Лабораторная работа номер 13

Индивидуальный проект

Тема: «Роль человека в процессе машинного перевода: редактирование текста»

* Предредактирование текста: задачи и методы
* Постредактирование текста: задачи и подходы
* Роль лингвистов и переводчиков в процессе редактирования
* Использование автоматических систем в процессе редактирования
* Перспективы улучшения качества машинного перевода через редактирование

1. **Введение**

Машинный перевод (МП) активно развивается, но без участия человека он часто дает неточные или неестественные результаты. Предредактирование (адаптация текста перед переводом) и постредактирование (коррекция перевода) значительно улучшают качество. В проекте исследуются методы редактирования, роль лингвистов и автоматических инструментов в этом процессе.

1. **Предредактирование текста: задачи и методы**

Предредактирование – процесс модификации проблемных фрагментов исходного текста за счет изменения его структуры, состава лексических единиц и т.п. при сохранении семантической эквивалентности с целью оптимизации сложности переводческой задачи с учетом влияния такой модификации на качество перевода.

Предредактирование выполняется при средней и высокой сложности переводческой задачи с использованием разнообразных алгоритмов, например, по критериям: минимизации СлЗПtxtiTXT, минимаксимизации отдельных свойств и параметров исходного текста и др.

Диапазоны значений TTCtxtiTXT, соответствующие низкому, среднему и высокому уровням определяются на основе предварительного анализа требований к переводу и с учетом способа перевода (ручной/машинный). Цель предредактирования – снизить сложность переводческой задачи TTCtxtiTXT до низкой.

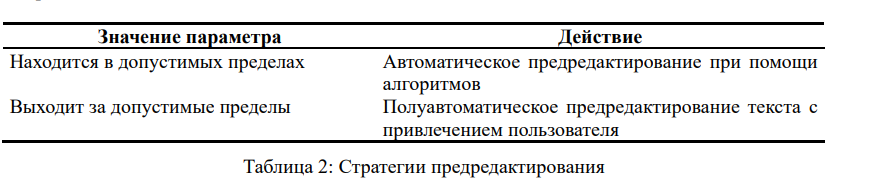
На начальном этапе предредактирования происходит отбор фрагментов текста, которым соответствуют высокие значения TTCi, то есть



где TTCalw – допустимое значение сложности задачи перевода, т.е. низкое.

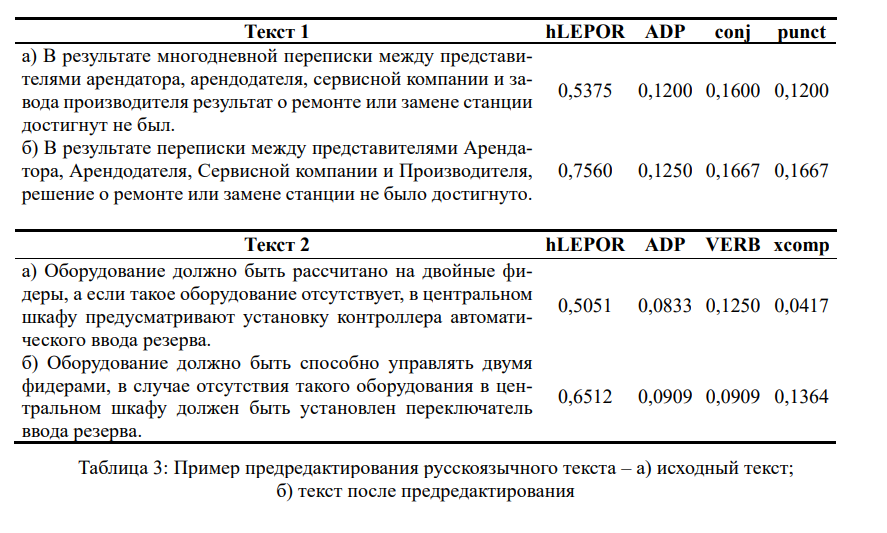
Для тех фрагментов текста, для которых сложность задачи перевода является средней или высокой, определяем элементы вектора Cfei, значения которых повышают сложность переводческой задачи данного фрагмента:

Для каждого найденного элемента матрицы Cfei, соответствующего условию  в зависимости от его значения и ограничений алгоритмов предредактирования определяется стратегия предредактирования (таблица 2). Выбор алгоритма зависит от критерия оптимизации выбранного значения.



Далее производится редактирование текста в соответствии с имеющимися методами и алгоритмами предредактирования, формируется текст txt′iTXT и происходит переход к этапу перевода.

Предредактирование позволяет повысить качество МП. Пример предредактирования и его влияния на значения параметров текста и качество перевода на английский язык представлены в таблице 3. В таблице представлены наиболее значимые параметры для соответствующего текста.



1. **Постредактирование текста: задачи и подходы**

Несомненно, [машинный перевод](https://lingvanex.com/ru/blog/machine-translation-in-natural-language-processing/) позволяет как небольшим частным компаниям, так и крупным международным корпорациям значительно экономить время. Но насколько надежны его результаты? Сложные фразы зачастую превращаются в бессмыслицу, теряются стилистические нюансы, игнорируются культурные особенности. Особенно остро это проявляется в литературном переводе, где важна творческая интерпретация. Что делать, если почти каждое пятое предложение нуждается в исправлении? Здесь на помощь приходит постредактирование машинного перевода.

Сегодня постредактирование — это не просто ручная корректировка ошибок профессиональными переводчиками и редакторами. Это автоматизированный процесс, где ключевую роль играют большие языковые модели (LLMs). Они не только исправляют очевидные ошибки, но и улучшают стиль, делая текст более естественным и читаемым. Обученные на больших объемах данных, эти модели «понимают» язык, а не просто сопоставляют слова.

Существуют различные подходы к росту редактирования, каждый из которых имеет свои особенности и преимущества:

1. Расширение деталей  
Этот подход предполагает детальное описание объектов, процессов или явлений, которые ранее были упомянуты лишь поверхностно. Например, если исходный текст описывает событие кратко, автор может добавить больше подробностей, цитат участников, статистические данные или исторические справки.  
Пример: Исходный текст:«Сегодня прошел митинг против повышения налогов.»Ростредактированный вариант:«Сегодня тысячи жителей города вышли на улицы, протестуя против планируемого правительством повышения налогов. Участники митинга подчеркивали негативное влияние данной меры на уровень жизни населения, приводили конкретные цифры роста цен на продукты первой необходимости и коммунальных услуг.»

2. Использование примеров и иллюстраций  
Добавление наглядных примеров помогает лучше раскрыть тему и делает материал понятнее читателю. Это особенно полезно в научных статьях, учебниках или аналитических обзорах.  
Пример: Исходный текст:«Процесс фотосинтеза важен для всех растений».Ростредактированный вариант:«Фотосинтез является ключевым процессом, обеспечивающим растения энергией. Например, зеленые листья деревьев поглощают солнечный свет, преобразуют углекислый газ и воду в глюкозу и кислород. Без этого процесса жизнь на Земле была бы невозможна, поскольку именно благодаря фотосинтезу атмосфера насыщается кислородом, необходимым животным и людям для дыхания.»

3. Пересказ и переформулировка  
Иногда текст кажется сухим или перегруженным сложными терминами. Переписывание отдельных фрагментов текста простым языком позволяет сделать материал доступнее и привлекательнее для широкой аудитории.  
Пример: Исходный текст:«Методы оптимизационной теории позволяют минимизировать затраты ресурсов предприятия.»Ростредактированный вариант:«Оптимизационная теория предлагает набор методов, позволяющих компаниям сократить расходы и повысить эффективность производства. Эти методы помогают грамотно распределять ресурсы, снижая издержки и увеличивая прибыль предприятия.»

4. Добавление комментариев экспертов  
Включив мнения специалистов или авторитетных лиц, можно значительно расширить объем текста и придать ему дополнительную значимость. Экспертные комментарии повышают доверие читателей и делают статью более информативной.  
Пример: Исходный текст:«Проблема глобального потепления требует срочных мер от мирового сообщества.»Ростредактированный вариант:«По мнению ведущих климатологов, проблема глобального потепления становится критической угрозой для планеты Земля. Ученые предупреждают, что повышение температуры даже на один градус Цельсия приведет к катастрофическим последствиям: засухам, наводнениям и массовой миграции животных и людей.»

5. Обобщение и систематизация  
Если статья охватывает большое количество фактов или материалов, можно объединить их в тематические блоки или главы, сгруппировав по определенному принципу. Такой подход упрощает восприятие большого объема информации и улучшает структурированность текста.  
Пример: Исходный текст:«Российская культура богата традициями и фольклором».Ростредактированный вариант:«Культура России представляет собой уникальное сочетание различных традиций и фольклорных элементов. К ним относятся народные праздники, такие как Масленица и Ивана Купала, традиционные ремесла, такие как гжельская керамика и хохломская роспись, а также богатое музыкальное наследие, представленное композиторами Чайковским, Рахманиновым и Шостаковичем.»

1. **Роль лингвистов и переводчиков в процессе редактирования**

Лингвист может уловить слова, которые не совсем понятны в контексте, созданном машиной. Он замечает, что текст может потребоваться локализовать, чтобы он был приемлемым в целевой культуре. Лингвист также следит за тем, чтобы в тексте использовалась правильная терминология и она соответствовала целевой аудитории.

Переводчик способен проверять правильность выбора, применяя логические и языковые правила, делать логические выводы, обобщать разрозненные материалы. Он может выбирать из ряда источников на языке перевода наиболее подходящее соответствие для непереведённой части фрагмента и приводить его в соответствие с контекстом и лингвистическими требованиями конкретного языка.

1. **Использование автоматических систем в процессе редактирования**

*Grammarly, LanguageTool — проверка грамматики*

Опечатки и грамматические ошибки случаются, но проверка своей работы «на глаз» не даёт стопроцентной гарантии, так как некоторые ошибки могут ускользнуть от внимания или их легко пропустить.

