



DOI: <https://doi.org/10.15688/jvolsu2.2021.3.8>

UDC 81'322.4
LBC 81.184

Submitted: 11.12.2020
Accepted: 30.03.2021

EFFICIENCY OF MACHINE TRANSLATION IN URBAN DISCOURSE

Svetlana A. Korolkova

Volgograd State University, Volgograd, Russia

Anna A. Novozhilova

Volgograd State University, Volgograd, Russia

Abstract. This article aims to analyze the use of *Yandex.Translate*, an online machine translation system, in translating urban discourse texts on the web. The authors use integrative linguistic-and-pragmatic approach to assess machine translation quality in a global digital setting. The aim is to show the efficiency of a state-of-the-art machine translation system and to investigate its usefulness in practical application. The authors perform a detailed analysis of the Paris city website content, which is automatically translated from French into Russian with *Yandex.Translate*. The data selection is justified by the absence of official foreign versions of this website, which points to the need of machine translation engines integrated in a web browser. Less than 20% of the analysed machine-translated texts demonstrate high language quality, whereas 60% can be referred to as acceptable – the text preserves the meaning of the source but contains some errors and inaccuracies in the target language. About 20% of the machine-translated text contains blunders, which violate Russian language norms. It causes source text contents distortion and communication failures. In the end, a classification of the system errors is presented. It is also concluded that machine translation would substitute middle-skilled human translators in the future. However, the use of such systems will enforce standardisation and simplification of the target language.

Key words: information technology, machine translation, *Yandex.Translate*, translation quality, language norm, semantic error, urban discourse.

Citation. Korolkova S.A., Novozhilova A.A. Efficiency of Machine Translation in Urban Discourse. *Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 2. Yazykoznanie* [Science Journal of Volgograd State University. Linguistics], 2021, vol. 20, no. 3, pp. 87-98. (in Russian). DOI: <https://doi.org/10.15688/jvolsu2.2021.3.8>

УДК 81'322.4
ББК 81.184

Дата поступления статьи: 11.12.2020
Дата принятия статьи: 30.03.2021

ЭФФЕКТИВНОСТЬ СИСТЕМ МАШИННОГО ПЕРЕВОДА В УРБАНИСТИЧЕСКОМ ДИСКУРСЕ

Светлана Азадовна Королькова

Волгоградский государственный университет, г. Волгоград, Россия

Анна Алексеевна Новожилова

Волгоградский государственный университет, г. Волгоград, Россия

Аннотация. Настоящая статья посвящена анализу системы машинного перевода «Яндекс.Переводчик» и ее использованию в сети Интернет для перевода текстов урбанистического дискурса. Для оценки качества работы системы в условиях гиперглобализации цифрового пространства используется интегративный лингвопрагматический подход. Цель работы – показать уровень развития современных систем машинного перевода и выявить целесообразность их использования для потребительских целей. Представлены результаты анализа перевода контента официального городского интернет-портала г. Парижа, выполненного программой «Яндекс.Переводчик» с французского языка на русский. Выбор материала исследования обус-

ловлен отсутствием переводных версий данного портала на других языках, что подтверждает факт востребованности встроенных в браузер систем машинного перевода. Выявлено, что менее 20 % текста, переведенного машинным переводчиком, показывают высокое качество продукта, около 60 % материала переведено на уровне «приемлемого качества», при котором допускаются ошибки и неточности, не искажающие смысла исходного текста, такой речевой продукт не представляет собой полностью корректного текста на языке перевода; около 20 % переведенного контента содержат семантические ошибки и грубые нарушения языковой нормы русского языка, которые приводят к искажению содержания исходного текста и сбою в коммуникации. Представлена классификация ошибок, допускаемых системой, и сделаны прогностические выводы о том, что использование систем машинного перевода способствует стандартизации и упрощению языка и вытеснению с рынка переводчиков среднего уровня квалификации.

Ключевые слова: информационные технологии, машинный перевод, «Яндекс.Переводчик», качество перевода, языковая норма, семантические ошибки, урбанистический дискурс.

Цитирование. Королькова С. А., Новожилова А. А. Эффективность систем машинного перевода в урбанистическом дискурсе // Вестник Волгоградского государственного университета. Серия 2, Языкознание. – 2021. – Т. 20, № 3. – С. 87–98. – DOI: <https://doi.org/10.15688/jvolsu2.2021.3.8>

Введение

Информационные технологии уже давно и прочно вошли практически во все сферы науки и техники, и современный мир уже невозможно представить без них. В последнее десятилетие наблюдается бурное развитие искусственного интеллекта и в этой связи все чаще звучат прогнозы о том, что роботы и «умные системы» вытеснят ряд профессий. Среди таких профессий иногда называют профессию переводчика, заявляя, что машинный перевод (МП) сможет заменить ее в ближайшем будущем. Данное утверждение может показаться на первый взгляд весьма обоснованным, поскольку в эпоху глобализации и распространения сети Интернет объемы информации на разных языках стремительно увеличиваются и потребность в их переводе постоянно растет, а системы машинного перевода совершенствуются и справляются с переводом огромного контента буквально за несколько секунд.

Лингвисты и ученые-переводоведы участвуют в развитии машинного перевода, анализируют и оценивают качество текстов, полученных при его помощи. Такие исследования способствуют выявлению языковых ошибок и проблем в работе машинных переводчиков и имеют практическую значимость для создателей лингвистического программного обеспечения. В настоящей статье исследуются возможности выполнять качественный перевод одной из наиболее используемых на сегодняшний день российских систем машинного перевода – «Яндекс.Переводчик».

