

# ВОСПРИЯТИЕ СЕТЕВОГО КОНТЕНТА: НА МАТЕРИАЛЕ ДАННЫХ ВКОНТАКТЕ

**Харламов А.А.,**

*доктор технических наук старший научный сотрудник Института высшей нервной деятельности и нейрофизиологии РАН, Москва, профессор Департамента программной инженерии НИУ ВШЭ, Москва, профессор Кафедры прикладной и экспериментальной лингвистики МГЛУ, Москва*

**Пильгун М.А.,**

*доктор филологических наук, ведущий научный сотрудник Института языкознания РАН, г. Москва  
pilgunm@yandex.ru*

## Аннотация

Статья посвящена исследованию специфики восприятия контента в русскоязычных социальных сетях. Материалом послужила база данных социальной сети ВКонтакте. В работе использовался кроссдисциплинарный мультимодальный подход, для интерпретации данных применялся психолингвистический анализ. В качестве основного инструментария была выбрана система для автоматического смыслового анализа текста TextAnalyst 2.0., также использовались: инструмент для интеллектуального анализа текста Automap, программное обеспечение для визуализации сетевых структур Gephi (алгоритм Force Atlas 2), платформа для создания визуальной аналитики, интерактивной визуализации данных Tableau. В ходе исследования было установлено, что восприятие сетевого контента определяется сложным комплексом факторов: эффективность контента, степень реакции сетевого сообщества зависит от уровня символического капитала источника распространения контента, учета особенностей коммуникативной ситуации и успешности использования императивной стратегии.

**Ключевые слова:** социальные медиа, медиадискурс, политический дискурс, социальные акторы, семантическая роль, русский падеж.

## ВВЕДЕНИЕ

Важность изучения языка и коммуникации в русскоязычных социальных сетях, проблем эффективности контента определяется ростом количества и активности пользователей Рунета. Так, в сентябре 2018 года аудитория Рунета составила 1,773,385,284 пользователей (по данным LiveInternet). Данные, иллюстрирующие статистику за октябрь и ноябрь месяц представлены в Таблице 1.

Таблица 1

Основные показатели посещаемости сайтов Рунета

Наименование социальной сети	Авторы	Авторы	Сообщения
Просмотры	2,277,233,933	2,334,978,477	2,334,523,814
Сессии	424,100,001	411,191,791	407,634,892
Посетители	149,451,614	149,936,458	147,830,604
В среднем online	1,194,984	881,392	1,194,555
В среднем активных online	573,665	633,563	710,352
Средняя длительность (мин.)	5.6	5.9	5.9
Просмотров на посетителя	15	16	16

Самый популярный русскоязычный ресурс «ВКонтакте» (see Table 2) позиционирует себя как платформу, предназначенную для объединения людей, сервисов, компаний с помощью создания простых и удобных инструментов коммуникации. Технические и коммуникативные характеристики ресурса свидетельствуют об очень высокой степени востребованности у пользователей: пользователей в месяц — 97 млн; сообщений в сутки — 6,5 млрд; отметок «Нравится» в сутки — 1 млрд; пользователей мобильных платформ — 77 %; просмотров видео в сутки — 500 млн; языковых версий — 86. (см. <https://vk.com/about>).

Таблица 2

Статистика социальных сетей (январь 2018)

Наименование социальной сети	Авторы	Сообщения
ВКонтакте	26 633 778	410 168 754
Instagram	10 396 161	89 995 624
Facebook	1 824 256	53 266 232
Twitter	1 010 690	41 184 629
Мой мир (my.mail.ru)	66 957	4 336 042
LiveJournal	59 523	3 387 579

([https://br-nalytics.ru/statistics/author?hub\\_id=16048anddate=201708andcountry\\_id=20andperiod\\_type=mont](https://br-nalytics.ru/statistics/author?hub_id=16048anddate=201708andcountry_id=20andperiod_type=mont))

С расширением виртуального пространства в функционировании речевых структур стали происходить трансформации, вызванные необходимостью адаптации к новым коммуникативным условиям. М. Мартин считает, что «новый язык» был сформирован в интернете для того, чтобы компенсировать отсутствие невербальных и паравербальных средств общения [1] Действительно, форматы веб-коммуникаций, замещающие в определенной степени устное общение, даже при наличии аудиальных компонентов должны компенсировать отсутствие информации, которая передается с помощью паравербальных, мимических, пантомимических,

проксомимических средств. Кроме того, виртуальное пространство накладывает дополнительные специфические модусы на конкретную коммуникативную ситуацию. Конвергентность и многоаспектность онлайн-коммуникаций требуют кроссдисциплинарного подхода, mixed methods которые и применяются во многих исследованиях, представляющих различные аспекты взаимодействия в веб-среде [2–6].

Анализ контента, сетевых взаимодействий имеет уже обширную исследовательскую традицию. В частности, специфика коммуникативных процессов в социальных сетях, воздействие на политические и социальные процессы исследованы в работах [7–10].

В последнее время анализ сетевых коммуникаций представлен исследованиями в сфере мультимодальных динамических сетей [11], а также бимодальными и трехмодальными сетями [12–14]. Одномодальные коммуникации типа “друг-друг” уже мало привлекают внимание исследователей.

Широкое распространение в последние десятилетия приобретают исследования в мультимодальной перспективе [15–17]. Именно мультимодальный подход представляется наиболее адекватным при анализе сетевого контента, поскольку позволяет конвертировать данные, информацию, поступающие по различным каналам.

Анализ медиапространства как мультимодальной сферы получает все большее распространения в различных исследованиях [18–19]. Между тем, сложно не согласиться с мнением Д. Кристала, что, какая бы ни была культура интернета, она по-прежнему в значительной степени основана на текстах [20].

Проблема оценки контента остается важнейшей задачей, как для образовательных, так и для научных, социальных, общественных и пр. целей. На сегодняшний день существует два подхода к решению этой проблемы: с помощью технических ресурсов и непосредственного восприятия человека, экспертной оценки. В частности, автоматическая оценка текста становится достаточно популярной в сфере оценки письменных работ L2 при обучении иностранному языку и при тестировании большого количества студентов [21–24]. Так, достаточно широкое распространение и положительные оценки экспертов получила программа Coh-Metrix, хотя исследователи отмечают еще ряд проблем, связанных с ее использованием [24–25].

Важным критерием адекватности и объективности является сравнение результатов, полученных с помощью программного обеспечения и непосредственного человеческого участия, например, профессионалами-экспертами (см. например, [26]). Именно контаминация данных двух подходов, сравнение результатов, полученных автоматическими средствами и при непосредственном участии человека, с помощью аналитических процедур, представляется наиболее корректным решением проблемы выявления точности и адекватности оценки контента.

