

## II. КОРПУСНЫЙ АНАЛИЗ В КОГНИТИВНЫХ ИССЛЕДОВАНИЯХ ТЕКСТА

УДК 81'33

### WORD2VEC- И ГРАФОСЕМАНТИЧЕСКИЕ МОДЕЛИ ИСПОЛЬЗОВАНИЯ ЭМОТИКОНОВ И ЭМОДЗИ В ТЕКСТАХ ИНТЕРНЕТ-КОММУНИКАЦИИ\*

К.И. Белоусов, И.А. Обухова, И.А. Лабутин

Пермский государственный национальный исследовательский университет (Пермь, Россия)  
belousovki@gmail.com

В статье изучается использование эмодзи и эмодзи в интернет-коммуникации. Материалом послужил массив комментариев пользователей соцсети ВКонтакте. С помощью метода Word2Vec построена модель контекстуальной близости использования эмодзи в текстах; она показывает более тесную контекстуальную связь отрицательных эмодзи разных типов друг с другом. В качестве гипотезы, объясняющей большую вариативность выбора негативных эмодзи при описании похожих контекстов, рассмотрено влияние психологических параметров личности на использование знаков разной тональности в схожих ситуациях общения. Гипотеза проверена с помощью размеченного веб-корпуса комментариев пользователей, прошедших опросник BFI. Установлено, что для разных психологических типов личности в большей степени характерна вариативность совместного использования отрицательных знаков; знаки же положительной тональности демонстрируют более устойчивые структурные паттерны.

*Ключевые слова:* эмодзи, эмодзи, WORD2VEC, BFI, социальные сети, интернет-коммуникация.

Для цитирования: Белоусов К.И., Обухова И.А., Лабутин И.А. WORD2VEC и графосемантические модели использования эмодзи и эмодзи в текстах интернет-коммуникации // Вопросы когнитивной лингвистики. 2023. № 2. 47-62 С.

DOI: 10.20916/1812-3228-2023-2-47-62

#### 1. Введение

Эмодзи и эмодзи как графические символы электронной коммуникации используются пользователями социальных сетей и мессенджеров для обозначения эмоционального состояния или настроения, имитации невербальных сигналов в виртуальном пространстве (мимики, жестов, поз) и дополнения основного смысла высказывания [Derks et al. 2008; Китова 2016; Hu et al. 2016; Riordan 2017]. Пользователи также могут применять эти знаки для замены отдельных слов или словосочетаний в тексте или создания в нем визуального эффекта за счет изображения. Ср.: Эмодзи – это «изображение, состоящее из знаков препинания, используемое в SMS-сообщениях и электронных письмах для выражения определенной эмоции» (перевод наш) [CALD URL: <https://dictionary.cambridge.org/dictionary/english/e>

moticon]. Эмодзи – это пиктограммы и идеограммы, представляющие собой цветные изображения, обозначающие эмоции и те или иные объекты действительности: здания, продукты питания, виды активности и др. «В эпоху взрывного распространения Интернета естественно стремиться к <...> такому расширению естественного языка, который включал бы не только обычные слова, но и символы-сокращения, эффективно кодирующие стандартные смыслы» [Браудо, Фрумкина 2006: 34]. К таким символам авторы причисляют смайлики.

Статистика показывает, что более 90% интернет-аудитории использует в текстах эмодзи и эмодзи [Павлов 2017: 133]. Их присутствие в текстах часто расценивается участниками онлайн-коммуникации в качестве нормы, т.е. эти знаки становятся неотъемлемым атрибутом вербальной электронной коммуникации, поскольку способны сопровождать, конкретизировать и частично или полностью заме-

\* Работа выполнена в рамках государственного задания ПГНИУ (проект FSNF-2020-0023).

щать вербальную часть сообщения. Пиктограммы используются в сфере торговли: покупателям интернет-магазинов предлагают оценить работу персонала или качество доставки с помощью эмоциональной шкалы эмодзи. Эмодзи также используются для привлечения внимания аудитории в мета-тегах веб-страниц с целью поисковой оптимизации веб-сайтов. Популярность эмодзи и эмодзи находит отражение и в семиотическом перекодировании знаков на естественном языке в эмодзи-знаки в мессенджерах и соцсетях (с помощью функции автозамены).

В этой связи представляет интерес изучение использования графических средств обозначения эмоций в текстах интернет-коммуникации, в частности выявление взаимосвязей между обращением авторов сообщений к эмодзи и эмодзи как маркерам эмоционального состояния субъекта речи и психологическими характеристиками личности. В статье рассматривается влияние параметров психологического теста ВФИ на использование эмодзи и эмодзи при написании комментариев в социальной сети ВКонтакте.

## 2. Эмодзи в пространстве Big Data

Изучение использования графических знаков, обозначающих эмоции (в широком смысле слова), может сочетать корпусные и экспериментальные методы. Эксперимент «работает» с выборочными данными и удобен в тех случаях, когда корпусные методы сложны в использовании. Например, когда стоит задача проведения лингвистического и/или психологического опроса, при этом результаты экспериментального сбора материала также могут быть представлены в виде корпуса. Корпусные методы уместнее применять во многих эпистемологических ситуациях, связанных, например, с анализом частотности единиц, их дистрибуции, контекстуальным анализом и др. В последние годы в гуманитарных науках начинают использоваться методы, основанные на анализе так называемых Больших данных (Big Data, BD). Эти методы могут работать как с размеченными корпусами, так и большими массивами текстов. Достоинством работы с BD является репрезентативность данных и воспроизводимость результатов.

В нашем исследовании был осуществлен анализ моделей использования эмодзи в текстах для последующего анализа в рамках экспериментальных методов. Материалом послужили 23 млн.

постов и комментариев 2,5 млн. пользователей соцсети ВКонтакте.

С использованием библиотек *Natasha* (URL: <https://github.com/natasha/natasha>), *Razdel* (URL: <https://github.com/natasha/razdel>), *Slovnet* (URL: <https://github.com/natasha/slovnet>) и *Pymorphy* [Korobov 2015] была выполнена предварительная обработка корпусов: произведена токенизация текстов (разбиение на предложения и слова); исключены стоп-слова и знаки препинания (оставлены эмодзи для дальнейшего анализа сообщений); выполнено приведение слов к словарной форме с учетом их вхождения в словосочетания.

Далее на основе предобработанных корпусов с использованием библиотеки *Gensim* [Rehůřek, Sojka 2011] было произведено построение векторных моделей по алгоритму *Word2Vec* (W2V) [Mikolov et al. 2013]. В качестве гиперпараметра размерности векторного пространства исходя из соображений относительно компактного размера корпуса было выбрано значение 50.

В качестве метрики расстояния использовался косинус угла между векторами, представляющими анализируемые единицы (в нашем случае эмодзи). Полученные матрицы размером  $N \times N$  ( $N$  – количество отобранных эмодзи) визуализировались в виде графов в приложении *SciVi* (<https://scivi.semograph.com>). Порог вхождения вершин и ребер в граф был ограничен: по вершинам  $> 0,01\%$  вхождения единицы в корпус, по ребрам  $> 0,7$  семантической близости, что соответствует углу, равному  $45^\circ$  между векторами (эмодзи).

Иными словами, каждый эмодзи был представлен в виде вектора размерности 50. Близость векторов (показатель косинуса угла между ними) свидетельствует о близости контекстного использования эмодзи: при величине равной 1 два эмодзи могут полностью замещать друг друга. Заметим, что данная модель не отражает фактов совместной встречаемости знаков в одном контексте, а только их контекстуальную близость.

Для построения модели были отобраны 46 наиболее частотных эмодзи, передающих разные эмоциональные состояния, жесты и действия пишущих (Табл. 1). Для определения значений отобранных эмодзи использовалась электронная библиотека *Emojipedia* [URL: <https://emojipedia.org>].

