

Раздел I. ЯЗЫКОЗНАНИЕ

УДК 81.255.2

DOI: 10.15593/2224-9389/2021.1.1

И.А. Барина¹, И.Г. Овчинникова²

Получено: 14.02.2021

Принято: 06.03.2021

Опубликовано: 28.04.2021

¹ Пермский национальный
исследовательский политехнический
университет, Пермь, Российская Федерация,
² Первый Московский государственный
медицинский университет им. И.М. Сеченова,
Москва, Российская Федерация

ВЛИЯНИЕ НОВЫХ ПЕРЕВОДЧЕСКИХ ТЕХНОЛОГИЙ НА РАСПОЗНАВАНИЕ И КЛАССИФИКАЦИЮ ОШИБОК ПЕРЕВОДА

Обсуждаются различные подходы к классификации ошибок перевода в науке и переводческой индустрии, обусловленные разной степенью внимания к новым переводческим технологиям. Новые технологии предлагают различные системы машинного перевода, компьютерно-переводческие инструменты, автоматические системы распознавания ошибок и оценки качества перевода. Приводится общая характеристика работы переводчика на облачной платформе, оснащенной САТ-инструментами, что подразумевает выполнение проекта и общение с коллегами в сфере компьютерно-опосредованной коммуникации (КОК). Доступность разнообразных инструментов, возможность интегрировать машинный перевод (МП) в транслят и редактировать его, необходимость использовать английский язык для сетевого поиска нужной информации обуславливают расширение набора компетенций современного переводчика. Усложнение рабочей среды и многозадачность создают условия для появления новых переводческих ошибок. Анализ основных классификаций ошибок перевода позволяет определить, насколько полно в научных и прикладных классификациях представлены потенциальные несовершенства текста перевода (ТП). Прикладные классификации ориентированы на стандарты МП и разработку автоматических систем распознавания ошибок. В такого рода классификациях не рассматриваются ошибки дискурсивные, логические и фактические, поскольку для обнаружения этих ошибок необходимо обращение к фоновым знаниям и общей картине мира, что пока не доступно для действующих автоматических систем. Мы отличаем ошибки, распознаваемые автоматическими системами оценки качества перевода, от тех, с которыми приходится иметь дело профессиональному переводчику и пост-редактору машинного перевода. Основываясь на сопоставлении классификаций, можно выделить: 1) ошибки внешней адекватности текста, отражающие несоответствие общей картине мира и законам логики; 2) ошибки внутритекстовой адекватности, искажающие содержание исходного текста в ТП, среди которых выделяются ошибки в терминологии, лексическом выборе; 3) дискурсивные ошибки, вызванные нарушением правил и норм дискурса принимающей культуры; 4) ошибки связности ТП, разрушающие стиливое разнообразие, грамматическую правильность, нормативное оформление ТП (пунктуацию, орфографию, форматирование). Поскольку автоматические системы не распознают ошибки связности, терминологические и лексические ошибки, профессиональный переводчик является незаменимым специалистом в подготовке высококачественного ТП.

Ключевые слова: *ошибки переводчиков, машинный перевод, компьютерные инструменты перевода, компьютерно-опосредованная коммуникация, классификация ошибок.*

I.A. Barinova¹, I.G. Ovchinnikova²

Received: 14.02.2021

Accepted: 06.03.2021

Published: 28.04.2021

¹ Perm National Research Polytechnic University,
Perm, Russian Federation

² I.M. Sechenov First Moscow State Medical
University, Moscow, Russian Federation

THE IMPACT OF THE COMPUTER-ASSISTED TRANSLATION ON THE RECOGNITION AND CLASSIFICATION OF TRANSLATION ERRORS

The article discusses various approaches to the classification of translation errors offered by academic literature and translation industry, based on differing perspectives on new translation technologies. The latter provide various machine translation systems, computer translation tools, automatic error recognition systems and translation quality assessment. The article gives a general description of the translator's work on a cloud platform equipped with CAT tools, which implies the implementation of the project and communication with colleagues in the field of computer-mediated communication (CMC). The availability of various tools, the ability to integrate machine translation (MT) into translates and edit them, and the need to use English for online search of the necessary information determine the expansion of the set of competencies of a modern translator. The complexity of the working environment and multitasking create conditions for the appearance of new translation errors. The analysis of the main classifications of translation errors allows us to determine how fully the potential imperfections of the target text (TT) are presented in pure and applied classifications. Applied classifications are focused on MT standards and the development of automatic error recognition systems. Such classifications ignore discursive, logical, and factual errors, since in order to detect ones, it is necessary to refer to the background knowledge and the general picture of the world, which is not yet available for existing automatic systems. We distinguish errors that are recognized by automatic translation quality assessment systems from those that a professional translator and a post-editor of machine translation have to deal with. Based on the comparison of classifications, we can differentiate between: (1) errors of external adequacy of the text, reflecting the discrepancy between the general picture of the world and the laws of logic; (2) errors of internal adequacy, distorting the content of the source text in the TT, there being errors in terminology, lexical choice, inter alia; (3) discursive errors caused by violation of the rules and norms of the discourse of the target culture; (4) errors in the coherence of the TT that destroy the stylistic diversity, grammatical correctness, and normative design of the TT (punctuation, spelling, and formatting). Since automatic systems do not recognize connectivity errors, terminological and lexical errors, a professional translator is an indispensable specialist in producing a high-quality TT.

