



оригинальная статья

<https://elibrary.ru/tssnmb>

Использование методов психодиагностики и векторной семантики для изучения влияния эмоционального интеллекта и пола информантов на понимание эмодзи

Обухова Ирина Андреевна

Пермский государственный национальный исследовательский университет, Россия, Пермь

eLibrary Author SPIN: 9448-2814

<https://orcid.org/0000-0002-7574-7833>

Scopus Author ID: 58309332000

ireneobukhova@gmail.com

Аннотация: В статье исследуется влияние социальных и психологических параметров информантов на интерпретацию эмодзи – графических знаков, отображающих эмоции и настроение пользователей социальных интернет-сервисов. Цель – выявить возможную зависимость интерпретаций эмодзи, полученных в группах информантов, от пола опрашиваемых и разных показателей их эмоционального интеллекта. Актуальность работы определяется растущей ролью эмодзи в онлайн-коммуникации, а практическая значимость видится в возможности учета изучаемых факторов информантов в задачах автоматического анализа тональности сообщений. Новизна исследования заключается в том, что сочетание методов психодиагностики и векторной семантики ранее не применялось для изучения взаимного влияния социальных и психологических параметров пользователей на понимание эмодзи. Методология исследования включает векторизацию текстов с использованием статистической меры TF-IDF для выделения набора ключевых слов, характеризующих каждый эмодзи, а также статистический анализ ключевых слов с применением U-критерия Манна-Уитни и методов коррекции р-значений. В результате выявлено, что гендерные различия в понимании эмодзи наиболее выражены в интерпретации положительных знаков, а уровень эмоционального интеллекта влияет на понимание как положительных, так и отрицательных эмодзи. Метод Монте-Карло, применяемый для проверки устойчивости выявленных различий между группами информантов, подтвердил устойчивость обнаруженных эффектов.

Ключевые слова: эмодзи, эмоциональный интеллект, пол, TF-IDF, векторизация текста, электронная коммуникация

Цитирование: Обухова И. А. Использование методов психодиагностики и векторной семантики для изучения влияния эмоционального интеллекта и пола информантов на понимание эмодзи. *Виртуальная коммуникация и социальные сети*. 2025. Т. 4. № 2. С. 115–125. <https://doi.org/10.21603/2782-4799-2025-4-2-115-125>

Поступила в редакцию 05.02.2025. Принята после рецензирования 13.03.2025. Принята в печать 17.03.2025.

full article

Effect of Emotional Intelligence and Gender on Emoji Interpretation: Psychodiagnostics and Vector Semantics

Irina A. Obukhova

Perm State University, Russia, Perm

eLibrary Author SPIN: 9448-2814

<https://orcid.org/0000-0002-7574-7833>

Scopus Author ID: 58309332000

ireneobukhova@gmail.com

Abstract: Emoji are easily comprehensible graphical symbols that represent emotions and moods in digital communication. However, internet users may interpret them according to their own social and psychological background. This research featured the effect of gender and emotional intelligence on emoji interpretation. The survey

results were processed by the methods of psychodiagnostics and vector semantics to reveal the co-effect of social and psychological parameters on semiotic comprehension. The method of text vectorization employed the TF-IDF statistics to extract the keywords for each emoji, which were then subjected to statistical analysis using the Mann-Whitney U-test and *p*-value correction. The gender differences appeared to be more relevant for positive symbols while the level of emotional intelligence was important for both positive and negative emoji. The Monte Carlo method confirmed the reliability of the correlations obtained. As the relevance of emoji in online communication keeps growing, these results may facilitate the automatic sentiment analysis of digital messages.

Keywords: emoji, emotional intelligence, gender, TF-IDF, text vectorization, digital communication

Citation: Obukhova I. A. Effect of Emotional Intelligence and Gender on Emoji Interpretation: Psychodiagnostics and Vector Semantics. *Virtual Communication and Social Networks*, 2025, 4(2): 115–125. (In Russ.) <https://doi.org/10.21603/2782-4799-2025-4-2-115-125>

Received 5 Feb 2025. Accepted after review 13 Mar 2025. Accepted for publication 17 Mar 2025.

Введение

Современная коммуникация все чаще осуществляется в цифровой среде, что подтверждается ростом числа интернет-пользователей¹, увеличением объема интернет-трафика [Бикеева 2022], а также развитием различных форм социального взаимодействия [Гавриленко 2021], таких как социальные сети, мессенджеры, платформы знакомств, видеоплатформы и др. В условиях расширяющегося цифрового взаимодействия особое значение приобретают невербальные средства коммуникации. Одним из таких инструментов стали эмодзи – картинки, иллюстрирующие «любой предмет, идею, состояние, в том числе и эмоции» [Турко 2024: 3010]. Эмодзи дополняют текстовые сообщения, позволяют передавать эмоции, оттенки смысла и намерения собеседников. Однако понимание этих знаков может варьироваться ввиду используемых устройств, культурных факторов, контекста, социальных и психологических параметров пользователей (пола, возраста, образования, эмоционального интеллекта и пр.).

В зависимости от производителя электронных устройств одни и те же эмодзи могут отображаться по-разному, что может стать причиной недопонимания между пользователями. Например, рассматривались различия в интерпретации эмодзи при использовании платформ Apple, Google, Microsoft, Samsung и LG [Miller et al. 2016]. Знаки одной семантики, предложенные пятью платформами, имеют разное графическое воплощение, однако набор черт, формирующих эмодзи, одинаков. Некоторые знаки могут восприниматься как более или менее эмоциональные и семантически богатые. «Даже

если эмодзи не отличаются на разных платформах, это еще не означает, что его эмоциональное значение интерпретируется одинаково» [Иванова, Пальмина 2021: 80]. Пользователи могут путать эмодзи разной семантики в силу их графического сходства друг с другом. Например, эмодзи с семантикой удивления (😮) похож на эмодзи с семантикой страха (😱), поскольку используются одни и те же акцентные элементы – широко открытый рот и округленные глаза.

