

УДК 81'33

КОМПАРАТИВНОЕ ИССЛЕДОВАНИЕ ТЕКСТОВ УЧАСТНИКОВ ЭКСТРЕМИСТСКОГО ФОРУМА И ЛИЦ С ИЗВЕСТНЫМИ ПСИХОЛОГИЧЕСКИМИ ХАРАКТЕРИСТИКАМИ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ СТИЛЕМЕТРИЧЕСКОГО АНАЛИЗА¹

ЛИТВИНОВА Татьяна Александровна,
кандидат филологических наук,
заведующая научно-исследовательской лабораторией корпусной идиолектологии,
Воронежский государственный педагогический университет

АННОТАЦИЯ. Широкое распространение в Сети большого числа экстремистских материалов вызвало необходимость в разработке эффективной стратегии, направленной на противодействие экстремизму. Одной из задач, которую следует решать при разработке такой стратегии, является изучение личностных особенностей экстремиста. Несмотря на большое число исследований, посвященных данному вопросу, он далек от своего разрешения вследствие трудностей, связанных с получением объективных данных о психологических характеристиках экстремистов. Одним из средств доступа к психологическим особенностям человека является анализ его речевой продукции. В работе представлены результаты экспериментов, направленных на сравнительное исследование текстов участников форума, внесенного в Федеральный список экстремистских материалов РФ, и текстов лиц с известными психологическими характеристиками. В исследовании использовались методы компьютерной лингвистики и многомерного анализа данных (кластерный анализ, метод главных компонент, консенсусные сети и т.д.). Сделаны выводы, касающиеся как особенностей экстремистского дискурса, так и вопросов выбора лингвистических признаков.

КЛЮЧЕВЫЕ СЛОВА: корпус текстов, экстремистский текст, стилеметрия, корпусная лингвистика, кластерный анализ.

**A COMPARATIVE STUDY OF TEXTS OF PARTICIPANTS OF EXTREMIST FORUM
AND PERSONS WITH KNOWN PSYCHOLOGICAL CHARACTERISTICS
USING METHODS OF STYLOMETRIC ANALYSIS**

LITVINOVA T. A.,
Cand. Philolog. Sci., Head of Corpus Idiialectology Lab,
Voronezh State Pedagogical University

ABSTRACT. The widespread dissemination of a large number of extremist materials on the Web necessitated the development of an effective strategy to counter extremism. One of the tasks that should be addressed in the course of development of such a strategy is to study the personal characteristics of an extremist. Despite the large number of studies devoted to this issue, it is far from being resolved due to the difficulties associated with obtaining objective data on the psychological characteristics of extremists. One of the means of access to the psychological characteristics of a person is the analysis of his speech. The paper presents the results of an experiment aimed at a comparative study of the texts of forum participants included in the Federal List of Extremist Materials of the Russian Federation and texts of individuals with known psychological characteristics performed with the use of computer linguistics and multivariate data analysis methods (cluster analysis, principal component analysis, consensus networks). Conclusions are drawn regarding both the characteristics of extremist discourse and selection of linguistic features.

KEY WORDS: text corpus, extremist text, stylometry, corpus linguistics, gender linguistics, text classification.

Введение

Появление в Сети большого числа групп, поддерживающих и распространяющих экс-

тремистскую идеологию, вызвало необходимость в исследованиях, направленных на установление психологического портрета данной категории лиц, их моральных установок и ценностей, с целью разработки комплекса мер, направленных на противодействие экстремизму. За последние годы как отечественными, так и зарубежными учеными было проведено множество исследований, посвященных анализу индивидуально-личностных особенностей и социально-биографических характеристик указанной группы лиц. При проведении таких исследова-

¹ Исследование выполнено при поддержке гранта Президента РФ № МК-5718.2018.6 «Речевой портрет экстремиста: корпусно-статистическое исследование (на материале экстремистского форума «Кавказ-чат»)».

© Литвинова Т.А., 2020

Информация для связи с автором: centr_rus_yaz@mail.ru

ний, однако, возникают методологические проблемы, связанные со сложностью получения объективных данных о психологических характеристиках лиц, придерживающихся экстремистских взглядов.

Ценным методологическим инструментом, который может быть использован для получения новых данных о личностных особенностях указанной группы лиц, является анализ речи. Целью настоящей статьи является описание результатов сопоставительного исследования текстов авторов экстремистского форума (ЭФ), открыто выражавших экстремистские взгляды, и текстов лиц схожей возрастной группы с известными психологическими характеристиками, проведенного с использованием стилеметрических методов, основанных на анализе поддающихся квантификации элементов текста.

1. Состояние проблемы

Проблема анализа личности экстремиста изучается в многочисленных работах как отечественных, так и зарубежных авторов [1; 3; 8; 9; 11; 12; 14; 23]. Заметим, что отдельной научной задачей является изучение личности террориста (см., например, [13]), однако мы большей частью в своем обзоре коснемся работ, посвященных именно изучению личности экстремиста.