На рис. 1 представлен скриншот фрагмента результатов запроса к W2V-модели. В поле (1) выбирается модель W2V, построенная на основе контента ВКонтакте. В поле (2) помещаются знаки-стимулы – отобранные эмодзи. В поле (3) вы-

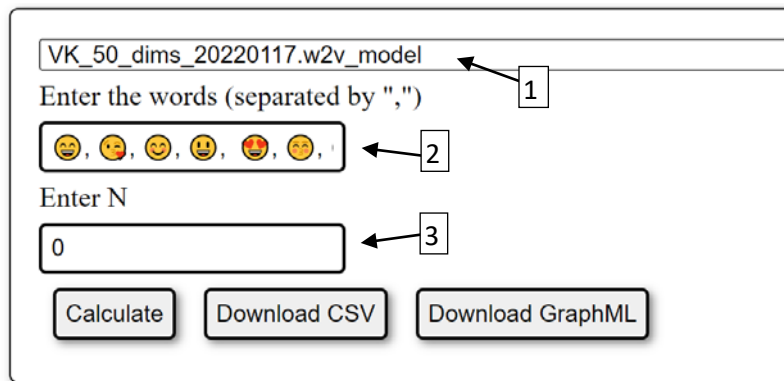
Таблица 1

Эмодзи	Значение эмодзи
	Радость: знаки передают положительные эмоции и хорошее настроение: радость, веселье, удовольствие, возбуждение, искренний смех и т.п.
	Любовь: знаки выражают чувство любви, привязанности, обожания, восторга и т.п.
	Печаль: знаки передают различные оттенки грусти: разочарование, обиду, одиночество, печаль, боль, грусть, досаду, разочарование и т.п.
	Подмигивающее лицо (шутка, флирт, передача скрытого значения)
	Сарказм: ухмылка, лукавое, самодовольное или озорное лицо; знак также может использоваться для передачи флирта или сексуальных намеков
	Одобрение: жесты «большой палец вверх» и «ОК» (супер, отлично) указывают на одобрение
	Удивление: знаки передают удивление, потрясение или шок; могут также передавать страх или сомнение
	Благодарность: сложенные руки (жест означает «спасибо» или «пожалуйста»; альтернативное использование знака – молитва)
	Безразличие: равнодушное лицо (знаки передают нейтральное настроение, но используются и для передачи легкого раздражения или беспокойства)
	Недоумение: знак выражает замешательство или нерешительность, обычно используется для обозначения легкой грусти или разочарования
	Страх: знаки означают ужас, испуг, а также шок, страх, недоверие и сильное волнение
	Вина: гримаса (знак передает ряд негативных эмоций, например, вину, нервозность, смущение или неловкость)
	Смущение: застенчивое лицо (обычно изображает смущение; альтернативные значения – лесть, удивление, недоверие, восхищение, привязанность и волнение)
	Злость: «Пар из носа» (знак передает раздражение, гнев и презрение). Ярость (знак передает сильную степень гнева). Злость (знак передает разную степень гнева – от раздражения до отвращения и возмущения).
	Задумчивость: используется для задавания вопроса или пренебрежительного отношения к кому-то или чему-то.

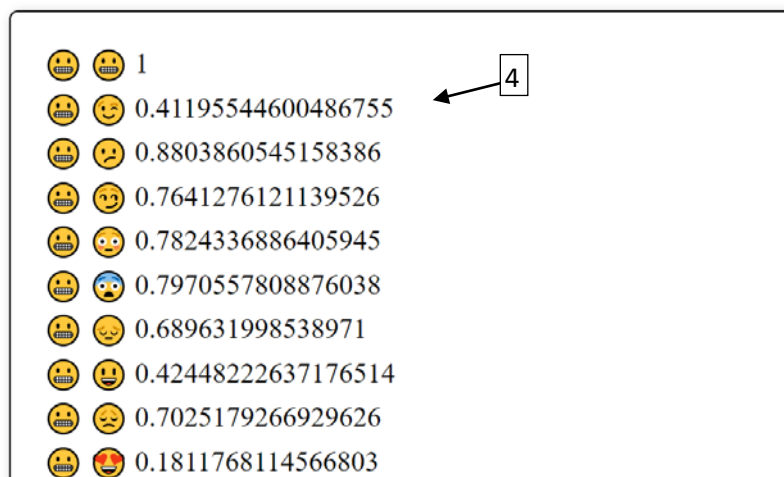
бирается количество знаков между стимулами; в случае  $N = 0$  в модели остаются только знаки-стимулы (дополнительные знаки не включаются). В поле (4) выводится список расстояний между всеми знаками-стимулами (46 эмодзи дают 2116 комбинаций; на рисунке видно, что одинаковые эмодзи имеют косинусное расстояние, равное 1).

На следующем этапе осуществлялась визуализация результатов запроса к W2V-модели корпуса ВКонтакте в виде графовой модели (см. рис. 2), вершинами которой являются эмодзи – знаки разной тональности, а ребрами – показатели контекстуальной близости между ними. Графически частотность употребления каждой пиктограммы передается размером. С помощью

метода модулярности [Lambiotte, Delvenne, Barahona 2009] осуществлялось разбиение графа на классы модулярности – подграфы; а затем производилась укладка графа. В результате данных действий вершины, относящиеся к одному кластеру, группировались в пространстве и отделялись от вершин, образующих другой кластер. При этом каждый класс визуально представлен отдельным цветом. Заметим, что выделенные классы модулярности (кластеры) как фрагменты гиперсети обособлены друг от друга условно: любая из вершин данного класса может иметь связи с другими вершинами гиперсети, относящимися к разным классам модулярности.



Result:



😬 😬	1
😬 😊	0.41195544600486755
😬 😞	0.8803860545158386
😬 😏	0.7641276121139526
😬 😬	0.7824336886405945
😬 😡	0.7970557808876038
😬 😞	0.689631998538971
😬 😊	0.42448222637176514
😬 😞	0.7025179266929626
😬 😡	0.1811768114566803

**Рис. 1.** Результат запроса к W2V-модели корпуса ВКонтакте

*Примечание.* Цифры представляют показатели косинусов угла между векторами (контекстной близости всех пар) эмодзи

Наиболее частотными являются знаки, передающие положительные эмоции («фиолетовый» (2) кластер): 😊, 😊, ❤️, 😊 и 🙌. Особенностью данных эмодзи является их контекстная обособленность друг от друга; обособленны друг от друга и знаки любовной тематики («розовый» (1) кластер). На рисунке 2 видно, что данные кластеры имеют относительно небольшую связность вершин. Иными словами, положительные эмодзи и эмодзи с семантикой «любовь» более четко определены в коммуникативно-семантическом аспекте, некоторые из них, например, выражение благодарности 🙌, контекстуально отделены от всех остальных и ближайшая связь с эмодзи 😊 и ❤️ имеет пороговое значение вхождения 0,5. Контекстуальная определенность некоторых положительных эмодзи проявляется более четко, нежели у их словесных аналогов. Например, слова с семантикой благодарности часто использу-

ются для выражения сарказма («Ну, спасибо тебе! Удружили!»); в то же время использование эмодзи благодарности сложно представить в данной функции.

Интересно, что положительные эмодзи («фиолетовый» кластер) на графе не имеют непосредственной связи с отрицательными («зеленый» (4) кластер), передающими спектр состояний: печаль, злость, раздражение, разочарование, безразличие и вина. «Желтый» (5) кластер, состоящий из знаков с семантикой «удивление», «страх» и «смущение», тоже не имеет непосредственной связи с «фиолетовым» кластером. «Буферной» зоной служит «голубой» (3) кластер, представляющий экспрессивно выраженные положительные (смех) или амбивалентные состояния (обдумывание 🤔). В «буферный» кластер также попали эмодзи, выражающие ироническую или саркастическую ухмылку (😏 и 😊); в зависимости от контекста употребления эти знаки можно трактовать

по-разному. На графе видно, что «голубой» кластер имеет довольно большую связность вершин по сравнению с кластерами 1–2, 5.

Близость негативных эмодзи друг другу высока – об этом свидетельствует наличие многочисленных недифференцированных связей между отдельными эмодзи в «негативном» («зеленом») кластере. Таким образом, **схожий или один и тот же контекст (поведенческая речевая реакция коммуникантов) может передавать эмоцию злости (😡), раздражения (😠), разочарования (😞), а также другие оттенки грусти** (см. связи на графе) в зависимости от используемых знаков. По нашему мнению, выбор конкретного эмодзи может быть обусловлен психологическими параметрами языковых личностей. При этом позитивные эмодзи лучше структурированы, что означает

меньшую вариативность выбора знака для описания похожих или даже тождественных ситуаций разными типами языковых личностей. В то же время реакция на одну и ту же негативную ситуацию может включать целый спектр эмоциональных состояний: злость, грусть, разочарование и безразличие.

Отметим, что «негативный» («зеленый») кластер является многочисленным по количеству знаков. Использование в текстах большого арсенала негативных эмодзи в качестве синонимов можно объяснить тем, что на практике спектр отрицательных эмоций шире, чем положительных: «проявление отрицательных эмоций мы встречаем чаще, чем положительных, т.к. приятное человек считает нормальным и естественным для себя, а отрицательные воздействия стремится преодолеть».