Вариативность в понимании эмограмм рассматривается и в межкультурной коммуникации [Stanton 2018; Takahashi et. al 2017]. Так, камерунцы, танзанийцы и японцы по-разному толковали список одних и тех же эмотиконов, выражающих всего 3 эмоции: грусть, равнодушие и счастье [Takahashi et al. 2017]. Разное понимание эмотиконов среди представителей одной и разных культур подтверждает идею о том, что эмотикон не является универсальным знаком: «Несмотря на универсальный характер базовых эмоций и унификацию интерпретации эмотиконов носителями разных лингвокультур, обнаруживается ошибочность в декодировании из-за разницы восприятия экспрессии эмоции в том или ином эмотиконе в разных этноязыковых сообществах, что может служить маркером в определении коммуникантов на оси *свой – чужой* [Иванова-Якушко и др. 2024: 296]. Такая ошибочность в декодировании знаков свойственна и эмодзи.

В исследовании A. Oleszkiewicz и коллег проверялась способность детей в возрасте 4–8 лет распознавать эмоции людей, изображенных на фотографиях, и интерпретировать эмотиконы, репрезентирующие

¹ Факты и цифры 2023. Измерение цифрового развития. *Вестник связи*. 2024. № 2. С. 32–39. <https://elibrary.ru/dutkpe>

разные эмоции. Информанты практически без ошибок определяли, какие эмоции испытывают люди на фотографиях и какие эмоции отображаются с помощью эмодзи, причем точность распознавания эмоций в обоих случаях была выше у девочек и у детей старшего возраста. Эмоции счастья и печали распознавались детьми лучше, чем отвращение и страх [Oleszkiewicz et al. 2017].

Данная статья посвящена анализу влияния социальных и психологических факторов информантов, а именно пола и эмоционального интеллекта, на понимание эмодзи. Цель – выявить возможную зависимость интерпретаций эмодзи, полученных в группах информантов, от пола опрашиваемых и разных показателей их эмоционального интеллекта.

Для достижения цели работы были поставлены следующие задачи, соответствующие этапам ее проведения:

- 1) обработка текстового материала;
- 2) распределение материала по 4 группам информантов;
- 3) векторизация текстов для выделения ключевых слов, отражающих интерпретации каждого эмодзи, в каждой группе информантов;
- 4) проведение статистического анализа для проверки гипотезы о различии интерпретаций эмодзи в группах информантов;
- 5) выполнение проверки устойчивости различий.

Методы и материалы

Материалом исследования послужили ответы информантов, полученные в результате проведения онлайн-опроса², посвященного изучению эмодзи в медиакоммуникации. В опросе приняли участие 320 человек. Для диагностики эмоционального интеллекта информантов использовалась методика Д. В. Люсина – опросник ЭМИн [Люсин 2006; 2009; Люсин и др. 2004]. Выбор этой методики объясняется тремя причинами: 1) опросник ЭМИн является стандартом для измерения эмоционального интеллекта; 2) опросник разработан на российской выборке; 3) предварительные версии опросника были успешно апробированы на большом количестве испытуемых [Люсин 2006: 5–6].

Первая часть опроса состояла из 46 утверждений, проверяющих способность к пониманию своих и чужих эмоций и управлению ими. По отношению к каждому утверждению информантам требовалось

выразить степень своего согласия (предлагались следующие ответы: *совсем не согласен*, *скорее не согласен*, *скорее согласен* и *полностью согласен*). Результатом этой части опроса стал подсчет баллов ответов каждого испытуемого по 10 шкалам опросника ЭМИн. В работе использовалась интегральная шкала Общий уровень эмоционального интеллекта (далее в тексте будет использоваться аббревиатура этой шкалы – ЭМИн). Во второй части опроса относительно 13 эмодзи был задан следующий вопрос с открытым вариантом ответа: *Что означает этот эмодзи?* Всего на этот вопрос было получено 5163 ответа, из них 1742 уникальных. Ответы испытуемых автоматически сохранялись в электронную таблицу. Распределение количества информантов по полу и показателям шкалы ЭМИн представлено в таблице 1).

Как видно из таблицы 1, информанты были сбалансированы по параметру *пол*. Баллы, полученные по шкале ЭМИн, соотносились с квартилями – значениями, которые делят упорядоченную выборку на 4 примерно равные части. Первый квартиль приходится на 72 балла, второй (медиана) – 81 балл; третий – 92 балла. Так, были установлены следующие уровни ЭМИн информантов: 1) низкий (0–72 баллов); 2) средний (73–91 баллов); 3) высокий (92–132 баллов). Ответы самой многочисленной группы (информантов со средним ЭМИн) не будут анализироваться в статье, поскольку выявляется зависимость толкований эмодзи от взаимного сочетания пола и контрастных психологических параметров информантов, а именно низкого и высокого уровней ЭМИн мужчин и женщин. Отметим, что информантов с низким ЭМИн всего на 1 % больше,

Табл. 1. Распределение количества информантов по изучаемым параметрам
Tab. 1. Number of informants per group for each parameter

Параметры информантов	Группы информантов / уровни ЭМИн	Количество информантов	
		абс.	отн., %
Пол	мужчины	160	50,0
	женщины	160	50,0
ЭМИн (баллы)			
0–72	низкий ЭМИн	88	27,5
73–91	средний ЭМИн	147	46,0
92–132	высокий ЭМИн	85	26,5

² Онлайн-опрос. URL: <https://forms.gle/nCFNyBgVnZ4is7vv9> (дата обращения: 13.01.2025).

чем информантов с высоким ЭМИн (ср. 27,5 % и 26,5 % соответственно).

