

ЛЕКСИЧЕСКИЙ ПОРТРЕТ СОВРЕМЕННЫХ ОТЕЧЕСТВЕННЫХ СЕРИАЛОВ: КОЛИЧЕСТВЕННЫЙ АНАЛИЗ

LEXICAL PORTRAIT OF CONTEMPORARY RUSSIAN TV SERIES:A QUANTITATIVE ANALYSIS

D. Vasilev
I. Guryanov

Summary: This article explores linguistic features in Russian TV series across different genres using computational linguistics methods. The study analyzes textual transcripts from four popular series released within the last five years: Fisher (2023), Slovo patsana. Krov' na asfalte (2023), Trigger (2021), and Inspector Gavrilov (2024). The methodology incorporates structural analysis, readability assessment, lexical diversity metrics, and sentiment analysis using natural language processing (NLP) models. Findings reveal genre-specific patterns: detective and drama series exhibit negative sentiment and taboo lexis; thrillers show the highest lexical diversity with neutral emotional tones; comedies feature minimal text volume and predominantly positive evaluations.

Keywords: audiovisual discourse, TV series discourse, serial text, quantitative analysis, lexical diversity, sentiment analysis.

Васильев Данил Олегович
аспирант, Казанский федеральный университет
danielcrow2000@mail.ru
Гурьянов Игорь Олегович
кандидат филологических наук, доцент, Казанский
федеральный университет
IOGuryanov@yandex.ru

Аннотация: Статья посвящена исследованию языковых особенностей российских сериалов различных жанров с применением методов компьютерной лингвистики. В работе анализируются текстовые транскрипты четырёх популярных сериалов за последние 5 лет – «Фишер» (2023), «Слово пацана. Кровь на асфальте» (2023), «Триггер» (2021) и «Инспектор Гаврилов» (2024). Методология исследования включает структурный анализ, оценку удобочитаемости, лексическое разнообразие и сентимент-анализ с использованием моделей естественной обработки языка. Результаты выявили жанровую специфику: детективы и драмы отличаются негативной тональностью и табуированной лексикой; триллер – максимальным лексическим разнообразием и нейтральной эмоциональной окраской; комедия – минимальным объемом текста и преобладанием позитивных оценок.

Ключевые слова: аудиовизуальный дискурс, сериалный дискурс, сериалный текст, количественный анализ, лексическое разнообразие, сентимент-анализ.

В последние годы рынок российских сериалов демонстрирует стремительный рост. Согласно данным аналитического сервиса «Кинопоиск Pro», апрель 2023 года ознаменовался безоговорочным лидерством отечественных сериалов в рейтингах популярности. [9].

Текущая ситуация демонстрирует значительное укрепление позиций российского контента на внутреннем рынке. Особенно показательной стала динамика марта 2023 года, когда российские проекты достигли 43% представленности в рейтинге, в то время как доля американского контента сократилась до 22%. Среди других стран-лидеров по производству популярных сериалов выделяются Япония, Турция и Южная Корея [9].

По данным Telegram-канала «Новости кинопроизводства» 2024 год ознаменовался существенными изменениями в структуре производства телевизионного контента в России: наблюдается заметная переориентация рынка в сторону изменения традиционных форматов сериалного производства. Так, в прошлом году количество произошло резкое сокращение производства сериалов среднего метража, которые традиционно составляли основу программирования как цифровых платформ, так и эфирного телевидения. Эта тенденция затронула как новые проекты, так и продолжения успешных франшиз [7].

Рост популярности отечественных сериалов, отмечаемый аналитиками в 2023–2024 годах, сопровождается важными трансформациями не только в количественных показателях производства, но и в языковой организации сериалных текстов. Это актуализирует вопрос о специфике речевого построения современных отечественных сериалов.

Материал и методы исследования

Основу методологии исследования составил компьютерный лингвистический анализ текста с использованием языка программирования Python и специализированных библиотек для обработки естественного языка (NLP).

В качестве материала анализа выступили транскрипты диалогов из четырех популярных российских сериалов различных жанров: детектива «Фишер», криминальной драмы «Слово пацана. Кровь на асфальте», психологического триллера «Триггер» и комедии «Инспектор Гаврилов».

