

МЕТОДОЛОГИЯ ПЕРЕВОДА



НАУЧНАЯ СТАТЬЯ

УДК 82-1: 81'322.4

DOI: 10.55959/MSU2074-6636-22-2024-17-2-7-25

ПРОБЛЕМА ТИПОЛОГИИ ОШИБОК МАШИННОГО ПЕРЕВОДА: СТРЕМЛЕНИЕ К УНИВЕРСАЛЬНОСТИ VS ТАРГЕТИРОВАННОСТЬ

Андрей Владимирович Гребенюк

Санкт-Петербургский государственный университет,
г. Санкт-Петербург, Россия

Для контактов: yax-light@yandex.ru

Аннотация. Работа посвящена выявлению проблем, имеющих в существующих подходах к типологизации ошибок машинного перевода (МП), а также поиску путей к её совершенствованию. Несмотря на то, что подобные классификации могут быть основаны на типологии, разработанной на материале переводов, выполненных человеком, ошибки МП имеют свою специфику, которая также должна учитываться в системе классификации. В первой части статьи обсуждается понятие машинного перевода и переводческой ошибки и даётся обзор основных подходов к классификации переводческих ошибок в традиционном, «человеческом» переводе. Во второй части статьи обсуждаются классификации, предлагаемые для анализа перевода, выполненного системами автоматического перевода, и некоторые их ограничения. Изучался материал, преимущественно связанный с переводами в паре с русским языком и выполненными с помощью наиболее популярных в России сервисов машинного перевода «Яндекс Переводчик», “Google Translate” и “Promt”. В частности, обсуждаются выделяемые различными авторами основные классы ошибок, их частотность, а также их «вес» с точки зрения их способности приводить к коммуникационному сбою. Констатируется отсутствие в настоящее время единого подхода к построению подобных классификаций, а также зависимость существующих типологий от типа текста, языковой пары и конкретной системы автоматического перевода. Делается вывод о том, что

оптимальным подходом к классификации ошибок МП является подход не универсальный, а таргетированный, то есть зависящий от определенных параметров перевода. Также обсуждается возможность практического применения машинного перевода при обучении иностранному языку и в работе профессиональных переводчиков, а также необходимость пост-редактуры переведённых текстов.

Ключевые слова: машинный перевод, системы автоматического перевода, переводческие ошибки, типология ошибок, постредактирование

Для цитирования: Гребенюк А.В. Проблемы типологии ошибок машинного перевода: стремление к универсальности vs таргетированность // Вестник Московского университета. Серия 22. Теория перевода, 2024, № 2. С. 7–25. DOI: 10.55959/MSU2074-6636-22-2024-17-2-7-25

Статья поступила в редакцию 27.04.2024;
одобрена после рецензирования 08.07.2024;
принята к публикации 08.07.2024.

TOWARDS BUILDING A TAXONOMY OF MACHINE TRANSLATION ERRORS: VERSATILITY VS TARGETED APPROACHES

Andrey V. Grebenyuk

St. Petersburg State University, St. Petersburg, Russia

For contacts: yax-light@yandex.ru

Abstract. This study examines problems in existing approaches to classifying machine translation (MT) errors. Despite the fact that such classifications can be based on a taxonomy developed on the material of traditional translations, MT errors have their own specifics, which should also be considered in the classification system. The first part of the presented paper discusses the concepts of machine translation and translation error per se and provides an overview of the main approaches to building a taxonomy of translation errors in traditional, “human” translation. In the second part, we discuss the existing classification systems proposed for the analysis of machine translation output and some of their limitations. The research material in the discussed papers is mainly focused on translations in pairs with the Russian language and obtained through the use of the most popular in Russia machine translation services, viz. Yandex Translator, Google Translate and Promt. In particular, we discuss the main classes of errors identified by various authors, the frequency of the errors of the said classes, as well as their gravity in terms of their ability to cause a communication failure. This study argues that there is currently no unified approach to the construction

© Grebenyuk A.V., 2024

of such classifications, as well as that they are inevitably dependent on the type of the analysed text, the language pair and the chosen automated translation system. It is concluded that the optimal approach to classifying MT errors is not a universal approach, but a targeted one, that is, depending on certain translation parameters. The possibility of practical applications of machine translation in teaching foreign languages and in the work of professional translators, as well as the need for post-editing of translated texts, are also discussed.

Keywords: machine translation, automated translation systems, translation errors, taxonomy of errors, error analysis, post-editing

For citation: Grebenyuk A. V. (2024). Towards building a taxonomy of machine translation errors: versatility vs targeted approaches. *Vestnik Moskovskogo Universiteta. Seriya 22. Teorija Perevoda — Moscow University Bulletin on Translation Studies*. 2. P. 7–25. DOI: 10.55959/MSU2074-6636-22-2024-17-2-7-25

The article was received on April 27, 2024;
approved after reviewing on July 08, 2024;
accepted for publication on July 08, 2024.

Введение

В соответствии с одним из определений перевода (по данным И.С. Алексеевой), перевод — это деятельность (и результат этой деятельности), которая «заключается в вариативном перевыражении, перекодировании текста, порождённого на одном языке, в текст на другом языке, осуществляемая переводчиком, который творчески выбирает вариант в зависимости от вариативных ресурсов языка, вида перевода, задач перевода, типа текста и под воздействием собственной индивидуальности» (Алексеева, 2008: 7). Несмотря на то, что в приведённое определение явным образом входит указание на человеческий характер данной деятельности, с середины XX века существует и иной вид перевода — машинный (Воронович, 2013: 3; Wang, 2022). Вызывая скепсис ряда специалистов на протяжении многих десятилетий, методы машинного перевода (МП) тем не менее непрерывно развивались и в настоящее время являются инструментом, хоть и требующим, как правило, постредактуры, но всё же позволяющим решать широкий ряд практических задач (Wang, 2022; Беляева, 2022).