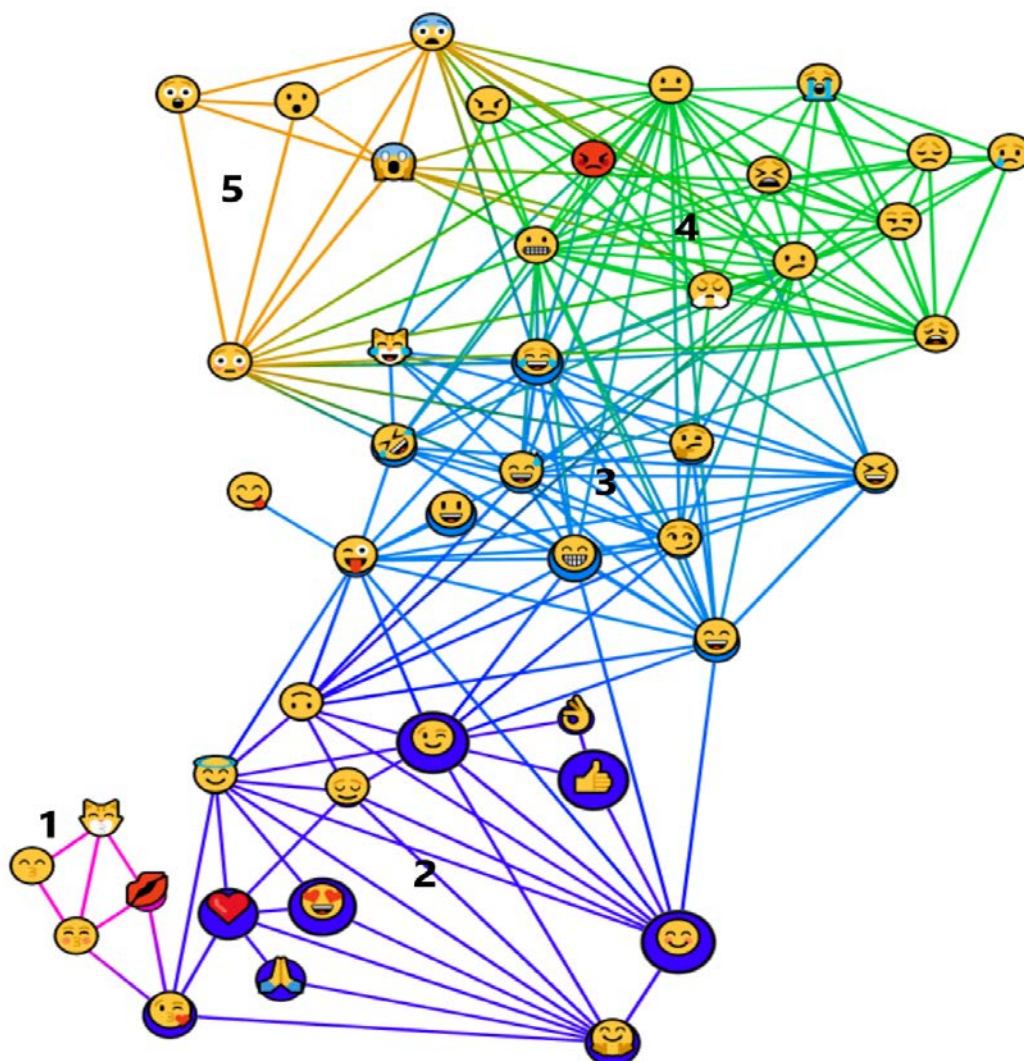


Рис. 2. Графовая модель эмодзи

леть и показать, насколько они неприятны» [Величкова, Киршинова 2015: 166]. Отрицательные переживания обычно ярче и продолжительнее положительных [Винарская 2003], вероятно, поэтому контекстуальная близость «отрицательных» эмодзи при описании схожего контекста равна вариативности знаков «положительной» тональности (ср. количество эмодзи в «зеленом» и «фиолетовом» кластерах).

Психологические характеристики личности «определяют поведение человека и проявляются в мотивах, действиях и реакциях человека на различные жизненные ситуации» [Коженина 2016: 3]. Установлено, что психологические особенности поведения в условиях конфликтной ситуации значительно отличаются у экстравертов, интровертов и амбивертов (смешанный психотип): экстраверты прибегают к стратегии соперничества; для интровертов характерны пассивные стратегии поведения (избегание, приспособление и компромисс); амбиверты ориентированы на сотрудничество [Дубровина, Климонтова, Чепурко 2016; Ионова, Пятаева 2020]. В стрессовой ситуации экстраверты склонны к истерико-психопатическим реакциям, а интроверты – к депрессии, срыву или ухудшению результатов деятельности [Апуневич, Смирнова 2016]. Влияние особенностей темперамента и личных качеств на эмоциональное поведение личности в разнообразных ситуациях выявлено в работах [Маренчук и др. 2016; Куудар и др. 2017; Зекерьяев, Летючая 2021].

Речевое поведение личности может зависеть от биологических, социальных, культурных и психологических факторов, например, пола, возраста, социального статуса, воспитания, национально-культурной идентичности, ситуации и цели общения, стиля взаимодействия с собеседником, а также психологических характеристик. «Любая речевая деятельность эвристична, т.е. каждый раз осуществляется выбор одного из возможных путей реализации высказывания, и этот выбор осуществляется в зависимости от личностных особенностей говорящего» [Доброва 2008:5]. О существовании определенных корреляций между психологическим типом личности и предпочитаемым речевым поведением см. [Горбунова 2017, 2021; Доброва 2013; Попова 2018; Анкудинова, Непочатых 2018].

Таким образом мы полагаем, что выбор конкретного (в первую очередь отрицательного) эмодзи при написании комментариев может быть обусловлен психологическими параметрами ин-

дивидов. Заявленная проблема является масштабной и не может быть решена в одном исследовании. В данной работе будет представлен один из возможных подходов к выявлению особенностей использования эмодзи в письменной речи людьми с разными психологическими чертами.

### 3. Методика сбора и обработки экспериментального материала

В экспериментальной части исследования используется размеченный веб-корпус комментариев пользователей соцсети ВКонтакте «Речевые и неречевые параметры пользователей социальной сети» (часть корпуса зарегистрирована в Федеральной службе по интеллектуальной собственности в качестве базы данных) [Речевые и неречевые... 2018]), прошедших психологический опрос The Big Five Inventory (BFI) в адаптации С.А. Щебетенко<sup>2</sup> [Shchebetenko 2017]). Исходный материал представлен комментариями, которые были размечены с помощью многоуровневого классификатора, включающего в себя семантику (модальность и дейксис), стилистику (жанры и лексические слои), графику (эмодзи/эмотиконы) и др. При этом комментарии были соотнесены с пользователями, каждый из которых имел показатели по всем шкалам BFI (все данные об информантах были обезличены). Объем корпуса составляет 19179 комментариев 93 мужчин и 206 женщин. Лингвистическая разметка корпуса и генерация таблиц осуществлялись в ИС «Семограф» (<https://semograph.org>). Подробное описание исследования представлено [Белоусов и др. 2020].

В данной статье рассмотрен только один раздел классификатора – *графические параметры* реплик (использование эмотиконов и эмодзи). Классификатор этих знаков представлен набором категорий (или в терминах ИС «Семограф» – *полей*). Каждая категория эмотиконов/эмодзи образовывалась множеством комментариев, объединенных общим семантическим признаком.

На рис. 3 видно, столбец «Поля» представляет собой классификатор со счетчиками статистики; в столбце «Термы» размещаются реплики пользователей, а в столбце «Контексты» каждая реплика связана со своим контекстом. На заключительном этапе с помощью инструментария ИС «Семограф» проводилось автоматическое вычис-

<sup>2</sup> Авторы статьи благодарят д. психол. н. С.А. Щебетенко за предоставленные данные психологического опроса информантов.

ление объема выделенных категорий эмотиконов/эмодзи.

#### 4. Психологические параметры информантов

В качестве психологической дифференциации информантов используются бинарные параметры теста BFI: 1) экстраверсия – интроверсия, 2) доброжелательность – враждебность, 3) добросовестность – недобросовестность, 4) эмоциональная стабильность – нейротизм и 5) открытость – консерватизм. Данные параметры представлены во многих психологических тестах и типологиях личности, однако использование опросника BFI объясняется тем, что эта модель широко применяется для изучения психологических черт личности в интернет-коммуникации и социальных медиа. Опросник BFI имеет признание со стороны научного сообщества; эта модель воспроизводится и адаптируется для разных лингвокультур (см. [Angleitner et al. 1990; Saucier et al. 2005] и др.).

Психологические показатели вычислялись для каждого испытуемого. Например, отнесение к полюсам «экстраверсии-интроверсии» осуществлялось на основе показателей среднего значения

(M) и стандартного отклонения (SD), отбрасываемого влево и вправо от среднего. Если у испытуемого показатели по шкале «экстраверсия-интроверсия» превышали  $M+SD$ , полагалось, что у него выражены черты экстраверсии; в случае нахождения значений на шкале влево от  $M-SD$  – черты интроверсии. Если значения располагались в интервале между  $[M-SD, M+SD]$ , полагалось невыраженность признака (смешанный тип личности). Таким образом, «экстраверсия-интроверсия» описывалась с помощью трех психологических характеристик: интроверсия ( $M-SD$ ), невыраженность признака (или «0») и экстраверсия ( $M+SD$ ). Аналогичным образом вычислялись показатели по остальным психологическим параметрам. Значения M и SD для каждой шкалы см. в [Белоусов и др. 2020: 14].