Материалы и методы

Материалом исследования послужил контент официального городского интернет-портала Парижа (www.paris.fr) и его перевод на русский язык, выполненный машинным переводчиком «Яндекс.Переводчик» в автоматическом режиме (около 16 000 печатных знаков). Данная функция встроена в Яндекс.Браузер и предлагается системой при переходе пользователей на сайты на любом иностранном языке, поддерживаемом «Яндекс.Переводчиком».

Городской интернет-портал выбран в качестве материала исследования не случайно. Такие ресурсы становятся все более востребованными среди большого количества пользователей Интернета. Они привлекают внимание разных целевых групп: потенциальных туристов, предпринимателей и бизнесменов, инвесторов, специалистов в области истории, культуры, рекламы и т. д. Одной из основных задач городских интернет-порталов является реализация PR-функции, которая направлена на повышение аттрактивности города и формирование его позитивного бренда и имиджа [Гуреева, 2017, с. 129–130]. Городские интернет-порталы, по мнению Э.Ю. Новиковой, – это социокультурное пространство, в котором сочетаются признаки и элементы различных институциональных дискурсов: медийного, виртуального, административного, туристического, маркетингового, культурного и др. [Новикова, 2015, с. 468].

На наш взгляд, большинство текстов городских порталов можно отнести к так на-

зываемому урбанистическому дискурсу, который, как отмечают некоторые исследователи, сформировался в результате «пересечения большого количества дискурсов, от постфордистского дискурса до дискурса глобализации и оформления мирового города, общественного дискурса, административного дискурса и т. д.» [Зарубина, 2018, с. 178]. Урбанистический дискурс все чаще становится предметом исследования отечественных и зарубежных лингвистов, поскольку привлекает своим разнообразием жанров, стилей и яркой палитрой средств вербальной и невербальной коммуникации [Chen Fachun, Leontovich, 2020; Aiello, Tosoni, 2016; и др.]. В Интернете именно урбанистический дискурс становится «ключевым каналом трансляции территориального бренда, формирующим образ и ассоциации из набора символов, пресуппозиций, знаний, которые связаны с территорией как в глобальном, так и локальном сетевых контентх [Попова, Королькова, Степанова, 2020, с. 124].

Анализируя машинный перевод такого разнопланового контента, можно проследить, как далеко продвинулся искусственный интеллект в сфере перевода и оценить, насколько эффективна его работа с текстами разных типов и какое потенциальное влияние может оказать машинный перевод на развитие переводческой отрасли и самого языка.

Исследование проводилось с применением следующих лингвистических методов исследования: описательный, сопоставительный методы, структурно-семантический, стилистический анализ, прием количественных подсчетов. Данные методы и приемы позволили выделить ряд преимуществ и недостатков использования программ машинного перевода при осуществлении межкультурной коммуникации в рамках урбанистического дискурса.

Результаты и обсуждения

До 20-х гг. XXI в. история развития и применения МП имела поступательный характер. Не углубляясь в историю становления и развития МП, можно смело утверждать, что наше время – бум МП как в сфере его развития, так и его применения и распространения.

М. Меняйлова, руководитель направления машинного перевода в компании ЭГО Транслейтинг, констатирует, что «начиная с 2000-х развитие МП ускоряется, с 2019 года мы можем наблюдать постоянные изменения в работе с МП каждые полгода, что вскоре может перерасти в глобальные обновления раз в месяц» [Меняйлова, 2020].

С одной стороны, появление коронавируса в 2020 г. поставило под сомнение сам феномен глобализации из-за мирового карантина и закрытия границ, с другой стороны, он же привел к гиперглобализации цифрового пространства. Человечество было вынуждено уйти на удаленную работу, и коммуникация в цифровом пространстве многократно увеличилась. Возникла насущная потребность в быстром переводе значительных объемов текста, что вызвано даже не только интернационализацией бизнеса, сколько гиперускорением технологического процесса и его повсеместным проникновением, в том числе и на бытовом уровне. Без систем машинного перевода мы не сможем общаться с «умными предметами» с голосовым управлением, которые разрабатываются в одной стране и продаются по всему миру. Очевидно, что интернационализация бизнеса, обучения, культурного и социального взаимодействия будет только усиливаться, и технологический прогресс в сфере искусственного интеллекта и МП будет только ускоряться. Публикация издательством Individuum в конце 2020 г. книги Дэниела Сасскинда «Будущее без работы. Технологии, автоматизация и стоит ли их бояться» [Susskind, 2020], перевод которой был выполнен «Яндекс.Переводчиком» за 40 секунд и отредактирован профессиональным переводчиком А. Дунаевым за 3 месяца (обратите внимание на разницу во времени!) [Тарарак, 2020], как никогда актуализирует вопрос, задаваемый профессиональным переводческим сообществом с 50-х гг. прошлого века: заметит ли МП переводчиков. При этом, как отмечает Д. Беллос в своей книге «Что за рыбка в вашем ухе?», нужно помнить: МП не появился ниоткуда, на ровном месте. Все варианты перевода, выражения и т. д. придумал не МП, это результат бесконечного и кропотливого труда «человеческих переводчиков»: именно их переводы и создали корпуса тек-

стов, именно они вместе с программистами разрабатывали и развивали МП [Беллос, 2019, с. 152].