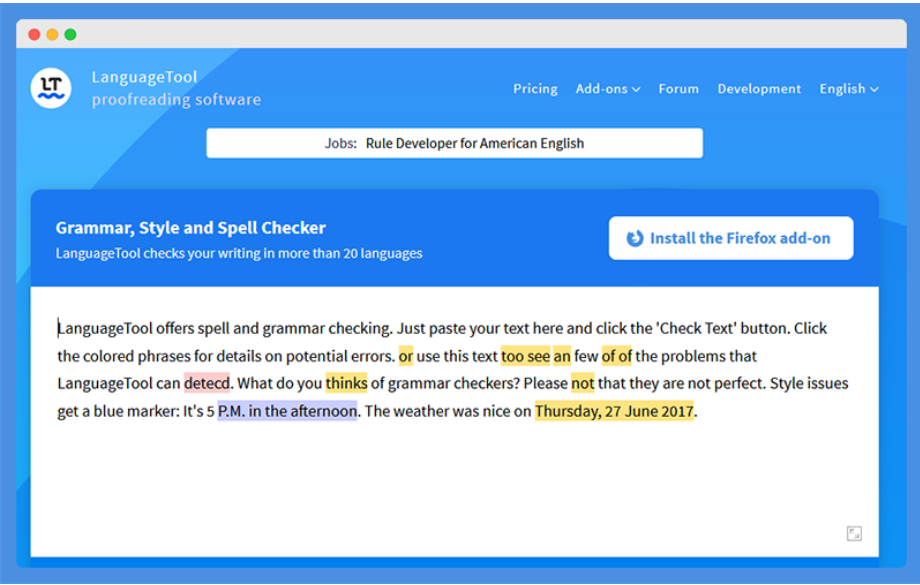
Что вам нужно добавить в свой личный набор инструментов для письма, так это программу проверки грамматики, которая поможет вам редактировать свои работы быстрее и эффективнее.

LanguageTool и Grammarly — одни из [лучших программ для проверки грамматики](https://rigorousthemes.com/blog/best-free-grammar-checker-online/), которые помогут вам заранее выявлять ошибки и улучшать свою письменную речь. Задача состоит в том, чтобы выбрать между ними, поскольку у обеих есть свои преимущества и недостатки.

Мы сравниваем LanguageTool и Grammarly, чтобы вы могли выбрать тот вариант, который вам больше нравится, и использовать его в своей повседневной работе.

LanguageTool против Grammarly: что это?

Языковой инструмент

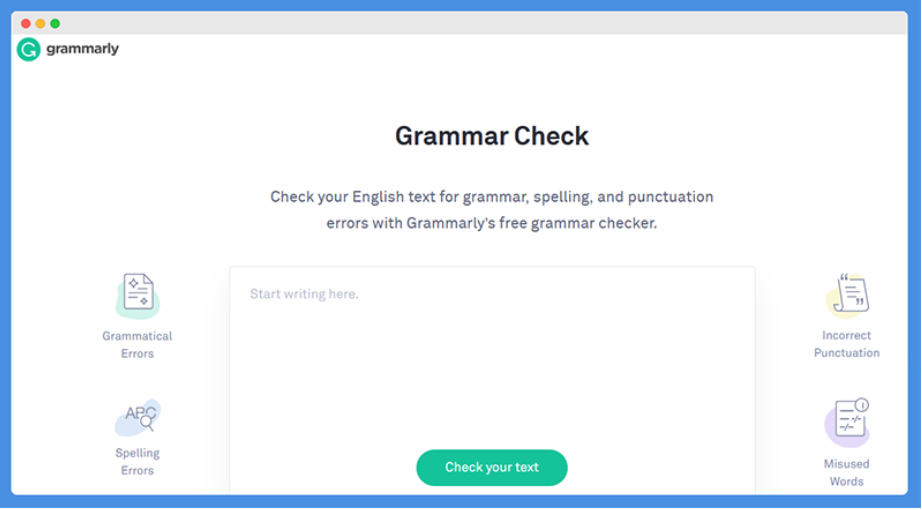


LanguageTool — это программа проверки грамматики с открытым исходным кодом, которая доступна в виде настольного приложения, веб-сервиса и расширения для браузера.

Этот инструмент был разработан группой программистов и любителей языка, которые создали его, чтобы помочь частным и корпоративным пользователям проверять тексты, корректировать свою работу и устранять грамматические и стилистические ошибки.

LanguageTool известен своей способностью выявлять грамматические, орфографические и пунктуационные ошибки в более чем 20 языках и их вариантах. К ним относятся английский, немецкий, португальский, французский, испанский, польский, русский, итальянский и несколько других.

Grammarly



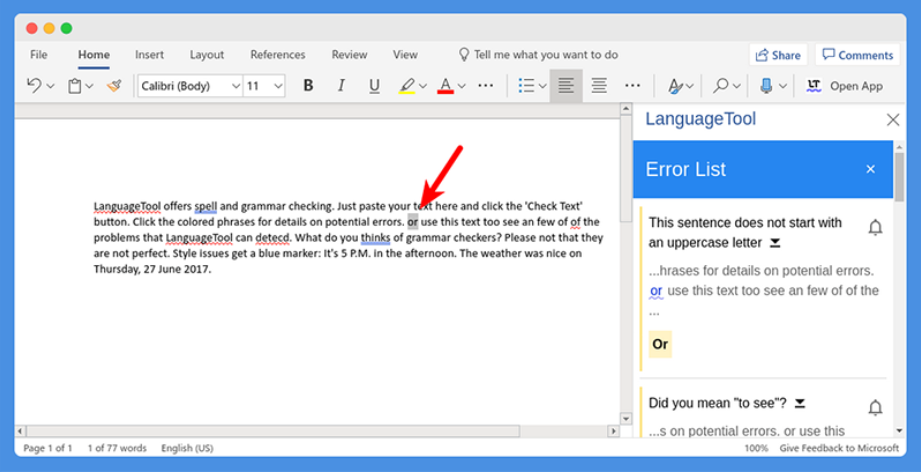
Grammarly — это ведущий онлайн-инструмент для проверки грамматики, которым пользуются профессионалы, студенты и частные лица в разных сферах и отраслях. Этот инструмент может определять стиль и предлагать расширенные пояснения к вашим текстам.

Кроме того, Grammarly — это простой в использовании, настраиваемый инструмент, который помогает вам писать без ошибок и улучшать свою манеру письма, чтобы вы могли говорить то, что думаете.

Grammarly использует инновационные подходы, в том числе передовые технологии, машинное обучение, глубокое обучение и искусственный интеллект (ИИ), чтобы предоставлять непревзойденную помощь в написании текста без ошибок. Вы можете проверять свой текст на наличие распространенных и сложных ошибок в любом используемом вами приложении, включая Gmail, Google Docs, LinkedIn, Twitter, Facebook и другие.

Grammarly vs LanguageTool: как они работают?

Языковой инструмент



LanguageTool — это простая и удобная в использовании программа для проверки грамматики. Всё, что мне нужно сделать, — это ввести текст в текстовое поле на веб-странице, и она обнаружит и подчеркнёт все грамматические и орфографические ошибки, которые могут содержаться в текстовом поле.

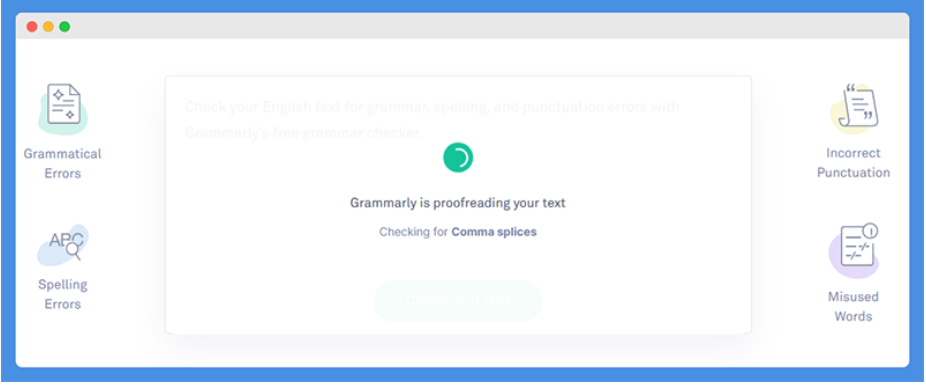
Когда я нажимаю на подчеркнутую ошибку, LanguageTool отображает всплывающее окно с предложениями по замене слова или фразы, и я могу нажать на предложение, чтобы заменить ошибку правильным написанием или фразой.

Кроме того, инструмент позволяет мне нажать на значок в нижней части страницы, чтобы увидеть потенциальные ошибки, которые выявляет LanguageTool, и их исправления.

Самое приятное в использовании LanguageTool — это то, что это инструмент с открытым исходным кодом, поэтому мне не нужно регистрироваться или подписываться, чтобы использовать его для проверки орфографии и грамматики в моём тексте.

Кроме того, я могу использовать персонализированные настройки для разных сайтов с помощью расширения LanguageTool для браузера, которое также позволяет переключать регистр текста и отключать дополнение на сайте.

Grammarly



Что мне нравится в Grammarly, так это то, что он следит за тем, чтобы всё, что вы пишете, было понятным, легко читаемым и правильным. Он использует алгоритмы, чтобы выявлять возможные ошибки в тексте, и предлагает варианты, соответствующие контексту, чтобы помочь с правописанием, грамматикой, использованием слов, стилем, многословием, пунктуацией и плагиатом.

Программа объясняет каждое предлагаемое решение, чтобы я мог решить, стоит ли исправлять проблему.

Мне нравится бесплатное расширение для браузера Grammarly, особенно когда я пишу на разных сайтах в интернете. Расширение также позволяет копировать или вставлять текст в настольное приложение или редактор Grammarly.

