

– ТЕОРИЯ И ПРАКТИКА ПЕРЕВОДА В АЛТАЙСКИХ ЯЗЫКАХ –

УДК 81'322.4

<https://doi.org/10.25587/2782-6627-2025-2-39-52>*Оригинальная научная статья*

Анализ ошибок машинного перевода текстов разных жанров в паре «русский и якутский языки»

А. А. Васильева✉, О. В. Бурцева

Северо-Восточный федеральный университет им. М.К. Аммосова,
г. Якутск, Российская Федерация✉ vaaperevod@mail.ru, olgamekbur@mail.ru

Аннотация

В статье на примере трех отрывков из текстов разных жанров анализируются ошибки, допущенные онлайн-переводчиками Яндекс.Переводчик и Google Translate в паре «русский и якутский языки». Якутский язык в онлайн-переводчики был внедрен несколько лет назад, и массовое их использование только начинается. Анализ ошибок машинного перевода в паре «русский и якутский языки» может способствовать выработке рекомендаций для эффективного обучения и машинных систем, и переводчиков-людей, прежде всего постредактированию автоматизированного перевода. Поэтому на данном этапе развития теории и практики русско-якутского перевода выработка классификации ошибок машинного перевода, представленная в данной статье, является весьма актуальной. В экспериментальных целях с помощью указанных сервисов были сделаны переводы следующих текстов: а) отрывок из научного доклада на русском языке; б) отрывок из перевода Библии на якутском языке (обратный перевод на русский язык); в) стихотворение на якутском языке на тему нравственности. Классификация ошибок машинного перевода в этой паре языков основывается на существующих в других парах языков типологиях: а) лексико-семантические и лексико-стилистические ошибки; б) нормативно-узусальные ошибки; в) грамматические ошибки; г) орфографические ошибки; д) пунктуационные ошибки. Описывается характер ошибок и поясняются причины их возникновения. Основными причинами ошибок оказываются: а) разный объем заложенных в машины исходных данных для осуществления анализа языка текста, предназначенного к переводу; б) специфика сходств и различий контактирующих языков, связанная с полисемией и грамматическими лагунами. При сравнении качества переводов обоих сервисов отмечаются их сильные и слабые стороны, проявленные при переводе с якутским языком. В конце статьи делаются выводы о совершенствовании лингвистических баз этих сервисов и о необходимости приобретения соответствующих новых компетенций профессиональными переводчиками, работающими в паре «русский и якутский языки».

Ключевые слова: машинный перевод, русский язык, якутский язык, переводческие ошибки, лексика, грамматика, искажение, неточность, Google Translate, Яндекс.Переводчик, постредактирование машинного перевода

Финансирование. Исследование не имело финансовой поддержки

Для цитирования: Васильева А. А., Бурцева О. В. Анализ ошибок машинного перевода текстов разных жанров в паре «русский и якутский языки». *Алтаистика. Altaistics*. 2025, № 2 (17), С. 39-52. <https://doi.org/10.25587/2782-6627-2025-2-39-52>

Analyzing machine translation errors in texts of different genres in the Yakut-Russian language pair

Akulina A. Vasilieva✉, Olga V. Burtseva

M.K. Ammosov North-Eastern Federal University, Yakutsk, Russian Federation

✉vaaperevod@mail.ru, olgamekbur@mail.ru

Abstract

The article analyzes errors of web translators Google Translate and Yandex Translate in the Yakut-Russian language pair in three texts of different genres. The Yakut language was introduced to web translators several years ago, and their mainstream use is at an early stage. Analyzing machine translation in the Yakut-Russian language pair can contribute to the development of recommendations for the effective training of both machine systems and human translators, primarily in machine translation post-editing. Thus, the classification of machine translation errors in this article is relevant in the current development stage of the Yakut-Russian theory and practice. For the experiment's purposes, the following texts were translated via the specified services: a) an excerpt from a scientific paper in Russian; b) an excerpt from a translation of the Bible into Yakut (reverse translation into Russian); c) a poem in Yakut on morality. The classification of machine translation errors in this language pair is based on typologies of other language pairs: a) lexico-semantic and lexico-stylistic errors; b) errors in standards and usage; c) grammar errors; d) spelling errors; e) punctuation errors. The nature of errors is described and the reasons for their occurrence are explained. The main causes of errors are: a) the difference in the volume of initial data in the machine systems for analyzing the language of the text being translated; b) the special nature of similarities and differences of contacting languages, associated with polysemanticity and grammar gaps. Upon comparing the quality of translations of both services, their strengths and weaknesses in translation with the Yakut language are noted. The article concludes that the linguistic databases of these services need improvements, and human translators who work with the Yakut-Russian language pair need to acquire corresponding expertise.

Keywords: machine translation, Yakut language, Russian language, translational errors, vocabulary, grammar, alterations, infelicity, Google Translate, Yandex Translate, machine translation post-editing

Funding. The study had not financial support

For citation: Vasilieva A.A., Burtseva O.V. Analyzing machine translation errors in texts of different genres in the Yakut-Russian language pair. *Altistics*. 2025, No 2 (17), Pp. 39-52. <https://doi.org/10.25587/2782-6627-2025-2-39-52>

Введение

Из множества имеющихся сервисов по переводу текстов якутский язык используется в Google Translate и Яндекс.Переводчик. Эти два сервиса представляют собой ведущие платформы для машинного перевода, каждая из которых имеет свои уникальные особенности и подходы. История создания Google Translate и Яндекс.Переводчика весьма интересна и отражает эволюцию технологий машинного перевода [1].

Google Translate – это бесплатный сервис, разработанный компанией Google в апреле 2006 г. Изначально сервис использовал статистический машинный перевод (далее – МП), основанный на анализе больших объемов параллельных текстов. Статистический подход позволял компьютеру находить закономерности в текстах и делать предположения о переводе, основываясь на частоте встречаемости тех или иных словосочетаний. В ноябре 2016 г. Google представил революционную технологию – нейронный МП (NMT). Вместо традиционного дословного перевода NMT использует глубокую нейронную сеть, которая обрабатывает целые предложения одновременно, учитывая контекст. Это позволило значительно улучшить качество перевода, особенно для длинных и сложных предложений. В настоящее время Google Translate обеспечивает перевод на более чем 240 языков. Якутский язык в него был загружен в 2024 г. в составе последнего крупнейшего обновления [2].