Для совершенствования машинного перевода необходимо как можно более отчётливое понимание специфики возникающих в нём ошибок. Классификации этих ошибок был посвящён значи-

тельный ряд исследований, но этими же исследованиями выявлен целый ряд проблем, затрудняющих такую классификацию в принципе. В настоящей работе критически рассматриваются результаты этих исследований: основные классы ошибок, их частотность, а также их «вес» с точки зрения их способности приводить к коммуникационному сбою. Констатируется зависимость существующих типологий от типа текста, языковой пары и конкретной системы автоматического перевода. По итогам данного исследования автор отдаёт предпочтение таргетированной классификации ошибок МП перед классификацией универсальной. Помимо теоретических выводов, настоящее исследование содержит практические наблюдения и рекомендации, связанные с работой наиболее распространённых в России средств МП.

Типологизация ошибок традиционного перевода

Общим местом современного переводоведения является признание недостижимости «абсолютно точного» перевода без каких-либо потерь. Вместо этого принято говорить об эквивалентности и адекватности — мерах соответствия переведённого текста тексту исходному и цели перевода соответственно, и о необходимости передачи инварианта перевода — тех, компонентов содержания, которые не могут быть опущены или заменены (Алексеева, 2008: 145–146; Шамова, 2005). Не останавливаясь подробно на различных подходах к соотношению и строгому определению данных понятий, разграничению уровней эквивалентности и степеней инвариантности, отметим, что содержание данных понятий будет также зависеть от типа текста.

Упомянутые потери информации при переводе часто называются ошибками перевода. Мы будем понимать их в широком смысле, необязательно проистекающими из некомпетентности переводчика (в некоторых работах предлагается разделение на «ошибки» и «деформации» по критерию обоснованности/необоснованности (Гарбовский, 2007: 535–536). Приведём более строгое, данное Р.К. Миньяр-Белоручевым (Миньяр-Белоручев, 1996: 131–132), определение ошибки перевода: «некоторое количество переданной или добавленной информации, вычлняемой в виде либо непереведённого речевого отрезка в исходном тексте, либо добавленного речевого отрезка в переводном тексте». Соответственно возникает вопрос о создании классификации подобных ошибок.

Обзор литературы показывает, что в качестве признака классификации как правило выбирается или источник (причина) воз-

никновения ошибки, или уровень языковой системы (и/или аспект коммуникации), на котором проявляется данная ошибка (Федорова, 2020). В качестве классификации первого типа можно привести пример одной из типологий, предложенных Н.К. Гарбовским (Гарбовский, 2007: 514):

- 1) недостаточное владение языком оригинала;
- 2) недостаточный когнитивный опыт, недостаток знаний об описываемой в исходном тексте области окружающей действительности;
- 3) невнимательное отношение к системе смыслов, заключённой в исходном тексте, непонимание того, что автор говорит о предмете;
- 4) неумение различить особенности индивидуального стиля автора исходного речевого произведения.

Н.К. Гарбовский в своей работе приводит и классификацию второго типа, в которой ошибки соответствуют уровням семантики слова, словосочетания, предложения и предметной ситуации в целом, и, дополнительно, в отдельную категорию выделяются стилистические ошибки (Там же: 517–535). Ко второму типу также относятся типологии таких теоретиков советского переводоведения, как Я.И. Рецкер, В.Н. Комиссаров и Л.С. Бархударов, сравнительный анализ которых представлен в обзоре (Терентьева, 2020). В качестве примера приведём классификацию Л.С. Бархударова, во многом основанную на семиотической триаде Ч. Морриса (семантика, прагматика, синтактика):

- 1) лексические:
 - 1.1) референциальные (денотативное содержание);
 - 1.2) прагматические (регистр, коннотации, эмоциональная окраска);
 - 1.3) внутрилингвистические (рифма, ритм, звукопись);
- 2) грамматические (ошибки передачи грамматических категорий числа, модальности, времени и т.д.).

Разумеется, существуют и классификации, едва ли однозначно попадающие в одну из двух рассмотренных групп. Так, встречаются классификации, основанные на этапах процесса перевода, учитывающие уже упоминавшийся критерий обоснованности (мотивированности), а также разнообразные гибридные типологии (Самсонов, 2022; Гу Цзюньлин, 2016). Представляется интересной классификация, предложенная Д.М. Бузаджи и соавторами (Бузаджи, 2009), распределяющая ошибки по четырём категориям в

порядке убывания значимости и, далее, по нескольким группам в каждой категории:

- 1) ошибочная передача денотативного содержания текста:
 - 1.1) опущение, добавление, замена информации (значимые ошибки);
 - 1.2) неточная передача денотативного содержания текста (менее значимые: инвариант перевода сохраняется);
- 2) стилистические ошибки:
 - 2.1) нарушения в передаче функционально-стилевых или жанровых особенностей;
 - 2.2) калькирование;
 - 2.3) нарушения узуса переводящего языка;
- 3) ошибки передачи авторской оценки:
 - 3.1) ошибки, связанные с ослаблением или усилением экспрессии оригинала;
 - 3.2) собственно ошибки авторской оценки;
- 4) очевидные нарушения нормы и узуса переводящего языка:
 - 4.1) нарушение орфографических и пунктуационных норм;
 - 4.2) неверная передача имён собственных;
 - 4.3) нарушения стилистических норм переводящего языка;
 - 4.4) ошибочная передача специфических видов цифровых данных;
 - 4.5) несоответствие нормам оформления текстов данного типа.

Типологизация ошибок машинного перевода

На протяжении последнего десятилетия внимание исследователей было также обращено и к проблеме типологизации ошибок машинного перевода. Нужно отметить, что к анализу соответствующих публикаций следует подходить с осторожностью, потому что полученные результаты могут зависеть не только от рассматриваемой языковой пары и типа текста, как это было ранее, но и от конкретной системы автоматического перевода. Кроме того, все подобные системы непрерывно совершенствуются, и переводы одного текста, разделённые значимым промежутком времени, могут существенно отличаться. Ниже предлагается анализ основных публикаций, посвящённых анализу ошибок машинного перевода, преимущественно в наиболее популярных (Кочеткова, 2017) нейросетевых системах «Яндекс Переводчик» (также использует статистический подход) и Google Translate.