После установления уровня ЭМИн информанты были распределены на 4 группы с учетом взаимного сочетания пола и контрастных психологических показателей. Таблица 2 отражает количество информантов, образующих каждую группу, а также количество ответов, связанных с толкованием эмодзи, полученных от каждой группы. Отметим, что распределение количества информантов по выделенным группам вполне равномерное, как и распределение количества ответов в каждой группе.

Табл. 2. Количественные данные каждой группы информантов
Tab. 2. Quantitative data for each group

Группы информантов	Количество информантов		Количество ответов	
	абс.	отн.,%	абс.	отн.,%
женщины с высоким ЭМИн	43	13,4	600	11,6
женщины с низким ЭМИн	44	13,7	649	12,5
мужчины с высоким ЭМИн	42	13,1	572	11,0
мужчины с низким ЭМИн	44	13,7	628	12,0

В качестве стимульного материала исследования были взяты эмодзи, отображающие мимику и жесты, т.е. невербальные в лингвистическом и внешние в психологическом планах формы проявления тех или иных чувств человека. К принципам отбора стимулов относятся следующие аспекты:

- 1) интуитивная понятность знака;
- 2) разная частота употребления эмодзи в текстах электронной коммуникации;
- 3) разная тональность знаков (положительные, отрицательные и нейтральные);
- 4) культурная универсальность знака (исключались эмодзи, которые имеют специфическое значение в российской культуре, а также авторские эмодзи-неологизмы).

Для отбора эмодзи-стимулов привлекался корпус, который содержит векторные представления текстового массива социальной сети ВКонтакте (см. подробнее о корпусе в [Белоусов и др. 2023а: 3–4; 2023б: 48–50]. Этот корпус является частью платформы, интерфейс которой позволяет делать

запросы к хранящимся в репозитории векторным представлениям текстовых массивов³. В результате запроса к корпусу генерируется таблица встречаемости его единиц в текстовом массиве, семантическая близость запрашиваемых единиц корпуса в виде косинусных расстояний, а также информация о частоте использования этих единиц [Там же]. В нашем случае в качестве запрашиваемых единиц корпуса выступили эмодзи. Так, отбор эмодзи основывался на частотном анализе их использования в корпусе текстов ВКонтакте, что подтверждает факт их реального применения в онлайн-коммуникации. Значения отобранных знаков представлены в таблице 3 (для определения значений использовались две электронные библиотеки эмодзи: emojiworld и emojis⁴).

Для обработки и анализа собранного материала применялись библиотеки Python. **Первый этап** исследования заключался в предварительной обработке текстов: преобразование текстов в единый регистр; удаление знаков пунктуации и стоп-слов – слов, не несущих значимой смысловой нагрузки. С использованием библиотеки Natasha⁵ выполнены токенизация (разбиение текста на слова) и лемматизация – процесс приведения словоформ текста к их словарной форме, т.е. лемме (см. подробнее о предобработке текста на естественном языке у [Рудак, Зори 2024: 39–40; Пальмов, Салихов 2024: 45]. Далее лемматизированные тексты были представлены в формате словаря, в котором каждому эмодзи соответствовал список его лемм. На рисунке 1 показан фрагмент словаря эмодзи до этапа удаления из текстов стоп-слов.

На **втором этапе** лемматизированные тексты были отсортированы на 4 группы с учетом контрастных параметров информантов: 1) женщины с высоким ЭМИн; 2) женщины с низким ЭМИн; 3) мужчины с высоким ЭМИн и 4) мужчины с низким ЭМИн. Для каждой группы информантов составлен отдельный словарь, представленный 13 эмодзи и списками их лемм.

На **третьем этапе** происходила векторизация текстов с целью выделения набора ключевых слов для каждого эмодзи в 4 группах испытуемых. В качестве метода векторизации была выбрана статистическая мера TF-IDF (*Term Frequency – Inverse Document Frequency*), которая оценивает важность

³ Корпус. URL: <https://concept.semograph.com/concept> (дата обращения: 13.01.2025).
⁴ Библиотека emojiworld. URL: <https://www.emojiworld.ru/>; Библиотека emojis. URL: <https://emojis.wiki/ru/> (дата обращения: 10.01.2025).
⁵ Документация библиотеки Natasha. URL: <https://github.com/natasha/natasha> (дата обращения: 10.01.2025).

Табл. 3. Значения эмодзи-стимулов
Tab. 3. Official emoji interpretation

Эмодзи	Значения
😊	Улыбка, радость, счастье, удовольствие, благодарность. Может употребляться в саркастическом или ироническом смысле
😏	Действие <i>подмигивать</i> , юмор, флирт, одобрение, признательность и понимание
👍	Жест <i>большой палец вверх</i> ; выражает одобрение, симпатию, поддержку и согласие
😍	Любовь, влюбленность, симпатия, обожание, счастье, благодарность. Может использоваться в противоположных значениях
😄	Радостное лицо с ухмылкой; обозначает беспечность, смех, гордость; оба ряда зубов олицетворяют счастье
🙏	Сложенные вместе ладони; обозначает молитву, благодарность, просьбу, оказание поддержки, а также жест <i>дай пять</i> (приветствие или выражение дружеских чувств)
😱	Испуганное лицо; обозначает реальный или притворный (наигранный) страх, а также шок, испуг, стресс
😭	Сильный плач; выражает обиду, расстройство, горе, печаль, травму, боль, поражение, разочарование. Может использоваться в ироническом смысле
😲	Удивленное лицо; обозначает удивление, неожиданность, ошеломление, потрясение, шок
😐	Лицо без эмоций (<i>pokerface</i>); может выражать безразличие, скуку, усталость, молчаливое неодобрение
😡	Красное от злости лицо; выражает недовольство, раздражение, гнев. Может использоваться в ироническом смысле
😞	Разочарованное лицо; выражает грусть, печаль, сожаление, раскаяние, моральное перегорание, боль
🤢	Может выражать болезнь / отравление, рвоту, отвращение, неприязнь, безразличие

