

УДК 81`33

Doi 10.26456/vtfilol/2026.2.161

КОРПУСНЫЙ И НЕЙРОСЕТЕВОЙ ПОРТРЕТ КОНСТРУКЦИИ «КОМУ ВСЕ РАВНО НА ЧТО»¹⁰

Е.Ю. Скороходова¹, Т.М. Обухова^{2,3}

¹Московский физико-технический институт, г. Москва

²Московский государственный университет имени М.В. Ломоносова,
г. Москва

³Государственный институт русского языка им. А.С. Пушкина, г. Москва

Рассматриваются возможности исследования, когда данные корпусной лингвистики проходят верификацию большими языковыми моделями. Для апробации анализируются трансформационные процессы в русском языке на примере изменений выражения «Мне все равно». Возникшее как ошибочное расширение конструкции дополнение *на + вин.п.* со временем получило широкое распространение. Задачей стало определение масштабов такого распространения методами корпусной лингвистики с последующей экспериментальной верификацией большими языковыми моделями (LLM).

Ключевые слова: *корпусные методы, Национальный корпус русского языка, НКРЯ, большие языковые модели, нейросети, грамматическая вариативность.*

Изменения в языке и речи нередко связаны с ошибками носителей языка. При речепорождении случайно или сознательно нарушается норма, и со временем частота определенных нарушений становится такой, что формирует новую норму. Изменение нормы постепенно фиксируется словарями, что требует значительного периода определения и осмысления той или иной новации. Орфоэпические словари фиксируют новации произношения и ударения, толковые словари уточняют значение и грамматические показатели неологизмов и пр. Но современные методики и технологии, в частности корпусные [6; 7; 11] и нейросетевые [12; 13], позволяют проследить за внутренними процессами развития языка и быстрее определить направление возможных изменений.

На протяжении ряда лет считается, что основные изменения в современной речевой практике связаны с влиянием разговорной речи [1; 3: 126; 5, 10]. В целом этот процесс выглядит следующим образом:

¹⁰ Исследования проведены в рамках соглашения с Минобрнауки России от 17.01.2025 № 075-03-2025-662 (шифр FSMG-2025-0086, тема проекта: «Прикладные исследования по внедрению технологий искусственного интеллекта в высшем образовании»).

1. Появление некоей новации (произносительной особенности, слова, фразеологизма, грамматической конструкции) и ее использование в пределах разговорной речи, где допустимо отступление от нормы. Новация может возникнуть также как результат языковой игры.
2. Использование данной новации в социальных сетях и текстах СМИ – сначала постепенное, с оговорками и уточнениями («как принято сейчас говорить»; «можно сказать» и пр.) [4]. Современные СМИ довольно быстро реагируют на любые новшества, особенно если они способны привлечь внимание аудитории.
3. Постепенная легитимизация новшества за счет его повторяемости; в перспективе – фиксация словарями.

Таким же образом легитимизируются и те единицы, которые изначально расценивались как нарушения нормы. Это приводит к постепенному изменению нормы или формированию нескольких вариантов нормы на всех уровнях языковой системы. Эти процессы длятся десятилетиями.

Как же разворачиваются данные процессы сегодня? Рассмотрим грамматическую и лексическую трансформацию выражения *всё равно*. Например, «Словарь живого великорусского языка» В. И. Даля (1882) приводит следующие выражения: *Коли все равно, так и делай по-моему; Бывает добро, да не всякому равно* [2: 6]. «Словарь русского языка» под ред. А. П. Евгеньевой (1985) указывает: все равно – 1) безразлично, одинаково; 2) несмотря ни на что; 3) с союзами «как, будто» в качестве сравнения [8: 157]. В иллюстративном примере – *[Петрушке] было совершенно все равно...* уже фиксируется датив. «Фразеологический словарь современного русского литературного языка» под ред. А. Н. Тихонова (2004) практически повторяет статью из словаря 1985 года; один из примеров повторяется, во втором – *Вам, товарищ, все равно* – обозначен дативный субъект [9: 118].

