



оригинальная статья

<https://elibrary.ru/iygjua>

Исследование сложности, понятности и переводимости русского текста традиционными методами и методом обратного машинного перевода

Мельникова Виктория Сергеевна

Кемеровский государственный университет, Россия, Кемерово

eLibrary Author SPIN: 9334-8632

<https://orcid.org/0000-0002-2480-1967>

melnikovavika2017@yandex.ru

Аннотация: Контекст исследования заключается в растущей потребности в эффективных методах оценки текстов в аспекте их сложности, понятности и переводимости. Один из таких методов – обратный машинный перевод, позволяющий выявить потенциальные точки напряженности в интерпретации текста. Теоретической основой статьи выступает гипотеза, согласно которой переводимость и понятность являются соположенными лингвокогнитивными понятиями, которые могут взаимно корректировать друг друга в текстовых и лексических измерениях текстов в целом и их отдельных компонентов. Подтверждение данной гипотезы осуществляется в работе путем сопоставления прямых апелляций к языковому и метаязыковому сознанию носителей языка, выявленных экспериментальным способом (опросом). Важное место в названном сравнении занимает исследование Э. С. Клышинского, который осуществил лингвокогнитивный эксперимент на понимаемость текстов на славянских языках носителями русского языка. Научная новизна исследования заключается в разработке алгоритмов использования цифрового метода, включающего обратный машинный перевод и программы сравнения текстов на схожесть. Это обстоятельство является основанием для использования обратного машинного перевода текста для его измерения и оценки с точки зрения простоты / сложности, ясности, доступности. Машинный перевод является одним из способов реализации цифровой методики обработки текстов. Цель – доказать, что цифровые методы способны выступать в роли инструментов для оценки сложности, понятности и переводимости текстов. Помимо обратного машинного перевода используются программы сравнения текстов на схожесть. В работе его результаты верифицируются результатами опроса-эксперимента. Такая логика исследования позволяет сопоставить две модели понимания: одну, существующую в рамках естественного языкового сознания, и другую, функционирующую в контексте искусственного интеллекта. Материалом послужило стихотворение А. С. Пушкина «Клеветникам России». Результаты показали, что тексты с высокой сложностью восприятия и понимания становятся менее понятными при обратном машинном переводе. Полученные данные позволяют сделать вывод о прямо пропорциональной соотнесенности понятности и переводимости. Результаты опроса-эксперимента совпали с экспериментом, построенным на использовании метода обратного машинного перевода: респонденты не поняли те же слова / словосочетания во фрагментах стихотворения, что и искусственный интеллект. Выводы исследования подтверждают возможность использования обратного машинного перевода для выявления помех в понимании текстов среднестатистическим носителем языка.

Ключевые слова: сложность текста, транслятивная лингвистика, переводимость, обратный машинный перевод, сложность, понятность

Цитирование: Мельникова В. С. Исследование сложности, понятности и переводимости русского текста традиционными методами и методом обратного машинного перевода. *Виртуальная коммуникация и социальные сети*. 2025. Т. 4. № 1. С. 23–33. <https://doi.org/10.21603/2782-4799-2025-4-1-23-33>

Поступила в редакцию 20.12.2024. Принята после рецензирования 15.01.2025. Принята в печать 20.01.2025.

full article

Complexity, Comprehensibility, and Translatability from Russian: Traditional vs. Digital Methods

Victoria S. Melnikova

Kemerovo State University, Russia, Kemerovo

eLibrary Author SPIN: 9334-8632

<https://orcid.org/0000-0002-2480-1967>

melnikovavika2017@yandex.ru

Abstract: Reverse machine translation is an effective method of text evaluation in terms of complexity, comprehensibility, and translatability. It highlights potential translation difficulties at the pre-translation stage. Translatability and comprehensibility are juxtaposed linguacognitive concepts that can correct each other in the textual and lexical dimensions. Using E. S. Klyshinsky's linguacognitive experiment on the comprehensibility of Slavic tests by native Russian speakers, the authors checked this hypothesis using the survey method to appeal to the linguistic and metalinguistic consciousness of native speakers. They developed a new algorithm for comparing texts by reverse machine translation based on points of simplicity, complexity, clarity, and accessibility. As a digital text processing method, machine translation proved able to evaluate textual complexity, comprehensibility, and translatability. The results obtained with the new algorithm were verified with a survey, which made it possible to compare two models of understanding, i.e., that of human linguistic consciousness and that of the artificial intelligence. Alexander Pushkin's *To the Slanderers of Russia*, a poetic text of high complexity, became less understandable after reverse machine translation. Comprehensibility and translatability appeared to be directly proportional. The survey yielded the same results as the reverse machine translation: the respondents and the artificial intelligence failed to understand the same words and phrases. The method of reverse machine translation proved effective in detecting textual comprehension challenges for native speakers.

Keywords: text complexity, translation linguistics, translatability, reverse machine translation, complexity, comprehensibility

Citation: Melnikova V. S. Complexity, Comprehensibility, and Translatability from Russian: Traditional vs. Digital Methods. *Virtual Communication and Social Networks*, 2025, 4(1): 23–33. (In Russ.) <https://doi.org/10.21603/2782-4799-2025-4-1-23-33>

Received 20 Dec 2024. Accepted after review 15 Jan 2025. Accepted for publication 20 Jan 2025.