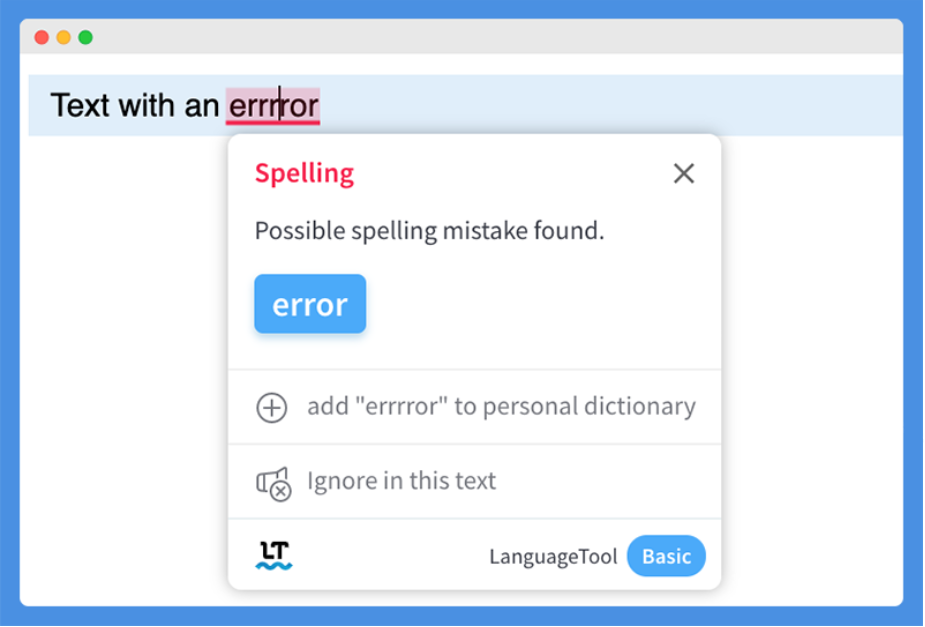
Кроме того, когда я работаю над документом Microsoft Word, Grammarly для Microsoft Office или Grammarly для Microsoft Word на Mac очень удобен, так как он выявляет ошибки по мере их появления.

Существуют приложения для устройств на iOS и Android, с помощью которых я могу редактировать и проверять свою работу на ходу и быть уверенным, что в том, что я печатаю, нет ошибок. На самом деле, приложение Grammarly для iPad предлагает редактор Grammarly без необходимости открывать его в интернете.

LanguageTool против Grammarly: Особенности

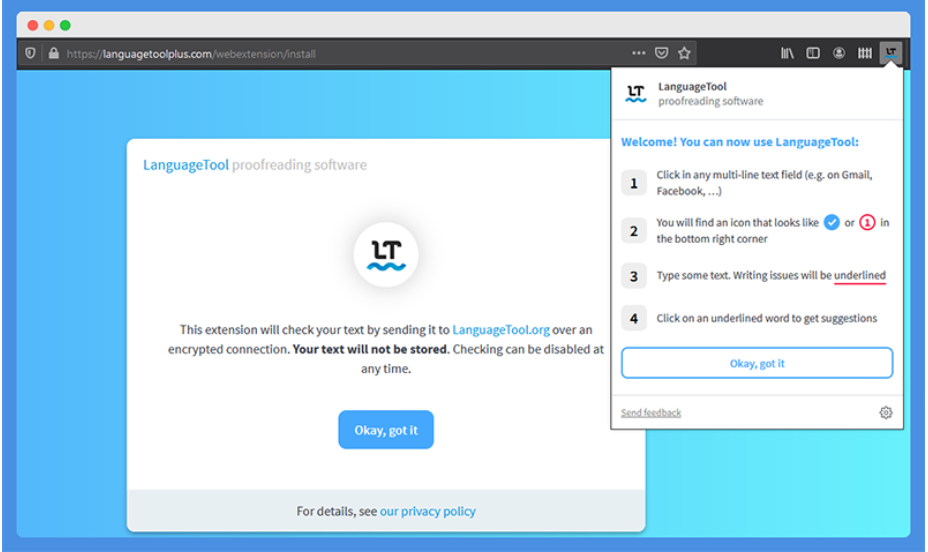
Языковой инструмент

ПЕРСОНАЛЬНЫЙ СЛОВАРЬ



С помощью личного словаря мне легко находить слова или исключения, которые я использую почти каждый день, но которые я не обязательно найду в обычном словаре и которые могут быть частью моего родного языка. Инструмент сканирует и проверяет мой текст на основе определённых правил, хотя я всегда могу отключить эти правила, а также добавлять слова в словарь.

РАСШИРЕНИЕ LANGUAGETOOL

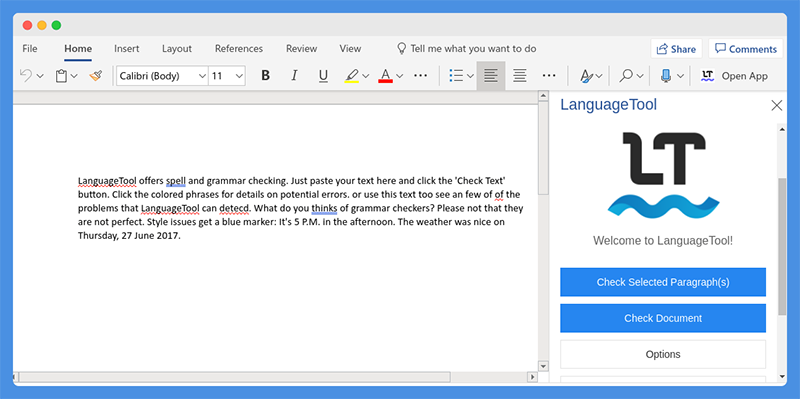


Это бесплатное расширение для браузера Chrome или Firefox позволяет мне проверять текст, и инструмент выявляет ошибки, которые не могут обнаружить обычные средства проверки орфографии, а также некоторые грамматические ошибки.

Как и полнофункциональное приложение, это расширение поддерживает более 25 языков и их диалекты, включая английский, французский, русский, испанский, немецкий, голландский и польский. Это делает его незаменимым расширением для тех, для кого эти языки не являются родными.

Я также использую LanguageTool в качестве дополнения для Google Docs, Microsoft Word или LibreOffice, чтобы проверять текст непосредственно в этих приложениях и находить неправильные слова или фразы с пояснениями и предложениями.

РАБОЧИЙ СТОЛ LANGUAGETOOL



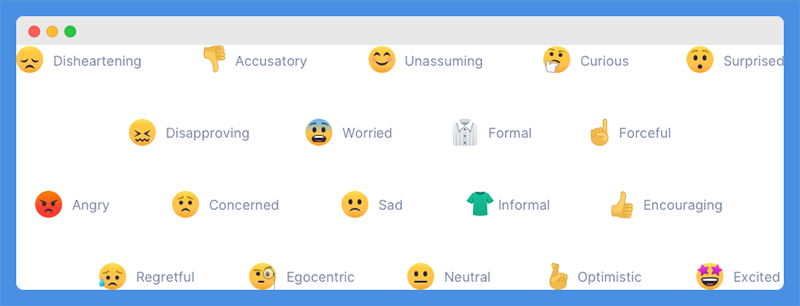
[Настольная версия LanguageTool](https://languagetool.org/download/LanguageTool-5.1.zip) [идеально подходит для использования в автономном режиме](https://rigorousthemes.com/blog/best-offline-grammar-checker-software/). Однако для этого требуется Java 8 или более поздняя версия, и она будет работать только в том случае, если вы установите дополнительные данные, в противном случае она будет работать так же, как и онлайн-версия. После загрузки zip-файла распакуйте его и дважды щелкните по languagetool.jar, чтобы запустить приложение.

СОЧЕТАНИЯ КЛАВИШ

Когда мне нужно быстрее проверить текст, сочетания клавиш в LanguageTool — настоящая находка.

Grammarly

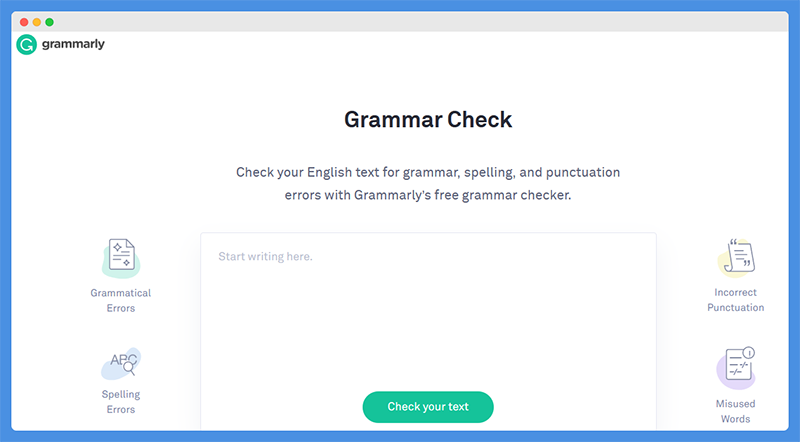
ДЕТЕКТОР ТОНОВ



Детектор интонации Grammarly помогает мне донести сообщение именно так, как я хочу его донести. Я могу говорить обеспокоенно, дружелюбно, неформально или в любом другом тоне в зависимости от контекста, чтобы направить своё сообщение в нужное русло и донести его эффективно. Таким образом, вы можете быть уверены, что моя аудитория отреагирует так, как я ожидаю.

Проверка тональности анализирует мои формулировки, выбор слов, использование заглавных букв и знаков препинания, чтобы определить тон моего сообщения перед его отправкой.