Первоначально это простое двусоставное предложение с косвенным дополнением (*мне, обществу, человеку*) и возможным расширением через придаточное изъяснительное: *Ему все равно, что будет с этим домом.*

Грамматически эта конструкция схожа с безличными предложениями типа *Мне плохо*, лексически – с вульгарно-просторечным выражением *мне наплевать на... + вин.п.*. Сходство это подкрепляется тем фактом, что в последнее время выражение *мне всё равно* нередко используется с предлогом *на* и дополнением в винительном падеже. Механизм объединения данных выражений представляется следующим: привыкнув в разговорной речи использовать вариант *мне наплевать на...*, при смене коммуникативной ситуации говорящий вынужден заменить его более приемлемым, используя конструкцию *мне всё равно*. Однако в процессе спонтанной речи далеко не всегда удается изменить и грамматический строй высказывания, поэтому сохраняется управление

на + вин.п. Постепенно за счет частотности конструкция *мне всё равно на + вин.п.* начинает рассматриваться как допустимый вариант для использования в ситуации устного общения, даже официального.

Анализ источников показывает, что эта ошибка не нова: первое её появление в публичном поле (в НКРЯ) относится к 1936 году: *Тогда было заметно, как ему все равно на случайных собутельников, и на то, что он сам кричал наемни об искусстве, музыке, России. [И. С. Лукаш. Вьюга. От автора (1936)].*

Целью настоящего исследования является комплексное описание конструкции *кому всё равно на что*, выступающей в функции выражения эмоционально-окрашенного безразличия, и описание её системных характеристик в современном русском языке на основе корпусных данных с последующей верификацией при помощи больших языковых моделей.

Для достижения данной цели в работе решаются следующие задачи:

1. установить и формализовать актантную структуру конструкции;
2. охарактеризовать семантические классы ее обязательных компонентов – дативного субъекта и объекта-стимула;
3. описать выявленные закономерности на материале Национального корпуса русского языка;
4. проверить, используют ли данную конструкцию большие языковые модели (GPT-5, Perplexity, DeepSeek, Grok, Qwen, GigaChat, YandexGPT).

Материалом для исследования послужила выборка из 164 контекстов из Национального корпуса русского языка. Мы остановили свой выбор на газетном подкорпусе и подкорпусе социальных сетей. Во-первых, социальные сети представляют собой пространство максимально естественной, неотредактированной письменной речи, обладающей характеристиками спонтанности и аутентичности. В этой коммуникативной среде стигматизированные и разговорные конструкции, такие как *плевать*, проявляются в своём наиболее органичном виде. Во-вторых, важным преимуществом является прагматический контекст, позволяющий анализировать не только грамматическую структуру, но и коммуникативные интенции, стоящие за выбором конструкции. Наконец, вопрос частотности имеет особую значимость в этих подкорпусах – именно в них конструкция представлена шире всего (так, в устном подкорпусе нашлось всего 2 примера¹¹, в основном – 23¹², в мультимедийном – 1¹³, релевантных из которых в общей сложности осталось 4).

¹¹ <https://ruscorpora.ru/s/316yx>

¹² <https://ruscorpora.ru/s/2R87N>

¹³ <https://ruscorpora.ru/s/nZRLD>

Опишем структуру и параметры поискового запроса. Поиск был основан на строгой последовательности четырех ключевых элементов: *слово I + всё + равно + на*.

Для слова *I* был задан грамматический признак дательного падежа. Поиск ограничивался теми частями речи из предложенных в НКРЯ, которые способны субстантивироваться в дативе – существительное, местоимение-существительное, прилагательное, числительное, числительное-прилагательное, местоименное прилагательное.

Для компонентов *всё*, *равно* и *на* поиск осуществлялся по конкретным словоформам для точного попадания в предикативное ядро конструкции.