В исследовании проведен структурный анализ текстов сериалов с расчетом ключевых показателей. Плотность текста оценивалась по количеству слов и предло-

жений в минуту экранного времени. Сложность языка определялась адаптированными индексами Флеша и SMOG, автоматически рассчитываемыми на основе слоговой структуры и длины предложений. Лексическое разнообразие анализировалось через метрики TTR, RTTR и CTTR, учитывающие соотношение уникальных и общих слов. Для оценки эмоциональной окраски применялась модель rubert-base-cased-sentiment, классифицирующая предложения на позитивные, нейтральные и негативные.

Сериал как форма аудиовизуального дискурса

Аудиовизуальный дискурс (далее АВД) представляет собой сложное явление, объединяющее вербальные и невербальные элементы в единое смысловое пространство. В отличие от естественной речи, которая возникает спонтанно в процессе непосредственного общения, речь в аудиовизуальном произведении является результатом комплексной совместной работы сценаристов, режиссёров и актёров. Эта речь создаёт у зрителя ощущение повседневного общения, однако при этом она подчиняется специфике жанра, нарратива и ожиданиям зрителей.

Сериалы представляют собой уникальный формат аудиовизуального дискурса, сочетающий вербальные и невербальные средства коммуникации. В отличие от кинофильмов, их разворачивающийся формат повествования позволяет детально развивать сюжетные линии и характеры персонажей, формируя особую связь со зрителем. Ключевой особенностью сериального дискурса является необходимость поддержания речевой идентичности персонажей через характерные языковые маркеры (специфическая лексика, синтаксические конструкции, профессиональный жаргон), сохраняющиеся на протяжении всего повествования. Этот комплексный феномен объединяет устную речь, визуальные образы и музыкальное сопровождение в единое смысловое пространство.

Создание художественных диалогов, воспринимающихся как естественные и правдоподобные, представляет собой одну из центральных проблем как в киносценаристике. Трудность заключается не только в попытке точного воспроизведения спонтанной разговорной речи, сколько в отборе и умелом использовании специфических дискурсивных характеристик, которые зритель способен однозначно идентифицировать как признаки естественной коммуникации [15]. Согласно Д.В. Жабину организация спонтанной речи на фонетическом, лексическом и синтаксическом уровнях существенно отличается как от письменной речи, так и от других видов устной коммуникации [1], а также характеризуется незавершёнными речевыми конструкциями, грамматическими колебаниями, самокоррекцией, хезитацией, повторами и оговорками [11]. Эти элементы от-

ражают естественный процесс речепорождения, когда говорящий одновременно формулирует и корректирует высказывание. Однако в рамках аудиовизуального дискурса наблюдается существенный парадокс: достижение эффекта спонтанной устной коммуникации требует тщательной предварительной текстуальной организации.

Важность изучения сериальной речи заключается в её влиянии на языковые нормы и коммуникативные практики. В отличие от письменной речи, где выбор языковых средств осознан и подчинён строгим правилам, или естественной устной речи, где коммуникация про текает непреднамеренно [5], кинодиалог занимает промежуточное положение. С одной стороны, кинотекст, как лингвистический компонент АВД, стремится к убедительности [4, 10], воспроизводя особенности живой речи, которая может быть как спонтанной, так и подготовленной [2, 3, 13]. С другой – эти элементы подражания гипертрофированы, отобраны и стилизованы под конкретный жанр. Стоит также отметить, что каждый лингвистический компонент речевого портрета персонажа обладает определенной семиотической нагрузкой [4].

В аудиовизуальном дискурсе любое произведение не просто передаёт информацию, а конструирует «новую реальность», предлагая зрителю поверить в правдоподобие происходящего, которое может быть далёким от реальной действительности. Аутентичность в аудиовизуальных произведениях создаётся через тщательный подбор опорных подробностей [4, 16]. Эти детали, которые могут быть физическими, социальными, эмоциональными, языковыми или поведенческими, служат своеобразными якорями, позволяющими зрителю достраивать недостающие элементы «кинореальности».

Как было указано ранее, в отличие от естественной устной речи, возникающей спонтанно в процессе коммуникации, тексты аудиовизуальных произведений тщательно конструируются на письме, но должны восприниматься как неподготовленные [6]. В современной лингвистике данное явление получило название псевдоустности – искусственного воспроизведения устной речи, которая лишь имитирует спонтанность, оставаясь заранее продуманной и отредактированной. Однако до сих пор не было выработано единого подхода к определению термина «псевдоустная речь» или «псевдоустный кинотекст». Эта терминологическая неопределенность требует тщательного концептуального осмысливания.