Наименее дискуссионным является вопрос, связанный с демографическими характеристиками экстремистов. Как отмечается в большинстве соответствующих работ (см., например, [1]), современный экстремист – это лицо от 18 до 29 лет, имеющее, как правило, среднее или среднее специальное образование, при этом большинство экстремистов – мужчины. Относительно психологического портрета экстремиста единства мнений не наблюдается, однако отечественные исследователи отмечают среди характеристик экстремистов, как правило, такие как агрессивность, хладнокровие, мстительность, повышенная импульсивность [9], эгоизм, демонстративность поведения, целеустремленность [11] и т.д.

В работе [3] представлены результаты психологического обследования 915 лиц, осужденных за терроризм и экстремизм, в результате которого было установлено, что лица, осужденные за экстремистские преступления, характеризуются снобизмом, агрессивностью, упорством и самоуверенностью, являющимися «лишь компенсацией неустойчивости и эмоциональной лабильности подобных личностей» [3, с. 146].

Обзор, проведенный в работе [23], показывает, что низкий уровень образования и низкий уровень дохода не являются предикторами экстремистского поведения, в отличие от религиозности; сложно также определить единый психологический профиль лиц, склонных к экстремизму.

Другая линия исследований связана с психологическим обследованием групп лиц (как правило, лиц молодого возраста как наиболее склонных к экстремистскому поведению) с целью выявления стремления к экстремизму и личностных предикторов такой склонности. Так, в упомянутой выше работе [23] представлены результаты психологического обследования 295 румынских подростков (15-18 лет), в ходе которого было выявлено, что комплексный фактор («мегафактор») «низкий интеллект/низкий уровень экстраверсии/высокий уровень склонности к согласию» является предиктором склонности к экстремистскому поведению.

Работа российских авторов [8], которые обследовали 64 школьника старших классов, показала, что

подростки со склонностью к экстремистскому поведению обладают низким уровнем эмпатии, менее толерантны, менее коммуникабельны и имеют заниженную самооценку.

Противоречия в результатах работ, посвященных анализу психологических характеристик экстремистов, связанные преимущественно со сложностью (либо же невозможностью) получения объективных данных, неоднородностью самой исследуемой группы лиц, а также появлением большого числа текстов лиц, склонных к экстремистскому поведению, привели к активизации исследований, ориентированных на лингвистический анализ таких текстов, направленных на выявление особенностей мышления и психики их авторов. Большинство подобных исследований в настоящее время выполнены на материале английского языка (см. подробнее: [6]). При этом в них, как правило, анализируются или лексика, связанная с эмоциями, или так называемые «показатели языковой сложности текста», для измерения которой в разных работах используются разные маркеры (доля длинных слов; доля предлогов; доля слов, описывающих когнитивные процессы; доля союзов тех или иных тематических групп и т.д.) [13; 20]. Большое внимание в подобных работах уделяется анализу местоимений (так, в большинстве работ, выполненных в том числе и на материале русского языка [7; 18], отмечается повышенная частотность местоимения «они» в текстах экстремистски настроенных авторов).

В связи с активным развитием интернет-коммуникации в работах последних лет широко исследуются твиты фолловеров различных экстремистских и террористических организаций в сравнении с твитами других пользователей [12; 16; 19; 22]. Интересно, что такие не связанные с темой параметры текста, как доля длинных слов, доля слов, написанных прописными буквами, относятся к признакам, вносящим самый большой вклад в модель, различающую твиты «радикальных» и «нерадикальных» пользователей [19].

Общий вывод, который можно сделать на основании анализа подобных работ, заключается в необходимости дальнейших исследований для выявления особенностей мышления и эмоциональной сферы экстремистки настроенных пользователей. Перспективным направлением таких исследований является сравнение текстов лиц, пропагандирующих экстремистскую идеологию, с текстами авторов, психологические характеристики которых известны, и именно данная цель является основной для нашей работы.

2. Материал и методы исследования

2.1. Материал исследования

Материалом исследования являются посты участников форума «Кавказ-Чат», входящие в состав датасета 'kavkazchat' (подробнее о датасете см. в работах [6; 7]). Форум «Кавказ-Чат» внесен в Федеральный список экстремистских материалов РФ. Ряд участников данного ЭФ прямо говорят о ненависти к России и ее гражданам. Исследователи относят данный форум к одному из инструментов террористической агитации и пропаганды [10].

Следует, однако, отметить, что не все участники указанного форума открыто выражают экстремистские взгляды. Для формирования специального исследовательского корпуса нам были применены два основных критерия отбора авторов, ранее успешно использованны для анализа текстов указанного форума в гендерном аспекте [7]: 1) наличие у данного

автора не менее 50 постов в ЭФ; 2) наличие в постах данного автора случаев совместной встречаемости лемм: *рун** + *куфр** и/или *рун** + *джихад**; *джихад** + *к*фир** (с частотностью каждой леммы не менее 3).

Всего нами были отобраны тексты 104 пользователей, из них 16 женщин (пол авторов определялся на основании информации о пользователе, представленной в постах как самого пользователя, так и других участников). Все посты каждого пользователя были объединены в один файл. Данный корпус назван нами корпусом участников экстремистского форума (далее – КУЭФ).

