI 10.26170/pl20-01-13.*

**ACKNOWLEDGMENTS.** Research is accomplished with financial support of the Russian Foundation for Basic Research (RFBR) Project № 20-012-00415.

#### REFERENCES

1. Araeva, L. A. Forensic linguistic examination of the criminal manifestations of extremism / L. A. Araeva, M. A. Osadchiy. — Text: direct // Criminal process. — 2006. — No. 4 (16). — P. 45—56. [Sudebno-lingvisticheskaya ekspertiza po kriminal'nym proyavleniyam ekstremizma / L. A. Araeva, M. A. Osadchiy. — Tekst : neposredstvennyy // Ugolovnyy protsess. — 2006. — № 4 (16). — S. 45—56]. — (In Rus.)
2. Baranov, A. N. Linguistic examination of the text: theory and practice / A. N. Baranov. — Moscow : Flinta : Science, 2007. — 592 p. — Text: direct. [Lingvisticheskaya ekspertiza teksta: teoriya i praktika / A. N. Baranov. — Moskva : Flinta : Nauka, 2007. — 592 s. — Tekst : neposredstvennyy]. — (In Rus.)
3. Bidova, B. B. Crimes Qualified by Extremist Motives / B. B. Bidova. — Text : direct // Young scientist. — 2013. — No. 11. — P. 515—518. [Prestupleniya, kvalifitsirovannyye nalichiem ekstremistskikh motivov / B. B. Bidova. — Tekst : neposredstvennyy // Molodoy uchenyy. — 2013. — № 11. — S. 515—518]. — (In Rus.)
4. Voroshilova, M. B. Russian Lotto 18 of 1488: the experience of a comprehensive psychological and linguistic examination / M. B. Voroshilova, K. V. Zlokazov. — Text: direct // Ural Philological Bulletin / FSBEI HPE "Ural State Ped. Univ."; ed-in-chief T. A. Gridina. — Ekaterinburg, 2013. — Issue. 4. — P. 69—82. [Russkoe loto 18 iz 1488: opyt kompleksnoy psikhologo-lingvisticheskoy ekspertizy / M. B. Voroshilova, K. V. Zlokazov. — Tekst : neposredstvennyy // Ural'skiy filologicheskyy vestnik / FGBOU VPO «Ural. gos. ped. un-t.»; gl. red. T. A. Gridina. — Ekaterinburg, 2013. — Vyp. 4. — S. 69—82]. — (In Rus.)
5. Gagina, O. V. Psychological and linguistic study of short texts (including creolized) in cases related to countering extremism (based on demotivators) / O. V. Gagina, V. O. Kuznetsov, Yu. N. Zubkova. — Text : direct // Theory and Practice of Forensics. — 2013. — No. 2 (30). — P. 43—44. [Psikhologo-lingvisticheskoe issledovanie kratkikh tekstov (v tom chisle kreolizovannykh) po delam, svyazannym s protivodeystviem ekstremizmu (na materiale demotivatorov) / O. V. Gagina, V. O. Kuznetsov, Yu. N. Zubkova. — Tekst : neposredstvennyy // Teoriya i praktika sudebnoy ekspertizy. — 2013. — № 2 (30). — S. 43—44]. — (In Rus.)
6. Gal'perin, I. R. Text as an Object of Linguistic Research / I. R. Gal'perin. — Moscow : Science, 1981. — Text: direct. [Tekst kak ob'ekt lingvisticheskogo issledovaniya / I. R. Gal'perin. — M. : Nauka, 1981. — Tekst : neposredstvennyy]. — (In Rus.)