Ключевым методологическим решением являлось сознательное неприменение грамматического ограничения на словоформу, следующую за предлогом *на*. Это позволило количественно оценить соотношение релевантных (с винительным падежом) и нерелевантных контекстов в общей выборке, а также выявить и описать потенциальные синтаксические вариации и случаи синтаксической омонимии, обеспечив полноту лингвистического описания, зафиксировав все контексты, где формально присутствует исследуемая последовательность элементов.

Для изучения результатов поиска использовался метод сплошной выборки. Каждый контекст ($n = 164$) подвергался квалификационному анализу для выявления релевантных случаев реализации искомой синтаксической конструкции. Из общего массива данных в газетном подкорпусе ($n = 80$) было отобрано 11 релевантных контекстов (13,75% от общего числа)¹⁴, в подкорпусе социальных сетей ($n = 84$) – 59 (70,24% от общего числа)¹⁵. Критерием релевантности выступало наличие целостной актантной структуры «датель + всё равно + на + винительный падеж», где все компоненты находятся в непроницаемом синтаксическом единстве; остальные примеры были исключены из дальнейшего анализа.

Для верификации продуктивности исследуемой синтаксической модели была сформирована репрезентативная выборка больших языковых моделей (LLM), в которую вошли как технологические лидеры мирового уровня (GPT-5, Perplexity, Deep Seek, Grok, Qwen), так и ведущие русскоязычные разработки (GigaChat и YandexGPT).

Промт для всех моделей был сформулирован следующим образом: *Придумай 50 предложений с оборотом «кому + всё равно»*. Такой запрос позволил минимизировать влияние исследователя на результат, поскольку не содержал прямого указания на предложное управление с *на*. Это дало возможность проверить не явные знания модели о правилах,

¹⁴ Для центральных СМИ – <https://ruscorporation.ru/s/486zn>, для региональных СМИ – <https://ruscorporation.ru/s/587Rv>

¹⁵ <https://ruscorporation.ru/s/66WB1>

а её имплицитное «чувство языка» и ассоциативные связи, связывающие предикат «всё равно» с управлением винительного падежа. Также для обеспечения чистоты эксперимента использовался новый сеанс чата без истории сообщений, что исключило влияние контекста предшествующего обсуждения.

Результаты корпусного анализа позволили выявить основные закономерности употребления конструкции *кому всё равно на что*.

Распределение по жанровому разнообразию представлено в рисунке 1.

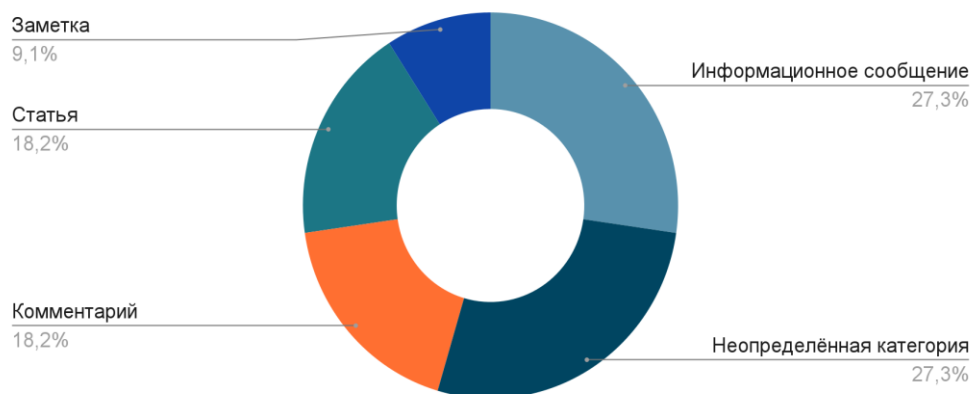


Рис. 1. Жанровая принадлежность контекстов с конструкцией *кому всё равно на что* в газетном подкорпусе НКРЯ

Как мы видим, конструкция наиболее характерна для оперативных жанров (информационное сообщение, комментарий), что согласуется с её разговорной и оценочной природой, позволяющей передать непосредственную реакцию говорящего.