Для сравнительного анализа мы отобрали из БД *RusIdiolect* [5], содержащей тексты лиц, представляющие собой, как и тексты КУЭФ, образцы «непрофессиональной» [2] письменной речи, а также данные о продуцентах текстов, релевантные для текущего исследования: пол, возраст, психологические характеристики по интегральным шкалам Фрайбургского личностного опросника: 1) экстраверсия – интроверсия (E); 2) эмоциональная лабильность – эмоциональная устойчивость (Emo); 3) маскулинность – фемининность (Masc). Все авторы текстов из группы сравнения относятся к возрастной группе 18–25 лет.

Высокие оценки по указанным шкалам говорят соответственно о выраженной экстраверсии, высокой эмоциональной лабильности (неустойчивости эмоционального состояния), маскулинности (протекании психической деятельности преимущественно по мужскому типу); низкие – о выраженной интровертированности, эмоциональной стабильности, протекании психической деятельности преимущественно по женскому типу. Авторы группы сравнения были разделены на 3 категории (высокая, средняя, низкая степень выраженности признака). Все тексты авторов с высокой степенью выраженности той или иной черты были объединены в файлы, содержащие в названии обозначение High, со средней –

Medium, с низкой – Low. Авторы группы сравнения («контрольный корпус») названы нами User, авторы ЭФ – Extr.

Поскольку объем текстов ЭФ от одного автора варьируется в значительных пределах (от 1909 до 209 000 слов), необходимо минимизировать влияние авторского сигнала, поскольку нашей задачей является выявление языковых особенностей, присущих группе лиц. Изначально все тексты авторов ЭФ были разделены нами на отрезки длиной в 1000 слов (тексты авторов, общий объем постов которых не достигал 1000 слов, были удалены). Как показали предварительные исследования, «групповой» (User/Extr) сигнал оказывается наиболее сильным, за ним следует гендерный сигнал, который, в свою очередь, оказывается сильнее «психологического», в связи с чем нами было принято решение объединить тексты мужчин и женщин для нивелирования гендерного сигнала. В связи с этим мы сформировали 5 файлов примерно равного объема, в каждый из которых было добавлены отрезки текстов авторов-мужчин (не более 1000 слов от каждого автора) и тексты авторов-женщин (в связи с меньшим числом авторов-женщин использовались несколько отрезков по 1000 слов, которые добавлялись к текстам авторов-мужчин в каждый из файлов), при этом авторы в файлах не совпадали. Тексты женщин и мужчин также объединялись в один файл при формировании контрольного корпуса. Таким образом, влияние авторского и гендерного сигнала было устранено.

Нами были сформированы три подкорпуса: «Экстраверсия», «Эмоциональная лабильность», «Маскулинность», в каждый из которых вошли тексты Extr, а также тексты авторов из группы сравнения, разделенные на три подгруппы по степени выраженности соответствующей психологической характеристики.

Итоговая статистика корпуса представлена в табл. 1.

Таблица 1 – Статистика корпусов текстов

Авторы форума				Авторы из группы сравнения					
Имя файла	Длина в словах	Имя файла	Длина в словах	Имя файла	Длина в словах	Имя файла	Длина в словах	Имя файла	Длина в словах
				Подкорпус «Экстраверсия»		Подкорпус «Эмоциональная лабильность»		Подкорпус «Маскулинность»	
Extr_1	29000	Extr_4	27000	User_High_E	22862	User_High_Emo	24171	User_High_Masc	21219
Extr_2	31000	Extr_5	27000	User_Medium_E	24506	User_Medium_Emo	23847	User_Medium_Masc	22763
Extr_3	27000			User_Low_E	18758	User_Low_Emo	19965	User_Low_Masc	20089
		Всего	141 000	Всего	66 126	Всего	67 983	Всего	64 071

2.2. Методы исследования

В качестве методов исследования нами были использованы методы многомерного анализа, направленные на поиск естественной структуры в данных, – метод главных компонент (PCA); иерархический кластерный анализ с использованием различных мер расстояния и валидацией результатов путем построения консенсусных деревьев; консенсусные сети (подробнее об указанных методах см.: [18]). Анализ проводился при помощи пакета RStylo [15].

Данные методы позволяют наглядно оценить степень близости анализируемых текстов, что непосредственно связано с целью нашего анализа. Кроме того, указанные методы дополняют друг друга: так, кластерный анализ позволяет визуализировать большое число текстов, но результаты визуализации сложно интерпретировать, поскольку требуются дополнительные процедуры для выявления признаков, вносящих вклад в разделение текстов на кластеры. Реализация метода PCA в RStylo в свою

очередь позволяет отобразить на графике веса признаков, вносящие вклад в компоненту ('loadings').