Одной из ранних и уже классических работ в области типологизации ошибок машинного перевода является работа 2006 года Дэвида Вилара и соавторов (Vilar, 2006). Классификация, предложенная в статье, была разработана на четырёх языковых парах: английский/испанский и английский/китайский. Для машинного перевода использовались различные алгоритмы, основанные на статистическом подходе. Классификация является преимущественно двухуровневой:

- 1) Пропущенные слова
 - 1.1) знаменательные слова (7,9%; 7,2%)
 - 1.2) служебные слова (12,0%; 18,8%)
- 2) Ошибка порядка слов
 - 2.1) уровень слов (13,3%; 13,3%)
 - 2.2) фразовый уровень (2,1%; 7,1%)
- 3) Неправильные слова
 - 3.1) неверное лексическое значение (21,9%; 28,2%)
 - 3.2) неверная форма (33,9%; 9,9%)
 - 3.3) лишнее слово (0,0%; 1,1%)
 - 3.3) стилистическая ошибка (7,9%; 9,9%)
 - 3.4) передача идиом (0,7%; 1,7%)
- 4) Неизвестные слова (в исходном тексте)
 - 4.1) неизвестная основа (0,3%; 1,1%)
 - 4.2) неизвестная форма (0,0%; 1,6%)
- 5) Пунктуация

Некоторые подгруппы этой классификации имеют третий уровень. Так, подкатегории «уровень слов» и «фразовый уровень» категории «порядок слов» далее делятся на неправильный порядок ближнего и дальнего расположения в тексте (*local / long range*), а в ошибках лексического значения отдельно выделяется группа, связанная с неверным выбором контекстуального значения. Сами авторы указывают на трудность построения логически последовательной классификации и зачастую невозможности однозначного отнесения ошибки в ту или иную группу. В схеме классификации указаны частотности ошибок разных категорий (в процентах от общего числа; пунктуационные ошибки не учитывались) для перевода с английского на испанский и с испанского на английский соответственно. Как видно из статистики, самыми частотными оказываются ошибки подкатегорий «неверное лексическое значение», «неверная форма» и «стилистическая ошибка», при этом распределение зависит от направления перевода в паре: в силу большей

морфологической сложности испанского языка подкатегория «неверная форма» составляет треть всех ошибок (против одной десятой в переводе на английский).

В уже упоминавшейся работе Л.Н. Беляевой (Беляева, 2022) рассматриваются возможные проблемы машинного перевода в контексте постредактирования: неправильный перевод фрагмента текста в целом; пропуск или вставка дополнительной информации; некорректный выбор терминологии; проблемы, относящиеся к грамматике, орфографии и пунктуации (объединены в одну группу); несоблюдение стилистических норм переводящего языка; нарушение стандартов страны языка перевода (наименований валюты и разделителей десятичных чисел, неверное использование прописных букв); несогласованность (использование нестандартизированной терминологии). Особый интерес здесь представляет последний пункт. Действительно, при автоматическом переводе одно и то же слово может быть переведено несколькими разными способами в пределах одного абзаца (в зависимости лишь от локального контекста), что, хотя и не исключено в переводе человеком, является в значительной степени специфичными именно для машинного перевода.

Перевод немецкоязычных СМИ на русский системами Яндекс и Google исследуется в работе (Иванченко, 2021). Первый класс, который выделяют авторы — класс лексико-семантических ошибок, совершенных вследствие: омонимии, полисемии, наличия нетривиальных контекстуальных значений, эллиптических конструкций, имён собственных, реалий и иноязычных включений. Ошибки данного класса оказываться самыми частотными. Далее в порядке уменьшения частотности идут: нормативно узувальные, грамматические (неверная передача грамматических категорий падежа, времени, рода) и самые малочисленные — пунктуационные и орфографические. В работе также указывается, что, хотя обе системы в целом хорошо справляются с орфографией и пунктуацией, большее количество ошибок данного типа совершает «Яндекс Переводчик».

В работе Н.С. Кочетковой и Е.В. Ревинной (Кочеткова, 2017) также исследуют автоматический перевод с немецкого на русский сервисами компаний Яндекс, Google, и Promt, но уже на материале научно-технического доклада посвящённого электроэнергетике. Авторы выделяют три категории ошибок: лексико-семантические, синтаксические и морфологические. К первой группе относились случаи пропуска и неверного перевода лексической единицы, а также ситуации, когда смысл предложения в целом трудно понять.

Отметим, что последнее — это тоже проблема специфичная именно для автоматического перевода. Категория лексико-семантических ошибок оказалась самой частотной (52% всех ошибок), затем идут морфологические (неправильная передача категорий рода, числа и т.д., 27%) и синтаксические (неверный порядок слов, 23%).

Схожая классификация использовалась в работе Л.Э. Щекиной (Щекина, 2021), посвящённой особенностям работы системы «Яндекс Переводчик» в переводе с французского языка на русский. В ней дополнительно выделяется четвёртая категория пунктуационных ошибок. Как и в предыдущей работе, самой частотной оказалась категория лексико-семантических ошибок (63% всех ошибок), далее идут синтаксические (20%), морфологические (15%) и пунктуационные ошибки (2%). Различие результатов может быть обусловлено различием как языковой пары, так и анализируемых систем машинного перевода. Кроме того, работа Н.С. Кочетковой вышла на 4 года позже и застала другое состояние сервиса «Яндекс Переводчик».

В работе М.А. Зайцевой (Зайцева, 2023) также исследовалась работа сервиса Яндекса, но уже в переводе с английского на русский. Для классификации обнаруженных ошибок перевода использовалась система из двух основных категорий (лексико-семантические ошибки и ошибки передачи культурных реалий), которые затем делились на группы. В лексико-семантические вошли ошибки перевода отдельных слов и словосочетаний и, отдельными группами, ошибки перевода сленга и метафор. Во второй категории выделялись ошибки передачи исторических, культурных, географических реалий и культурно-обусловленные ошибки, вызванные недостатком контекста. Частотности ошибок различного вида в работе не приводится.