Анализ тематической принадлежности контекстов демонстрирует широкую сочетаемость исследуемой конструкции с различными сферами. Наиболее репрезентативной областью оказалась *политика и общественная жизнь* (3 контекста), где конструкция используется для выражения позиции как персональных, так и институциональных акторов («*Не секрет, что России все равно на собственные подписи на международных актах...*» — сказала эстонский лидер. [Президент Эстонии заявила, что Россия является «самым трудным партнером» // *Московский комсомолец*, 21.11.2019]). Часть употреблений (2 контекста) относится к неопределенной категории, что может говорить о переходе конструкции в разряд общеупотребительных средств, не связанных со специализированными регистрами. При этом конструкция проявляет значительную семантическую гибкость, встречаясь в таких разнородных тематических зонах, как *администрация и управление; спорт; досуг, зрелища и развлечения; здоровье и медицина; бизнес, коммерция,*

экономика, финансы; искусство и культура (по 1 контексту). Это распределение подтверждает способность анализируемой модели функционировать в любом тематическом поле современного публицистического дискурса.

В найденных контекстах субъект, занимающий дативную позицию, представлен тремя основными группами: личные местоимения (*мне, им*), имена собственные (*Вернблуму*), обобщенно-личные и именные группы (*России, такому человеку*).

Объекты-стимулы, вводимые предлогом *на*, характеризуются морфологическим и семантическим разнообразием. По количеству слов в составе группы они варьируют от односоставных (*штрафы, возраст, него, пациенты*) до многокомпонентных именных групп (*свой внешний вид; собственные подписи; снег, грязь, дождь*).

Семантически объекты распределяются между несколькими устойчивыми типами: социальные и правовые нормы, оценочные и социальные характеристики, физические условия и природные явления, поведенческие характеристики, а также конкретные лица. Таким образом, конструкция демонстрирует высокую полисемичность объекта.

Корпусный анализ подкорпуса социальных сетей показывает, что конструкция *кому всё равно на что* широко функционирует в неформальном онлайн-дискурсе, охватывая каналы Telegram, сообщества «ВКонтакте» и личные блоги. Для этих жанров характерны разговорность, эмоциональность и ситуативность, что способствует активному использованию модели в спонтанных высказываниях.

Субъект дативной позиции в большинстве случаев выражен личными местоимениями (*мне, ему, им, тебе*), что соответствует межличностному характеру коммуникации. Встречаются и коллективные или институциональные акторы (*Сберу, майкам, областному правительству*), а также обобщенные группы (*тем, кому всё равно*). Объекты-стимулы демонстрируют семантическую и морфологическую вариативность: от конкретных лиц и материальных объектов до абстрактных понятий (*мнение, обстоятельства, безопасность, цена, происходящее*). Такая неоднородность подтверждает отсутствие тематических ограничений на валентную позицию объекта.

Тематическая зона употреблений при этом значительно шире, чем в газетном подкорпусе. Помимо общественных и социальных вопросов, представлены бытовые ситуации, отношения, обсуждения технологий, гейминга, контента, животных и морали.

Если говорить о позиции конструкции *кому всё равно на что* в предложении, то в обоих подкорпусах чаще всего она стоит либо в начале основной фразы, либо во встроеной части предложения, но никогда не завершает предложение.

Анализ хронологического распределения (см рис. 2) выявляет устойчивую динамику роста частотности целевой конструкции, при этом

активизация ее употребления в социальных сетях происходит несколько позже, чем в газетном дискурсе.