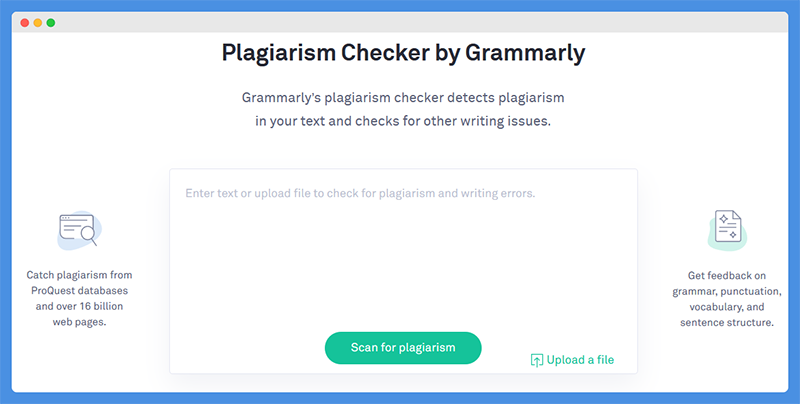
ПРОВЕРКА ГРАММАТИКИ



Grammar Check — это ещё один бесплатный онлайн-инструмент от Grammarly, который проверяет ваш текст на наличие орфографических, грамматических и пунктуационных ошибок. Он также проверяет наличие опечаток, ошибок, проблем со структурой предложений и многое другое.

Таким образом, вы можете исправить и устранить эти сложные ошибки, а также получить предложения по исправлению неправильно расставленных знаков препинания или часто путаемых слов, используемых в неподходящем контексте. С Grammarly Premium вы получаете доступ к более продвинутым функциям и проверкам.

ПРОВЕРКА НА ПЛАГИАТ

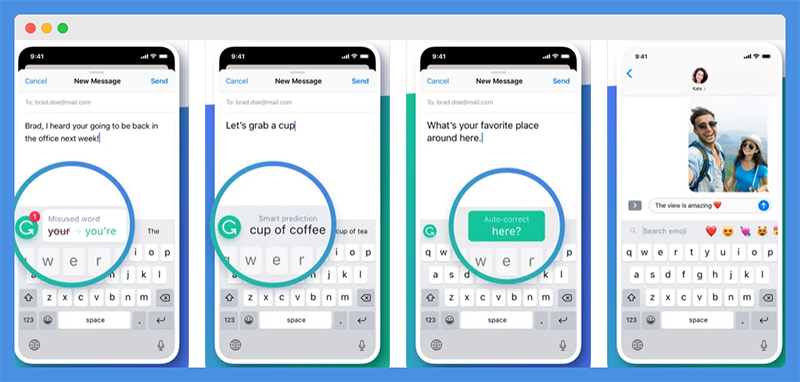


Как писатель, я очень беспокоюсь о своей работе и о том, как поисковые системы воспринимают каждую её часть. Один из самых важных инструментов, который я считаю полезным в Grammarly, — это проверка на плагиат.

Это один из тех инструментов, за которые нужно платить отдельно, но в Grammarly есть встроенный инструмент для проверки на плагиат, который обнаруживает любое дублирующееся содержимое в моём тексте и проверяет его на наличие других проблем.

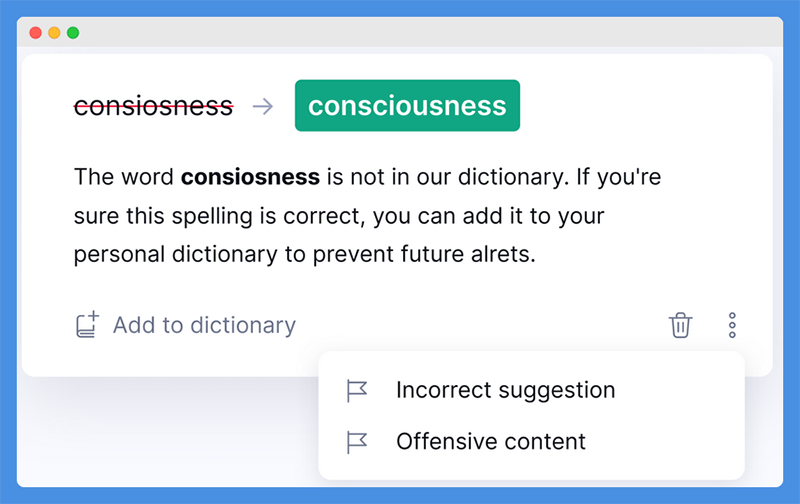
Самое приятное то, что эта функция выявляет плагиат на миллиардах веб-страниц и в академических базах данных и сообщает мне, есть ли в моём тексте дублированный материал. Она также выделяет области, требующие цитирования, и предлагает ресурсы, которые помогут мне правильно указать источники.

ГРАММАТИЧЕСКАЯ КЛАВИАТУРА



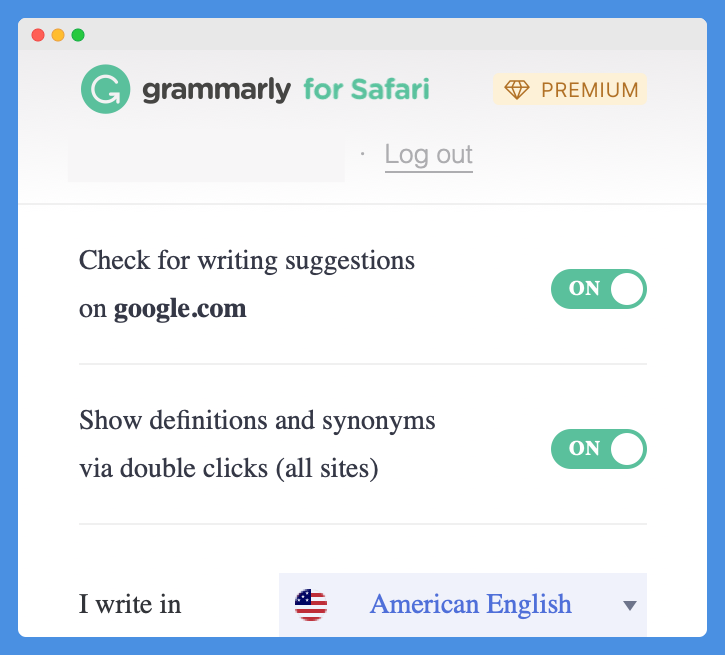
Как и в LanguageTool, в Grammarly тоже есть клавиатура, но для её установки мне пришлось отсканировать QR-код с помощью телефона. Клавиатура помогает мне создавать более качественные электронные письма, тексты без ошибок для социальных сетей и более понятные тексты. Она доступна для устройств [iOS](https://apps.apple.com/app/id1158877342?mt=8) и [Android](https://play.google.com/store/apps/details?id=com.grammarly.android.keyboard&referrer=adjust_reftag%3DcHJoMN4eU3OiV%26utm_source%3Dnavigation).

ГРАММАТИЧЕСКИЙ РЕДАКТОР



В Grammarly есть интуитивно понятный текстовый редактор, который я считаю чрезвычайно полезным для быстрой проверки текстов онлайн. Это удобное место, где я могу писать и настраивать предложения и тексты в интернете.

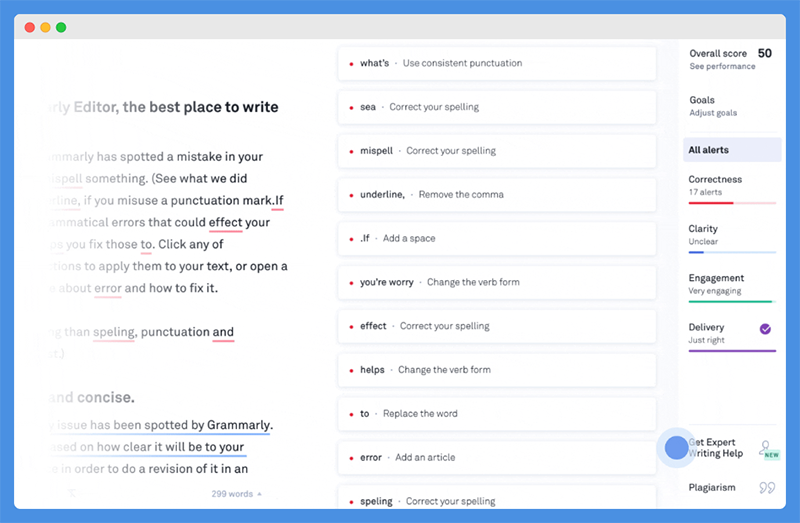
РАСШИРЕНИЯ БРАУЗЕРА



У Grammarly также есть [расширения для браузеров](https://support.grammarly.com/hc/en-us/articles/115000091552-Install-the-free-Grammarly-Spelling-and-Grammar-Checker-Extension) для четырёх основных браузеров, включая Safari, Chrome, Edge и Firefox, что очень удобно, потому что LanguageTool ограничивает меня только двумя браузерами.

Кроме того, расширение Grammarly предлагает исправления на многих сайтах, включая Twitter, Medium и Google Docs, исправляя сотни типов ошибок в моём тексте с помощью мощных алгоритмов Grammarly.

ЭКСПЕРТНЫЕ ПИСЬМЕННЫЕ УСЛУГИ

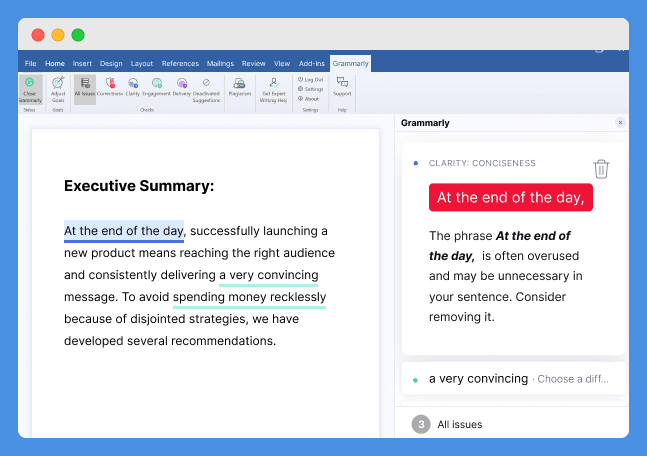


Иногда мне нужен дополнительный глаз, чтобы проверить свою работу, вот почему мне нравится сервис экспертного письма Grammarly. Эта дополнительная функция, доступная с тарифным планом Premium, дает мне доступ к команде экспертов Grammarly по написанию текстов, которые просматривают и исправляют любые обнаруженные ошибки.

