

УДК 81.331

URL: https://ptsj.bmstu.ru/catalog/hum/ling_/984.html

АНАЛИЗ ОШИБОК МАШИННОГО ПЕРЕВОДЧИКА ПРИ ПЕРЕВОДЕ ТЕХНИЧЕСКИХ ТЕКСТОВ С АНГЛИЙСКОГО ЯЗЫКА НА РУССКИЙ

А.Ф. Шаталина

shura.shatalina@mail.ru

SPIN-код: 9026-3690

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

Статья посвящена анализу результатов машинного перевода узкоспециализированных научно-технических текстов. Рассмотрено понятие переводческой ошибки, описаны ее виды. Проведен анализ результатов машинного перевода узкоспециализированных текстов по авиации и космонавтике. Выделены ошибки, сделанные машинным переводчиком. Ошибки определялись на основе эталонного перевода, выполненного профессиональным переводчиком. В результате исследования классифицированы виды ошибок, сделанных в процессе машинного перевода. Отмечено, что при переводе узкоспециализированных текстов машинным способом наиболее часто встречаются смысловые, стилистические и переводные ошибки.

Ключевые слова: перевод, машинный перевод, машинный переводчик, технический текст, узкоспециализированный текст, переводческая ошибка, язык перевода, язык источника

Введение. Несмотря на то что в настоящее время в развитии машинного перевода произошли большие изменения, в коммерческой сфере системы машинного перевода не используются в полной мере, поскольку данные системы могут облегчить работу «живому» переводчику, но не заменить его. Однако работы по совершенствованию систем машинного перевода не останавливаются и ведутся каждый день.

В настоящее время сфера использования машинного перевода сводится к промышленному переводу. Машинный перевод может помочь в рутинной переводческой работе, когда скорость выполнения перевода важнее качества. Например, он может оказаться решением для специалиста, не владеющего тем или иным иностранным языком, на котором ему нужно прочесть интересующий его текст. Хотя машинный перевод, как правило, оказывается полон ошибок и неувязок, человек вполне может получить общее представление о содержании текста, что поможет ему решить, стоит ли заказывать более качественный перевод или попытаться отредактировать полученный. Таким образом, результаты машинного перевода почти всегда требуют редактирования. А то, насколько адекватными можно считать результаты перевода с по-

мощью компьютера, определяется не только качеством системы машинного перевода, но и качеством последующего редактирования.

Оценка качества перевода. Под переводческой ошибкой принято понимать необоснованное отступление от нормативного требования эквивалентности (т. е. содержательного соответствия перевода оригиналу) [1, с. 415] как меру несоответствия перевода оригиналу [2, с. 208], меру дезинформирующего воздействия на читателя [3, с. 424].

«Содержательные», смысловые ошибки подразделяют на два принципиальных вида — буквализмы и вольности. Буквализм — ошибка переводчика, заключающаяся в передаче формальных или семантических компонентов слова, самого слова, словосочетания или фразы в ущерб смыслу или информации о структуре [2], передача коммуникативно нерелевантных элементов оригинала, что приводит к нарушению нормы и узуса языка перевода, либо оказывается искаженным действительное содержание оригинала [4]. Вольность — передача ключевой информации без учета формальных и семантических компонентов исходного текста.

Выделение вольностей и буквализмов — наиболее традиционный и вместе с тем узкий подход к пониманию переводческой ошибки. С одной стороны, он свидетельствует о нарушении только лишь одной из норм перевода — нормы эквивалентности, а с другой — практически непригоден для оценки качества перевода, не дает представления о степени дезинформирующего воздействия перевода.

В методических целях для оценки качества перевода используется более широкое понятие переводческой ошибки. Оно основано не только на учете нарушений в передаче содержания исходного текста, но и на учете ошибок, связанных с «адаптацией этого содержания к новым лингвоэтническим условиям». Другими словами, второй тип ошибок — это так называемые ошибки выражения, представляющие собой как вынужденные (спровоцированные оригиналом), так и невынужденные нарушения нормы и узуса языка перевода.