Keywords: *translation errors, machine translation, computer-assisted translation, computer-mediated communication, error classification.*

Современные технологии трансформируют рабочую среду переводчиков в компьютерно-опосредованную коммуникацию (КОК) с помощью облачных платформ, оснащенных компьютерными переводческими инструментами (САТ-инструменты). Такие платформы помогают организовать взаимоотношения переводчика с заказчиком, предлагая инструменты для поиска проекта, выполнения перевода и локализации текста перевода (ТП), обсуждения результатов с заказчиком и перевода платежей. Выполненный на таких платформах перевод обозначают термином САТ (сокращение англоязычного термина computer-assisted translation – перевод, выполненный при помощи

компьютерных инструментов). Платформы CAT объединяют все переводческие инструменты: словари, тезаурусы, конкордансы, переводческую память, а также предполагают возможность интеграции машинного перевода (МП) в транслят по ходу выполнения работы.

Поскольку в современном обществе зачастую необходимо быстро предоставить точный и качественный перевод в различных отраслях промышленности и сферах социального взаимодействия, платформа CAT предлагает приемлемое практическое решение стоящих перед переводчиками задач. Переводческие инструменты пополняются автоматическими системами распознавания ошибок и оценки качества перевода [1]. Такие системы используют как в системе МП перед выводом результата, так и отдельно, для оценки CAT. Современные автоматические системы распознавания ошибок пока не позволяют выявить все типы ошибок перевода. Результаты МП и CAT нуждаются в постредактировании профессиональными переводчиками, чтобы ТП отвечал требованиям заказчика, правилам и нормам общения в принимающей культуре. Интеграция сегментов МП в документы, переведенные вручную, увеличивает вероятность ошибок из-за различий в стратегиях перевода. В новой технологической среде переводчик приобрел дополнительную специальность – постредактора МП и CAT. Согласно образовательному стандарту ЕС, постредактирование и доработка всех ТП перед их отправкой заказчику входят в набор профессиональных компетенций переводчика наряду с умением вести переговоры с клиентами [2].

КОК и новые технологии оказывают давление на переводчиков, поскольку добавляют новые виды деятельности и обязанности при работе на платформах CAT. Профессиональные переводчики, работающие в глобальных корпорациях и получающие поддержку от своих коллег в соответствующей сфере (юридической, финансовой, промышленной и т.п.), умеют пользоваться CAT, применяют автоматические системы оценки и инструменты постредактирования. Тем не менее они в основном предпочитают переводить с нуля и редактировать выполненный вручную перевод, а не использовать CAT-инструменты [3]. Практикующие переводчики, связанные с переводческими агентствами, работают в основном с CAT, в том числе и на облачных платформах; они нередко заняты в коллективной работе над онлайн-проектами.

КОК ограничивает выбор коммуникативных средств. Профессиональная деятельность онлайн, проектная работа в команде привносят новые нормы и стандарты в профессиональное общение [4]. Работая на платформе CAT, переводчики сталкиваются с проблемами при сотрудничестве с членами рабочего коллектива, которые зачастую представляют разные культуры,

с одной стороны; с другой стороны, переводчикам не хватает обратной связи с руководителем проекта и объяснения обоснованности принимаемых решений [5]. Таким образом, платформа CAT облегчает перевод, но может спровоцировать новые ошибки из-за различий в культуре профессионального общения переводчиков, работающих над одним проектом, из-за недостаточной обратной связи, необходимости интегрировать результаты МП и сегменты, сохраненные переводческой памятью, в ТП и многозадачности.

Ошибки в CAT также возникают из-за неудачного дизайна платформ и недостаточной компетентности переводчика в компьютерных технологиях [6]. Ошибки, вызванные неудачным дизайном различных компьютерных программ и некомпетентным использованием компьютерных технологий, описываются и классифицируются в работе Д.Д. Вудз с соавторами, вышедшей в 1994 году [7]. Стремительное распространение компьютерных инструментов, программного обеспечения, технологий, облегчающих взаимодействие компьютера и человека, приводит к необходимости пересмотреть классификацию ошибок перевода. Среди категорий и классов ошибок перевода, возникающих под влиянием интеграции МП и CAT, важно определить соотношение ошибок, распознаваемых автоматическими системами, и тех, которые может заметить профессиональный переводчик.

При редактировании CAT или МП переводчик проверяет информацию, представленную в исходном тексте (ИТ). Необходимые данные можно обнаружить в Интернете. Однако ознакомление с данными и их критический анализ предполагают использование английского языка, поскольку 54 % из 10 миллионов лучших веб-сайтов предлагают свой контент на английском языке. Таким образом, английский язык становится *lingua franca* в КОК. При редактировании ТП переводчику важно глубоко понимать содержание исходного текста, которое необходимо соотносить с принимающей культурой. Передача сообщения, заключенного в исходном тексте, не только напрямую зависит от точности, связности и читабельности ТП, но и включает в себя культурные ассоциации, обычаи и жанровые модели принимающей культуры [8]. Таким образом, для практикующего переводчика владение CAT-инструментами, английским языком и навыками распределенной деятельности в КОК играет решающую роль.

Новые технологии перевода, предлагая новые инструменты, вынуждают переводчика осваивать работу на облачных платформах (интегрируя МП и фрагменты, извлеченные из коллективной переводческой памяти), расширять репертуар рабочих языков, адаптироваться к условиям компьютерно-опосредованной коммуникации.

Цели, материал и методы исследования

Основная цель статьи – выявление ошибок перевода, возникающих в новых технологических условиях, неуловимых для систем автоматического распознавания ошибок, на основе сопоставления различных классификаций переводческих ошибок. Мы намерены ответить на два исследовательских вопроса:

1) Насколько существенны различия в классификациях ошибок перевода, предложенных специалистами по переводоведению и используемыми в академической среде, от классификаций прикладного характера, разработанных для оценки качества МП, переводческой памяти, САТ и автоматических систем распознавания ошибок.

2) Какие ошибки не распознаются автоматическими системами и в силу этого требуют пристального внимания переводчиков и постредакторов.