👍: ['супер', 'согласие', 'с', 'что-то', 'одобрение', 'согласие', 'ок']...
 😏: ['смех', 'смех', 'счастье', 'смех', 'радость', 'смех', 'я']...
 😄: ['подмигивание', 'подмигивание', 'в', 'значение', 'согласие', 'одобрение', 'согласие']...
 😍: ['улыбка', 'счастье', 'довольство', 'улыбка', 'радость', 'удовлетворение', 'очень']...
 😊: ['что-то', 'сильно', 'нравиться', 'полный', 'согласие', 'влюбленность', 'вау']...
 😐: ['безразличие', 'отсутствие', 'слово', 'легкий', 'негодование', 'покер', 'фэйс']...
 😞: ['расстройство', 'разочарование', 'печаль', 'расстройство', 'грусть', 'печаль', 'огорчение']...
 😡: ['злость', 'недовольство', 'злость', 'гнев', 'агрессия', 'злость', 'негодование']...
 😭: ['плач', 'сильный', 'печаль', 'горе', 'рыдание', 'расстройство', 'печаль']...
 😱: ['страх', 'шок', 'ужас', 'страх', 'шок', 'удивительный', 'неожидано']...
 😲: ['шок', 'шок', 'неверие', 'удивление', 'удивительный', 'удивление', 'сильный']...
 🙏: ['благодарность', 'хлопок', 'по', 'ладошка', 'согласие', 'с', 'собеседникам']...
 🤢: ['отвращение', 'тошнота', 'блевотина', 'в', 'значение', 'фу', 'гадость']...

Рис. 1. Фрагмент словаря эмодзи, составленного на основе лемм
Fig. 1. Lemma-based emoji thesaurus

слова в документе относительно всего корпуса документов. Параметр TF (частота слова) измеряет, как часто слово встречается в конкретном документе, т.е. представляет отношение числа вхождений некоторого слова к общему числу слов документа. Параметр IDF (обратная частота документа) измеряет, насколько редко это слово встречается в других документах корпуса. Общая формула TF-IDF умножает эти два показателя, повышая вес слов, которые часто встречаются в одном документе,

но редки в других, что помогает выделить наиболее значимые слова (см. подробнее о мере TF-IDF в [Булыга, Курейчик 2023: 218–219; Курганская, Степанова 2024: 1600; Бенгфорт и др. 2019: 90–91, 92]. Векторизация выполнена с применением библиотеки Scikit-Learn⁶.

Каждый текст был представлен как вектор, где каждому слову соответствует его вес (значимость). Была построена TF-IDF-матрица, где строки представляли собой отдельные ответы

⁶ Документация библиотеки Scikit-Learn. URL: <https://scikit-learn.org> (дата обращения: 10.01.2025).

информантов, а столбцы – уникальные слова. Таким образом, в каждой группе информантов для каждого эмодзи был рассчитан набор слов на основе TF-IDF-значений. На рисунке 2 приведен фрагмент словаря эмодзи, который составлен на текстовом материале группы женщин с высокими показателями ЭМИн.

Как видно из рисунка 2, количество слов для каждого эмодзи разное, что объясняется установленным пороговым значением TF-IDF $\geq 0,05$. Этот порог позволяет исключить из дальнейшего анализа слова с нулевыми TF-IDF-метриками и уменьшить выборку ключевых слов. В качестве примера представим фрагмент таблицы слов с их TF-IDF-метриками для эмодзи с семантикой удивления / шока (рис. 3). Таблица отражает слова, полученные сразу в 4 группах информантов. Если слово отсутствовало для какой-то группы, его TF-IDF-значение приравнивалось к нулю.

На **четвертом этапе** список слов с их TF-IDF-метриками для каждого эмодзи подвергался статистическому анализу для проверки гипотезы о различии интерпретаций эмодзи в группах информантов с учетом вероятности ошибок при множественных сравнениях. Применялся U-критерий Манна-Уитни, реализованный библиотекой SciPy⁷. Для корректировки p-значений при множественных сравнениях

использовались 3 метода: 1) Бонферрони, 2) Холма, 3) Беньямини-Хохберга.

U-критерий Манна-Уитни применялся для сравнения ключевых слов двух независимых групп с целью выявления значимых различий между их распределениями. Сравнивались следующие пары групп:

- 1) женщины с высоким ЭМИн и женщины с низким ЭМИн;
- 2) мужчины с высоким ЭМИн и мужчины с низким ЭМИн;
- 3) женщины с высоким ЭМИн и мужчины с высоким ЭМИн;
- 4) женщины с низким ЭМИн и мужчины с низким ЭМИн.

Первые две пары учитывают одинаковый пол информантов, при контрастном уровне их ЭМИн. Это дает возможность проследить влияние ЭМИн на интерпретацию эмодзи. Вторые две пары позволяют проследить влияние пола на интерпретацию эмодзи, поскольку учитывают одинаковый уровень ЭМИн и разный пол информантов. U-критерий Манна-Уитни с поправкой Бонферрони и двух аналогичных ему методов «выдают» в качестве результата «сырые» и скорректированные p-значения (p).