Приведённая выше классификация Д.М. Бузаджи легла в основу исследования В.Г. Красильниковой (Красильникова, 2021), посвящённого сравнительному анализу работы систем компаний Google и Яндекс в переводе англоязычного научно-популярного текста медицинской тематики на русский язык. Специфика текста обусловила отказ от категории ошибок передачи авторской оценки. В статью приведена подробная статистика ошибок: в обеих системах самыми частотными оказались нарушения при передаче денотативного содержания текста, 46,7% и 42,8% от всех ошибок для Google и Яндекс переводчиков соответственно. При этом основной является подкатегория неточной передачи информации (34,3% и 32,6% всех ошибок), а опущения, добавления и замены, напротив,

относительно маловероятны. Далее в обеих системах идут стилистические ошибки (40,0% и 42,4%) и ошибки нарушения нормы и узуса переводящего языка (13,3% и 14,8%). Из приведённых данных видно, что обе системы демонстрируют схожее распределение ошибок.

В следующей публикации Ф.Б. Ситдикова и соавторы (Ситдикова, 2019) исследуют особенности работы системы Google Translate в переводе англоязычных субтитров к фильмам на русский язык. К первой категории (53% ошибок) относились лексико-семантические ошибки (неспособность выбрать верное контекстуально-обусловленное значение отдельной лексической единицы, неверный перевод словосочетаний с непрямым значением, перевод имён собственных). Ко второй категории (28%) принадлежат морфологические ошибки, к третьей (19%) — синтаксические и стилистические. В работе также отмечается, что именно лексико-семантические ошибки затрудняли понимание в наибольшей степени.

В исследовании Г.А. Шушариной и Ж.В. Петруниной (Шушарина, 2021) проводится сравнительный анализ работы сервисов Google и Promt в переводе на русский язык англоязычных текстов различных функциональных стилей. Интересно отметить, что в основе классификации ошибок лежит классификация переводческих трансформаций В.Н. Комиссарова: лексические, грамматические и лексико-грамматические. Авторы не приводят частотности ошибок каждой группы, ограничиваясь конкретными примерами перевода. В работе также делается вывод о сопоставимом качестве перевода обеими системами (Google всё же демонстрирует меньшее количество грамматических ошибок).

Интересная типология приводится в работе О.Ю. Павловой и М.Ю. Рябовой (Павлова, 2022). В ней исследовался перевод системой Google Translate русскоязычных научно-технических текстов на английский. Они дали следующее определение ошибке перевода: «любой вариант машинного перевода, который был подвергнут редактированию профессиональным переводчиком» и предложили следующую классификацию: группа А — передача смысла (точность, полнота), группа Б — содержание (логика, факты), группа В — язык и стиль (плавность, адаптация, подъязык, идиоматичность, технические ошибки). Не останавливаясь подробно на каждом из подразделов, укажем, что 88% ошибок относились к группе «Язык и стиль». Для ряда подразделов не было выявлено ни одной ошибки: полнота (опущение/добавление информации) и подраздел «факты». Технические ошибки (орфография и пунктуация) также

оказались малочисленными — 4% от общего числа ошибок. В работе также перечислены основные причины нарушения плавности: «1) рамочная структура предложения; 2) чрезмерное использование предлога “of” при переводе многокомпонентных генитивных словосочетаний; 3) многословие; 4) неестественная тема-рематическая структура предложения; 5) избыточное синтаксическое членение; 6) тавтология». В публикации Е.В. Шевчук (Шевчук, 2021), также посвященной переводу на английский язык, отмечалась высокая частотность пропуска или ошибочного использования артиклей и в целом значительная языковая интерференция на уровне синтаксиса.

Выводы

Машинный перевод на современном этапе своего развития прочно вошёл в арсенал инструментов переводчиков и, в сочетании с последующей постобработкой, позволяет решать широкий ряд задач в сфере профессионального, в т.ч. коммерческого, перевода (см. результаты опросов переводчиков и сотрудников переводческих компаний (Нечаева, 2018)), а также в сфере обучения иностранному языку и теории перевода. Среди основных достоинств МП отмечаются его непрерывное развитие и улучшение качества, доступность, простота в использовании, скорость, конфиденциальность и возможность встраиваться в другие программы и ресурсы и в автоматическом режиме переводить, например, сайты и субтитры к видео (Федорова, 2020; Кочеткова, 2017; Ситдикова, 2019; Шушарина, 2021; Шевчук, 2021). Вместе с тем, подавляющее большинство авторов отмечают необходимость осторожного отношения к результатам такого перевода и, следовательно, постредактуры с целью исправления возможных ошибок перевода.

Проблеме типологии ошибок перевода (в т.ч. машинного) посвящена обширная литература (см. также (Бадулин, 2022; Краснова, 2015; Лютковская, 2021; Переходько, 2017; Daems, 2017; Freitag, 2021)), при этом единого подхода на данный момент не существует. Общими ограничениями, свойственными ряду предложенных типологий и, безусловно, отражающими сложность данной проблемы, являются: невозможность выделить единый классификационный признак, пересечение предложенных классов (и, следовательно, невозможность однозначно отнести ошибку к одному из них) и неполное покрытие предложенными классами всего множества возможных ошибок. Тем не менее представляется возможным сделать следующие выводы:

1. Классификация ошибок перевода в значительной степени определяется языковой парой (и направлением перевода в этой паре), типом текста и целью перевода. Иными словами, представляется целесообразным строить классификации, привязанные к определённым параметрам перевода и, возможно, в целях изучения конкретного языкового явления (пример подобной, таргетированной, классификации см. в (Гончаров, 2019; Кирсанова, 2023)).

2. В случае машинного перевода влияние на классификацию ошибок теоретически может также оказывать рассматриваемая система МП, однако, на практике, распределение ошибок в наиболее популярных в России системах Google Translate, «Яндекс Переводчик» и Promt (Красильникова, 2021; Шушарина, 2021; Арестова, 2015) оказывается весьма близким.

3. Классификация ошибок машинного перевода может быть основана на типологии, разработанной на материале переводов, выполненных человеком, но должна содержать существенные изменения, связанные со спецификой МП.

4. Большинство предложенных классификаций являются двухуровневыми. На первом уровне чаще всего выделяются группы лексических (лексико-семантических) ошибок, грамматических ошибок (морфологических и/или синтаксических) и ошибок, связанных с нарушением узуса переводящего языка. Классификации ошибок по причине их происхождения для машинного перевода малочисленны (см., например, классификацию П.Н. Хроменкова (Семенов, 2008: 67–76), в которой все ошибки делятся на ошибки автоматического анализа и автоматического синтеза, и работу (Al Sharou, 2022), в которой обсуждаются ошибки и особенности исходного текста, приводящие к ошибкам перевода).