Яндекс.Переводчик был запущен в 2011 г. и изначально использовал статистический подход, аналогичный Google Translate. Основной целью разработки было создание эффективного инструмента для перевода текстов на русском языке и с русского на другие языки. В отличие от Google, Яндекс сделал акцент на адаптивность и локализацию. Команда разработчиков активно работала над улучшением качества перевода для русского языка и языков постсоветского пространства. Это включало использование собственных корпусов текстов и развитие уникальных алгоритмов, адаптированных под особенности этих языков. В последние годы Яндекс.Переводчик перешел на использование нейронных сетей, что повысило качество перевода. Сервис также внедрил функцию персонализации, позволяющую пользователям создавать собственные словари и настраивать предпочтения для определенных тематик. Якутский язык в Яндекс.Переводчик был внедрен в 2020 г. [3].

Эти два сервиса представляют собой ведущие платформы для машинного перевода, каждая из которых имеет свои уникальные особенности и подходы. Многие исследователи и языковые активисты убеждены, что с помощью систем машинного перевода, в том числе развивающихся алгоритмов распознавания и синтеза речи, повышаются шансы на сохранение исчезающих языков, что определяет актуальность лингвистического исследования продукции машинного перевода, актуальность выработки практических рекомендаций для обучения искусственного интеллекта (далее – ИИ), а также подготовки переводчиков со специальными навыками постредактирования машинного перевода.

Для оценки качества перевода немаловажной является типология переводческих ошибок с учетом специфики МП. Так, Т.А. Иванченко отмечает, что «Причины ошибок в переводах систем машинного перевода можно связать, с одной стороны, со спецификой естественных языков и речевых произведений, порождаемых на этих языках, с другой – с несовершенством алгоритмов работы систем машинного перевода в настоящее время, которые не учитывают эту специфику и не способны, в отличие от человека, выдавать приемлемый по качеству перевод, не требующий дальнейшего редактирования» [4, с. 32]. Однако эти алгоритмы задаются человеком, поэтому для обучения ИИ в системах машинного перевода с якутским языком необходимо сотрудничество программистов и лингвистов, которые сообщая могли бы приспособить результаты сопоставительных исследований русского и якутского языков для улучшения качества машинного перевода.

За многие десятилетия развития теории перевода было выработано много разных типологий переводческих ошибок, которые используются не только для описания самого процесса перевода, но и в дидактике перевода. Обзорам типологий переводческих ошибок посвящены главы в монографиях и отдельные статьи исследователей как машинного, так и человеческого перевода [5–8].

В теории русско-якутского перевода была выработана следующая типология переводческих ошибок, уходящая корнями в типологии, разработанные Р. К. Миньяром-Белоручевым (1988), Л. К. Латышевым (2000), Н. К. Гарбовским (2004): 1. Ошибки (итэбэс, т. е. недостатки) содержания текста перевода: а) искажение (токурутуу); б) неточность (халбан тылбаас); в) неясность (мунаах тылбаас); г) потеря информации (көтүтүү); 2. Ошибки (итэбэс, т. е. недостатки) языка перевода: а) игнорирование особенностей языковой системы (тыл систиэмэтин кэһэр); б) нарушение литературных норм (тыл литературнай нуорматын кэһэр); в) нарушение узуса (тыл үөрүйэбин кэһэр). Эти ошибки рассматриваются в вузовском учебном пособии в связке с их причинами: содержательные ошибки возникают из-за недостаточной компетентности переводчика, т. е. имеют субъективный характер, а языковые и речевые ошибки возникают не только из-за недостаточной компетентности переводчика, но и по условно объективным причинам, например, из-за системных расхождений русского и якутского языков (отсутствие того или

инного грамматического явления или их функциональное различие, узуальные особенности и т.п.) [9].

В данной типологии внимание уделяется причинам возникновения переводческих ошибок с тем, чтобы будущий переводчик понимал, как нельзя переводить и почему. Поэтому типология ошибок при русско-якутском переводе, разработанная в дидактических целях вузовской подготовки переводчиков с якутским языком, полностью соответствует следующему обобщению Е. А. Княжевой: «В классических работах по переводоведению, выполненных в теоретическом ключе, интерес к переводческим ошибкам зачастую связан с изучением проблем асимметрии языковых систем, анализом причин и характера девиатов и т. д. Исследования прикладного характера, как правило, ориентированы на оптимизацию оценки переводов в дидактических или узкопрофессиональных целях и поэтому соответствующие классификации являются очень детальными» [5, с. 190].

Лингвисты, исследующие МП, также обращаются к типологиям переводческих ошибок, адаптируют их для того, чтобы анализировать ошибки МП и улучшать его качество [6, 10–13]. Из выводов указанных работ для оценки качества МП с якутским языком актуальны следующие: 1) «Машинный перевод позволяет существенно сэкономить время и, в целом, современные системы нейронного перевода прекрасно справляются с задачей передачи общего смысла сообщения. Однако, они еще не достигли того уровня совершенства, когда участие человека из переводческого процесса полностью исключается. Именно по этой причине сегодня набирает популярность профессия постредактора: человека, который «доводит перевод до приемлемого уровня» [10, с. 52]; 2) для систем МП с русским языком, в частности с немецкого на русский, характерны следующие типы ошибок: «лексико-семантические и лексико-стилистические ошибки; нормативно-узуальные ошибки, то есть ошибки, связанные с нарушением речевой нормы языка перевода; грамматические ошибки; пунктуационные ошибки; орфографические ошибки» [4, с. 33]; 3) «наиболее частотными оказываются лексико-семантические ошибки. Относительно редкими в современных системах МП являются ошибки в орфографии и пунктуации и ошибки, связанные с опущением отдельных слов или фрагментов текста» [6, с. 18]; 4) «У различных ошибок разный «вес» с точки зрения их способности приводить к коммуникативному сбою. Наиболее значимыми являются ошибки, относящиеся к группе лексико-семантических» [6, с. 18]. С увеличением количества исследований, посвященных качеству МП с якутским языком, эти выводы могут быть скорректированы с учетом специфики русского и якутского языков, и объема баз данных, на которых основывается МП в указанной паре языков.