По результатам опроса информанты были распределены на десять групп (Табл. 2), при этом влияние смешанного типа личности, не противопоставленного полюсам психологических шкал, на использование эмотиконов/эмодзи при написании комментариев в дальнейшем анализе не учитывалось.

Охарактеризуем психологические характеристики теста BFI. Интроверсия и экстраверсия

Название	T	F	FR
ЭМОДЗИ СЕМАНТИКА	12693	274	13799
любовы/флирт	2008	113	2083
сарказм/ирония	97	41	102
печаль	277	66	278
одобрение	54	21	55
удивление	79	23	80
недоумение	10	7	10
смущение	2	2	2
безразличие	14	9	15
страх	6	5	6
злость	1	1	1
подмигивание	134	43	143
вина	5	3	5

Название	Тип	K	P
Аня, ну я не так крут	COM	1	5
Аня, ну я не так крут [2]	COM	1	5
Ахаха...у меня ещё такая фотка с тараханом есть	COM	1	6
Божечки	COM	1	6
Бохиня дискотеки	COM	1	5
Вячеслав, 😊😊😊 похож	COM	1	5
Да, первого января погода нас поредала! ЦЕЛЫЙ день ясно было, СОЛНЦЕ! Плюс 3 примерно на улице 😊 сейчас уже опять снег выпал, немного пасмурно.	COM	1	10
Денис, 😊 (подколоп)	COM	1	7
Дерьмище каков то дегенератское	COM	1	6

Терм	Контекст	Предпросмо	date
Александр, проверить оригинальность	740	17.03.2017 13:59	
Александр, я когда первый раз читал, тоже не дочитал	740	17.03.2017 14:10	
Алексей, слишком прилично выглядит,лоз тому не узнаешь	60	24.12.2016 18:12	
Аня, администратор	216	25.12.2016 18:32	

Рис. 3. Фрагмент классификатора (раздел ЭМОДЗИ)

Таблица 2

## Выборка информантов

Психологическая характеристика	Информанты	
	абс.	отн., %
Экстраверсия	45	15,0%
Интроверсия	55	18,3%
Доброжелательность	53	17,7%
Враждебность	32	10,7%
Добросовестность	48	16,0%
Недобросовестность	44	14,7%
Эмоциональная стабильность	52	17,3%
Нейротизм	48	16,0%
Открытость	45	15,0%
Консерватизм	55	18,3%

рассматриваются с точки зрения жизненных установок: экстраверсия предполагает направленность на взаимодействие с другими людьми и ориентацию личности на внешний мир, а интроверсия – ориентацию на внутренний мир и его самосохранение [Юнг 2001]. Если экстраверты характеризуются как общительные, дружелюбные и энергичные личности, то интроверты – как замкнутые, неуверенные и инертные [Thompson 2008].

Параметр «доброжелательность» характеризует отзывчивых, внимательных и оптимистичных людей, стремящихся избегать разногласий. Для таких людей характерны стремление к сотрудничеству и альтруизму, а также готовность в ущерб своим интересам идти на компромисс. Личности, склонные к доброжелательности, умеют сопереживать, ответственно подходят к выполнению обязательств. Личности с низким уровнем доброжелательности эффективно отстаивают свои интересы; они эгоцентричны и равнодушны к чужим потребностям. Склонные к враждебности люди стараются дистанцироваться от других и отказываются от сотрудничества [Орел, Сенин 2008].

Личности с высокими показателями по шкале «добросовестность» представляют собой высокоорганизованных людей, стремящихся хорошо выполнить свою работу. Эта черта характерна для людей, серьезно относящихся к обязательствам перед другими. Такие люди трудолюбивы, систематичны, аккуратны; демонстрируют спланированное поведение [Thompson 2008]. Личности с низкими показателями по шкале «добросовестность» обычно спокойные; они менее ориентированы на успех и более склонны к антиобщественному поведению [Ozer 2006].

Личностям, склонным к нейротизму, свойственно испытывать негативные аффективные состояния (грусть, страх, тревогу, гнев, раздражение,

чувство вины и т.п.). Для них характерны заниженная самооценка, пониженная стрессоустойчивость, совершение иррациональных поступков, неспособность эффективно контролировать эмоции. Эмоционально стабильные личности, наоборот, отличаются высокой стрессоустойчивостью и хорошей адаптивностью; обычно они хорошо контролируют эмоции [Орел, Сенин 2008].

Личности с характеристикой «открытость» склонны поддерживать новые идеи в разных сферах жизни и подвергать сомнению авторитеты. Такие люди легко переключаются с одного вида деятельности на другой; их эмоциональные реакции обычно интенсивны [там же 2008]. Личности с низким уровнем открытости обычно консервативны в своих взглядах, традиционны в своем мировоззрении и поведении. Их можно охарактеризовать как практичными людьми с узкой сферой интересов; они комфортно чувствуют себя в знакомой обстановке, а их эмоциональные реакции обычно приглушены [Harris 2004; DeYoung 2015].

### 5. Классификация эмотиконов и эмодзи: первичный количественный анализ

Основу семантической классификации изучаемых нами графических средств составляют основные эмоции человека и их внешние проявления (мимика, позы, жестикация). Классификация разрабатывалась с учетом особенностей собранного речевого материала, основных видов эмотиконов и эмодзи, а также существующих классификаций эмоций (см. [Изард 1999; Симонов 1983; Ekman 1973; Woodworth, Schlosberg 1955] и др.). Для выявления значений и рассмотрения примеров употребления графических знаков привлекались интернет-словари и электронные библиотеки эмотиконов и эмодзи: словарь



английского сленга Urban Dictionary (URL: <https://www.urbandictionary.com>), словарь акронимов Netlingo (URL: <https://www.netlingo.com>), энциклопедия Emojipedia (URL: <https://emojipedia.org>) и др.

Созданный классификатор включает в себя 15 категорий эмодзи и эмодзи; подробнее о классификаторе в [Белоусов, Обухова 2019]. Классификацию составляют следующие разнородные элементы: чувства, базовые и вторичные эмоции, физические состояния и конкретные действия. Разнородность классификации объясняется особенностью собранного материала – «активным использованием в электронной коммуникации этих знаков для выражения своего эмоционального состояния» [там же: 9].

Остановимся на количественных данных, характеризующих использование графических средств. Всего в классификатор было отнесено 12693 комментария 274 пользователей, что составило 66% всех реплик в корпусе и 92% всех пользователей, принявших участие в исследовании. Данные первичного количественного анализа представлены в Табл. 3.

Прокомментируем первую строку таблицы. 265 пользователей использовали в своих текстах эмодзи/эмодзи категории «радость». В 10876 репликах этих пользователей встречается хотя бы один знак этой категории, что составляет 85,7% комментариев от числа всех комментариев в корпусе. В среднем на одного информанта приходится 41 реплика. Остальные строки интерпретируются таким же образом.

Как видим, графические знаки с семантикой

радости являются самыми частотными в корпусе. Все остальные категории относительно категории «радость» в целом являются низкочастотными. Исключением можно назвать только категорию «любовь», на которую приходится почти шестая часть всех реплик в корпусе (16%).

Далее рассмотрим совместное употребление (совместную встречаемость) представленных категорий в комментариях пользователей в зависимости от их психологических параметров.