Исследование сериальной речи имеет важное социолингвистическое значение, поскольку медиа не только отражают, но и формируют языковые тенденции. Речевые модели из популярных сериалов активно проникают в повседневное общение, особенно среди молодежи. Яркий пример – сериал «Друзья» (1994), который не только отразил языковые особенности американской

молодёжи 1990-х, но и сформировал новые речевые паттерны, вышедшие далеко за пределы экрана. Одним из наиболее показательных случаев стало введение в массовый обиход термина «френдзона» (friendzone) – понятия, описывающего ситуацию, когда односторонние романтические чувства одного человека встречают лишь дружеское отношение со стороны объекта симпатии.

Количественный анализ лексического уровня сериальных текстов

В рамках настоящего исследования был проведён количественный анализ лексического состава сериальных текстов на материале четырёх популярных отечественных сериалов, отобранных по критериям высокого рейтинга на платформе «Кинопоиск» и жанрового разнообразия: «Слово пацана. Кровь на асфальте», «Фишер», «Триггер» и «Инспектор Гаврилов». Общий объём проанализированного видеоматериала составил более 30 часов сериального контента.

Базовые статистики текстов данных сериалов включают в себя подсчёты среднего количества слов и предложений на один эпизод, среднее количество слов в предложении. Был произведен расчёт плотности текста во времени: количество слов и предложений в минуту за один эпизод. (Таб. 1.)

Сериал «Фишер» лидирует по объёму текста, но отличается разумеренным темпом – 16 предложений/минуту. «Слово пацана» и «Триггер» при схожем количестве слов демонстрируют более динамичное повествование. Выраженные отличия демонстрирует «Инспектор Гаврилов» с минимальным объёмом текста, но максимальной плотностью речи.

Рассмотрим характеристики текстов сериалов с точки зрения удобочитаемости. Для этого мы воспользуемся следующими метриками: индекс удобочитаемости Флеша и индекс SMOG. (Таб. 2.)

Анализ удобочитаемости проводился по адаптированной формуле Флеша [8]. Полученные значения свидетельствуют о высокой доступности текстов для широкой аудитории. Параллельно рассчитанный индекс SMOG указывает на среднюю сложность, характерную для образованной аудитории. Оба индекса отражают структурные характеристики текста, но оценивают разные аспекты сложности восприятия.

Представленные результаты анализа лексического разнообразия текстов сериалов с использованием метрик TTR, RTTR и CTTR позволяют оценить богатство словарного запаса и разнообразие языковых средств, используемых в сериальной речи. (Таб. 3.)

Таблица 1.

Общая статистика сериальных текстов.

Сериал	Среднее количество слов в эпизоде	Среднее количество предложений в эпизоде	Средняя продолжительность эпизода (минуты)	Плотность слов	Плотность предложений
Фишер	5000	826	52	96	16
Слово пацана. Кровь на асфальте	4882	1181	51	95	23
Триггер	4840	1153	52	93,08	23
Инспектор Гаврилов	2100	697	25	84,00	28

Таблица 2.

Анализ удобочитаемости текстов сериалов.

Индекс	Название сериалов			
	Фишер	Слово пацана. Кровь на асфальте	Триггер	Инспектор Гаврилов
Индекс Флеша	86	98	87	95
SMOG	9	8	9	8

Таблица 3.

Коэффициенты лексического разнообразия.

Сериал	Метрики		
	TTR	RTTR	CTTR
Фишер	0,21	40,88	28,91
Слово пацана. Кровь на асфальте	0,24	39,07	27,627
Триггер	0,14	41,64	29,44
Инспектор Гаврилов	0,31	35,13	24,84

Анализ лексического разнообразия показал, что сериал «Инспектор Гаврилов» имеет наибольший базовый показатель TTR, свидетельствующий о частом использовании уникальных слов. Однако более точную оценку дают скорректированные метрики RTTR и CTTR, по которым сериал «Триггер» демонстрирует наибольшее лексическое разнообразие. «Фишер» также показал высокие значения, подтверждающие значительное разнообразие словаря. Эти данные указывают, что при интерпретации TTR необходимо учитывать длину текста, используя дополнительные метрики для объективной оценки.

В ходе исследования эмоциональной окраски текстовых материалов популярных сериалов был применен метод сентимент-анализа с использованием технологий обработки естественного языка (NLP). Основой для анализа послужила предобученная модель rubert-base-cased-sentiment от платформы Hugging Face. (Рис. 1.)