Следует остановиться на проблеме эффективности контента. Сегодня уже достаточное количество исследований, подтверждает, что компьютерные модели, используя цифровые следы людей, способны с высокой степенью точности диагностировать характер личности. Так, предиктивная аналитика на основе построения психологических портретов

и моделей поведения акторов представлена в работах [27]. Универсальным цифровым следом исследователи считают «лайк» [27–29]. Кроме того, репосты и комментарии также являются яркими маркерами востребованности контента в сетевой среде.

Таким образом, в рамках данного исследования как показатели эффективности и востребованности сетевого контента будут рассматриваться: количество лайков, репостов, и комментариев, а также достижение коммуникативной цели (в данном случае: переход из онлайн в офлайн коммуникацию, организация оппозиционного митинга).

### ВЫЯВЛЕНИЕ ПРИЧИН, ВЛИЯЮЩИХ НА ВОСПРИЯТИЕ И РАСПРОСТРАНЕНИЕ СЕТЕВОГО КОНТЕНТА

Целью данного исследования является выявление причин, влияющих на восприятие и распространение сетевого контента.

Исследовательские вопросы:

- Что определяет восприятие сетевого контента акторами?
- Что влияет на эффективность сетевого контента?
- Какой контент оказывается наиболее востребованным в сетевой среде и привлекает внимание пользователей?

Материалом для исследования послужила база данных социальной сети ВКонтакте, (март 2017 г.), связанная с оппозиционными митингами, которые прошли 26 марта под условным названием «Он вам не Димон» (#ДимонОтветит). Акции протеста были вызваны распространением в сети фильма-расследования А. Навального о деятельности премьер-министра Дмитрия Медведева.

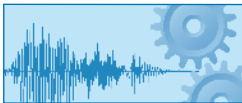
- контент вовлеченных акторов — n 43 712,
- контент активных акторов — n 15 021,
- релевантные посты — n 23 602,
- число слов — n 470 893,
- число знаков — n 3 569 442.

Контент используется на основании правил ВКонтакте<sup>1</sup> п 7.1.3., а также ст. 1274 ГК.

#### Метод исследования

В исследовании использовался кроссдисциплинарный мультимодальный подход. Для корректной интерпретации восприятия контента применялся также контент-анализ, семантический и психолингвистический анализ.

<sup>1</sup> П. 7.1.3. Пользователь, размещая на Сайте принадлежащий ему на законных основаниях Контент, предоставляет другим пользователям неисключительное право на его использование исключительно в рамках предоставляемого Сайтом функционала, путем просмотра, воспроизведения (в том числе копирования) и иные права исключительно с целью личного некоммерческого использования, кроме случаев, когда такое использование причиняет или может причинить вред охраняемым законом интересам правообладателя.



Формальный анализ был выполнен поэтапно, сочетая количественные и качественные методы анализа.

### Процедуры

Последовательность действий при проведении анализа соответствовала следующему алгоритму работы:

- I. Выделение пользователей, вовлеченных в семантическое поле «Он вам не Димон» (#ДимонОтветит).
  - I.1. Выделение активных акторов, генерирующих контент в рамках данного семантического поля.
  - I.2. Анализ коммуникативных сетевых действий и цифровых следы активных акторов.
- II. Анализ контента.
  - II.1. Выявление общей сети понятий, характеризующей всю собранную базу данных, и тематической структуры анализируемого контента (с помощью TextAnalyst 2.0. и Automap).
  - II.2. Выделение двух типов постов: датасет 1 с максимальным количеством цифровых следов и датасет 2 с нулевыми показателями цифровых следов. Был проведен анализ обеих групп с помощью TextAnalyst2.0.
  - II.3. Проведение сравнительного анализа датасета 1 и датасета 2 с помощью метода семантического дифференциала.

На этапе исследования контента с помощью программы TextAnalyst были сформированы общая для всего корпуса семантическая сеть, и извлеченная из сети тематическая структура анализируемого контента.

Полученный из семантической сети перечень наиболее значимых понятий и концептов (имеющих наивысший ранг), несущих основную смысловую нагрузку, позволил определить ядро информации, наиболее важные для акторов семантические акценты.

Выделение тематической структуры выбранного сетевого контента позволило описать содержание в виде иерархии связанных тем и подтем, которые отражают основные концепты и соответствуют узлам сети понятий. Используемый инструментарий позволил создавать иерархическую тематическую структуру, выявить базовую тему и раскрывать связи с подтемами, которые формируют разноуровневые сематические сети.

Кроме того, контент был проанализирован с точки зрения связности тематического дерева. Были выделены тематические кусты, появившиеся в результате изменения порога по весу связей в сети понятий (разрыв более или менее сильных связей). Подобный анализ собранного датасета позволил проанализировать структуру сводного медиатекста в различных срезах и на разных уровнях семантической глубины.

### Инструментарий

*TextAnalyst 2.0.* — система для автоматического смыслового анализа текста. Технология TextAnalyst предназначена для автоматического

статистического формирования однородной (ассоциативной) семантической сети текста (являющейся смысловым портретом текста), описывающего некоторую ситуацию. Семантическая сеть является графом, множество вершин которого соответствует концептам текста, а дуги соответствуют связям этих концептов в тексте. И вершины сети и дуги имеют весовые характеристики, отражающие ранги концептов и их связей в тексте.

Такой смысловой портрет позволяет выявить смысловое ядро текста, в наибольшей степени характеризующее ситуацию, описанную в тексте.

Сформированная семантическая сеть текста может быть использована для выявления тематической структуры текста. Минимальный древовидный подграф семантической сети, полученный из нее путем разрыва слабых связей и выявления наиболее весомой вершины, оказывается оглавлением текста, характеризующим его разделы с точки зрения соотношения взаимосвязанных родительских и дочерних тем разных уровней.

Семантическая сеть текста формируется в несколько шагов (1) На первом шаге из текста удаляется информация, порождающая информационный шум: удаляются стоп-слова, рабочие и общепотребимые слова, проводится лемматизация, что делает последующий анализ устойчивым. (2) На втором шаге формируется частотная сеть текста, объединением в цепочку пар слов, встречающихся в предложениях текста. Частота встречаемости слов в тексте в этой сети характеризует веса вершин, а частота встречаемости пар слов в предложениях текста характеризует веса дуг. (3) На третьем шаге весовые характеристики вершин пересчитываются с использованием итерационной процедуры с учетом связности вершин в сети. Вершины, связанные с большим числом вершин с большими весами, получают больший, по сравнению с другими вершинами, вес.