Была сформулирована нулевая (H_0) гипотеза, согласно которой интерпретация эмодзи в группах информантов не зависит от пола или уровня

'🙏': [('спасибо', 0.37), ('доброжелательность', 0.28), ('дружеский', 0.28), ('здорово', 0.28), ('окей', 0.28), ('поддержка', 0.28), ('флирт', 0.28), ('да', 0.14), ('знать', 0.14), ('иметь', 0.14), ('класс', 0.14), ('конфуз', 0.14), ('нехотеть', 0.14), ('ок', 0.14), ('отмечать', 0.14), ('помочь', 0.14), ('понять', 0.14), ('прикольно', 0.14), ('расположение', 0.14), ('согласие', 0.14), ('текст', 0.14), ('отвращение', 0.95), ('гадость', 0.17), ('бесить', 0.09), ('неприятный', 0.09), ('получиться', 0.09), ('сделать', 0.09), ('крайний', 0.08), ('недовольство', 0.08), ('нет', 0.08), ('улыбка', 0.65), ('радость', 0.58), ('доброжелательность', 0.19), ('рад', 0.13), ('удовлетворение', 0.13), ('хороший', 0.13), ('просто', 0.11), ('одобрение', 0.1), ('чувство', 0.09), ('благодарность', 0.06), ('взрослепожилой', 0.06), ('видеть', 0.06), ('вкус', 0.06), ('воспринимать', 0.06), ('жуткий', 0.06), ('испуг', 0.06), ('недовольство', 0.52), ('эмоция', 0.47), ('сообщение', 0.44), ('текст', 0.29), ('чувство', 0.21), ('выражение', 0.15), ('зависеть', 0.15), ('картинка', 0.15), ('контекст', 0.15), ('любовь', 0.15), ('симпатия', 0.15), ('шуточный', 0.15), ('постироничный', 0.12), ('человек', 0.11), ('использовать', 0.1)], ('благодарность', 0.73), ('спасибо', 0.54), ('бог', 0.19), ('бесить', 0.12), ('воспринимать', 0.12), ('неиспользовать', 0.12), ('отмечать', 0.12), ('помочь', 0.12), ('считать', 0.12), ('значение', 0.11), ('отрицательный', 0.11), ('только', 0.11), ('отличный', 0.1)], ('сильный', 0.56), ('расстройство', 0.52), ('грусть', 0.46), ('печаль', 0.35), ('разочарование', 0.17), ('очень', 0.12), ('капец', 0.06), ('сделать', 0.06), ('сильно', 0.06), ('сказать', 0.06), ('событие', 0.06), ('ужас', 0.06), ('хотеться', 0.06), ('крутой', 0.05), ('отрицательный', 0.05), ('плохой', 0.05), ('просто', 0.05)],

Рис. 2. Фрагмент словаря эмодзи, составленного на основе текстов группы женщин с высоким ЭМИн
Fig. 2. Emoji thesaurus compiled by women with high emotional intelligence

⁷ Документация библиотеки SciPy. URL: <https://scipy.org> (дата обращения: 10.01.2025).

all_results['😱']

Emoji	fem_high	fem_low	ml_high	ml_low
Word				
испуг	0.06	0.00	0.09	0.00
молодой	0.06	0.00	0.00	0.00
неприятный	0.00	0.07	0.00	0.00
сильный	0.10	0.15	0.06	0.08
удивление	0.77	0.61	0.96	0.82
ужас	0.25	0.00	0.00	0.00
шок	0.55	0.76	0.23	0.55

Прим.: fem_high – женщины с высоким ЭМИн; fem_low – женщины с низким ЭМИн; ml_high – мужчины с высоким ЭМИн; ml_low – мужчины с низким ЭМИн.

Рис. 3. Фрагмент таблицы ключевых слов и их TF-IDF-значения для эмодзи с семантикой удивления / шока
Fig. 3. Keywords and their TF-IDF-scores for emoji

ЭМИн информантов (статистических различий нет). Определен уровень значимости $\alpha = 0,05$, что означает 5 %-ую вероятность ошибочно отвергнуть H_0 , а также рассчитано р-значение: 1) если $p \leq \alpha$, отвергаем H_0 в пользу H_1 (есть статистически значимые различия); 2) если $p > \alpha$, нет оснований отвергнуть H_0 (различия незначимы).

На пятом этапе исследования применялся метод Монте-Карло, реализованный с помощью библиотеки Scikit-Learn, для проверки устойчивости выявленных различий между группами информантов. Была проведена тысяча итераций случайного разбиения данных, в каждой из которых 70 % выборки использовалось для статистического анализа. Оставшиеся 30 % выборки использовались для оценки стабильности результатов на разных подмножествах данных. На каждом

разбиении данных проводился U-критерий Манна-Уитни для сравнения TF-IDF-значений между сравниваемыми группами внутри каждого эмодзи. Это позволило оценить, насколько стабильно сохраняется различие между группами при изменении состава выборки.

По результатам каждой итерации зафиксировано р-значение для каждой пары групп. Рассчитана доля итераций, в которых различия между группами были статистически значимыми (при $p \leq 0,05$). Высокая доля значимых итераций свидетельствовала о надежности различий, тогда как низкие значения указывали на возможную нестабильность эффекта.

Результаты

В таблице 4 отражены эмодзи, для которых нулевая гипотеза, предполагающая отсутствие различий в интерпретации знаков в сравниваемых группах, отклоняется при установленном пороговом значении TF-IDF $\geq 0,05$ и $p \leq \alpha$. Были представлены «сырые» (3 колонка) и скорректированные (3 последних колонки) р-значения.

Исходя из приведенных данных (табл. 4), для всех трех корректирующих методов для эмодзи, означающие действие *подмигивать*, и эмодзи с семантикой *улыбка / радость* различия оказались статистически значимыми между группами женщин и мужчин с высоким ЭМИн. Скорректированные р-значения, которые меньше 0,05, подтверждают наличие значимых гендерных различий в интерпретациях этих эмодзи.