Во многих вузах, занимающихся подготовкой переводчиков, используется следующая классификация ошибок, основанная на работах Л.К. Латышева [5, 6]:

– ошибки в трансляции исходного содержания — функционально не мотивированные отклонения от содержания оригинала, различающиеся по степени дезинформирующего воздействия;

– искажения — субъективно обусловленное отклонение содержания перевода от оригинала, в результате которого возмущенный потенциал переводного текста не соответствует возмущенному потенциалу исходного текста (вводит в заблуждение получателя относительно предмета сообщения, вызывает неадекватные представления);

– неточности — отклонения от содержания оригинала, характеризующиеся меньшей степенью дезинформирующего воздействия; немотивированное опущение или добавление информации, не искажающее полностью содержание оригинала, но требующее уточнения;

– неясности — функционально-содержательный изъян перевода, затемняющий смысл высказывания, достаточно четко выраженный в исходном тексте, часто возникают из-за неудобоваримости синтаксических конструкций. Может быть связан с неудачным выбором слова или структуры фразы либо обусловлен неадекватным использованием или неиспользованием переводческого приема;

– ошибки в адаптации содержания и форм его выражения к новым лингвоэтническим условиям восприятия, их причина — неумение компенсировать расхождения лингвоэтнической коммуникативной компетенции носителей исходного языка и языка перевода, так называемые языковые ошибки, ошибки в оформлении текста на языке перевода. Классифицируются по видам единиц языка и речи;

– лексическая ошибка — ошибка, связанная с неправильным использованием основного или контекстуального значения слова, а также нарушение норм сочетаемости слов в языке перевода;

– грамматическая ошибка — нарушение грамматических (в том числе синтаксических) норм языка перевода, не приводящее к искажению смысла оригинала;

– стилистическая ошибка — использование слова, конструкции или стилистического средства, не соответствующих по своим функционально-языковым свойствам той жанрово-стилистической разновидности текстов, к которой принадлежит перевод;

– орфографическая ошибка — ошибка в правописании слов на языке перевода;

– пунктуационная ошибка — несоблюдение правил пунктуации языка перевода.

Анализ ошибок машинного переводчика при переводе узкоспециализированных текстов. В результате анализа машинного перевода текстов по аэрокосмической технике [7, 8] были выделены следующие типы ошибок: смысловые, трансформационные, стилистические, и отдельная группа ошибок, вызванных работой самой машинного перевода.

Смысловые ошибки — это ошибки, которые вводят получателя в заблуждение относительно предмета сообщения, вызывают неадекватные представления или субъективно обусловленное отклонение содержания перевода от оригинала, например:

Nasa says it has learnt the safety lessons of the Columbia disaster, which cost the lives of seven astronauts, and has made this the most scrutinized shuttle flight ever.

НАСА говорит, что оно усвоило уроки безопасности катастрофы в Колумбии, которая стоила жизни семи астронавтов, и сделало этот полет самым тщательно изученным шаттлом.

The main wing of the finished plane was more than thirty-three meters across. That is wider than the main wing on today's big passenger planes. The center part of the plane held the crew. And on either side of this body were two long fuel tanks.

Главное крыло готового самолета было более тридцати трех метров в поперечнике. Это шире, чем основное крыло на сегодняшних больших пассажирских самолетах. Центральную часть самолета удерживал экипаж. И по обе стороны от этого тела были два длинных топливных бака.

Смысловые ошибки, допущенные системой машинного перевода при переводе текстов по аэрокосмической технике с английского языка на русский, можно классифицировать на ошибки с полной потерей смысла и ошибки с частичной потерей смысла.

При ошибке с полной потерей смысла понимание текста не представляется возможным, например:

It had completed a forty-thousand-kilometer flight around the world. It had not stopped. And it had not refueled.

Он совершил сорок тысяч километров кругосветного полета. Она не остановилась на достигнутом. И она не подпитывается.

