JOURNALFLEDU.COM
Foreign Languages in Uzbekistan

Scientific Electronic Journal "Foreign Languages in Uzbekistan", 2026, vol. 12, No 5 (70), pp. 46–62
https://journalfledu.com
ISSN: 2181-8215 (online)

STRUCTURAL ANALYSIS OF HATER COMMENTS: A COMPARATIVE STUDY ACROSS FOUR LANGUAGES

Nikita Sergeevich Subbota

Lecturer of the Department of English Philology
Brest State University named after Alexander Sergeyevich Pushkin
Brest, Belarus
sns-englang@list.ru
https://orcid.org/0000-0001-5569-4167
https://doi.org/10.36078/1794210768

(received: September 11, 2026; revised: October 09, 2026; accepted: October 20, 2026)

Abstract. The article presents a comparative analysis of the structural organization of negatively evaluative user comments. The dataset comprises 4,000 public Instagram comments, with 1,000 comments in each of four languages: Russian, English, German, and Polish. The annotation scheme covers text length, graphic and paralinguistic devices, repetition, relations between verbal and visual modes, and erratives. Absolute and relative frequencies are supplemented with Pearson chi-square tests and Cramer's V. Short, single-sentence contributions predominate in all four subsets. Russian comments are the most compact, whereas the German subset has the highest mean number of words and characters. The most frequent paralinguistic resource in every subset is the co-occurrence of emoji and verbal text. This pattern reaches its maximum in German and its minimum in English. English comments are also the most likely to remain exclusively verbal. The strongest cross-language effects concern emoji accompanying text and the semiotic composition of a comment, although the effect sizes remain small. The findings describe the four sampled subsets and should not be interpreted as direct evidence of language- or culture-level causation without controlling topics, source posts, and platform contexts.

Keywords: internet comment; verbal aggression, hater discourse, contrastive linguistics, multimodality, emoji, corpus annotation.

СТРУКТУРНЫЙ АНАЛИЗ ХЕЙТЕРСКИХ КОММЕНТАРИЕВ: СРАВНИТЕЛЬНОЕ ИССЛЕДОВАНИЕ НА МАТЕРИАЛЕ ЧЕТЫРЕХ ЯЗЫКОВ

Никита Сергеевич Суббота

Преподаватель кафедры английской филологии
Брестский государственный университет
имени Александра Сергеевича Пушкина
Брест, Республика Беларусь

Аннотация. Статья посвящена сопоставительному анализу структурной организации отрицательно-оценочных пользовательских комментариев. Материал составил корпус из 4000 единиц, собранных в открытых профилях Instagram: по 1000 комментариев на русском, английском, немецком и польском языках. Разметка охватывала протяженность текста, графические и паралингвистические средства, типы повторов, соотношение вербального и визуального кодов, а также эрративы. Абсолютные и относительные частоты дополнены критерием хи-квадрат Пирсона и коэффициентом V Крамера. Установлено, что все подкорпусы характеризуются преобладанием однофразовых и коротких реплик, однако русскоязычные комментарии наиболее компактны, а немецкоязычные имеют наибольшее среднее число слов и знаков. Главным паралингвистическим ресурсом во всех выборках выступает сочетание текста с эмодзи; его доля максимальна в немецком подкорпусе и минимальна в английском. Англоязычные единицы чаще сохраняют исключительно вербальную форму. Наиболее выраженные межъязыковые эффекты связаны с включением эмодзи в вербальный текст и с семиотическим составом комментария, но величины эффектов остаются небольшими. Результаты описывают различия между конкретными подкорпусами и не позволяют без контроля тематики и стимулов непосредственно приписывать их языку или лингвокультуре.

Ключевые слова: интернет-комментарий; речевая агрессия; хейтеринг; сопоставительная лингвистика; мультимодальность; эмодзи; корпусная разметка.

Введение

Обращение к структурной организации хейтерских комментариев целесообразно начать с жанровой природы самого интернет-комментария. Как отмечает Т. Н. Колокольцева, интернет-комментарий представляет собой «реактивный (ответный) интернет-жанр», текст которого тематически, структурно-семантически и прагматически связан с исходным сообщением и образует с ним «своего рода диалогическое единство» [Колокольцева, 2024: 167]. Иными словами, комментарий не существует в коммуникации изолированно: его краткость, степень развёрнутости и выбор выразительных средств во многом определяются тем, на какой пост он отвечает и какое место занимает в разворачивающемся обсуждении. Именно поэтому при анализе комментария недостаточно учитывать только количество предложений или слов. Не менее важен способ организации его материальной формы: сочетание вербального текста с эмодзи, графическое выделение, повтор, изменение регистра, нарушение орфографической нормы. Совокупность этих параметров позволяет увидеть, каким образом отрицательная оценка получает внешнее оформление в цифровой коммуникации.