Введение

Сложность и понятность текста являются важными аспектами в различных областях знаний. М. Д. Бабкина отмечает важность использования *ясного языка* для обеспечения доступности информации [Бабкина 2021]. Как указывает Л. В. Катыхева, тренд простого и понятного языка становится все более актуальным в России [Катыхева 2022: 92]. Таким образом, это подтверждает значительный интерес к изучению простого и понятного языка. Л. В. Щерба также подчеркивал значимость понятности речи, утверждая, что любое высказывание, которое трудно понять, не достигает своей цели и становится «отрицательным языковым материалом» [Щерба 1974: 39]. Идея использования обратного машинного перевода для исследования языка была представлена Н. Д. Голевым в статье

«Источниковый потенциал обратного машинного перевода» [Голев 2018]. Особое внимание исследователь уделяет различным видам перевода, включая обратный, машинный и обратный машинный перевод [Голев 2022]. Современная эпоха характеризуется стремительным развитием информационных технологий, которые оказывают значительное влияние на различные сферы человеческой деятельности, в том числе лингвистическую науку. Одним из ключевых аспектов этого процесса является внедрение искусственного интеллекта (ИИ), который становится незаменимым инструментом для филологов и лингвистов. Развитие слабого и сильного ИИ привело к появлению новых возможностей для анализа текстовой информации, включая художественные произведения [Колмогорова, Залевская 2023: 22].

Объект исследования – русский текст, подвергнутый обратному машинному переводу, с точки зрения его сложности, понятности и переводимости; предмет – транслятивное функционирование текста, опосредованное методом обратного машинного перевода. Цель – доказать, что цифровые методы способны выступать в роли инструментов для оценки сложности, понятности и переводимости текстов. Задача исследования заключается в разработке и обосновании метода обратного машинного перевода как инструмента для комплексной оценки сложности, понятности и переводимости текстов, с учетом того, что эти параметры относятся к разным аспектам восприятия и обработки текста, но при этом находятся в отношениях взаимозависимости.

Традиционные подходы к оценке текстов не всегда адекватно отображают действительную степень сложности их восприятия. Методы компьютерных технологий, в ряду которых находится обратный машинный перевод, позволяет предложить новый подход к этой проблеме [Мельникова 2021; 2024]. Было установлено, что с помощью метода обратного машинного перевода можно не только оценивать сложность и понятность текста, но и выявлять потенциальные проблемы в аспекте доступности его восприятия [Мельникова 2021]. В настоящей статье на первый план выдвигается проблема соотносимости разных форматов квалификации простоты / сложности и понятности текста с параметром его переводимости на разные мировые языки, в результате этого актуализируется такая лингвокогнитивная категория единиц языка, как переводимость.

В нашем исследовании сложность текста и переводимость коррелируют. За последние десятилетия благодаря теории текста был разработан обширный набор категорий, в терминах которых можно описать любой текст [Kolmogorova, Margolina 2024]. Необходимо строго различать и интерпретировать понятия лингвистической сложности текста, которая определяется объективно [Ibid.], и трудности текста, которая определяется субъективно [Солнышкина 2023: 425]. Также следует учитывать категорию понятности, связанную с языковым контактированием, и понятие переводимости.

Переводимость текста зависит от внутренней ценности текста [Benjamin 1972: 11]. В рамках исследования выдвигается гипотеза о зависимости переводимости от понятности. При таком подходе уровень переводимости языковой единицы может

служить индикатором ее понятности. Данная корреляция открывает новые горизонты для анализа перевода, позволяя рассматривать перевод как процесс, в котором важную роль играют когнитивные аспекты восприятия. Полагаем, что обоснование и / или опровержение выдвинутой гипотезы о соотношении переводимости и понятности является теоретически значимым для решения общих вопросов лингвокогнитологии. Научная новизна исследования заключается в разработке алгоритмов использования цифрового метода, включающего обратный машинный перевод и программы сравнения текстов на схожесть. Мы предлагаем новый систематизированный подход к оценке сложности и понятности текстов, основанный на анализе результатов обратного машинного перевода. Метод позволяет учитывать как лексическую, так и синтаксическую сложность текста, представляя собой значительное расширение существующих методик. В отличие от предыдущих исследований, наш подход фокусируется на выявлении сложностей, возникающих при обратном машинном переводе. Это дает возможность глубже понять причины недопонимания текста как человеком, так и ИИ.

Актуальность исследования связана с проблемой соотношения лингвокогнитивных категорий понятности и переводимости, а также с активным вхождением в повседневную жизнь ИИ, с помощью которого мы проверяем нашу гипотезу. С развитием технологий ИИ и нейронных сетей возможности применения обратного машинного перевода значительно расширились. Практическая значимость статьи: предложен и описан метод, позволяющий определить уровень сложности и понятности текста с целью дальнейшего повышения его качества. Этот метод обратного машинного перевода может применяться в различных областях, таких как корректура, переводческая практика, образование, юрислингвистика, разработка программного обеспечения. Таким образом, метод может быть применен в исследованиях, связывающих лингвистику и компьютерные технологии.

Данная работа тесно связана с активно развивающимся направлением в языкознании – когнитивной лингвистикой, обладающей определенными теоретико-методологическими установками [Болдырев 2020: 6]. Е. М. Масленникова рассматривала динамический компонент переводческой деятельности, призывая к развитию теории перевода как таковой в более широком контексте. Она отмечала, что переводчик не ограничивается простой

передачей сообщения с одного языка на другой, но также интерпретирует его в соответствии со своими собственными представлениями и культурным контекстом. Интерпретация в процессе перевода обусловлена рядом факторов: языковыми и культурными различиями между исходным и целевым языком; индивидуальным пониманием текста переводчиком; эстетическими и этическими соображениями [Масленникова 2007: 53].