Исследование выполнено на материале научных публикаций на русском и английском языках. Примеры ошибок перевода извлечены из авторской базы переводческих ошибок [6], а также из опубликованных материалов [9, 10]. Мы применяли дискурсивный и семантический анализ, а также сопоставительный метод.

Различные подходы к оценке ошибок

Анализируя тексты переводов на платформе САТ, переводчик или постредактор комментирует (аннотирует) ошибки и предлагает правильный вариант. В теоретических работах по переводоведению, в публикациях разработчиков систем МП и практикующих переводчиков представлены различные основания для аннотирования и классификации ошибок. Ошибки в переводе принадлежат человеку и системам МП; в теоретических работах проанализирован профессиональный перевод, выполненный вручную, в то время как в прикладных публикациях обсуждается качество МП и САТ. Таким образом, классификации ошибок перевода отражают материал, с которым имеют дело исследователи.

Классификации переводческих ошибок в академических кругах

В академических кругах различия в теоретических подходах к подготовке профессиональных переводчиков приводят к расхождениям в классификациях ошибок. Классификации отражают три процедуры в процессе перевода и локализации: анализ ИТ, выбор стратегий его преобразования в ТП и редактирование текста перевода [11]. Рассмотрим соотношение ошибок и первых двух процедур – анализа ИТ и выбора стратегии преобразования. Последняя процедура – редактирование текста – подразумевает обнаружение ошибок связности и оформления семантической целостности ТП.

Анализ ИТ включает проверку информации, выделение жанровых особенностей, стилистических приемов и фигур речи. Сбои в этой процедуре

приводят к неадекватному ТП и искажению исходного сообщения. Вертикальный подход к переводу предполагает понимание исходного текста перед тем, как переформулировать заключенное в нем сообщение средствами языка перевода (ЯП) и выбрать необходимую степень локализации для ТП. Вертикальный подход связан с нисходящей и нелинейной стратегиями перевода. Горизонтальный подход предполагает, что единицы исходного языка (ИЯ) вызывают соответствующие структуры ЯП в памяти переводчика, а процесс транскодирования сообщения приводит к пониманию исходного текста. Горизонтальный подход соответствует восходящей и линейной стратегиям перевода. Между тем горизонтальный подход соответствует базовому алгоритму МП, который способен обрабатывать исходный текст на уровне целого документа только посредством его последовательного посегментного перекодирования на ЯП, от первого сегмента до последнего. Таким образом, человек-переводчик имеет преимущество в использовании как горизонтального, так и вертикального подхода в переводе.

Выбор стратегии трансформации ИТ зависит от степени необходимой локализации ТП. Выбор нерелевантной стратегии наносит ущерб восприятию ТП в принимающей культуре, что приводит к непониманию исходного сообщения [12]. Ущерб обнаруживается в связности, точности и адекватности ТП. Подклассы ошибок связности обозначаются в соответствии с характеристикой исходного и целевого языков с позиций контрастивной лингвистики. Ошибки в анализе ИТ и ошибочная стратегия перевода влияют на точность и адекватность ТП. Все исследователи учитывают сбой анализа ИТ и выбора стратегии его трансформации при классификации переводческих ошибок.

Классификация ошибок имеет важное значение при оценке качества перевода. Оценка качества предполагает различение ошибок в точности, стиле, грамматике и форматировании [13]. Выделенные классы ошибок включены в типологию ошибок перевода, однако их подклассы различаются для разных языковых пар. Сравнивая ошибки в переводах, выполненных вручную, с результатами машинного перевода, исследователи выделяют ошибки, приводящие к неприемлемому результату, и ошибки адекватности [14]. Ошибки приемлемости отражают расхождение в особенностях дискурса – ТП не соответствует дискурсивным нормам принимающей культуры. Ошибки адекватности выявляются путем сопоставления ИТ и ТП, они показывают несоответствие между исходным сообщением и его передачей в принимающей культуре. В [11] показано, что ошибки адекватности проявляются как в искажении отношения текста к культуре и фоновым знаниям, так и в разрушении внутренней семантической структуры ТП. «Внешние» ошибки адекватности приписывают ИТ ложные логические и культурные связи, разрушают связь его содержания с фоновыми знаниями, искажают авторскую точку зрения. Например, при переводе с иврита на русский переводчик заменяет

мнение исследователей о происхождении жителей древней крепости на ‘варианты жителей’, тем самым опуская важную информацию и нарушая лексическую сочетаемость:

1) יושבי לזהות באשר סברות שתי העלו החוקרים. בגבעה נמצאו ששרידיה המצודה

(Подстрочник: *Исследователи поднимали два мнения идентификации жителей крепости, что руины были найдены на холме.*)

Исследователи приводят два возможных варианта жителей крепости, руины которой находятся на холме.

В результате русский читатель узнает о том, что в крепости могли обитать две разных категории жителей; в тексте на русском языке искажены авторская точка зрения и реальное положение вещей.

«Внутренние» ошибки адекватности искажают содержание текста (перевод ‘варианты жителей’ нарушает внутритекстовую адекватность). Ошибки внутритекстовой адекватности перекрывают ошибки точности перевода. Следует отметить, что ошибки внешней адекватности текста соответствуют ошибкам приемлемости, выделенным в работе J. Daems [14].