Оценочность в данном случае выступает не случайным сопутствующим признаком, а одним из жанрообразующих свойств комментария. Е. Ю. Викторова и К. В. Пантеева, анализируя

комментарии к постам в Instagram, показывают, что анонимность и дистантность интернет-общения создают условия для более свободного и эксплицитного выражения отрицательной оценки. По их наблюдению, косвенные средства — риторические вопросы, риторические восклицания, метафоры и неологизмы — помогают автору «привлечь больше внимания к комментарию» и усилить речевое воздействие на адресата [Викторова & Пантеева, 2023: 71]. Для настоящего исследования это наблюдение принципиально важно. Отрицательная оценка может быть выражена не только лексически, но и структурно: через краткость реплики, графическое выделение, повтор или включение визуального компонента. Поэтому далее рассматривается не содержание отдельных инвектив, а формальная организация комментария как целостного поликодового сообщения.

При этом необходимо сразу обозначить терминологические границы исследования. В настоящей статье словосочетание «хейтерский комментарий» используется как рабочее обозначение отрицательно-оценочной реакции на публичное лицо или связанное с ним событие. Данный термин не является синонимом понятия *hate speech* (язык вражды).

Как показывает систематический обзор Ф. Полетто и его соавторов, в исследованиях цифровой агрессии используется целый ряд терминов *(hate speech, abusive language, offensive language и aggression)*, однако их определения, объекты и схемы аннотирования не совпадают. Авторы подчёркивают, что *hate speech* представляет собой сложное и многогранное понятие, а размытость границ между смежными понятиями приводит к неоднородности корпусов и различиям в результатах классификации [Poletto et al., 2021: 478, 481].

Сходную проблему с другой стороны рассматривают З. Васим и соавторы, предлагая различать задачи обнаружения агрессивной речи по типу адресата, содержанию и коммуникативной функции [Waseem et al., 2017]. М. Эль-Шериф и коллеги, в свою очередь, подчёркивают необходимость учитывать целевой объект высказывания и развивают целевой подход (target-based approach) к анализу языка вражды [ElSherief et al., 2018]. Поэтому исследуемая здесь выборка, сформированная вокруг комментариев о звёздах и публичных личностях, не должна автоматически интерпретироваться как корпус языка вражды, так как принадлежность адресата к защищаемой социальной группе не являлась критерием отбора.

Уточнив терминологию, мы неизбежно сталкиваемся с вопросом о том, как именно негативная оценка «упаковывается» в цифровом сообщении. Для ответа на него нелишним будет вспомнить о

гибридной природе сетевой речи. С. Ю. Имбер, анализируя Instagram-комментарии, характеризует их как реализацию «новой устно-письменной формы общения» и связывает с фонетико-графическими и синтаксическими сдвигами [Имбер, 2020: 31, 42]. Именно эта двойственность объясняет, почему на экране вдруг появляются графические суррогаты интонации, жестов или мимики. Капслок, намеренное удлинение гласных, шквал восклицательных знаков, эмодзи — всё это работает на уровне, функционально близком к невербалике.

Развивая эту мысль в более широком контексте, М. В. Иванова и Н. И. Клушина говорят о системности языковых изменений и делают акцент на особой роли креативной функции в пользовательском контенте [Иванова & Клушина, 2021: 368]. Что это значит для нашего исследования? Прежде всего, что любое отклонение от нейтральной нормы в комментарии — вовсе не обязательно случайная оплошность автора. Часто это осознанная стилизация или способ эмоционального усиления. Впрочем, в рамках данной работы мы не пытаемся автоматически реконструировать все эти скрытые функции. Мы фиксируем их строго как формальные признаки, помня об ограничениях самой разметки.

Отсюда вытекает одно критически важное методологическое правило. Графический маркер нельзя слепо приравнивать к агрессии. Да, верхний регистр, эрративы или повторы знаков препинания способны передавать конфликтную интонацию, но их реальная функция всегда диктуется контекстом. О. Ю. Гукосьянц, например, прямо включает капслок и эрративы в список маркеров конфликтного поведения. Но, и это важно, рассматривает их только в связке с лексикой и прагматикой [Гукосьянц, 2023: 313]. Еще более осторожную позицию занимает Т. А. Иванченко. Исследователь подчеркивает: агрессивный характер немецкоязычного комментария «выявляется в первую очередь в ходе интерпретации плана содержания», ведь враждебные смыслы далеко не всегда прорываются на поверхностный уровень [Иванченко, 2026: 2478].

Формальная разметка в нашей работе описывает лишь «поверхность» сообщения, и она ни в коем случае не должна подменять прагматический анализ. Это различие становится буквально критичным, когда мы переходим к интерпретации статистически значимых межъязыковых различий.

Отдельного разговора заслуживает визуальный компонент. Пожалуй, именно он сильнее всего связывает структурный и прагматический уровни. Е. И. Галяшина обращает внимание на полифункциональность

пиктограмм: они могут как дополнять текст, так и радикально менять его смысл, трансформировать эмоцию или даже маскировать истинную интенцию автора [Галяшина, 2022]. В зарубежной науке схожие идеи продвигают Л. Логи и М. Заппавинья. Опираясь на социальную семиотику, они доказывают, что эмодзи не существуют в вакууме: их значение рождается только во взаимодействии с языковым окружением. Авторы справедливо подчеркивают высокую контекстуальную зависимость эмодзи, рассматривая их как ресурс для выражения эмоций [Logi & Zappavigna, 2023].