С поправкой Беньямини-Хохберга дополнительно выявляются значимые различия для других парных сравнений, например, между группами мужчин с высоким и с низким ЭМИн для эмодзи с семантикой *улыбка / радость* и *плач / разочарование*, что подтверждает влияние эмоционального интеллекта на интерпретацию знаков. С помощью

Табл. 4. Эмодзи со значимыми различиями в интерпретации среди парных групп информантов
Tab. 4. Emoji with significant statistical interpretation differences for paired groups

Эмодзи	Сравниваемые пары групп информантов	U-критерий Манна-Уитни	Поправка Бонферрони	Поправка Холма	Поправка Беньямини-Хохберга
😊	женщины с высоким ЭМИн	0,0003	0,0136	0,0133	0,0068
😊	и мужчины с высоким ЭМИн	0,0001	0,0034	0,0034	0,0034
😄	мужчины с высоким ЭМИн	0,0035	–	–	0,0335
😭	и мужчины с низким ЭМИн	0,0011	–	–	0,0143
🙏	женщины с высоким ЭМИн	0,0039	–	–	0,0335
😓	и женщины с низким ЭМИн	0,0011	–	–	0,0143

этого же метода выявлены значимые различия между группами женщин с высоким и с низким ЭмИн для знаков с семантикой *благодарности* / *просьбы* и *отвращения* / *неприязни*, что также подтверждает влияние эмоционального интеллекта на понимание эмодзи. Скорректированные р-значения подтверждают значимые статистические различия между указанными группами информантов в контексте понимания других эмодзи, которые не были выявлены при использовании поправки Бонферрони и метода Холма.

Далее с помощью метода Монте-Карло оценивалась устойчивость различий в интерпретации эмодзи, выявленных с помощью U-критерия Манна-Уитни и 3 корректирующих методов, между группами информантов. Доля итераций с р-значением $\leq 0,05$ показывает, насколько стабильно различия сохраняются при случайном разбиении выборки (табл. 5).

Стабильные различия. Различия в понимании эмодзи между группами женщин и мужчин с высоким ЭмИн оказались наиболее устойчивыми (строка 3): для эмодзи, отражающего действие *подмигивать*, значимость выявлялась в 98,4 % итераций, а для знака с семантикой *улыбка* / *радость* – в 99,7 % итераций (табл. 5). Поскольку у этих информантов одинаковый уровень ЭмИн, но разная половая принадлежность, можно утверждать, что пол значимо влияет на понимание этих двух эмодзи. Стабильные различия для этих эмодзи, выявленные методом Монте-Карло, подтверждают результаты U-критерия Манна-Уитни и корректирующих методов, где эти различия были статистически значимыми при всех 3 поправках.

Различия по уровню ЭмИн у групп мужчин и женщин. Значимость различий между мужчинами с высоким и низким ЭмИн подтверждается для эмодзи с семантикой *улыбки* / *радости* (77,1 %) и *плач* / *разочарование* (92,6 %), что согласуется с поправкой Беньямини-Хохберга, выявившей

их значимость. Аналогичным образом методом Монте-Карло подтверждены различия между группами женщин с высоким и низким ЭмИн для эмодзи с семантикой *благодарность* / *просьба* (77,7 %) и *отвращения* / *неприязни* (94,2 %). Эти различия также совпадают с результатами поправки Беньямини-Хохберга.

Менее устойчивые различия. Для эмодзи с семантикой *злости* / *гнева* менее устойчивые различия выявлены в 60,8 % итераций между мужчинами с высоким и с низким ЭмИн, что указывает на умеренную стабильность эффекта. А для эмодзи с семантикой *безразличия* различия между женщинами с высоким и низким ЭмИн были значимыми в 53,7 % итераций, что также свидетельствует о неустойчивости различий.

Соотношения метода Монте-Карло с результатами U-критерия Манна-Уитни и корректирующих методов. Метод Монте-Карло подтвердил значимость различий, выявленных U-критерием Манна-Уитни и поправкой Беньямини-Хохберга для 5 эмодзи: 😊, 😄, 🙄, 🙏 и 🙃. Различия, значимые только по U-критерию Манна-Уитни, но не подтвержденные методом Монте-Карло (например, знаки 😡 и 😞), могут быть менее устойчивыми и зависеть от состава выборки.

Как видно из таблицы 5, для сравниваемой пары *женщины с низким ЭмИн и мужчины с низким ЭмИн* не было выявлено статистических различий в интерпретации эмодзи.

Заключение

Пол и уровень эмоционального интеллекта информантов влияют на интерпретацию эмодзи. Гендерные различия наиболее выражены для понимания положительных знаков, а именно эмодзи с семантикой *улыбки* / *радости* и эмодзи, передающим действие *подмигивать*. Уровень эмоционального интеллекта информантов влияет на понимание как положительных, так и отрицательных знаков:

Табл. 5. Доля значимых ($p \leq 0,05$) итераций для парных сравнений
Tab. 5. Statistically significant ($p \leq 0.05$) iterations for pairwise comparisons

Парные группы информантов	😊	😄	🙄	🙏	🙃	😡	😞
женщины с высоким ЭмИн и женщины с низким ЭмИн	–	–	–	0,777	0,942	–	0,537
женщины с высоким ЭмИн и мужчины с высоким ЭмИн	0,984	0,997	–	–	0,723	–	–
мужчины с высоким ЭмИн и мужчины с низким ЭмИн	–	0,771	0,926	–	–	0,608	–
женщины с низким ЭмИн и мужчины с низким ЭмИн	–	–	–	–	–	–	–

среди положительных – эмодзи с семантикой улыбки /радости и благодарности / просьбы, а среди отрицательных – эмодзи с семантикой плача / разочарования, отвращения / неприязни и злости.

Анализ с использованием U-критерия Манна-Уитни и корректирующих методов подтвердил значимые различия между группами информантов в понимании изучаемых эмодзи. Метод Монте-Карло выявил устойчивость этих эффектов, особенно для положительных эмодзи – знака с семантикой улыбки /радости и знака, передающего действие подмигивать. В то же время нейтральный эмодзи безразличие продемонстрировал наибольшую вариативность интерпретаций и слабую зависимость от пола и эмоционального интеллекта.