Материалы и методы

Для оценки качества перевода Google Translate и Яндекс.Переводчик были использованы следующие категории текстов: а) текст № 1 – отрывок из научного доклада одного из авторов данной статьи, местами со сложными грамматическими конструкциями на русском языке; б) текст № 2 – отрывок из перевода Библии на якутском языке (был проделан обратный перевод на русский язык); в) текст № 3 – стихотворение на якутском языке на тему нравственности. Переводы были сделаны в феврале 2025 г., предварительные результаты эксперимента представлены на Всероссийской НПК в г. Якутске 14 марта 2025 г.

Для анализа возможностей двух онлайн-переводчиков, содержащих в себе якутский язык, была применена классификация переводческих ошибок МП из статьи Т. Иванченко, которая в том или ином виде также встречается в работах других авторов: а) лексико-семантические и лексико-стилистические ошибки; б) нормативно-узуальные ошибки; в) грамматические ошибки; г) пунктуационные ошибки; д) орфографические ошибки. Кроме этого, в данной статье сравниваются и пользовательские возможности сервисов.

Целью эксперимента стало определение способности сервисов справляться с различными уровнями сложности языкового материала. Критериями качества перевода сервисов были следующие ожидания:

1) сохранение семантической точности – сравнивается, насколько точно передается смысл оригинального текста, особенно в случаях с идиоматическими выражениями и культурными особенностями;

2) грамматическая корректность – сравнивается правильность построения предложений, согласование времен, родов и чисел, а также использование предлогов и союзов;

3) стилистическое соответствие – сравнивается, насколько стиль перевода соответствует стилю оригинала;

4) контекстуальная адаптация – определяется то, учитывают ли сервисы контекст при переводе многозначных слов и т. п.;

5) скорость обработки – сравнивается время, необходимое сервисам для выполнения перевода каждого текста. Это поможет определить их производительность и удобство для пользователя;

6) интерактивность и удобство интерфейса – сравнивается простота использования сервисов, наличие дополнительных функций (например, встроенные словари, возможность исправления ошибок вручную).

Результаты и обсуждение

Итак, текст № 1 – отрывок из доклада «Нравственные основы Библии для детей: перевод заповедей на язык саха», в котором наряду с мыслями о важности родного языка при формировании нравственных основ человека через библейские заповеди, упоминается другой механизм формирования светских законов в современном обществе.

Как представлено в табл. 1, перевод этого отрывка сервисом Google Translate получился более высокого качества, чем у Яндекс.Переводчик. Из 21 предложения 8 переведены корректно обоими сервисами, поэтому эти предложения не включены в таблицу. Можно сделать вывод, что перевод простых нераспространенных предложений онлайн-переводчиками выполняется без особого труда, на примере предложений: «Не произноси имени Господа Бога твоего напрасно»; «Почитай отца твоего и мать»; «Не убий»; «Не укради».

Таблица 1

Результат перевода сервисами

Table 1

Translation results by the services

№	Исходный текст №1	Яндекс.Переводчик	Google Translate
1	Установленные государством законодательные акты призваны обеспечить стабильность в обществе, а также защитить права человека посредством принятых норм морали.	Государство олохтоммут сокуон аакталара общества бигэ туруктаах буолууну хааччыйарга, ону таһынан моральный нуормаларынан киһи быраабын көмүскүүргэ ынырылынна.	Судаарыстыба олохтообут сокуоннай аакталара уопастыбаҕа туруктаах буолууну хааччыйарга, ону тэнгэ ылыллыбыт сизр-майгы нуормаларынан киһи быраабыгар көмүскэллээх буолаллар.
2	Правовые нормы регулируют поведение людей с помощью системы законов и подзаконных актов, принятых компетентными правотворческими органами.	Правовой нуормалар дьон сокуоннар тиһигин уонна сокуоннай аакталар көмөлөрүнэн ылыныллыбыт сокуоннар тиһиктэринэн бэрээдэктииллэр.	Сокуон нуормалара дьон быһыытын-майгытын сөптөөх сокуону онгорор уорганнар ылыммыт сокуоннарын уонна быраабылаларын систиэмэтинэн хонтуруоллууллар.