У неспециалиста, коим является заказчик, обратившийся в переводческое бюро, или любой рядовой потребитель МП, термин «машинный перевод» вызывает две реакции: либо ужас, либо полный восторг и убеждение, что машинный перевод вполне может заменить человеческий. Для профессионала, работающего в сфере перевода, МП – это инструмент, который необходимо использовать, поскольку он позволяет получить выгоду во времени и, соответственно, сократить финансовые затраты, но при этом следует быть готовым получить на выходе перевод «приемлемого качества», согласно Международному переводческому стандарту 2017 г. [ISO 18587: 2017]. Стоит отметить, что этот стандарт стал результатом ускорения прогресса в сфере МП и понимания, что спрос клиентов на использование МП будет только расти. Как отмечается во введении Стандарта, МП позволяет сократить затраты на перевод, ускорить выход товаров на конкретные рынки, но главное – увеличить потоки информации, сохраняя конкурентоспособность специалиста-переводчика. При этом в стандарте подчеркивается, что «скорость, с которой меняются системы машинного перевода, делает нецелесообразным разработку всеобъемлющего международного стандарта для этих систем, который может сдерживать инновации или игнорироваться индустрией разработки технологий перевода» [ISO 18587: 2017].

С 2018 г. все системы МП, предлагаемые ведущими IT-игроками на рынке переводческих услуг, например, российские компании Яндекс и Промт, являются гибридными, то есть объединяют базовый SMT-модуль перевода и модуль NMT, которые выполняют перевод, а третий модуль, в системе Яндекса (CatBoost), осуществляет на основании ряда факторов – длина предложения (короткие фразы лучше переводит статистическая модель), терминология и т. д. вплоть до синтаксиса, оценивание вариантов и выбирает лучший, который и предьявляется пользователю. Кроме этого, в коммерческих версиях есть возможность подключать уже на-

работанную Память переводов клиента для обучения системы, как показано в исследовании Р. Лука, что обеспечивает более эффективный перевод, поскольку качество инструментов МП напрямую зависит от качества корпуса текстов, с которым он работает [Loock, 2016]. По мнению М. Меняйловой, «современное развитие кастомизации движков машинного перевода под организацию, проект, тематику и любые задачи значительно расширяет границы применения МП и возможно даже стирает их вовсе» [Меняйлова, 2020]. Современные системы МП – это инструменты для облегчения труда переводчика, поскольку их можно использовать, с одной стороны, для выполнения рутинной работы переводчика, с другой стороны, для сокращения затрат (временных и финансовых) заказчика.

Использование МП на профессиональном переводческом рынке уже имеет свои четкие ориентиры. Профессионалы знают, для каких типов текстов и в какой предметной сфере его использование даст максимальный результат и выгоду. Они осознают пределы его использования и понимают, в каких случаях его использование оправдано. Для профессионалов очевидно, что МП не устраняет человека-переводчика, а создает новые сферы деятельности, например, постредактирование МП (МТРЕ – machine translation post edition). Следовательно, важно сфокусировать исследование на использовании МП в так называемых «бытовых» ситуациях, когда потребитель «неспециалист», не владеющий иностранным языком, прибегает к использованию МП для решения сиюминутных потребностей: ознакомиться со статьей с иноязычного ресурса, с иноязычным сайтом во время путешествий и т. д. В таких ситуациях потребитель использует общедоступные ресурсы – «Яндекс.Переводчик», «Google.Translate» или аналогичные им ресурсы в социальных сетях, например, в Фейсбуке. Французское подразделение Фейсбука, разрабатывая собственную систему МП (уже сейчас в этой сети осуществляется более 4 миллиардов переводов в день в режиме «timeline»), делает ставку на «потребительский» МП для редких языков (пушту, суахили, кхмерский или непальский), на которых практически не существует корпусов текстов, при

этом разработчики пытаются создать практически универсальные модели, применимые к любой языковой паре [Lupieri, 2019].

Обратимся к контенту официального городского интернет-портала столицы Франции. Такие сайты, как уже отмечалось выше, изобилуют текстами разных жанров и стилей, и на примере работы системы «Яндекс.Переводчик» дадим оценку эффективности ее использования при переводе этих ресурсов для «потребительских» целей.

Официальный портал Парижа www.paris.fr (site officiel de la Ville de Paris) имеет отличительную черту: он – одноязычный, то есть любой пользователь Интернета, пытающийся найти здесь нужную для себя информацию, сталкивается с языковым барьером и необходимостью воспользоваться услугами МП. При этом внизу презентационной страницы находится ссылка «Visit Paris. Tourist Office» на многоязычную (11 языков – все основные европейские, а также русский, китайский, японский и корейский языки) туристическую страницу французской столицы www.parisinfo.com. Иными словами, не владеющий французским языком пользователь Интернета может ознакомиться только с туристической информацией: достопримечательности и краткая историческая справка, мероприятия, музеи, рестораны, транспорт и т. д., но получить практические сведения и информацию о повседневной жизни парижан, воспользоваться полезными для повседневной жизни ссылками и т. д. с официальной страницы города без МП он не сможет.