Переранжирование вершин, таким образом, учитывает связность слов в тексте на глубину  $n$  шагов — где  $n$  соответствует числу итераций в процедуре переранжирования (в программе TextAnalyst  $n=10$ ). Учет связей слов в тексте на глубину  $n$  шагов отличает алгоритм построения тематического дерева от алгоритмов тематического моделирования, в которых используется т.н. монгограммная модель текста (мешок слов: bag-of-words), в которой слова рассматриваются как несвязанные между собой в тексте.

*Automap* — инструмент для интеллектуального анализа текста, позволяющий извлекать информацию с использованием методов сетевого анализа, поддерживает извлечение нескольких типов данных из неструктурированных документов.

*Gephi* (алгоритм Force Atlas 2) — программное обеспечение для визуализации сетевых структур [30].

*Tableau* — платформа для создания визуальной аналитики, интерактивной визуализации данных.

### Результаты и дискуссия

I. На первом этапе исследования были выделены все пользователи, проявившие интерес к оппозиционному митингу. Вовлеченные в лингвистический модус анализировались по вербальному контенту, собранному по хештегам #димонответит, содержащему оригинальный авторский контент и контент пользователей об активных акторах в рамках указанного семантического поля.



### Схема анализа вербального контента

1. Выделение и анализ семантического поля.
  - 1.1. Эксплицитные средства выражения интенции, оценки, мнения.
  - 1.2. Прямые средства воздействия.
    - 1.2.1. Орфографические и графические средства.
    - 1.2.2. Речевые средства:
      - 1.2.2.1. Фонетические.
      - 1.2.2.2. Лексические.
      - 1.2.2.3. Морфологические.
      - 1.2.2.4. Синтаксическое.
      - 1.2.2.5. Стилистические.
    - 1.3. ИмPLICITное выражение интенции, оценки, мнения:
      - 1.3.1. Орфографические и графические средства.
      - 1.3.2. Речевые средства:
        - 1.3.2.1. Лексические.
        - 1.3.2.2. Морфологические.
        - 1.3.2.3. Синтаксическое
        - 1.3.2.4. Стилистические.
      - 1.3.3. Риторические средства.
      - 1.3.4. Средства речевого воздействия.
      - 1.3.5. Ассоциативные связи.

Результаты сбора данных, анализа и кластеризации вовлеченных пользователей представлены в виде визуализации (Рис. 1).

Алгоритм кластеризации вершин графа был выполнен по [31]. При анализе использовался только контент, генерированный реальными пользователями. Выявление искусственных сущностей проводилось с помощью анализа профиля: отсутствие друзей, отсутствие или незначительное количество визуальных данных (фото без лица конкретного пользователя), отсутствие постов других пользователей со ссылками на актора, активность только в определенные периоды, соответствующие знакомым временным периодам, связанными с политическими событиями, а также вхождение в уже известные бот-сети. В ходе исследования использовалась технология SocialDataHub, которая позволяет автоматически анализировать профиль в социальной сети.

- I.1. Далее был проведен анализ, позволивший выделить активных акторов, генерировавших контент в рамках данного семантического поля и которые приняли непосредственное участие в протестных мероприятиях. Анализ проводился с использованием вербального модуля (авторского контента) и визуального модуля — по данным статичных визуальных материалов с помощью алгоритма SocialDataHub, который позволяет достигать точности до 85 %.  
Анализ основных характеристик акторов и трех кластеров участников по целевым установкам.
- I.2. Анализ коммуникативных сетевых действий позволил выделить и ранжировать цифровые следы активных акторов для дальнейшего исследования.

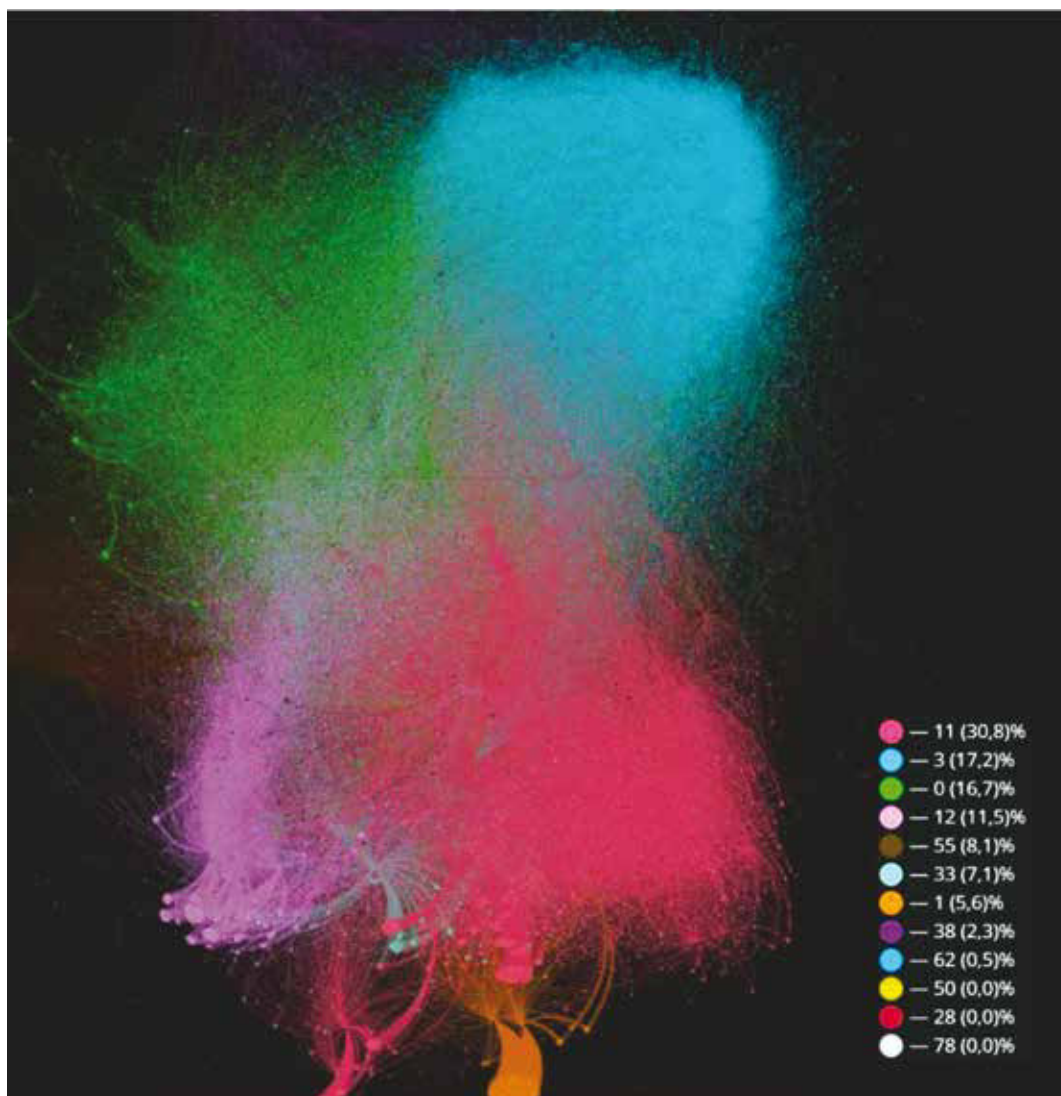


Рис. 1. Вовлеченные пользователи, отреагировавшие на митинг 26 марта.