Поскольку нами ставится задача выявления степени близости языковых профилей участников ЭФ с профилями лиц с известными психологическими характеристиками, но при этом анализируемые тексты отличались по тематике, мы использовали в качестве языковых признаков максимально не зависящие от темы, к которым традиционно относятся: частоты самых частотных слов корпуса (основную их часть составляют служебные слова), части речи и их последовательности, меры сложности текста. Отметим, что к часто используемым в диагностическом моделировании автора текста признаком относятся также п-граммы символов, однако в нашем исследовании мы не использовали их в силу низкой интерпретируемости результатов (подробнее о языковых признаках, используемых в диагностическом моделировании автора текста, см. в работе [4]).

Список самых частотных слов (СЧС) корпуса был составлен нами следующим образом (лемматизация не проводилась):

1) был построен список слов, встречающихся во всех текстах объединенного корпуса. Таких слов оказалось 510;

2) далее из них были отобраны 200 СЧС (после этого значения в частотном списке находится большое число полнозначных слов);

3) из 200 СЧС были удалены полнозначные слова.

Частеречная разметка проводилась с использованием морфоанализатора TreeTagger (<http://corpus.leeds.ac.uk/mocky/>) с полным набором тэгов. Также вычислялись частоты биграмм частей речи.

Маркеры сложности вычислялись при помощи программы QUITA (всего анализировалось 15 параметров: TTR, h-Point, Entropy, Average Tokens Length, R1, RR, RRmc, Lambda, Adjusted Modulus, G, R4, Нарах Percentage, L, WritersView, CurveLength RIndex; подробнее о параметрах см.: [17]). Известно, что все параметры данной группы зависят в той или иной степени от длины текста. В

нашем исследовании мы выбрали решение сократить длину текста, сделав ее равной длине самого короткого фрагмента, – 18758 слов.

Тексты Extr сравнивались с текстами User в трех сериях экспериментов (отдельно для каждой психологической характеристики), при этом эксперименты проводились как в режиме «без сэмплирования» (no sampling), так и с сэмплированием, то есть с разделением файлов на отрезки равной длины (normal sampling) (были проведены эксперименты с разными значениями длины текста).

В общей сложности нами было проведено несколько сотен экспериментов.

3. Результаты и обсуждение

Все без исключения проведенные нами эксперименты с различными настройками (как с применением сэмплинга, так и без него) и с использованием разных методов анализа данных показали, что самые частотные слова корпуса, традиционно считающиеся контекстно-независимыми (напомним, что нами использовались только слова из верхней зоны частотного списка; единичные слова из этой зоны списка, так и иначе связанные с тематикой текста, были удалены), разделяют тексты лиц группы сравнения и тексты авторов ЭФ. Указанный результат может быть объяснен фактором жанра (форум является диалогическим жанром, в то время как в корпусе для сравнения содержатся различные типы текстов: письмо другу, описание картины и т.д.).

На рис. 1 приведен типичный пример результатов кластерного анализа с использованием в качестве признаков 160 СЧС, отобранных по описанной выше схеме, для текстов подкорпуса «Экстраверсия». Все тексты объединяются в два крупных кластера: кластер текстов авторов ЭФ и авторов группы сравнения. Примечательно, что внутри кластера текстов авторов группы сравнения выделяются кластер текстов лиц с высоким уровнем экстраверсии и кластер текстов лиц с низким и средним уровнями выраженности данной характеристики, что может свидетельствовать о высоком потенциале данной группы языковых признаков для определения уровня экстраверсии автора текста.

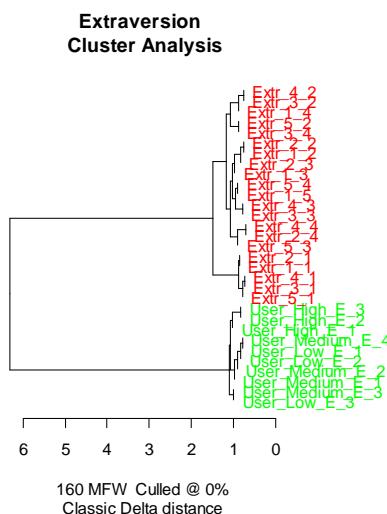


Рис. 1 – Результаты кластерного анализа для подкорпуса «Экстраверсия»
(признаки – самые частотные слова корпуса)

Результаты РСА представлены на рис. 2. На этом рисунке также показаны слова, вносящие наибольший вклад в компоненты, при этом расстояния между словами отражают их совместную встречаемость в тексте. Слова, расположенные слева и спра-

ва от горизонтальной оси, взаимно противоположны [21]: когда частотность одной группы слов растет в данном тексте, частотность другой падает (то же относится и к вертикальной оси).