Итак, проанализируем контент и МП главной страницы городского портала Парижа, вкладки которой отправляют к текстам разного содержания, жанров и адресованы разнообразной целевой аудитории. Перевод страницы «Яндекс.Переводчиком» выполнен без пропусков: все рубрики и все микротексты переведены. Анализ переведенной страницы позволяет сделать вывод, что 15–20 % контента переведено на хорошем уровне и соответствует запросу требовательного в языковом плане получателя. Это в первую очередь названия рубрик в форме простых предложений, не содержащих культурно маркированных и многозначных единиц, понимание которых вне контекста затруднено.

(1) *Visite virtuelle: la Bibliothèque historique à 360°.*

Виртуальный *тур*: Историческая библиотека на 360.

(2) *Cours d'adultes de Paris.*

Курсы для взрослых в Париже.

(3) *Les restaurateurs peuvent réserver des stands sur les marchés.*

Рестораторы могут забронировать *киоски* на рынках.

При этом следует отметить, что «Яндекс.Переводчик» находит для выделенных единиц наиболее адекватный контексту вариант перевода, который не зафиксирован в словарях, то есть движок учитывает смысловые связи внутри предложения, а качество параллельных текстов статистической модели движка высокое.

Перевод 50–60 % контента можно характеризовать как текст «приемлемого качества» («good enough») в терминологии центра TAUS (Translation Automation User Society) и Международного стандарта 2017. При таком качестве текст перевода точно передает смысл исходного сообщения, но не является грамматически, синтаксически и стилистически корректным с точки зрения языковой нормы [TAUS, 2010]. Иными словами, перевод является правильным с точки зрения семантики, в нем нет случайных добавлений или опущений, и в результате МП не появилось неприемлемого с точки зрения культуры переводного языка контента. Однако такой переводной текст вызывает, как правило, «когнитивный диссонанс» у профессиональных переводчиков, лингвистов и других получателей, трепетно относящихся к языку. Рассмотрим недочеты, вызванные самыми разнообразными причинами.

Для этой группы характерны в первую очередь недочеты, возникающие при калькировании, когда не учитывается контекст и нарушается лексическая сочетаемость. Чаще всего они появляются при переводе многозначных слов или контекстов. Ср:

(4) *Couvre-feu, fermetures... toutes les mesures à Paris.*

Комендантский час, *закрытие*... все меры в Париже.

(5) *Signaler une anomalie avec DansMaRue.*

Сообщить *об аномалии* с DansMaRue.

(6) Une nouvelle *halte* de nuit pour les femmes sans abri.

Новая ночная *остановка* для бездомных женщин.

(7) En raison du *contexte sanitaire*, les cours des mineurs sont suspendus à partir du 16 janvier.

Из-за санитарного контекста занятия для несовершеннолетних приостанавливаются с 16 января.

Во всех приведенных примерах получатель без больших усилий способен понять содержание высказываний: в (4) речь идет о *списке организаций, закрытых* в связи с эпидемией COVID-19; в (5) – о *технических неполадках* в указанной рубрике; в (6) – о *приюте* для женщин, попавших в сложную ситуацию; в (7) – об *эпидемиологической ситуации*.

К недочетам этого же уровня, то есть характерным для «приемлемого качества», можно отнести и грамматические ошибки:

– несогласование падежа после числительных:

(8) 13 «arbres remarquables» dans Paris;

13 замечательные деревья в Париже;

(9) 15 expos à ne pas manquer en 2021;

15 выставки, которые нельзя пропустить в 2021 году;

– неправильный выбор предлога из нескольких потенциальных вариантов в переводном тексте:

(10) les musées parisiens travaillent déjà à vous proposer *pour* 2021;

музеи уже работают над тем, чтобы предложить вам *для* 2021;

– потеря анафорических связей и эллипса:

(11) La vaccination *des plus de 75 ans* débute ce lundi.

Вакцинация *старше 75 лет* начинается в понедельник.

(12) *Vous avez jusqu'au 7 février* pour vous inscrire en ligne.

У вас есть до 7 февраля, чтобы зарегистрироваться в Интернете.

В приведенных примерах получатель на языке перевода легко восстанавливает пропущенные машиной смысловые связи и понимает, что речь идет о вакцинации *людей старше 75 лет* и *возможности до 7 февраля* получить какую-то услугу.

Наиболее частотным является такой недочет, как дословный перевод, при котором происходит нарушение лексической и грамматической сочетаемости.

(13) A partir du 16 janvier, seuls les équipements sportifs extérieurs restent ouverts jusqu'à 18 h.

С 16 января до 18 часов *остается только открытый* спортивный инвентарь.

(14) *Sa sépulture est visible* au cimetière de Montmartre.

Его захоронение видно на Монмартрском кладбище.

Получатель понимает, что в высказывании (13) речь идет о том, что *с 16 января спортивное оборудование на открытом воздухе доступно для занятий до 18 часов*, в (14) – о том, что *могилу можно увидеть на кладбище Монмарта*. Таких примеров достаточно много, что может отразиться на прагматическом уровне: получатель контента может испытывать определенное раздражение из-за того, что понимание текста затруднено по причине узуально и нормативно некорректного языкового оформления.