Исследование эмоциональной окраски четырёх российских сериалов выявило чёткую жанровую специфику. «Фишер» и «Слово пацана. Кровь на асфальте» демонстрируют преобладание мрачной тональности, соответствующей их тяжёлой тематике - расследованию убийств и жизни криминальных группировок. В противоположность этому, комедийный «Инспектор Гаврилов» отличается ярко выраженной позитивной окраской. Выделяется «Триггер» с доминированием нейтрального тона, что отражает его психологическую направленность.

Проведенное исследование выявило значительные различия в использовании грубой лексики в четырех популярных российских сериалах. (Таб. 4.)

Анализ выявил 403 случая употребления вульгарной лексики в исследуемых сериалах. Наибольшая частота зафиксирована в «Триггере», тогда как «Фишер» показал наиболее агрессивный профиль с максимальной концентрацией табуированной лексики при меньшем общем количестве. В «Слове Пацана» преобладают вульгаризмы с умеренной стилистической окраской. Комедийный «Инспектор Гаврилов» продемонстрировал наименьшее использование грубой лексики среди всех анализируемых проектов. (Таб. 5.)

Заключение

Квантитативное исследование лексики российских сериалов выявило взаимосвязь между жанром, структурой текста и его эмоциональной окраской. Анализ четырёх популярных сериалов показал, что детектив «Фишер» отличается максимальным объёмом текста (5000 слов/эпизод) при низкой плотности предложений (16/мин), тогда как комедия «Инспектор Гаврилов» демонстрирует обратную картину (2100 слов, 28 предложений/мин).

Лексический анализ выявил наибольшее разнообразие в триллере «Триггер», в то время как комедия показала неожиданно высокий базовый показатель TTR. Сентимент-анализ зафиксировал жанровую специфику: от негативных

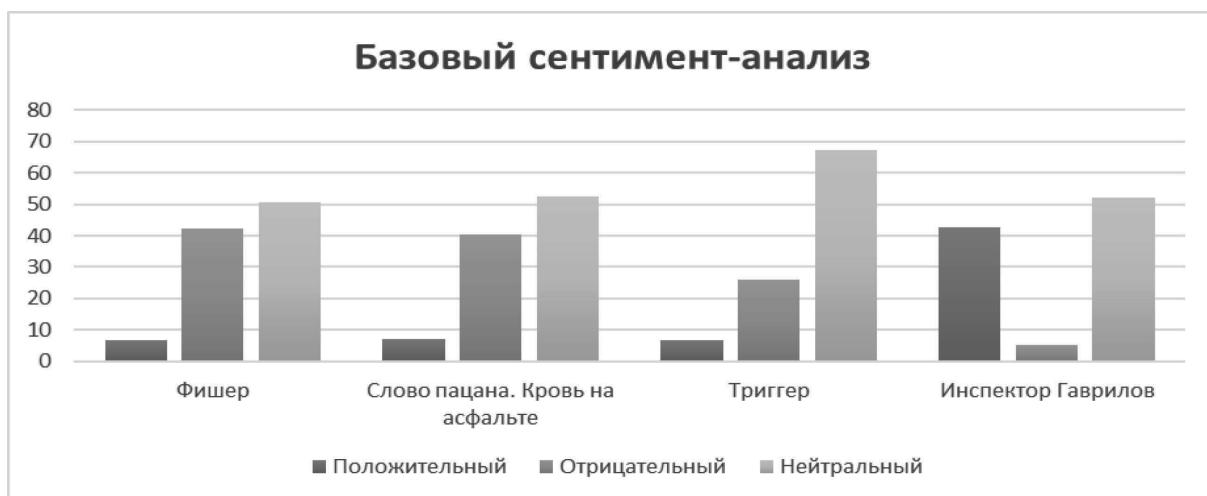


Рис. 1. Сентимент-анализ текстов современных отечественных сериалов

Таблица 4.

Статистический анализ вульгаризмов в российских сериалах.

Категория	Всего	Слово пацана	Триггер	Фишер	Инспектор Гаврилов
Общее количество	403	103	167	102	31
Уникальные типы	44	8	16	14	6
Средняя частота на эпизод	24,4	23,4	37,1	28,3	6,8

Таблица 5.

Классификация вульгаризмов по И. Носеку [12].