Именно поэтому, когда в ходе разметки мы используем тег *emoji intens*, мы просто фиксируем факт совместного присутствия эмодзи и вербального текста. В таких случаях мы не всегда имеем дело с усилением негатива. Количественные данные лишь показывают, насколько часто эти два кода взаимодействуют. А вот какую именно функцию несёт знак в каждой конкретной реплике — вопрос для отдельного, глубокого контекстного анализа. Более того, сопоставление подкорпусов требует предельной осторожности. Наблюдаемая разница в частотах может быть связана вовсе не с языком как таковым, а с темой поста, фигурой адресата, негласными правилами конкретного сетевого сообщества или вообще характеристиками исходного стимула.

Показательна в этом плане работа Н. А. Климковой, где русские и немецкие комментарии сопоставлялись на материале одного и того же новостного сюжета. Автор резонно замечает, что одинаковый стимул даёт «симметричные данные для проведения анализа» [Климкова, 2024]. В нашем корпусе такого жесткого контроля нет: комментарии собраны из разных профилей, под разными публикациями. Следовательно, полученные результаты нужно трактовать именно как различия между сформированными языковыми выборками, а не как некие прямые и абсолютные свойства русского, английского, немецкого или польского языков.

Эта важная оговорка задаёт и всю дальнейшую логику нашего анализа. Сначала мы выстраиваем общий структурный фон для всего корпуса. Затем точечно рассматриваем конкретные межъязыковые расхождения. И только после этого оцениваем их реальный масштаб и, главное, содержательную интерпретируемость.

Цель исследования — выявить общие и различающиеся структурные характеристики хейтерских комментариев в русском, английском, немецком и польском подкорпусах. Для этого сопоставляются протяженность реплик, графические и паралингвистические средства,

типы повторов, соотношение вербального и визуального кодов и частота эрративов.

Материалы и методы

Материал исследования составили негативно-оценочные комментарии из открытых профилей Instagram, специализирующихся на освещении событий из жизни звёзд и публичных личностей. Данные собраны методом автоматизированного парсинга с последующей ручной выборкой. Общий объём корпуса — 4000 комментариев: по 1000 единиц на русском, английском, немецком и польском языках. Равный размер подкорпусов обеспечивает сопоставимость относительных частот.

Включение двух славянских и двух германских языков создаёт симметричную основу для сопоставления, но генетическая принадлежность языка сама по себе не объясняет пользовательское поведение. На структуру комментария могут влиять тема исходного поста, медиастатус адресата, нормы конкретного сообщества, интерфейсная версия платформы и характеристики аудитории. Поэтому единицей вывода в работе является языковой подкорпус, а не лингвокультура в целом.

На первом этапе для каждого комментария автоматически подсчитывались число предложений, слов и знаков. Подсчёт выполнялся с помощью авторского скрипта на языке Python; сводные показатели рассчитывались стандартными функциями Microsoft Excel. Для каждой языковой выборки определялись минимальное, среднее и максимальное значения. Дополнительно комментарии распределялись по числу предложений, по интервалам длины в словах, а также по числу знаков в целом.

На втором этапе вручную кодировались графические и паралингвистические признаки согласно заданным тегам. Так, тег *CAPS* присваивался при наличии фрагмента в верхнем регистре; *punct intens* — при явной интенсификации пунктуации; *mock quoting* — при псевдоцитировании или насмешливом выделении чужих слов кавычками; *code switch* — при переключении языкового кода внутри комментария; *emoji intens* — при сочетании эмодзи с вербальным текстом; *emoji only* — при отсутствии вербального текста и наличии только эмодзи; *emoji wordrep* — при замещении слова или морфемы графическим знаком. Важно отметить, что указанные теги не образуют взаимоисключающую классификацию, поэтому один комментарий мог одновременно иметь несколько признаков.

Отдельный блок разметки описывал повторы. Фиксировались повторы знаков препинания, букв, слов или словосочетаний, а также повторы эмодзи. Дополнительно фиксировалось количество комментариев, в которых повторы отсутствуют. Подтипы повторов рассматривались как потенциально совместимые признаки в виду того, что один комментарий может содержать одновременно несколько разновидностей повтора.

Семиотический состав кодировался четырьмя категориями с использованием заданного набора тегов. Категория *textual* — включала исключительно вербальные комментарии; *textual_visual* — единицы, где вербальная часть предшествовала визуальному компоненту или структурно доминировала; *visual_textual* — единицы с начальной либо доминирующей визуальной частью при наличии текста; *visual* — сообщения без вербального компонента. Дополнительно отмечались эрративы, то есть орфографические, пунктуационные и грамматические отклонения.