Сопоставительный анализ оригинальных и переводных высказываний на портале обнаруживает, что 20–35 % контента – это перевод с ошибками, приводящими к искажению смысла или его полному отсутствию в переводном тексте. Все эти ошибки, на наш взгляд, вызваны различными причинами, но они всегда являются следствием того, что искусственный интеллект пока еще не в состоянии выстраивать сложные смысловые связи и обращаться к самым неожиданным контекстам для извлечения смыслов и обличения инферированных подтекстов в вербальную форму. Прежде всего отметим трудности, которые испытывает движок при переводе коротких названий, поскольку их интерпретация требует большого объема фоновых знаний и понимания узкого и широкого контекста. В таких ситуациях МП дает кальку исходного слова или словосочетания, так как заложенные алгоритмы принятия решения не позволяют прибегнуть к адекватной замене, раскрывающей содержание этого явления / события.

(15) *Stationnement résidentiel*.

Жилая парковка.

(16) *Encombrants*.
Громоздкие.

Если перевод фразы в (15) благодаря фоновым знаниям получателя и семантике лексических единиц, входящих в словосочетание, позволяет получателю догадаться, что речь идет о *парковке для жителей квартала*, то перевод в примере (16) создает сложности для понимания того, что речь идет о рубрике *утилизировать громоздкие вещи*.

Далее следует отметить ошибки, вызванные особенностями французского языка. Его носители часто используют аббревиатуры, сокращают слова или образуют новые слова за счет слияния сокращенных основ словосочетания. При переводе таких единиц машина начинает что-то «выдумывать» в надежде угадать правильный ответ [Как победить морников, 2017].

(17) *Expos*.
Экспо.

Данный пример, исходя из контекста, абсолютно прозрачен для переводчика-специалиста: *Expos* = *выставки*, но машина, ориентируясь на форму, выбирает другой вариант, который создает смысловую ошибку, поскольку соотносится с иным контекстом.

(18) *Facil'Familles*.
Facil'Семьи.

(19) *La campagne de vaccination se poursuit dans les Ehpad publics*.

Кампания вакцинации продолжается в общественных *Ehpad*.

В обоих случаях машина приняла решение сохранить частично или полностью исходный вариант сокращения. В примере (18) движок оставил первую часть сокращения в оригинале, а вторую часть перевел; при таком варианте перевода содержание рубрики осталось полностью затемненным для получателя, хотя хороший переводчик, скорее всего, поймет, что речь идет о портале для семей и предложит вариант *все для семьи*. В примере (19) машина перенесла сокращение на языке оригинала в перевод, хотя само сокращение, его расшифровка и перевод зафиксированы во всех онлайн-словарях: *Établissement*

d'Hébergement pour Personnes Agées Dépendantes – дом для престарелых. Эта ошибка повлекла за собой следующую – неправильный выбор варианта перевода для многозначного прилагательного, что привело к серьезному смысловому искажению в переводе, поскольку в русскоязычной версии транслируется информация о вакцинации в общественных местах, а не *государственных домах для престарелых*. К тому же окончание *-pad* может соотноситься русскоязычным получателем с продукцией компанией Apple, например, *айпад*, что полностью искажает смысл исходного высказывания.

Далее рассмотрим случаи, когда предложенный движком перевод семантически полностью неправильный.

(20) *On connaît le plus emblématique de la capitale : l'arc de triomphe de la place de l'Etoile*.

Известна самая знаменитая столица – Триумфальная арка на площади Этуаль.

Ошибка, допущенная машиной в этом примере, вероятно, вызвана тем, что МП до сих пор испытывает определенные трудности с восстановлением анафорических связей в тексте. В данном случае движок не смог отнести артикли *le* и *la*, грамматическая и анафорическая связь которых очевидна переводчику человеку. Предлагаем вариант перевода:

Известна самая знаменитая арка столицы:
Триумфальная арка на площади Этуаль.

(21) *Le 15 janvier 1200, le roi Philippe Auguste reconnaît officiellement l'Université de Paris, l'une des premières universités d'Europe*.

15 января 1200 года король Филипп Август официально признал Парижский университет одним из первых университетов Европы.

В примере (21) смысловая ошибка в переводе вызвана многозначностью глагола и недостаточной способностью машины правильно интерпретировать знаки пунктуации. Для специалиста-переводчика запятая в оригинале сигнализирует о том, что часть является приложением к существительному *l'Université*, об этом же сигнализирует субстантивированное числительное *l'une*. К тому же специалист может ориентироваться на фоновые знания о том, что первый официаль-

ный документ о создании первого *Universitas* – сообщества преподавателей и студентов со своим уставом, правилами и гарантией автономии для этого сообщества – был издан королем Филиппом II Августом. Предлагаем вариант перевода:

15 января 1200 года король Филипп Август официально *основывает* Парижский университет, один из первых университетов Европы.

(22) *Paris est placé en vigilance jaune neige et verglas, aujourd’hui, par Météo France. La vigilance est prévue pour cette nuit.*

Париж находится в желтом снегу и гололеде, сегодня погода Франция. Бдительность назначена на эту ночь.

Пример (22) иллюстрирует недостаточную способность искусственного интеллекта членить фразу на смысловые единицы в отсутствие знаков препинания и верно интерпретировать написание слова с заглавной буквы. Перевод, предложенный движком, не имеет смысла, поскольку в оригинале развернутое словосочетание *en vigilance jaune neige et verglas* было неверно разбито движком на базовые смысловые единицы, хотя даже начинающему переводчику известно, что прилагательное цвета во французском языке всегда стоит после существительного, к которому оно относится. Кроме того, словосочетание *vigilance jaune*, за которым следует *neige et verglas* и *Météo France*, в сознании переводчика-человека активирует единицы, связанные с уровнями погодной опасности, а заглавные буквы в словосочетании *Météo France* сигнализируют о том, что это название национальной метеорологической службы. Некорректный перевод движком первой фразы обуславливает семантически неверный перевод второго предложения. Предлагаем вариант перевода:

В Париже объявлен желтый уровень опасности из-за снега и гололеда на сегодняшнюю ночь, – сообщает сегодня Метео Франс.