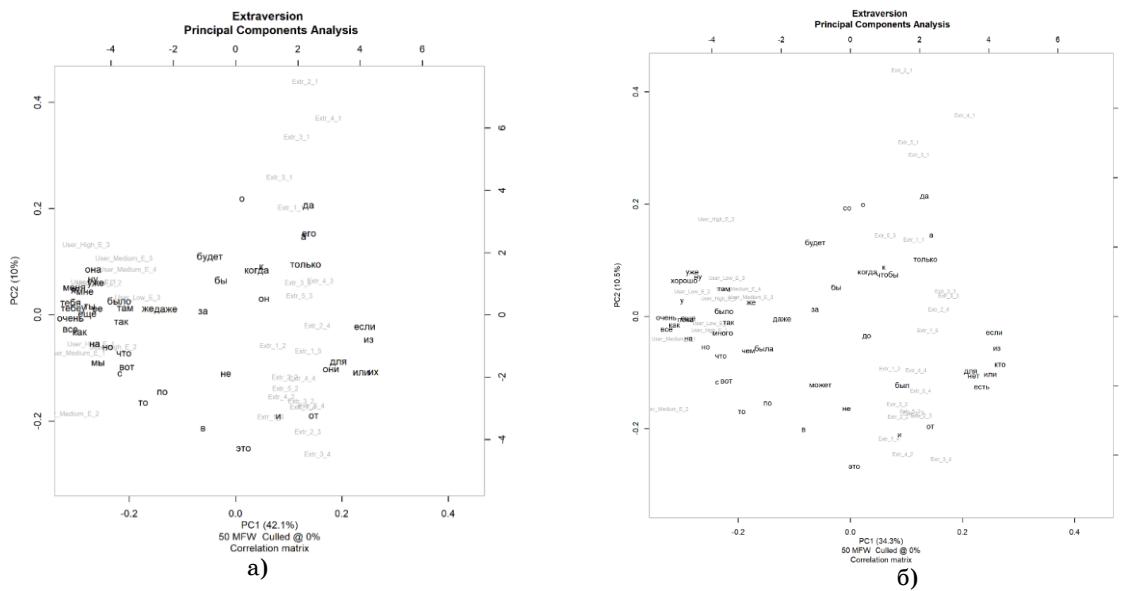


Рис. 2 – Результаты РСА для 50 самых частотных слов корпуса:
а – с личными местоимениями; б – без личных местоимений

Как мы видим (рис. 2, а), первые две компоненты объясняют в сумме 52,1% вариации, причем первая компонента, которая разделяет тексты авторов ЭФ и группы сравнения, объясняет 42,1% вариации. Несмотря на то, что форум относится к диалогическим жанрам, личные местоимения второго лица (ты, тебе, тебя) более частотны в текстах лиц контрольной группы, как и местоимения 1-го лица (я, меня, мне, мы) и 3-го лица единственного числа (она). Для текстов авторов ЭФ характерны местоимения 3-го лица множественного числа (они, их) – неоднократно отмечавшаяся в работах различных авторов черта экстремистского дискурса.

Как указывается в работах специалистов по стилеметрии [15], личные местоимения могут быть тесно связаны с коммуникативной ситуацией, поэтому рекомендуется проводить дополнительные эксперименты с исключением этой категории слов. Результаты такого анализа представлены на рис. 2,

б. Мы видим, что процент вариации при сохранении того же числа СЧС снизился, указывая на то, что местоимения вносят важный вклад в разделение групп авторов, однако разделение по первой компоненте сохраняется. Результаты кластерного анализа при этом не меняются.

Второй эксперимент с подкорпусом «Экстраверсия» был проведен с использованием в качестве признаков частей речи (униграмм и биграмм, т.е. последовательностей из двух элементов). И для униграмм, и для биграмм частей речи наблюдается разделение текстов не по психологическим характеристикам авторов, а по их групповой принадлежности (User vs Extr). На рис. 3 представлены результаты анализа с использованием метода консенсусных деревьев для 20-100 самых частотных униграмм и биграмм частей речи, встречающихся не менее чем в 10% текстов.

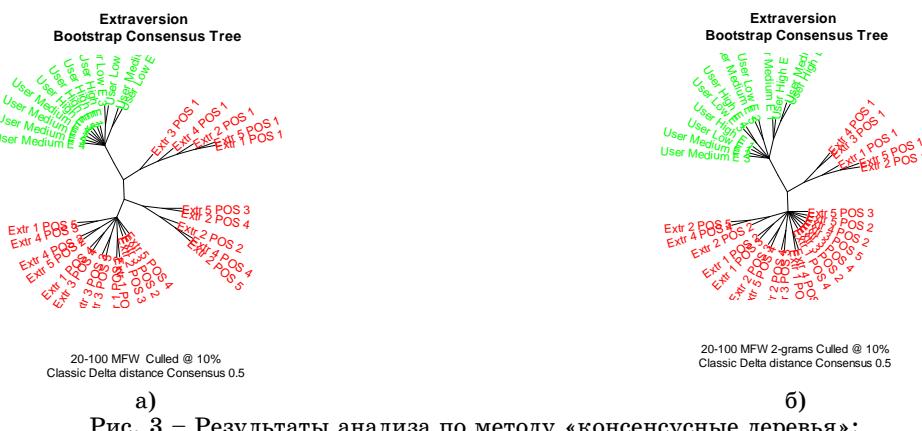


Рис. 3 – Результаты анализа по методу «консенсусные деревья»:
а – для униграмм частей речи; б – для биграмм частей речи

PCA (рис. 4) показал отчетливое разделение текстов User vs Extr на основе их морфологического состава (первая компонента). В целом следует отметить, что в экспериментах с морфологическими признаками не прослеживается тенденция разделе-

ния авторов группы сравнения с разным уровнем экстраверсии. Возможно, указанное обстоятельство связано с тем, что нами был использован относительно большой набор морфологических тегов.