5. 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. У различных ошибок разный «вес» с точки зрения их способности приводить к коммуникативному сбою. Наиболее значимыми являются ошибки, относящиеся к группе лексико-семантических.

Список литературы:

Алексеева И.С. Введение в переводоведение. СПб.: Филологический факультет СПбГУ, 2008. 368 с.

Арестова А.А. Сравнительный анализ систем машинного перевода // Вестник Волгоградского государственного университета. Серия 9, 2015.

№ 13. С. 105–109. URL: <https://yrw.jvolsu.com/attachments/article/340/Arestova.pdf> (дата обращения: 21.11.2023).

Бадулин Д.Е., Бужинский В.В. Типология ошибок систем машинного перевода // Материалы XXIV Всероссийской студенческой научно-практической конференции Нижневартовского государственного университета (05–06 апреля 2022 года, Нижневартовск), 2022. Т. 7. С. 76–80. URL: <https://www.elibrary.ru/item.asp?id=48570053> (дата обращения: 21.11.2023).

Беляева Л.Н. Машинный перевод в современной технологии процесса перевода // Известия РГПУ имени А.И. Герцена, 2022. № 203. С. 22–30. URL: <https://cyberleninka.ru/article/n/mashinnyu-perevod-v-sovremennoy-tehnologii-protssessa-perevoda/viewer> (дата обращения: 03.12.2023).

Бузаджи Д.М., Гусев В.В., Ланчиков В.К., Псурцев Д.В. Новый взгляд на классификацию переводческих ошибок. М.: Всерос. центр переводов, 2009. 120 с.

Воронович В.В. Машинный перевод: конспект лекций по специальности «Компьютерная лингвистика, Компьютерное обучение языкам». Минск: Белорусский государственный университет, 2013. 39 с.

Гарбовский Н.К. Теория перевода. М.: Изд-во Моск. ун-та, 2007. 544 с.

Гончаров А.А., Бунтман Н.В., Нуриев В.А. Ошибки в машинном переводе: проблемы классификации // Системы и средства информатики, 2019. № 3. С. 92–103. URL: https://www.researchgate.net/publication/340792802_MACHINE_TRANSLATION_ERRORS_PROBLEMS_OF_CLASSIFICATION (дата обращения: 20.11.2023).

Гу Цзюньлин, Хуан Чжунлянь. Система классификации переводческих ошибок // Вестник Московского университета. Серия 22. Теория перевода, 2016. № 3. С. 26–40. URL: <https://www.elibrary.ru/item.asp?id=44541285> (дата обращения: 20.11.2023).

Зайцева М.А. О типологии ошибок гибридной системы машинного перевода «Яндекс Переводчик» // Молодой учёный, 2023. № 45 (492). С. 38–46. URL: <https://moluch.ru/archive/492/107574/> (дата обращения: 20.11.2023).

Иванченко Т.А. Ошибки в машинном переводе с немецкого языка на русский (на материале статей немецкоязычных СМИ и текстов их переводов) // Учёные записки Санкт-Петербургского университета технологий управления и экономики, 2021. № 4 (76). С. 30–41. URL: <https://www.elibrary.ru/item.asp?id=47400297> (дата обращения: 20.11.2023).

Кирсанова М.А. Машинный перевод: а нужен ли переводчик? // Сборник материалов конференции «Иностранный язык в профессиональной сфере: педагогика, лингвистика, межкультурная коммуникация», 2023. Т. 1. С. 84–90. URL: <https://www.elibrary.ru/item.asp?id=53761410> (дата обращения: 21.11.2023).

Кочеткова Н.С., Ревина Е.В. Особенности машинного перевода // Филологические науки. Вопросы теории и практики, 2017. № 6 (72), Ч. 2. С. 106–109. URL: https://www.gramota.net/articles/issn_1997-2911_2017_6-2_30.pdf (дата обращения: 20.11.2023).

Красильникова В.Г. Анализ качества машинного перевода системами Google Translate и Яндекс. Переводчик (на материале отрывка из научно-популярного издания по медицине) // Молодой учёный, 2021. № 23 (365). С. 492–494. URL: <https://moluch.ru/archive/365/81991/> (дата обращения: 20.11.2023).

Краснова Т.И., Ванюшин И.С. Machine translation error analysis // Молодой учёный, 2015. № 9 (89). С. 124–126. URL: <https://moluch.ru/archive/89/18290/> (дата обращения: 21.11.2023).

Лютковская В.Л. Типичные ошибки в машинном переводе и оптимизация процесса постредактирования // Языковая личность и перевод: материалы VI Междунар. науч.-образоват. форума молодых переводчиков имени Д.О. Половцева, 2021. С. 55–58. URL: <https://elib.bsu.by/handle/123456789/276847> (дата обращения: 21.11.2023).

Миньяр-Белоручев Р.К. Теория и методы перевода. М.: Московский лицей, 1996. 208 с.

Нечаева Н.В., Светова С.Ю. Постредактирование машинного перевода как актуальное направление подготовки переводчиков в вузах // Вопросы методики преподавания в вузе, 2018. Т. 7. № 25. С. 64–72. URL: <https://cyberleninka.ru/article/n/postredaktirovanie-mashinnogo-perevoda-kak-aktualnoe-napravlenie-podgotovki-perevodchikov-v-vuzah/viewer> (дата обращения: 21.11.2023).

Павлова О.Ю., Рябова М.Ю. Типология ошибок в машинном переводе // Полилингвальность современной культуры. Сборник статей Международной научной конференции «Филологическая наука и образование в Кузбассе», 2022. С. 247–258. URL: <https://www.elibrary.ru/item.asp?id=50357977> (дата обращения: 21.11.2023).

Переходько И.В., Мячин Д.А. Оценка качества компьютерного перевода // Вестник Оренбургского государственного университета, 2017. № 2. С. 92–96. URL: http://vestnik.osu.ru/2017_2/18.pdf (дата обращения: 21.11.2023).