**Анализ коммуникативного сетевого поведения проводился по схеме:**

1. Количество лайков.
2. Количество постов по релевантной теме.
3. Коммуникативные ресурсы:
  - 3.1. Предпочтительные стратегии (ритуальные, убеждающие, императивные, провокационные).
  - 3.2. Тактики прямого коммуникативного воздействия.
  - 3.3. Тактики косвенного коммуникативного воздействия.
4. Количество репостов по релевантной теме.
5. Количество комментариев.
6. Количество ссылок.
7. Количество и характеристика групп.
8. Количество сообщений.
9. Интегрированные коммуникативные ресурсы косвенного воздействия.



Результаты представлены на Рис. 2.

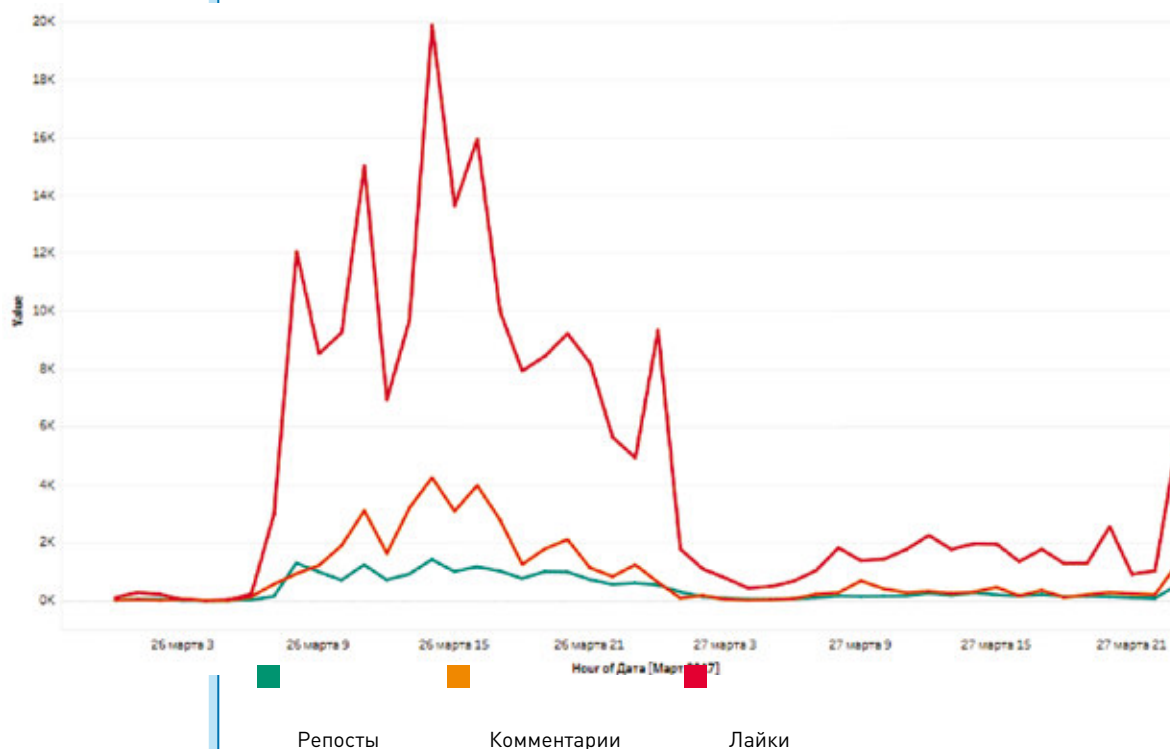


Рис. 2. Соотношение цифровых следов (лайки, репосты, комментарии)

## III. АНАЛИЗ КОНТЕНТА.

III.1. На начальном этапе исследования контента были выявлены частотные концепты, общая сеть понятий, характеризующая всю собранную базу данных, и тематическая структура анализируемого контента (см. Рис. 3).

Набор наиболее значимых понятий и концептов, несущих основную смысловую нагрузку, позволил определить ядро информации, наиболее важные для акторов семантические акценты. Выделение тематической структуры выбранного сетевого контента позволило описать содержание в виде иерархии связанных тем и подтем, которые отражают основные концепты и соответствуют узлам сети понятий. Используемый инструментарий позволил создавать иерархическую тематическую структуру, выявить базовую тему и раскрывать связи с подтемами, которые формируют разноуровневые сематические сети.

Кроме того, контент был проанализирован с точки зрения связности тематического дерева. Также были выделены тематические кусты, появившиеся в результате изменения порога по весу связей в сети понятий (разрыв более или менее сильных связей). Подобный анализ собранного датасета позволил проанализировать структуру сводного медиатекста в различных срезах и на разных уровнях семантической глубины (см. рис. 4).