Эксперты тщательно отбираются и регулярно проверяются, чтобы моя работа была краткой, точной и удобной для чтения.

Я могу выбрать один из двух вариантов: только правильность (доступно через 30 минут, 24 часа и 3 дня) или правильность и ясность (доступно через 6 часов, 24 часа или 3 дня).

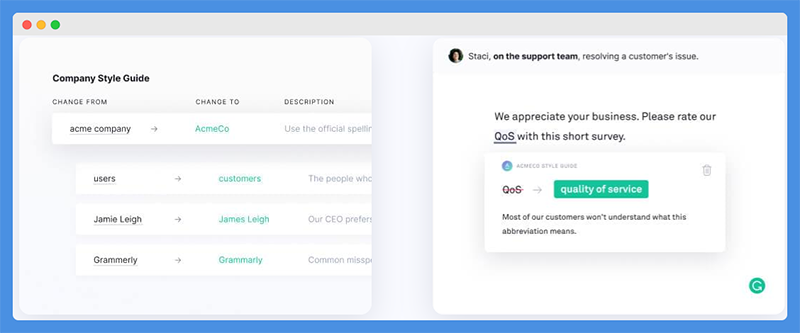
ГРАММАТИКА ДЛЯ MICROSOFT OFFICE



Когда я работаю над документом в Microsoft Office, иногда не все ошибки выделяются, пока я работаю или просматриваю документ. В таких случаях Grammarly для Microsoft Office очень удобен, так как он предлагает исправления, пока я пишу в Outlook или Word.

Как и LanguageTool, это настольное приложение, являющееся копией онлайн-редактора Grammarly, удобно, когда я не хочу заходить в Grammarly через браузер. Я также могу перетаскивать документы на значок Grammarly, чтобы мгновенно улучшить свой текст.

ГРАММАТИЧЕСКИЙ БИЗНЕС



Grammarly Business разработан для команд, и мне нравится, что я могу делиться этим инструментом со своими коллегами, когда мы работаем над документами. Таким образом, легко создавать тексты без ошибок, которые выглядят убедительно и производят впечатление на клиентов.

Этот инструмент предлагает в режиме реального времени рекомендации по стилю письма, читабельности, тону и выбору слов, а также проверяет наличие грамматических, орфографических и пунктуационных ошибок, чтобы каждое предложение было понятным и лаконичным.

Кроме того, как команда, мы можем согласовывать свою работу с руководством по стилю компании и разъяснять всем сотрудникам принципы взаимодействия с брендом. Этот инструмент работает на разных устройствах и платформах, включая Gmail, социальные сети, Google Docs, Salesforce и WordPress.

*OpenNMT, DeepL — гибридные системы перевода с подсказками.*

OpenNMT и DeepL — это эффективные системы машинного перевода, которые используют различные технологии, включая гибридные подходы.  
OpenNMT  
Это открытая платформа, которая поддерживает нейронные модели и может быть адаптирована для конкретных нужд. Она позволяет интегрировать подсказки, что улучшает качество перевода. Пользователи могут добавлять свои данные для обучения моделей, повышая точность в специфических областях.  
DeepL  
Занимает лидирующие позиции благодаря своим алгоритмам глубокого обучения и языковым моделям. DeepL использует контекстуальные подсказки для повышения качества перевода, что делает его более естественным. Пользователи оценили его способность передавать нюансы и стилистические особенности.  
Обе системы продолжают развиваться, внедряя новые технологии и методы, чтобы улучшить качество и адаптацию переводов.

1. **Перспективы улучшения качества машинного перевода через редактирование**

Улучшение качества машинного перевода (МТ) через редактирование представляет собой многообещающий подход, способствующий значительному повышению точности и естественности переведенных текстов. Традиционные системы МТ часто сталкиваются с проблемами сохранения смысла оригинала, адекватности передачи стилистики и культурных особенностей языка. Однако использование постредактирования человеком и автоматическое редактирование открывает новые перспективы для решения этих проблем.  
Рассмотрим основные направления развития технологий улучшения качества машинного перевода через редактирование:

*1. Автоматизированная поддержка редактора*

Один из ключевых аспектов заключается в создании инструментов поддержки редактирования, которые облегчают работу человека-переводчика. Такие инструменты предлагают варианты исправлений ошибок, выявленных системой машинного перевода, подсказывают возможные альтернативы лексических единиц и конструкций, сокращая временные затраты и повышая производительность труда.  
Примеры реализации включают системы автоматической проверки грамматики и стиля, адаптивные интерфейсы редактирования и средства автоматического сопоставления предложений оригинального и целевого текстов.

*2. Глубокая интеграция NLP-технологий*Интеграция передовых моделей обработки естественного языка (NLP), таких как трансформеры (например, BERT, GPT), позволяет решать проблемы анализа контекста, понимания семантики и выбора правильных форм выражения мыслей. Использование глубоких нейронных сетей обеспечивает возможность учитывать культурные аспекты, метафоры и идиомы, что повышает точность перевода сложных высказываний.  
Пример: система машинного перевода на основе модели Transformer учитывает связи между словами в предложении и генерирует переводы, соответствующие нормам русского языка.

*3. Адаптация к специфике предметных областей*Важным направлением является разработка специализированных решений для конкретных отраслей, таких как медицина, юриспруденция, наука и техника. Машинный перевод, настроенный на узкую область, способен обеспечить высокий уровень терминологической согласованности и соответствия стилю профессионального общения.  
Например, специализированная медицинская система машинного перевода должна уметь точно передавать значения терминов, относящихся к заболеваниям, лекарствам и процедурам лечения.

*4. Активное обучение на примерах человеческого редактирования*Алгоритмы активного обучения позволяют автоматически обновлять систему машинного перевода на основе внесённых людьми поправок. Когда человек вручную правит полученный машинный перевод, изменения фиксируются и используются для тренировки новых версий модели. Таким образом, система постепенно совершенствуется, адаптируясь к предпочтениям пользователей.

*5. Интерактивные гибридные технологии*Современные гибридные системы объединяют возможности автоматизированного перевода и ручного вмешательства. Они предоставляют человеку удобный интерфейс для интерактивного взаимодействия с результатами перевода, позволяя быстро вносить необходимые изменения и получать обратную связь.  
Одним из перспективных направлений здесь являются виртуальные ассистенты, предлагающие рекомендации по улучшению переводов и предоставляющие расширенные сведения о возможных вариантах выражений.

*Заключение*  
Подходы к улучшению качества машинного перевода через редактирование открывают большие возможности для дальнейшего прогресса в области межъязыковой коммуникации. Современные технологии глубокого обучения, интеграции NLP-моделей и интерактивных интерфейсов способствуют созданию эффективных решений, способных успешно справляться с трудностями языкового перевода и обеспечивать высокое качество результатов.

**Вопросы**

1. Что такое предредактирование текста и какие задачи оно решает?
2. Какие методы предредактирования могут быть использованы для улучшения качества машинного перевода?
3. В чем заключается процесс постредактирования и почему он важен для машинного перевода?
4. Какие подходы к постредактированию текста существуют? Приведите примеры.
5. Какую роль играют лингвисты и переводчики в процессе редактирования машинного перевода?
6. Какие автоматические системы используются для редактирования текста? Назовите их преимущества и недостатки.
7. Как работают гибридные системы перевода, такие как OpenNMT и DeepL?
8. Какие перспективы улучшения качества машинного перевода через редактирование существуют?
9. Как NLP-технологии, такие как BERT и GPT, влияют на качество машинного перевода?
10. Какие инструменты поддержки редактора могут быть использованы для автоматизации процесса редактирования?

**Проверочные задания**

|  |  |
| --- | --- |
| **Термин** | **Определение** |
| 1. Предредактирование | A. Процесс улучшения машинного перевода путем внесения изменений в полученный текст. |
| 2. Постредактирование | B. Использование специалистов для анализа, исправления ошибок и улучшения стиля машинного перевода. |
| 3. Лингвистическая экспертиза | C. Процесс изменения исходного текста для упрощения задачи машинного перевода и повышения его качества. |
| 4. Автоматическая проверка грамматики | D. Программа или инструмент, который автоматически обнаруживает грамматические и стилистические ошибки в тексте. |

1. Соотнесите термин с определением **Ответ:** (1-C, 2-A, 3-B, 4-D)
2. В каждой строчке найдите лишнее слово.

* а) Предварительное редактирование, б) Последующее редактирование, в) Перевод, г) Алгоритм BLEU
* а) Лингвист, б) Переводчик, в) Программист, г) Редактор
* а) LanguageTool, б) Grammarly, в) OpenNMT, г) Microsoft Word

**Ответ:** г) Алгоритм BLEU (это метрика оценки, а не этап подготовки/редактирования текста).

в) Программист (остальные - специалисты, работающие с текстом).

г) Microsoft Word (это текстовый редактор, остальные - инструменты редактирования текста).

3.  Задания на установление последовательности:

1. Анализ результатов машинного перевода
2. Внесение изменений в текст
3. Оценка качества отредактированного текста
4. Выявление ошибок и неточностей

**Ответ:** 1, 4, 2, 3

4. Определите, является ли утверждение верным или неверным.

1. Предварительное редактирование всегда необходимо для достижения высокого качества машинного перевода. (Верно/Неверно)
2. Автоматические системы полностью заменяют необходимость участия человека в процессе редактирования. (Верно/Неверно)
3. BLEU — это метрика, используемая для оценки качества предварительного редактирования. (Верно/Неверно)

**Ответы:**Неверно (зависит от сложности текста и требований к переводу), Неверно, Неверно (BLEU оценивает качество перевода)

5.Определите, являются ли следующие утверждения верными или неверными.