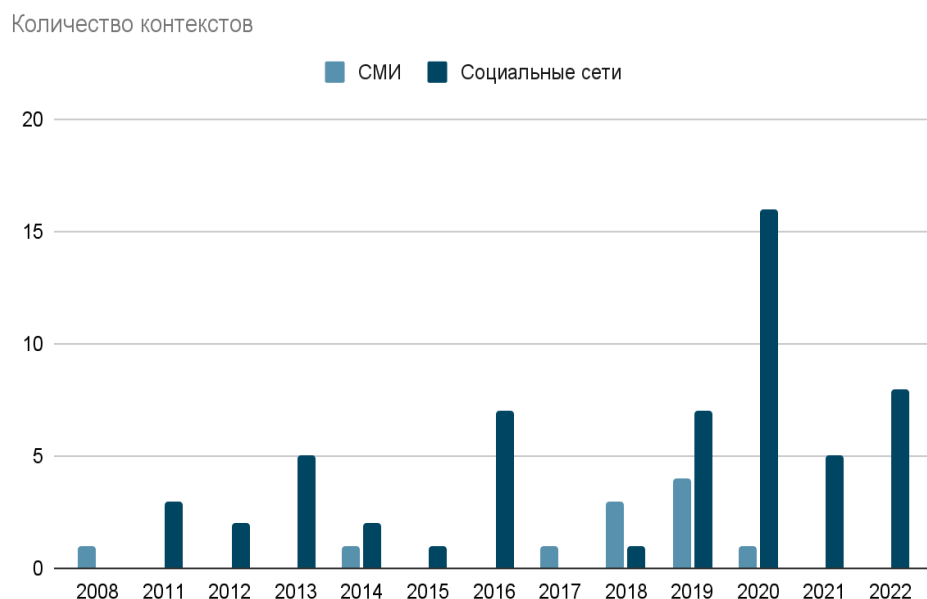


Рис. 2. Динамика употребления конструкции *кому всё равно на что* по годам в подкорпусе социальных сетей и в газетных подкорпусах НКРЯ

Исследование на предложенном наборе данных имеет ряд ограничений. Во-первых, подкорпуса НКРЯ пополняются неравномерно, поэтому в выборку не попадают самые свежие примеры; кроме того, среди релевантных контекстов встречаются формально одинаковые данные с разными датами. Во-вторых, в работу не был включён Генеральный Интернет-корпус русского языка (ГИКРЯ)¹⁶, который потенциально дал бы больше примеров сетевого употребления. В-третьих, нами были просмотрены корпус «Один речевой день» (ОРД)¹⁷, Корпус русской устной речи¹⁸ и другие корпуса звучащей речи¹⁹, однако релевантных употреблений конструкции там не обнаружилось. Это не столько ограничение, сколько указание на перспективу: дальнейший анализ более широких аудиокорпусов или сбор собственных данных позволил бы проверить, насколько модель распространена в устном общении.

¹⁶ <http://www.webcorpora.ru/>

¹⁷ <https://ord.spbu.ru/>

¹⁸ <https://russpeech.spbu.ru/search/trn-search.php?wf=%E2%F1%B8+%F0%E0%E2%ED%EE>

¹⁹ <https://spokencorpora.ru/>

Для оценки статуса конструкции в современном языковом сознании был проведён эксперимент с использованием генеративных языковых моделей. Результаты представлены в Таблице 1.

Таблица 1

Способность языковых моделей к генерации контекстов с конструкцией *кому всё равно на что*

языковая модель	релевантные контексты (все найденные) ²⁰	типичная структура предложений для нерелевантных контекстов
GPT-5	1. Кому всё равно на сроки, тот не торопится.	1. Кому теперь всё равно, тот и спорить не будет. 2. Мне всё равно, кому задавать вопросы.
Perplexity	–	1. Кому всё равно, как учиться, тот редко добивается успеха. 2. Ей всё равно, что на неё смотрят.
DeepSeek	1. Она сказала, что ей всё равно на мое мнение. 2. После стольких лет предательства мне стало на них всё равно. 3. После его слов мне стало на неё всё равно. 4. Она поняла, что ей всё равно на его богатство и статус.	1. Пусть я буду неправ, мне всё равно, я всё равно выскажу своё мнение. 2. Ему всё равно, как ты это объяснишь коллегам. 3. Он бросил папку на стол и сказал: «Мне всё равно, решайте сами».
Grok	–	1. Влюблённым всё равно, пойдут они в кино или останутся дома. 2. Честно говоря, тебе всё равно, кто победит в этом споре.
Qwen	–	1. Энрике всё равно, пойдёшь ты со мной или останешься дома. 2. Ему всё равно, что о нём думают — он знает, кто он есть.