3	Они упорядочивают значительную область общественно-государственных отношений вместе с тем, они оставляют вне сферы своего влияния отношения, регулируемые религиозными нормами (зависть, верность, внебрачные отношения).	Кинилэр общественнай-государственной сыһаннаһылар ураты уобарастарын бэрээдэктээн, сыһаныгар сабыдыаллаан (ордугурбааһын, итэҕэл, эйбэстик, эйбэстик сыһаннаһы) салаллар.	Кинилэр уопастыбаннай-судаарыстыба сыһаннаһыларын улахан эйгэтин тупсараллар, ону кытта бэйэлэрин сабыдыалларын эйгэтин таһыгар хааллараллар итэҕэл нуормаларынан бэрээдэктээх сыһаннар (ымсыы, бэриниилээх буолуу, дыиз кэргэн таһынаабы дымалалар).
4	В то время, когда Десять заповедей Библии затрагивают все сферы человеческих взаимоотношений.	Библия уонча кэриэстэрэ киһи аймах сыһанын бары эйгэтин таарыйар кэмнэрэ.	Библия Уон Кэриэһин киһи-аймах сыһанын бары эйгэтигэр дьайар кэмигэр.
5	Краткое содержание 10 заповедей:	10 кэриэс кылгас ис хоһооно:	Уон ыйыы түмүгэ:
6	2. Не сотвори себе кумира.	2. Кумиры айыма.	2. Бэйэҕин кумир оностума.
7	4. Помни день субботний, чтобы святить его (день седьмой — суббота Господу).	4. Субуотаҕа кини Сибэтиэй күнэ (сэттис-Айыы тойон субуота күн).	4. Субботаны сибэтиэйдик тутуһарга өйдөө (сэттис күн – Айыы Тойон субуотата).
8	7. Не прелюбодействуй.	7. Көссүүлэспэккин.	7. Ойохтонума.
9	9. Не произноси ложного свидетельства на ближнего твоего.	9. Чугас дьоннор сымыйанан туоһулаама.	9. Чугас киһигин утары сымыйа туоһулаама.
10	10. Не желай ничего, что у ближнего твоего.	10. Чугас дьоннор тугу да баҕарбаккын.	10. Чугас киһин туохха да баарын баҕарыма.
11	Первые четыре заповеди показывают как проявлять любовь к Богу, последние шесть заключают в себе обязанности любви к ближнему (т.е. ко всем людям).	Бастакы түөрт кэриэһи танара тапталын быһыытынан көрдөрөллөр, тиһэх алта киһи чугастаабытыгар (т.е. дьонно барытыгар).	Бастакы түөрт кэриэс Танараҕа тапталы хайдах көрдөрөргө көрдөрөр, бүтүһүк алта кэриэһигэр чугастаабы киһини (ол аата бары дьону) таптыыр эбээһинэс баар.
12	В Новом Завете Иисус Христос разъяснил содержание Десяти заповедей сведя их к двум величайшим: о любви к Богу и о любви к ближнему.	Саҥа кэс тылга Иисус Христос уон кэриэстэрин ис хоһоонун икки улуу кэриэс икки сүнгэн суолталаабын быһаарбыта: Танараҕа таптал уонна чугастаабыга таптал туһунан.	Саҥа Кэс Тылга Иисус Христос Уон Кэс Тыл ис хоһоонун быһааран, кинилэри икки улахан Кэс Тылга кыччаппыт: Танараҕа таптал уонна чугастаабы киһиэхэ таптал.
13	Словами «возлюби Господа Бога твоего всем сердцем твоим, и всею душою твоею, и всем разумением твоим: сия есть первая и наибольшая заповедь; вторая же подобная ей: возлюби ближнего твоего, как самого себя». Иисус Христос не упразднил, а углубил и утвердил их важное значение.	«Айыы Тойон Танараҕын бүтүн сүрэххинэн таптаа, туох баар өйгүнэн-санаабынан; бу бастакы уонна саамай улахан кэриэс; оттон иккиһин манньык этэр: чугас дьоннун бэйэтин курдук таптаа». Иисус Христос үрдэппэтэҕэ, дирин суолталаабын бигэргэппитэ.	Тылларынан эттэххэ, «Айыы Тойон Танараҕын бүтүн сүрэххиттэн, бүтүн дууһабынан, бүтүн өйгүттэн-санаабыттан таптаа: бу бастакы уонна саамай улахан кэриэс, иккис кэриэс: Чугас киһигин бэйэҕин курдук таптаа». Иисус Христос суох онорботоҕо, кинилэр суолталарын дирингэтэн, бигэргэппитэ.

Далее обратим внимание на ошибки, выявленные при переводе остальных 13 предложений, представленных в табл. 1:

а) лексико-семантические и лексико-стилистические ошибки:

– искажения, вызванные пропуском перевода слов: в предложении №2 Яндекс.Переводчик пропустил перевод слова *поведение* ‘быһы-майгы’, в результате исказился смысл предложения, будто бы люди регулируют что-то, однако неправильный выбор падежей (грамматическая ошибка) усугубил ситуацию, в итоге перевод предложения оказался непонятным набором слов. В предложении № 11 пропущен перевод слова *как* ‘хайдах’, в результате любовь к Богу превратилась в любовь Божью;

– искажения, вызванные неправильным подбором синонимов, словарных соответствий: Яндекс.Переводчик в предложении № 1 использовал ‘ыгырылынна’ в качестве перевода слова *призван*, хотя по контексту соответствовало бы слово ‘анаммыт’ или ‘аналлаах’ (т. е. ‘предназначается, предназначен’). В предложении № 3 для слова *область* в смысле *сфера* подобран немотивированный перевод ‘убарас’ (‘образ’); для понятия *верность* подобрано слово ‘итэбэл’ (‘вера, религия’); а для понятия *внебрачные отношения* – словосочетание ‘эйэбэстик сыһыаннаһы’ (т. е. ‘вежливое отношение’). Несмотря на то, что перевод предложения грамматически построен правильно, большое количество неверных лексических соответствий превращает предложение в абсурд. Google Translate в данном отрывке допустил больше подобных ошибок. Например, в предложении №2 глагол *регулирует* переведен как ‘хонтуруоллуулар’, т. е. ‘контролируют’, а не ‘салайаллар’ (‘управляют’), или не ‘сааһылылар’ (‘упорядочивают’), как в Яндекс.Переводчике. В предложении № 13 глагол *не упразднил* Яндекс.Переводчик перевел как ‘үрдэппэтэбэ’ (т. е. ‘не возвысил’), в то время как Google Translate подобрал более подходящий глагол ‘суох онорбото’ (т. е. ‘не уничтожил / не исключил’);

– неясности, вызванные пропуском перевода слов: неправильно переведенный Яндекс.Переводчиком глагол *призваны* в предложении № 1 в Google Translate пропущен, видимо, из-за грамматической формы перевода однородного глагола *защитить* ‘көмүскэллээх буолаллар’ в конструкции *призваны обеспечить..., а также защитить...* Таким образом, данную ошибку можно отнести и к грамматическим;

б) нормативно-узусальные ошибки: В предложении № 3 Google Translate допустил узусальные ошибки, вызванные дословным (пословным) переводом *улахан эйгэ* (т. е. ‘большая сфера’) вм. *киэн эйгэ* (т. е. ‘широкая сфера / область’) при переводе словосочетания *значительная область*; ‘дьэз кэргэн таһынаабы сыһыаннаһылар’ (т. е. ‘отношения вне семьи’) вм., например, ‘көссүүлэһии’ или ‘хос ойохтонуу’ при переводе словосочетания *внебрачные отношения*.