## 6. Результаты исследования

Один и тот же комментарий пользователя мог относиться к нескольким категориям классификатора при условии, если в одной реплике встречались эмодзи/эмодзи, выражающие разные эмоции/состояния. Так, комментарий «Спасибо, Олеся! И отдельная благодарность за помощь в учебе! Что бы я без тебя делала! 😊🙏» относится сразу к двум категориям: эмодзи 🙏 относится к категории «благодарность», а эмодзи 😊 – к категории «радость». Связи между категориями устанавливаются на основе совместного присутствия в репликах одного пользователя двух и более знаков, относящихся к разным категориям графических средств. Выявленные связи между категориями во всем корпусе данных можно представить в виде семантической карты (матрицы размера  $N \times N$ ), в столбцах и строках которой располагаются названия категорий эмодзи и эмодзи. Такая карта строится автоматически на стадии построения классификации в ИС «Семограф». Частота совместной встречаемости двух категорий в репликах корпуса (или его отдельной

Таблица 3

Категории эмодзи и эмодзи: данные первичного количественного анализа

Категории эмодзи и эмодзи	Информанты, абс.	Комментарии информантов		
		абс.	отн., %	встречаемость
Радость	265	10876	85,700	41,00
Любовь	113	2008	16,000	17,80
Печаль	66	277	2,200	4,20
Подмигивание	43	134	1,050	3,11
Сарказм/ирония	41	97	0,760	2,36
Одобрение	21	54	0,420	2,57
Удивление	23	79	0,620	3,43
Благодарность	5	22	0,170	4,40
Безразличие	9	14	0,110	1,55
Недоумение	7	10	0,078	1,42
Страх	5	6	0,040	1,20
Вина	3	5	0,030	1,66
Смущение	2	2	0,015	1,00
Злость	1	1	0,007	1,00
Задумчивость	1	1	0,007	1,00

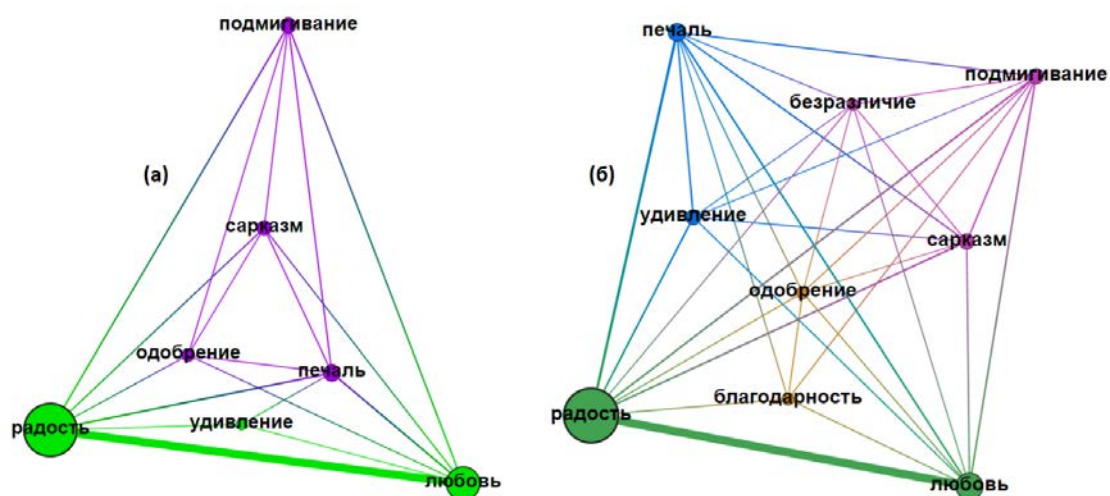


Рис. 4. Графовая модель категорий эмотиконов и эмодзи:  
(а) параметр «Экстраверсия», (б) параметр «Интроверсия»

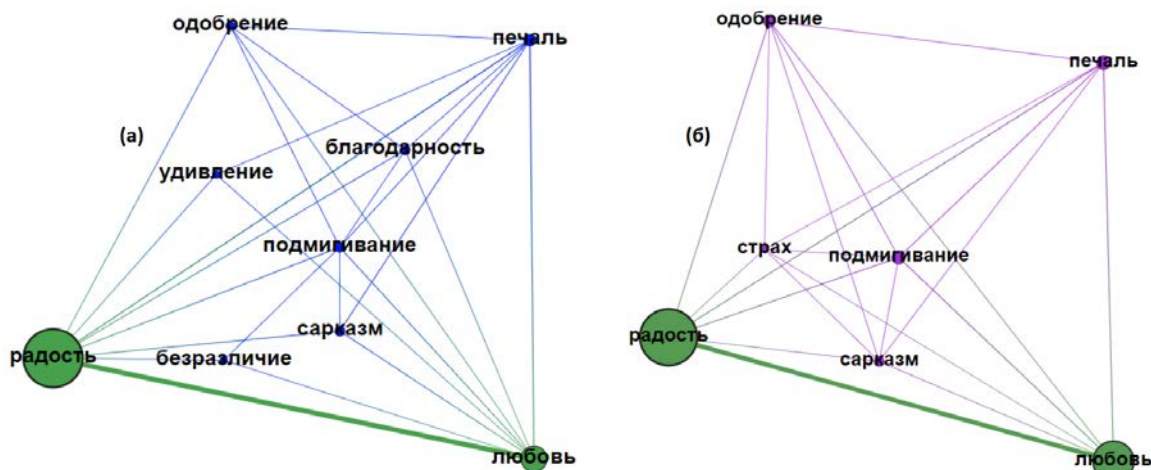


Рис. 5. Графовая модель категорий эмотиконов и эмодзи:  
(а) параметр «Доброжелательность», (б) параметр «Враждебность»

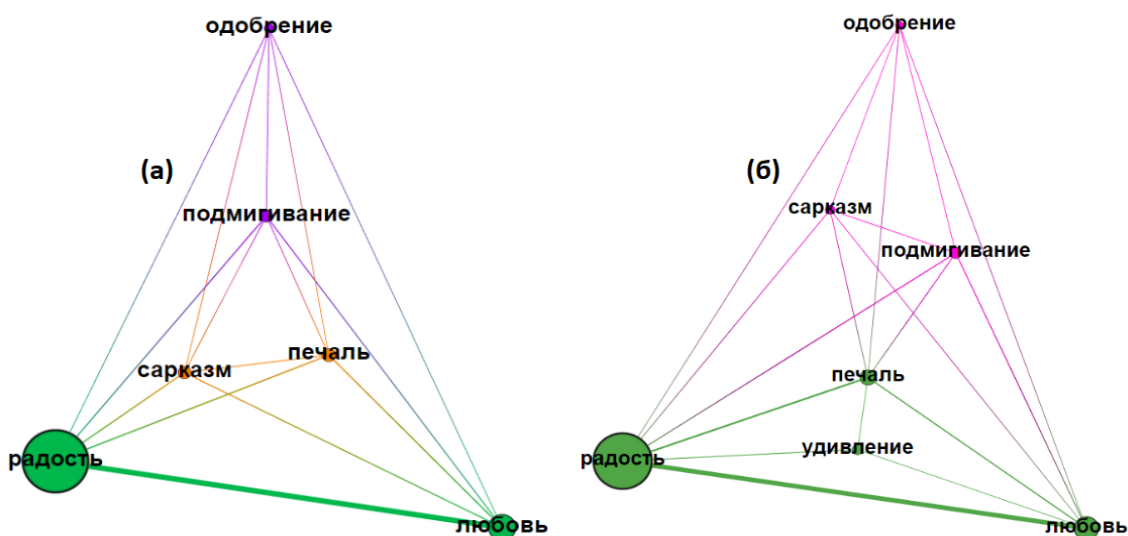


Рис. 6. Графовая модель категорий эмотиконов и эмодзи:  
(а) параметр «Добросовестность», (б) параметр «Недобросовестность»

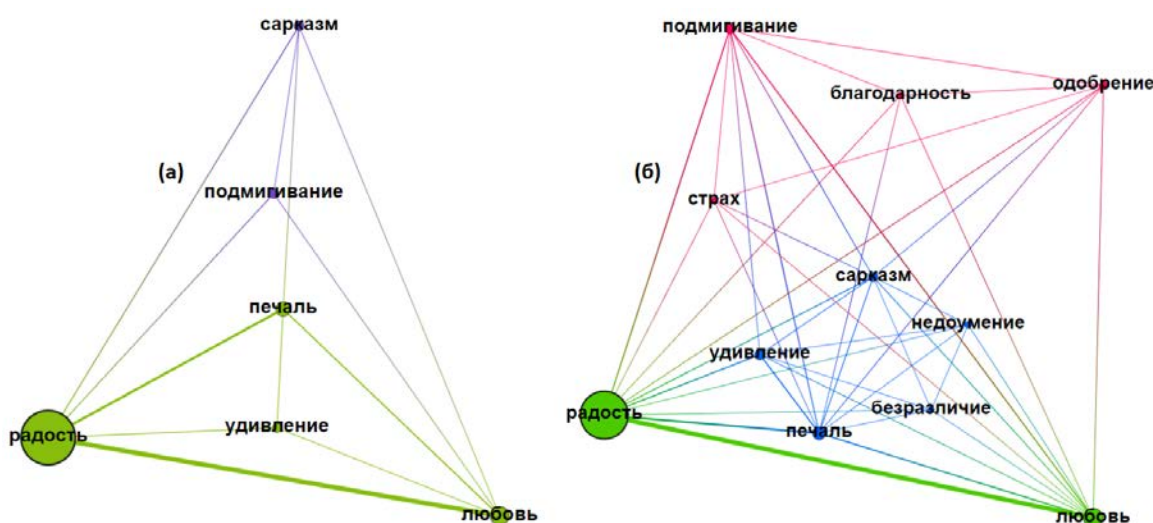


Рис. 7. Графовая модель категорий эмотиконов и эмодзи:  
(а) параметр «Эмоциональная стабильность», (б) параметр «Нейротизм»