Количественный анализ основывался на абсолютных и относительных частотах. Для сопоставления распределений комментариев по числу предложений и слов применялся критерий *хи-квадрат Пирсона (χ^2)*. Он позволяет проверить, согласуются ли наблюдаемые различия с предположением об одинаковом распределении исследуемого признака в четырёх подкорпусах: чем меньше соответствующее *p-значение*, тем меньше совместимость данных с этим предположением. Порог статистической значимости был установлен на уровне 0,05. При этом общий результат проверки характеризует совокупность четырёх подкорпусов и сам по себе не означает, что различия значимы между каждой парой языков.

Аналогичным образом отдельно проверялись десять бинарных признаков, фиксируемых по принципу «есть / нет»: семь графических и паралингвистических тегов, наличие хотя бы одного типа повтора, исключительно вербальный состав комментария и наличие эрративов. Для этой серии проверок применялась *поправка Холма*, ограничивающая риск ошибочно признать различия значимыми вследствие одновременного анализа нескольких признаков. Выводы о статистической значимости в данной серии делались по скорректированным p-значениям.

Выраженность выявленных различий оценивалась с помощью *коэффициента V Крамера*, принимающего значения от 0 до 1: чем ближе значение к нулю, тем слабее связь между признаком и принадлежностью комментария к определённому языковому подкорпусу. Таким образом, *p-значение* использовалось для оценки

статистической значимости, а *V Крамера* — для характеристики силы связи. Эти показатели интерпретировались совместно с частотами, поскольку при объёме корпуса 4000 единиц статистически значимыми могут оказаться и небольшие различия. Отсутствие статистической значимости, в свою очередь, не рассматривалось как доказательство полного совпадения подкорпусов.

Исследование использует опубликованные в открытом доступе пользовательские тексты, однако в статье представлены только агрегированные показатели. Имена пользователей, адреса профилей и дословные примеры не приводятся. Структурная разметка описывает форму сообщения. Вывод о коммуникативной функции отдельного графического элемента или о наличии имплицитной агрессии требует контекстного анализа исходной публикации и ветви обсуждения.

Результаты исследования

Протяженность и сегментация комментариев. Во всех четырех подкорпусах преобладают короткие одночастные реакции, но степень компрессии различается (табл. 1). Среднее число предложений варьируется от 1,18 в русскоязычной выборке до 1,42 в англоязычной. Одним предложением ограничиваются 86,7 % русских, 80,4 % польских, 74,6 % немецких и 71,9 % английских комментариев. Максимальные значения подтверждают ту же последовательность на верхней границе: шесть предложений в русском подкорпусе, семь в польском, восемь в немецком и девять в английском.

Распределения по числу предложений статистически различаются: $\chi^2(24)=92,07;\ p<0,001;\ V=0,088$. Коэффициент V указывает на небольшой эффект. Следовательно, различие устойчиво внутри данного массива, но не отменяет общего для всех четырех языков доминирования однофразовой модели.

Таблица 1.

Показатели протяженности комментариев

Язык	Среднее число предложений	Одно предложение, %	Среднее число слов	Интервал ≤5 слов, %	Среднее число знаков с пробелами
Русский	1,18	86,7	5,24	65,6	33,85
Английский	1,42	71,9	8,57	44,7	48,05
Немецкий	1,36	74,6	9,18	38,0	58,17
Польский	1,28	80,4	7,80	50,5	49,24

Различия проявляются сильнее при подсчёте слов. Русскоязычный комментарий содержит в среднем 5,24 слова, польский — 7,80, английский — 8,57, немецкий — 9,18. Доля первого интервала составляет 65,6 % в русском подкорпусе, 50,5 % в польском, 44,7 % в английском и 38,0 % в немецком. В диапазон 11–20 слов, напротив, входит 22,4 % немецких, 18,6 % английских, 15,1 % польских и 10,0 % русских комментариев. Доля единиц длиннее 20 слов невелика во всех случаях, но достигает 8,8 % в немецкой и 7,6 % в английской выборках против 6,2 % в польской и 2,1 % в русской.

Распределения по словесным интервалам также неодинаковы: $\chi^2(24)=207{,}64;\ p<0{,}001;\ V=0{,}132$. Этот эффект остаётся небольшим, но превосходит эффект сегментации по предложениям. Среднее число знаков с пробелами согласуется со словесной длиной лишь частично: немецкий подкорпус имеет максимум 58,17 знака, английский и польский близки друг к другу — 48,05 и 49,24, русский минимум равен 33,85. Такое соотношение зависит одновременно от числа слов и их графической длины, поэтому показатели знаков нельзя сводить к синтаксической развёрнутости.

Полученные данные позволяют перейти от общей характеристики протяжённости к вопросу о том, какими средствами краткая реакция компенсирует отсутствие развёрнутого вербального описания. Если длина комментария отражает степень его текстовой компрессии, то графические средства показывают, каким образом эта компрессия может сопровождаться эмоциональным и визуальным усилением.