Классификация в [8] была разработана для переводчиков. Авторы предлагают четыре различные категории ошибок: 1) ошибки в содержании и семантике, 2) дискурсивные ошибки, 3) ошибки в оценке (стратегии перевода) и 4) ошибки узуса (словоупотребления и формообразования). Первая категория соответствует ошибкам точности, включенным в класс ошибок внутритекстовой адекватности, а также некоторым из ошибок «внешней» адекватности. К категории ошибок внутренней адекватности текста относятся искажения жанровых признаков. Ошибки в стратегии перевода задевают внешнюю и разрушают внутритекстовую адекватность. ТП может оказаться неприемлемым для принимающей культуры. ТП может оказаться не приемлемым для принимающей культуры. Ошибочная стратегия перевода проявляется в непоследовательном выборе способа передачи названий в различных разделах ТП, что порой встречается в САТ, выполненных командой переводчиков на облачной платформе [6]:

1) הכחולה המרינה

Голубая марина / Синяя марина / Голубая пристань для яхт

2) דגים ידי

Иди фиш / Иди дагим (название популярного рыбного ресторана)

Ошибки узуса охватывают грамматику и неправильный лексический выбор, соотносясь с ошибками связности и точности соответственно. Поскольку грамматические ошибки влияют на стиль (например, синтаксическая конструкция из разговорной речи неуместна в официальном стиле), некоторые из грамматических ошибок можно отнести к разряду дискурсивных.

Как видим, ошибки точности, проявляющиеся в семантических искажениях, входят в каждую из четырех категорий. Ошибки точности включают

в себя опущение и добавление информации, семантические сдвиги из-за ошибок в терминологии, неправильный выбор именованя существей и т. п. Логические и фактологические ошибки внешней адекватности текста отличаются от ошибок точности и рассматривают как ошибки в передаче содержания. Главным достоинством этой классификации мы считаем выделение дискурсивных ошибок, определить которые можно на уровне целого документа, а не его отдельного сегмента.

Подробные классификации ошибок на материале переводов на русский язык были созданы для аннотации корпусов русского языка, используемых в обучении профессиональному переводу в университетах (Russian Learner Translator Corpus (Корпус несовершенных переводов) [15; 16] и в системах машинного перевода (Corpus for Russian data-to-text generation (Корпус для генерирования текстов на русском языке в системе «от данных – в тексту») [17]). Классификации содержат категории лингвистических, дискурсивных и семантических (неверные логические связи и пресуппозиции) ошибок. Таким образом, в классификациях различают ошибки связности, точности и адекватности текста перевода. Категория лингвистических ошибок включает в себя классы ошибок в грамматике, лексическом выборе, стилистике, орфографии и пунктуации. Пропуски и дополнения относятся к категории дискурса [17]. Классы и подклассы показывают результаты нисходящего просмотра (мониторинга) текста перевода.

Диаграмма (рис. 1) представляет собой комбинацию академических взглядов на классификацию ошибок перевода без учета расхождений в позициях исследователей.



Рис. 1. Категории и классы ошибок в академических классификациях ошибок перевода

При распознавании ошибок необходимо сочетать нисходящие и восходящие стратегии мониторинга текста перевода; ошибки оцениваются в системе дискурса, что позволяет распознавать семантические и логические связи между сегментами на уровне документа и значимость каждого сегмента в тексте перевода [18]. Лингвистические ошибки и ошибки в семантике можно распознать при посегментном сравнении ТП с ИТ (то есть снизу вверх, с использованием восходящей стратегии). Семантические сдвиги, пропуски и дополнения информации можно заметить благодаря применению стратегий нисходящего и нелинейного мониторинга ТП. Семантические сдвиги, пропуски и дополнения, а также смешение стилей могут представлять собой дискурсивные ошибки, если эти ошибки приводят к нарушению норм дискурса принимающей культуры в ТП.

Таким образом, в научных кругах исследователи выделяют категории ошибок внутритекстовой и внешней адекватности. Внутритекстовые ошибки снижают беглость и точность ТП; подклассы ошибок связности варьируют в зависимости от типа языка перевода (аналитического или флективного). Внешнетекстовые ошибки искажают приемлемость и функционирование ТП в принимающей культуре.

Категории ошибок перевода в классификациях прикладного характера

В переводческой индустрии исследователи рассматривают возможность создания систем автоматического распознавания ошибок, применимых для оценки машинного перевода в процессе его выполнения. Один из подходов к разработке подобных систем основан на переборе набора рабочих гипотез перевода предложения и выборе оптимальной версии. Для распознавания ошибок необходимо задать системе четкую непротиворечивую классификацию возможных ошибок.

Классификация ошибок в инструкции по оценке качества перевода по различным параметрам [19] включает две категории: ошибки точности и связности. К категории ошибок точности относят неверный лексический выбор, терминологические ошибки, пропуски, дополнения и непереведенные отрезки [20]. Различие между неверным лексическим выбором и ошибками в терминологии отражает важность терминологии для переводческой практики, поскольку переводчик часто имеет дело с различными документами, принадлежащими узкой профессиональной области, где широко используют отредактированный МП. В руководстве для постредакторов МП [21] авторы рекомендуют изучить несоответствия в терминологии и предложить экспертам в соответствующей профессиональной области варианты для их устранения. МП нередко оставляет сегменты ИТ непереведенными на ЯП, так как

алгоритм МП может пропускать последние позиции в длинных цепочках синтаксических зависимостей [6]. Классификацию ошибок связности разработчики автоматических систем корректируют с учетом лингвистических категорий ЯП. Благодаря ориентации на морфологию ЯП для морфологически богатых языков выделяют дополнительные подклассы в категории ошибок связности (например, ошибки в роде, лице, числе, падеже для славянских языков) [19]. Это позволяет обучать автоматические системы распознаванию грамматических ошибок, которые обычно встречаются во флективных языках. Хотя дополнительные подклассы ошибок не охватывают полностью контраст между грамматическими системами контактирующих языков [1], обучение на подклассах грамматических ошибок заметно улучшает способность алгоритмов и нейронных сетей обеспечивать связность МП. Клубичка и др. показали, как корректировка классификации ошибок в зависимости от морфологической типологии ЯП повышает качество МП и производительность автоматических систем распознавания ошибок [19].