Кроме этого, важно отметить, что Google Translate использовал три синонима перевода слова *заповедь* на протяжении всего текста – ‘ыйыы’ (т. е. ‘указание’) в предложении № 5, ‘кэриэс’ (т. е. ‘завет, заповедь’) в предложениях № 11 и № 13, ‘Кэс Тыл’ (т. е. ‘Завет’) в предложении № 12, тогда как Яндекс.Переводчик использовал единственный вариант – ‘кэриэс’;

в) грамматические ошибки:

– неправильный подбор формы числа, несогласованность в числе: Яндекс.Переводчик в предложении № 2 – *сокуоннар тиһигин* вм. *сокуоннар тиһиктэрин* (согласование во мн. ч. – аффикс -лар имеется в обоих компонентах изафетного словосочетания) или *сокуон тиһигин* (согласование в ед. ч. обоих компонентов изафетного словосочетания). В предложениях №4 и № 12 Яндекс.Переводчик использовал форму мн. ч. существительного в словосочетании с числительным больше 1, что является нарушением литературной нормы. В предложении № 10 Яндекс.Переводчик немотивированно заменил форму ед.ч. *ближний твой* на форму мн. ч. ‘чугас дьоннор’ (т. е. ‘близкие люди’). Необходимо

отметить, что в данном сервисе форма мн. ч. используется слишком часто, вслед за оригиналом, тогда как по литературным и узуальным нормам якутского языка во многих случаях используется форма ед. ч., а значение множественности выражается различными лексическими и грамматическими способами. В этом отношении перевод Google Translate больше соответствует нормам якутского языка;

- неправильный подбор формы лица лично-возвратного местоимения. В предложении № 13 лично-возвратное местоимение от основы *бэйэ* стоит в форме III лица *бэйэтин* вместо формы II лица *бэйэбин*;

- неправильный подбор отрицательной формы глагола. В предложениях № 8 и № 10 вместо отрицательной формы повелительного наклонения II лица ед. ч. на *-ыма* Яндекс.Переводчик использовал отрицательную форму изъявительного наклонения настоящего-будущего времени II л., ед. ч. на *-баккын*, которая используется «в качестве одного из средств выражения пожелания к совершению действия» [14, с. 207], но не соответствует по категоричности оригиналу. Эти же предложения Google Translate перевел семантически неправильно, но грамматически верно;

- неправильный подбор эквивалентов падежно-предложных форм. В предложении № 6 возвратное местоимение *себе* стоит в дательном падеже. Оба онлайн-переводчика перевели это предложение неверно. Яндекс.Переводчик опустил местоимение, т. е. не перевел его, чем вызвал искажение смысла оригинала. Google Translate использовал форму винительного падежа ‘бэйэбин’ (т. е. ‘себя’) вместо дательного ‘бэйэбэр’ (т. е. ‘себе, для себя’), тем самым исказил смысл оригинала. В предложении № 9 Яндекс.Переводчик отразил формальное соответствие винительного падежа русского языка и дательного падежа якутского языка, проигнорировав, в отличие от Google Translate контекст оригинала, в котором предлогу *на* соответствует послелог ‘утары’ (т. е. ‘против’). Предложение № 10 обоими сервисами переведено неправильно из-за того, что онлайн-переводчики не смогли установить по контексту опущенное в оригинале слово *есть/имеется* («Не желай ничего, что [есть] у ближнего твоего»). В переводе на якутский язык переводчик-человек восстановил бы опущенное слово и перевел бы словом наличия ‘баар’ (т. е. ‘есть, имеется, существует’) в форме винительного падежа ‘баары’ (т. е. ‘то, что имеется’ (в вин. п.);

- нарушение порядка слов в предложении. В предложении № 2 Google Translate неправильно определил место слова ‘сөптөөх’ (т. е. ‘подходящие / соответствующие требованиям чего-либо’), выступающего как соответствие слова *компетентные*. Из-за этого возникла неясность в переводе, которую можно было бы устранить путем перемещения слова ‘сөптөөх’ ближе к определяемому им слову ‘уорганнар’ (т. е. ‘органы’);

г) орфографические ошибки. К орфографическим ошибкам можно отнести написание строчными буквами слов, которые традиционно пишутся с заглавной буквы в текстах данной тематики (Яндекс.Переводчик, предложения № 7, № 12). Сюда же можно отнести и случаи отказа от транскрибирования заимствуемых слов по фонетическим законам якутского языка (Яндекс.Переводчик, предложения № 1, № 2, № 3).

Следующий текст для сравнительного анализа – отрывок из перевода Библии «Аан дойду айыллыта» (‘Сотворение неба и земли’) [15, с. 9–10]. Для демонстрации в данной статье выбраны только те предложения, которые переведены некорректно. Результат перевода обоими системами представлен в таблице № 2.

Таблица 2

Результат перевода сервисами

Table 2

Translation results by the services

Исходный текст	Яндекс.Переводчик	Google Translate
6. Онтон Танара: «Уу ортотугар кытаанах арана баар буолан ууну икки аныгы араардын», – диэбитэ.	6. И сказал Бог: раздели воду на две части, потому что между ними есть твердый слой.	6. И сказал Бог: да будет твердь посреди воды, и да отделит она воду от воды.
7. Танара кытаанах арананы онгорон, арана аннынаабы ууну арана үрдүнээби ууттан араарбыта. Эппитин курдук буолбута.	7. Бог создал твердый слой и отделил воду, находящуюся под слоем, от воды, находящейся над слоем. Так и случилось.	7. И создал Бог твердь, и отделил воду, которая под твердью, от воды, которая над твердью: и стало так. Все произошло так, как он сказал.
8. Танара кытаанах арананы «халлаан» диэн ааттаабыта. Кизһэ буолбута, сарсыарда буолбута – иккис күн бүппүтэ.	8. Бог называл твердый слой «небом». Произошло это вечером, произошло утром – закончился второй день.	8. И назвал Бог твердь небом. И был вечер и утро дня второго.
9. Онтон Танара: «Халлаан аннынаабы уу бииргэ мустан, кураанах сир көһүннүн», – диэбитэ. Эппитин курдук буолбута.	9. И сказал Бог: пусть вода Под небом соберется вместе, и земля станет пустой. Так и случилось.	9. И сказал Бог: да соберётся вода, которая под небом, в одно место, и да явится суша. И стало так. Все произошло так, как он сказал.
10. Танара кураанах сири «сир», оттон мустубут ууну «байбал» диэн ааттаабыта. Итиэннэ Танара итини барытын көрөн сөбүлээбитэ.	10. Бог назвал сушу «землей», а воду «морем». И Бог, видя все это, согласился.	10. И назвал Бог сушу Землей, а собрание вод назвал Морями. И увидел Бог все, что Он создал, и вот, хорошо весьма.