²⁰ Все примеры представлены в оригинальном виде, полученные непосредственно от языковых моделей. Орфография и пунктуация сохранены без изменений.

GigaChat	–	1. Мне всё равно, куда мы пойдем гулять вечером. 2. Какое платье надеть на праздник? Всем всё равно!
YandexGPT	–	1. Мне всё равно, что подумают соседи о нашем ремонте. 2. Им всё равно, сколько это будет стоить — главное, чтобы получилось красиво.

Мы видим, что из семи протестированных моделей только две продемонстрировали способность к генерации релевантных контекстов – DeepSeek (4 примера) и GPT-5 (1 пример). Большинство моделей демонстрируют тенденцию к использованию шаблонных структур, преимущественно с местоимениями в дательном падеже и придаточными предложениями с союзами. Лишь Deep Seek и Grok показали способность к варьированию синтаксических структур, использовали не только местоимения, но и существительные в дательной позиции. Несмотря на изначальную гипотезу о потенциальном преимуществе русскоязычных разработок (GigaChat, YandexGPT), ни одна из них не предложила ни одного релевантного контекста, ограничиваясь конструкциями с союзным присоединением. Это значит, что современные продвинутые языковые модели обладают различной степенью усвоения периферийных синтаксических конструкций, причем языковая специализация не гарантирует лучших результатов в области синтаксической генерации.

Таким образом, нейросети, заранее обученные на большом количестве текстов, уже допускают использование выражения *кому всё равно на что* как возможное и потенциально правильное.

Сочетание методов корпусной лингвистики и возможностей искусственного интеллекта оказалось продуктивным для решения определенных задач, что открывает широкие перспективы для научных исследований. Современные корпусные системы позволяют создавать и совершенствовать языковые модели, а применение нейросетей обеспечивает масштаб и скорость анализа. Сочетание работы с корпусами и с большими языковыми моделями позволяет взаимно верифицировать результаты исследований.

Результаты таких научных исследований могут быть востребованы в теории перевода, сравнительном языкознании, стилистике. Определение частотности и воспроизводимости различных речевых единиц позволяет не только составить наиболее полную картину особенностей современной речевой практики, но и оперативно актуализировать учебные программы.

Список литературы

1. Валгина Н.С. Пути вхождения разговорных конструкций в письменную речь // Филологические науки. 1997. № 6. С. 22–26.
2. Даль В.И. Толковый словарь живого великорусского языка: в 4 т. М.: Русский язык, 1991. Т. 4. С. 6.
3. Дегальцева А. В., Сиротинина О. Б. Отношение носителей русского языка к его нормам: динамический аспект // Изв. Саратов. ун-та Нов. сер. Сер. Филология. Журналистика. 2024. №2. С. 124–130. URL: <https://cyberleninka.ru/article/n/otnoshenie-nositeley-russkogo-yazyka-k-ego-normam-dinamicheskiiy-aspekt> (дата обращения: 25.11.2025). DOI: <https://doi.org/10.18500/1817-7115-2024-24-2-124-130>
4. Завадская Ю.О. Оговорки как явление устной спонтанной речи: основные аспекты анализа // Коммуникативные исследования. 2018. №2 (16). URL: <https://cyberleninka.ru/article/n/ogovorki-kak-yavlenie-ustnoy-spontanoy-rechi-osnovnye-aspekty-analiza> (дата обращения: 30.11.2025). DOI: 10.25513/2413-6182.2018.2.37-51
5. Лейчик В.М. Тенденции интеллектуализации и демократизации в современном русском языке и их выражение в публицистических текстах // Актуальные проблемы лингвистики и терминоведения: международный сборник научных трудов, посв. юбилею проф. З. И. Комаровой. Екатеринбург: УГПУ, ИЯ РАН, 2007. С.133–148.
6. Плунгян В.А. Зачем нужен Национальный корпус русского языка? Неформальное введение // Национальный корпус русского языка: 2003–2005. М.: Индрик, 2005. URL: <https://ruscorpora.ru/media/uploads/2022/04/21/r32.pdf> (дата обращения: 30.11.2025).
7. Плунгян В.А. Национальный корпус русского языка в лингвистических исследованиях // Integrum: точные методы и гуманитарные науки. М., 2006, С. 76–84.
8. Словарь русского языка: в четырех томах. Под ред. А. П. Евгеньевой. М.: Русский язык. 1985. Том 1. С. 157.
9. Фразеологический словарь современного русского языка. Под ред. А. Н. Тихонова. М.: Флинта-Наука, 2004. Том 1. С.118.
10. Шамылова О. Синтаксические особенности разговорного русского: норма и отклонения // Образование и наука в XXI веке. 2025. Т. 1. № 62-4. URL: <https://mpcareer-google.ru/index.php/journal/article/view/2547> (дата обращения: 25.11.2025).
11. Шмелёв А.Д. Языковые факты и корпусные данные // Русский язык в научном освещении. 2010. № 19(1). С. 236–265.
12. Torrent T.T., Hoffmann T., Almeida A.L., Turner M. Copilots for Linguists: AI, Constructions, and Frames // Cambridge University Press. 2024. DOI: <https://doi.org/10.1017/9781009439190>
13. Uchida S. Using early LLMs for corpus linguistics: Examining ChatGPT's potential and limitations // Applied Corpus Linguistics. 2024. Vol. 4. No. 1. P. 100089. DOI: <https://doi.org/10.1016/j.acorp.2024.100089>