Russia's space program has a bad case of the Red Planet blues.

У российской космической программы плохой случай блюза Красной планеты.

Report decries lack of UK equivalent to Microsoft.

Отчет осуждает отсутствие Великобритании эквивалентно Microsoft.

Частичная потеря смысла уводит заказчика перевода от полного понимания текста, например:

First of all, despite the Cold War being over for 20 years, Russia is not all that forthcoming with specific details of Phobos-Grunt's construction.

Во-первых, несмотря на то что «холодная война» закончилась 20 лет назад, Россия не всегда готова к конкретным деталям строительства Фобоса-грунта.

Instead, Phobos-Ground added to a list of embarrassing space failures over the last year.

Вместо этого Phobos-Ground добавил к списку смущающих космических сбоев за последний год.

Russia's doomed Mars probe may not be a total waste.

Обреченный зонд Марса в России не может быть полным отходом.

Например, в рассматриваемых примерах машинный переводчик переводит время как дату, что также искажает смысл перевода:

The European Space Agency (ESA) probe lifted off at 1902 local time (1602 BST) on Saturday from the Plesetsk Cosmodrome in Russia.

Зонд Европейского космического агентства (ЕКА) стартовал в 1902 году по местному времени (1602 BST) в субботу с космодрома Плесецк в России.

Перечисленные ошибки представляют собой грубое искажение содержания оригинала текста об аэрокосмической технике. Они возникают из-за отсутствия правильного понимания содержания оригинала, так как в настоящее время современные системы машинного перевода используют алгоритмы перевода, которые не учитывают общую тематику текста; отсутствует база знаний предметной области перевода. Обращение к такой базе могло бы возможным решением ряда проблем, связанных с выбором лексических единиц языка перевода. Смысловые ошибки недопустимы при переводе, поскольку их наличие препятствует успешной коммуникации.

Трансформационными ошибками можно считать те ошибки, которые приводят к неточной передаче смысла оригинала, но не искажают его полностью. Трансформационные ошибки включают в себя лексические, грамматические, неверный перевод фразовых глаголов, аббревиатур, терминологии.

Лексические ошибки подразумевают неверный или некорректный перевод слова. Вот несколько примеров:

SHUTTLERETURNTOFLIGHT

ТРАНСФЕР ВОЗВРАЩЕНИЕ
В РЕЙС

The main wing of the finished plane was more than thirty-three meters across.

Главное крыло готового самолета было более тридцати трех метров в поперечнике.

Грамматические ошибки — это ошибки в образовании форм слов, ошибки в употреблении причастных и деепричастных оборотов, ошибки в согла-

совании и управлении, ошибки в построении сложных предложений, смешение прямой и косвенной речи, например:

"I feel extremely sorry for the teams of scientists who have staked whole careers on preparing experiments," he said.

«Мне очень жаль команды ученых, которые всю свою карьеру делали ставку на подготовку экспериментов», — сказал он.

"Space is a risky business, it always has been, it doesn't always go perfectly," he told the BBC News website.

«Космос — это рискованный бизнес, он всегда был, он не всегда идет идеально», — сказал он на веб-сайте BBC News.

Фразовые глаголы представляют сложность для систем машинного перевода при переводе технического текста, например:

He eventually plans to reduce the size of the apparatus, taking it down to a small, lightweight vest for divers.

В конце концов он планирует уменьшить размер аппарата, опустив его до небольшого легкого жилета для дайверов.

Its hold on the top slot was revealed in the latest list of the Top 500 supercomputers on Earth.

Его оставание на вершине было выявлено в последнем списке 500 лучших суперкомпьютеров на Земле.

Неверный перевод аббревиатур и сокращений, в большинстве случаев машинный переводчик вообще не переводит аббревиатуры и сокращения, что является большим недостатком, поскольку они могут содержать в себе полезные сведения, влияющие на общее понимание смысла.

Discovery is due to touch down at Kennedy Space Center on 7 August at 0546 EDT (0946 GMT; 1046 BST).