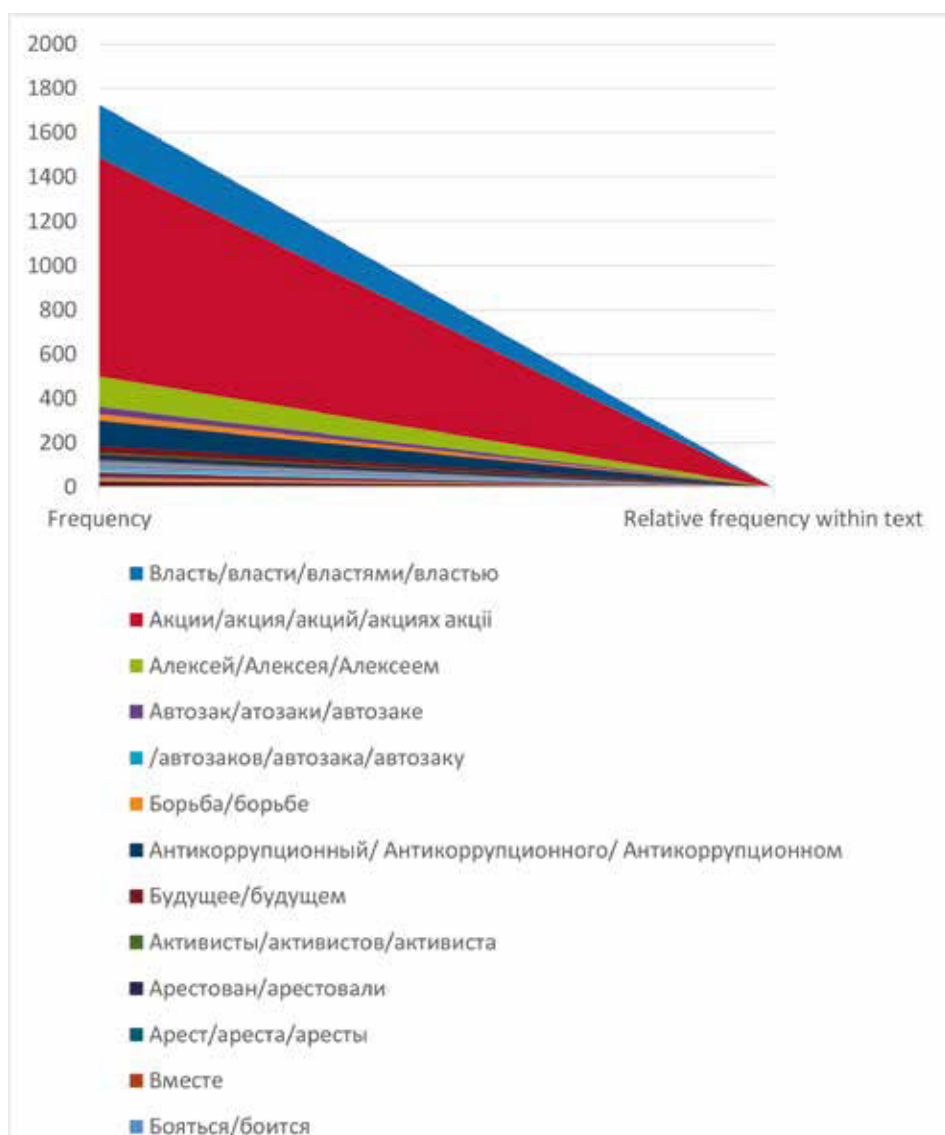


Рис. 3. Частотные концепты.

Закономерно, что вся информационная иерархия в данном контенте подчинена единой теме, имеет вид дерева с единственным корнем: «ДимонОтветит». Максимальный вес понятия (100) оказался у следующей номинаций: ДимонОтветит/Медведев/Димон/ОнВамНеДимон/ ОнНамНеДимон, митинг/протест против коррупции, Россия/ страна, Тверская, города, задержан.

Максимальный вес связи для концепта коррупция в общей базе данных у номинаций: митинги против коррупции (74), города (45), Медведев (42)/Димон (38), страна (41).

Анализ тематической структуры собранного контента отражает основное содержание интегрированных медиатекстов, связанное с проведением по всей стране митингов против коррупции, и персоной чиновника из высшего эшелона власти, вызвавшего возмущение общества и послужившего катализатором протестного движения.

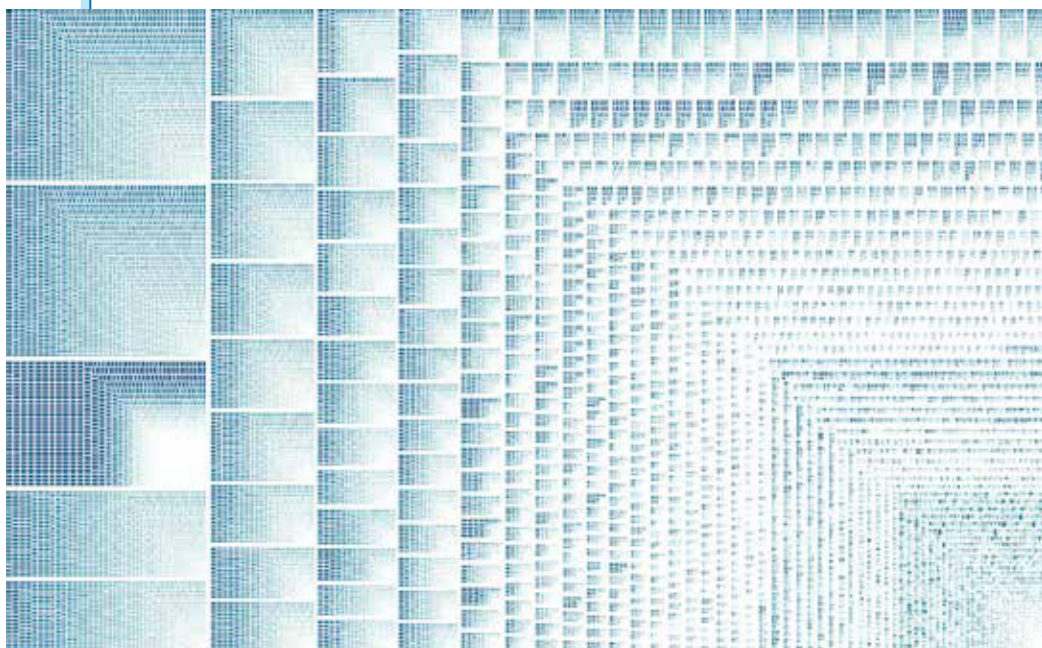


Рис. 4. Тематическая структура общей базы данных.

II.2. В соответствии с выбранными критериями эффективности и востребованности контента было проведено ранжирование постов по количеству цифровых следов.

II.2.1. Далее были выделены медиатексты, вызвавшие наибольший резонанс в сетевой среде (датасет 1).

Характеристики датасета 1:  
размер памяти — 98 КБ;  
число постов — 40;  
число слов — 2 226;  
число знаков — 16 031;  
число лайков — 52 040;  
число комментариев — 7 295;  
число репостов — 3 813.

А также медиатексты, не вызвавшие вообще никакой реакции у сетевого сообщества (датасет 2).

Характеристики датасета 2:  
размер памяти — 97 КБ;  
число постов — 119;  
число слов — 1 830;  
число знаков — 14 763;  
лайков — 0;  
комментариев — 0;  
репостов — 0.

II.2.2. Построение семантических сетей обоих датасетов позволило выявить наиболее значимую информацию, на которой делают акцент авторы сообщений.

В семантической сети 1 наибольший вес имеют понятия, обозначающие инициатора протестный движений. Центр семантической сети, актуальная информация персонафицирована, внимание пользователей концентрируется на политическом лидере, инициировавшем митинги, который становится безусловным лидером мнений, значительно расширяет виртуальный электорат.

Любопытно, что семантическая сеть 2 представляет более широкий набор информационных кластеров. Максимальный вес имеют понятия, характеризующие причины, вызвавшие протестные митинги. Центр семантической сети, актуальная информация связаны также не только с лидером оппозиции, но и представляют организацию (ФБК), которая провела антикоррупционное расследование, семантический акцент делается также на реакции властей, негативном отношении к информационной политике СМИ и др.