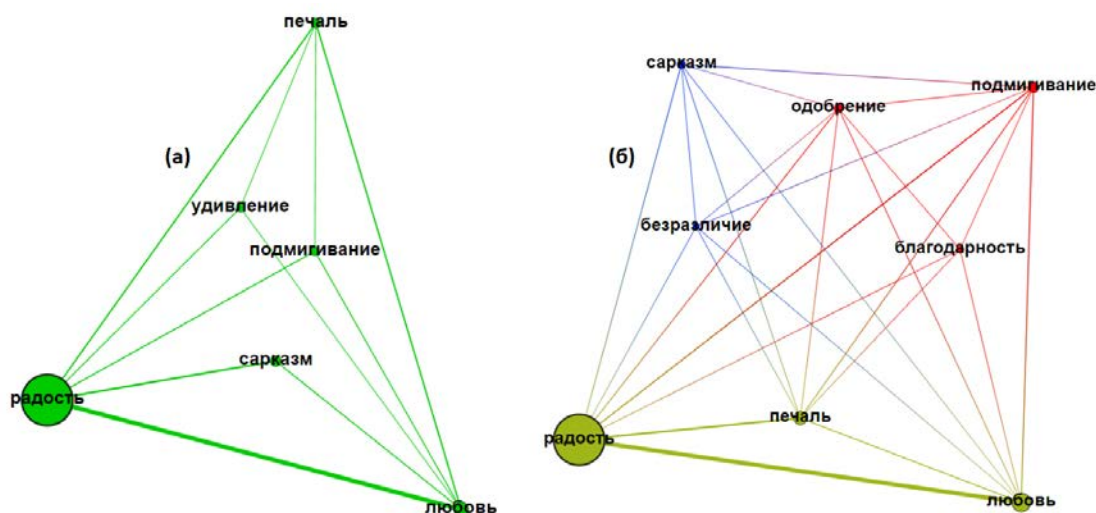


Рис. 8. Графовая модель категорий эмотиконов и эмодзи:  
(а) параметр «Открытость», (б) параметр «Консерватизм»

пересечении столбца и строки. Граф, построенный на основе семантической карты, представляет собой визуализацию семантической карты. Вершинами такой модели являются категории эмотиконов и эмодзи, а ребрами – связи между ними. Графически объемы категорий и частоты связей передаются размером (для вершины) и толщиной (для ребер), а принадлежность категории к тому или иному кластеру передается цветом. Образование кластера свидетельствует о более частотном факте совместной встречаемости разных категорий в репликах тех или иных информантов. Графы строились с помощью визуализатора SciVi.

На рис. 4–8 представлены графовые модели, отражающие совместное употребление категорий

эмотиконов и эмодзи в комментариях информантов с контрастными психологическими характеристиками. Как видим, в структуре всех графов четко выделяются ядерная, предъядерная и периферийная зоны, между которыми наблюдаются резкие границы. Ядро образует категория «радость», предъядерную зону – категория «любовь», а периферию – оставшиеся выделенные категории.

Простую структуру имеют графы, построенные на материале реплик пользователей с характеристиками «экстраверсия», «враждебность», «добросовестность», «эмоциональная стабильность» и «открытость» (см. рис. 4 (а), рис. 5 (б), рис. 6 (а), рис. 7 (а) и рис. 8 (а)). Указанные графы

состоят из 6–7 вершин, а количество ребер варьируется от 11 и более. Эти графы образуются с помощью 2–3 кластеров, однако граф, построенный на материале реплик информантов с характеристикой «открытость», состоит только из одного «зеленого» кластера (рис. 8 (а)). Более сложную структуру имеют графы, построенные на материале комментариев информантов с чертами «интроверсия», «доброжелательность», «недобросовестность», «нейротизм» и «консерватизм» (см. рис. 4 (б), рис. 5 (а), рис. 6 (б), рис. 7 (б) и рис. 8 (б)). Эти графы состоят уже из 7–11 вершин, а количество ребер варьируется от 24 и более. Указанные графы образуются с помощью 2–4 кластеров.

Прокомментируем наиболее интересные факты совместного употребления категорий эмодзи и эмотиконов в текстах информантов с контрастными психологическими характеристиками. Отметим, что для всех информантов характерна связь между категориями «радость» и «любовь» (положительные знаки), при этом данные категории относятся к одному кластеру, а совместная встречаемость знаков указанной семантики довольно частотна (ср. цвета кластеров и размер ребер для этих категорий на рис. 4–8). Как видим, все остальные связи между категориями являются низкочастотными, о чем свидетельствуют наличие тонких ребер на рис. 4–8. Таким образом, связь между категориями «радость» и «любовь» можно считать устойчивой для всех информантов.

Наиболее сложными для интерпретации оказываются графы с 3–4 кластерами (см., например, характеристики «нейротизм» и «консерватизм»). Например, у информантов, склонных к нейротизму, один и тот же контекст (поведенческая речевая реакция) соединяет и положительные, и отрицательные эмоции: страх, подмигивание, благодарность и одобрение («розовый» кластер на рис. 7(б)). Самый многочисленный кластер у людей с высокими показателями нейротизма образуется с помощью пяти категорий разной тональности: «печаль», «безразличие», «удивление», «недоумение» и «сарказм», при этом указанные связи между категориями низкочастотны (см. «синий» кластер на рис. 7(б)).

У информантов с характеристикой «консерватизм» в одном кластере также оказываются знаки разной тональности, например, категории «печаль», «радость» и «любовь» (см. «зеленый» кластер на рис. 8(б)).

Большее количество вершин, ребер и кластеров свидетельствует о том, что информанты

более склонны комбинировать в одном и том же или схожем контексте знаки разной тональности.

## 7. Заключение

Модель, отражающая контекстуальную близость эмодзи, обнаруженную на материале 23 млн. текстов более 3 млн. пользователей, показывает, что позитивные эмодзи (знаки с семантикой «радость», «любовь», «одобрение», «благодарность») обособлены друг от друга; они более четко определены в коммуникативно-семантическом аспекте по сравнению с отрицательными эмодзи или эмодзи, выражающими экспрессивные положительные или амбивалентные эмоциональные состояния. Позитивные эмодзи лучше структурированы, что означает меньшую вариативность выбора знака для описания похожих или тождественных ситуаций разными типами языковых личностей. Положительные эмодзи напрямую не связаны с отрицательными; связь между этими знаками образуется через «буферную» зону – знаки, выражающие экспрессивно выраженные положительные и амбивалентные состояния. Контекстуальная близость разных типов отрицательных эмодзи высока, т.е. схожий или один и тот же контекст может передавать сразу несколько типов отрицательных тональностей. Иными словами, поведенческая речевая реакция на тождественные негативные ситуации включает широкий спектр эмоциональных состояний (например, печаль, злость, страх, безразличие и т.п.), т.е. предполагает большую вариативность выбора знака. Мы предположили, что такая структура может быть обусловлена разной реакцией пишущих на схожие ситуации (определены сходными наборами слов, окружающими исследуемые эмодзи), происходящей из-за различий в психологических характеристиках самих пишущих. Это обусловило обращение к использованию эмодзи в комментариях пользователей в сетевом общении. В первую очередь, нас интересовали случаи использования нескольких графических знаков одновременно (т.е. в одном контексте), что моделировало ситуацию первого исследования, проведенного на больших данных.

Графовые модели, построенные на материале размеченного веб-корпуса текстов, выполняют поставленную задачу – они отражают совместное употребление категорий эмодзи и эмотиконов в комментариях пользователей. Эти модели показывают, что употребление графических средств зависит от психологических параметров пользователей. Зависимость выражается: а) в

усложнении либо упрощении формальных параметров структуры совместного использования эмотиков и эмодзи в речи (сокращения/увеличения количества категорий, а также в количестве кластеров категорий, показывающих реальное совместное использование классов графических средств в текстах), и б) в видах композиции эмотиконов и эмодзи, представляющей разные виды отношения пишущего к возникающим ситуациям или темам диалогов. Так, например, информанты с психологической характеристикой «открытость» легко комбинируют знаки разных тональностей в одном сообщении (все категории относятся к одному кластеру); напротив, у информантов с характеристикой «консерватизм» использование эмотиконов и эмодзи укладывается в три крупных фрейма (радость – любовь – печаль; безразличие – сарказм; одобрение – подмигивание – благодарность).

Отметим, что среди знаков положительной тональности для всех информантов характерна устойчивая частотная связь между категориями «радость» и «любовь», а использование категорий «благодарность» и «одобрение» оказывается вариативным (эти категории попадают в другой кластер). У информантов с разными психологическими характеристиками самыми вариативными комбинациями оказываются знаки отрицательной тональности (категории «печаль», «безразличие» и «страх»), а также знаки, отражающие амбивалентные состояния (категории «подмигивание», «сарказм» и «удивление»). Таким образом, можно говорить о большей вариативности структурных связей отрицательных типов эмотиконов и эмодзи, вызванных психологическими характеристиками коммуникантов.