В анализируемом контенте было выявлено, что «Яндекс.Переводчик» испытывает сложности с переводом англицизмов во французском языке. Это вполне предсказуемо: движок обучен работать с двумя языками, и

все, что выходит за пределы программы, вызывает у машины сбой.

(23) *Adeptes des dribbles et des dunks, vous avez envie de travailler votre shoot dans un endroit sympa ? On vous propose un petit tour d’horizon des plus beaux playgrounds de basket de la capitale.*

Вы хотите работать *в приятном месте*? Вам предлагается небольшой обзор самых красивых баскетбольных *игр* столицы.

В примере (23) движок пропустил англицизмы (кстати, широко известные для любителей баскетбола) в первом предложении, а во втором предложении дал неправильный вариант перевода заимствования, что привело к абсолютно некорректному по смыслу переводу. Предлагаем вариант перевода:

Любите постучать мячом и побросать в корзину, хотите поработать над своим броском в крутом месте? Для вас обзор самых красивых баскетбольных площадок столицы.

Однако при переводе имени мэра города Парижа движок дает перевод с английского варианта звучания имени:

(24) *Anne Hidalgo présente ses vœux 2021 en vidéo. Эnn Идальго представляет свои пожелания в видео.*

Несомненно, для рядового получателя это всего лишь сбой, не влияющий на понимание текста, но человек-переводчик знает, что подобные ошибки могут приводить к серьезным конфликтам. Предлагаем вариант перевода:

Аnn Идальго: видеопоздравление с 2021 годом.

В качестве примера курьезных случаев, которые в целом стали гораздо реже встречаться в современных МП, можно привести переводы, где имена нарицательные были даны как имена собственные:

(25) *Fontaine, raconte-moi des histoires...*

Фонтейн, расскажи мне истории ...

(26) *puis un net redoux en soirée et la nuit suivante;*

а затем *чистый Реду* вечером и следующей ночью.

Если в примере (25) есть определенные орфографические предпосылки для неверно-

го перевода – заглавная буква и возможность интерпретировать слово как обращение, то в примере (26) остается непонятным, почему машина интерпретировала имя нарицательное как имя собственное. Предлагаем варианты переводов для (25), (26) соответственно:

Фонтан, расскажи мне истории;
а затем *явное потепление* вечером и ночью.

Иными словами, абдуктивный поиск движка приводит пока к нелепостям в переводе, которые условно можно назвать «машинными ляпами».

Заключение

Анализ контента презентационной страницы городского портала Парижа показал, что «Яндекс.Переводчик» на 75–80 % справляется с задачей предоставить рядовому пользователю возможность ознакомиться с иноязычным материалом в хорошем или приемлемом качестве текста. Если учесть, что машина может это сделать за сотые доли секунды, то, по нашему мнению, это очень хороший результат, демонстрирующий, насколько глубоко искусственный интеллект вторгается в очень непростую сферу языка и сознания.

Выявленные ошибки можно квалифицировать следующим образом:

1. Переводческие недочеты без потери семантического смысла (75 %):

- нарушение грамматической сочетаемости;
- нарушение лексической сочетаемости;
- потеря анафорических связей;
- калькирование.

2. Семантические ошибки (25 %):

- потери и деформации сокращений;
- трудности с пониманием англицизмов;
- трудности с пониманием многозначных слов и игнорирование контекста (как широкого, так и узкого);
- трудности с пониманием сложного синтаксиса и анафорических связей;
- «машинные ляпы».

При этом следует отметить, что качество МП значительно повышается в связи с развитием систем искусственного интеллек-

та. Те ошибки, которые были отмечены нами, возможно будут преодолены в ближайшее время за счет того, что «Яндекс.Переводчик» быстро обучается и активно использует тексты, порождаемые человеком-переводчиком. Например, российская переводческая онлайн-платформа SmartCat, популярная как среди фрилансеров, так и корпоративных пользователей, использует МП Яндекса, тем самым способствуя его прогрессу.

Однако основные проблемы создания МП не относятся к сфере семантических ошибок, которых в целом становится меньше, а лежат в плоскости стандартизации языка, его однообразия и упрощения. Статистическая модель построена на оперировании уже готовыми соответствиями, которые выбираются машиной из значительного количества параллельных текстов, и не предполагает порождение «адекватных замен». Несмотря на то что внедрение нейросетей и переход на гибридную систему дает возможность более свободного порождения вариантов перевода, на сегодняшнем этапе развития МП текст, создаваемый машиной, – это повторяющийся контент «приемлемого качества». Частое его воспроизведение создает опасность того, что человек, потребляющий такой контент, начнет порождать тексты «приемлемого качества». Следовательно, сегодня не теряет актуальности классическая постановка проблемы: МП сами по себе не представляют опасности, она заключается в том, насколько разумно и ответственно этот инструмент используется человеком. Как показали результаты нашего исследования, на данный момент не удастся достичь качественного абдуктивного машинного поиска, соответствующего уровню хорошего человека-переводчика.