При сопоставлении обсуждаемых классификаций ошибок с теми, что разработаны с позиций теории перевода в академических исследованиях, обнаруживается соответствие ошибок связности лингвистическим ошибкам, ошибок точности – лексическим и семантическим ошибкам. Категория ошибок связности включает грамматические, стилистические, орфографические и пунктуационные ошибки. Класс грамматических ошибок является наиболее типичным: более 80 % ошибок МП нарушают нормы грамматики ЯП [20]. Морфологические ошибки регулярно встречаются в МП в языках с богатой морфологией [20]. В переводах, выполненных профессиональными переводчиками, грамматические ошибки также оказываются наиболее типичными для категории ошибок связности [22].

Обсуждаемые классификации не включают категорию дискурсивных ошибок. Автоматические системы распознавания ошибок не способны учитывать особенности дискурса. Благодаря грамматике текста и правилам дискурса возможно разрешение анафоры, снятие многозначности, устранение омонимии [23]. К сожалению, современный уровень формального описания правил и особенностей дискурса не позволяет обучать автоматические системы распознавания ошибок находить дискурсивные ошибки. Соответственно, данная категория отсутствует в прикладных классификациях.

Кроме того, прикладные классификации не учитывают фактические и логические ошибки, поскольку системы МП обрабатывают естественный язык, но не проверяют корреляцию между содержанием текста и общей картиной мира. Ошибки внешней адекватности не находят отражения в классификациях.

Выявленные различия в «теоретических» и прикладных классификациях ошибок перевода отражают готовность прикладных специалистов обеспе-

чивать возможность проверять МП на связность и точность, оставляя профессиональным постредакторам и переводчикам выявление дискурсивных и функциональных ошибок – тех ошибок, что видны на уровне целого документа, рассматриваемого не линейно, а сверху вниз, от общей мысли текста к ее развитию в отдельных его частях. Прикладные классификации оказываются более прозрачными и универсальными в силу необходимости соответствовать стандартам систем МП и технологий разработки автоматических систем распознавания ошибок и оценки качества перевода.

Автоматическое распознавание ошибок и оценка качества перевода

Для распознавания и оценки языковых отклонений в МП разрабатываются специальные системы и инструменты. Параметры, принятые в системах оценки качества, позволяют перейти от субъективных пристрастий к объективным суждением, однако только человек способен определить, является ли обнаруженное отступление от нормы ошибкой.

Автоматические системы в основном обучены распознавать, оценивать и исправлять ошибки связности в текстах на английском языке, что имеет важное значение для глобальной коммуникации. С. Брайант и др. обсуждают достоинства автоматических систем исправления грамматических ошибок, обученных на новом расширенном корпусе английского языка [24]. Однако эффективность автоматического распознавания и исправления ошибок при переводе текста на морфологически богатый язык не достигает того уровня надежности, который получен для английского языка [25].

Одно из возможных решений проблемы автоматического распознавания ошибок связности и точности основано на определении контрольного набора языковых форм и категорий («контрольных точек»), различающих язык исходного текста и ЯП; контрольный набор можно сформировать на основе параллельного двуязычного корпуса [1]. Такой прикладной подход в определенной степени соответствует академической теории критических точек для принятия переводческого решения профессиональным переводчиком [26]. Теория критических точек учитывает потенциальную межъязыковую интерференцию и вмешательство переводчика, а также контрастивное описание контактирующих языков. База данных «контрольных точек» включает в себя список языковых единиц исходного языка и их возможные переводы на целевой язык. Чтобы создать такую базу данных, разработчику необходимо определить лингвистические категории, наиболее точно описывающие контраст между ЯО и ЯП. Лингвистические категории характеризуют различия, охватывающие лексику, фразеологию, морфологию и синтаксис ЯП. Описанные различия указывают на потенциальные ошибки при МП. Од-

нако лингвистическое обоснование «контрольных точек» все еще далеко от того, чтобы обеспечить релевантные результаты в автоматическом распознавании ошибок и оценке качества перевода МП.

Распределение категорий и классов ошибок в аспекте их распознавания автоматическими системами показано на рис. 2.



Рис. 2. Классы ошибок МП для автоматического распознавания ошибок и оценки качества перевода

Основываясь на сопоставлении классификаций, можно выделить: 1) ошибки внешней адекватности текста, отражающие несоответствие общей картине мира и законам логики; 2) ошибки внутритекстовой адекватности, искажающие содержание исходного текста в ТП, среди которых выделяются ошибки в терминологии, лексическом выборе; 3) дискурсивные ошибки, вызванные нарушением правил и норм дискурса принимающей культуры; 4) ошибки связности ТП, разрушающие стилевое разнообразие, грамматическую правильность, нормативное оформление ТП (пунктуацию, орфографию, форматирование). Автоматические системы не распознают ошибки внешней адекватности ТП и его приемлемости для принимающей культуры. Профессиональный переводчик, работающий на платформе CAT, и постредактор МП отвечают за анализ и устранение таких ошибок.