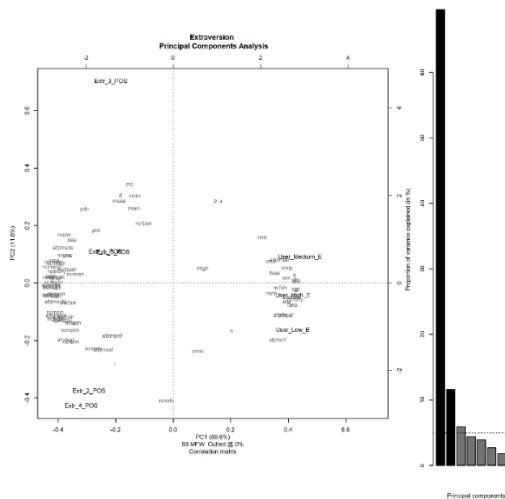


Рис. 4 – PCA с частями речи в качестве признаков

При анализе подкорпуса «Экстраверсия» с использованием маркеров сложности также было отмечено разделение текстов авторов ЭФ и авторов группы сравнения (рис. 5) по первой компоненте, а

также тенденция к кластеризации большей части текстов авторов ЭФ с текстами авторов группы сравнения с низким уровнем интроверсии.

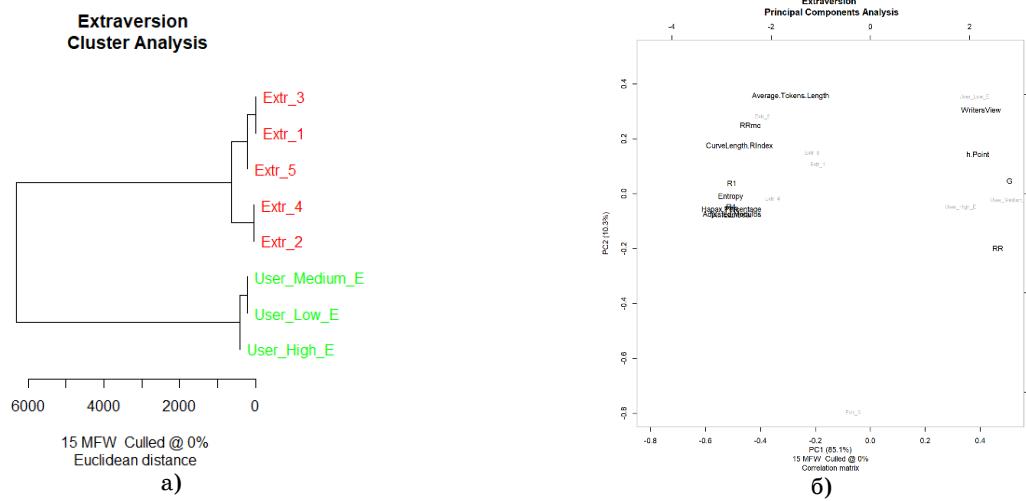


Рис. 5 – Кластерный анализ (а) и PCA (б) с маркерами сложности текста

Анализ подкорпуса «Эмоциональная лабильность» показал, как и в случае с подкорпусом «Экстраверсия», разделение авторов по группам (как при использовании местоимений в частотном списке, так и с их удалением), однако выраженных тенденций в аспекте близости текстов авторов ЭФ к текстам авторов с тем или иным уровнем эмоциональной лабильности не наблюдалось. Кроме того, в отличие от экстраверсии, по данной характеристике не удалось обнаружить разделения текстов авторов с использованием СЧС в качестве признаков.

Анализ морфологического уровня также не обнаружил разделения на группы авторов по уровню эмоциональной стабильности. На рис. 6, а приведен пример консенсусной сети – удобного инструмента

визуализации степени сходства между объектами (в данном случае – текстами) (чем выше толщина линии на рисунке, тем устойчивее связь между узлами). При построении данной консенсусной сети в качестве признаков использовались морфологические теги (от 50 до 200 самых частотных). Как видим, наиболее сильные связи обнаруживаются внутри групп (тексты авторов ЭФ наиболее близки к текстам других авторов ЭФ; то же относится и к группе сравнения).

PCA с использованием мер сложности (рис. 6, б) также показал разделение по группам авторов по первой компоненте, объясняющей большую часть вариации (83,4%), однако по второй компоненте происходит разделение на средний/низкий и высокий/высокий.

кий уровня, при этом большая часть текстов авторов ЭФ обнаруживают схожесть с текстами лиц со

средним/низким уровнем эмоциональной лабильности.