1. Постредактирование может включать исправление стилистических ошибок.  
   **Ответ:** Верно.
2. Гибридные системы перевода, такие как DeepL, не требуют участия человека.  
   **Ответ:** Неверно.
3. NLP-технологии, такие как BERT, используются только для машинного перевода, но не для редактирования.  
   **Ответ:** Неверно.
4. Лингвисты играют ключевую роль в локализации текста для целевой аудитории.  
   **Ответ:** Верно.

**Практическая часть**

1. Оценить влияние пред- и постредактирования на качество машинного перевода с использованием автоматических инструментов.

Три текстовых фрагмента (5-7 предложений каждый) разной сложности:

Фрагмент 1: Простой текст (например, описание товара).

Фрагмент 2: Текст средней сложности (например, статья из новостей).

Фрагмент 3: Сложный текст (например, отрывок из научной статьи или юридического документа).

**Фрагмент 1**

Представляем вам набор мужских футболок из хлопка Smite, который станет идеальным дополнением к вашему стильному и комфортному гардеробу! Особенностью наших футболок является плотность 180 гр/м2, а также высокое качество. Наши футболки мужские смайт изготовлены из мягкой и эластичной ткани, что обеспечивает максимальный комфорт в любое время года. Стильный однотонный дизайн с удлиненным силуэтом делают футболки модными и современными. Базовые цвета в комплекте, позволят собирать разные образы каждый день. Прямой крой и широкие плечи придадут вашему образу брутальность и стиль.

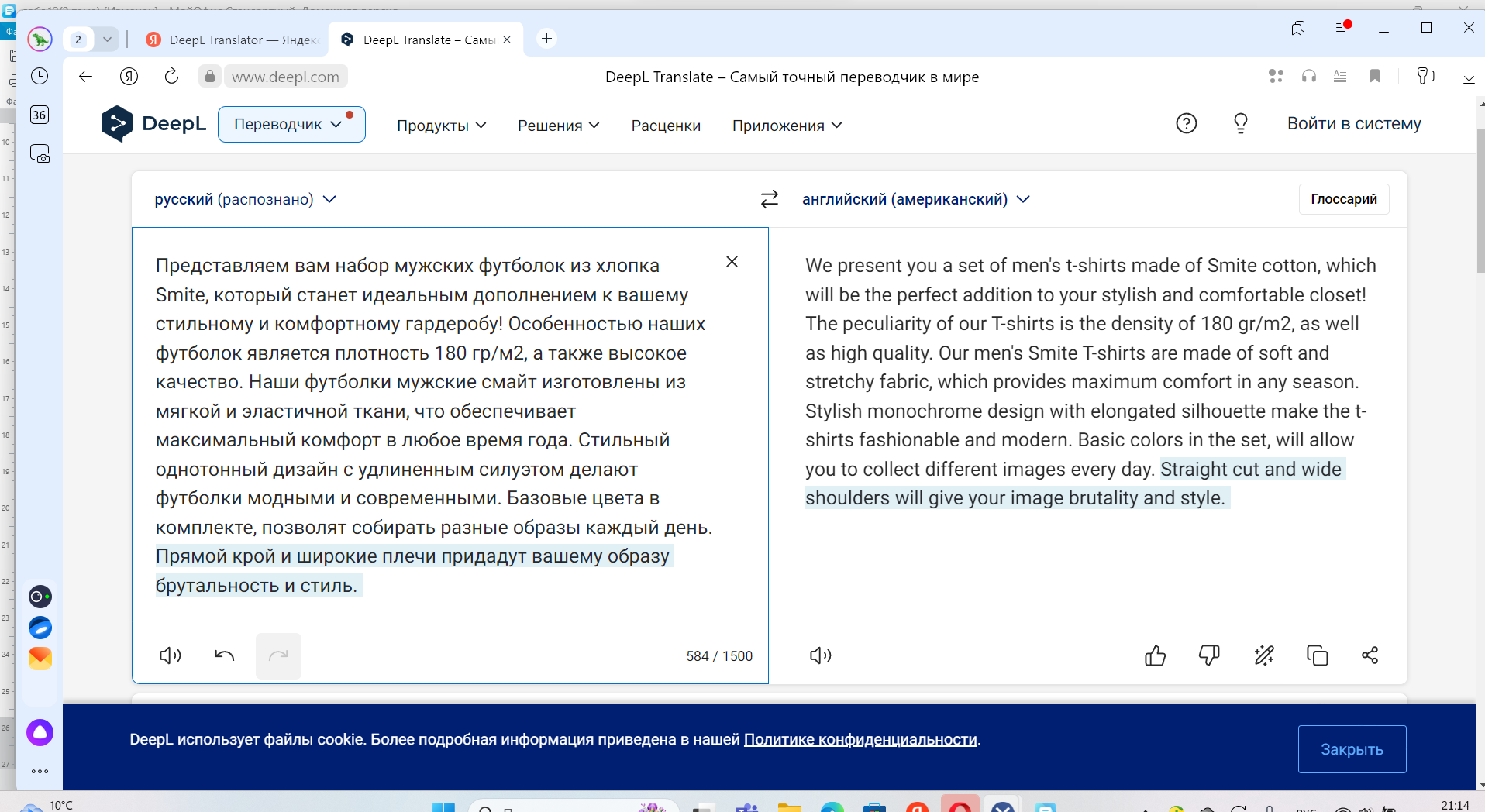
**Фрагмент 2**

Проезд по улице Софьи Перовской в Воронеже закроют до 11 июня. Об этом сообщил мэр города Сергей Петрин в понедельник, 12 мая. Ограничение начнет действовать в 0:00 вторника, 13 мая.Как сообщало [РИА «Воронеж»](https://riavrn.ru/news/v-voronezhe-s-12-po-15-maya-zakroyut-dvizhenie-po-ulice-sofi-perovskoj/), изначально планировалось, что проезд будет закрыт с 12 по 15 мая от улицы Чернышевского до Старомосковского переулка. Однако теперь его закроют от улицы Чернышевского до улицы Декабристов. Корректировки были внесены, так как второй этап устройства инженерных сетей для Петровской набережной конструктивно более сложный, чем первый, вдвое протяженнее и включает большое количество различных коммуникаций.

**Фрагмент 3**

При поступлении локально на левой ягодичной области отмечался обширный некроз кожи размерами 24х30 см, формирующийся некроз кожи по задненаружной поверхности левого бедра и в левой паховой области. Пальпаторно флюктуации не отмечалось. После осмотра была госпитализирована в отделении реанимации и интенсивной терапии для дообследования и дальнейшего лечения. При КТ органов брюшной полости и малого таза от 19.12.24г. Заключение: КТ-признаки панникулита левой ягодичной области. Проведена некрэктомия кожи и мягких тканей левой ягодичной области, левого бедра, паховой области слева. Проведено бактериологическое исследование отделяемого и гистологическое исследование резецированного материала: на уровне эпителиального слоя зоны некроза со слабо выраженной нейтрофильной инфильтрацией, отек на уровне дермы, неравномерное кровенаполнение сосудов.

**Этап 1: Машинный перевод (без редактирования)**  
Берем первый фрагмент и переводим с помощью **DeepL Translator.**



We present you a set of men's t-shirts made of Smite cotton, which will be the perfect addition to your stylish and comfortable closet! The peculiarity of our T-shirts is the density of 180 gr/m2, as well as high quality. Our men's Smite T-shirts are made of soft and stretchy fabric, which provides maximum comfort in any season. Stylish monochrome design with elongated silhouette make the t-shirts fashionable and modern. Basic colors in the set, will allow you to collect different images every day. Straight cut and wide shoulders will give your image brutality and style.