Изучение интерпретирующей функции переводческого понимания расширяет знания о когнитивных механизмах перевода, способствующих гармонизации смыслов в переводческом пространстве [Кушнина и др. 2019: 55]. Представленный тезис находит конкретное воплощение в исследованиях, посвященных анализу понятности переведенных русскоязычных текстов и отдельных слов носителями русского языка [Клышинский 2021: 76]. Э. С. Клышинский стремится измерить понятность текстов на основе опросов, которые, на наш взгляд, представляют собой формат стихийной переводимости. Такой подход позволяет перевести абстрактное понятие понятности в конкретные показатели, доступные для измерения. Исследование интерпретативных процессов в переводе остается актуальным в рамках современной когнитивной парадигмы. Переводчик, взаимодействуя с когнитивными пространствами различных языковых личностей, создает интерпретативное понимание, порождающее новые смыслы. Таким образом, интерпретация в авторизованном переводе представляет собой сложный и многоаспектный процесс, подверженный влиянию множества факторов [Раздобудько-Чович 2012: 76]. Итак, в рассмотренных выше работах акцентируется основная идея нашей статьи: перевод эквивалентен пониманию (интерпретации).

Методы и материалы

Основным методом настоящего исследования выступает метод обратного машинного перевода¹. Инструмент исследования – переводческий сервис Google для машинного и обратного машинного перевода. Исходный текст выбирается в зависимости от его жанра. Для анализа было выбрано стихотворение А. С. Пушкина «Клеветникам России»².

Являясь «сильными» текстами культуры, знаковые художественные тексты обладают неоднозначностью эстетической информации и высокой степенью интерпретативности, что обеспечивает генерирование вторичных текстов различной семиотической природы. С. В. Микони пишет: «Связь между образами и словами национальных языков неоднозначна. Это порождает проблему понимания иностранных слов и слов, обозначающих новые образы» [Микони 2021: 66].

Текст переводится на другой язык с использованием машинного перевода (в данном эксперименте было взято 10 языков (табл. 1)). Всего проанализировано 10 переводов. Стихотворение было поделено на 6 фрагментов по принципу выделения строф в стихотворении. Выделение одной строфы из контекста стихотворения считаем возможным, поскольку мы хотим провести детальный анализ каждого фрагмента произведения с целью сосредоточения на конкретных элементах текста (лексика, структура). Стихотворение представляет собой целостное произведение, где каждый элемент связан с другими и работает на общий замысел автора, поэтому результаты экспериментов с отдельными строфами мы интерпретируем без учета общего контекста, а так, как показал результат обратного машинного перевода. Переведенный текст возвращается на исходный русский язык. Сравниваются исходный и обратно переведенный тексты для выявления трансформаций. Метод обратного машинного перевода использовался в связке с программой сравнения текстов на схожесть³ и когнитивным опросом-экспериментом людей, говорящих на русском языке. Сервис автоматического сравнения текстов применялся для количественной оценки различия между исходным текстом и его машинным переводом. Этот инструмент позволил получить объективные данные о степени изменения смысла текста после перевода.

Цель опросов – выявить связи между восприятием текста респондентами и их оценками сложности и переводимости текста. Исследование направлено на изучение влияния различных лингвистических и когнитивных факторов на понимание и интерпретацию текста. Эксперимент состоял из двух этапов:

¹ В настоящем исследовании мы не учитывали особенности алгоритмов программ машинного перевода, проявляющиеся на фоне перевода с помощью нейронного метода, т. к. считаем это отдельной задачей.

² Пушкин А. С. Клеветникам России. URL: <https://www.culture.ru/poems/4966/klevetnikam-rossii> (дата обращения: 20.11.2024).

³ Программа сравнения текстов на схожесть. URL: <http://text.num2word.ru> (дата обращения: 20.11.2024).

Этап 1: реализуется холистический подход. Респондентов просили сначала прочесть стихотворение А. С. Пушкина «Клеветникам России» целиком, а затем интуитивно оценить фрагменты по шкале понятности / непонятности. Были разработаны вопросы, направленные на выявление отношения информантов к тексту⁴. Подсчет процентов: за 100 % принималась сумма всех баллов, выставленных всеми участниками по каждому фрагменту. Затем вычислялся средний балл для каждого фрагмента, который и представлял процент понятности.

Этап 2: реализуется элементаристский подход. Произведено членение текста на фрагменты, каждый из которых был проанализирован на наличие затруднений в понимании. Участники должны были указать слова, словосочетания или фразы, вызвавшие у них затруднения.

Вопрос, заданный информантам: укажите слова, словосочетания или фразы, которые вызвали у вас затруднение в понимании данного фрагмента (для каждого из 6 фрагментов). Шкала оценивания не использовалась. Оценивание производилось на основе качественных данных – конкретных слов и выражений, указанных / выбранных респондентами. Процент затруднений рассчитывался как доля ответов, содержащих упоминание конкретного слова или выражения, от общего числа ответов по определенному фрагменту. Результаты сравнивались с данными обратного машинного перевода.

Технологическая сторона исследования заключалась в проведении опроса-эксперимента, в котором приняли участие 100 человек. Однако 1048 человек начали заполнять анкету, но не завершили процесс, закрыв ее до окончания (данные были собраны за 63 дня, или 9 недель). Эти результаты указывают на высокую сложность предложенного для анализа текста, а также на отсутствие мотивации у большинства участников углубляться в анализ представленного материала. Важно отметить, что исследование не предполагало формирования специфической выборки респондентов: в эксперименте участвовали жители России, владеющие русским языком как родным. Отказ от создания целенаправленной выборки обусловлен тем, что используемая в исследовании модель машинного перевода запрограммирована и настроена на работу со среднестатистическим пользователем. Таким образом, ценность мнений опрашиваемых в настоящем исследовании

заключается в том, что они представляют собой отражение восприятия среднестатистического носителя русского языка. Поскольку опрос проходил среди людей, проживающих на территории РФ и говорящих на русском языке как на родном, полученные данные отражают общие тенденции в понимании и интерпретации текстов на русском языке.