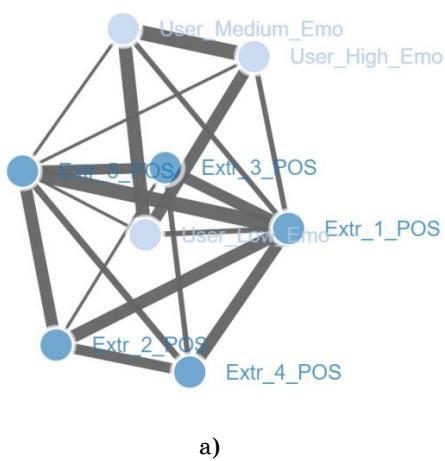
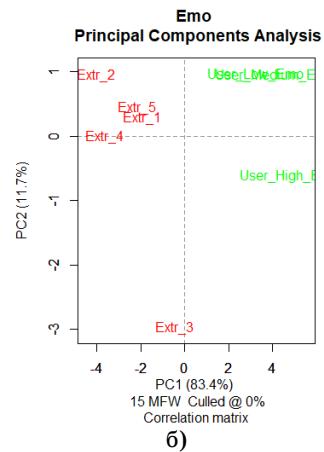


Рис. 6 – Результаты анализа подкорпуса «Эмоциональная лабильность»:
а – консенсусная сеть; б – PCA



При анализе подкорпуса «Маскулинность» был также обнаружен общий тренд разделения текстов по типам авторов (для всех групп языковых признаков). Различные типы анализа показали, что тексты авторов ЭФ объединяются в группы с текстами авторов с разным уровнем маскулинности, что может свидетельствовать о наличии среди авторов ЭФ лиц с разным уровнем данного признака.

Выводы

Проведенный анализ позволил сделать следующие выводы.

В первую очередь, следует отметить, что проблема степени независимости от контекста различных групп признаков текста, традиционно использующихся в стилеметрических исследованиях, требует дальнейшего всестороннего исследования. Так, в нашей работе в качестве признаков использовались такие считающиеся контекстно независимыми параметры, как частотности строевых слов (как с местоимениями, так и без), частотности слов разных частей речи и их биграмм, показатели сложности текста. Однако во всех сериях экспериментов при использовании данных признаков тексты разделяются на группы на основании источника (форум либо же корпус сравнения), несмотря на то, что в обоих случаях анализу подвергались тексты «непрофессиональной» письменной речи (хотя тексты авторов ЭФ, безусловно, могут содержать заимствования из различных источников). Кроме того, с высокой вероятностью авторы ЭФ и авторы текстов из корпуса сравнения принадлежат одной возрастной группе. Таким образом, различия в коммуникативной ситуации и целеустановке авторов могут отражаться и на уровне считающихся контекстно независимыми признаков текста.

Различные типы языковых признаков обладают различным диагностирующим потенциалом в отношении тех или иных психологических характеристик авторов. Так, экстраверсия проявляется на уровне самых частотных слов корпуса, а также на уровне показателей сложности. Эмоциональная стабильность в нашем исследовании оказалась связанной только с показателями сложности текста. Морфологические признаки, широко использующиеся в работах по диагностированию личности по тексту, в нашем анализе разделяли тексты на основании принадлежности автора к группе участников ЭФ либо же к контрольной группе. Возможно, данное обстоятельство связано с широким набором тегов, используемых в нашем исследовании. Очевидно, вопрос выбора параметров текста для диагностирования той или иной личностной характеристики с учетом коммуникативной ситуации должен быть исследован более тщательно. Кроме того, отметим необходимость более широкого использования показателей сложности текста в диагностическом моделировании его автора (в настоящее время данный тип признаков используется в небольшом числе работ в указанном направлении).

Наш анализ показал близость текстов авторов ЭФ к текстам лиц с низким уровнем экстраверсии и средним и низким уровнями эмоциональной лабильности.

Несмотря на то, что в нашем исследовании отсутствуют в силу объективных причин данные психологического обследования авторов ЭФ, проведенный анализ позволил получить результаты, которые могут быть использованы для дальнейшей разработки проблемы диагностирования личности по тексту, а также проблемы обнаружения экстремистски настроенных авторов.

СПИСОК ЛИТЕРАТУРЫ:

1. Абдулганиев, Р. Р. Криминологический портрет личности современного экстремиста [Текст] / Р. Р. Абдулганиев, С. Н. Миронов // Вестник Казанского юридического института МВД России. – 2018. – №1 (31). – С. 42–47.