**Этап 2: Предредактирование (до перевода)**

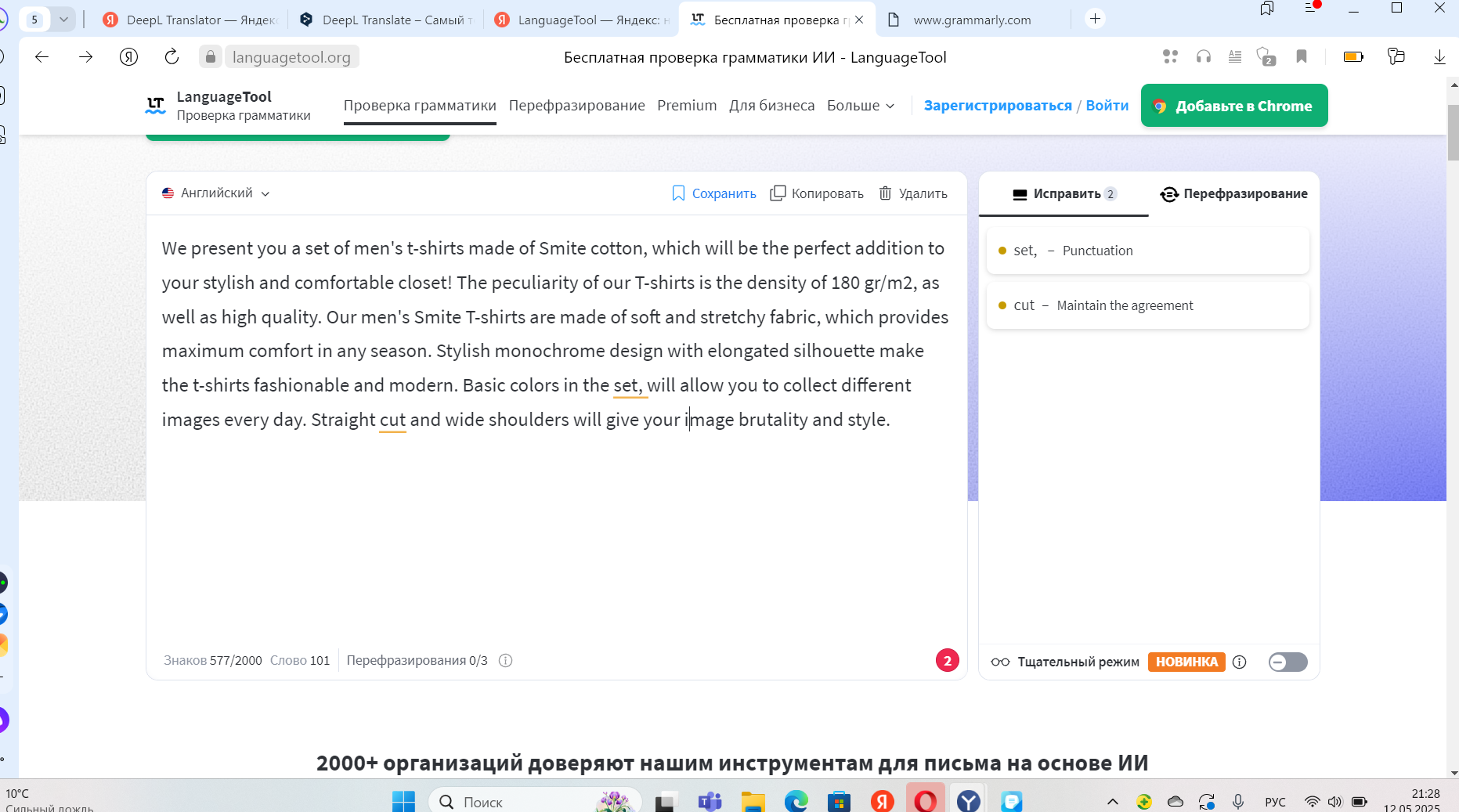
Берем наш первый фрагмент: Представляем вам набор мужских футболок из хлопка Smite, который станет идеальным дополнением к вашему стильному и комфортному гардеробу! Особенностью наших футболок является плотность 180 гр/м2, а также высокое качество. Наши футболки мужские смайт изготовлены из мягкой и эластичной ткани, что обеспечивает максимальный комфорт в любое время года. Стильный однотонный дизайн с удлиненным силуэтом делают футболки модными и современными. Базовые цвета в комплекте, позволят собирать разные образы каждый день. Прямой крой и широкие плечи придадут вашему образу брутальность и стиль.

Анализируем текст на наличие сложный конструкций, неоднозначных фраз, которые могут вызвать трудности при переводе и редактируем.  
Представляем вам набор мужских футболок из хлопка Smite, который станет идеальным дополнением к вашему стильному и комфортному гардеробу! Особенностью наших футболок является плотность 180 гр/м2, а также высокое качество. Наши футболки мужские Smite изготовлены из мягкой и эластичной ткани, что обеспечивает максимальный комфорт в любое время года. Стильный однотонный дизайн с удлиненным силуэтом делают футболки модными и современными. Базовые цвета в комплекте, позволят собирать разные образы каждый день. Прямой крой и широкие плечи придадут вашему образу брутальность и стиль.

В данном тексте тексте нет конструкций, которые могут вызвать трудности при переводе. Потому перевод текста остается тем же: We present you a set of men's t-shirts made of Smite cotton, which will be the perfect addition to your stylish and comfortable closet! The peculiarity of our T-shirts is the density of 180 gr/m2, as well as high quality. Our men's Smite T-shirts are made of soft and stretchy fabric, which provides maximum comfort in any season. Stylish monochrome design with elongated silhouette make the t-shirts fashionable and modern. Basic colors in the set, will allow you to collect different images every day. Straight cut and wide shoulders will give your image brutality and style.

**Этап 3: Постредактирование (после перевода)**

Загружаем наш текст в LanguageTool и смотрим какие ошибки нужно исправить.



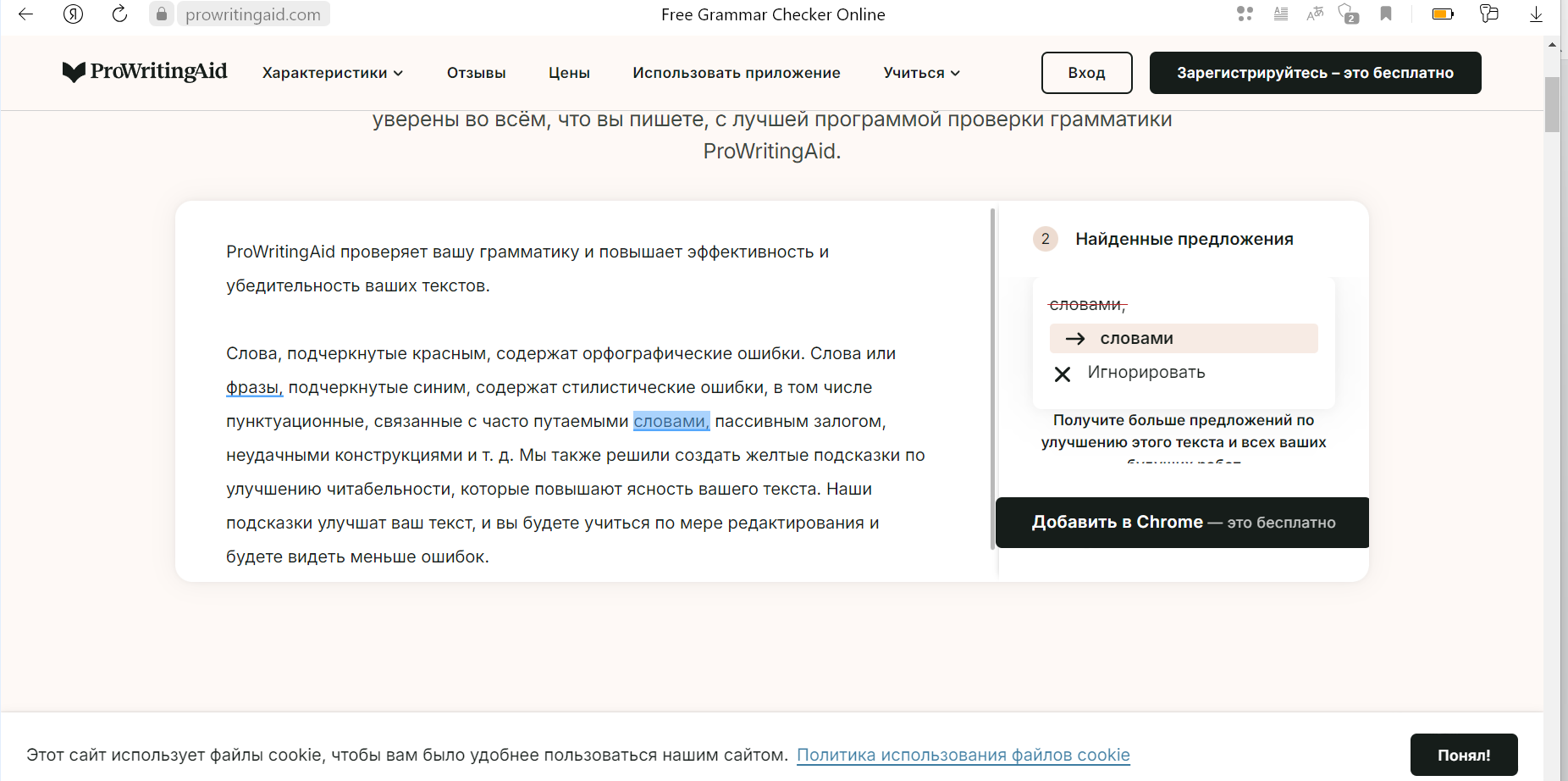
Ошибки, предлагаемые приложением, некорректны. Текст имеет тот же перевод.

**Этап 4: Сравнение и анализ**

Сравните три варианта перевода (машинный перевод без редактирования, машинный перевод после предварительного редактирования и машинный перевод после последующего редактирования) для оставшихся двух фрагментов.

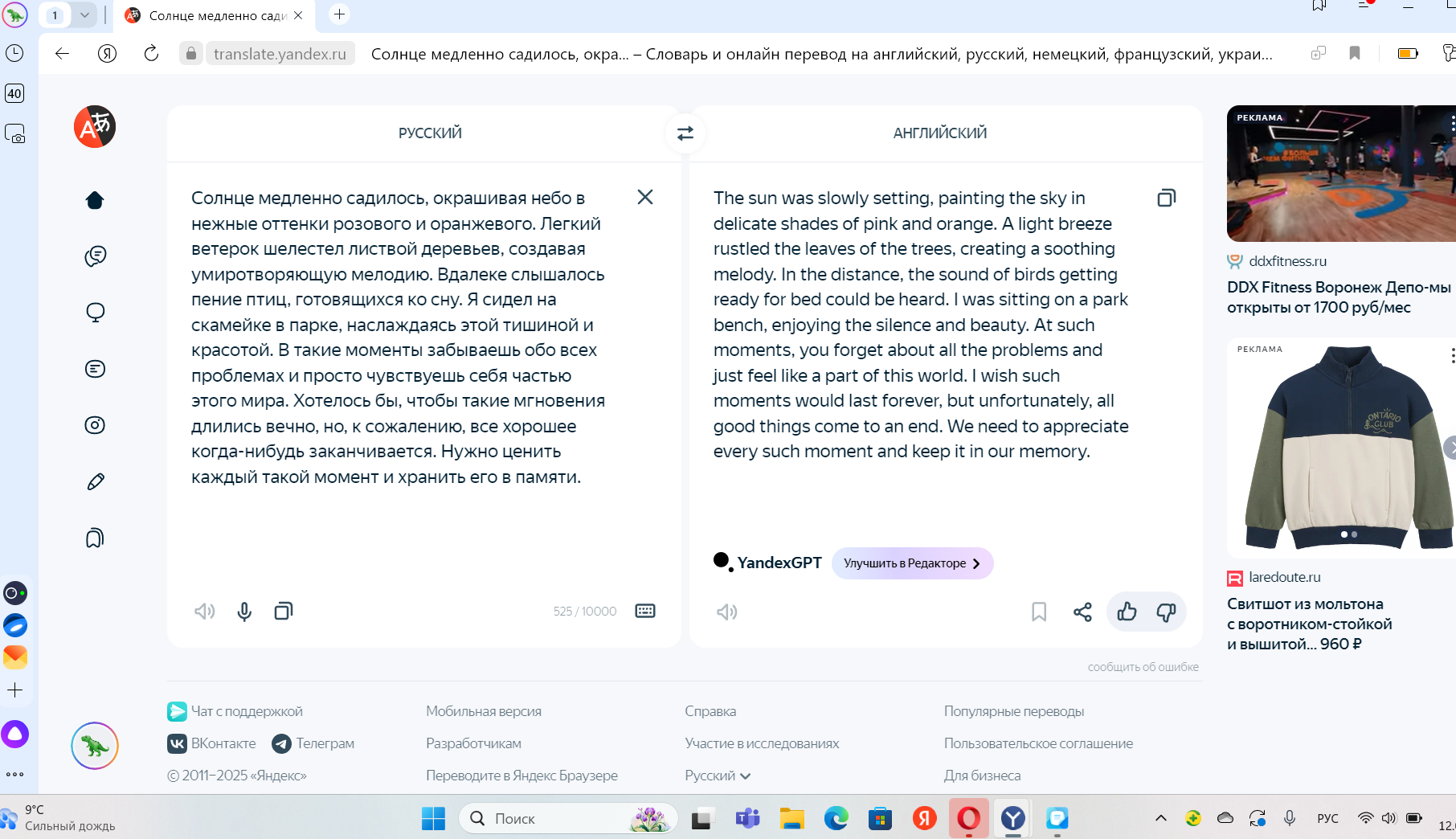
1. Выберите фрагмент русскоязычного текста среднего уровня сложности, выполните автоматический перевод, применяя программу автоматической корекции грамматики и пунктуации в ProWritingAi.