Фрагменты стихотворения, которые были взяты для проверки предложенного метода:

- (1) О чем шумите вы, народные витии?
Зачем анафемой грозите вы России?
Что возмутило вас? волнения Литвы?
Оставьте: это спор славян между собою,
Домашний, старый спор, уж взвешенный судьбою,
Вопрос, которого не разрешите вы.
- (2) Уже давно между собою
Враждуют эти племена;
Не раз клонилась под грозою
То их, то наша сторона.
Кто устоит в неравном споре:
Кичливый лях, иль верный росс?
Славянские ль ручьи сольются в русском море?
Оно ль иссякнет? вот вопрос.
- (3) Оставьте нас: вы не читали
Сии кровавые скрижали;
Вам непонятна, вам чужда
Сия семейная вражда;
Для вас безмолвны Кремль и Прага;
Бессмысленно прельщает вас
Борьбы отчаянной отвaga –
И ненавидите вы нас...
- (4) За что ж? ответствуйте: за то ли,
Что на развалинах пылающей Москвы
Мы не признали наглой воли
Того, под кем дрожали вы?
За то ль, что в бездну повалили
Мы тяготеющий над царствами кумир
И нашей кровью искупили
Европы вольность, честь и мир?..
- (5) Вы грозны на словах — попробуйте на деле!
Иль старый богатырь, покойный на постеле,
Не в силах завинтить свой измайловский штык?
Иль русского царя уже бессильно слово?
Иль нам с Европой спорить ново?
Иль русский от побед отвык?
Иль мало нас? Или от Перми до Тавриды,
От финских хладных скал до пламенной
Колхиды,

⁴ Опрос в рамках исследования соотнесенности лингвокогнитивных категорий понятности и переводимости. URL: <https://forms.yandex.ru/u/6682f044068ff0035bb48908/> (дата обращения: 20.11.2024).

(6) *От потрясенного Кремля
До стен недвижного Китая,
Стальной щетиною сверкая,
Не встанет русская земля?..
Так высылайте ж к нам, витии,
Своих озлобленных сынов:
Есть место им в полях России,
Среди нечуждых им гробов.*

Результаты

Был проведен анализ переводимости фрагментов через идиш⁵ и другие языки: китайский, испанский, английский, хинди, французский, немецкий, потому что указанные языки входят в десятку самых распространенных языков в мире⁶, а также украинский, польский, белорусский (эти языки взяты для проверки гипотезы о естественной близости родственных (славянских) языков, создающей ситуацию большей степени понятности, чем неродственных) – с целью идентификации и регистрации случаев, вызывающих затруднения при машинном переводе. Обратные машинные переводы с дополнительных языков делались для верификации исследования.

Создана сводная таблица, отражающая гипотезу – по степени понятности фрагментов и понятности

языков-посредников. В таблице 1 представлена статистическая обработка результатов обратных машинных переводов названных фрагментов, установлен среднестатистический процент сходства исходных русских текстов по степени их переводимости на другие языки. Полученные результаты (обратный машинный перевод, русский язык) проведены через программу сравнения текстов на схожесть⁷.

Вслед за Э. С. Клышинским полагаем, что при наличии небольшого материала более корректно рассуждать о понятности конкретного текста, а не о понятности языка в целом [Клышинский 2021: 64], поэтому в таблице 1 проанализированы среднестатистические данные по каждому фрагменту и языку. Также анализ переводимости конкретного фрагмента, а не всего текста был проведен грузинскими учеными [Сичинава, Данелия 2022: 21], именно поэтому нами было принято решение анализировать не только общий смысл стихотворения, но и провести эксперименты по каждому фрагменту стихотворения отдельно.

Результаты эксперимента показали (табл. 1), что самым понятным является фрагмент 5 (40,33 %). Второе место по степени понятности занимает фрагмент 6 (34,58 %), третье – фрагмент 2 (33,91 %),

Табл. 1. Результаты сопоставления обратных переводов 6 строф стихотворения А. С. Пушкина «Клеветникам России», %
Tab. 1. Comparative analysis of reverse translations of six stanzas from Alexander Pushkin's To the Slanderers of Russia, %

Языки	Ф1	Ф2	Ф3	Ф4	Ф5	Ф6	Весь текст	Среднестатистическое по каждому языку
Китайский (традиционный)	15	18,87	18,6	15,69	14,49	18	16,14	16,94
Испанский	23,68	23,4	21,95	26,09	40,38	33,33	29,22	28,13
Английский	41,18	40,48	25	26,09	47,06	37,5	37,07	36,22
Хинди	28,21	39,53	25	18,37	40,74	27,91	29,84	29,96
Французский	27,78	28,26	21,95	24,44	42,31	37,5	32,08	30,37
Немецкий	21,95	29,79	24,39	18,37	38,89	37,5	28,92	28,48
Украинский	63,33	69,44	59,38	61,76	68,18	55,56	63,02	62,94
Польский	27,03	29,55	23,81	28,26	33,33	30,95	28,8	28,82
Белорусский	35,29	30	25	31,11	36,36	36,59	32,78	32,39
Идиш	27,78	29,79	16,28	17,39	41,51	30,95	27,71	27,28
Среднестатистическое по каждому фрагменту	31,12	33,91	26,14	26,76	40,33	34,58	32,56	32,14

Прим.: Ф – фрагмент.

⁵ Данный язык является языком народов России. В РФ идиш относится к группе малых языков. Этот язык распространен среди жителей Еврейского автономного округа. В ЮНЕСКО его включили в список языков, находящихся под угрозой исчезновения. Идиш представлен в работе как материал для специализированного (конкретного) анализа, но мог бы быть взят любой язык.

⁶ Самые распространенные языки мира. URL: <https://infografics.ru/all/samye-populjarnye-jazyki-mira/> (дата обращения: 20.11.2024).