Об авторах:

СКОРОХОДОВА Елена Юрьевна – доктор филологических наук, доцент, профессор Департамента иностранных языков Московского физико-технического института (141701, Московская обл., г. Долгопрудный, Институтский пер., дом 9); e-mail: aentar@gmail.com

ОБУХОВА Татьяна Михайловна – кандидат педагогических наук, старший преподаватель Высшей школы перевода (факультета) Московского государственного университета имени М.В. Ломоносова (119991, г. Москва, Ленинские горы, д. 1), ведущий эксперт лаборатории когнитивных и лингвистических исследований Государственного института русского языка им. А.С. Пушкина (117485, г. Москва, ул. Академика Волгина, д. 6); e-mail: t.m.obukhova@gmail.com

A CORPUS-BASED AND NEURAL PORTRAIT OF THE CONSTRUCTION “КОМУ ВСЁ РАВНО НА ЧТО”

E.Yu. Skorokhodova¹, T.M. Obukhova^{2,3}

¹Moscow Institute of Physics and Technology, Moscow

²Lomonosov Moscow State University, Moscow;

³Pushkin State Russian Language Institute, Moscow

This article presents a methodology for studying language change that combines corpus analysis with experimental verification using large language models (LLMs). Using the spread of the construction "Мне все равно на + accusative case" as a case study, it demonstrates how frequency data from corpora is supplemented by assessing the naturalness and generation probability of variants via LLMs. The results confirm that LLMs effectively capture current norms and variation, proving to be a powerful tool for diachronic linguistics.

Keywords: *corpus methods, Russian National Corpus, RNC, large language models, neural networks, grammatical variation.*

About authors:

SKOROKHODOVA Elena Yurievna – Doctor of Philology, Professor of the Department of Foreign Languages, Moscow Institute of Physics and Technology, MIPT (Institutskiy all., 9, Moscow reg. Dolgoprudniy city, 141701), e-mail: aentar@gmail.com

ОБУКHOVA Tatiana Mikhailovna – Senior Lecturer of Higher School of Translation and Interpreting, Lomonosov Moscow State University (Leninskiye Gory, 1, Moscow, 119991); Leading Expert of Language and Cognition Laboratory, Pushkin State Russian Language Institute (Academic Volgin Str., 6, Moscow, 117485); e-mail: t.m.obukhova@gmail.com

Статья поступила в редакцию 21.03.26
Подписана в печать 10.04.26

© Скороходова Е.Ю.,
Обухова Т.М., 2026