Дискавери должен приземлиться в Космическом центре Кеннеди 7 августа в 0546 EDT (0946 GMT; 1046 BST).

The adapted hummer averaged 28 mph (45 kph) and hit a top speed of 36 mph (57 kph).

Адаптированный Хаммер в среднем 28 mph (45 kph) и ударил максимальную скорость 36 mph (57 kph).

Неверный перевод терминологии также является недостатком машинного перевода, поскольку в технических текстах терминология — одна из самых важных составляющих.

Ocean fall

Падение океана

THE EARTH'S ICE COVER

Земной ледяной покров

В статье представлены различные типы ошибок машинного перевода, среди которых: неверный перевод терминов, неверный перевод слов (как следствие — изменение смысла), не переведены даты, сокращения, аббревиатуры. Также машинный переводчик допускает грамматические ошибки (неверный перевод времен), не различает степени прилагательных и не переводит фразовые глаголы. В результате допущения таких ошибок в переводе искажаются лишь отдельные детали, в то время как ситуация остается той же. Подобные ошибки возникают из-за неточного понимания значения некоторых слов в оригинале или неверной оценки переводчиком степени соответствия значений английского и русского терминов.

Стилистические ошибки подразумевают использование слов, конструкций или стилистических средств, не соответствующих по своим функционально-языковым свойствам той жанрово-стилистической разновидности текстов, к которой принадлежит перевод, например:

By road and by rail, Sweden has a lot to teach the rest of us.

По дороге и по железной дороге, Швеция имеет много, чтобы научить остальных из нас.

And if methane doesn't light your fire, you can choose to have your car run on a high-grade biofuel mix.

И если метан не зажигает огонь, вы можете выбрать, чтобы ваш автомобиль работал на высококачественной смеси биотоплива.

They clapped and cheered as the rocket was seen clearing the launch pad in a cloud of smoke, but their jubilation was short-lived.

Они хлопали и ликовали, когда ракету видели очищающей стартовую площадку в облаке дыма, но их ликование было недолгим.

Surface measurements and overflights suggest Greenland's ice is undergoing melting.

Поверхностные измерения и пролеты показывают, что лед Гренландии переходит в таяние.

Scientists said the crash was a "tragedy" and it would be years before they could launch a similar mission, even if more funding were available.

Ученые сказали, что авария была «трагедией», и пройдут годы, прежде чем они смогут запустить подобную миссию, даже если будет больше финансирования.

That in turn makes further pushing less fruitful so that eventually nothing can be accelerated beyond the speed of light.

Это, в свою очередь, делает дальнейшее продвижение менее плодотворным, так что в конечном итоге ничто не может быть ускорено сверх скорости света.

No equation is anywhere near as recognizable as $E = mc^2$

Нет уравнение так узнаваем, как $E = mc^2$.

Instead, Phobos-Ground added to a list of embarrassing space failures over the last year.

Вместо этого Phobos-Ground добавил к списку смущающих космических сбоев за последний год.

A robotic Humvee has managed to drive itself for seven hours without crashing on a race course in the US.

A robotic Humvee has managed to drive itself for seven hours without crashing on a race course in the US.

Nations have long competed to have the tallest edifice.

Народы уже давно соревновались, чтобы иметь самое высокое здание.

He supervised two students — Matthias Hahnen and Robert Doerr — as they developed their concept for an everyday object that could be enriched by computers.

Он руководил двумя студентами — Маттиас и Роберт Дорр — как они разработали свои концепции на обычные вещи, которые могут быть обогащены с помощью компьютеров.

The machines in the top two slots of the Top 500 are unchanged from the last list produced in June 2005 and both are IBM Blue Gene machines.

Машины в первых двух слотах Top 500 не изменились по сравнению с последним списком, выпущенным в июне 2005 года, и оба являются машинами IBM BlueGene.

Примеры демонстрируют основной тип стилистических ошибок — несоответствие нормам литературного русского языка.