Сериал	Слёгкой стилистической окраской	С повышенной грубоостью	Табуированная лексика
Слово пацана	58	33	12
Триггер	79	19	49
Фишер	31	16	55
Инспектор Гаврилов	16	12	0

тонов в драмах до позитивных в комедии, с нейтральным профилем у «Триггера». Все сериалы обладают высокой удобочитаемостью при средней сложности текста (SMOG 8–9), адаптированной под целевую аудиторию.

Перспективы исследования включают расширение корпуса, разработку мультимодальных моделей, диахронический анализ и изучение влияния языковых характеристик на рейтинги.

ЛИТЕРАТУРА

1. Жабин Д.В. Формальные признаки спонтанной речи говорящего в ситуации стресса: дис. ... канд. филол. наук / ВГУ – Воронеж, 2006. – 330 с.
2. Земская Е.А. Русская разговорная речь: лингвистический анализ и проблемы обучения: учеб. пособие. 5-е изд., стер. – М.: ФЛИНТА, 2016. – 241 с.
3. Исаева А.А. Маркеры дифференциации спонтанной и подготовленной звучащей речи // Вестник ВГУ. Серия: Лингвистика и межкультурная коммуникация. – 2020, № 4. – С. 37–48.
4. Козуляев А.В. Интегративная модель обучения аудиовизуальному переводу: дис. ... кан. пед. наук / Москва РУДН – Москва, 2019–234 с.
5. Лумэй В. Понятие и особенности естественной речи современного русского языка // Современное педагогическое образование. – 2023. – №12 // URL: <https://cyberleninka.ru/article/n/ponyatie-i-osobennosti-estestvennoy-rechi-sovremennoogo-russkogo-yazyka> (дата обращения: 22.05.2025)
6. Любаева А.А. Системный анализ средств реализации принципа псевдоустности в аудиовизуальных произведениях на русском языке / А.А. Любаева // Collegium Linguisticum - 2022 : Сборник научных статей Ежегодной конференции Студенческого научного общества МГЛУ, Москва, 16–19 марта 2022 года / Редакционная коллегия: И.А. Гусейнова [и др.]. – Москва: Московский государственный лингвистический университет, 2022. – С. 1047–1055.
7. Новости Кинопроизводства // Telegram Channel // URL: <https://t.me/s/filmres> (дата обращения 21.05.2025)
8. Оборнева И.В. Автоматизация оценки качества восприятия текста // Вестник МГПУ. Серия: Информатика и информатизация образования. – 2005. – №5. // URL: <https://cyberleninka.ru/article/n/avtomatizatsiya-otsenki-kachestva-vospriyatiya-teksta> (дата обращения: 25.05.2025)
9. Россия вышла в лидеры по производству наиболее популярных сериалов для отечественного зрителя // Коммерсантъ // URL: <https://www.kommersant.ru/doc/6011513> (21.05.2025)
10. Стогниева О.Н. Индивидуализация как необходимое условие процесса обучения иностранному языку // Вектор науки Тольяттинского государственного университета. Сер. Педагогика, психология. – 2015. – № 3 (22). – С. 153–158.
11. Тезекбаева Г.А. Спонтанная речь как объект лингвистики // Вестник ВГУ. Серия: Филология. Журналистика. – 2011, – №1. – С. 76–79.
12. Цирихова С.А. Термин «Вульгаризм» в трактовке Эрика Партириджа // Актуальные проблемы филологии и педагогической лингвистики. – 2008. – №10. // URL: <https://cyberleninka.ru/article/n/termin-vulgarizm-v-traktovke-erika-partridzha> (дата обращения: 25.05.2025)
13. Шведова Н.Ю. Очерки по синтаксису русской разговорной речи. – М.: Азбуковник, 2003. – 281 с.
14. Baños-Rocío. Creating credible and natural-sounding dialogue in dubbing: Can it be taught? The Interpreter and Translator Trainer. – 2021. – pp. 13-33
15. Martínez Sierra J.J. Introducción a la traducción audiovisual / J.J. Martínez Sierra. – Murcia: Universidad de Murcia, 2012. – 164 p.
16. Matamala A. Designing a Course on Audio Description and Defining the Main Competences of the Future Professional / A. Matamala, P. Orero // LANS – TTS. – 2007. – Vol. 6. – pp. 329–344.17.

© Васильев Данил Олегович (danielcrow2000@mail.ru), Гурьянов Игорь Олегович (IOGuryanov@yandex.ru).

Журнал «Современная наука: актуальные проблемы теории и практики»