2. Загоровская, О.В. Естественная (непрофессиональная) письменная речь как модус существования современного русского языка и «зеркало» идиолекта его носителя [Текст] / О.В. Загоровская // Известия Воронежского государственного педагогического университета. – 2019. – №2. – С. 202-206.
3. Казберов, П.Н. Психологический профиль террориста [Текст] / П.Н. Казберов, Б.Г. Бовин, А.А. Фасоля // Психология и право. – 2019. – Т. 9. – №3. – С. 141–157.
4. Литвинова, Т.А. Идентификация и диагностирование автора письменного текста [Текст] / Т.А. Литвинова, О.А. Литвинова. – Воронеж : Изд-во ВГПУ, 2015. – 332 с.
5. Литвинова, Т.А. Идиолект как объект корпусной идиолектологии: к становлению нового лингвистического направления [Текст] / Т.А. Литвинова // Ученые записки Новгородского государственного университета имени Ярослава Мудрого. – 2019. – № 7 (25). – С. 1–4.
6. Литвинова, Т.А. Построение речевого портрета участника экстремистского форума с использованием методов корпусной лингвистики [Текст] / Т.А. Литвинова, О.А. Литвинова // Известия Воронежского государственного педагогического университета. – 2019. – №1. – С. 217-222.
7. Литвинова, Т.А. Стилеметрическое исследование текстов участников экстремистского форума: гендерный аспект [Текст] / Т.А. Литвинова // Известия Воронежского государственного педагогического университета. – 2019. – №4. – С. 227-236.
8. Майкова, Е.С. Психологические особенности подростков, склонных к экстремистскому поведению [Текст] / Е.С. Майкова, Т.И. Бонкало // Ученые записки Российского государственного социального университета. – 2013. – №2(113). – С. 122-125.
9. Петрянин, А. В. Личность экстремиста: криминологический аспект [Текст] / А. В. Петрянин // Юридическая наука и практика: Вестник Нижегородской академии МВД России. – 2012. – № 20. – С. 105-107.
10. Стёпин, Д. С. Особенности осуществления террористической агитации и пропаганды с использованием интернет-ресурсов (на примере форума «Кавказ-Чат») [Текст] / Д. С. Стёпин // Проблемы теории и практики борьбы с экстремизмом и терроризмом: материалы научно-практической конференции. – М. : Российская криминологическая ассоциация; Ставрополь : Изд-во СКФУ, 2015. – С. 25-32.
11. Хомяченкова, О.В. О структуре личности преступника, совершившего преступление экстремистской направленности [Текст] / О.В. Хомяченкова // Вестник Московского университета МВД России. – 2016. – № 6. – С. 78-80.
12. Psychology and morality of political extremists: evidence from Twitter language analysis of alt-right and Antifa [Text] / M. Alizadeh [et al.] // EPJ Data Sci. – 2019. – № 8(17). – P. 1-35.
13. Baele, S. J. Lone-Actor Terrorists' Emotions and Cognition: An Evaluation Beyond Stereotypes [Text] / S. J. Baele // Political Psychology. – 2016. – № 38(3). – P. 449–468.
14. The Dark Tetrad and radicalization: personality profiles in young women [Electronic resource] / H. Chabrol [et al.] // Behavioral Sciences of Terrorism and Political Aggression. – 2019. – URL: <https://doi.org/10.1080/19434472.2019.1646301>.
15. Eder, M. Stylometry with R: a package for computational text analysis [Text] / M. Eder, J. Rybicki, M. Kestemont // R Journal. – 2016. – № 8(1). – P. 107-121.
16. Quantifying salient concepts discussed in social media content: An analysis of tweets posted by ISIS fangirls [Text] / S. Ghajar-Khosravi, P. Kwanten, N. Derbentseva [et al.] // Journal of Terrorism Research. – 2016. – № 7(2). – P. 79–90.
17. Kubát, M. QUITA – Quantitative Index Text Analyzer [Text] / M. Kubát, V. Matlach, R. Čech. – Lüdenscheid : RAM-Verlag, 2014. – 114 pp.
18. Litvinova, T. Analysis and Detection of a Radical Extremist Discourse Using Stylometric Tools [Text] / T. Litvinova, O. Litvinova // Advances in Intelligent Systems and Computing. – 2020. – Vol. 1114. – P. 30-43.
19. Nouh, M. Understanding the Radical Mind: Identifying Signals to Detect Extremist Content on Twitter [Text] / M. Nouh, J. R. Nurse, M. Goldsmith // Proceedings of ISI. – 2019. – P. 98-103.
20. Pennebaker, J. W. Using computer analyses to identify language style and aggressive intent: The secret life of function words [Text] / J. W. Pennebaker // Dynamics of Asymmetric Conflict. – 2011. – № 4(2). – P. 92–102.
21. Tabata, T. Dickens's Narrative Style: A Statistical Approach to Chronological Variation [Text] / T. Tabata // Revue, Information et Statistique Dans Les Sciences Humaines. – 1994. – № 30. – P. 165–182.
22. Linguistic analysis of pro-ISIS users on Twitter [Electronic resource] / J. Torregrosa [et al.] // Behavioral Sciences of Terrorism and Political Aggression. – 2019. – P. 1–15. – URL: [doi:10.1080/19434472.2019.1651751](https://doi.org/10.1080/19434472.2019.1651751).
23. Irrational beliefs and personality traits as psychological mechanisms underlying the adolescents' extremism mind-set [Electronic resource] / S. Trip [et al.] // Frontiers in psychology. – 2019. – T. 10. – URL: <https://doi.org/10.3389/fpsyg.2019.01184>.