Семенов А.Л. Современные информационные технологии и перевод: учебное пособие. М.: Академия, 2008. 224 с.

Ситдикова Ф.Б., Хисамова В.Н., Усманов Т.Р. «Передняя часть правого телёнка», или ошибки машинного перевода (на основе анализа автоматических субтитров) // Балтийский гуманитарный журнал, 2019. Т. 8. № 2 (27), Ч. 2. С. 324–327. URL: <https://cyberleninka.ru/article/n/perednyaya-chast-pravogo-telenka-ili-oshibki-mashinnogo-perevoda-na-osnove-analiza-avtomaticheskikh-subtitrov/viewer> (дата обращения: 20.11.2023).

Терентьева Д.М., Васильев Л.Г. К проблеме классификации переводческих ошибок // Вестник Калужского университета, 2020. № 3 (48). С. 56–60. URL: <https://www.elibrary.ru/item.asp?id=44541285> (дата обращения: 20.11.2023).

Федорова Н.Ю. Типология переводческих ошибок как основа для разработки методики обучения переводу научных текстов // Профессионально-ориентированное обучение языкам: реальность и перспективы. Сборник статей участников Ежегодной всероссийской научно-практической

конференции с международным участием, 2020. С. 49–55. URL: <https://www.elibrary.ru/item.asp?id=42990073> (дата обращения: 20.11.2023).

Шамова Н.В. Разграничение понятий «эквивалентность» и «адекватность» в переводе // Вестник МГУ. Серия 19. Лингвистика и межкультурная коммуникация, 2005. № 2. С. 171–180. URL: <https://www.elibrary.ru/item.asp?id=42990073> (дата обращения: 07.12.2023).

Шевчук Е.В., Никифорова Ж.А. Постредактирование и типичные ошибки в автоматизированном переводе научно-публицистических текстов // Вопросы методики преподавания в вузе, 2021. Т. 10. № 39. С. 46–54. URL: <https://cyberleninka.ru/article/n/postredaktirovanie-i-tipichnye-oshibki-v-avtomatizirovannom-perevode-nauchno-publitsisticheskikh-tekstov/viewer> (дата обращения: 21.11.2023).

Шушарина Г.А., Петрунина Ж.В. Сопоставительный анализ текстов онлайн-переводов // Вестник Челябинского государственного университета, 2021. № 4 (450). Филологические науки. Вып. 124. С. 185–192. URL: <https://cyberleninka.ru/article/n/sopostavitelnyu-analiz-tekstov-onlayn-perevodov> (дата обращения: 21.11.2023).

Щекина Л.Э. Особенности онлайн-перевода с французского языка на русский (на примере работы сервиса Яндекс-переводчик) // Сборник материалов конференции “Research work” (Кемерово, 31.08.2021), 2021. С. 31–34. URL: <https://www.rgph.vsu.ru/ru/science/sss/reports> (дата обращения: 01.12.2023).

Al Sharou K., Specia L. (2022) A Taxonomy and Study of Critical Errors in Machine Translation. Proceedings of the 23rd Annual Conference of the European Association for Machine Translation, pp. 171–180. URL: <https://aclanthology.org/2022.eamt-1.20/> (дата обращения: 21.11.2023).

Daems J., Vandepitte S., Hartsuiker R.J., Macken L. (2017) Identifying the Machine Translation Error Types with the Greatest Impact on Post-editing Effort. *Frontiers in Psychology*. Vol. 8. URL: <https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01282/full> (дата обращения: 21.11.2023).

Freitag M., Foster G., Grangier D., Ratnakar V., Tan Q., Macherey W. (2021) Experts, Errors, and Context: A Large-Scale Study of Human Evaluation for Machine Translation. *Transactions of the Association for Computational Linguistics*. Vol. 9, pp. 1460–1474. URL: https://direct.mit.edu/tacl/article/doi/10.1162/tacl_a_00437/108866/Experts-Errors-and-Context-A-Large-Scale-Study-of (дата обращения: 21.11.2023).

Vilar D., Xu J., D’Haro L., Ney H. (2006) Error analysis of statistical machine translation output. Proceedings of the 5th Conference (International) on Language Resources and Evaluation. Italy, Genoa: European Language Resources Association, pp. 697–702. URL: https://pemt.ru/wp-content/uploads/2023/03/Error_analysis_of_statistical_machine_translation_output.pdf (дата обращения: 20.11.2023).

Wang H., Wu H., He Z., Huang L., Ward Church K. (2022) Progress in Machine Translation // *Engineering*. Vol. 18, pp. 143–153. URL: <https://www.sciencedirect.com/science/article/pii/S2095809921002745> (дата обращения: 03.12.2023).

References

Alekseeva I.S. (2008) *Vvedenie v perevodovedenie = Introduction to translation studies*. St. Petersburg: Faculty of Philology, SPBU. 368 p. (In Russian).

Al Sharou K., Specia L. (2022) A Taxonomy and Study of Critical Errors in Machine Translation. // Proceedings of the 23rd Annual Conference of the European Association for Machine Translation, pp. 171–180. Available at: <https://aclanthology.org/2022.eamt-1.20/> (accessed: 21.11.2023).

Arestova A.A. (2015) Sravnitel'nyy analiz sistem mashinnogo perevoda = Comparative analysis of machine translation systems. *Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 9*. No. 13, pp. 105–109. Available at: <https://yrv.jvolsu.com/attachments/article/340/Arestova.pdf> (accessed: 21.11.2023) (In Russian).

Badulin D.E., Buzhinskiy V.V. (2022) Tipologiya oshibok sistem mashinnogo perevoda = Typology of machine translation system errors. Proceedings of the XXIV All-Russian Student Scientific and Practical Conference of Nizhnevartovsk State University (April 05–06, 2022, Nizhnevartovsk). Vol. 7, pp. 76–80. Available at: <https://www.elibrary.ru/item.asp?id=48570053> (accessed: 21.11.2023) (In Russian)

Belyaeva L.N. (2022) Mashinnyy perevod v sovremennoy tekhnologii protsessa perevoda = Machine translation and modern translation technology. *Izvestiya RGPU im. A.I. Gertsena*. No. 203, pp. 22–30. Available at: <https://cyberleninka.ru/article/n/mashinnyy-perevod-v-sovremennoy-tehnologii-protsess-a-perevoda/viewer> (accessed: 03.12.2023) (In Russian).