Заключение

Следует признать, что технологии облегчают перевод. Платформы CAT и системы МП позволяют быстро переводить большие объемы документов, обнаруживать некоторые ошибки, а также направлять конечный продукт за-

казчику. Переводчики концентрируют свои усилия на передаче содержания исходного текста в принимающую культуру. Тем не менее новая технологическая среда ставит новые проблемы и раскрывает важность целостного взгляда на особенности текста, предлагаемого для перевода. Различные типы ошибок перевода распознаются благодаря сочетанию различных стратегий. Чтобы обнаружить дискурсивные ошибки, переводчик должен уметь переключать стратегии, анализируя нелинейно МП или перевод, выполненный на платформе САТ. Переводчик и постредактор несут ответственность за исправление дискурсивных, логических и фактических ошибок, поскольку прикладные разработки пока не могут обеспечить современную технологию распознавания и оценки дискурсивных ошибок для морфологически богатых языков, а также сопоставление содержания ТП с общей картиной мира. Системы автоматического распознавания ошибок не анализируют соотношение языкового выражения с референтной ситуацией, в то время как профессиональный переводчик в состоянии оценить соответствие содержания ТП фоновым знаниям. Тем не менее автоматические системы умеют распознавать основные классы ошибок в категориях связности и точности для многих языковых пар. Прикладные классификации ошибок не включают дискурсивные ошибки и ошибки внешней адекватности ТП, поскольку разработчики систем МП и платформ САТ только пытаются создать надежный инструмент для предотвращения этих ошибок.

Разнообразные стратегии переработки текстов при переводе и редактировании может использовать профессиональный переводчик, который по-прежнему является ключевой фигурой в переводческой индустрии.

Список литературы

1. Popović M. Error classification and analysis for machine translation quality assessment / J. Moorkens, S. Castilho, F. Gaspari, S. Doherty (Eds), *Translation Quality Assessment of Machine Translation // Technologies and Applications: Springer, Cham.* – 2018. – Vol. 1. – P. 129–158.
2. European Master's in Translation Competence Framework [Электронный ресурс] // European Commission. – 2017. – URL: https://ec.europa.eu/info/sites/info/files/emt_competence_fw_k_2017_en_web.pdf (дата обращения: 10.02.2021).
3. Blagodarna O. Insights into post-editors' profiles and post-editing practices // *Revista Tradumàtica: Traducció i tecnologies de la informació i la comunicació.* – 2018. – No. 16. – P. 35–51.
4. Cagiltay K., Bichelmeyer B., Kaplan A.G. Working with multicultural virtual teams: critical factors for facilitation, satisfaction and success // *Smart Learning Environment.* – 2015. – No. 2. – P. 1–16.
5. Sakamoto A., Foedisch M. No news is good news // *Translation Spaces.* – 2017. – No. 6(2). – P. 333–352.

6. Овчинникова И.Г. Использование компьютерных переводческих инструментов: новые возможности, новые ошибки // Вестник Рос. ун-та дружбы народов. Сер. Лингвистика. – 2019. – № 23(2). – С. 544–561.

7. Behind human error: Cognitive systems, computers and hindsight [Электронный ресурс] / D.D. Woods, L.J. Johannesen, R.I. Cook, N.B. Sarter // A.W. Schopper (Ed.). State-of-the art report (SOAR) N 94-01, Crew System Ergonomic Information Analysis Center, Ohio, Wright-Patterson Air Force Base, 1994. – URL: <https://apps.dtic.mil/dtic/tr/fulltext/u2/a492127.pdf> (дата обращения: 10.02.2021).

8. Новый взгляд на классификацию переводческих ошибок / Д.М. Бузаджи, В.В. Гусев, В.К. Ланчиков, Д.В. Псурцев. – М.: Изд-во Всерос. центр переводов, 2009.

9. Овчинникова И.Г., Павлова А.В. Переводческий билингвизм: по материалам ошибок письменного перевода. – М.: Флинта-Наука, 2016.

10. Павлова А.В. Типология переводческих ошибок билингвов (на материале ошибок студентов русского отделения переводческого факультета в Гермерсхайме) // Иностранные языки в высшей школе. – Рязань, 2013. – Вып. 1 (24). – С. 80–89.

11. Reinke U. State of the art in translation memory technology // G. Rehm, F. Sasaki, D. Stein, A. Witt (Eds.). Language technologies for a multilingual Europe: TC3 III (Translation and Multilingual Natural Language Processing 5). – Berlin: Language Science Press, 2018. – P. 55–85.

12. Nord C. Translating as a purposeful activity: Functionalist approaches explained. – London: Routledge, 2018.

13. Thelen M. Translation quality assessment or quality management and quality control of translation // B. Lewandowska-Tomaszczyk, M. Thelen (Eds.). Translation and meaning. Part 8: International Maastricht-Lodz Duo Colloquium on “Translation and Meaning” (Lodz, Poland, Universitaire Pers Maastricht UPM, Maastricht, 23–25 September 2005). – P. 411–424.

14. Daems J. A translation robot for each translator? A comparative study of manual translation and post-editing of machine translations: process, quality and translator attitude [Электронный ресурс]: PhD thesis. – Ghent University, Belgium, 2016. – URL: <https://biblio.ugent.be/publication/8058017> (дата обращения: 10.02.2021).

15. Kunilovskaya M., Morgoun N., Pariy A. Learner vs. professional translations into Russian: lexical profiles // The International Journal for Translation and Interpreting Research. – 2018. – No. 10. – P. 33–52.

16. Куниловская М.А. Классификация переводческих ошибок для создания разметки в учебном параллельном корпусе Russian Learner Translator Corpus // *Lingua mobilis*. – Челябинск, 2013. – № 1 (40). – С. 141–159.

17. Shimorina A., Khasanova E., Gardent C. Creating a Corpus for Russian Data-to-Text Generation using neural machine translation and post-editing / T. Erjavec, M. Marcińczuk, P. Nakov, J. Piskorski, L. Pivovarova, J. Šnajder, J. Steinberger, R. Yangarber (Eds.) // Proceedings of the 7th Workshop on Balto-Slavic Natural Language Processing. – August, 2019. – Florence, Italy. – P. 44–49.