В анализ интерпретаций эмодзи были включены слова как с высокими, так и с низкими TF-IDF-значениями. Слова с высокими TF-IDF-метриками соответствовали лексическим единицам, передающим ядерное, т.е. основное, первичное значение эмодзи, тогда как слова с низкими TF-IDF-метриками относились к ближней и дальней периферии, отражая вторичные, дополнительные значения. Включение в анализ периферийных значений может означать, что статистические различия между группами информантов проявляются не столько в понимании наиболее очевидных значений эмодзи, сколько в их дополнительных интерпретациях.

Различия в понимании эмодзи могут объясняться тем, что разные по полу и уровню эмоционального интеллекта информанты вариативнее толкуют эмодзи словами из ближней или дальней

периферии, тем самым смещая акценты интерпретаций. Вероятно, информанты с высокими показателями эмоционального интеллекта в тех или иных лингвистических контекстах лучше понимают разнообразие оттенков эмоций, передаваемых с помощью эмодзи. По нашему мнению чем выше эмоциональный интеллект, тем более разнообразна семантическая периферия толкований эмодзи, тогда как у информантов с низким эмоциональным интеллектом интерпретации могут быть более устойчивыми.

Перспективы исследования: необходимо рассмотреть вариативность толкований эмодзи в каждой группе информантов, а также выявить, толкования каких зон – ядерной, приядерной или периферийных – вносят наибольший вклад в статистические различия между группами в понимании этих знаков.

Конфликт интересов: Автор заявил об отсутствии потенциальных конфликтов интересов в отношении исследования, авторства и / или публикации данной статьи.

Conflict of interest: The author declared no potential conflict of interests regarding the research, authorship, and / or publication of this article.

Благодарности: Автор выражает благодарность своему научному руководителю – доктору филологических наук, профессору Белоусову Константину Игоревичу за советы в процессе подготовки статьи.

Acknowledgment: The author expresses her gratitude to Professor Konstantin I. Belousov, Doctor of Philology, for supervising the research.

Литература / References

- Белоусов К. И., Баширов Р. К., Зелянская Н. Л., Лабути И. А., Рябинин К. В., Чумаков Р. В. Профилирование концептуальных систем на основе комплекса методов психосемантики и машинного обучения. *Научно-техническая информация. Серия 2: Информационные процессы и системы*. 2023а. № 7. С. 1–14. [Belousov K. I., Bashirov R. K., Zelyanskaya N. L., Labutin I. A. Ryabinin K. V., Chumakov R. V. Profiling of conceptual systems based on a complex of methods of psychosemantics and machine learning. *Nauchno-tehnicheskaja informacija. Serija 2: Informacionnye processy i sistemy*, 2023a, (7): 1–14. (In Russ.)] <https://doi.org/10.36535/0548-0027-2023-07-1>
- Белоусов К. И., Обухова И. А., Лабути И. А. WORD2VEC- и графосемантические модели использования эмодзи и эмодзи в текстах интернет-коммуникации. *Вопросы когнитивной лингвистики*. 2023б. № 2. С. 47–62. [Belousov K. I., Obukhova I. A., Labutin I. A. WORD2VEC- and semantic graph modeling of emoticons and emoji use in Internet communication texts. *Issues of Cognitive Linguistics*, 2023b, (2): 47–62. (In Russ.)] <https://doi.org/10.20916/1812-3228-2023-2-47-62>
- Бенгфорт Б., Билбро Р., Охеда Т. Прикладной анализ текстовых данных на Python. Машинное обучение и создание приложений обработки естественного языка. СПб.: Питер, 2019. 368 с. [Bengfort B., Bilbro R.,