МП уже изменяет и будет еще радикальнее трансформировать ситуацию на переводческом рынке. «Человеческий» перевод постепенно станет роскошью, для его выполнения будут привлекаться только высококвалифицированные переводчики, а специалисты среднего уровня будут постепенно вытеснены МП, при этом функции квалифицированных переводчиков в сфере нехудожественных текстов будут смещены в сторону редактирования МП.

СПИСОК ЛИТЕРАТУРЫ

- Беллос Д., 2019. «Что за рыбка в вашем ухе?». М. : КоЛибри. 416 с.
- Гуреева А. А., 2017. PR-функция текста туристического брендинга: лингвистические средства как способ реализации // Брендинг как коммуникативная технология XXI века : материалы III Всерос. науч.-практ. конф., г. Санкт-Петербург, 1–3 марта 2017 г. СПб. : Изд-во С.-Петерб. гос. экон. ун-та. С. 128–131.
- Зарубина А. Б., 2018. Урбанистический дискурс в аспекте мультидисциплинарной дискурсологической практики как драйвер позитивных изменений городской коммуникативной среды // Филологический аспект. № 1 (33). С. 172–181.
- Как победить морников: Яндекс запустил гибридную систему перевода, 2017. URL: <https://yandex.ru/blog/company/kak-pobedit-mornikov-yandeks-zapustil-gibridnuyu-sistemu-perevoda> (дата обращения: 14.01.2021).
- Меняйлова М., 2020. Машинный перевод: глобализация, коронавирус или почему мы все еще сомневаемся? URL: <https://www.topt.ru/library/translation-as-service/mashinnyj-perevod-globalizacziya,-koronavirus-ili-pochemu-myi-vse-eshhe-somnevaemysya.html> (дата обращения: 10.01.2021).
- Новикова Э. Ю., 2015. Дискурсивные характеристики «глобального города» на городском интернет-портале // Вестник Московского государственного лингвистического университета. № 6 (717). С. 467–476.
- Попова О. И., Королькова С. А., Степанова Е. В., 2020. Стратегии брендинга малых городов в интернет-пространстве // Вестник Волгоградского государственного университета. Серия 2, Языкознание. Т. 19, № 2. С. 123–138. DOI: <https://doi.org/10.15688/jvolsu2.2020.2.11>.
- Тарарак Е., 2020. Машина vs человек: отберет ли искусственный интеллект хлеб у переводчиков? URL: <https://novayagazeta.ru/articles/2020/12/13/88357-mashina-vs-chelovek> (дата обращения: 10.01.2021).
- Chen Fachun, Leontovich O. A., 2020. A Tale of two Cities: Historical Narratives in the Russian and Chinese Urban Landscapes // Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 2, Yazykoznanie [Science Journal of Volgograd State University. Linguistics], 2020, vol. 19, no. 2, pp. 78–85. DOI: <https://doi.org/10.15688/jvolsu2.2020.2.7>.
- Aiello G., Tosoni S., 2016. Going About the City: Methods and Methodologies for Urban Communication Research // International Journal of Communication. № 10. P. 1252–1262.
- International Standard on Translation Services ISO 18587: 2017. URL: https://standartgost.ru/g/ISO_18587:2017 (date of access: 31.01.2020).
- Loock R., 2016. Using electronic corpora by professional translators: When? How? What for? // Revue de l'Institut des langues et cultures d'Europe, Amérique, Asie et Australie. No. 27. URL: <http://journals.openedition.org/ilcea/3835> (date of access: 5.01.2021).
- Lupieri S., 2019. La traduction automatique fait des pas géant. URL : <https://www.lesechos.fr/weekend/business-story/la-traduction-automatique-fait-des-pas-de-geant-1211737> (date of access: 15.01.2021).
- Susskind D., 2020. A World Without Work: Technology, Automation, and How We Should Respond. New York : Metropolitan Books. 320 p.
- TAUS, 2010. Machine Translation Post-editing Guidelines. URL: <https://www.taus.net/think-tank/best-practices/postedit-best-practices/machine-translation-post-editing-guidelines> (date of access: 5.01.2021).

ИСТОЧНИК

Site officiel de la Ville de Paris. URL: <https://www.paris.fr/> (date of access: 5.01.2021).

REFERENCES

- Bellos D., 2019. «*Chto za rybka v vashem uhe?*» [“Is That a Fish in Your Ear?”]. Moscow, KoLibri Publ. 416 p.
- Gureeva A.A., 2017. PR-funkcija teksta turisticheskogo brendinga: lingvisticheskie sredstva kak sposob realizacii [Branding PR-Function in Travel Discourse Texts: Ways to Perform]. *Branding kak kommunikativnaya tekhnologiya XXI veka: materialy III Vseros. nauch.-prakt. konf., g. Sankt-Peterburg, 1–3 marta 2017 g.* [Branding as a Communication Technology of the 21st Century. Proceedings of the 3rd All-Russian Scientific and Practical Conference. Saint-Petersburg, 1–3 March, 2017]. Saint Petersburg, Izd-vo Sankt-Peterburgskogo gosudarstvennogo ekonomicheskogo universiteta, pp. 128–131.
- Zarubina A.B., 2018. Urbanisticheskij diskurs v aspekte multidisciplinarnoy diskursologicheskoy praktiki kak drayver pozitivnykh izmenenij gorodskoy kommunikativnoj sredy [Urban Discourse in the Aspect of Multidisciplinary Discursive Practices as a Driver of Positive Changes in the City Communicative Environment]. *Filologicheskij aspekt*, vol. 1 (33), pp. 172–181.