### Список литературы / References

*Анкудинова С.А., Непочатых И.А.* Влияние психотипа личности на выбор стратегии коммуникации студентами-будущими учителями // Евразийский Союз Ученых. 2018. № 3-5 (48). С. 5-8. [Ankudinova S.A., Nepochatyh I.A. Vlijanie psihotipa lichnosti na vybor strategii kommunikacii studentami-budushhimi uchiteljami // Evrazijskij Sojuz Uchenyh. 2018. № 3-5 (48). S. 5-8.]

*Апуневич О.А., Смирнова Е.С.* Особенности эмоционального выгорания у работников УВД с разным уровнем экстра-интроверсии // Психология и педагогика: достижения и тенденции развития. 2016. С. 16-24. [Apunevich O.A., Smirnova E.S. Osobennosti jemocional'nogo vygoraniya u rabotnikov UVD s raznym urovnem jekstra-

introversii // Psihologija i pedagogika: dostizhenija i tendencii razvitija. 2016. S. 16-24.]

*Белоусов К.И. и др.* Многопараметрический анализ лингвистических данных в информационной системе «Семограф» (на примере исследования речевого поведения пользователей социальной сети) / Белоусов К.И., Ерофеева Е.В., Баранов Д.А., Зелянская Н.Л., Щebetenko С.А. // Вестник Томского государственного университета. Филология. 2020. № 64. С. 5-29. [Belousov K.I. i dr. Mnogoparametricheskij analiz lingvisticheskikh dannyh v informacionnoj sisteme «Semograf» (na primere issledovanija rechevogo povedenija pol'zovatelej social'noj seti) / Belousov K.I., Erofeeva E.V., Baranov D.A., Zeljanskaja N.L., Shhebetenko S.A. // Vestnik Tomskogo gosudarstvennogo universiteta. Filologija. 2020. № 64. S. 5-29.]

*Белоусов К.И., Обухова И.А.* Влияние пола и самооценки пользователей социальной сети на использование эмотиконов и эмодзи в процессе речевой коммуникации // Вестник Пермского университета. Российская и зарубежная филология. 2019. № 11 (3). С. 5-18. [Belousov K.I., Obuhova I.A. Vlijanie pola i samoocenki pol'zovatelej social'noj seti na ispol'zovanie jemotikonov i jemodzi v processe rechevoj kommunikacii // Vestnik Permskogo universiteta. Rossijskaja i zarubezhnaja filologija. 2019. № 11 (3). S. 5-18.]

*Браудо Т.Е., Фрумкина Р.М.* О знаковых системах, замещающих естественный язык // Культурно-историческая психология. 2006. № 2 (3). С. 28-37. [Braudo T.E., Frumkina R.M. O znakovyh sistemah, zameshhajushhih estestvennyj jazyk // Kul'turno-istoricheskaja psihologija. 2006. № 2 (3). S. 28-37.]

*Величкова Л.В., Киришинова О.В.* Психолингвистический подход к исследованию эмоциональности речи // Вестник Ленинградского государственного университета им. А.С. Пушкина. 2015. Т. 1. № 3. С. 165-171. [Velichkova L.V., Kirshinova O.V. Psiholingvisticheskij podhod k issledovaniju jemocional'nosti rechi // Vestnik Leningradskogo gosudarstvennogo universiteta im. A.S. Pushkina. 2015. T. 1. № 3. S. 165-171.]

*Винарская Е.Н.* Выразительные средства текста (на материале русской поэзии). 2-е изд. Воронеж: ВГУ, 2003. [Vinarskaja E.N. Vyrazitel'nye sredstva teksta (na materiale russkoj poezii). 2-e izd. Voronezh: VGU, 2003.]

*Горбунова Д.А.* О чем говорит психотип: функциональное распределение прагматических

маркеров в речи интровертов и экстравертов // Коммуникативные исследования. 2021. Т. 8. № 1. С. 55-69. [Gorbunova D.A. O chem govorit psihotip: funkcional'noe raspredelenie pragmaticheskikh markerov v rechi introvertov i jekstravertov // Kommunikativnye issledovanija. 2021. T. 8. № 1. S. 55-69.]

*Горбунова Д.А.* Спонтанная устная речь с точки зрения психолингвистики: прагматемы в речи информантов с различным психотипом // Коммуникативные исследования. 2017. № 4 (14). С. 73-82. (На англ. яз.). [Gorbunova D.A. Spontannaja ustnaja rech' s točki zrenija psiholingvistiki: pragmatemy v rechi informantov s razlichnym psihotipom // Kommunikativnye issledovanija. 2017. № 4 (14). S. 73-82. (Na angl. jaz.).]

*Доброва В.В.* Семантические и психологические особенности функционирования сокращенных форм общения в речи // Вестник Самарского государственного технического университета. Серия: Психолого-педагогические науки. 2013. № 1 (19). С. 26-32. [Dobrova V.V. Semanticheskie i psihologicheskie osobennosti funkcionirovanija sokrashhennykh form obshhenija v rechi // Vestnik Samarskogo gosudarstvennogo tehničeskogo universiteta. Serija: Psihologo-pedagogicheskie nauki. 2013. № 1 (19). S. 26-32.]

*Дубровина С.В., Климонтова Т.А., Чепурко Ю.В.* Стратегия поведения в конфликтной ситуации при различной направленности личности (по параметру экстраверсия-интроверсия) // Азимут научных исследований: педагогика и психология. 2016. Т. 5. № 4 (17). С. 335-337. [Dubrovina S.V., Klimontova T.A., Čepurko Ju.V. Strategija povedenija v konfliktnoj situacii pri razlichnoj napravlenosti lichnosti (po parametru jestraversija-introversija) // Azimut nauchnyh issledovanij: pedagogika i psihologija. 2016. T. 5. № 4 (17). S. 335-337.]

*Зекер'яев Р.И., Летючая В.Г.* Психологические особенности поведения в конфликте у людей с разными типами темперамента // Заметки ученого. 2021. № 5-1. С. 322-329. [Zeker'jaev R.I., Letjuchaja V.G. Psihologicheskie osobennosti povedenija v konflikte u ljudej s raznymi tipami temperamenta // Zametki učenogo. 2021. № 5-1. S. 322-329.]

*Изард К.Э.* Психология эмоций. СПб.: Питер, 2012. [Izard K. Je. Psihologija jemocij. SPb.: Peter, 2012.]

*Ионова М.С., Пятаева Е.В.* Стратегии поведения в конфликте студентов с экстравертным и интровертным типом личности // Огарёв-Online.

2020. № 11 (148). С. 3. [Ionova M.S., Pjataeva E.V. Strategii povedenija v konflikte studentov s jekstravertnym i introvertnym tipom lichnosti // Ogarjov-Online. 2020. № 11 (148). S. 3.]

*Китова Е.Б.* Общение в интернет-среде и «универсальный язык» эмодзи // Вопросы теории и практики журналистики. 2016. Т. 5. № 4. С. 654-664. [Kitova E.B. Obshhenie v internet-srede i «universal'nyj jazyk» jemodzi // Voprosy teorii i praktiki zhurnalistiki. 2016. T. 5. № 4. S. 654-664.]

*Кожевина А.П.* Структурная организация доверия при различной направленности личности (по параметру экстраверсия-интроверсия): дис. Ярославль, 2016. [Kozhevina A.P. Strukturnaja organizacija doverija pri razlichnoj napravlenosti lichnosti (po parametru jekstraversija-introversija): dis. Jaroslavl, 2016.]

*Куулар Ш.В. и др.* Психофизиологические особенности студентов с разным типом стратегии поведения в конфликтных ситуациях // Science for Education Today. 2017. Т. 7. № 5. С. 67-80. [Kuular Sh.V. i dr. Psihofiziologicheskie osobennosti studentov s raznym tipom strategii povedenija v konfliktnyh situacijah // Science for Education Today. 2017. T. 7. № 5. S. 67-80.]

*Маренчук Ю.А. и др.* Особенности конфликтного поведения студентов в условиях учебно-профессиональной деятельности // Kant. 2016. № 3 (20). С. 37-42. [Marenchuk Ju.A. i dr. Osobennosti konfliktnogo povedenija studentov v uslovijah uchebno-professional'noj dejatel'nosti // Kant. 2016. № 3 (20). S. 37-42.]

*Орел В.Е., Сенин И.Г.* Личностные опросники NEO PI-R, NEO FFI: руководство по применению / В.Е. Орел, И.Г. Сенин, 2-е изд., Ярославль: НПЦ «Психодиагностика», 2008. [Orel V.E., Senin I.G. Lichnostnye oprosniki NEO PI-R, NEO FFI: rukovodstvo po primeneniju / V.E. Orel, I.G. Senin, 2-e izd., Jaroslavl': NPC «Psihodiagnostika», 2008.]