Тематическая структура датасета 1 (митинг — 100) несколько также не уступает датасету 2 (митинги — 100; огнетушитель — 99).

Ассоциативные сети с ядерной номинацией «Митинг», выявленные по материалам 1-го и 2-го датасета в целом идентичны (см. рис. 5, 6). Между тем, следует отметить, что ассоциативный поиск по базе датасета 1 позволил выявить большую значимость эффекта освещения проведенных митингов, их территориальное распространение, задержания и пр.

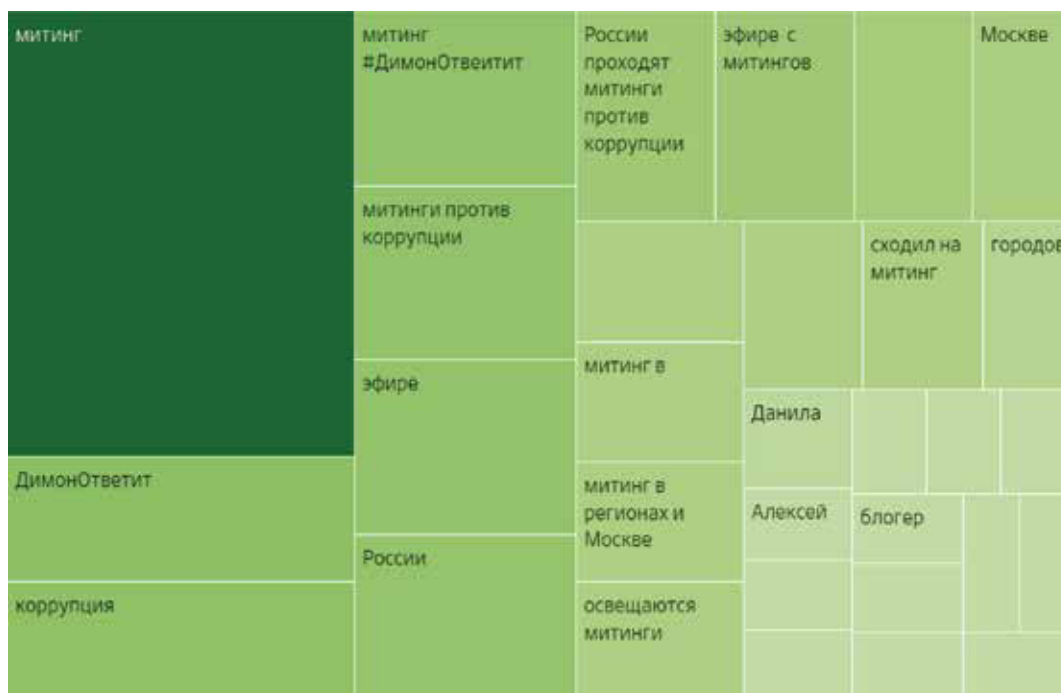


Рис. 5. Ассоциативный поиск по базе данных 1.

II.3. На заключительном этапе исследования восприятие контента проверялось при непосредственном участии испытуемых с помощью методики семантического дифференциала. Фокус-группа состояла из 20 участников (возраст 20–25 лет, 60 % — жен., 40 % — муж.). Результаты также не показали значительного расхождения в восприятии базы данных 1 и базы данных 2 (см. рис. 7).

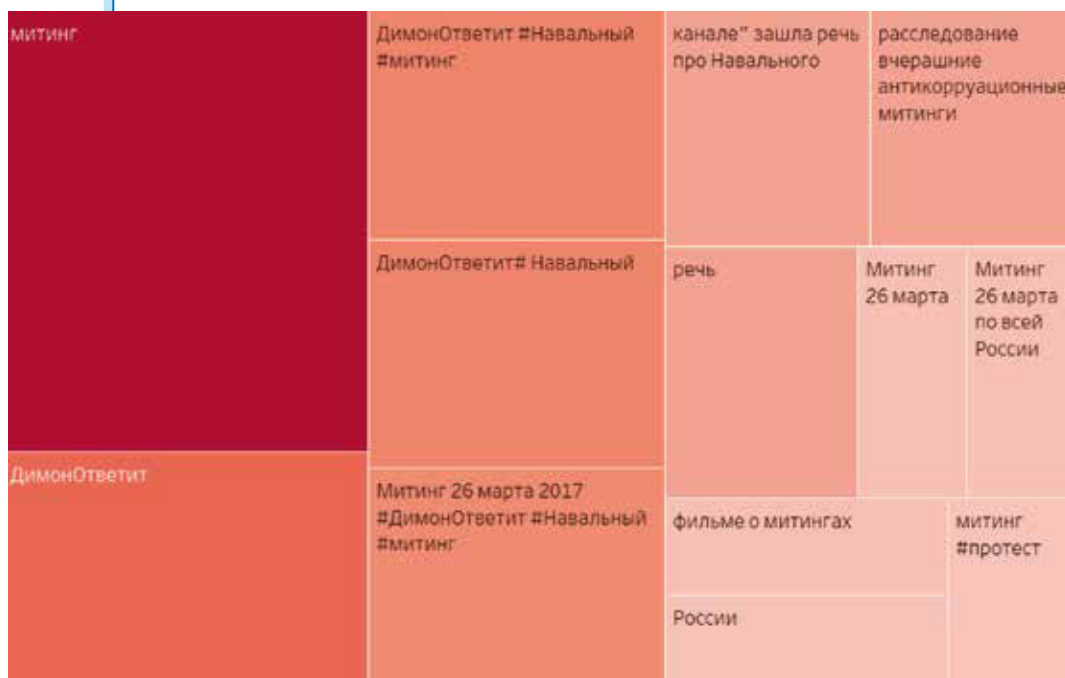


Рис. 6. Ассоциативный поиск по базе данных 2.