- Kak pobedit' mornikov: Jandeks zapustil gibridnuju sistemu perevoda*, 2017 [How to Defeat Morniki: Yandex has Launched a Hybrid Translation System]. URL: <https://yandex.ru/blog/company/kak-pobedit-mornikov-yandeks-zapustil-gibridnuyu-sistemu-perevoda> (accessed 14 January 2021).
- Menyaylova M., 2020. *Mashinnyy perevod: globalizacija, koronavirus ili pochemu my vse eshhe somnevaemsja?* [Machine Translation: Globalization, Coronavirus, or Why Do We Still Have Doubts?]. URL: <https://www.toptr.ru/library/translation-as-service/mashinnyj-perevod-globalizacziya,-koronavirus-ili-pochemu-myi-vse-eshhe-somnevaemysya.html> (accessed 10 January 2021).
- Novikova E.Yu., 2015. Diskursivnye harakteristiki «glokal'nogo goroda» na gorodskom internet-portale [Discourse of the “Glocal City” on the Urban Website]. *Vestnik Moskovskogo gosudarstvennogo lingvisticheskogo universiteta* [Vestnik of Moscow State Linguistic University], no. 6 (717), pp. 467-476.
- Popova O.I., Korol'kova S.A., Stepanova E.V., 2020. Strategii brendirovaniya malyh gorodov v internet-prostranstve [Branding Strategies for Small Cities in the Internet Space]. *Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 2, Yazykoznanie* [Science Journal of Volgograd State University. Linguistics], vol. 19, no. 2, pp. 123-138. DOI: <https://doi.org/10.15688/jvolsu2.2020.2.11>.
- Tararak E., 2020. *Mashina vs chelovek: otberet li iskusstvennyj intellekt hleb u perevodchikov?* [Machine vs Human: Will Artificial Intelligence Take the Bread Away From Translators?]. URL: <https://novayagazeta.ru/articles/2020/12/13/88357-mashina-vs-chelovek> (accessed 10 January 2021).
- Chen Fachun, Leontovich O.A., 2020. A Tale of Two Cities: Historical Narratives in the Russian and Chinese Urban Landscapes. *Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 2, Yazykoznanie* [Science Journal of Volgograd State University. Linguistics], 2020, vol. 19, no. 2, pp. 78-85. DOI: <https://doi.org/10.15688/jvolsu2.2020.2.7>.
- Aiello G., Tosoni S., 2016. Going About the City: Methods and Methodologies for Urban Communication Research. *International Journal of Communication*, no. 10, pp. 1252-1262.
- International Standard on Translation Services ISO 18587:2017*. URL: https://standartgost.ru/g/ISO_18587:2017 (accessed 31 January 2020).
- Loock R., 2016. Using Electronic Corpora by Professional Translators: When? How? What for? *Revue de l'Institut des langues et cultures d'Europe, Amérique, Asie et a Australie*, no. 27. URL: <http://journals.openedition.org/ilcea/3835> (accessed 5 January 2021).
- Lupieri S., 2019. *La traduction automatique fait des pas géant*. URL: <https://www.lesechos.fr/weekend/business-story/la-traduction-automatique-fait-des-pas-de-geant-1211737> (accessed 15 January 2021).
- Susskind D., 2020. *A World Without Work: Technology, Automation, and How We Should Respond*. New York, Metropolitan Books. 320 p.
- TAUS, 2010. *Machine Translation Post-Editing Guidelines*. URL: <https://www.taus.net/think-tank/best-practices/postedit-best-practices/machine-translation-post-editing-guidelines> (accessed 5 January 2021).

SOURCE

Site officiel de la Ville de Paris. URL: <https://www.paris.fr/> (accessed 5 January 2021).

Information About the Authors

Svetlana A. Korolkova, Candidate of Sciences (Pedagogy), Associate Professor, Department of Translation Theory and Practice, Volgograd State University, Prosp. Universitetsky, 100, 400062 Volgograd, Russia, tipp@volsu.ru, korolkova@volsu.ru, <https://orcid.org/0000-0003-2083-4101>

Anna A. Novozhilova, Candidate of Sciences (Philology), Associate Professor, Department of Translation Theory and Practice, Volgograd State University, Prosp. Universitetsky, 100, 400062 Volgograd, Russia, tipp@volsu.ru, novozhilova@volsu.ru, <https://orcid.org/0000-0001-7601-9048>

Информация об авторах

Светлана Азадовна Королькова, кандидат педагогических наук, доцент кафедры теории и практики перевода, Волгоградский государственный университет, просп. Университетский, 100, 400062 г. Волгоград, Россия, tipp@volsu.ru, korolkova@volsu.ru, <https://orcid.org/0000-0003-2083-4101>

Анна Алексеевна Новожилова, кандидат филологических наук, доцент кафедры теории и практики перевода, Волгоградский государственный университет, просп. Университетский, 100, 400062 г. Волгоград, Россия, tipp@volsu.ru, novozhilova@volsu.ru, <https://orcid.org/0000-0001-7601-9048>