- Ojeda T. *Applied text analysis with python. Enabling language-aware data products with machine learning*. St. Petersburg: Piter, 2019, 368. (In Russ.)]
- Бикеева М. В. Оценка цифровой активности населения в контексте теории поколений. *Статистика и экономика*. 2022. Т. 19. № 5. С. 25–34. [Bikeeva M. V. Assessment of digital activity of the population in the context of the theory of generations. *Statistics and Economics*, 2022, 19(5): 25–34. (In Russ.)] <https://doi.org/10.21686/2500-3925-2022-5-25-34>
- Булыга Ф. С., Курейчик В. М. Сравнительный анализ методов векторизации текстовых данных большой размерности. *Известия ЮФУ. Технические науки*. 2023. № 2. С. 212–226. [Bulyga Ph. S., Kureichik V. M. Comparative analysis of methods of vectorization of high dimensional text data. *Izvestiya SFedU. Engineering Sciences*, 2023, (2): 212–226. (In Russ.)] <https://doi.org/10.18522/2311-3103-2023-2-212-226>
- Гавриленко О. В. Цифровизация как фактор трансформации новых форм социального взаимодействия и механизмов контроля. *Коммуникативный капитал цифровой эпохи*: науч. онлайн-конф. с Междунар. уч. (Москва, 12 февраля 2021 г.) М.: МАКС Пресс, 2021. С. 57–64. [Gavrilenko O. V. Digitalization as a factor in the transformation of new forms of social interaction and control mechanisms. *The communication capital of the digital age*: Proc. Sci. Online Conf. with Intern. participation, Moscow, 12 Feb 2021. Moscow: MAKSS Press, 2021, 57–64. (In Russ.)] <https://elibrary.ru/ubiulk>
- Иванова И. В., Пальмина К. С. Влияние эмодзи и эмотиконов на итоговую эмоциональную окраску текста при сентимент-анализе. *Научные исследования в современном мире. Теория и практика*: Всерос. (нац.) науч.-практ. конф. (Санкт-Петербург, 10 мая 2021) СПб.: ГНИИ Нацразвитие, 2021. С. 79–82. [Ivanova I. V., Palmina K. S. Effect of emojis and emoticons on the final emotional coloring of the text in the sentimental analysis. *Scientific research in the modern world. Theory and practice*: Proc. All-Russian (National) Sci.-Prac. Conf., St. Petersburg, 10 May 2021. St. Petersburg: HNRI National development, 2021, 79–82. (In Russ.)] <https://elibrary.ru/ztqzce>
- Иванова-Якушко М. М., Анумян К. С., Карабулатов М. Н. Эмотикон как пример развития графической эмо-тиологии в виртуальном пространстве. *Виртуальная коммуникация и социальные сети*. 2024. Т. 3. № 3. С. 296–308. [Ivanova-Yakushko M. M., Anumyan K. S., Karabulatov M. N. Emoticon in development of virtual graphic emotionology. *Virtual Communication and Social Networks*, 2024, 3(3): 296–308. (In Russ.)] <https://doi.org/10.21603/2782-4799-2024-3-3-296-308>
- Курганская Е. В., Степанова Н. В. Выявление «токсичности» в социальных сетях на основании критерия семантической близости. *Филологические науки. Вопросы теории и практики*. 2024. Т. 17. № 5. С. 1596–1603. [Kurganskaia E. V., Stepanova N. V. Identification of "toxicity" in social networks based on the semantic proximity criterion. *Philology. Theory & Practice*, 2024, 17(5): 1596–1603. (In Russ.)] <https://doi.org/10.30853/phil20240231>
- Люсин Д. В. Новая методика для измерения эмоционального интеллекта: опросник ЭмИн. *Психологическая диагностика*. 2006. № 4. С. 3–22. [Lyusin D. V. A new technique for measuring emotional intelligence: The EmIn questionnaire. *Psikhologicheskaya diagnostika*, 2006, (4): 3–22. (In Russ.)] <https://elibrary.ru/vzmifn>
- Люсин Д. В. Опросник на эмоциональный интеллект ЭмИн: новые психометрические данные. *Социальный и эмоциональный интеллект: от процессов к измерениям*, ред. Д. В. Люсин, Д. В. Ушаков. М.: Институт психологии РАН, 2009. С. 264–278. [Lyusin D. V. EmIn emotional intelligence questionnaire: New psychometric data. *Social and emotional intelligence: From processes to measurements*, eds. Lyusin D. V., Ushakov D. V. Moscow: Institut psikologii RAN, 2009, 264–278. (In Russ.)] <https://elibrary.ru/slanqv>
- Люсин Д. В., Марютина О. О., Степанова А. С. Структура эмоционального интеллекта и связь его компонентов с индивидуальными особенностями: эмпирический анализ. *Социальный интеллект: Теория, измерение, исследования*, ред. Д. В. Люсин, Д. В. Ушаков. М.: Институт психологии РАН, 2004. С. 129–140. [Lyusin D. V., Maryutina O. O., Stepanova A. S. The structure of emotional intelligence and the relationship of its components to individual characteristics: An empirical analysis. *Social Intelligence: Theory, measurement, and research*, eds. Lyusin D. V., Ushakov D. V. Moscow: Institut psikologii RAN, 2004, 129–140. (In Russ.)] <https://elibrary.ru/idtctm>
- Пальмов С. В., Салихов Р. Р. Сравнительный анализ библиотек Pymorphy3 Pymystem3. *Наука и бизнес: пути развития*. 2024. № 6. С. 45–49. [Palmov S. V., Salikhov R. R. A comparative analysis of Pymorphy3

- and Pymystem3 libraries. *Science and Business: Development Ways*, 2024, (6): 45–49. (In Russ.)] <https://elibrary.ru/wwjfkx>
- Рудак Л. В., Зори С. А. Обработка текста методами естественного языка. *Информатика и кибернетика*. 2024. № 3. С. 39–44. [Rudak L. V., Zori S. A. Text processing using natural language methods. *Informatics & Cybernetics*, 2024, (3): 39–44. (In Russ.)] <https://elibrary.ru/vhnwmf>
- Турко У. И. Смайлик, эмотикон и эмодзи как невербальные средства общения в интернет-коммуникации. *Филологические науки. Вопросы теории и практики*. 2024. Т. 17. № 9. С. 3005–3011. [Turko U. I. Smiley, emoticon, and emoji as nonverbal communication tools in internet communication. *Philology. Theory & Practice*, 2024, 17(9): 3005–3011. (In Russ.)] <https://doi.org/10.30853/phil20240427>
- Miller H., Thebault-Spieker J., Chang S., Johnson I., Terveen L., Hecht B. "Blissfully happy" or "ready to fight": Varying interpretations of emoji. *Tenth International AAAI Conference on Web and Social Media (ICWSM 2016)*: Proc. Conf., Cologne, 17–20 May 2016. 2016, 10(1): 259–268. <https://doi.org/10.1609/icwsml.v10i1.14757>
- Oleszkiewicz A., Frackowiak T., Sorokowska A., Sorokowski P. Children can accurately recognize facial emotions from emoticons. *Computers in Human Behavior*, 2017, 76: 372–377. <https://doi.org/10.1016/j.chb.2017.07.040>
- Stanton A. L. Islamic emoticons and religious authority: Emerging practices, shifting paradigms. *Contemporary Islam*, 2018, 12(2): 153–171. <https://doi.org/10.1007/s11562-017-0412-8>
- Takahashi K., Oishi T., Shimada M. Is 😊 smiling? Cross-cultural study on recognition of emoticon's emotion. *Journal of Cross-Cultural Psychology*, 2017, 48(10): 1578–1586. <https://doi.org/10.1177/0022022117734372>