⁷ Программа сравнения текстов на схожесть...

четвертое – фрагмент 1 (31,12 %), пятое – фрагмент 4 (26,76 %). Самым непонятным оказался фрагмент 3 (26,14 %).

Э. С. Клышинский пишет: «Традиционно считается, что русский язык, относящийся к восточнославянской подгруппе, наиболее близок к другим языкам этой же подгруппы – белорусскому и украинскому» [Клышинский 2021: 62]. В нашем исследовании подтвердился этот тезис: украинский язык показал самый высокий процент понятности – 62,94 %, а белорусский язык занял третье место – 32,39 %.

Кроме того, был проведен опрос-эксперимент с целью сравнения результатов обратного машинного перевода с ответами респондентов. Ход проведения опроса: участникам эксперимента было предложено 6 фрагментов стихотворения⁸. С данными фрагментами испытуемым необходимо было ознакомиться и ответить на следующие вопросы: понятен ли общий смысл текста; какой из фрагментов представляется наиболее сложным для понимания; респондентам было предложено указать, какие слова и выражения фрагмента затрудняли понимание; как они поняли конкретную фразу; также было предложено выбрать из конкретного списка слова, которые затрудняют понимание. Таким образом, опрос-эксперимент включал в себя вопросы, опирающиеся на фрагменты 6 текстов стихотворения. Благодаря этим вопросам в анкете мы выяснили, какие именно фрагменты трудны для понимания и интерпретации, какие конкретные элементы представляли сложность.

Проанализируем ответы на задаваемые вопросы, учитывая количественные и качественные закономерности. Результаты, касающиеся первого вопроса (Напишите, понятен ли Вам общий смысл текста: к кому оно обращено, и что хотел сказать поэт клеветникам России?), показали, что для 64 % текст стихотворения в целом был больше непонятен, чем понятен. Приведем несколько ответов респондентов: *во-первых, текст достаточно длинный, пока прочитаешь, сбиваешься с мыслей, а, во-вторых, для меня в нем встретилось много необычных слов, значения которых я не знаю⁹; Может, недругам России посвящено...; Нет; К недрузьям России. Не стоит нас трогать; Затрудняюсь ответить, вроде да; Да, в целом понятен. Поэт хочет сказать, чтобы клеветники не лезли не в свое дело; Понятен. Обращался к мужам России, призывая подняться на защиту Родины,*

а не оставаться «диванными критиками»; Да, понятен. Оно обращено к недругам России. Поэт хотел донести до них, что грозить России – опасно и глупо; Общий смысл текста непонятен, нужно сильно вдумываться и вспоминать факты из истории; Про понятность общего смысла текста трудно говорить.

Вопрос о сложности текста выявил, что самым непонятным оказался фрагмент 3 (26,14 %), а самым понятным – фрагмент 5 (40,33 %). Второе место по степени понятности занимает фрагмент 6 (34,58 %), третье – фрагмент 2 (33,91 %), четвертое – фрагмент 1 (31,12 %), пятое – фрагмент 4 (26,76 %).

Выявлены слова и выражения фрагмента 3, которые затрудняли его понимание: *кровавые скрижали, это про что речь?; Витии, анафема; Вития, анафема, взвешанный; Иль много раз повторяется; Щетиною сверкая; О какой семейной вражде идет речь? Почему Кремль и Прага?; Самая первая строка, потому что не знаю значение слова «витии»; Тяготеющий над царствами кумир; Колхида; Потрясенный Кремль и недвижный Китай – что это вообще, непонятно! и про недвижные гробы непонятно.*

Далее были получены развернутые ответы на вопрос, как респонденты поняли фразу **О чем шумите вы, народные витии?**: *Ну, шумите – типа тревожитесь, громко что-то говорите... А вот витии – вообще хз!; Не знаю значения слова такого; Не понял; О чем разговариваете? О чем сплетничаете?; О чем вы поднимаете бучу, ажиотаж; Среди народа есть ораторы, которые возмущают окружающих, высказывают мнения; Ну, шумите, типа говорите, а вот значение слова витии я не знаю; О чем хотите сказать?*

Также респондентам был задан вопрос на понимание фразы **Зачем анафемой грозите вы России?** Представим десять развернутых ответов: *ну, грозите – типа угрожаете, а вот анафемой... Что такое анафема?; Затрудняюсь ответить; Не знаю, что такое анафема; Сыпете угрозы России; Я не знаю, что такое слово «анафема», так бы было понятнее, думаю; Как отлучение России от чего-либо или кого-либо; Мне кажется, здесь имеется ввиду разделение России и Польского народа; «Зачем проклятием и всяческим злом грозите вы России»; Те, кто рассуждают о судьбе России, грозят ей отлучением от церкви, если она пойдет по такому пути; Я не знаю, что такое анафемой.*

⁸ См. фрагменты стихотворения, представленные в разделе Методы и материалы.

⁹ Здесь и далее в примерах сохранены авторские орфография и пунктуация.

Далее респондентам предлагалось выбрать из списка слова / словосочетания, которые им показались трудными для восприятия; при этом список слов был составлен неслучайно: в нем оказались те слова / словосочетания, которые не были корректно обработаны системой машинного перевода.

В таблице 2 приведены слова / словосочетания / фразы, которые вызвали затруднения в понимании (которые оказались непонятными, или которые вызвали трудность в первоначальном понимании), по каждому из 6 фрагментов и их процентное соотношение. Напомним, что процент затруднений рассчитывался как доля ответов, содержащих упоминание конкретного слова или выражения, от общего числа ответов по данному фрагменту.