Частные случаи. К частным случаям ошибок относятся проблемы машинного переводчика (не переведено слово; один раз перевел слово, в другой не перевел вообще; неверный перевод названий, имен и фамилий); причина данного явления неизвестна.

Опущение слов при переводе в некоторых случаях может привести к искажению смысла, иногда является необязательным, например:

The complex will house offices, luxury flats, a hotel and leisure facilities.

В комплексе разместятся офисы, квартиры, отель и удобства

Также встречаются случаи, когда машинный переводчик в одном предложении (абзаце) переводит слово (словосочетание), а в следующем — нет, пример:

The bio-powered version of the Saab 95 costs around \$1,000 (f500) more than the normal model.

Биоэнергетическая версия Сааб 95 стоит около 1000 долларов (f500) больше, чем обычная модель.

But with pump prices for the E85 mix a third cheaper than normal petrol, company car tax breaks, and exemptions for parking and congestion charges, Saab reckons you get that \$1,000 back within the first year.

The mission will take the astronauts on board the Shenzhou VI spacecraft on a five-day orbit around Earth, during which they will carry out experiments.

Shenzhou VI, like Shenzhou V, is based on Russia's Soyuz spacecraft, a model developed in the late 1960s.

Еще одной ошибкой, допускаемой машинным переводчиком, является неверный перевод имен и фамилий или вообще отсутствие перевода:

William Touche

Fei Junlong

Уильям туш

Fei Junlong

Также могут быть не переведены или неверно транскрибированы названия географических объектов, например:

The test was part of preparations for a robot vehicles race across the Mojave.

Тест был частью подготовки к гонке роботов по пустыне Можаве.

Результаты анализа текстов. Всего было исследовано 18 текстов, содержащих 402 предложения. Распределение количества ошибок по типам (абсолютные значения и доли в общем числе ошибок) представлено ниже:

- смысловые ошибки — 62 (21 %);
- стилистические ошибки — 49 (17 %);
- переводные ошибки — 147 (50 %);
- частные случаи — 37 (12 %).

Полученные в результате анализа данные позволяют утверждать, что полностью автоматизированный перевод (независимый от человека) на сегодняшний день невозможен. Но результаты машинного перевода могут быть использованы для поверхностного ознакомления с содержанием при условии, что текст используется как сигнальная информация и не требует тща-

тельного редактирования [9, с. 98–103; 10, с. 46–55]. Для полноценного перевода необходимо постредактирование текста с учетом выявленной типологии ошибок машинного перевода. Отметим, что были выделены лишь некоторые наиболее частотные ошибки машинного перевода и данный список может быть, несомненно, продолжен.

Заключение. Современные компьютерные программы перевода достаточно совершенны, но они до сих пор не могут разрешить самую важную задачу процесса перевода: выбор контекстуально-верного варианта, который в каждом тексте обусловлен многими причинами. Именно по этой причине качественный машинный перевод без участия человека возможен лишь в теории: машины, даже несмотря на их быстротечный прогресс, все еще слишком слабы, чтобы знать точные перечни всех синтаксических конструкций, возможных на том или ином языке, а также многое другое, что требуется знать любому уважающему себя переводчику. А пока что лишь мозг человека может различать разнообразные ситуации, видеть темы и ремы, учитывать стилистику, особенности языков и многое-многое другое.

Таким образом, в результате исследования собран теоретический материал для анализа и создания классификации, т. е. тексты по аэрокосмической отрасли в количестве 18 штук. Выявлены ошибки машинного переводчика при переводе текстов по аэрокосмической тематике с английского на русский язык в количестве 295 (всего). Ошибки, допущенные системой машинного перевода, были классифицированы и выделены 4 основные группы: смысловые (ошибки, влияющие на общее понимание текста), стилистические (ошибки, которые влияют на общее восприятие информации), переводные (ошибки, включающие в себя лексические и грамматические) и частные случаи.