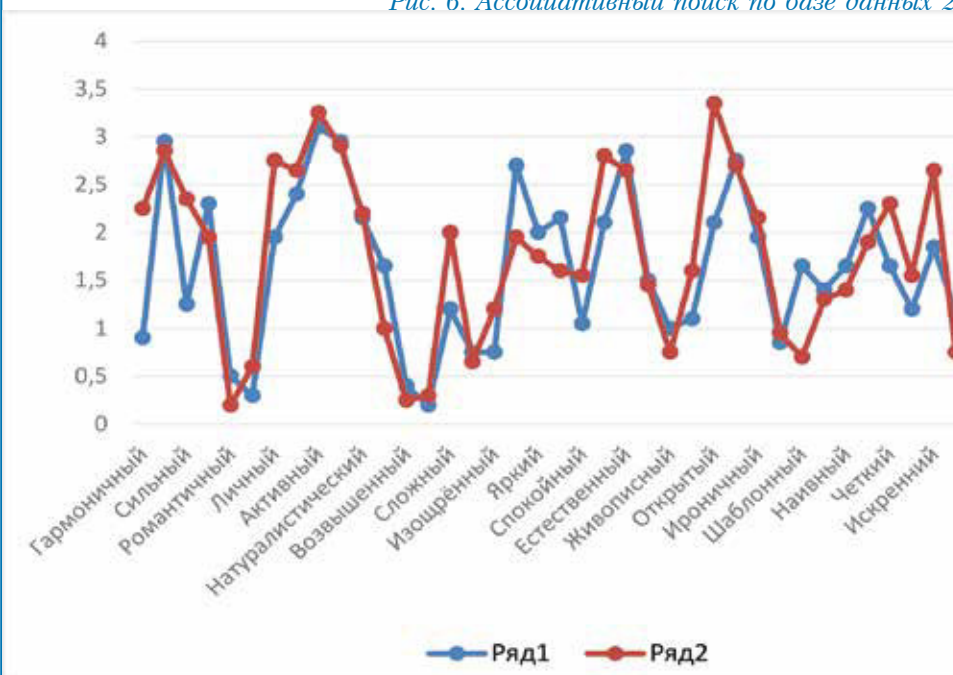


Рис. 7. Результаты, полученные с помощью семантического дифференциала.

Таким образом, семантические и лингвистические характеристики базы данных 1 и базы данных 2 практически не различаются, в отличие от характеристик сетевой эффективности, которые противопоставляют два данных кластера как максимально эффективный и минимально эффективный.

## ЗАКЛЮЧЕНИЕ

Результаты исследования позволяют сделать вывод, что восприятие и эффективность сетевого контента определяется комплексом причин.

Эффективность контента, степень реакции сетевого сообщества зависит от уровня символического капитала источника распространения контента, учета особенностей коммуникативной ситуации и успешности использования императивной стратегии.

Значимость сетевого контента определяется, в первую очередь, актором, который опознается как автор данного сообщения. Харизматичность оппозиционного лидера, а также круга его соратников, распространявших контент, составляющий базу данных 1, сыграла большую роль в интенсивности информационных волн. Востребованность контента, генерируемого лидером мнений (включая близкий круг акторов), успешность постов (максимальное количество лайков, репостов, комментариев) обеспечивает во многом символический капитал актора, который повышает значимость текста при его восприятии сетевым сообществом. Базу данных 2 генерировали обычные пользователи сети с незначительным символическим капиталом.

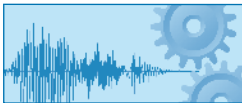
Наиболее востребованным в сетевой среде, привлекающим внимание пользователей становится контент, созданный актором с наибольшим символическим капиталом. Разумеется, следует отметить, что данное положение справедливо в большей степени в определенном сегменте сетевой среды и круге акторов, на которых распространяется влияние определенного лидера мнений.

Выбор адекватных коммуникативных инструментов, успешное решение актуальных задач позволяет лидеру мнений активно влиять на общественное мнение в сети и разрушать границы онлайн и офлайн коммуникаций, переводить виртуальные интенции в реальные действия. На акции протестов против коррупции 26 марта 2018 года вышли по официальным данным 70 000 человек (неофициальные данные значительно выше) в 100 городах России.

Контент, выделенный и проанализированный в ходе исследования, можно охарактеризовать как реализующий императивную стратегию, направленную на иррациональную сферу, несет сильный эмоциональный заряд.

Данная коммуникативная стратегия соответствовала острой потребности в российском обществе (у молодежи особенно) социальной справедливости, большой протестный потенциал деструктивного типа. В частности, несмотря на четкую адресацию акции (#ДимонОтветит), протест носил по большей части недифференцированный характер («Достало!»).

Основные характеристики адресата (активных акторов) также были учтены при реализованы в контенте: социальная пассивность акторов сочетается с прагматизмом, стремлением к комфорту и устойчивому благополучию, которые должны быть обеспечены внешними силами, а не самими участниками. Кроме того, следует выделить стремление к самореализации, гипертрофированное чувство самоуважения, эгоцентризм на фоне общей неудовлетворенности. Показательно, что приоритеты участников данной протестной акции находятся в сфере хобби (музыка, кино, технические инновации и пр.) и спорта (киберспорта).



Еще одна важная характерная черта — стремление к геймификации, которая становится ведущим коммуникативным трендом. Можно говорить о геймификации как типе языкового сознания (кеды на провода — символ, замствованный Алексеем Навальным в фильме «Хвост виляет собакой»), игрушечные уточки, поведение с ОМОНОм. В медиатексте это выражается саркастической тональностью, постмодернистским дискурсом.

Можно предположить, что рост протестного потенциала в данной ситуации связан скорее с эффективной работой команды Навального с определенной группой акторов, а не с устойчивой политической активностью.

В отношении мотивации акторов протестного медиаконтента можно заключить следующее: анализ данных позволяет свидетельствовать об отсутствии устойчивых политических предпочтений, о широком спектре политических взглядов акторов. Причем эксплицитную поддержку позиции Навального выражали только самые юные акторы, составляющие кластер «неофиты». «Идейные борцы» и «гедонисты» выражали широкий спектр оценок от безразличия до неодобрения, разочарования и сарказма в отношении оппозиционного лидера.

Отчасти подобное положение можно объяснить отсутствием общей повестки. Кроме антикоррупционного пафоса, участников митинга объединяло негативное отношение к федеральным СМИ, этическая интерпретация коррупции и рефлексия по поводу сформировавшейся оппозиция «молодежь — власть».

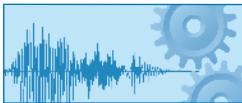
Следует заключить, что восприятие контента в сети в значительной степени зависит от актора, который опознается как автор сообщения, пропорционально объему его символического капитала и степени влияния как лидера мнений, релевантности актуальной коммуникативной ситуации, представленной в сетевой среде и эффективности использования императивной стратегии.

### Литература

1. *Martin M.* (2018) Dictionary of Digital Pictograms and Glossary for Internet Use and Portable Telephones. Cambridge Scholars Publishing, Lady Stephenson Library, Newcastle upon Tyne.
2. *Sauter T.* (2014) 'What's on your mind?' Writing on Facebook as a tool for self-formation. *J New Media & Society* 16: 823–839. doi:10.1177/1461444813495160.
3. *Lipschultz J.H.* (2014) Social Media Communication: Concepts, Practices, Data, Law and Ethics. Routledge, New York and London.
4. *Verboord M.* (2014) The impact of peer-produced criticism on cultural evaluation: A multilevel analysis of discourse employment in online and offline film reviews. *J New Media & Society*, 16: 921–940. doi:10.15405/epsbs.2018.09.02.61.
5. *Dunbar RIM, Arnaboldi V, Conti M., Passarella A.* (2015) The structure of online social networks mirrors those in the offline world. *J Social Networks* 43: 39–47. doi:10.1016/j.socnet.2015.04.005.