Buzadzhi D.M., Gusev V.V., Lanchikov V.K., Psurtsev D.V. (2009) *Novyy vzglyad na klassifikatsiyu perevodcheskikh oshibok = A new look at the classification of translation errors*. Moscow: Vserossiyskiy tsentr perevodov. 120 p. (In Russian).

Daems J., Vandepitte S., Hartsuiker R.J., Macken L. (2017) Identifying the Machine Translation Error Types with the Greatest Impact on Post-editing Effort. *Frontiers in Psychology*. Vol. 8. Available at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01282/full> (accessed: 21.11.2023).

Fedorova N.Yu. (2020) Tipologiya perevodcheskikh oshibok kak osnova dlya razrabotki metodiki obucheniya perevodu nauchnykh tekstov = Typology of translation errors as a basis for developing methods of teaching translation of scientific texts. Proceedings of the annual All-Russian scientific and practical conference with international participation, pp. 49–55. Available at: <https://www.elibrary.ru/item.asp?id=42990073> (accessed: 20.11.2023) (In Russian).

Freitag M., Foster G., Grangier D., Ratnakar V., Tan Q., Macherey W. (2021) Experts, Errors, and Context: A Large-Scale Study of Human Evaluation for Machine Translation. *Transactions of the Association for Computational Linguistics*. vol. 9. pp. 1460–1474. Available at: https://direct.mit.edu/tacl/article/doi/10.1162/tacl_a_00437/108866/Experts-Errors-and-Context-A-Large-Scale-Study-of (accessed: 21.11.2023).

Garbovskiy N.K. (2007) *Teoriya perevoda = Translation Theory*. Moscow: Moscow St. Univ. Publ. 544 p. (In Russian).

Goncharov A.A., Buntman N.V., Nuriev V.A. (2019) Oshibki v mashinnom perevode: problemy klassifikatsii = Machine translation errors: problems of classification. *Sistemy i sredstva informatiki*. No. 3, pp. 92–103. Available at: https://www.researchgate.net/publication/340792802_MACHINE_TRANSLATION_ERRORS_PROBLEMS_OF_CLASSIFICATION (accessed: 20.11.2023) (In Russian).

Gu Tzzyun'lin, Khuan Chzhunlyan'. (2016) Sistema klassifikatsii perevodcheskikh oshibok = Towards the problem of classification of translation mistakes. *Vestnik Moskovskogo universiteta. Seriya 22. Teoriya perevoda*. No. 3, pp. 26–40. Available at: <https://www.elibrary.ru/item.asp?id=44541285> (accessed: 20.11.2023) (In Russian).

Ivanchenko T.A. (2021) Oshibki v mashinnom perevode s nemetskogo yazyka na russkiy (na materiale statey nemetskoazychnykh SMI i tekstov ikh perevodov) = Errors in machine translation from german into russian (based on articles of german-language media and their translations). *Uchenye zapiski Sankt-peterburgskogo universiteta tekhnologiy upravleniya i ekonomiki*. No. 4 (76), pp. 30–41. Available at: <https://www.elibrary.ru/item.asp?id=47400297> (accessed: 20.11.2023) (In Russian).

Kirsanova M.A. (2023) Mashinnyy perevod: a nuzhen li perevodchik? = Machine translation: do we need a translator? Proceedings of conference “Foreign language in the professional sphere: pedagogy, linguistics, intercultural communication”. Vol. 1, pp. 84–90. Available at: <https://www.elibrary.ru/item.asp?id=53761410> (accessed: 21.11.2023) (In Russian).

Kochetkova N.S., Revina E.V. (2017) Osobennosti mashinnogo perevoda = The features of machine translation. *Filologicheskie nauki. Voprosy teorii i praktiki*. No. 6 (72), part 2, pp. 106–109. Available at: https://www.gramota.net/articles/issn_1997-2911_2017_6-2_30.pdf (accessed: 20.11.2023) (In Russian).

Krasil'nikova V.G. (2021) Analiz kachestva mashinnogo perevoda sistemami Google Translate i Yandeks.Perevodchik (na materiale otryvka iz nauchno-populyarnogo izdaniya po meditsine) = Analysis of the quality of machine translation by Google Translate and Yandex.Translate systems (based on an excerpt from a popular scientific publication on medicine). *Molodoy uchenyy*. No. 23 (365), pp. 492–494. Available at: <https://moluch.ru/archive/365/81991/> (accessed: 20.11.2023) (In Russian).

Krasnova T.I., Vanyushin I.S. (2015) Machine translation error analysis. *Molodoy uchenyy*. No. (89), pp. 124–126. Available at: <https://moluch.ru/archive/89/18290/> (accessed: 21.11.2023).

Lyutkovskaya V.L. (2021) Tipichnye oshibki v mashinnom perevode i optimizatsiya protsessa postredaktirovaniya = Typical errors in machine translation and optimization of the post-editing process. Proceedings of the VI International Scientific and Educational Forum for Young Translators, pp. 55–58. Available at: <https://elib.bsu.by/handle/123456789/276847> (accessed: 21.11.2023) (In Russian).

Min'yar-Beloruhev R.K. (1996) Teoriya i metody perevoda = Theory and methods of translation. Moscow: Moskovskiy litsey. 208 p. (In Russian).

Nechaeva N.V., Svetova S.Yu. (2018) Postredaktirovanie mashinnogo perevoda kak aktual'noe napravlenie podgotovki perevodchikov v vuzakh =

Post-editing machine translation as a new activity for teaching translation at universities. *Voprosy metodiki prepodavaniya v vuze*. Vol. 7. No. 25, pp. 64–72. Available at: <https://cyberleninka.ru/article/n/postredaktirovanie-mashinnogo-perevoda-kak-aktualnoe-napravlenie-podgotovki-perevodchikov-v-vuzah/viewer> (accessed: 21.11.2023) (In Russian).