Таким образом, результаты опроса-эксперимента совпали с изначальным экспериментом, построенным на методе обратного машинного перевода: респонденты испытывали затруднения в понимании тех же слов и словосочетаний, что и ИИ. Эксперимент показал прямую связь между понятностью и переводимостью текста. Сложная лексика снижает переводимость текста. Большинство респондентов отметили, что тексты с культурными реалиями сложнее воспринимать. Итак, языковые единицы, вызывающие затруднения у машинного переводчика, часто оказываются сложными и для носителей языка. В этом видим закономерность, поскольку люди разрабатывают программы машинного перевода и базы данных к ним.

Обсуждение

Полученные результаты подтверждают предположение о том, что обратный машинный перевод может служить маркером для анализа заявленных параметров, что согласуется с выводами ряда предыдущих исследований. Так, А. Д. Шмелев и А. А. Зализняк подчеркивают, что обратный (реверсивный) перевод может служить инструментом лингвистического анализа слов [Шмелев, Зализняк 2017: 395]. Однако наш подход отличается тем, что мы используем этот метод не просто для анализа отдельных слов или фрагментов текста, а для выявления помех в понимании текста. Этот аспект делает наше исследование особенно актуальным, т.к. позволяет глубже понять причины снижения качества перевода и, соответственно, улучшить процессы адаптации текста к целевым языкам.

С помощью обратного машинного перевода А. В. Иркова изучала степень юридизированности лексики текстов закона [Иркова 2022], которые

Табл. 2. Результаты опроса-эксперимента: процентное соотношение непонятных слов / словосочетаний / фраз по каждому из 6 фрагментов стихотворения

Tab. 2. Survey results: incomprehensible words and phrases per stanza, %

Фрагменты стихотворения		%
Ф1	народные витии	43,6
	Что возмутило вас?	4,3
	уж взвешенный судьбою	21,7
	Зачем шумите	17,4
	славян между собою	8,7
	Которого не разрешите вы	4,3
Ф2	кичливый лях	21,9
	иль верный росс	18,8
	Оно ль иссякнет?	18,8
	не раз клонилась под грозою	15,6
	иль ручьи	15,6
	иль	6,3
Ф3	враждуют	3
	сии кровавые скрижали	28,6
	вам непонятна вам чужда	14,3
	борьбы отчаянной отвага	14,3
	прельщает вас	25
	и ненавидите вы нас	10,7
Ф4	безмолвны	7,1
	Мы тяготеющий над царствами кумир	36
	на развалинах пылающей Москвы	12
	За что ж?	4
	наглой воли	8
	За то ль, что в бездну повалили	8
Ф5	ответствуйте: за то ли	4
	И нашей кровью искупили	8
	Не в силах завинтить свой измайловский штык?	33,3
	Иль русского царя уже бессильно слово?	11,1
	Иль нам с Европой спорить ново?	11,1
	От финских хладных скал до пламенной Колхиды	14,8
Ф6	Иль старый богатырь, покойный на постеле	11,1
	Вы грозны на словах	7,4
	Иль русский от побед отвык?	7,4
	Иль мало нас?	3,7
	Стальной щетиною сверкая	18,8
	От потрясенного Кремля	15,6
	До стен недвижимого Китая	15,6
	Не встанет русская земля?	12,5
	Так высылайте ж к нам, витии	12,5
	Своих озлобленных сынов	12,5
	Есть место им в полях России	6,3
	Среди нечуждых им гробов	6,3

Прим.: Ф – фрагмент.

демонстрируют значительное снижение точности при обратном переводе, подтверждая необходимость упрощения языка для повышения переводимости. Аналогичные тенденции прослеживаются и в нашей работе, но мы фокусируемся на анализе возможных причин возникновения этих трудностей.

Кроме того, в работе Н. А. Фатеевой акцентируется внимание на важности понятности текста для успешного перевода, а также для создания понятного метафорического образа [Фатеева 2019: 70]. В ходе нашего исследования было установлено, что фрагменты, которые были оценены как понятные на исходном языке, демонстрируют более высокие показатели точности при обратном машинном переводе. Это согласуется с выводами Н. А. Фатеевой о том, что понятность текста на исходном языке является ключевым фактором, влияющим на качество адекватного перевода. В нашей работе анализируется влияние различных факторов на понимание текста, таких как сложная структура предложений, терминология, с целью выявления новых аспектов проблемы. Таким образом, результаты исследования предлагают новый метод оценки текста через призму обратного машинного перевода.

Заключение

Исследование подтвердило гипотезу о том, что обратный машинный перевод может служить эффективным методом, позволяющим выявить скрытые трудности восприятия текста и его потенциальные проблемы при переводе. Логика этого утверждения основана на следующем наблюдении: при обратном машинном переводе тексты с высокой степенью сложности в значительной мере ослабляют уровень своей понятности. Сигналы трудности понимания, посылаемые обратным машинным переводом, свидетельствуют о наличии сложной структуры текста, которая требует дополнительных усилий для интерпретации со стороны среднестатистического человека. В результате метод обратного машинного перевода предоставляет возможность

объективно оценить сложность и понятность текста, а также выявить его слабые места.

Установлена прямая зависимость между понятностью и переводимостью фрагментов текста, что подтверждает гипотезу: информантам в тексте трудны для понимания те же единицы, что и машинному переводчику. Чем сложнее текст для восприятия, тем больше вероятность возникновения проблем при его интерпретации. Следовательно, высокая степень сложности текста препятствует его эффективной передаче на другие языки.

Сравнение результатов эксперимента с изначальной задачей продемонстрировало, что метод обратного машинного перевода действительно позволяет не только оценить лексическую и синтаксическую структуру текста, но и выявить его сложные участки, которые могут затруднять понимание. Экспериментальные данные подтвердили, что тексты с высоким уровнем сложности заметно теряют свою понятность при обратном машинном переводе, что дополнительно подтверждает обоснованность предложенного метода.