Яндекс.Переводчик не смог передать торжественности стиля оригинала, и допустил много ошибок, многословие. К лексико-стилистическим ошибкам можно отнести следующие: *баар буолан* – ‘потому что между ними есть’; *кизһэ буолбута, сарсыарда буолбута* – ‘произошло это вечером, произошло утром’. Есть ошибки в подборе синонимичных эквивалентов: *кураанах сир* – ‘земля пустая’, вм. ‘сухая’; *сөбүлээбитэ* – ‘согласился’, вм. ‘понравилось/принял’, хотя в русском оригинале глаголу *сөбүлээбитэ* соответствует ‘[и увидел Бог], что это хорошо’. Также допущены грамматические ошибки: (*кени*) *араардын* (т. е. ‘(он) разделит+пусть’) (III л.) – ‘(ты) раздели’ (II л.); некорректно передана видо-временная форма глагола: *ааттаабыта* – ‘называл’ (несов. в.). В пяти предложениях перевода отсутствуют пунктуационные и орфографические ошибки.

В отличие от Яндекс.Переводчика, Google Translate смог сразу идентифицировать текст из Писания – перевод этого сервиса имеет почти полное соответствие тексту Синодального перевода Библии. Возможно, переводы Библии на разные языки содержатся в корпусе Google Translate в качестве параллельных текстов. Эти переводы привлекают внимание лингвистов именно благодаря своей эквивалентности. Например, известно, что в 2016 г. германскими лингвистами была создана база из переводов Библии на 1169 языках для автоматического анализа языковых маркеров грамматической категории времени [16]. В исследуемом якутско-русском переводе в предложениях №7 и №9 онлайн-переводчик добавил дословный перевод якутской фразы «Эппитин курдук буолбута» – ‘Все произошло

так, как он сказал'. На самом деле эта фраза является якутским пояснительным переводом фразы 'И стало так'. Из этого следует, что Google Translate стремится делать полный перевод всего заданного текста.

Также допущена пунктуационная ошибка в предложении №9 при передаче прямой речи – отсутствуют кавычки, прямая речь не начинается с заглавной буквы. Это также объясняется тем, что Google Translate идентифицировал Синодальный перевод и скопировал форму предложений. В предложении №10 имеется орфографическая ошибка – написание слов с заглавной буквы вместо строчной как в тексте-оригинале.

Третий текст для сравнительного анализа перевода Google Translate и Яндекс. Переводчик – стихотворение на якутском языке известной поэтессы Елизаветы Мигалкиной [17], в котором осуждаются человеческие грехи. В таблице №3 наряду с оригиналом и двумя переводами, сделанными машинами, в четвертой колонке дается подстрочный перевод, выполненный одним из авторов данной статьи. В подстрочном переводе в квадратных скобках даются слова, раскрывающие семантику и эмоциональные оттенки якутских слов, подобранных поэтессой с учетом гармонии звуков для выражения своего отношения к человеческим грехам.

Таблица 3

Результат перевода сервисами

Table 3

Translation results by the services

Исходный текст	Яндекс.Переводчик	Google Translate	Подстрочный перевод авторов статьи
Киһи барахсан	Человек	Человек почти	Человек [милый / слабый]
Орто дойдуга оһоһуллаҥ кэлиэбиттэн	С появлением средневековья	Так как он был создан в Среднем мире	С момента прихода в Срединный мир
Олуона да олобу	Личная жизнь	Даже жизнь	Странную [бестолковую] жизнь
Олорорго күһэллэр:	Вынужденный жить:	Их заставляют сидеть:	Проживать вынужден:
Суугунуур солкобо сууланар,	Обертывается в шелк,	Обернутый прохладным шелком,	В [дорогие] шелестящие шелка оборачивается,
Хачыгырас харчыга хараҥарар,	Темнеет на скрипучих деньгах,	Крик омрачен деньгами,	К шуршащим деньгам [одержимо] стремится,
Чалбаран аска чачайар,	Чалый едит,	Твои молитвы разбросаны по еде,	[Чрезмерно] изобильной едой давится,
Суон сурахха суудайар,	Судит на слух,	По данным Swan Learning,	Известным стать [тщеславно] пытается,
Сүүс сыл олоруон саантыыр,	Ему грозит до ста лет лишения свободы,	Он хочет прожить сто лет,	Желает сто лет прожить,
Аҕала сатаан араастаан саатыыр...	Он стыдится своего отца...	Его отец пытается отличить...	Развлекается как может...
Өссө таптал татакыйдаах,	Еще есть любовь,	А еще любовь привлекательна,	А еще любовь [плотская] обурекает,
Төрүүр-ууһуур айдааннаах...	Рождается и размножается...	Репродуктивное насилие...	Плодится и размножается [неукротимо]...

Сравнительный анализ показал, что оба сервиса сталкиваются с существенными трудностями при переводе художественных текстов, особенно стихотворений. Онлайн-переводчики оказались неспособными сохранить оригинальные смыслы и стилистику, что привело к значительным искажениям и неточностям вплоть до бессмысленности. Оба онлайн-переводчика запутались в омонимах (*олор* – ‘жить, сидеть, быть заключенным под стражу’) и в многозначных словах (*айдаан* – ‘шум, скандал’, в данном контексте – ‘неблаговидные занятия’ (значение №4 в Большом толковом словаре якутского языка, том 1, стр. 340)). В их словарях (лексических базах), видимо, не оказалось звукоподражательных глаголов (*суугунаа* – ‘шелестеть’, *хачыгыраа* – ‘хрустеть, шуршать’). Кроме того, неправильно интерпретировано деепричастное сочетание со служебным глаголом *аҕала сатаан* (т. е. ‘через силу, притворно’): данное словосочетание принято онлайн-переводчиками как отдельные слова. Так, слово *аҕала* (деепричастная форма от глагола *аҕал* – ‘приносить, приводить’) переведено как *отец* (т. е. ‘машина усмотрела корень’ *аҕа* – ‘отец’), а перевод ‘стыдится’, видимо, мотивировано служебным глаголом *сатаан* (т. е. вспомогательный глагол (деепричастие), обозначающий попытку совершить действие, выраженное деепричастием на *-а*) и глаголом *саатыыр* (т. е. ‘развлекается’), внешне похожих, но не однокоренных со словом *саат* (т. е. ‘стыд, позор’). Запрограммированность на полный перевод видна не только из выдачи неправильно проанализированной морфологии аффиксов и корней, но и выдачей немотивированной лексики, например, выражение «*суон сурахха суудайар*» (т. е. ‘пытается стать известным’, досл.: ‘толстой вести жаждать [о себе]’) в Google Translate заменено на ‘По данным Swan Learning’. Этот результат подчеркивает необходимость как обучения постредакторов-людей, так и необходимость дальнейшего совершенствования алгоритмов ИИ для работы с такими специфичными жанрами, как поэзия.