6. Ryan L.D'Angelo A. (2018) Changing times: Migrants' social network analysis and the challenges of longitudinal research. *J Social Networks* 53: 148–158. doi:10.1016/j.socnet.2017.03.003.
7. Tufekci Z., Wilson Ch. (2012) Social Media and the Decision to Participate in Political Protest: Observations From Tahrir Square. *J Communication* 62(2): 363–379. doi:10.1111/J.1460-2466.2012.01629.X.
8. Dunbar RIM, Arnaboldi V., Conti M., Passarella A. (2015) The structure of online social networks mirrors those in the offline world. *J Social Networks* 43: 39–47. doi: 10.1016/j.socnet.2015.04.005.
9. González-Bailón S., Wang N. (2016) Networked discontent: The anatomy of protest campaigns in social media. *J Social Networks* 44: 95–104. doi:10.1016/j.socnet.2015.07.003.
10. Brusco M.J., Doreian P. (2019) Partitioning signed networks using relocation heuristics, tabu search, and variable neighborhood search. *J Social Networks* 56:70–80. doi:10.1016/j.socnet.2018.08.007.
11. Roth C., Cointet J-P. (2010) Social and semantic coevolution in knowledge networks. *J Social Networks* 32: 16–29. doi.10.1016/j.socnet.2009.04.005.
12. Latapy M., Magnien C., Vecchio N.D. (2008) Basic notions for the analysis of large two-mode networks. *J Social Networks* 30(1): 31 — 48. doi:10.1016/j.socnet.2007.04.006.
13. Murata T. (2010) Detecting communities from tripartite networks. In: Rappa M, Jones P, Freire J, Chakrabarti S. WWW, ACM, pp1159–1160. doi:10.1145/1772690.1772853.
14. Opsahl T. (2013) Triadic closure in two-mode networks: Redefining the global and local clustering coefficients. *J Social Networks* 35. doi: 10.1016/j.socnet.2011.07.001.
15. Kress G. (2010) *Multimodality. A Social Semiotic Approach to Contemporary Communication.* Routledge, Londres.
16. Lutkewitte C. (2013) *Multimodal Composition: A Critical Sourcebook.* Bedford/St. Martin's, Boston — New York.
17. Waciewicz S., Zywickzynski P. (2017) The multimodal origins of linguistic communication. *J Language & Communication* 54: 1–8. doi:10.1016/j.langcom.2016.10.001.
18. Alexander J., Rhodes J. (2014) On multimodality: New media in composition studies. Conference on College Composition and Communication/National Council of Teachers of English, Urbana.
19. Velkova J. (2018) *Studying Emerging Data Practices: Creating a Cultural Biography of Objects Through Using the Web as an Ethnographic Resource.* Sage research methods, London. <http://methods.sagepub.com/case/emerging-data-practices-creating-cultural-biography-ethnographic-resource>.
20. Crystal D. (2004) *Language and the Internet.* Cambridge University Press, Cambridge.
21. Ericsson P.F., Haswell R.H. (2006) *Machine scoring of student essays: Truth and consequence.* Utah State University Press, Logan.
22. Deane P. (2013) On the relation between automated essay scoring and modern views of the writing construct. *J Assessing Writing* 18(1): 7–24. doi:10.1016/j.asw.2012.10.002.
23. Weigle S.C. (2013) English language learners and automated scoring of essays: Critical considerations. *J Assessing Writing* 18(1): 85–99. doi:10.24093/awej/vol9no2.11.
24. Matthews J., Wijeyewardene I. (2018) Exploring relationships between automated and human evaluations of L2 texts. *J Language Learning & Technology* 22(3): 143–158. doi:10.1177/0265532216676851.
25. McNamara D.S., Graesser A.C., McCarthy P., Cai Z. (2014) *Automated evaluation of text and discourse with Coh-Metrix.* Cambridge University Press, New York.
26. Crossley S.A., McNamara D.S. (2012) Predicting second language writing proficiency: The roles of cohesion and linguistic sophistication and their relations to judgments of essay quality. *J Research in Reading* 35(2): 115–135.
27. Kosinski M., Stillwell D., Graepel T. (2013) Private traits and attributes are predictable from digital records of human behavior. *Proc Natl Acad Sci USA* 110(15): 5802–5805.
28. Bachrach Y., Kohli P., Stillwell D., Graepel T. (2013) Manifestations of user personality in website choice and behaviour on online social networks. *J Mach Learn* 95(3):357–380.





29. Youyou Wu., Kosinski M., Stillwell D. (2015) Computer-based personality judgments are more accurate than those made by humans. J PNAS 112 (4): 1036–1040. doi:10.1073/pnas.1418680112.
30. Jacomy M., Venturini T., Heymann S., Bastian M. (2014) ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. J PLoS ONE 9(6): e98679.
31. Girvan M., Newman MEJ. (2002) Community structure in social and biological networks. Proceedings of the National Academy of Sciences 99(12): 7821–7826. doi: 10.1073/pnas.122653799.

## ANALYSIS OF PERCEPTION OF NETWORK CONTENT: ON THE BASIS OF VKONTAKTE DATA

***Kharlamov A.A.,***

*Doctor of Technical Sciences, Senior Researcher, Institute of Higher Nervous Activity RAS, Moscow, Professor, Department of Applied and Experimental Linguistics, MSLU, Moscow, Professor, School of Software Engineering HSE, Moscow*

***Pilgun M.A.,***

*Doctor of Philology, Leading Researcher, Institute of Linguistics, RAS, Moscow, pilgunm@yandex.ru*

### Abstract

The article is devoted to the study of the specifics of content perception in Russian-speaking social networks. The material formed the basis of the social network Vkontakte. We used crossbows multidisciplinary multimodal approach to data interpretation psycholinguistic analysis was applied. As chief in automatic semantics system analysis of the text, TextAnalyst 2.0., also used: a tool for Autocart text analysis software gephi network structure visualization tool (ForceAtlas2 algorithm), visual Analytics platform, interactive VI-storing table data. The study found, that the perception of network content is determined by a complex factors: content effectiveness, degree of network message response it depends on the level of symbolic capital of the source — enrichment of content, taking into account the peculiarities of the communicative situation and the success of the imperative strategy.

**Keywords:** social media, media discourse, political discourse, social actors, semantic role, russian case.