Исследование научно значимо в систематизации методологического подхода к оценке текстов с использованием обратного машинного перевода. Предложен новый эффективный инструмент для анализа сложности и понятности текста. Применение данного метода будет полезно редакторам, законодателям, авторам и педагогам, стремящимся повысить качество своих материалов в аспекте понятности. В дальнейшем с помощью предложенного метода нами будет проведен глубокий анализ юридического текста.

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

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

Литература / References

- Бабкина М. Д. Ясный язык как средство обеспечения доступности информации. Рекомендации по применению. Предварительный национальный стандарт. М.: Наш солнечный мир, 2021. 144 с. [Babkina M. D. *Easy language as a means of making information accessible. Recommendations for use. Preliminary national standard*. Moscow: Nash solnechnyi mir, 2021, 144. (In Russ.)]
- Болдырев Н. Н. Когнитивный подход в лингвистике и смежных областях знания. *Вопросы когнитивной лингвистики*. 2020. № 2. С. 5–25. [Boldyrev N. N. Cognitive approach in linguistics and related areas of research. *Issues of Cognitive Linguistics*, 2020, (2): 5–25. (In Russ.)] <https://doi.org/10.20916/1812-3228-2020-2-5-25>

- Голев Н. Д. Источниковый потенциал обратного машинного перевода. *Вестник Кыргызско-Российского Славянского университета*. 2018. Т. 18. № 1. С. 36–45. [Golev N. D. Source potential of the back machine translation. *Herald of Kyrgyz-Russian Slavic University*, 2018, 18(1): 36–45. (In Russ.)] <https://elibrary.ru/yspbud>
- Голев Н. Д. Транслятивная лингвистика (аспектуализированный обзор исходных положений). Часть 1. Гносеология перевода. *Вестник Кемеровского государственного университета*. 2022. Т. 24. № 6. С. 717–734. [Golev N. D. Translative linguistics: An aspectualized review of initial provisions. Part 1. Gnoseology of translation. *Vestnik Kemerovskogo gosudarstvennogo universiteta*, 2022, 24(6): 717–734. (In Russ.)] <https://doi.org/10.21603/2078-8975-2022-24-6-717-734>
- Иркова А. В. Изучение степени юридизированности лексики текстов закона цифровыми методами. *Филология, иностранные языки и медиакоммуникации: XVII (XLIX) Междунар. науч. конф. студентов, аспирантов и молодых ученых*. (Кемерово, 1–30 апреля 2022 г.) Кемерово: КемГУ, 2022. С. 545–549. [Irkova A. V. Studying the degree of juridization of law text vocabulary by digital methods. *Philology, foreign languages, and media communications: Proc. XVII (XLIX) Intern. Sci. Conf. students, postgraduates and young scientists*, Kemerovo, 1–30 Apr 2022. Kemerovo: KemSU, 2022, 545–549. (In Russ.)] <https://elibrary.ru/lcacya>
- Катышева Л. В. «Простой и понятый язык» как новый тренд коммуникации государства и общества. *Коммуникология*. 2022. Т. 10. № 2. С. 85–94. [Katysheva L. V. Plain and easy language as a new trend of communication between the state and society. *Communicology*, 2022, 10(2): 85–94. (In Russ.)] <https://doi.org/10.21453/2311-3065-2022-10-2-85-94>
- Клышинский Э. С. Статистический метод количественной оценки понятности иностранных славянских языков для русскоязычного читателя. *Вестник Новосибирского государственного университета. Серия: Информационные технологии*. 2021. Т. 19. № 1. С. 61–79. [Klyshinsky E. S. Quantitative estimation of intelligibility of foreign Slavic languages: Case of Russian native speakers. *Vestnik Novosibirskogo gosudarstvennogo universiteta. Seriya: Informacionnye tekhnologii*, 2021, 19(1): 61–79. (In Russ.)] <https://doi.org/10.25205/1818-7900-2021-19-1-61-79>
- Колмогорова А. В., Залевская Е. Д. Компьютерное моделирование как инструмент анализа художественного текста. *Филологический класс*. 2023. Т. 28. № 2. С. 22–33. [Kolmogorova A. V., Zalevskaya E. D. Computer-assisted modeling as an instrument for fiction text analysis. *Philological Class*, 2023, 28(2): 22–33. (In Russ.)] <https://elibrary.ru/lthcot>
- Кушникова Л. В., Литвинова С. В., Перлова И. В., Пермякова К. В. Интерпретирующая функция переводческого понимания как отражение индивидуальной переводческой картины мира. *Лингвориторическая парадигма: теоретические и прикладные аспекты*. 2019. № 24. С. 48–56. [Kushnina L. V., Litvinova S. V., Perlova I. V., Permyakova K. V. Interpretative function of translation understanding as reflection of translator's individual worldview. *Linguistic & rhetorical paradigm: Theoretical and applied aspects*, 2019, (24): 48–56. (In Russ.)] <https://elibrary.ru/vqswwj>
- Масленникова Е. М. Традиционные подходы к переводу: новое о старом. *Вестник ТвГУ. Серия: Филология*. 2007. № 9. С. 46–53. [Maslennikova E. M. Traditional approaches to translation: New about the old. *Vestnik TvGU. Series: Philology*, 2007, (9): 46–53. (In Russ.)] <https://elibrary.ru/jvdfid>
- Мельникова В. С. Измерение степени понятности текста с помощью обратного машинного перевода (на материале центрального стихотворения сборника «Пейзаж с наводнением» (1993) И. А. Бродского). *Инновационный конвент «Кузбасс: образование, наука, инновации»: XII Инновационный конвент*. (Кемерово, 8 февраля 2024 г.) Кемерово: КемГУ, 2024. С. 513–517. [Melnikova V. S. Measuring the degree of text understandability using reverse machine translation (based on the central poem of the collection "Landscape with flood" (1993) I. A. Brodsky). *Innovation Convention Kuzbass: Education, Science, and Innovation: Proc. XII Innovation Convention*, Kemerovo, 8 Feb 2024. Kemerovo: KemSU, 2024, 513–517. (In Russ.)] <https://elibrary.ru/fxiykd>
- Мельникова В. С. Измерение степени сложности текста закона с помощью компьютерных программ: к постановке проблемы. *Филология, иностранные языки и медиакоммуникации: XVI (XLVIII) Междунар. науч. конф. студентов, аспирантов и молодых ученых*. (Кемерово, 1–30 апреля 2021 г.) Кемерово: КемГУ, 2021. С. 31–33. [Melnikova V. S. Measuring the degree of difficulty of the text of the law with the help of computer programs: To the statement of the problem. *Philology, foreign languages, and media communications: Proc. XVI*