Pavlova O.Yu., Ryabova M.Yu. (2022) Tipologiya oshibok v mashinnom perevode = Proceedings of the International Scientific Conference “Philological science and Education in Kuzbass”, pp. 247–258. Available at: Typology of errors in machine translation (accessed: 21.11.2023) (In Russian).

Perekhod'ko I.V., Myachin D.A. (2017) Otsenka kachestva komp'yuternogo perevoda = Evaluation of the quality of machine translation. *Vestnik Orenburgskogo gosudarstvennogo universiteta*. No. 2, pp. 92–96. Available at: http://vestnik.osu.ru/2017_2/18.pdf (accessed: 21.11.2023) (In Russian).

Shamova N.V. (2005) Razgranichenie ponyatiy “ekvivalentnost” i “adekvatnost” v perevode = Differentiation of the concepts of “equivalence” and “adequacy” in translation. *Vestnik MGU. Seriya 19. Lingvistika i mezhkul'turnaya kommunikatsiya*. No. 2, pp. 171–180. Available at: <https://cyberleninka.ru/article/n/razgranichenie-ponyatiy-ekvivalentnost-i-adekvatnost-v-perevode/viewer> (accessed: 07.12.2023) (In Russian).

Shevchuk E.V., Nikiforova Zh.A. (2021) Postredaktirovanie i tipichnye oshibki v avtomatizirovannom perevode nauchno-publitsisticheskikh tekstov = Post-editing and typical mistakes in the computer-aided translation of academic, scientific, and journalistic texts. *Voprosy metodiki prepodavaniya v vuze*. Vol. 10. No. 39, pp. 46–54. Available at: <https://cyberleninka.ru/article/n/postredaktirovanie-i-tipichnye-oshibki-v-avtomatizirovannom-perevode-nauchno-publitsisticheskikh-tekstov/viewer> (accessed: 21.11.2023) (In Russian).

Shusharina G.A., Petrunina Zh.V. (2021) Sopostavitel'nyy analiz tekstov onlayn-perevodov = Comparative analysis of online translations texts. *Vestnik Chelyabinskogo gosudarstvennogo universiteta*. No. 4 (450). *Filologicheskie nauki*. Issue 124, pp.185–192. Available at: <https://cyberleninka.ru/article/n/sopostavitelnyy-analiz-tekstov-onlayn-perevodov> (accessed: 21.11.2023) (In Russian).

Shchekina L.E. (2021) Osobennosti onlayn-perevoda s frantsuzskogo yazyka na russkiy (na primere raboty servisa Yandeks-perevodchik) = Features of online translation from French into Russian (using Yandex Translator service). Proceedings of the conference “Research work”, pp. 31–34. Available at: <https://www.rgph.vsu.ru/science/sss/reports> (accessed: 01.12.2023) (In Russian).

Semenov A.L. (2008) Sovremennye informatsionnye tekhnologii i perevod = Modern information technologies and translation. Moscow: Akademiya. 224 p. (In Russian).

Sitdikova F.B., Khisamova V.N., Usmanov T.R. (2019) “Perednyaya chast' pravogo telenka”, ili oshibki mashinnogo perevoda (na osnove analiza avtomaticheskikh subtitrov) = Machine translation errors (on the basis of analysis of automatic subtitles). *Baltiyskiy gumanitarnyy zhurnal*. Vol. 8. No. (27), part 2, pp. 324–327. Available at: <https://cyberleninka.ru/article/n/perednyaya-chast-pravogo-telenka-ili-oshibki-mashinnogo-perevoda-na-osnove-analiza-avtomaticheskikh-subtitrov/viewer> (accessed: 20.11.2023) (In Russian).

Terent'eva D.M., Vasil'ev L.G. (2020) K probleme klassifikatsii perevodcheskikh oshibok = Towards the problem of classification of translation mistakes. Vestnik Kaluzhskogo universiteta. No. 3 (48), pp. 56–60. Available at: <https://www.elibrary.ru/item.asp?id=44541285> (accessed: 20.11.2023) (In Russian).

Vilar D., Xu J., D'Haro L., Ney H. (2006) Error analysis of statistical machine translation output. Proceedings of the 5th Conference (International) on Language Resources and Evaluation. Italy, Genoa: European Language Resources Association, pp 697–702. Available at: https://pemt.ru/wp-content/uploads/2023/03/Error_analysis_of_statistical_machine_translation_output.pdf (accessed: 20.11.2023).

Voronovich V.V. (2013) Mashinnyy perevod: konspekt lektsiy po spetsial'nosti 'Komp'yuternaya lingvistika, Komp'yuternoe obuchenie yazykam' = Machine translation: lectures on Computational Linguistics, Computer language learning. Minsk: BSU. 39 p. (In Russian).

Wang H., Wu H., He Z., Huang L, Ward Church K. (2022) Progress in Machine Translation. Engineering. Vol. 18, pp. 143–153. Available at: <https://www.sciencedirect.com/science/article/pii/S2095809921002745> (accessed: 03.12.2023).

Zaytseva M. A. (2023) O tipologii oshibok gibridnoy sistemy mashinnogo perevoda "Yandeks Perevodchik". Molodoy uchenyy. No. 45 (492), pp. 38–46. Available at: <https://moluch.ru/archive/492/107574/> (accessed: 20.11.2023) (In Russian).

ИНФОРМАЦИЯ ОБ АВТОРЕ:

Гребенюк Андрей Владимирович — аспирант кафедры математической лингвистики филологического факультета СПбГУ; Россия, 199034, Санкт-Петербург, Университетская наб., д. 7–9; yax-light@yandex.ru

ABOUT THE AUTHOR:

Andrey V. Grebenyuk — Postgraduate Student at the Department of Mathematical Linguistics, St. Petersburg State University; 7–9 Universitetskaya nab., St Petersburg 199034, Russia; yax-light@yandex.ru

Конфликт интересов: положения и точки зрения, представленные в данной статье, принадлежат автору и не обязательно отражают позицию какой-либо организации или российского научного сообщества. Автор заявляет об отсутствии конфликта интересов.

Conflict of interests: the ideas and opinions presented in this article entirely belong to the author and do not necessarily reflect the position of any organization or the Russian scientific community as a whole. The author states that there is no conflict of interests.