Оценка интерактивности и удобства интерфейса выявила практически одинаковые параметры работы, однако существенным отличием является скорость обработки текста. Так, Google Translate может сразу перевести текст, состоящий лишь из нескольких предложений, содержащих, как указано в сервисе, 3900 символов (на деле – еще меньше), в то время как Яндекс.Переводчик переводит текст любого размера. Соответственно, увеличивается время на перевод при работе с Google Translate, когда приходится разбивать текст на несколько отрывков, а потом объединять отрывки в цельный текст.

Заключение

На основе представленного анализа можно сделать следующие выводы. Основные типы ошибок машинного перевода для любых пар языков одинаковы – лексические, грамматические, узуальные, пунктуационные, орфографические. Судя по обзору статей по типологии ошибок машинного перевода с русским языком одинаковы, видимо, и специфические ошибки машины на уровне анализа входного текста и на уровне синтеза выходного текста. То есть совершенствование общих алгоритмов перевода обоих сервисов должно повысить качество обработки языкового материала. Наибольшее количество ошибок связано с лексикой и семантикой, выбором из ряда синонимов в соответствии с контекстом, и в паре «русский и якутский языки».

Вместе с тем уникальное соотношение сходств и различий контактирующих при переводе пар языков требует разработки и усовершенствования алгоритмов перевода текстов разных жанров в конкретных языках – в данном случае русско-якутского, якутско-русского перевода. Здесь необходимо согласованно работать и программистам, и лингвистам якутского языка. Так же важна задача обогащения словаря онлайн-переводчиков, т. е. расширение баз данных, корпусов, параллельных текстов. Качество перевода может

быть связано с содержанием корпусов параллельных текстов, что продемонстрировано в материале данной статьи при анализе обратного перевода отрывков из Библии.

Немаловажным представляется роль постредактора-человека, особенно при коррекции узуальных ошибок. В связи с тем, что якутский язык в онлайн-переводчиках появился совсем недавно, перед профессиональным сообществом переводчиков должна быть поставлена задача овладения техниками редакционной обработки машинного перевода, а также должно быть уделено внимание дидактике подготовки переводчиков с соответствующими профессиональными компетенциями.

Несмотря на существующие вызовы, МП якутского языка активно развивается благодаря современным технологиям ИИ. Улучшение качества перевода требует дальнейших исследований, расширения языковых корпусов и внедрения более сложных моделей обработки естественного языка. В перспективе автоматизированный перевод сможет не только облегчить коммуникацию, но и внести вклад в сохранение и развитие якутского языка.

Литература

1. Хохряков А. Виды и обзор движков машинного перевода: Яндекс переводчик, Google Translate и DeepL. <https://habr.com/ru/articles/852810/> (Дата обращения: 11.03.2025)
2. Лактюшин Н.В. «Гугл Переводчик» добавили 110 языков: среди них языки регионов России <https://hi-tech.mail.ru/news/111645-v-gugl-perevodchik-dobavili-110-yazykov-sredi-nih-yazyki-regionov-rossii/> (Дата обращения: 11.03.2025)
3. Борисов Н. Яндекс.Переводчик освоил якутский язык <https://ysia.ru/yandeks-perevodchik-osvoil-yakutskij-yazyk/> (Дата обращения: 11.03.2025)
4. Иванченко Т.А. Ошибки в машинном переводе с немецкого языка на русский (на материале статей немецкоязычных СМИ и текстов их переводов). *Ученые записки Санкт-Петербургского университета технологий управления и экономики*. 2021;22(4):30-41. DOI: 10.35854/2541-8106-2021-4-30-41
5. Княжева Е.А. *Оценка качества перевода: история, теория, практика*. Москва: Флинта; 2018:248.
6. Гребенюк А.В. Проблема типологии ошибок машинного перевода: стремление к универсальности vs таргетированность. *Вестник Московского университета*. 2024;22(2):7-25. DOI: 10.55959/MSU2074-6636-22-2024-17-2-7-25
7. Цзюньлин Гу, Чжунлянь Хуан. Система классификации переводческих ошибок. *Вестник Московского университета*. 2016;22(3):26-39.
8. Гончаров А.А., Бунтман Н.В., Нуриев В.А. Ошибки в машинном переводе: проблемы классификации. *Системы и средства информатики*. 2019;29(3):92-103. DOI: 10.14357/08696527190308
9. Петрова Т.И. *Тылбаас уопсай теорията*. Якутск: Издательский дом Якутского госуниверситета; 2007:38 (На якут. языке).
10. Шевчук Е.В., Никифорова Ж.А. Постредактирование и типичные ошибки в автоматизированном переводе научно-публицистических текстов. *Вопросы методики преподавания в вузе*. 2021;10(39):46-54. DOI: 10.18720/HUM/ISSN 2227-8591.39.04
11. Панасенков Н.А. Опыт обучения студентов-лингвистов постредактированию машинного перевода (на материале англо-русского перевода с помощью систем «Google Translate», «Яндекс Переводчик» и «Promt»). *Педагогическое образование в России*. 2019;(1):55-60. DOI 10.26170/po19-01-08
12. Митин П.С. Типичные ошибки машинного перевода общеполитической лексики (на примере перевода прессы Арабской Республики Египет). *Вестник Московского университета. Серия 22. Теория перевода*. 2021;22(1):7-34.
13. Шушарина Г.А., Петрунина Ж.В. Сопоставительный анализ текстов онлайн-переводов. *Вестник Челябинского государственного университета*. 2021;450(4):185-192. DOI 10.47475/1994-2796-2021-10425
14. Грамматика современного якутского литературного языка. Москва: Наука; 1982:496.