- (XLVIII) Intern. Sci. Conf. students, postgraduates and young scientists, Kemerovo, 1–30 Apr 2021. Kemerovo: KemSU, 2021, 31–33. (In Russ.)] <https://elibrary.ru/nxrxii>
- Микони С. В. Оценивание понятности научных русскоязычных работ. *Компьютерная лингвистика и вычислительные онтологии*. 2021. № 5. С. 61–69. [Mikoni S. V. Assessment of the comprehensibility of scientific Russian-language works. *Kompyuternaya lingvistika i vychislitelnye ontologii*, 2021, (5): 61–69. (In Russ.)] <https://doi.org/10.17586/0000-0000-2020-4-00-00>
- Раздобудько-Чович Л. И. Перевод как интерпретация и импровизация. На материале сербских переводов мемуаров и В. Набокова с двух «оригиналов»: английского (Speak, memory) и русского («Другие берега»). *Вестник Московского университета. Серия 22: Теория перевода*. 2012. № 2. С. 72–80. [Razdobudko-Čović L. I. Translation as interpretation and improvisation: A case study of authorized translations of Vladimir Nabokov's memoirs into Serbian. *Moscow University Translation Studies Bulletin*, 2012, (2): 72–80. (In Russ.)] <https://elibrary.ru/pemzhh>
- Сичинава Н. Г., Данелия Б. Лексические особенности «Сказки о рыбаке и рыбке» А. С. Пушкина и ее грузинских переводов. *Journal of Science. Lyon*. 2022. № 33. С. 17–22. [Sitchinava N. G., Danelia B. Lexical features of "The tale of the fisherman and the fish" by A. S. Pushkin and its Georgian translations. *Journal of Science. Lyon*, 2022, (33): 17–22. (In Russ.)] <https://doi.org/10.5281/zenodo.6967190>
- Солнышкина М. И. Сложность текста: status quo или движение вперед? *Технологическая независимость и конкурентоспособность Союзного Государства, стран СНГ, ЕАЭС и ШОС: VI Междунар. науч.-техн. конф.* (Минск, 6–8 декабря 2023 г.) Мн.: БГТУ, 2023. С. 421–430. [Solnyshkina M. I. Text complexity: Status quo or forward movement? *Technological independence and competitiveness of the Union State, the CIS countries, the EAC and the SCO: Proc. VI Intern. Sci.-Techn. Conf.*, Minsk, 6–8 Dec 2023. Minsk: BSTU, 2023, 421–430. (In Russ.)] <https://elibrary.ru/vjubal>
- Фатеева Н. А. Передача образности в переводе поэзии как результат интерпретации контекста (на примере перевода стихотворения Э. Багрицкого «Осень» (1915) на английский язык). *Иностранные языки в высшей школе*. 2019. № 4. С. 65–73. [Fateeva N. A. Interpreting context to convey imagery in translation of poetry (illustrated by translating E. Bagritsky's poem "Autumn" (1915) into English). *Foreign languages in tertiary education*, 2019, (4): 65–73. (In Russ.)] <https://elibrary.ru/jfzkqk>
- Шмелев А. Д., Зализняк А. А. Реверсивный перевод как инструмент лингвистического анализа дискурсивных слов. *Компьютерная лингвистика и интеллектуальные технологии: Ежегодная Междунар. конф. «Диалог»*. (Москва, 31 мая – 3 июня 2017 г.) М.: РГГУ, 2017. С. 394–406. [Shmelev A. D., Zaliznyak A. A. Reverse translation as a tool for analysis of discursive words. *Computational linguistics and intelligent technologies: Proc. Intern. Conf. "Dialog"*, Moscow, 31 May – 3 Jun 2017. Moscow: RSUH, 2017, 394–406. (In Russ.)] <https://elibrary.ru/xnfzzf>
- Щерба Л. В. О тройном аспекте языковых явлений и об эксперименте в языкознании. In: Щерба Л. В. *Языковая система и речевая деятельность*. М.: Наука, 1974. С. 24–39. [Shcherba L. V. The threefold aspect of linguistic phenomena and experimental linguistics. In: Shcherba L. V. *The language system and speech activity*. Moscow: Nauka, 1974, 24–39. (In Russ.)]
- Benjamin W. Die Aufgabe des Übersetzers. In: Benjamin W. *Gesammelte Schriften*. Suhrkamp, 1972, IV.1, 9–21.
- Kolmogorova A. V., Margolina A. V. Written vs generated text: "Naturalness" as a textual and psycholinguistic category. *Research result. Theoretical and Applied Linguistics*, 2024, 10(2): 71–99. <https://doi.org/10.18413/2313-8912-2024-10-2-0-4>