Графические и паралингвистические средства. Наиболее частотным графическим ресурсом во всех подкорпусах является сочетание эмодзи с вербальным текстом. Тег *emoji intens* присвоен 49,7 % немецких, 39,3 % польских, 37,8 % русских и 25,0 % английских комментариев (табл. 2). Межъязыковое различие по этому признаку имеет наибольшую величину среди проверенных бинарных параметров: $\chi^2(3)=130{,}63;$ *p с поправкой Холма* $<0{,}001;\ V=0{,}181$. Эффект остаётся небольшим, но показывает, что визуальное сопровождение текста является главным источником структурной дифференциации подкорпусов.

Таблица 2.

Графические и паралингвистические признаки в процентах

Язык	CAPS	PI	MQ	CS	EI	EO	EWR
Русский	2,3	7,2	1,3	7,9	37,8	5,6	1,5
Английский	7,0	9,7	0,7	4,2	25,0	4,4	4,3
Немецкий	3,9	9,6	2,0	4,7	49,7	3,9	0,8
Польский	1,4	8,0	2,3	5,5	39,3	3,9	0,9

Примечание. *PI — интенсификация пунктуации; MQ — псевдоцитирование; CS — переключение кода; EI — эмодзи при вербальном тексте; EO — только эмодзи; EWR —замещение слова или морфемы эмодзи. Признаки могли сочетаться в одном комментарии.*

Верхний регистр чаще отмечен в английской выборке — 7,0 %, затем в немецкой — 3,9 %, русской — 2,3 % и польской — 1,4 %. Различия подтверждены после поправки на множественные сравнения: $\chi^2(3)=51,67;\ p<0,001;\ V=0,114$. Прописные буквы указывают на графическое выделение, но сами по себе не доказывают повышенную агрессивность английских комментариев. Они могут маркировать акцент, имитацию громкости, цитирование или название. Достоверно установить их прагматический статус можно только в связи с содержанием.

Переключение кода наиболее часто встречается в русском подкорпусе — 7,9 %; в польском оно составляет 5,5 %, в немецком 4,7 %, в английском 4,2 %. Эффект статистически значим после поправки, но всё равно остаётся небольшим: $\chi^2(3)=15,33;\ p=0,006;\ V=0,062$. Направление переключения, длина и функция иноязычного включения отдельно не фиксировались. Очевидно, что данный показатель свидетельствует о неодинаковой частоте смешения кодов. Однако без обращения к содержанию определить, связано ли оно с заимствованием, цитированием, обращением к иностранной аудитории или экспрессивной стилизацией не представляется возможным.

Замещение слова или морфемы эмодзи заметно выделяет английскую выборку: 4,3 % против 1,5 % в русской, 0,9 % в польской и 0,8 % в немецкой. Различия значимы: $\chi^2(3)=44,17;\ p<0,001$ *после поправки;* $V=0,105$. Эта конструкция отличается от простого добавления эмодзи: визуальный знак занимает позицию внутри вербальной структуры и участвует в построении пропозиционального смысла. Такое взаимодействие соответствует социальной семиотической трактовке, согласно которой языковая и визуальная части совместно реализуют значение [Logi & Zappavigna, 2023].

Интенсификация пунктуации имеет близкие частоты — от 7,2 % до 9,7 % и не даёт значимого межъязыкового различия после поправки Холма: $\chi^2(3)=5,74;\ p=0,249;\ V=0,038$. Не различается и частота комментариев только с эмодзи: от 3,9 % до 5,6 %; $\chi^2(3)=4,54;\ p=0,249;\ V=0,034$. Для псевдоцитирования исходное p равно 0,019, но скорректированное значение равно 0,056. Эти результаты представляются важными для интерпретации, поскольку дают основания говорить о том, что не каждый визуально заметный приём формирует специфический профиль отдельного подкорпуса.

Повторы. Отсутствие повторов доминирует во всех подкорпусах: 78,1 % английских, 77,2 % польских, 72,8 % русских и 71,2 % немецких комментариев (табл. 3). Соответственно, хотя бы один тип повтора встречается в 21,9 % английских, 22,8 % польских, 27,2 % русских и 28,8 % немецких единиц. Общая зависимость наличия повтора от языка статистически значима, но величина эффекта мала: $\chi^2(3) = 17{,}84$; $p = 0{,}002$ *после поправки Холма;* $V = 0{,}067$.

Таблица 3.

Повторы в структуре комментариев в процентах

Язык	Повторы отсутствуют	Пунктуация	Буквы	Лексические единицы	Эмодзи
Русский	72,8	8,0	2,5	1,1	16,0
Английский	78,1	9,9	2,1	1,0	8,8
Немецкий	71,2	9,4	1,8	1,5	17,2
Польский	77,2	7,6	1,7	0,3	13,8

Примечание. *Подтипы повторов могли сочетаться. Поэтому их сумма не используется как оценка общей доли комментариев с повторами.*

Главный источник различий внутри блока — повтор эмодзи. Он отмечен в 17,2 % немецких, 16,0 % русских, 13,8 % польских и 8,8 % английских комментариев. Повтор пунктуации значительно ровнее: от 7,6 % в польском до 9,9 % в английском подкорпусе. Повтор букв не отличается высокой частотой и находится в диапазоне 1,7–2,5 %; лексический повтор ещё менее частотен — 0,3–1,5 %. Следовательно, редупликация визуального элемента выполняет в корпусе более заметную структурную функцию, чем удлинение букв или повтор вербального фрагмента.