15. Төрүттэни кинигэтэ. Тахсы кинигэтэ. Москва: Институт перевода Библии; 2022:279 (На якут. языке).
16. Библия позволит осуществлять машинный перевод на тысячи языков <https://indicator.ru/humanitarian-science/bibliya-mashinnyj-perevod-11-05-2017.htm> (Дата обращения: 11.06.2025)
17. Мигалкина Е.И. *Сарсынньа сурук: хоһооннор*. Якутск: Айар; 2023:96 (На якут. языке)

References

1. Khokhryakov A. Types and review of machine translation engines: Yandex Translator, Google Translate and DeepL <https://habr.com/ru/articles/852810/> [Accessed: 11.03.2025] (in Russian).
2. Laktyushin NV. languages have been added to Google Translate: among them are the languages of the regions of Russia <https://hi-tech.mail.ru/news/111645-v-gugl-perevodchik-dobavili-110-yazykov-sredi-nih-yazyki-regionov-rossii/> [Accessed: 11.03.2025] (in Russian).
3. Borisov N. Yandex.Translator mastered the Yakut language <https://ysia.ru/yandeks-perevodchik-osvoil-yakutskij-yazyk/> [Accessed: 11.03.2025] (in Russian).
4. Ivanchenko TA. Errors in machine translation from German into Russian (based on articles of German-language media and their translations). *Uchenye zapiski Sankt-Peterburgskogo universiteta tekhnologii upravleniya i ekonomiki*. 2021;22(4) (in Russian). DOI: 10.35854/2541-8106-2021-4-30-41
5. Knyazheva EA. *Assessing the quality of translation: history, theory, practice*. Moscow: Flint; 2018:248 (in Russian).
6. Grebenyuk AV. Towards building a taxonomy of machine translation errors: versatility vs targeted approaches. *Vestnik Moskovskogo Universiteta. Seriya 22. Teoriya Perevoda = Moscow University Bulletin on Translation Studies*. 2024;22(2):7-25 (in Russian). DOI: 10.55959/MSU2074-6636 22-2024-17-2-7-25
7. Gu Jun-ling, Huang Zhong-lian. Study into classifications of translation mistakes. *Vestnik Moskovskogo Universiteta. Seriya 22. Teoriya Perevoda = Moscow University Bulletin on Translation Studies*. 2016;22(3):26-39 (in Russian).
8. Goncharov AA, Buntman NV, Nuriev VA. Machine translation errors: problems of classification. *Systems and Means of Informatics*. 2019;29(3):92-103 (in Russian). DOI:10.14357/08696527190308
9. Petrova TI. *General theory of translation*. Yakutsk: Publishing house of the Yakutia State University; 2007:38 (In Yakut).
10. Shevchuk EV., Nikiforova ZhA. Post-editing and typical mistakes in the computer-aided translation of academic, scientific, and journalistic texts. *Teaching Methodology in Higher Education*. 2021;10(39):46–54 (in Russian). DOI: 10.18720/HUM/ISSN 2227-8591.39.04
11. Panasenkov NA. Experience of teaching linguistics students how to post-edit machine-generated translation (based on english-russian translations via Google Translate, Yandex.Translate and Prompt systems). *Pedagogical Education in Russia*. 2019;(1):55-60 (in Russian). DOI 10.26170/po19-01-08
12. Mitin PS. Typical errors in machine translation of sociopolitical vocabulary: a case study of translations of Egypt's mass media. *Vestnik Moskovskogo Universiteta. Seriya 22. Teoriya Perevoda = Moscow University Bulletin on Translation Studies*. 2021;22(1):7-34 (in Russian).
13. Shusharina GA., Petrunina ZhV. Comparative analysis of online translations texts. *Bulletin of Chelyabinsk State University*. 2021;450(4):185-192. (in Russian). DOI 10.47475/1994-2796-2021-10425
14. Grammar of the modern Yakut literary language. Moscow: Science; 1982:496 (in Russian).
15. Genesis. Exodus. Moscow: Institute for Bible Translation; 2022:279 (In Yakut).
16. The Bible will allow machine translation into thousands of languages <https://indicator.ru/humanitarian-science/bibliya-mashinnyj-perevod-11-05-2017.htm> [Accessed: 01.06.2025] (in Russian).
17. Migalkina EI. *Letter to tomorrow: poems*. Yakutsk: Ayar; 2023:96 (In Yakut).

Об авторах

ВАСИЛЬЕВА Акулина Александровна – кандидат филологических наук, доцент, доцент кафедры «Стилистика якутского языка и русско-якутского перевода» Института языков и культуры народов Северо-Востока РФ, ФГАОУ ВО «Северо-Восточный федеральный университет имени М.К. Аммосова», ORCID: ID 0000-0003-2692-7026, ResearcherID: AAB-1348-2019, e-mail: vaaperevod@mail.ru

БУРЦЕВА Ольга Валерьевна – магистрант кафедры стилистики якутского языка и русско-якутского перевода, ФГАОУ ВО «Северо-Восточный федеральный университет имени М.К. Аммосова», e-mail: olgamekbur@mail.ru

About the authors

Akulina A. VASILIEVA – Cand. Sci. (Philology), Associate Professor, Yakut Stylistics and Russian-Yakut Translation Department, Institute of Languages and Culture of the Peoples of the Northeast, M.K. Ammosov North-Eastern Federal University, ORCID: ID 0000-0003-2692-7026, ResearcherID: AAB-1348-2019, e-mail: vaaperevod@mail.ru

Olga V. BURTSEVA – Master's student, Yakut Stylistics and Russian-Yakut Translation Department, M.K. Ammosov North-Eastern Federal University, e-mail: olgamekbur@mail.ru

Вклад авторов

Васильева А. А. – разработка концепции, методология, администрирование данных, создание черновика рукописи, редактирование рукописи, руководство исследованием, администрирование проекта

Бурцева А. В. – верификация данных, проведение статистического анализа, проведение исследования, создание черновика рукописи, визуализация

Authors' contribution

Akulina A. Vasilieva – conceptualization, methodology, data curation, writing – original draft, writing – review & editing, supervision, project administration

Olga V. Burtseva – data curation, validation, formal analysis, investigation, writing – original draft, visualization

Конфликт интересов

Один из авторов – Васильева А.А., является членом редколлегии сетевого издания «Алтаистика. Altaistics». Авторам неизвестно о каком-либо другом потенциальном конфликте интересов, связанном с этой рукописью.

Conflict of interest

One of the authors – Akulina A. Vasilieva Cand. Sci. (Philology), Chief Researcher is a member of editorial board of Altaistics. The authors are not aware of any other potential conflict of interest relating to this article.

Поступила в редакцию / Submitted 10.06.2025

Принята к публикации / Accepted 23.06.2025