18. Vandepitte Daems S., Hartsuiker R.J., Macken L. Identifying the machine translation error types with the greatest impact on post-editing effort // *Frontiers in psychology*. – 2017. – No. 8. – P. 1282.

19. Klubička A. Toral, Sánchez-Cartagena V.M. Fine-grained human evaluation of neural versus phrase-based machine translation // *The Prague Bulletin of Mathematical Linguistics*. – 2017. – No. 108 (1). – P. 121–132.

20. Tools and guidelines for principled machine translation development [Электронный ресурс] / N. Aranberri, E. Avramidis, A. Burchardt, O. Klejch, M. Popel, M. Popovic // N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, S. Piperidis (Eds.) // *Proceedings of the 10th International Conference on LREC*. (Portorož, Slovenia, 2016). – P. 1877–1882. – URL: <https://www.aclweb.org/anthology/volumes/L16-1/> (дата обращения: 10.02.2021).

21. MT post-editing guidelines / I. Massardo, J. van der Meer, S. O'Brien, F. Hollo-wood, N. Aranberri, K. Drescher. – The Netherlands: TAUS Signature Editions, 2016.

22. The First Automatic Translation Memory Cleaning Shared Task / E. Barbu, C.P. Escartín, L. Bentivogli, M. Negri, M. Turchi, C. Orasan, M. Federico // *Machine Translation*. – 2016. – No. 30 (3–4). – P. 145–166.

23. Smith K.S. On integrating discourse in machine translation // B. Webber, A. Popesku-Belis, J. Tiedemann (Eds.). *Proceedings of the Third Workshop on Discourse in Machine Translation*. – Copenhagen, Denmark, 2017. – P. 110–121.

24. The BEA-2019 shared task on grammatical error correction [Электронный ресурс] / C. Bryant, M. Felice, Ø.E. Andersen, T. Briscoe // H. Yannakoudakis, E., Kochmar, C. Leacock, N. Madnani, I. Pilan, T. Zesch (Eds.). *Proceedings of the Fourteenth Workshop on Innovative Use of NLP for Building Educational Application*. – Florence, Italy, 2019. – P. 52–75. – URL: <https://www.aclweb.org/anthology/W19-44> (дата обращения: 10.02.2021).

25. Translation Quality and Productivity: A Study on Rich Morphology Languages [Электронный ресурс] / L. Specia, K. Harris, A. Burchardt, M. Turchi, M. Negri, I. Skadina // S. Kurohashi, P. Fung (Eds.). *Proceedings of Machine Translation Summit XVI*. Vol. 1: Research Track. – Nagoya, Aichi, Japan, 2017. – P. 55–71. – URL: http://aamt.info/app-def/S102/mtsummit/2017/wpcontent/uploads/sites/2/2017/09/MTSummitXVI_ResearchTrack.pdf (дата обращения: 10.02.2021).

26. Munday J. *Evaluation in translation: Critical points of translator decision-making*. – London: Routledge, 2012.

References

1. Popović M. Error classification and analysis for machine translation quality assessment. *Translation quality assessment of machine translation. Technologies and applications*. Eds. J. Moorkens, S. Castilho, F. Gaspari, S. Doherty. Springer, Cham, 2018, vol. 1, pp. 129–158.

2. European master's in translation competence framework. European Commission, 2017, available at: https://ec.europa.eu/info/sites/info/files/emt_competence_fwk_2017_en_web.pdf (accessed 10.02.2021).

3. Blagodarna O. Insights into post-editors' profiles and post-editing practices. *Revista Tradumàtica: Traducció i tecnologies de la informació i la comunicació*, 2018, no. 16, pp. 35–51.

4. Cagiltay K., Bichelmeyer B., Kaplan A.G. Working with multicultural virtual teams: critical factors for facilitation, satisfaction and success. *Smart Learning Environments*, 2015, no. 2, pp. 1–16.
5. Sakamoto A., Foedisch M. No news is good news. *Translation Spaces*, 2017, no. 6(2), pp. 333–352.
6. Ovchinnikova I.G. Ispol'zovanie komp'iuternykh perevodcheskikh instrumentov: novye vozmozhnosti, novye oshibki [Working on computer-assisted translation platforms: New advantages and new mistakes]. *Vestnik Rossiiskogo universiteta druzhby narodov. Seriya: Lingvistika*, 2019, no. 23(2), pp. 544–561.
7. Woods D.D., Johannesen L.J., Cook R.I., Sarter N.B. Behind human error: Cognitive systems, computers and hindsight. State-of-the art report (SOAR) Number 94-01. Ed. A.W. Schopper, Crew System Ergonomic Information Analysis Center, Ohio, Wright-Patterson Air Force Base, 1994, Available at: <https://apps.dtic.mil/dtic/tr/fulltext/u2/a492127.pdf> (accessed 10.02.2021).
8. Buzadzi D.M., Gusev V.V., Lanchikov V.K., Psurtsev D.V. Novyi vzgliad na klassifikatsiiu perevodcheskikh oshibok [New view on the classification of translation errors]. Moscow, All-Russian translation center, 2009.
9. Ovchinnikova I.G., Pavlova A.V. Perevodcheskij bilingvizm: Po materialam oshibok pis'mennogo perevoda [Translation bilingualism: based on errors in text translation]. Moscow, Flinta-Nauka, 2016.
10. Pavlova A.V. Tipologiya perevodcheskikh oshibok bilingvov (na materiale oshibok studentov russkogo otdeleniia perevodcheskogo fakul'teta v Germerskhaime) [Types of translation errors made by bilingual students (exemplified by mistakes made by students of Russian at the translation department of university of Mainz)]. *Inostrannye iazyki v vysshei shkole*, 2013, no. 1(24), pp. 80–89.
11. Reinke U. State of the art in translation memory technology. In: Language technologies for a multilingual Europe: TC3 III (Translation and multilingual natural language processing 5). Eds. G. Rehm, F. Sasaki, D. Stein, A. Witt. Berlin, Language Science Press, 2018, pp. 55–85.
12. Nord C. Translating as a purposeful activity: Functionalist approaches explained. London, Routledge, 2018.
13. Thelen M. Translation quality assessment or quality management and quality control of translation. Eds. B. Lewandowska-Tomaszczyk, M. Thelen. *Proc. of the 4th International Maastricht-Lodz Duo Colloquium on Translation and Meaning. Part 8* (Lodz, Sep. 23-25 September, 2005). Lodz, Poland, Universitaire Pers Maastricht UPM, pp. 411–424.
14. Daems J. A translation robot for each translator? A comparative study of manual translation and post-editing of machine translations: process, quality and translator attitude. Ph. D. thesis. Ghent University, Belgium, 2016, available at: <https://biblio.ugent.be/publication/8058017> (accessed 10.02.2021).
15. Kunilovskaya M., Morgoun N., Pariy A. Learner vs. professional translations into Russian: Lexical profiles. *The International Journal for Translation and Interpreting Research*, 2018, no. 10, pp. 33–52.