7. Grishchenko, A. I. Expressive ethnonyms as signs of hate speech / A. I. Grishchenko, N. A. Nikolina. — Text: direct // Hate speech and language of harmony in the sociocultural context of modernity : collective monograph. / resp. ed. I. T. Vepreva, N. A. Kupina, O. A. Mikhailova. — Ekaterinburg : Publishing House of the Ural University, 2006. — P. 175—187. — (Works / Ural MION; issue 20). [Ekspressivnye etnonimy kak primety yazyka vrazhdy / A. I. Grishchenko, N. A. Nikolina. — Tekst : neposredstvennyy // Yazyk vrazhdy i yazyk soglasiya v sotsiokul'turnom kontekste sovremennosti : kolektivnaya monogr. / otv. red. I. T. Vepreva, N. A. Kupina, O. A. Mikhailova. — Ekaterinburg : Izdatel'stvo Ural'skogo universiteta, 2006. — S. 175—187. — (Trudy / Ural. MION ; vyp. 20)]. — (In Rus.)
8. Zlokazov, K. V. Representation of nationalist ideas in a creolized text / K. V. Zlokazov. — Text: direct // Political Linguistics. — 2014. — No. 3. — P. 236—241. [Reprezentatsiya idey natsionalisticheskoy napravlenosti v kreolizovannom tekste / K. V. Zlokazov. — Tekst : neposredstvennyy // Politicheskaya lingvistika. — 2014. — № 3. — S. 236—241]. — (In Rus.)
9. Mitina, O. V. Formalized text research methods: experience of applying technical documentation to the analysis / O. V. Mitina, A. S. Evdokimenko // Tomsk State University Bulletin. Philology. 2010. No. 1. Vol. 9. P. 60—69. [Formalizovannyye metody issledovaniya tekstov: opyt primeneniya k analizu tekhnicheskoy dokumentatsii / O. V. Mitina, A. S. Evdokimenko. — Tekst : neposredstvennyy // Vestnik Tomskogo gosudarstvennogo universiteta. Filologiya. — 2010. — № 1. — T. 9. — S. 60—69]. — (In Rus.)
10. On counteracting extremist activity: Federal law: July 25, 2002 No. 114-ФЗ: (as amended on December 25, 2012) // Consultant Plus: legal reference system: website. [O protivodeystvii ekstremistskoy deyatel'nosti : Federal'nyy zakon : ot 25 iyulya 2002 g. № 114-FZ : (red. ot 25.12.2012) // Konsultant Plyus: spravochno-pravovaya sistema : sayt]. — URL: <http://base.consultant.ru/cons/cgi/online.cgi?req=doc;base=LAW;n=139786> (date of access: 01.05.2019). — Text : electronic.
11. Semenova, A. V. Media Content Analysis: Problems and Application Experience / A. V. Semenova, M. V. Korsunskaya ; ed. V. A. Mansurova. — Moscow : Institute of Sociology RAS, 2010. 332 p. [Kontent-analiz SMI: problemy i opyt primeneniya / A. V. Semenova, M. V. Korsunskaya ; pod red. V. A. Mansurova. — M. : Institut sotsiologii RAN, 2010. — 324 s. — Tekst : neposredstvennyy]. — (In Rus.)
12. Extremist text and a destructive personality: a monograph / Yu. A. Antonova, L. E. Vesnina, M. B. Voroshilova, K. V. Zlokazov, Yu. R. Tagiltseva, Karapetyan A. A. . Yekaterinburg: Ural State Pedagogical University, 2014. 272 p. [Ekstremistskiy tekst i destruktivnaya lichnost' : monogr. / Yu. A. Antonova, L. E. Vesnina, M. B. Voroshilova, K. V. Zlokazov, Yu. R. Tagiltseva, A. A. Karapetyan ; Ural. gos. ped. un-t. — Ekaterinburg : [b. i.], 2014. — 272 s. — Tekst : neposredstvennyy]. — (In Rus.)
13. Feldman, R. KDT — knowledge discovery in texts / R. Feldman, I. Dagan. — Text : unmediated // Procedia of the First International Conference on Knowledge Discovery (KDD). — 1995. — P. 112—117.
14. Kodratoff, Y. Knowledge discovery in texts: A definition and applications / Y. Kodratoff. — Text : unmediated // Lecture Notes in Computer Science. — 1999. — No 1609. — P. 16—29.