Сочетание высокой доли *emoji intens* и повторов эмодзи особенно характерно для немецкой выборки. Этот результат согласуется с данными о графических средствах агрессивных немецких комментариев, где повтор описывается как способ передачи высокой эмоциональности и экспрессии [Иванченко, 2026: 2480], но не доказывает одинаковую функцию для каждого конкретного случая. Чтобы различить усиление, серийную реакцию, иронию и визуальное замещение, необходим анализ последовательности конкретных знаков и их позиции относительно текста.

Семиотический состав и эрративы. Наконец, необходимо определить, как отдельные графические приёмы соотносятся с общей

организацией сообщения. Для этого частные теги сопоставляются с категориями семиотического состава: только текст, текст с визуальным компонентом, визуальный компонент с последующим текстом и только визуальная реакция.

Исключительно вербальная организация преобладает во всех подкорпусах, но её доля различается. Максимальное значение зафиксировано для английских комментариев — 72,4 %, далее следуют польские — 60,0 %, русские — 55,8 % и немецкие — 50,2 % (табл. 4). По дополнению к доле текстовой информации визуальный компонент участвует в 49,8 % немецких, 44,2 % русских, 40,0 % польских и 27,6 % английских единиц. Зависимость текстовой монокодовости от языка значима: *$\chi^2(3)=110,80$; $p<0,001$ после поправки Холма; $V=0,166$*.

Таблица 4.

Семиотический состав комментариев и эрративы в процентах

Язык	Только текст	Текст и визуальный компонент	Визуальный и текстовый компоненты	Только визуальный компонент	Эрративы
Русский	55,8	36,6	2,0	5,6	5,6
Английский	72,4	22,7	0,7	4,2	4,2
Немецкий	50,2	44,4	1,3	4,1	1,4
Польский	60,0	35,2	0,9	3,9	2,0

Примечание. *Категории семиотического состава взаимоисключающие и в каждом подкорпусе суммируются до 100 %. Эрративы учитывались отдельно.*

Во всех смешанных сообщениях преобладает последовательность, при которой текст предшествует визуальному компоненту или структурно доминирует: 44,4 % немецких, 36,6 % русских, 35,3 % польских и 22,7 % английских комментариев. Обратная модель *visual_textual* не отличается высокой частотой — от 0,7 % до 2,0 %. Комментарии только с визуальным компонентом составляют 3,9–5,6 %. Таким образом, различия связаны прежде всего с включением эмодзи в вербальную реплику, а не с вытеснением текста самостоятельными визуальными реакциями.

Распределение дополняет данные отдельных графических тегов. В немецком подкорпусе максимальны и доля *textual_visual*, и частота *emoji intens*; в английском одновременно максимальна доля *textual* и минимальна частота эмодзи при тексте. Однако данные категории не являются полностью взаимозаменяемыми. Визуальный компонент может включать разные типы знаков, а ручная семиотическая

классификация учитывает последовательность и доминирование, тогда как тег фиксирует наличие конкретного приёма.

Эрративы чаще отмечены в русской выборке — 5,6 %, следом в английской — 4,2 %, польской — 2,0 % и немецкой — 1,4 %. Различия значимы: $\chi^2(3)=35,72$; $p<0,001$ *после поправки Холма;* $V=0,095$.

Обсуждение результатов

Совокупность полученных данных заставляет нас переосмыслить саму природу хейтерского комментария. Это не просто «длинный» или «короткий» текст. Фактически перед нами компактная реактивная реплика, где предельная вербальная экономия буквально наслаивается на разнообразное графическое оформление. Почему так происходит? Однофразовая модель, общая для всех четырех подкорпусов, прямо указывает на жесткую жанровую привязку комментария к исходному посту. Стимул уже взял на себя часть предметного контекста. Комментатору не нужно описывать ситуацию — он может целиком сосредоточиться на эмоции, оценке или резком несогласии.

Межъязыковые различия при этом не складываются в простой набор изолированных «национальных» черт. Скорее, мы видим несколько устойчивых конфигураций. Возьмём русскую выборку: здесь мы фиксируем максимальную компрессию текста, но парадоксальным образом она компенсируется активным визуальным сопровождением и частым переключением кода. Английский материал идёт другим путем. Тексты здесь отличаются большей развернутостью, преобладают чисто вербальные сообщения. Однако любопытная деталь: именно в английском подкорпусе эмодзи чаще выступают не как дополнение, а как прямой заменитель слова или даже морфемы. Немецкая выборка выдаёт иную тенденцию — это самые длинные комментарии в среднем, плюс рекордная вовлечённость визуального кода. А что с польским языком? Формально он занимает промежуточное положение, но сводить его к некой «усредненной» позиции между остальными языками было бы явной ошибкой. Конфигурация получается специфической.