Знакомство с программой.



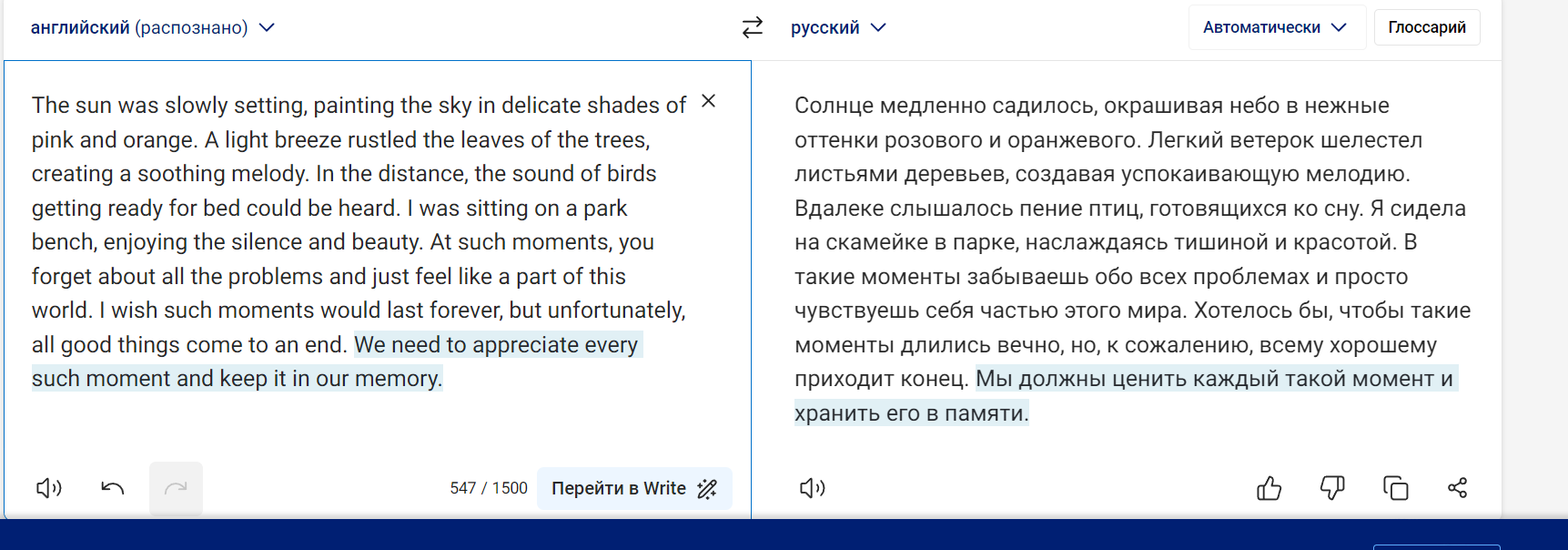
Выберем текст:

Солнце медленно садилось, окрашивая небо в нежные оттенки розового и оранжевого. Легкий ветерок шелестел листвой деревьев, создавая умиротворяющую мелодию. Вдалеке слышалось пение птиц, готовящихся ко сну. Я сидел на скамейке в парке, наслаждаясь этой тишиной и красотой. В такие моменты забываешь обо всех проблемах и просто чувствуешь себя частью этого мира. Хотелось бы, чтобы такие мгновения длились вечно, но, к сожалению, все хорошее когда-нибудь заканчивается. Нужно ценить каждый такой момент и хранить его в памяти.  
Переведем с помощью Яндекс переводчик:

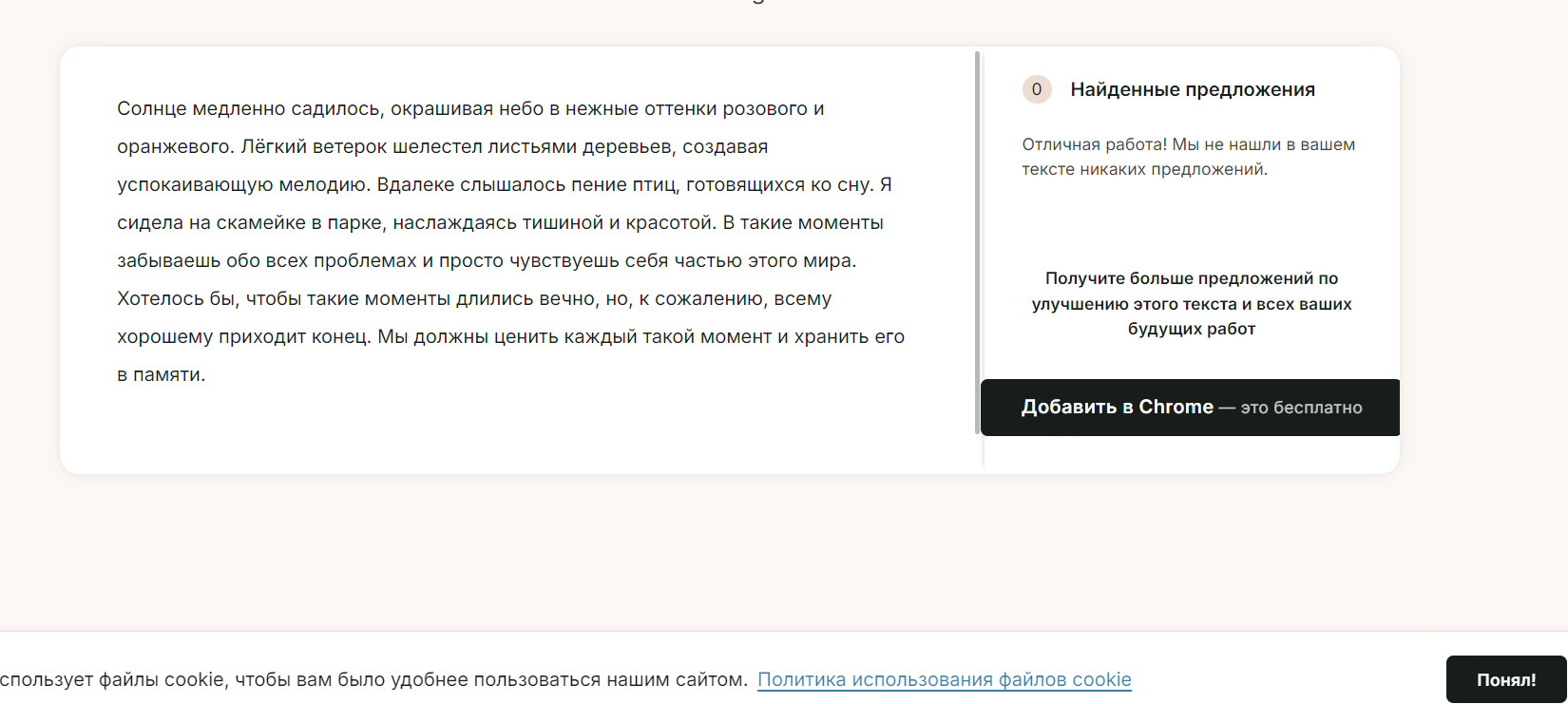


The sun was slowly setting, painting the sky in delicate shades of pink and orange. A light breeze rustled the leaves of the trees, creating a soothing melody. In the distance, the sound of birds getting ready for bed could be heard. I was sitting on a park bench, enjoying the silence and beauty. At such moments, you forget about all the problems and just feel like a part of this world. I wish such moments would last forever, but unfortunately, all good things come to an end. We need to appreciate every such moment and keep it in our memory.

Возьмем наш перевод и вставим в окошко **DeepL Translator.**



И теперь наш текст отправим в <https://prowritingaid.com/grammar-checker> , ошибок не найдено.



Проделайте тоже самое с вашим текстом.

**Источники:** <https://dialogue-conf.org/media/5881/zhivotovaaaplusberdonosovvd034.pdf> <https://belca.isu.ru/pluginfile.php/201907/mod_resource/content/1/Постредактирование%20МП.pdf> <https://habr.com/ru/articles/852810/>