*Павлов М.А.* Психолингвистический аспект использования невербальных средств коммуникации в сетевом Интернет-дискурсе // Актуальные вопросы филологии и переводоведения в свете современных исследований: сборник научных статей. Чебоксары: Чуваш. гос. пед. ун-т, 2017. С. 133-138. [Pavlov M.A. Psiholingvisticheskij aspekt ispol'zovanija neverbal'nykh sredstv kommunikacii v setevom Internet-diskurse // Aktual'nye voprosy filologii i perevodovedenija v svete sovremennyh issledovanij: sbornik nauchnyh statej. Čeboksary: Čuvash. gos. ped. un-t, 2017. S. 133-138.]

*Попова О.В.* Особенности речевого поведения личности интровертного и экстравертного типов в парламентском дискурсе // Вестник Волгоградского государственного университета. Серия 2: Языкознание. 2018. Т. 17. № 2. С. 134-142. [Popova O.V. Osobennosti rechevogo povedeniya lichnosti introvertnogo i jekstravertnogo tipov v parlamentskom diskurse // Vestnik Volgogradskogo gosudarstvennogo universiteta. Serija 2: Jazykoznanie. 2018. T. 17. № 2. S. 134-142.]

Речевые и неречевые параметры пользователей социальной сети [электронный ресурс]: свидетельство о государственной регистрации базы данных, охраняемой авторскими правами / Баранов Д.А., Белоусов К.И., Боронникова Н.В., Ерофеева Е.В., Зелянская Н.Л., Константинов И.М., Обухова И.А., Руденко Е.С., Русинова И.И., Худякова Е.С. М.: Федер. служба по интеллектуальной собственности. Внесено в реестр баз данных, рег. № 2018621839 от 20.11.2018. [Rechevye i nerechevye parametry pol'zovatelej social'noj seti [jelektronnyj resurs]: svidetel'stvo o gosudarstvennoj registracii bazy dannyh, ohranjaemoj avtorskimi pravami / Baranov D.A., Belousov K.I., Boronnikova N.V., Erofeeva E.V., Zeljanskaja N.L., Konstantinov I.M., Obuhova I.A., Rudenko E.S., Rusinova I.I., Hudjakova E.S. M.: Feder. sluzhba po intelektual'noj sobstvennosti. Vneseno v reestr baz dannyh, reg. № 2018621839 ot 20.11.2018]

*Симонов П.В.* Ответ профессору Б.И. Додонову (Еще раз о потребностно-информационном подходе к изучению эмоций) // Психологический журнал. 1983. № 4 (4). С. 119-133. [Simonov P.V. Otvet professoru B.I. Dodonovu (Eshhe raz o potrebnostno-informacionnom podhode k izucheniju jemocij) // Psihologicheskij zhurnal. 1983. № 4 (4). S. 119-133.]

Энциклопедия эмодзи Emojipedia. URL: <https://emojipedia.org> (дата обращения: 21.12.2021). [Enciklopedija jemodzi Emojipedia. URL: <https://emojipedia.org> (data obrashhenija: 21.12.2021).]

*Юнг К.Г.* Психологические типы. / пер. с нем. С. Лорие (под ред. В. Зеленского). СПб.: Азбука, 2001. [Jung K.G. Psihologicheskie tipy. / per. s nem. S. Lorie (pod red. V. Zelenskogo). SPb.: Azbuka, 2001.]

*Angleitner A., Ostendorf F., John O.P.* Towards a taxonomy of personality descriptors in German: a psycho-lexical study // European Journal of Personality. 1990. № 4 (2). P. 89-118.

*Belousov K. et al.* "Semograph" Information System as a Framework for Network-Based Science

and Education / Belousov K., Erofeeva E., Leshchenko Y., Baranov D. // International Conference on Smart Education and Smart E-Learning. Springer, Cham. 2017. P. 263-272.

*Derks D. et al.* Emoticons in computer-mediated communication: social motives and social context / Derks D., Bos A. E. R., Grumbkow J. von. // Cyberpsychology & behavior. 2008. № 11 (1). P. 99-101.

*DeYoung C.G.* Cybernetic big five theory // Journal of research in personality. 2015. Т. 56. P. 33-58.

*Ekman P.* Cross-cultural studies of facial expression // Darwin and facial expression: A century of research in review / ed. by P. Ekman. 1973. P. 169-222.

Emoticon // CALD – Cambridge Advanced Learner's Dictionary. URL: <https://dictionary.cambridge.org/dictionary/english/emoticon> (дата обращения: 21.12.2021).

*Harris J.A.* Measured intelligence, achievement, openness to experience, and creativity // Personality and individual differences. 2004. Т. 36. № 4. P. 913-929.

*Hu Y. et al.* Emoticon-based ambivalent expression: A hidden indicator for unusual behaviors in weibo / Hu Y., Zhao J., Wu J. // PloS one. 2016. Т. 11. № 1.

*Korobov M.* Morphological Analyzer and Generator for Russian and Ukrainian Languages // International Conference on Analysis of Images, Social Networks and Texts, AIST 2015. P. 320-332.

*Lambiotte R., Delvenne J.-C., Barahona M.* Laplacian Dynamics and Multiscale Modular Structure in Networks. 2009. URL: <https://arxiv.org/abs/0812.1770v3> (дата обращения: 06.10.2022).

*Mikolov T., Chen K., Corrado G., Dean J.* Efficient Estimation of Word Representations in Vector Space CoRR. 2013. URL: <https://arxiv.org/pdf/1301.3781.pdf>.

*Ozer D.J., Benet-Martinez V.* Personality and the prediction of consequential outcomes // Annual review of psychology. 2006. Т. 57. P. 401-421.

NetLingo: The Internet Dictionary. URL: <https://www.netlingo.com> (дата обращения: 19.12.2021).

*Riordan M.A.* Emojis as tools for emotion work: Communicating affect in text messages // Journal of Language and Social Psychology. 2017. Т. 36. № 5. С. 549-567.

*Řehůřek R., Sojka P.* Gensim-Python Framework for Vector Space Modelling //NLP Centre,

Faculty of Informatics, Masaryk University, Brno, Czech Republic. 2011. Т. 3. № 2.

*Saucier G. et.al* The factor structure of Greek personality adjectives / Saucier G., Georgiades S., Tsaousis I., Goldberg L.R. // *Journal of Personality and Social Psychology*. 2005. № 88(5). P. 856-875.

*Shchebetenko S.* Reflexive characteristic adaptations explain sex differences in the Big Five: But not in neuroticism // *Personality and Individual Differences*. 2017. № 111. P. 153-156.

*Thompson E.R.* Development and Validation of an International English Big-Five Mini-Markers // *Personality and Individual Differences*. 2008. № 45 (6). P. 542-548.

Urban Dictionary. URL: <https://www.urbandictionary.com> (дата обращения: 15.12.2021).

*Woodworth R.S., Schlosberg H.* *Experimental psychology*. N.Y.: Holt. 1955. P. 948.

## WORD2VEC- AND SEMANTIC GRAPH MODELING OF EMOTICONS AND EMOJI USE IN INTERNET COMMUNICATION TEXTS

**K.I. Belousov, I.A. Obukhova, I.A. Labutin**

*Perm State University (Perm, Russia)*

*belousovki@gmail.com*

The article deals with emoticons and emoji use in textual computer-mediated communication. The research material is based on comments left by users of the social networking service VK. Using machine learning methods (Word2Vec), we build a model of contextual proximity of emoji use in the comments. This model reveals a closer contextual relationship of negative-tone emoji of different types with each other.

The influence of personality's psychological parameters on the use of icons of different tonality in similar communicative situations is considered as a hypothesis that explains the greater variability in the choice of negative-tone emoji when writing similar contexts. This hypothesis was tested by using a tagged corpus of texts left by users that had passed the psychological test to determine the BFI degree.

We collected automatically and analysed the comments in the information system "Semograph". It was found that variability in the use of negative-tone types of emoticons and emoji (namely the icons with the semantics of sadness, fear, apathy and bewilderment) in the same context is more characteristic of different psychological types of personality. Conversely, positive-toned icons (namely the icons with the semantics of joy, love, thankfulness and approval) are less variable in the same context.

*Key words:* emoticons, emoji, WORD2VEC, BFI, social networks, internet communication.

**Acknowledgments:** The research is financially supported by Perm State University state assignment, project FSNF-2020-0023.

**For citation:** Belousov, K. I., Obukhova, I. A., Labutin, I.A. (2023). WORD2VEC- and semantic graph modeling of emoticons and emoji use in internet communication texts. *Voprosy Kognitivnoy Lingvistiki*, 2, 47-62 (In Russ.).