16. Kunilovskaia M.A. Klassifikatsiia perevodcheskikh oshibok dlia sozdaniia razmetki v uchebnom parallel'nom korpuse Russian learner translator corpus [Classification of translation errors for marking of education in parallel body Russian learner translator corpus]. *Lingua mobilis*, 2013, no. 1 (40), pp. 141–159.

17. Shimorina A., Khasanova E., Gardent C. Creating a corpus for Russian data-to-text generation using neural machine translation and post-editing. Eds. T. Erjavec, M. Marcińczuk, P. Nakov, J. Piskorski, L. Pivovarova, J. Šnajder, J. Steinberger, R. Yangarber. *Balto-Slavic Natural Language Processing*, Proc. 7th Workshop (Florence, Aug. 2019). Florence, Italy, pp. 44–49.

18. Daems J., Vandepitte S., Hartsuiker R.J., Macken L. Identifying the machine translation error types with the greatest impact on post-editing effort. *Frontiers in Psychology*, 2017, no. 8, pp. 1282.

19. Klubička F., Toral A., Sánchez-Cartagena A.M. Fine-grained human evaluation of neural versus phrase-based machine translation. *The Prague Bulletin of Mathematical Linguistics*, 2017, no. 108(1), pp. 121–132.

20. Aranberri N., Avramidis E., Burchardt A., Klejch O., Popel M., Popovic M. Tools and guidelines for principled machine translation development. Eds. N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik et al. *LREC*, Proc. 10th Int. Conf. (Portorož, 2016), Portorož, Slovenia, pp. 1877–1882, available at: <https://www.aclweb.org/anthology/volumes/L16-1/> (accessed 10.02.2021).

21. Massardo I., Meer J. van der, O'Brien S., Hollowood F., Aranberri N., Drescher K. MT post-editing guidelines. The Netherlands, TAUS Signature Editions, 2016.

22. Barbu E., Escartín C.P., Bentivogli L., Negri M., Turchi M., Orasan C., Federico M. The first automatic translation memory cleaning shared task. *Machine Translation*, 2016, no. 30 (3-4), pp. 145–166.

23. Smith K.S. On integrating discourse in machine translation. Eds. B. Webber, A. Popescu-Belis, J. Tiedemann. *Discourse in Machine Translation*, Proc. 3rd Workshop (Copenhagen, 2017). Copenhagen, Denmark, 2017, pp. 110–121.

24. Bryant C., Felice M., Andersen Ø.E., Briscoe T. The BEA-2019 shared task on grammatical error correction. Eds. H. Yannakoudakis, E. Kochmar, C. Leacock, N. Madnani, I. Pilan, T. Zesch. *Innovative Use of NLP for Building Educational Application*, Proc. 14th Workshop (Florence, 2019). Florence, Italy, pp. 52–75, available at: <https://www.aclweb.org/anthology/W19-44> (accessed 10.02.2021).

25. Specia L., Harris K., Burchardt A., Turchi M., Negri M., Skadina I. Translation quality and productivity: A study on rich morphology languages. Eds. S. Kurohashi, P. Fung. Proc. of *Machine Translation Summit XVI*, vol. 1. Nagoya, Japan, 2017, pp. 55–71, available at: http://aamt.info/app-def/S-102/mtsummit/2017/wpcontent/uploads/sites/2/2017/09/MTSummitXVI_ResearchTrack.pdf (accessed 10.02.2021).

26. Munday J. Evaluation in translation: Critical points of translator decision-making. London, Routledge, 2012.

Сведения об авторах

БАРИНОВА Ирина Александровна

e-mail: *barinova.i.a@yandex.ru*

Доцент кафедры иностранных языков, лингвистики и перевода, Пермский национальный исследовательский политехнический университет (Пермь, Российская Федерация)

ОВЧИННИКОВА Ирина Германовна

e-mail: *ira.ovchi@gmail.com*

Доктор филологических наук, профессор Института лингвистики и межкультурной коммуникации, Первый Московский государственный медицинский университет им. И.М. Сеченова (Москва, Российская Федерация)

About the authors

Irina A. BARINOVA

e-mail: *barinova.i.a@yandex.ru*

Associate professor, Department of Foreign Languages, Linguistics and Translation, Perm National Research Polytechnic University (Perm, Russian Federation)

Irina G. OVCHINNIKOVA

e-mail: *ira.ovchi@gmail.com*

Doctor of Philology, Professor, Institute of Linguistics and Intercultural Communication, I.M. Sechenov First Moscow State Medical University (Moscow, Russian Federation)