Литература

- [1] Швейцер А.Д. *Перевод и лингвистика*. Москва, Наука, 1988, 415 с.
- [2] Миньяр-Белоручев Р.К. *Как стать переводчиком?* Москва, Велби, 2007, 480 с.
- [3] Комиссаров В.Н. *Современное переводоведение*. Москва, ЭТС, 2001, 424 с.
- [4] Комиссаров В.Н. *Теория перевода (лингвистические аспекты)*. Москва, Высш. шк., 1990, 253 с.
- [5] Латышев Л.К. *Технология перевода*. Москва, Академия, 2005, 320 с.
- [6] Латышев Л.К. *Технология перевода*. Москва, НВИ-Тезаурус, 2001, 280 с.
- [7] Бездорожев С.В. Новый взгляд на роль переводчика: проблема использования машинного перевода. *Вестник Московского государственного областного университета*, 2013, № 2, 15 с.

- [8] Дроздова К.А. Машинный перевод: история, классификация, методы. *Вестник Омского государственного педагогического университета. Гуманитарные исследования*, 2015, № 3 (7), с. 156–158.
- [9] Бутенко Ю.И., Маргарян Т.Д., Кочеткова Е.Л. Переводческая деятельность в условиях цифровизации: обучение и трудоустройство. *Alma mater (Вестник высшей школы)*, 2020, № 9, с. 98–103. <https://doi.org/10.20339/AM.09-20.098>
- [10] Бутенко Ю.И., Авагян Н.А. Способы выражения модальности в параллельных текстах стандартов (на примере нормативной базы программной инженерии). *Вестник ВГУ: лингвистика и межкультурная коммуникация*, 2021, № 2, с. 46–55. <https://doi.org/10.17308/lic.2021.2/3414>

Поступила в редакцию 25.04.2024

Шаталина Александра Федоровна — студентка кафедры «Романо-германские языки», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

Научный руководитель — Бутенко Юлия Ивановна, кандидат технических наук, доцент кафедры «Романо-германские языки», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация. E-mail: iubutenko@bmstu.ru; SPIN код: 4763-8297.

Ссылку на эту статью просим оформлять следующим образом:

Шаталина А.Ф. Анализ ошибок машинного переводчика при переводе технических текстов с английского на русский язык. *Политехнический молодежный журнал*, 2024, № 03 (92). URL: https://ptsj.bmstu.ru/catalog/hum/ling_/984.html

ANALYZING ERRORS IN THE TECHNICAL TEXT TRANSLATION FROM ENGLISH INTO RUSSIAN

A.F. Shatalina

shura.shatalina@mail.ru

SPIN code: 9026-3690

Bauman Moscow State Technical University, Moscow, Russian Federation

The paper presents results of analyzing machine translation of the highly specialized scientific and technical texts. It considers the concept of translation error and describes its types. Results of machine translation of the highly specialized texts in aviation and astronautics are being analyzed. Errors made by the machine translator are highlighted. Errors are identified using a reference translation by a professional translator. The study results in classifying the types of errors made in machine translation. The paper notes that semantic, stylistic and translation errors are most often encountered in translating the highly specialized texts by the machine technique.

Keywords: translation, machine translation, machine translator, technical text, highly specialized text, translation error, translation language, source language

Received 25.04.2024

Shatalina A.F. — Student, Department of Romano-Germanic Languages, Bauman Moscow State Technical University, Moscow, Russian Federation.

Scientific advisor — Butenko Yu.I., Ph. D. (Eng.), Associate Professor, Department of Romano-Germanic Languages, Bauman Moscow State Technical University, Moscow, Russian Federation. E-mail: iubutenko@bmstu.ru; SPIN-code: 4763-8297.

Please cite this article in English as:

Shatalina A.F. Analyzing errors in the technical text translation from English into Russian. *Politekhnicheskij molodezhnyy zhurnal*, 2024, no. 03 (92). (In Russ.). URL: https://ptsj.bmstu.ru/catalog/hum/ling_/984.html