Но, пожалуй, самое важное открытие кроется не в самих параметрах экспрессивности. Максимальные межъязыковые эффекты проявляются в зоне взаимодействия вербального и визуального кодов. Это наблюдение позволяет нам существенно уточнить общий вывод. Структурная специфика каждого подкорпуса зависит не от того, насколько агрессивно используется графика «вообще». Ключевой фактор — это именно способы вживления визуального элемента в ткань текста. И эти способы, как показал анализ, кардинально разнятся.

Заключение

Подводя итог, мы можем констатировать: хейтерский комментарий во всех четырёх языковых средах реализуется прежде всего как короткая, однофразовая реакция. Она жёстко привязана к исходному посту. Любопытно, что межъязыковые различия, если оценивать их через значения V Крамера, не демонстрируют существенных различий. В основном это эффекты небольшого размера, которые вполне укладываются в единую жанровую модель.

Взглянем на объём текстов. Наиболее компактными, как и ожидалось, оказались русскоязычные комментарии. Немецкоязычные, напротив, демонстрируют наибольшую среднюю длину как в словах, так и в знаках. Английская выборка выделяется тем, что чаще сохраняет исключительно вербальную организацию, обходясь без визуальных вкраплений. Польский материал формально занимает промежуточное положение. Однако здесь необходимо сделать важную методологическую оговорку. Все эти выводы относятся строго к исследованным подкорпусам. Переносить их на языки или лингвокультуры в целом, без дополнительного контроля стимулов и тематики, было бы некорректно.

Но, пожалуй, самый интересный результат кроется не в длине текстов, а в соотношении вербального и визуального кодов. Это наш главный дифференцирующий параметр. Эмодзи в связке с обычным текстом чаще всего встречаются в немецком подкорпусе и реже всего — в английском. При этом английские комментарии демонстрируют иную тенденцию: там эмодзи гораздо чаще выступают не как дополнение, а как прямой заменитель слова или морфемы. Иными словами, языки различаются не просто количеством визуальных элементов, а самим способом их вживления в структуру сообщения.

Отдельно внимания заслуживают повторы. Наш анализ показал, что они вовсе не являются обязательным атрибутом хейтерской реплики. Более 70 % комментариев в каждом подкорпусе вообще не содержат зафиксированных повторов. Там же, где они встречаются, безоговорочно лидируют эмодзи. Повторы букв или целых лексических единиц явление гораздо более редкое. С пунктуацией история иная. Интенсификация знаков препинания демонстрирует близкие частоты во всех языках и, что важно, после поправки на множественные сравнения не формирует никакого значимого межъязыкового различия.

Многоуровневая разметка, безусловно, позволяет чётко очертить профиль исследуемого жанра. Но она ни в коем случае не отменяет необходимости глубокого семантико-прагматического анализа. Мы

ещё раз убедились в том, что формальный показатель никогда не равен коммуникативной функции. Один и тот же эмодзи, механический повтор или капслок способны как усиливать агрессию, так и оформлять иронию, маркировать цитату или вообще выполнять нейтральную роль. Именно поэтому перспективным вектором для дальнейшей работы нам видится объединение структурной разметки с тонкой классификацией: учёт объекта негативной оценки, степени её эксплицитности, конкретных речевых стратегий и, конечно, контекста исходного поста.

Использованная литература

Викторова Е. Ю., Пантеева К. В. Жанр интернет-комментария: аксиологический аспект // Жанры речи. — 2023. — Т. 18, № 1 (37). — С. 66–73. https://doi.org/10.18500/2311-0740-2023-18-1-37-66-73

Галяшина Е. И. Семиотика эмотиконов и анимационных картинок в аспекте судебной лингвистической экспертизы // Вестник Университета имени О. Е. Кутафина (МГЮА). — 2022. — № 2 (90). — С. 41–48. https://doi.org/10.17803/2311-5998.2022.90.2.041-048

Гукосьянц О. Ю. Некоторые маркеры конфликтопровоцирующей интенции авторов интернет-комментариев в период пандемии COVID-19 // Медиалингвистика. — 2023. — Т. 10, № 3. — С. 301–318. https://doi.org/10.21638/spbu22.2023.302

Иванова М. В., Клушина Н. И. Русский язык в современном интернет-пространстве: динамические процессы и тенденции развития // Русистика. — 2021. — Т. 19, № 4. — С. 367–382. https://doi.org/10.22363/2618-8163-2021-19-4-367-382

Иванченко Т. А. Вербализация агрессии в интернет-комментариях к общественно-политическим видеороликам (на материале телеканала «Второе немецкое телевидение») // Филологические науки. Вопросы теории и практики. — 2026. — Т. 19, вып. 6. — С. 2477–2485. https://doi.org/10.30853/phil20260314

Имбер С. Ю. Разговорная речь в социальных сетях (на примере комментариев к новостным публикациям в Instagram) // Вестник Московского государственного лингвистического университета. Гуманитарные науки. — 2020. — № 12 (841). — С. 31–45. — URL: https://www.vestnik-mslu.ru/Vest/12_841_H.pdf (дата обращения 15.10.2026).

Климкова Н. А. Визуальный компонент пользовательского комментария к новостному интернет-тексту в русской и немецкой

лингвокультурах // VERBA. Северо-Западный лингвистический журнал. — 2024. — № 4 (14). — С. 48–59. https://doi.org/10.34680/VERBA-2024-4(14)-48-59

Колокольцева Т. Н. Интернет-комментарий как сетевой жанр: общая характеристика, диалогический потенциал, проблемы классификации // Жанры речи. — 2024. — Т. 19, № 2 (42). — С. 164–173. https://doi.org/10.18500/2311-0740-2024-19-2-42-164-173

ElSherief M., Kulkarni V., Nguyen D., Wang W. Y., Belding E. Hate Lingo: A Target-Based Linguistic Analysis of Hate Speech in Social Media // Proceedings of the International AAAI Conference on Web and Social Media. — 2018. — Vol. 12, No. 1. — P. 42–51. https://doi.org/10.1609/icwsm.v12i1.15041

Logi L., Zappavigna M. A social semiotic perspective on emoji: How emoji and language interact to make meaning in digital messages // New Media & Society. — 2023. — Vol. 25, No. 12. — P. 3222–3246. https://doi.org/10.1177/14614448211032965

Poletto F., Basile V., Sanguinetti M., Bosco C., Patti V. Resources and benchmark corpora for hate speech detection: A systematic review // Language Resources and Evaluation. — 2021. — Vol. 55, No. 2. — P. 477–523. https://doi.org/10.1007/s10579-020-09502-8

Waseem Z., Davidson T., Warmsley D., Weber I. Understanding Abuse: A Typology of Abusive Language Detection Subtasks // Proceedings of the First Workshop on Abusive Language Online (Vancouver, Canada, 4 August 2017) / eds. Z. Waseem, W. H. K. Chung, D. Hovy, J. Tetreault. — Vancouver: Association for Computational Linguistics, 2017. — P. 78–84. https://doi.org/10.18653/v1/W17-3012

References

ElSherief, M., Kulkarni, V., Nguyen, D., Wang, W. Y., & Belding, E. (2018). Hate lingo: A target-based linguistic analysis of hate speech in social media, *Proceedings of the International AAAI Conference on Web and SocialMedia*, 12 (1), 42–51. https://doi.org/10.1609/icwsm.v12i1.15041

Galyashina, E. I. (2022). *Vestnik Universiteta imeni O. E. Kutafina*, No. 2 (90), 41–48. https://doi.org/10.17803/2311-5998.2022.90.2.041-048

Gukos'yants, O. Yu. (2023). *Medialingvistika*, vol. 10, No. 3, 301–318. https://doi.org/10.21638/spbu22.2023.302

Imber, S. Yu. (2020). *Vestnik Moskovskogo gosudarstvennogo lingvisticheskogo universiteta. Gumanitarnye nauki*, No. 12 (841), 31–45,

available at: https://www.vestnik-mslu.ru/Vest/12_841_H.pdf (accessed 15.10.2026)

Ivanchenko, T. A. (2026). *Filologicheskie nauki. Voprosy teorii i praktiki*, vol. 19, No. 6, 2477–2485. https://doi.org/10.30853/phil20260314

Ivanova, M. V., & Klushina, N. I. (2021). *Rusistika*, vol. 19, No. 4, 367–382. https://doi.org/10.22363/2618-8163-2021-19-4-367-382 (In Rus.)

Klimkova, N. A. (2024). *VERBA. Severo-Zapadnyi lingvisticheskii zhurnal*, No. 4 (14), 48–59. https://doi.org/10.34680/VERBA-2024-4(14)-48-59 (In Rus.)

Kolokol'tseva, T. N. (2024). *Zhanry rechi*, vol. 19, No. 2 (42), 164–173. https://doi.org/10.18500/2311-0740-2024-19-2-42-164-173

Logi, L., & Zappavigna, M. (2023). A social semiotic perspective on emoji: How emoji and language interact to make meaning in digital messages, *New Media & Society*, 25 (12), 3222–3246. https://doi.org/10.1177/14614448211032965

Poletto, F., Basile, V., Sanguinetti, M., Bosco, C., & Patti, V. (2021). Resources and benchmark corpora for hate speech detection: A systematic review, *Language Resources and Evaluation*, 55 (2), 477–523. https://doi.org/10.1007/s10579-020-09502-8

Viktorova, E. Yu., & Panteeva, K. V. (2023). *Zhanry rechi*, vol. 18, No. 1 (37), 66–73. https://doi.org/10.18500/2311-0740-2023-18-1-37-66-73 (In Rus.)

Waseem, Z., Davidson, T., Warmsley, D., & Weber, I. (2017). Understanding abuse: A typology of abusive language detection subtasks. In Z. Waseem, W. H. K. Chung, D. Hovy, & J. Tetreault (Eds.), Proceedings of the First Workshop on Abusive Language Online (Vancouver, Canada, August 4, 2017, pp. 78–84). Association for Computational Linguistics. https://doi.org/10.18653/v1/W17-3012